8-2021

Quantification of Compound and Cascading Hydroclimatic Extreme Events

Sourav Mukherjee
Clemson University

Follow this and additional works at: https://tigerprints.clemson.edu/all_dissertations

Part of the Civil Engineering Commons

Recommended Citation
Mukherjee, Sourav, "Quantification of Compound and Cascading Hydroclimatic Extreme Events" (2021).
All Dissertations. 3320.
https://tigerprints.clemson.edu/all_dissertations/3320

This Dissertation is brought to you for free and open access by the Dissertations at TigerPrints. It has been accepted for inclusion in All Dissertations by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.
QUANTIFICATION OF COMPOUND AND CASCADING HYDROCLIMATIC EXTREME EVENTS

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Civil Engineering

by
Sourav Mukherjee
August 2021

Accepted by:
Dr. Ashok Mishra, Committee Chair
Dr. Abdul A. Khan
Dr. Lawrence Murdoch
Dr. Charles Privette, III
ABSTRACT

Compound and cascading hydroclimatic extreme events have garnered much attention in recent studies. The combined effects of interconnected extremes can cause widespread damage, with a higher potential impact than individual extremes. Both anthropogenic warming and natural climate variability affect these extremes, which is why detecting past extreme events, understanding the underlying mechanisms, and assessing their future impacts can aid in mitigation efforts to reduce their overall impact. However, thus far, identifying such events is oversimplified and the propagation of their impact as cascades from the physical to human systems remains partly explored.

The overreaching goal of this thesis is to develop robust methodologies to quantify the compound and cascading extreme events, such as drought and heatwaves, extreme precipitation and atmospheric rivers, extreme heat and humidity, and flash droughts in the past and future climate. A suite of advanced statistical methods, system dynamics, and causality approaches are implemented to achieve the research goal. This thesis consists of ten chapters, and the objective of each chapter are summarized as follows.

1. Chapter 1 provides a brief introduction and examples of various compound and cascading hydroclimatic extremes.
2. Chapter 2 provides a perspective of drought indices and highlights the challenges in the context of climate change.
3. The objective of Chapter 3-Chapter 5 is to quantify the compound drought and heatwave characteristics (frequency, duration, and severity) and
investigate their association with natural climate variability, anthropogenic warming, land-climate feedback, and background aridity across the globe.

(4) **Chapter 6** is dedicated to quantifying the future changes in potential impact of heat-stress (combination of extreme heat and humidity) on the human population.

(5) The cascading influence of meteorological forcing on the moisture advection processes associated with extreme precipitation related to atmospheric rivers is discussed in **chapter 7**.

(6) The objective of **chapter 8** is to investigate and quantify the compound and cascading influence of different spatial drivers, such as precipitation, temperature, surface-energy portioning, soil moisture-temperature coupling strength, and vapor pressure deficit on the evolution and intensification of global flash droughts.

(7) **Chapter 9** proposes a methodology to quantify the compound and cascading effects in a dry-hot event network using a system dynamics approach. Finally, the conclusion and recommendations are provided in **chapter 10**.
DEDICATION

To my daughter
ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor Dr. Ashok Mishra for his support and guidance throughout my graduate studies. I would also like to thank my committee members Dr. Abdul Khan, Dr. Lawrence Murdoch, and Dr. Charles Privette, for their support and encouragement when I needed one.

I am deeply indebted to my mother, grandparents, uncles, brother, girlfriend, and her mother for their constant support and inspiration during this period. I would like to thank my colleagues, Dr. Anoop Valiyaveettil, Dr. Goutam Konapala, Dr. Tue Vue, Dr. Ali Omar Alnahit, Somnath Mondal, Ram Prasad Yaddanapudi, and Kumar Puran Tripathy, for their academic and philosophical insights on various topics. Lastly, I express my gratitude to Kristin Baker for her timely help with my graduation.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE PAGE</td>
<td>i</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xii</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION TO COMPOUND AND CASCADING HYDROCLIMATIC EXTREME EVENTS</td>
<td></td>
</tr>
<tr>
<td>Definition of Compound and Cascading Events</td>
<td>1</td>
</tr>
<tr>
<td>Examples of Compound and Cascading Hydroclimatic Extremes</td>
<td>4</td>
</tr>
<tr>
<td>Compound and Cascading Drought and Heatwave</td>
<td>4</td>
</tr>
<tr>
<td>Compound Heat Extreme and Humidity</td>
<td>6</td>
</tr>
<tr>
<td>Compound Precipitation and Wind Associated</td>
<td>8</td>
</tr>
<tr>
<td>with Atmospheric Rivers</td>
<td></td>
</tr>
<tr>
<td>Flash Drought Caused by Compound Changes in Temperature and Precipitation</td>
<td>10</td>
</tr>
<tr>
<td>Wildfires under pre-existing hot and dry conditions</td>
<td>11</td>
</tr>
<tr>
<td>Landslide due to precipitation after wildfire</td>
<td>12</td>
</tr>
<tr>
<td>Marine Heat waves, acidification, and deoxygenation</td>
<td>12</td>
</tr>
<tr>
<td>High-Ozone Concentration Associated with</td>
<td>13</td>
</tr>
<tr>
<td>Concurrent Heatwave and Stagnation</td>
<td></td>
</tr>
<tr>
<td>Concurrent Daytime and Nighttime Heatwave</td>
<td>14</td>
</tr>
<tr>
<td>Flooding due to the Convolution of Tropical Cyclone,</td>
<td></td>
</tr>
<tr>
<td>Precipitation, Storm Surge</td>
<td>15</td>
</tr>
<tr>
<td>Air pollution from combined drought and air chemistry</td>
<td>15</td>
</tr>
<tr>
<td>Concluding Remarks</td>
<td>16</td>
</tr>
<tr>
<td>References</td>
<td>16</td>
</tr>
</tbody>
</table>
Table of Contents (Continued)

II. A PERSPECTIVE OF DROUGHT INDICES UNDER CLIMATE CHANGE ................................................................. 39

   Overview of Drought Indices .......................................................... 39
   Relevance of Drought Indices in Climate Change .......................... 45
   Example of the Association Between Drought Indices and Land Surface Warming ........................................... 50
   Application and Limitations of Existing Drought Indices ............. 52
   Challenges Associated with Drought Indices in Climate Change Studies .......................................................... 58
   Summary and Conclusions ............................................................... 70
   References ...................................................................................... 74

III. COMPOUND DROUGHT AND HEATWAVES AT A GLOBAL SCALE: ROLE OF NATURAL CLIMATE VARIABILITY .............. 92

   Introduction ................................................................................... 92
   Methods .......................................................................................... 96
      Datasets ....................................................................................... 96
      Frequency of CDHW Events .......................................................... 97
      Association between the large-scale climate variability and the occurrence of CDHW events .......................... 99
   Results and Discussion ................................................................. 105
      Potential regional drivers at seasonal scale ................................. 105
      Relationship between CDHW events and large-scale climate variability ......................................................... 109
   Atmospheric anomalies associated with warm and cold phases ........................................................................ 114
   Surface energy budget anomalies associated with the warm and cold phases ................................................... 119
   Summary and Conclusion ............................................................... 124
   References ...................................................................................... 126
Table of Contents (Continued)

IV. COMPOUND DROUGHT AND HEATWAVE IN A WARMING WORLD .............................................................. 140
   Introduction .......................................................................................................................... 140
   Methods .............................................................................................................................. 142
   Data ...................................................................................................................................... 142
   Estimation of CDHW Event Characteristics ...................................................................... 143
   Results .................................................................................................................................... 144
   Changes in CDHW events .................................................................................................. 144
   Spatiotemporal changes in CDHW characteristics ......................................................... 148
   Effect of Background Aridity at Continental Scale ......................................................... 155
   Discussion and Concluding Remarks .............................................................................. 159
   References .......................................................................................................................... 161

V. RELATIVE CONTRIBUTION OF ANTHROPOGENIC WARMING AND NATURAL CLIMATE VARIABILITY TO CHANGES IN COMPOUND DROUGHT AND HEATWAVE ........................................ 175
   Introduction .......................................................................................................................... 175
   Materials and Methods ...................................................................................................... 177
   Data ...................................................................................................................................... 177
   Estimation of Compound Drought and Heatwave events .............................................. 179
   Measurement of degree of Susceptibility of HW (DSHW) towards drought ................. 179
   Measurement of Odds of occurrence of CHWD events ................................................. 180
   Results .................................................................................................................................... 182
   Annual Increase in the number of CDHW events ......................................................... 182
   Degree of Susceptibility of HW towards Drought ......................................................... 184
   Scaling factors associated with the Odds of having CDHW day .................................. 190
   Effect of 1.5°C and 2°C rise in Global Warming ............................................................... 193
   Discussion and Conclusion ............................................................................................... 195
   References .......................................................................................................................... 197
Table of Contents (Continued)  

VI. QUANTIFYING HEAT-STRESS IMPACT DUE TO ANTHROPOGENIC WARMING .......................................................... 210

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>210</td>
</tr>
<tr>
<td>Data and Methods</td>
<td>213</td>
</tr>
<tr>
<td>Climate data and Calculation of daily Heat Index</td>
<td>213</td>
</tr>
<tr>
<td>Population Estimates</td>
<td>215</td>
</tr>
<tr>
<td>Non-stationary GEV Framework to estimate High-end HS Severity</td>
<td>216</td>
</tr>
<tr>
<td>Estimation of Potential Impact of summer extreme HS</td>
<td>218</td>
</tr>
<tr>
<td>Estimating Changes in Risk and Relative Effect of Indicators</td>
<td>221</td>
</tr>
<tr>
<td>Results</td>
<td>222</td>
</tr>
<tr>
<td>Historical Analysis and Future Projections of extreme HS severity</td>
<td>222</td>
</tr>
<tr>
<td>Projected Extreme Summer HS Characteristics</td>
<td>225</td>
</tr>
<tr>
<td>Changes in Population Exposure</td>
<td>226</td>
</tr>
<tr>
<td>Future Changes in Risk of Potential Impact of High-End Summer HS</td>
<td>229</td>
</tr>
<tr>
<td>Summary and Conclusion</td>
<td>232</td>
</tr>
<tr>
<td>References</td>
<td>233</td>
</tr>
</tbody>
</table>

VII. INVESTIGATING THE CASCADING IMPACT OF METEOROLOGICAL FORCINGS ON EXTREME PRECIPITATION DRIVEN BY ATMOSPHERIC RIVERS .................................................................................. 246

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>246</td>
</tr>
<tr>
<td>Data and Methodology</td>
<td>251</td>
</tr>
<tr>
<td>Data</td>
<td>251</td>
</tr>
<tr>
<td>AR-Detection Methodology</td>
<td>252</td>
</tr>
<tr>
<td>AR Event and AR-EP Event Identification</td>
<td>253</td>
</tr>
<tr>
<td>Composite Analysis</td>
<td>254</td>
</tr>
<tr>
<td>Results and Discussion</td>
<td>256</td>
</tr>
<tr>
<td>Spatial distribution of EP and AR events in SEUS</td>
<td>256</td>
</tr>
<tr>
<td>AR-EP Characteristics in SEUS</td>
<td>258</td>
</tr>
</tbody>
</table>
Table of Contents (Continued)

Cascading effect of Synoptic Patterns on Moisture Availability and Transport ........................................... 265
Synthesis and Conclusion ...................................................... 275
References ........................................................................ 278

VIII. INVESTIGATING CLIMATE CONTROLS OF GLOBAL FLASH DROUGHTS ........................................ 293

Introduction ........................................................................ 293
Methodology ....................................................................... 296
Flash Drought Identification .................................................. 296
Estimation of Aridity Index .................................................... 300
Estimation of Vapor Pressure Deficit ...................................... 301
Estimation of Soil Moisture-Temperature Coupling Strength  ................................................................. 302
Determining the Variable Importance using Random Forest Algorithm ................................................. 303
Estimating Sensitivity of FD Intensity to Climate Variables ................................................................. 303
Results ................................................................................ 304
Hotspots of FD Frequency and Effect of Surface Energy Partitioning ......................................................... 304
Climate Controls Associated with Onset and Evolution of Flash droughts ............................................... 309
Relative Importance of Climate Controls on Flash Drought Intensity ..................................................... 312
Sensitivity of Flash Drought Intensity to Climate Variables ......................................................................... 316
Discussion and Conclusion ..................................................... 318
References ........................................................................... 320
Table of Contents (Continued)

IX. A CASCADE MODEL TO QUANTIFY COMPOUND AND CASCADING EFFECTS IN A DRY-HOT EVENT NETWORK .................................................................................. 335

Introduction ......................................................................................................................... 335
Methodology .......................................................................................................................... 339
Data ......................................................................................................................................... 339
Determining Compound and Cascading Dry and Hot Event Network .............................. 340
Model-based Estimation of Compound and Cascading Effects ........................................... 341
Results .................................................................................................................................... 344
Hotspots of Compound and Cascading Dry-Hot and Hot-Dry Events ............................... 344
Influence of Confounders ....................................................................................................... 347
Role of Background Aridity .................................................................................................... 351
Discussion and Conclusion ..................................................................................................... 353
References ............................................................................................................................... 356

X. CONCLUSION AND RECOMMENDATION ................................................................... 366

APPENDICES .......................................................................................................................... 373

A: Supplementary Information for Chapter 3 ................................................................. 373
B: Supplementary Information for Chapter 4 ................................................................. 399
C: Supplementary Information for Chapter 5 ................................................................. 415
D: Supplementary Information for Chapter 6 ................................................................. 422
E: Supplementary Information for Chapter 7 ................................................................. 440
F: Supplementary Information for Chapter 8 ................................................................. 444
G: Supplementary Information for Chapter 9 ................................................................. 455

xi
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CHAPTER 1</strong></td>
<td></td>
</tr>
<tr>
<td>Figure 1 Schematic diagram to represent risk from interconnected extreme events.</td>
<td>3</td>
</tr>
<tr>
<td>Figure 2 Schematic diagram explaining the cascading effect of ENSO, PDO, and NAO on moisture inhibition, and surface energy partitioning associated with compound drought and heatwaves</td>
<td>5</td>
</tr>
<tr>
<td>Figure 3 NOAA’s Heat Index values corresponding to relative humidity and temperature illustrating the heat warnings</td>
<td>8</td>
</tr>
<tr>
<td>Figure 4 AR making landfall during UTC 1200, 3rd May 2010 in the southeastern US; IVT intensity (shading), and IVT vectors (arrows) are calculated using the ERA5 dataset</td>
<td>9</td>
</tr>
<tr>
<td><strong>CHAPTER 2</strong></td>
<td></td>
</tr>
<tr>
<td>Figure 1 Schematic diagram showing the drought propagation under climate change.</td>
<td>44</td>
</tr>
<tr>
<td>Figure 2 (LEFT) Anomaly in spatially averaged yearly observed LSAT for the period, 1975-2014 with respect to the period, 1961-1990 (bar plot with positive (red) and negative (steelblue) anomaly) and spatially averaged annual PDSI_sc (line; orange) for the period 1975-2014, and (RIGHT) scatterplot (violet) and regression line (blue) of annually averaged PDSI_sc and LSAT Anomaly for (a) Africa, (b) Asia, (c) Australia, (d) Europe, (e) North America, and (f) South America. It is to be noted here that monthly PDSI_sc values are annually averaged and then correlation coefficients are estimated against anomaly based on yearly observed LSAT</td>
<td>53</td>
</tr>
<tr>
<td>Figure 3 Sensitivity of drought indices with change in LSAT for the six continents. a Africa. b Asia. c Australia. d Europe. e North America. f South America. Box plot showing median, interquartile range (IQR), outliers, and overall range excluding the outliers for the annual mean of continental averaged drought indices, SPI-1 (red); SPEI-1(green); and monthly PDSI_sc (blue) for every 0.25 °C change in LSAT during the period 1901–2013. To estimate the statistics related to box plot, values of drought indices are accumulated in bins corresponding to temperature anomalies nearest to every 0.25 °C change in global mean temperature ranging from − 0.5 to 0.75 °C</td>
<td>59</td>
</tr>
</tbody>
</table>
CHAPTER 3

Figure 1 (a) Time series of weekly sc_PDSI and daily Tmax based on the GPCC and CPC dataset for the whole period of analysis, 1982 to 2016, for a single grid cell (longitude= -100.5 , and latitude = 29.5 (in degrees)), and (b) CDHW events (enclosed within the shaded boxes) for the same grid cell during JJA in the year 2012 (shaded region in (a)). The right y-axis represents the sc_PDSI values and the left y-axis represents the Tmax……………………………………………………. 98

Figure 2 Flow diagram explaining the multi-stage exploration of model selection performed for a specific climate region. “NA” indicates none of the drivers are selected for the region or for any grid cell within the region………………………………………. 100

Figure 3 Regional Map showing combinations of potential regional drivers (RD1 to RD7; filled circles) for the 26 AR5-Climate regions based on the GPCC and CPC (left panel), and the ERA5 (right panel). The list of large-scale climate variabilities corresponding to each of the selected combinations of regional drivers (RD1 to RD7) is provided in Figure 2. ........................................................................ 106

Figure 4 Spatial map of the best models (M1 to M7; filled circles) explaining the inter-annual variability of CDHW events from 1982 to 2016 based on the GPCC and CPC (left panel) , and the ERA5 (right panel), which are identified by the Poisson GLM. The list of large-scale natural climate forcing corresponding to each of the selected models (M1 to M7) is provided in Figure 2. The bold boundaries indicate regions exhibiting the influence of a single large-scale driver, which is also consistent in both datasets........................................................................................................ 108

Figure 5 The statistically significant Poisson GLM regression coefficients for ENSO (Nino3.4) based on the GPCC and CPC dataset (left column), and the ERA5 (right column), explaining the interannual variability of CDHW events during (a, b) DJF, (c, d) MAM, (e, f) JJA, and (g, h) SON. The bold boundaries indicate regions exhibiting the influence of a single large-scale driver, which is also consistent in both datasets................................................................. 111
List of Figures (Continued)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 6 Same as in Figure 5 but for PDO</td>
<td>113</td>
</tr>
<tr>
<td>Figure 7 Same as in Figure 5 but for NAO</td>
<td>114</td>
</tr>
<tr>
<td>Figure 8 El Nino (left panel) and La Nina (right panel) based anomalies in divergent winds vectors (m/s) and velocity potential (color shading, unit:m²/s, scaled by 10⁶) contours at 200 mb with respect to the climatology during (a-b) DJF, (c-d) MAM, and (e-f) SON</td>
<td>116</td>
</tr>
<tr>
<td>Figure 9 Same as in Figure 8 but based on warm (left panel) and cold (right panel) phase of (a, b) PDO, and (c, d) NAO During the JJA season</td>
<td>118</td>
</tr>
<tr>
<td>Figure 10 El Nino (left panel) La Nina (right panel) based anomalies in sensible heat flux with respect to the climatology during (a, b) DJF, (c, d) MAM, and (e, f) SON. Stippling represents statistically significant at 95% confidence level. All units are in W/m². The sign convention implemented for the fluxes is positive upwards.</td>
<td>122</td>
</tr>
<tr>
<td>Figure 11 Same as in Figure 11 but based on the warm (left panel) and cold (right panel) phase of PDO. All units are in W/m². The sign convention implemented for the fluxes is positive upwards.</td>
<td>123</td>
</tr>
</tbody>
</table>

CHAPTER 4

Figure 1 (a-c) Linear and scatter plots show the annual timeseries of globally averaged CDHW_r (%) based on (a) GPCC-CPC dataset, (b) GPCC-BE dataset, (c) ERA5 dataset, and (d-f) bar-plots illustrate the percentage of global land area) affected by the CDHW events each year during the whole study period, 1983-2016 (%) based on (d) GPCC-CPC dataset, (e) GPCC-BE dataset, (f) ERA5 dataset. Linear annual trends in CDHW_r and in the global land area are estimated based on the Sen’s slope estimator, and statistical significance in trends are determined based on the MK test for the whole study period (1983-2016), past period (1983-1999), and recent warm period (2000-2016). The numbers in the bracket indicate the estimated slope, and the asterisks denote statistically significant trends (at 95% confidence level). | 146 |
List of Figures (Continued)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 2</td>
<td>Spatial map showing the changes in (a-c) CDHWf between two time periods (recent, 2000-2016 minus past, 1983-1999) based on (a) GPCC-CPC dataset, (b) GPCC-BE dataset, (c) ERA5 dataset, (d-f) same as in (a-c) but for the CDHWd (days/year), and (g-i) same as in (a-c) but for the CDHWs (per event/year). The numbers illustrated in (a-i) represent the percentage of area affected by an increase in the corresponding CDHW metrics for the NH (red), globe (blue), and SH (black).</td>
</tr>
<tr>
<td>Figure 3</td>
<td>(a-c) Line plots showing the latitudinal variation in the past period (1983-1999; in black) average and recent warm period (2000-2016; in red) average of CDHWf (events/year) based on (a) GPCC-CPC dataset, (b) GPCC-BE dataset, (c) ERA5 dataset, (d-f) interannual trends in CDHW characteristics for the past (in black) and recent warm period (in red) based on the three individual datasets, (g-l) same as in (a-f) but for CDHWd (days/year), and (m-r) same as in (a-f) but for the CDHWs (per event/year). The square symbols mark the latitudes showing statistically significant trends in the interannual variation of the CDHW metrics.</td>
</tr>
<tr>
<td>Figure 4</td>
<td>(a) Spatial map showing the arid, transitional, and humid regions identified based on the Aridity Index for the climatological period, 1983-2016, (b-d) Probability density of Past, 1983-1999 period (black), and recent warm, 2000-2016 period mean CDHWs (per event per year) for the arid, transitional, and humid regions of the Globe based on (b) GPCC-CPC, (c) GPCC-BE, and (d) ERA5 datasets.</td>
</tr>
</tbody>
</table>

CHAPTER 5

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Difference between the average number of CHWD events during the Pre-2000 period (1983-1999) and Post-2000 period (2000-2016) (d) Ks density plot for average number of CHWD events during the two periods for the globe, (c) spatial distribution of the ratio of the probabilities where probability of heat wave day conditioned on drought (pe) is significantly (at 5% significance level) greater than probability of heat wave day conditioned on drought (pe), and (d) percentage area of each climate region showing significantly (at 95% confidence level) greater probability of heat wave day conditioned on drought (pe) than that conditioned on no drought (pc).</td>
</tr>
</tbody>
</table>
List of Figures (Continued)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 2</td>
<td>186</td>
</tr>
<tr>
<td>(a) Correlogram showing the significant (at 5% significance level) partial correlation between the number of monthly CHWD days and the interannual variability of large-scale climate indices during the period 1982-2016 based on non-parametric Spearman’s rho. (b) Chord diagram showing the large-scale indices chosen based on the mechanistic explanation.</td>
<td></td>
</tr>
<tr>
<td>Figure 3</td>
<td>191</td>
</tr>
<tr>
<td>Scaling factors (coefficient of regression) and their corresponding 5-95% CI indicating the sensitivity of odds of occurrence of monthly CHWD days against the inter annual variability of large-scale climate variables and ANT obtained from the FLM for the 10 climate regions. The red color indicates the scaling factors for the ANT, and the blue color indicate the same for the large-scale climate indices. The green circles with a blue cross indicate the scaling factors that are not statistically significant (at 5% significant level).</td>
<td></td>
</tr>
<tr>
<td>Figure 4</td>
<td>195</td>
</tr>
<tr>
<td>Ratio of odds (OR) for 1.5°C and 2°C warming limits with respect to the current level of warming.</td>
<td></td>
</tr>
</tbody>
</table>

CHAPTER 6

Figure 1 Historical and CMIP5-model projections of extreme heat stress. (a-c) spatial average of historical observations and CMIP5 modeled projections (period, 2020-2100; shading represents the interquartile range of inter-model variability) of AM-HI representing the extreme HS severity for RCP8.5 (solid brown line) and RCP4.5 (solid red line) and corresponding Tmax (blue and light blue) for (a) 1d-HS, (b) 3d-HS, and (c) 7d-HS, (d-l) spatial maps showing the temporal mean of AM-HI for the CONUS for (d) present (1980-2019), (e) near-future (2020-2059), and (f) far-future (2060-2099) periods corresponding to 1d-HS for the RCP8.5 emission scenario; (g-i) same as in (d-f) but for 3d-HS events; (j-l) same as in (g-i) but for 7d-HS events. |
### List of Figures (Continued)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 2: Changes in Population Exposure and Extreme Summer HS Characteristics (a-c) Data ellipses (or envelops) representing the probability density function evaluated from the bivariate distribution of 40YHI and population estimates for the CONUS in the five climate scenarios, (d-f) probability density for spatial distribution of decadal trends in WTS of daily summer HS severity (summer annual maximum minus median) and bar plots indicating the total population exposure to specific trend magnitudes. Style of (a-c) is inspired by [59]. Note that the ellipses are approximations intended to show inter-scenario comparisons, rather than precisely calculated population counts.</td>
<td>228</td>
</tr>
<tr>
<td>Figure 3 Future Changes in Risk of Potential Impact of High-End Summer HS (a-l) Spatial distribution of RR showing the changes in risk of PI of HS in the RCP4.5-SSP2 and RCP8.5-SSP5 scenario in the near-future and far-future period due to the effect of (a-d) total effect from the interaction between climate change and population, (e-h) climate change only, (i-l) population only for 1d-HS events, and (m-o) boxplots showing the spatial distribution of these changes for the census regions (Mid-west, Northeast, South, and West of CONUS shown in (p)) based on these effects (indicated in a-l) for the 1d-, 3d-, and 7d-HS events.</td>
<td>230</td>
</tr>
</tbody>
</table>

### CHAPTER 7

**Figure 1** (a-b) Total EP95p event counts per year, (c-d) Total EP99p event counts per year, (e-f) AR-event counts per year for the period, 1979-2019 (after removing the TCs). ................................................................. 258

**Figure 2** (a-b) Spatial map showing the total number of (a) AR-EP95p and (b) AR-EP99p events observed per year during the 1979-2019 period over the SEUS region, (c-d) boxplots showing the spatial distribution of (c) AR-EP95p and (d) AR-EP99p event counts per year for the period 1979-2019 for the eight states of the SEUS region, and (e-f) spatial map showing the total precipitation magnitude (mm) and the IVT vectors for the ARs identified during the 3rd and 4th May 2010. The intensity of the IVT is represented by the length of the IVT vectors. ......................................................... 259

**Figure 3** The heat maps (a) show the total number and (b) average magnitude of AR-EP95 events for the overlapping seasons during the period 1979-2019, (c, and d) same as in (a, and b) but for AR-EP99 events. ........ 262
List of Figures (Continued)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 4 (a) Spatial map and (b) box plots showing the geographical distribution of maximum AR-EP magnitude for the top 100 severe AR-EP days during the extended cold (Nov-April) season for the 1979-2019 period, and (c), and (d) same as in (a), and (b), respectively, but for the extended warm (May-Oct) season.</td>
<td>266</td>
</tr>
<tr>
<td>Figure 5 Spatial map showing spatiotemporal evolution of (a-c) anomalies in composites of MSLP (shading; units hPa), and Z850 (contours; units: m) for cold (NDJFMA) season average centered on (a) 5 days before, and (b) 2 days before, and (c) on the day the top 100 severe AR-EP events are observed during the cold season; the location of the local maxima (minima) in Z850 anomaly are marked by the letter, “H” (”L”), (d-f) same as in (a-c) but for warm (MJJASO) season; striplings mark the grid points for which the anomalies in MSLP are found to be statistically significant (at 95% confidence level) based on the two-sample KS-test. The region corresponding to the top 100 severe AR-EP events is defined by the rectangle (black).</td>
<td>269</td>
</tr>
<tr>
<td>Figure 6 Spatial map showing composites of IVT field based on (a) 5 days before, (b) 2 days before, and (c) on the day the top 100 AR-EP events are observed during the cold (Nov-April) season, (d-f) same as in (a-c) but for composite anomalies in TCWV, (g- i) same as in (a-c) but for warm (May-Oct) season, and (j-l) same as in (d-f) but for the warm season. The anomalies are estimated based on the cold or warm season average, and the striplings indicate the grid points where statistically significant (95% confidence level) anomalies are obtained. The region corresponding to the top 100 severe AR-EP events is defined by the rectangle (black).</td>
<td>272</td>
</tr>
</tbody>
</table>
List of Figures (Continued)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
</table>

CHAPTER 8

Figure 1 Hotspots of FD Frequency and Effect of Available Surface Energy Partitioning (a-b) spatial map showing the percent of $F_{ESR}$, and $F_{RZSM}$ years between 1980 and 2018 in the GLEAM dataset, (c-d) contour plots showing the total number of (c) $F_{ESR}$ and (d) $F_{RZSM}$ events binned as a function of the climatological mean of yearly average precipitation (in mm/day/year) and PET (in mm/day/year) for the global 0.5x0.5° pixels, (e) contour plots showing the bivariate probability densities of FD intensity and (%) change in RZSM ($\delta RZSM$ (%)) calculated specific to all $F_{ESR}$ events between 1980 and 2018 over the arid, semi-arid, sub-humid, and humid regions, and (f) same as in (e) but for FD intensity and change in SESR calculated specific to all $F_{RZSM}$ events.308

Figure 2 Climate controls associated with onset and evolution of flash droughts (a) contour plots illustrating the total number of $F_{ESR}$ events binned as a function of standardized Pr anomalies (y-axis) for 1st, 2nd, 3rd, and 4th pentad before (denoted by -4 to -1 in the x-axis) and 1st, 2nd, and 3rd pentad after (denoted by 1 to 3 in the x-axis) the onset of FD events (denoted by 0 in the x-axis), (b-h) same as in (a) but for standardized anomalies of (b) Tmax, (c) VPD, and (d) pi, (e-h) same as in (a-d) but for the standardized anomalies corresponding to the $F_{RZSM}$ events. Spatial maps in (a-d) and (e-h) illustrate the mean of standardized anomalies of the climate variables for the 3rd and 1st pentad after the onset of the $F_{ESR}$ and $F_{RZSM}$ events.311

Figure 3 Relative importance of climate controls on flash drought intensity (a-d) bar-plots showing the order of relative importance of climate variables influencing the inter-annual variation of (a-b) $F_{ESR}$, and (c-d) $F_{RZSM}$ intensity for the different evaporation regimes. Note that both $F_{ESR}$ and $F_{RZSM}$ are calculated base on the GLEAM dataset, whereas the climate anomalies are calculated based on the (a, and c) ERA5 and (b, and d) MERRA2 datasets.314

Figure 4 Sensitivity of flash drought intensity to climate variables (a-d) line plots showing the sensitivity of $F_{ESR}$ and (e-h) $F_{RZSM}$ intensity to changes in standardized anomalies of climate variables based on ERA5 dataset and calculated for 0, 1, and 2 pentad lags (denoted as L0, L1, and L2) for the different evaporation regimes.317
List of Figures (Continued)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHAPTER 9</td>
<td></td>
</tr>
<tr>
<td>Figure 1 Compound and cascade model framework: (a) directed acyclic graph representing the dynamical system, and XY event network depicting the exposure and outcome variables (denoted by the binary sequency, $X_t Y_{t+T}$) and the association with the confounder, Z, (c) timeseries depicting association between global mean daily climatology of soil moisture (SM), maximum 2m air temperature (Tmax), and the confounder variables, total precipitation (Pr), total potential evaporation (PE), total actual evaporation (E), vapor pressure deficit (VPD), and surface net radiation (Rn) for 1980-2018 period.</td>
<td>338</td>
</tr>
<tr>
<td>Figure 2 Hotspots of Compound and Cascading Effect in Dry-Hot and Hot-Dry Event Network shown by spatial map of (a) maximum AF, (b) corresponding lags for the dry-hot cascade network, and (c-d) same as in (a-b) but for the hot-dry cascade network calculated for the period, 1980-2018. Note that only regions with a positive AF are shown in these spatial maps.</td>
<td>345</td>
</tr>
<tr>
<td>Figure 3 Influence of Confounders (a-h) Spatial distribution of statistically significant (at 5% significance level) odd ratios ($\exp(\beta)$) corresponding to standardized anomalies of VPD calculated by fitting the logistic regression model for the dry-hot event network for a lag of (a) 0 days, (b) 1 day, (c) 2 days, (d) 3 days, (e) 4 days, (f) 5 days, (g) 6 days, and (h) 7 days, and (i-p) same as in (a-h) but corresponding to standardized anomalies of ESR for the hot-dry event network.</td>
<td>349</td>
</tr>
<tr>
<td>Figure 4 Role of Background Aridity (a) Spatial map showing the classification of global evaporation regimes based on aridity index (AI) calculated as a ratio of mean annual precipitation and potential evaporation for 1980-2018 period, (b) Mean AF (%) binned as a function of AI and time-lags (0 to 7 days) for the dry-hot event network, and (c) same as in (b) but for the hot-dry event network.</td>
<td>352</td>
</tr>
</tbody>
</table>
CHAPTER ONE
INTRODUCTION TO COMPOUND AND CASCADING HYDROCLIMATIC EXTREME EVENTS

1. Definition of Compound and Cascading Events

The impact of extreme events is often challenging and compounding in nature. This is because, most of the time, the variables/drivers are mutually dependent and interact among themselves at different spatio-temporal resolutions in random order resulting in a series of natural hazards [1–5]. The combination of such multiple drivers and/or hazards contributes substantially to the societal and environmental risk and is generally defined as a compound and cascading extreme event [2]. A few meaningful definitions of compound events are available in previous literature. Some of the key definitions are provided below.

IPCC-SREX [6] defines compound events as “(1) two or more extreme events occurring simultaneously or successively, (2) combinations of extreme events with underlying conditions that amplify the impact of the events, or (3) combinations of events that are not themselves extremes but lead to an extreme event or impact when combined”.

Leonard et al., 2014 [2] proposed a more general definition which is, “a compound event is an extreme impact that depends on multiple statistically dependent variables or events”. This definition was proposed with the intention of eliminating ambiguity supposedly arising from the categorization in the SREX definition of compound events. Their definition focuses on the extremeness of the impact rather than the impact of
individual variables or events, the multiple variables on which the impact depends and the underlying factors or dependence that influence the severity of the impact [2]. For example, the severity of the 2010 Russian summer heatwave was exceptionally high [7] due to the pre-existing dry conditions during the first seven months of the year [8]. The combination of extreme dry and hot conditions contributed to widespread wildfires [9] with major societal and environmental consequences such as mortality [10], loss of crop, and massive air pollution in big cities like Moscow [11], thereby adding to the impact of heat wave [12]. In this case, the overall impact from the combination of drought, heatwave, and wildfire represents the compound extreme.

Another way of visualizing the compound extremes came from the idea of looking at the interaction relationships between the multiple hazards coming under various hazard groups (such as the geophysical, hydrological, near surface, atmospheric, biophysical, and space hazards) [13]. Keeping in mind the end user needs, Gill and Malamund, 2014 [13] synthesized and presented interactions between 21 natural hazards drawn from these hazard groups and identified 90 hazard interactions. Interactions involving a primary hazard triggering or increasing the likelihood of secondary hazards were considered in the study. Their study shows that a primary hazard trigger as well as increase the probability of a secondary hazard in 70% of the situation, can only trigger (but not increase the probability of) a secondary hazard in 17% of the situation, and can increase the probability of (but not trigger) a secondary hazard in 13% of the situation. Furthermore, it is observed that these interactions even have the potential to cause multiple hazard events.
Figure 1 Schematic diagram to represent risk from interconnected extreme events. This figure is reproduced from Zscheischler, et. al., 2018 [14].

To summarize, compound extreme definitions emphasize the interaction between multiple variables or hazards that increase the severity of their final impact. In some cases, even if one or more of the individual drivers or hazards are extreme, the combined effects may not be extreme. Therefore, the resulting event is not considered as a compound event. Relevant examples would be extreme precipitation in a coastal region accompanied by low ocean levels or extreme precipitation over a region with extreme shallow groundwater levels. In both cases, low-level flooding is expected and therefore, not considered as compound extremes. Similarly, there can be cases when the impact
from the multiple drivers or hazards may be extreme, although they are individually not extremes. In addition, the compound events can be represented as a combination of the Hazard types [13], such as climatic/hydrology events such as the flood, geophysical/technical events such as tsunami causing failure in a nuclear plant. Even climatic and non-climatic drivers interact with each other leading to hazards that affect the societal and environmental risk. The schematic illustrates this concept in Figure 1.

2. Example of Compound and Cascading Hydroclimatic Extremes

Specific examples of compound extremes are provided in the following section.

2.1. Compound and Cascading Drought and Heatwave

Simultaneous and sequential heatwave and drought episodes are among the most common compound, and cascading events studied extensively in many regions across the globe [15–18]. The compound and cascading effect of the extreme dry and hot conditions have been detrimental with consequences such as high mortality [10,19], loss of crop yield [20–24], and health hazards [25]. High precipitation variability induces a deficit in the soil moisture that gradually propagates into soil moisture drought. Due to the lack of moisture, cooling due to evaporation ceases to occur, which raises the sensible heat flux of the soil surface, thereby increasing the surface air temperature. This affects the surface energy portioning. The increase in the surface air temperature further increases the atmospheric demand for moisture, intensifying the drought (cascading effect). These land-atmospheric feedback processes are dominated by the surface energy budget and governed by a mutually reinforcing cycle often known as the soil-temperature coupling. They are very common in water-scarce or transitional regions [26].
Figure 2 Schematic diagram explaining the cascading effect of ENSO, PDO, and NAO on moisture inhibition and surface energy partitioning associated with compound drought and heatwaves.

Furthermore, these land-atmospheric feedback processes are enhanced mainly by the large-scale dynamics of the atmosphere such as the transient anticyclones [27]. A schematic diagram is presented in Figure 2 to illustrate the cascading effect of large-scale perturbations, such as El Nino Southern Oscillation (ENSO), Pacific Decadal Oscillation (PDO), and North Atlantic Oscillation (NAO) on moisture inhibition and surface energy partitioning associated with compound drought and heatwaves. Anticyclones are responsible for drastically decreasing the weather variability at a synoptic timescale. The lack of atmospheric moisture/humidity in these regimes is one of the main drivers responsible for the onset of meteorological drought. For example, anticyclonic weather
regime dominated over central Europe during the spring and summer of 2003, resulting in continued reduction of soil moisture content and elevated surface temperature driven by soil-temperature coupling [28,29].

Recent studies suggest a significant increase in simultaneous drought and heat wave events in regions such as western and southern US, northeastern China, western Europe, central Europe, Africa, southeastern Asia, southern India, northern South America, and eastern Australia [15,17,30–32]. Strong statistically significant relationship has been found between natural cycles like El Nino-Southern Oscillation (ENSO) and compound dry and hot extremes in various regions across the globe [27,33]. More interestingly, increased anthropogenic activities have also been cited as one of the key catalysts in the onset of hot and dry spells across the globe [17].

2.2. Compound Heat Extreme and Humidity

It is well established that the impacts from rise in temperature extremes are compounded with an increased atmospheric water vapor (or relative humidity) which slows down heat dissipation from human body, thereby leading to heat stress (HS) [34–36]. The extreme heat increases the body core temperature to dangerous levels and the added effect of increased humidity ceases evaporative cooling from the skin. Heat stress has led to massive human morbidity and mortality in recent years throughout the world [34,37,38]. Recently, many studies have reported the adverse effect of heat stress on public health and labor efficiency in the United States [39–41], much of which is attributed to the increased environmental exposure to such phenomena as an effect of
climate change [34,42–44]. Furthermore, extreme heat events in the Contiguous US (CONUS) have in recent decades been characterized by an earlier occurrence and increased severity, frequency, and areal extent. These trends appear connected to anthropogenic warming [45–48].

The HS is expected to become more severe due to an increase in frequency, duration, severity, and spatiotemporal extent of extreme warmth and moisture, which is an important, nonlinear, and often-neglected component. For example, severe heat stress (HS) events in the Midwest and the Gulf of Mexico coastal plains during the summers of 2019 and 2020, as well as many others during the 2010-19 decade, are representative of the types of extreme hot and humid events expected to become more common in the contiguous United States (CONUS) in future.

The US National Weather Service (NWS; https://www.weather.gov/) approximates apparent temperature (“feels-like” temperature) that translates the humidity effect with a metric, referred as heat index (HI). The HI successfully communicates risk associated with extreme heat stress events and as such, have been extensively used in previous studies [34,42]. In addition to that, most studies have communicated exposure to dangerous heat stress based on the operational definition of excessive heat warning recommended by the NWS, as characterized by HI value exceeding 105°F (or 40.6°C) [34,49] persisting for 2 hours or more (https://www.weather.gov/bgm/heat). The operational definition of heat warnings as recommended by NOAA’s National Weather Service is illustrated in Figure 3.
Figure 3 NOAA’s Heat Index values corresponding to relative humidity and temperature illustrating the heat warnings. This figure is downloaded from NOAA’s National weather services (https://www.weather.gov/phi/heat).

2.3. Compound precipitation and wind associated with Atmospheric Rivers

The concurrent occurrence of wind and precipitation extremes driven by atmospheric rivers (ARs) have significant societal impacts. The past few decades have witnessed a rapid intensification of extreme precipitation (EP) events triggered by atmospheric rivers (ARs) leading to massive flooding across the globe [50–55], including in the US [56–61]. ARs are long narrow corridors of strong poleward water vapor transport across the mid-latitudes [62–64]. They form a considerable part of the warm conveyor belt of extratropical cyclones and are often characterized by strong low-level winds and intense storm activities.
Figure 4 AR making landfall during UTC 1200, 3rd May 2010 in the southeastern US; IVT intensity (shading), and IVT vectors (arrows) are calculated using the ERA5 dataset.

Many global [65] and regional studies [57,58,66,67] have investigated various crucial aspects of landfalling ARs, including their climatology, contribution to snowpacks and other water resources [68], as drought busters [69] and flood producing drivers [70] across many parts of the globe. Numerous studies have adopted sophisticated algorithms and image processing tools to examine ARs and their associated precipitation [71–73]. The most robust metric used in the detection of ARs is the Integrated vapor transport (IVT) which is calculated in a Eulerian framework by performing a mass-weighted vertical integration within the pressure levels ranging from 1000 to 300 hPa as follows:
\[ IVT = \sqrt{\left( \frac{1}{g} \int_{0}^{300} qudp \right)^2 + \left( \frac{1}{g} \int_{0}^{300} qvdtp \right)^2} \] (1),

where \( g \) is the acceleration due to gravity (= 9.81 m s\(^{-2}\)), \( q \) is specific humidity (kg kg\(^{-1}\)), \( u \) is zonal wind (m/s), \( v \) is meridional wind (m/s), and \( dp \) is the pressure difference (Pa) between adjacent pressure levels. Figure 4 shows an example of AR making landfall in the southeastern US.

2.5. Flash Drought Caused by Compound Changes in Temperature and Precipitation

Drought is usually insidious that builds up gradually and persists for several months or years, producing a complex web of impacts that affects various sectors of the society [74,75]. The droughts are often investigated at more extended periods (e.g., multiyear period); however, more often than not droughts are accompanied by heatwaves that are characterized by unusually high temperatures sustaining over a few days to weeks [76]. Among the drought events, certain drought events are characterized by a sudden onset and rapid intensification within a few weeks, leaving the stakeholders with a limited amount of time to react and cope with the significant economic impacts. Such rapid intensification of drought is triggered by a complex interaction between elevated atmospheric demand, limited soil moisture, and high evaporation rates (ET) that often lead to the onset of rapidly intensifying moisture stress, a phenomenon termed as "flash drought" [77–80].

With the rise in global warming, the number of exceptional drought events has increased, which is likely to significantly impact socio-economic sectors, such as
agriculture, ecosystems, and energy production. Flash drought events are often difficult to predict, often resulting in devastating socio-economic impacts [81–84]. While studies based on regional observations have linked the onset and propagation of flash drought events to natural climate variability [85,86], future analysis has attributed possible amplification of such events to anthropogenic activities [87].

An abnormal decrease in precipitation, especially in a humid and sub-humid region, may lead to the rapid intensification of moisture stress leading to the onset of flash drought events [78,79,86]. Moreover, the unusual increase in temperature accelerates the terrestrial hydrological cycle that exacerbates atmospheric demand for evaporation. This favors the transition of a region from an energy-limited to a water-stressed regime, conditions favorable for the onset of flash droughts [88,89].

2.6. Wildfires under pre-existing hot and dry conditions:

The wildfire risk results from the combined effects of high precipitation variability leading to dry conditions, extreme temperatures, and below normal humidity. This type of compound events is very common in Mediterranean Europe [90], western US [91–94], and Australia [95,96]. During summer, these regions are generally characterized by transient anticyclones with clear sky, that enhances insolation and adiabatic warming of the surface air. These anticyclones are mostly slow moving or stationary (also known as atmospheric blocking) thereby allowing the heat wave to build gradually for a longer period of time [97] contributing to wildfire. The prolonged 2010 Russian heat wave is one such example that led to wildfire followed by massive air
pollution in the city of Moscow [11]. The compound impact of this heat wave, wildfire, and air pollution took more than 2000 lives [11,12].

2.7. Landslide due to precipitation after wildfire

High severity fires coupled with subsequent rainfall events has posed a major threat of landslides in the past [98]. Post-wildfire induced changes in the hydrological and physical properties of soil contributes to soil-hydraulics and composition [99]. The changes in the soil properties can modulate the watershed response, and sedimentation dynamics in the downstream region causing major losses such as mortality. Furthermore, due to fire suppression practices carried out in many regions, natural fire regimes have become more susceptible to fire of even greater intensity and coverage [100].

Post-fire ash and debris generated by wildfires gets deposited over the affected area or flown by wind in the surrounding regions. The deposited ash and debris contribute to the instability in the soil that makes it susceptible to landslides due to rainfall. Several post-wildfire landslides have occurred due to high as well as low rain-storms across many parts of the globe [101–103]. The 2003, 2007 and 2009 wildfire in the southern interiors of British Columbia triggered series of landslides in the region [101].

2.8. Marine Heat waves, acidification, and deoxygenation

Marine heat waves (MHWs) have shown profound ecological impacts by modulating the structure and dynamics of the ecosystem [104,105]. MHWs are generally caused by a combination of atmospheric and oceanographic processes [106]. MHWs can be defined as state of sustained (for days to months) and extremely high seas surface
temperatures (SSTs) over a coherent area [107]. A number of MHW has been reported since the beginning of the 21st century period including the 2003 northwestern Mediterranean [108], 2010-2011 west Australian [109], the northeast Pacific between 2013-2015 along the southern coast of California [110], and the 2016 high latitude MHW off the coast of Alaska [111].

The effect of MHWs coupled with the acidification and deoxygenation has been reported by model and observation-based studies [112]. Evidence suggests that the MHWs interact with the biogeochemical cycle affecting the biological processes, functioning and adaptation capabilities of the marine ecosystem [113]. Such impacts range from entire regime shift and composition of species resulting in harmful algal blooms, mass stranding mammals and mortalities of particular species [114]. For example, the 2011 western Australian MHW caused a massive geographical shift of the temperate reef ecosystem coupled with a reduction in the habitat of seaweeds and migration of tropical fish communities [115]. In addition, the Great Barrier Reef was reported to experience a massive (over 90%) coral bleaching event due to the 2016 MHWs triggered by global warming and 2015/2016 El Nino event [116,117].

2.9. High-Ozone Concentration Associated with Concurrent Heatwave and Stagnation

A recent study using the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem) was conducted to examine the present and future changes in ozone concentration as an impact of compound extreme weather events in the US [118]. The study revealed striking enhancement of higher ozone concentration due to the effect
of simultaneous heat wave and stationary circulation system characterized by low wind speed [118]. It is also suggested that during the future, it is most likely that the effect of compound events will lead to 10 to 13% increase in ozone concentration, compared to the present, with the most vulnerable nations being the US, Europe, and China where compound events are likely to intensify [118].

2.10. Concurrent daytime and nighttime heat wave

Climate change has resulted in an increased number of extreme diurnal heat events which is not only limited to the daytime hours but also observed during the night. Previous studies suggest a statistically significant rise in daily nighttime temperatures accompanied by a shift in the probability distribution of both daytime and nighttime temperatures during the last few decades [119–121]. Consequently, in a given time period, it is likely that a greater number of daytime extreme heat events overlap with the nighttime heat events over the calendar days. Those calendar days comprise of the concurrent daytime and nighttime heat wave events [119].

The compound impact of the daytime and nighttime heat events is alarming. It can be demonstrated in the context of human physiological limitations. The human body requires a break from the prolonged elevated daytime temperatures during the nighttime. However, if the night-time temperatures exceed the advisory danger limits, it prevents the body to acclimatize the body core temperature, thereby, exacerbating the pre-existing health disorders.
2.11. Flooding due to the Convolution of Tropical Cyclone, Precipitation, Storm Surge

Tropical cyclones [122,123] are one of the costliest natural hazards in the US [124–126], India and Bangladesh [127–129], UK [130,131], and Australia [132]. Flooding caused by the combination of wind and surge in tropical cyclones has a profound impact including substantial infrastructural damage, loss of property and lives [122,123,133]. In addition, tropical cyclone induced storm surge when accompanied with heavy precipitation at the same time can result in devastating floods in the coastal regions that habitats a major portion of the population [134].

2.12. Air pollution from combined drought and air chemistry

Climate change has the potential to increase the frequency and duration of drought across many parts of the world. Therefore, it is necessary to understand the impact of drought at a larger scale that includes interaction with other components of abnormalities in the atmospheric composition. For example, drought conditions have been known to have a major impact on the airborne fine dust pollution and public health risks [135–140]. Drought also influences the soil and vegetation health by hindering the movement of dusts [141], and reactive gases (such as the biogenic volatile organic compounds, or BVOCs, and NOx) from the surface into the upper atmosphere [142] as well as enhancing the dry deposition of gases and aerosols [143]. These modulations finally alters the chemical balance and the atmospheric lifetime of pollutants [140].

The compound effects of drought and atmospheric composition is further amplified by the increasing wildfire and anthropogenic activities [93]. However, Wang et
al., 2017 [140] reported that the increase in air pollution linked to drought is primarily driven by natural processes and has no significant influence from the anthropogenic emissions in the US. Observations from the study suggest that the rise in ground level ozone and PM2.5 concentration during the period 1990-2014 is primarily attributed to the combined effect of drought and air chemistry.

3. Concluding Remarks

Compound and cascading hydroclimatic extremes have severe to extreme socio-economic impact across the physical and human system, which is why quantification of these extremes are extremely important. This dissertation is focused on investigating, exploring, assessing, and devising nuanced techniques and their implementation in the realm of quantification of hydroclimatic extremes and their impacts. Specifically, four types of compound and cascading hydroclimatic extremes are studied in this thesis, (1) compound and cascading drought and heatwave, (2) combined heat and humidity (also known as heat stress), (3) extreme precipitation associated with ARs, and (4) flash droughts.

References


CHAPTER TWO

A PERSPECTIVE OF DROUGHT INDICES UNDER CLIMATE CHANGE

1. Overview of Drought Indices

Drought is an extreme climatic event that is insidious in nature because it develops slowly and often sneaks up on one [1]. As it gradually increases in intensity and duration, it can have major consequences, making it one of the costliest natural hazards [1]. Moreover, drought has multiple ecohydrological and socioeconomic impacts [2] including increased risk of wildfire [3], water scarcity [4], loss of crops [5] and livestock [6], increased food prices [7], migration [8], and indirect health effects [9]. The physical processes involved in drought are highly non-linear and involve feedbacks, and its impact propagates through multiple levels unequally that often cannot be quantified objectively [10]. Consequently, it is difficult to have a universal definition for drought [10].

However, drought definitions can be broadly categorized as either conceptual or operational [11]. Conceptual definition outlines the basic drought concepts with a general description of the physical processes involved, such as shortage of precipitation (meteorological drought), shortage of soil moisture (agricultural drought), shortage of water in lakes and streams (hydrological drought), and shortage of water for use by society related to water management [1, 12]. None of these are necessarily correct or wrong, and thus, all need to be recognized. On the other hand, operational definition focuses at identifying the onset, duration, and termination of drought episodes including their severity [1, 12]. Operational definitions aim at providing precise drought-related
information to support an effective early warning system [12]. Apart from the above definitions, a legal definition of drought is also available [13]. In addition to the effect of drought being context dependent, drought definitions such as that of operational drought [4] and socioeconomic drought [1, 11] are also in existence. Generalized definition of drought can be developed only through the aggregation of process-specific instantaneous droughts [10]. But this definition assumes that these processes are in equilibrium with the long-term climate, thereby overlooking the distinction between drought and water scarcity [10, 14]. Thus, numerous and diverse disciplines adopt different drought definitions depending on the stakeholder’s need as well as hydroclimatic variables included [1, 12].

Consistency among these drought definitions is a key to remove any ambiguity in framing drought policies and making decisions. The corresponding decision support tools rely on indicators and indices that are widely used to quantify the physical characteristics of drought (intensity, duration, and severity) [15]. Drought indicators and drought indices are formulated to track the hydrological cycles and are used interchangeably in drought-monitoring community [16]. Drought indicators are used in a broader sense that aggregate parameters such as precipitation, temperature, streamflow, groundwater levels, reservoir levels, snowpacks, soil moisture levels, and drought indices [16]. On the other hand, drought indices are single numeric values estimated from various hydroclimatic variables that influence drought and, therefore, it has a significant advantage over mere raw data in quantifying drought characteristics [16].
Drought assessment studies have made considerable progress so far in developing several drought indices applicable to various types of drought [1], such as Standardized Precipitation Index (SPI) for meteorological drought [17], Standardized Runoff Index (SRI) for hydrological drought [11, 18], and soil moisture percentiles for agricultural drought [1, 19]. However, the development and choice of drought indices should be specific to the primitive as well as newly emerging real-world problems and, therefore, it depends on several factors [1]. The following section provides an overview of some of the critical factors associated with formulation of drought indices:

1. **Types of drought**: The interconnection between various types of drought that occur simultaneously or sequentially makes it difficult to distinguish between one drought type from the other [20]. For example, the propagation of meteorological drought (which is caused mainly by precipitation deficit) to agricultural (caused by soil moisture deficit) and hydrological (deficit in water storage or streamflow) drought is non-linear in nature [21, 22]. In addition, the impact of meteorological drought shifts prominently towards soil moisture (agricultural drought) that further propagates to cause water storage deficits (hydrological drought) for even longer durations (Fig. 1) [23]. This complicates the formulation of drought indices with a view of quantifying a specific type of drought independent of the others.

2. **Drought characteristic**: Drought events have multiple and interrelated characteristics such as severity, duration, peak intensity, and recurrence interval [24, 25]. Each of these
characteristics may have a considerable influence on the impacts of drought. Consequently, monitoring natural and socioeconomic drought needs a joint assessment of individual drought characteristics as well as identifying the most dominant drought event specific to the impact being studied [21]. Moreover, in arid regions that naturally receive scanty or no rainfall, thereby always at the verge for water shortage, drought characteristics estimated in relative terms and absolute terms will be significantly different. In other words, the climatology of a specific region can influence drought characteristics significantly, especially if drought is measured in terms of anomalies.

3. Climate change: Impact of drought under a global warming scenario is more likely to aggravate in the future [26, 27]. Of course, droughts have always occurred, and the variability in sea surface temperature anomalies can cause global droughts [28, 29]. In addition, a change in regional climate such as slow-moving anticyclones that alters the climatology of a region by hindering the progress of synoptic weather systems can be responsible for enhancing the land-atmosphere feedback processes [30,31]. Due to the lack of available moisture in these regimes, the land-atmosphere feedback processes exacerbate the situation by increasing atmospheric temperatures and thus increasing the atmospheric demand for moisture, thereby leading to increased drying and heating of land surface at the same time, the impact of which is often alarming (such as wildfire risk) [3]. Figure 1 shows the connection between those processes that affects the propagation of drought under climate change.
Thus, underpinning the mechanisms behind such processes is relevant to formulate reliable drought indices that should incorporate all such participating processes, including the various human contributions that influence the drought characteristics and socioeconomic conditions [21]. In addition, the non-stationarity [32] in future climate may lead to large uncertainties in quantifying droughts [33]. Therefore, drought indices need to be robust and revised by including the nonstationary climate information.

4. The distinction between water scarcity and drought: Water scarcity and drought have separate implications [14]. Unsustainable use of water resources can lead to water scarcity and, therefore, can be controlled, while drought is a natural hazard, and its impacts can only be mitigated by adapting to the climate variability with prior measures [14]. In arid or semi-arid regions, dry conditions quickly lead to water scarcity, and this example emphasizes that the background climatology is also a factor. Hence, in water-scarce or arid regions, where drought and water scarcity usually occur simultaneously, drought situations are more severe and further aggravate water scarcity [12, 14]. Consequently, in such regions, the choice of a suitable indicator that makes a clear distinction between drought and water scarcity is necessary in making effective water management decisions [14].

5. Multivariate aspects of drought: Drought is influenced by multiple hydroclimatic variables such as precipitation, runoff, potential evapotranspiration (PET), and soil
moisture [1]. Thus, a single drought indicator may be insufficient to quantify drought and, therefore, such assessment requires drought indicators that blend more than one drought index or drought affecting variables [2].

The overall objective of this chapter is to provide a perspective on drought indices under climate change scenarios. The section “Relevance of Drought Indices in Climate

Fig. 1 Schematic diagram showing the drought propagation under climate change. (Note: this figure was revised with respect to original drought propagation concept proposed by Wilhite [12])
Change” presents the relevance of drought indices in climate change assessment, followed by a discussion on application and limitations of existing drought indices in the section “Application and Limitations of Existing Drought Indices”. The section “Challenges Associated With Drought Indices in Climate Change Studies” provides an overview of challenges associated with drought indices for climate change studies, and summary and conclusion are provided in the section “Summary and Conclusions”.

2. Relevance of Drought Indices in Climate Change

A number of drought indices have been developed to quantify a drought [1]. Most of the drought indices use either only precipitation or in combination with other meteorological variables. Also, numerous studies have investigated the effect of climate change on drying of global terrestrial surfaces. However, most of the studies on dryness fail to consider the background aridity [34–36] and thereby fail to incorporate the changes in available energy, air humidity, and wind speed [34]. Failure to account for such variables in formulating drought indices may lead to a spurious increase in drought under warming climate [34]. Therefore, instead of only considering contemporaneous anomalies to derive drought indices, it is important to also consider the factors that govern the background state [34]. On the other hand, it is evident that climate change-induced warming has accelerated hydrological processes, firstly, by increasing the energy available for evapotranspiration (ET) and, secondly, by increasing temperatures and thus the water holding capacity of the atmosphere [37]. Consequently, it results in more
intense, widespread, and persistent extreme climatic events like droughts. Therefore, temperature is likely to be an important variable for deriving appropriate drought indices under global warming. The following section provides an overview on the importance of temperature and anthropogenic forcings for drought assessment, followed by an example highlighting the role of drought under global warming.

Reconstructions [38] and instrumental observations [39] demonstrate that the Earth’s surface temperature has increased substantially over the past century, and by the end of the twenty-first century (2081–2100), it is expected to exceed the desirable limits of 1.5 and 2 °C above the pre-industrial level (1850–1900) [40]. Consequently, the intensity of precipitation has increased substantially, because as regulated by the Clausius-Clapeyron (C-C) relationship, there is an increase in atmospheric moisture holding capacity of approximately 7% per °C rise in temperature [41]. However, the surface energy available increases at a much slower rate and this governs the total precipitation amount through the availability of moisture [42]. Hence, there is also a considerable increase in longer dry periods [43]. Except for tropical hurricanes (characterized with large water vapor content), the troposphere is able to radiate away the energy released by condensed precipitation, and the distribution of relative humidity mostly remains relatively constant in both lower and higher latitudes under climate change [42]. Under such conditions, changes in mean precipitation depend on the water availability over both ocean and land surfaces [43]. However, land areas away from the oceans lack the adequate moisture supply to meet the evaporative demand of the atmosphere, leading to continental drying, high temperatures, and lower relative
humidity, as found in the model projections [35, 44]. Moreover, as the ocean surface tends to warm at a slower rate than the land and the atmosphere [35, 44], there is a considerable delay in the recharge process of the atmospheric moisture to finally reach the saturation level (necessary for precipitation), thereby resulting in longer dry periods over land [45]. Longer dry spells have direct influence in initiating long-term and severe droughts [46].

The extra heat due to global warming has accelerated the drying process in the recent past [27, 47], which is likely to cause more severe, persistent, and widespread droughts in the future with respect to the current climate [48, 49]. Furthermore, increases in severity of drought in future climates could be largely caused by the mean state change in the warming world. Previous studies have investigated the mean state aridity change due to global warming in terms of an aridity index defined by the ratio of annual precipitation to annual PET [34, 36, 50], and it is also shown that terrestrial climate would become drier as the Earth warms [34, 35], which leads to the expansion of the world drylands [36]. Furthermore, the anticyclonic regimes commonly present in setting up a drought are characterized by weather patterns that steer precipitating weather systems elsewhere and create a stable atmosphere that shuts down local convection. Hence, once the weather conditions are favorable for drought, climate change exacerbates the problem by adding small amounts of heat that accumulate over time, increasing temperatures and ET (drying) [26, 30]. Furthermore, due to limited moisture availability over land, such climate regimes experience a considerable rise in sensible heat fluxes (due to the absence of cooling by evaporation) during limited supply of latent energy
fluxes (due to soil moisture depletion), thereby further raising the land surface temperature [37, 51]. This coupling effect between soil moisture and temperature is commonly referred as soil-temperature coupling [31, 37] and can be a potential stressor for wildfire risk [52]. Observational studies confirm relationships between surface moisture deficit (leading to preceding drought conditions) and hot extremes in regional [53] as well as global [54] scales. Moreover, it is observed that higher correlation between warmer and dry conditions can increase the likelihood of concurrent heat and drought events [55]. Therefore, owing to the increasing exposure of heat events [56, 57], the compound effect of heat wave and drought will more likely have severe impacts in the future. Thus, temperature that directly controls evaporation and ET should be considered as an important contributor to drought events under the global warming scenarios [58]. Existing and popular indices used in drought studies under climate change incorporate the atmospheric demand (Standardized Precipitation Evapotranspiration Index (SPEI)) [59] and temperature effect with a crude approximation of potential evapotranspiration (Palmer Drought Severity Index (PDSI)) [60].

Interestingly, drought events during the last few decades, as well as projected in the future, are less likely to be comparable to the medieval droughts due to induced warming from greenhouse gas emissions, land cover, and land use changes from anthropogenic contributions [27, 49, 61, 62]. One such evidence of anthropogenic influence is the warming of the Indian Ocean that, coupled with the increase in sea surface temperature anomalies, caused the unprecedented Sahelian drought during the late twentieth century [63, 64]. Also, observed records indicate increased severity and
frequency of droughts over California during the past two decades related to anthropogenic warming [65, 66]. It is reported that early runoff due to early melting of snowpack in the region has affected the moisture content from the top-soil layer, thereby exacerbating hydrological drought during the summer [65]. Furthermore, anthropogenic contribution to recent and projected increase in drying trends in Syria has been reported by Kelley et al. [67]. The increasing and long-term drying trend has been attributed to the changes in precipitation driven by the increase in mean sea level pressure together with the long-term increase in warming over the Eastern Mediterranean Region for which no natural cause is apparent [67]. This is well supported by the positive response of the long-term drying to the increase in greenhouse gas emission-based on the model simulations that correlates well with the twentieth-century-observed precipitation trends in the Mediterranean Region [67, 68]. The combined effect of climate change on increased drying and land use changes has aggravated the drought impact in the region [69], causing migration of as many as 1.5 million people from rural to urban areas that contributed to the onset of Syrian civil war [67, 70].

Thus, drought quantification cannot be fully understood only based on the natural variability of climate as anthropogenic influence also plays a significant role in triggering as well as propagating drought events [1, 71]. Consequently, efforts have been made based on the existing climate models to detect anthropogenic contributions and attribute its influence on various climate extremes, including drought [71, 72]. In addition, the increase in population density further aggravates the human component influencing
drought [1]. For example, due to increased land use in overpopulated regions, runoff has increased substantially, thereby leaving little water to percolate into the soil [73]. Together with an increase in water demand for domestic [74], agricultural, and energy [75] sectors in highly populated regions, drought can pose a significant potential threat in the future. Therefore, a realistic assessment of drought also needs to incorporate such effects arising from the increase in anthropogenic influences.

From the above discussion, it can be inferred that variable associated with temperature (e.g., PET) play an important role in triggering droughts (dry spells); therefore, it must be considered in deriving drought indices for climate change assessment. In addition, the uncertainty associated with projected temperature is comparatively less with respect to precipitation based on the global climate model (GCM) outputs. In addition to temperature, other variables, such as precipitation, infiltration loss, and runoff, also significantly contribute to the occurrence of drought [26]. Drought indices and indicators should assimilate all these factors to quantify drought characteristics in the context of non-stationary climate [26].

3. Example of the Association Between Drought Indices and Land Surface Warming

In this section, we investigate the association between global warming and droughts. The self-calibrated PDSI (PDSI_sc) [60, 76] was selected for our analysis, as it is based on the physical water-balance and it incorporates the effects of precipitation, temperature, PET, and runoff. The PET is best estimated based on the Penman-Monteith (PM) method [77, 78] instead of the simple Thornthwaite (TH) method [79] that leads to
overestimation of drying in energy-limited areas [26]. The PDSI_sc can successfully capture long-term changes in drought with response to global warming, and it has been used in previous studies related to large-scale drying trends [26, 48, 80].

Therefore, we analyze the long-term temporal changes in drought using PDSI_sc as a measure of dryness to investigate whether overland droughts (drying) go hand-in-hand with rise in land surface air temperature (LSAT) by using the historical period (Fig. 2). Because a steady and sharp rise in global tropospheric temperature has been experienced since the mid-1970s [63, 81], this analysis is focused from 1975 onwards. Our analysis is based on continental averages; however, it is important to note that land-atmosphere feedback processes, which have major influence on drought, can be more accurately explained at finer scales.

We obtained a global gridded monthly observed PDSI_sc dataset [76, 82] (1850–2014) available at 2.5° resolution. Observed monthly LSAT was obtained from the updated CRUTEM4 dataset (1850–2017) at 0.5° resolution, as developed by the UK Meteorological Office Hadley Centre and the Climatic Research Unit at the University of East Anglia [83]. The gridded LSAT and PDSI_sc data are spatially averaged over the six continents. Finally, anomalies in continental mean annual LSAT are estimated for the period 1975–2014 with respect to the reference period 1961–1990.

The relationship between drying and rise in LSAT is well accounted by the negative correlation magnitude observed for all the six continents, with relatively stronger correlation for South America (−0.68), Africa (−0.48), and Australia (−0.48) (Fig. 2). Furthermore, the results from Fig. 2 clearly indicate that drought indices, such as
PDSI_sc, possess the required skill to capture the severe drying patterns due to increased loss of soil moisture by overland evaporation. However, there are limitations to the assessment of drought characteristics based on drought indices, as discussed in the following section.

4. Application and Limitations of Existing Drought Indices

Drought indices have evolved considerably through recent decades, keeping up with the evolution of drought itself under the changing climate. This section provides an overview of commonly used drought indices along with their limitations and skill to adapt to the climate change.

1. Palmer Drought Severity Index: PDSI was originally developed by Palmer [60] and is based on the primitive soil water balance that considers precipitation, runoff, and evaporative demand for a specific region. Nevertheless, the calibration period has a strong influence on the PDSI values, and it can be a limitation for its use in areas other than used for the calibration [84]. Guttman [84] showed that PDSI, being an autoregressive process, inherits a long-term memory owing to the temporal effect of the soil and atmospheric moisture conditions. Further scope of improvement in PDSI remains in the context of other shortcomings such as (i) fixed temporal scale and inherent autoregressive characteristic of PDSI over water-stressed regions [85], (ii) an inherent timescale that makes PDSI unsuitable for hydrological droughts [1], (iii) assumptions that any form of precipitation as rain leads to ambiguity in the application of PDSI in winter months and at high elevations [1]. For example, Sheffield et al. [86] found a marked difference in drought characteristics based on model-simulated and PDSI datasets over
the snow-dominated regions, which is attributed to the inadequate representation of winter processes in the calculation of PDSI; (iv) PDSI also inherits a negative bias in runoff estimations by assuming that runoff occurs only after all the soil layers are saturated [1], and (v) PDSI suffers from a considerable time lag in identifying developing and diminishing droughts [87].

Fig. 2 (LEFT) Anomaly in spatially averaged yearly observed LSAT for the period, 1975-2014 with respect to the period, 1961-1990 (bar plot with positive (red) and negative (steelblue) anomaly) and spatially averaged annual PDSI_sc (line; orange) for the period 1975-2014, and (RIGHT) scatterplot (violet) and regression line (blue) of annually averaged PDSI_sc and LSAT Anomaly for (a) Africa, (b) Asia, (c) Australia, (d)
Europe, (e) North America, and (f) South America. It is to be noted here that monthly PDSI_sc values are annually averaged and then correlation coefficients are estimated against anomaly based on yearly observed LSAT.

Moreover, Palmer [60] used an empirical approach and averaged the climatic characteristics and duration factors in the estimation of PDSI over very few regions, which limits the comparison of PDSI values among diverse climatological regions [88]. Overall, it can be said that PDSI is a relative measure of drought and the methods adopted to calibrate it are based on the previous climate scenario which is no longer valid in the context of the continuously changing climate [26,48]. To overcome this spatial inconsistency in PDSI, Wells et. al. [76] proposed PDSI_sc that self-calibrates (sc) the index at any location automatically by replacing the empirical climate characteristics and duration factors with dynamically derived values based on the historical climate data of that region.

Further improvement in PDSI has been made by replacing the TH [79] method with the PM [77, 78] method in the calculation of PET. PET based on the TH method [79] neglects climate variables such as solar and longwave radiations, humidity, and wind speed which affect the rate of moisture loss from the upper soil layers [26]. This leads to overestimation of drying in energy-limited areas [26]. The PM method [77, 78] can overcome these limitations for the estimation of PET. As a result, the self-calibrated PDSI based on the PM method (scPDSIpm) can be more appropriate to estimate large-scale changes in droughts (mainly agricultural droughts) in the context of global warming.
More recently, few other challenges have emerged, associated with the estimation of PET, as discussed in the section “Sparse Availability of Precipitation Data”. However, it is also important to note that PDSI actually tries to incorporate ET along with runoff, soil recharge, and moisture using precipitation, temperature, and available soil water data [60]. Despite several criticisms, PDSI gives a complete picture of the water cycle and remains as one of the most comprehensive drought indices [89]. Overall, PDSI_sc is a readily available standardized drought index and it can successfully capture long-term relative drying patterns in response to global warming [27, 48, 80, 82, 90].

2. Standardized Precipitation Index: The SPI [17] is one of the most popular indices used mainly to quantify meteorological drought. The SPI is based on a probabilistic approach, its estimation only requires precipitation data, and it is relatively easy to calculate. Nevertheless, exclusion of temperature, PET, wind speed, and soil moisture data as an input variable is a major limitation for generating reliable drought information under the warming climate [1, 59, 91].

3. Reconnaissance Drought Index (RDI) and Standardized Precipitation Evapotranspiration Index: (i) The RDI [92] is an improvement over the SPI, and it includes PET as one of the key variables. However, PET assesses the atmospheric demand for water but does not necessarily relate to ET because it needs to also assess the water availability. The RDI was used for drought monitoring and climate change impact assessment on water resources [93]. The RDI for a given time-period is estimated as a ratio between accumulated precipitation and PET [92]. However, the RDI lacks the ability to capture the variability of drought effectively with respect to change in
temperature [59]. Application of RDI may not be suitable when PET is equal to zero [59].

(ii) The SPEI [59] provides a relatively flexible approach that captures the combined effect of precipitation and PET [59]. Moreover, the SPEI performs adequately by considering equal sensitivity to precipitation and ET$_{\text{ref}}$ [94]. However, the SPEI may have few limitations in the case of comparing drought events between different climatic regions. For example, in semi-arid regions, the SPEI may be more sensitive towards the ET$_{\text{ref}}$, while in humid regions, it shows more sensitivity to precipitation [94]. Moreover, unlike the PDSI, the SPEI is not based on the water budget framework and fails to incorporate the soil moisture component for identifying agricultural droughts [59].

4. Multivariate Drought Index (MDI): MDIs are combinations of multiple hydroclimatic variables or drought indicators [95] that provide an alternative way to capture multiple aspects of drought conditions for efficient drought monitoring and early warning [96, 97]. Some of them can be listed as follows:

(a) Objective Blend of Drought Indicators (OBDI): Svoboda et al. [98] proposed OBDI based on the linear-weighted average of multiple drought indices.

(b) Aggregated Drought Index (ADI): The ADI [99] is constructed separately for each month using drought-affecting variables such as precipitation, streamflow, PET, reservoir storage, soil moisture, and snow water content. Principal component analysis is used to find the dominant hydrological signals corresponding to each drought type (meteorological, hydrological, and agricultural) [99]. However, PCA has limitations such as assumption of linearity in data transformation and dimensional reduction in the direction based on maximum variance.
(c) *Joint Drought Index* (JDI): The JDI [100] considers joint probabilities of precipitation and streamflow using multivariate probability distribution (e.g., copula).

(d) *Multivariate Standardized Drought Index* (MSDI): The MSDI [101] is introduced as a joint distribution of precipitation and soil moisture using a copula. Nevertheless, a copula has limitations such as lack of its ability to model high-dimensional dependence structure.

(e) Rajsekhar et al. [2] proposed the Multivariate Drought Index that uses kernel entropy component analysis (KECA) and incorporates variables such as precipitation, runoff, PET, and soil moisture. This index allows the user to extract higher information related to drought characteristics based on higher magnitude of entropy value [2]. However, soil moisture data are subjected to large uncertainties, and this reduces the confidence in the application of these indices.

5. *Relative Drought Indices*: Drought indices such as relative SPI (rSPI) and relative PDSI (rPDSI) are developed with an aim to provide an improvement in drought assessment under the non-stationary climate by providing an alternative way to compare drought between two or more time periods as well as between two or more stations. The former is achieved when drought indices are calibrated using aggregated observational data from all the stations based on a given reference period and then applied to future climate. This method can be applied to estimate the spatial shift of drought due to climate change [102]. On the other hand, the latter method is based on observational data from a given station, thereby allowing the user to capture the temporal changes of drought in the future with respect to the present climate [102]. However, the indices derived using the
second methodology may have shortcomings such as lack of comparability between different climate regions [102].

5. Challenges Associated with Drought Indices in Climate Change Studies

Although drought indices are useful to study climate change impact assessment, the following section discusses major challenges and limitations for such studies.

(a) Disagreement Among Drought Indices

The global mean temperature indirectly reflects the evaporative demand of the atmosphere in the absence of adequate moisture. Therefore, we estimated and compared the sensitivity of the abovementioned drought indices (SPI, SPEI, and PDSI_sc) with respect to rise in global mean temperature. The drought indices based on a shorter temporal window of 1 month were selected and derived for the entire globe: (i) SPI-1 was generated using precipitation dataset provided by the Global Precipitation Climatology Centre (GPCC) [103] (http://gpcc.dwd.de/) at 0.5° resolution, (ii) SPEI-1 data is downloaded at 0.5° resolution from Global SPEI dataset (available at http://spei.csic.es/database.html). This SPEI dataset is based on monthly precipitation and PET data available at the Climate Research Unit of the University of East Anglia that uses CRU TS version 3.23 dataset [104] (https://www.esrl.noaa.gov/psd/data/gridded/data.gpcc.html). The PET is estimated using the Penman-Monteith method [77, 78], and (iii) we use the same monthly dataset for PDSI_sc [76, 82] as in the previous analysis for Fig. 2.
The global gridded datasets of SPI-1, SPEI-1, and PDSI_sc are spatially averaged to generate time series at monthly scale. We analyze the sensitivity of mean annual drought indices at continental scale with respect to the change in the corresponding observed annual mean LSAT for the period 1901–2013 (Fig. 3). The LSAT is averaged at annual scale for estimating the anomalies so that any seasonal influence in the analysis is avoided. To perform the sensitivity analysis, we organized the magnitude of drought indices in temperature increments corresponding to temperature anomalies nearest to every 0.25 °C change in global mean temperature ranging from − 0.5 to 0.75 °C.
Fig. 3 Sensitivity of drought indices with change in LSAT for the six continents. a Africa. b Asia. c Australia. d Europe. e North America. f South America. Box plot showing median, interquartile range (IQR), outliers, and overall range excluding the outliers for the annual mean of continental averaged drought indices, SPI-1 (red); SPEI-1(green); and monthly PDSI_sc (blue) for every 0.25 °C change in LSAT during the period 1901–2013. To estimate the statistics related to box plot, values of drought indices are accumulated in bins corresponding to temperature anomalies nearest to every 0.25 °C change in global mean temperature ranging from −0.5 to 0.75 °C.

Figure 3 shows the box plots corresponding to an incremental change in temperature. We selected a shorter temporal scale that allows the drought indices to capture the influence of the warming on the loss of soil moisture leading to drying more effectively [1]. It can be noted that as compared to PDSI_sc, SPI-1 and SPEI-1 show a little change with rise in overland warming (Fig. 3). This may not be surprising as the SPI does not incorporate temperature or related variables as an input. On the other hand, the SPEI lacks the ability to produce comparable results between different climate regimes subjected to long-term drying [94]. Furthermore, the SPEI does not include the soil moisture information and, therefore, does not respond to the soil moisture drought adequately during the historic period [59].

However, PDSI_sc captures a consistent increase in drying with the rise in temperature (Fig. 3). This may be due to its ability to capture long-term droughts by incorporating soil moisture deficit or surplus from the previous months [105]. Thus,
while one drought index responds to the long-term drying with rise in temperature effectively, the other two indices seem to behave differently. This can be a major limitation among drought indices to adequately detect climate change impacts on drought characteristics under various climate regimes and temporal scale. Thus, apart from the disparity in defining drought objectively [10], drought indices can arrive at different results that leads to ambiguity in the decision or policy-making process related to impact assessment under climate change.

(b) Sparse Availability of Precipitation Data

It has been shown that sparse and poor quality of precipitation data [106] generate large uncertainties in quantifying spatiotemporal drought assessment under climate change [26, 27, 80, 107]. For example, Sheffield et al. [80] underestimated long-term drying based on PDSI_sc using NCEP/NCAR reanalysis data from four different precipitation datasets (CRUTS 3.10, DaiP, GPCC V4, and WilP). Out of these four products, CRUTS 3.10 has a poor spatial coverage since 1990 [107]. In other words, datasets based on poor gauge coverage can produce substantial uncertainty when gaps are filled with data from different sources (e.g., neighboring grid points) based on some climatology statistics [107]. Therefore, the compound effect of uncertainties in estimating topographical variables [108], coarse resolution of climate model outputs [109], and poor quality of precipitation dataset can generate large uncertainties in the calculation of drought indices.

Moreover, hydrological drought prediction requires high quality data to improve initial hydrological conditions based on which future droughts are estimated. Data
assimilation (DA) that merges observation (in situ or remotely sensed) with model output overcomes such limitation on data availability, and it improves the accuracy of drought prediction by providing accurate initial conditions [110]. Various Land Data Assimilation Systems (LDASs) have been developed so far, some of them are discussed as applicable to drought-related studies: the North America Land Data Assimilation System (NLDAS) [111], Global Land Data Assimilation System (GLDAS) [112], and Coupled Land and Vegetation Data Assimilation System (CLVDAS) [113] that can improve drought assessment under climate change. The land surface models (LSMs) provide improved parameterizations for seasonal and diurnal simulations of water fluxes, energy fluxes, and state variables that are essential for monitoring agricultural and hydrological droughts at hourly and daily timescales [114]. In addition, LSMs such as NLDAS-2 provide soil moisture for various depths and surface and sub-surface runoff data that enhance the accuracy to estimate agricultural and hydrological drought over North America, respectively. For example, top 2-m soil moisture anomaly can be indicative of agricultural drought, whereas the total runoff can indicate hydrological drought [114]. However, land surface models are still undergoing improvement in the applied physics to the horizontal and vertical distribution of soil hydraulic properties, incorporation of sink holes, and representation of the spatiotemporal distribution of precipitation [115].

(c) Estimation of PET

PET refers to the atmospheric evaporative demand and is extensively used in drought studies as a direct measure of relative dryness [48, 107] or as an input variable in the estimation of PDSI [60], RDI [92], and SPEI [59]. However, the selection of model
used to estimate PET is crucial in the reliable assessment of drying under the changing climate. For instance, the temperature-based model derived based on historical records to estimate PET is unlikely to reproduce reliable PET during the late twenty-first century. In other words, under the warming climate scenario, purely temperature-based models (TH method [79]) are likely to overestimate drying in the future climate [116]. Thus, climate variables such as radiation, wind speed, vapor pressure deficit, and humidity need to be considered. Consequently, the PM method that takes into account all of these climate variables is found to be more robust in the estimation of PET compared to other existing methods and has been extensively used in the context of studying the temporal and spatial variability of drought in the twenty-first century [48, 117]. However, large uncertainties can be seen due to the lack of reliable forcing data to calculate scPDSIpm [27, 107]. For example, changing cloud cover that controls the incoming solar radiation and wind speed variations that effect the rate of ET are more region specific [107, 118]. Along with spatiotemporal inhomogeneity of forcing data, these variables can trigger uncertainties in the global-scale assessment of drought under climate change [26, 27]. There are conflicting views if estimated drying under climate change will be significantly different, depending on whether precipitation or PET is used as the drought variable [119]. In addition to that, under high CO2 conditions, plants actually become more efficient and the resulting water savings that plants experience keeps higher amounts of water on land on average—i.e., the conventional drought indices might not account for this, leading to an overestimation of drought severity [120]. Furthermore, Milly and Dunne [121] reported discrepancies in the estimation of the change in PET that leads to bias in
continental drying trends. It is primarily attributed to the fact that stomatal conductance is not included as an input while estimating PET, and also due to the parameterization of sensible heat flux in terms of the gradient of potential temperature rather than temperature [121]. To avoid such discrepancies, an alternative method to estimate PET using the energy-based approach is proposed [121]. The proposed method assumes that long-term latent heat flux of PET is equal to the net radiation absorbed at the land surface [121]. However, the robustness of this approach requires more investigation and validation.

(d) Downscaling of Meteorological Variables

Temperature and precipitation are the primary meteorological variables of the hydrological processes [1] with higher uncertainty associated with precipitation in terms of its spatiotemporal distribution. Therefore, there is a need to develop robust downscaling methods to generate rainfall information at finer resolution to minimize the associated uncertainties [109, 122]. Consequently, drought indices derived from precipitation require effective downscaling techniques that can resolve discrepancies arising from scale issues [123], thereby helping the stakeholders to improve decision making [1]. However, drought assessments using GCM outputs are limited owing to considerable high bias associated with the precipitation estimates [124, 125], in addition to substantial intrinsic uncertainties originating from the inter-model variability [126–128]. This can be partly overcome by adopting simultaneous bias correction and spatial downscaling approaches [129]. In addition, GCMs do not exhibit a high degree of predictability especially over the extra-tropics owing to the limited physical
understanding of the ocean-atmosphere interactions in those regions [130]. This sets a major limitation to specify initial conditions for meteorological drought prediction [131].

Moreover, downscaling techniques face multiple challenges [132]. For example, (i) when a change factor method [122, 133] is applied to the coarser GCM outputs, it fails to include the local climate features, while transferring the relative change in signals directly to the scaled historical dataset. This limits the capability of this method to represent the change in climate, including time of occurrence and periodicity of events (such as drought) [132]. (ii) Although statistical downscaling methods are simpler than dynamical downscaling methods in terms of methodology and computational resources, it has considerable limitations. For example, statistical downscaling is done for each variable at individual grid points, thereby incorporating bias when applied to several variables or to several locations within one region [134]. In addition, the assumption of stationarity in the present climate will also be valid under the future climate scenarios, which implies overconfidence on the GCM’s ability to simulate the future climate variables (especially rainfall) [132]. Whereas (iii) a dynamical downscaling method clearly ignores the upscaled information from the local scale (sub-grid cells of GCMs) to the coarser grid cells and considers only one-way mode of transferring information (i.e., from the GCM to the nested RCM). Thus, the large-scale climate characteristics influenced by the local climate patterns may not be captured in the downscaled product [135].

Another challenge in downscaling methods arises from the lack of adequate hydrometric data in different parts of the world, especially developing countries [136].
However, with the advancement in DA techniques and land surface models, it is now possible to generate hydrological fluxes at finer resolutions [137]. For example, LDAS (https://ldas.gsfc.nasa.gov/) incorporates the high-resolution vegetation and soil coverages, and provides merged data products at 0.25° resolution and 0.125° resolution for global and regional (across central North America) analyses, respectively. Within this framework, the GLDAS [112] provides high-quality global land surface fields (implementing snow cover, water equivalent, soil moisture, surface temperature, and leaf area index) at 1° and 0.25° resolution from 1979 onwards that support several present and future climate predictions for various types of water resources applications.

(e) Choice of Baseline Period

The choice of baseline period plays an important role when comparing future drought under climate change with respect to historical drought as the reference period [27, 80, 90, 107, 138]. It is well known that by considering a longer (entire) period as the base period, the drought indices can be better calibrated and the future drought events can be compared with appropriate historical droughts [60, 90, 107]. However, the improper choice of base period with respect to which drought events are evaluated can produce considerable bias in the drought assessment under climate change. For example, Sheffield et al. [80] and Dai [48] used two different baseline (historical) periods (1950–2008 and 1950–1979, respectively) to quantify changes in drought under global warming. The average drought characteristics (e.g., duration, severity) were different based on two different baseline periods, which further led to difference in interpretation when future droughts (under climate change) were compared to historical drought characteristics. In
ideal scenarios, it is important to choose a baseline climatology that captures historical major drought events, for example, in this case, the inclusion of the Dust Bowl: the dry 1930s (1930–1931, 1934, 1936, and 1939–1940) [29] in the baseline period is likely to yield a different set of results [26]. In addition, the selection of 1950–2008 as the baseline period may include the effects of recent anthropogenic climate change that may be responsible to mask the climate change signals in the results of the analysis [26].

Similarly, the choice of different baseline periods can generate discrepancies in summarizing the results related to the same drought episode. For example, William et al. [71] and Luo et al. [138] investigated the causes behind the recent multiyear California drought (2012–2014). William et al. [71] reported that the anthropogenic warming trends account for 8–27% of the anomaly in 2012–2014 drought. On the other hand, Luo et al. [138] suggest that this multiyear drought most likely resulted from natural variability of climate and dominated by precipitation rather than temperature. The difference in results may be due to the usage of different drought indices, as well as the selection of different baseline periods: 1931–1990 [71] and 1979–2015 [138]. Thus, the baseline period should be appropriately chosen with caution by considering the drought aspect being studied.

(f) Non-stationary Climate and Choice of Probability Distribution

The appropriate selection of probability distribution plays an important role in deriving robust drought indices under climate change, especially considering stationary (historical) vs. nonstationary (future scenarios) patterns of climate variables. For example, calculation of SPI [17] is based on either a gamma distribution [17] or Pearson
type III distribution [139], whereas calculation of SPEI is based on a log-log distribution [59]. These distributions perform considerably well in fitting the time series of the hydroclimatic variables over a wide range of climatic region [140]. However, the selection of a single suitable probability distribution is challenging [141]. Vicente-Serrano et al. [141], while investigating best probability distributions to calculate the Standardized Streamflow Index (SSI), reported that most commonly used probability distributions (log-normal, Pearson type III, log-logistic, general extreme value, generalized Pareto, and Weibull) for flow frequency analysis provided good fits to streamflow data. However, none of the six probability distributions were able to adequately fit the streamflow series based on L-moment diagram. Therefore, the selection of distribution in developing a drought index is crucial and, if not done with caution, can generate large uncertainties.

Furthermore, it is well known that stationarity that implicates physical constancy of mechanisms involved in the hydrologic processes is no longer applicable due to the substantial anthropogenic changes in the present climate [32, 142]. Thus, drought characteristics will be different between stationary and non-stationary climate. Therefore, non-stationary statistics that are deterministic functions of time should be implemented in reliable assessment of hydrologic processes in the changing climate [32]. For example, the selection of probability distributions for precipitation is often challenged by significant zero values (mostly in dry climates), highly left skewed distributions, as well as limited data lengths [1]. Also, due to the non-stationary nature of climate variables under future climate scenarios, the probability distribution parameters of precipitation
will change over time. Therefore, it is important to consider non-stationarity by changing the probability distribution parameters over different timescales to improve drought assessment under climate change. Considering the strong association between precipitation and soil moisture, a similar assumption will also hold for soil moisture.

(g). Defining the Role of Anthropogenic Influence

Apart from the natural variability of climate, human activities have a significant control on drought initiation, propagation, and societal impacts [1, 66, 67, 69, 70, 74, 143]. Consequently, drought risk management is directed towards either adaptation to the natural causes of drought or mitigation of human-induced drought [143]. Identifying the anthropogenic causes of drought is crucial to assign proper weight to improve water management policies [74, 143]. However, the coupling of human components in hydrological models is in a preliminary stage for appropriately characterizing droughts under climate change. It is necessary to identify the associated challenges in distinguishing between natural and human influences due to the interplay between climate, soil, and vegetation dynamics [144, 145].

Detection and attribution (D&A) techniques [146], developed so far, use the combination of observation and GCMs in a virtual forcing scenario. This may allow the models to calculate drought characteristics in the absence of human influences [65, 71, 138]; however, the GCMs are vulnerable to uncertainties arising from boundary conditions, variability in the Earth system, parameter estimation, and model structure [147]. Furthermore, lack of observations for verification, and dependence on the model selection and the applied methodology, makes the existing D&A techniques less reliable.
in risk assessment of drought under the anthropogenic influence [148]. Therefore, quantifying uncertainties by estimating confidence intervals for risk ratios [148], and multimodel averaging rather than relying on individual model results [149], is necessary to avoid overconfidence in drought risk assessment based on drought indices. Moreover, uncertainties depend on the sample size of data and the severity of drought being studied; therefore, extra caution is needed while applying D&A methods [147].

Human-made infrastructures, such as dams and reservoirs, can also greatly affect the propagation of soil moisture and hydrological drought [74, 150]. Drought indices should capture such changes in drought propagation along with other human interactions such as dynamic changes associated with land use, irrigation efficiency, and rapid increase of population. However, such dynamics of human interactions is still in a preliminary stage in existing large-scale hydrologic modeling framework, and scientific advances are needed to overcome the aforementioned challenges.

8. Summary and Conclusions

A comprehensive discussion on the role of drought indices for climate change assessment is provided in this article. Existing drought indices were reviewed and compared based on their skill and limitations to capture drought characteristics in a non-stationary climate. Major shortcomings related to the formulation of drought indices under the changing climate, including the lack of robust approaches to separate the human component from the natural variability of climate, choice of baseline period, use of non-stationary climate information, and lack of observed data for validation, were
discussed. Significant progress is being made in drought research, and there is a scope to improve formulation of efficient drought indices with the hope of better drought preparedness by filling the gaps arising due to such shortcomings. The following conclusions can be drawn from this study:

1. The performance of drought indices, such as PDSI_sc, SPI, and SPEI, showed different degrees of sensitivity against the same level of observed warming at continental scale. Therefore, the formulation of drought indices without considering the factors that govern the background state may lead to drought artifacts under a warming climate.

2. Estimation of PET based on the energy budget framework can be a better physically based approach compared to only temperature-based equations. Also, uncertainties due to the spatial inhomogeneity in forcing data need to be considered to estimate PET for drought assessment under climate change [26, 27].

3. Major advancement in hydrologic modeling for drought assessment has been made with the development of LDAS. Land surface models have been successful in maintaining water and energy balance at macro-scale levels, thereby accurately capturing the components of hydrological fluxes in the top 1–2 m of the land surface at hourly and daily timescale, as well as at finer resolution [18]. These models have considerably improved the near real-time assessment of drought by providing modeled soil moisture, soil water equivalent, and runoff estimates at diurnal timesteps [18]. However, shortcomings need to be addressed in the existing LSMs by reducing uncertainties in hydrological fluxes by integrating in situ measurements and remotely sensed products [114, 115].
4. The choice of appropriate methodologies to develop drought indices for climate change assessment should consider projected climate variables with less uncertainty. This can be achieved by climate models simulating the best estimates of PET [107, 118, 121], and atmospheric demand or soil moisture [119, 151]. Specially, drought projections based on soil moisture-derived indices should be treated with extra caution owing to the lack of suitable observations for verification [33, 119]. Besides, there remains great uncertainty in what the future climate will be [152]; therefore, multimodel assessment is recommended against assessment using individual models [149].

5. Apart from the natural variability of climate, increase in severe and persistent droughts due to anthropogenic influence is reported in the last few decades [48]. Separating the natural causes from the human-induced factors is most likely to make drought assessment more realistic, thereby helping policy makers to simplify the complexities related to the water management decisions [153]. This can be achieved by objectively defining the role of human activity in drought assessment using drought indices.

6. Drought indices are widely used in multiple purposes by different stakeholders [154]. However, the actual usefulness and proper implementation of drought indicators/indices rely on how easily they can be interpreted by the stakeholders and serve the end user’s needs [1]. On the other hand, climate change affects a wider range of interconnected sectors [155], thereby further increasing the inherent complexity of quantifying socioeconomic droughts [1].
7. Climate model outputs as well as observed data are often available at coarser resolution, and it may limit our understanding on the hydrologic processes at finer scale [109]. Consequently, improved downscaling approaches should be developed to transform the information from coarser resolution to finer grid cells, thereby improving the assessment of drought impacts more realistically [1, 122]. For a good overview of different downscaling approaches for climate change assessment, see Maraun et al. [109]. Along with proper bias correction techniques, downscaling can provide quality data inputs for reliable drought assessment studies [156, 157]. Furthermore, an optimized model selection approach can be useful to select models with minimum uncertainty which should be adopted while downscaling drought indices based on climate models to capture their future variability [158].

8. The non-stationarity associated with climate change is likely to alter the parameters of the probability distributions of input variables in the formulation of drought indices. Therefore, adopting appropriate methods to capture nonstationary information for characterizing drought under climate change will likely generate reliable information for risk assessment and infrastructural management under the changing climate. Moreover, spatial drought risk can be investigated by integrating multiple drought characteristics (e.g., severity-duration-frequency (SDF)) [159, 160] that allows the user to compare historical major droughts with future scenarios under climate change [161].
9. References


CHAPTER THREE

COMPOUND DROUGHT AND HEATWAVES AT A GLOBAL SCALE: ROLE OF NATURAL CLIMATE VARIABILITY

1. Introduction

Recent decades have witnessed a notable increase in the concurrence of severe drought and heatwave events in many regions across the globe [1–6]. Such a spatiotemporal coexistence of these extremes is commonly known as a “compound events” [7], which has major implications for social-ecological systems such, as a reduction in crop yields [8–12], an increase in wildfires [13] enhanced tree mortality [14], and an exacerbation of human health hazards [15].

The occurrence of compound drought and heatwave (hereafter CDHW) events may be triggered by a variety of mechanisms ranging from local land-atmospheric feedback loops [16,17] to persistent large-scale ocean-atmosphere circulations anomalies (Hao et al., 2018; Seneviratne et al., 2012). The geographical regions under the influence of such extremes are characterized by prolonged clear skies and a lack of moisture in the lower atmosphere. The deficit in moisture results in the cessation of evaporative cooling, thereby limiting the latent heat fluxes over the land surface [20]. Subsequently, any further increase in diabatic heating in the region goes directly into increasing sensible heat fluxes and hence exacerbating pre-existing hot and dry conditions [21]. These anomalies in surface energy budget are often associated with atmospheric blocking and persistent lower-level divergent wind anomalies, which inhibit moisture convergence and lead to prolonged drier surface conditions, provides ample time for a heatwave to develop
and intensify over the region [22]. Many large-scale climate variabilities, including the El Niño Southern Oscillation (ENSO) [23], Pacific Decadal Oscillation (PDO) [24], and North Atlantic Oscillation (NAO) [25], are known to have a role in the formation of such high-pressure regimes or stationary blocking zones. Therefore, local-scale climate variability alone may not be able to effectively explain the occurrence of such events [26].

It is shown that large-scale climate oscillations, such as ENSO, play a significant role in the summertime occurrences of compound dry and hot events in tropical regions, particularly over Northern South America, Central and Southern Africa, Southeastern Asia, and Australia (Hao et al., 2018). In addition, a few regional studies have investigated the relationship between compound dry and hot events and the ocean-atmosphere circulation anomalies [27] with an aim to develop statistical models for improvement in the predictability of such events [28–30]. However, the calculation of drought (or dryness) index in these previous studies is either based on the monthly precipitation anomalies or on the Standardized Precipitation Index (SPI) (McKee et al., 1993). It is well established that drought quantification using only precipitation while ignoring the impact of temperature may lead to an underestimation of drying [32,33]. Therefore, the use of only precipitation-based indices has the potential to induce considerable methodological uncertainties, which may lead to unreliable estimates of CDHW characteristics. Furthermore, understanding the soil water properties during a hydrological cycle is relevant to the current physical understanding (of feedback processes) of the occurrence of CDHW events over a region (Hao et al., 2018; Mishra &
Singh, 2010). Therefore, the use of energy budget methods, which consider soil water as a variable in addition to precipitation and temperature in the identification of CDHW events can be more meaningful in the analysis of such compound occurrences. In addition, drought and heatwave events are interconnected due to the role of temperature as a common factor in triggering such events. While one of the earlier studies incorporates soil moisture anomalies in the estimation of drying (Hao et al., 2018), it makes use of monthly temperature anomalies in the estimation of CDHW characteristics, potentially compromising the valuable sub-monthly information needed to accurately identify the characteristics of heatwave events. The evaluation of CDHW at a monthly time scale may provide unreliable results as the heatwave events evolve over a period of days to a week. Therefore, in this study, we utilize an integrated temporal framework by incorporating droughts at a weekly timescale, and the heatwaves at a daily timescale to investigate the occurrence of CDHW events on a global scale.

While a plethora of drought indices are available in the scientific literature, the self-calibrated Palmer Drought Severity Index (sc_PDSI) [35] incorporates most of the hydroclimatic variables that exert critical controls on the drought characteristics [33,34]. Given the deficiencies in the earlier studies, we make use of daily observations and an empirical methodology to estimate the frequency of CDHW events for the entire globe. Therefore, our methodology utilizes the weekly self-calibrated Palmer Drought Severity Index (scPDSI) [35] for drought, and daily-scale heatwave events.

Our study focuses on each season (DJF, MAM, JJA, and SON) and on the 26 climate regions of the globe (as shown in Figure S1), which are defined in the AR5-
Overall, this study aims at providing a mechanistic understanding of the association between the seasonal occurrences of CDHW events and the major modes of climate variability (ENSO, NAO, and PDO) across these climate regions. Here, we quantify such associations based on a grid-based Poisson Generalized Linear Model (GLM), popularly used to model count data in climate science [37–39]. To achieve this objective, we (a) first select the large-scale climate indices as potential drivers for each climate region that influence the regional temperature and precipitation variability during each season by applying a non-parametric partial correlation technique; then (b) later use them as the explanatory variables to form candidate models that are subsequently applied to derive the grid-based GLM focusing on each season, separately; and (c) finally focus on the mechanistic understanding behind such associations by performing a composite analysis based on the relevant atmospheric and surface energy variables. It is important to note that reanalysis datasets are often used to provide mechanistic evidence supporting variations in the observations (precipitation and temperature). However, daily to weekly scale variations, such as CHDW events, in reanalysis may not align perfectly with the corresponding variations in the observations, which may result in inconsistencies when observed variations are explained in terms of anomalies in the reanalysis-based variables. Therefore, in our analyses, we make use of precipitation and temperature both from reanalysis and observations to highlight the inconsistencies that may arise due to this issue.

The paper is structured as follows. In Section 2.1, we describe the datasets used in the study. Section 2.2 describes the methodology applied in the computation of CDHW
events. The initial selection of potential drivers influencing the temporal variability of precipitation and temperature and final best model selection based on Poisson GLM are explained in Section 2.3. Section 3 provides results and discussion, followed by conclusion in Section 4.

2. Methods

2.1. Datasets

For this study, we select 1982 to 2016 as the analysis period. A daily observed precipitation dataset is obtained from the Global Precipitation Climatology Center (GPCC; https://opendata.dwd.de/climate_environment/GPCC/html/fulldata-daily_v2018_doi_download.html) at 1° horizontal grid spacing [40,41]. We select GPCC precipitation because it is fairly consistent in capturing precipitation patterns across different parts of the globe [42].

Available water content (AWC) is retrieved from the global database of texture derived AWC (Webb et al., 2000), available at 1° horizontal grid spacing. When integrated with the World Soil Data File (Zobler, 1986), the resulting dataset captures the soil properties relevant for determining the water storage in the individual soil horizons and for explaining the soil-atmosphere coupling in the lower atmosphere.

The daily observed maximum and minimum temperature \((T_{\text{max}}, \text{and } T_{\text{min}})\) data are obtained from the Climate Prediction Center (CPC). The CPC global temperature data is provided by the National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratory's Physical Sciences Division (PSD) (https://www.esrl.noaa.gov/psd/) at 0.5° horizontal grid spacing.
Monthly values of ENSO (Nino3.4), NAO, and PDO indices are retrieved from the NOAA Climate Prediction Centre (NOAA CPC; http://www.cpc.ncep.noaa.gov/). For the analysis of land-atmospheric dynamics, upper level (200 mb) zonal and meridional wind field, surface latent heat flux (SHF), and surface sensible heat flux (SHF) are obtained from the European Centre for Medium-Range Weather Forecasts Reanalysis 5 (ERA5; https://cds.climate.copernicus.eu/cdsapp#!/home). Additionally, we also make use of daily precipitation, Tmax and Tmin data from the ERA5 to investigate data related inconsistencies. We apply the Synergraphic Mapping System (SYMAP), described in Maurer et al., 2002, to regrid all dataset at 1° GPCC precipitation grid.

2.2. Frequency of CDHW Events

We define a CDHW event as a heatwave event occurring within the period of an extreme drought event. In this study, at each grid point, a heatwave event corresponds to three or more consecutive days with daily Tmax above a threshold [46,47]. We calculate unique threshold for each grid point as the 90th percentile of daily Tmax using the extended summer seasons (May to October in the Northern Hemisphere, and November to April in the Southern Hemisphere) of the study period. Drought conditions are based on the weekly sc_PDSI following Wells et al., (2004), which considers weekly total precipitation, weekly mean temperature (average of Tmax, and Tmin), and AWC in the energy-budget framework. A week is considered under extreme drought if the sc_PDSI for that week falls within the bottom ten percent of the sc_PDSI values (= extreme drought threshold) during the study period.
Figure 1. (a) Time series of weekly sc_PDSI and daily $T_{max}$ based on the GPCC and CPC dataset for the whole period of analysis, 1982 to 2016, for a single grid cell (longitude= -100.5 , and latitude = 29.5 (in degrees)), and (b) CDHW events (enclosed within the shaded boxes) for the same grid cell during JJA in the year 2012 (shaded region in (a)). The right y-axis represents the sc_PDSI values and the left y-axis represents the $T_{max}$.

A CDHW event occurs when daily $T_{max}$ during an extreme drought event exceeds the 90th percentile threshold for three or more consecutive days. The time-series
of CDHW events thus obtained are stratified based on four different seasons (DJF, MAM, JJA, and SON). Following the CHDW identifictaion criteria, Figure 1a depicts the time-series of the daily $T_{max}$ and weekly sc_PDSI for the whole period of the study at a single grid point, and Figure 1b illustrates three instances when a CDHW event takes place in 2012 JJA at the same location. The yearly total number of participating days (frequency) in the CDHW events during a season is considered as the yearly CDHW frequency (CDHWF) for that season. The spatial distribution of the yearly frequency of extreme drought, heatwave, and the CDHW events observed during all four seasons are illustrated in Figure S2, Figure S3 and Figure S4, respectively. For these and subsequent calculations, we perform separate analyses based on precipitation and temperature from the observations (GPCC and CPCC) and the reanalysis (ERA5). The yearly total number of participating days (frequency) in CDHW events during a season is considered as the yearly CDHW frequency for that season. The spatial distribution of the yearly frequency of extreme drought, heatwave, and CDHW events observed during all four seasons are illustrated in Figure S2, Figure S3 and Figure S4, respectively.

2.3. Association between the large-scale climate variability and the occurrence of CDHW events

As previously described, we select three major large-scale natural climate forcing (ENSO, PDO, and NAO) for this study. ENSO — an oscillation of the ocean-atmosphere system in the tropical Pacific — is one of the most important modes of variability impacting global precipitation and temperature on interannual timescales [23,48,49]. We use the Nino3.4 index to define ENSO. Similarly, NAO is a dominant mode of climate
variability over the Northern Hemisphere that ranges from intra-seasonal to multi-decadal time scales, and exerts influence on the variations in air temperature and precipitation across Europe, the Mediterranean Basin, and parts of North America and Asia especially during winter [25]. Likewise, PDO is an annual pattern of monthly sea surface temperature variability in the North Pacific region [24], which is known to have a major influence on precipitation variability over North America and Asia.

Figure 2 Flow diagram explaining the multi-stage exploration of model selection performed for a specific climate region. “NA” indicates none of the drivers are selected for the region or for any grid cell within the region.
Using a grid-based Poisson GLM, we quantify the association between these large scale modes of natural climate variability on the inter-annual variability of CDHW events for each season. We quantify such associations in two stages. First, a preliminary selection of the potential large-scale drivers is made based on the regional climate characteristics (represented as regional drivers (RDs), section 2.3.1.). Second, the selected RDs are used as guidelines to identify final predictors for developing the Poisson GLM (section 2.3.2.). The entire methodology employed in the selection of RDs and finally the GLMs for a specific climate region is shown as a flow diagram in Figure 2. A similar technique undertaking the preliminary selection of possible drivers has been employed in a previous study related to the association between climate variability and regional heatwave event characteristics [50].

2.3.1. Preliminary selection of potential large-scale drivers

The large-scale climate drivers are selected based on their potential influence on the seasonal precipitation and temperature variability over different geographical regions by using following steps:

(a) First, using non-parametric spearman’s rank correlation, we identify possible RDs that influence interannual precipitation and temperature variability at a seasonal time-scale. Monthly precipitation, temperature, and three large-scale climate indices (Nino3.4, PDO, and NAO) are seasonally averaged over the 1982 to 2016 period. Subsequently, correlations between the yearly values of the indices and the meteorological variables are estimated for each season at each grid point. In order to
account for the inter-dependence of different climate modes [50,51], we employ partial correlation technique (equation 8) to isolate the influence of individual forcing,

\[
r_{xyz} = \frac{r_{xy} - r_{xz}r_{yz}}{\sqrt{(1-r_{xz}^2)(1-r_{yz}^2)}} \tag{8}
\]

where \( r_{xyz} \) is the relative correlation between \( x \) (precipitation or temperature), and \( y \) (ENSO or NAO or PDO) with the effect of \( z \), either of the other indices are removed.

(b) Second, for each climate region, we estimate the percentage of grids where correlations are statistically significant at 95% confidence level, irrespective of the sign of \( r_{xyz} \). If the percentage of grids with significant correlation exceeds 10% of the climate region for either of the two variables, the corresponding RDs are selected as potential forcing factors influencing CDHW events in that region (as discussed in result, Section 3.1.).

2.3.2. Poisson Generalized Linear Model (GLM)

The selected RDs for a climate region are used as initial predictors in the Poisson GLM with CDHW events as a predictand. The Poisson GLM is a special form of the generalized linear model [52] that is very useful for modeling count data. A general approach is to assume that a random variable \( N_i \) follows a non-homogenous Poisson distribution with time-varying rate \( \lambda_i \), \( N_i = \text{Pois}(\lambda_i) \), if \( N_i \) takes on the values \( n = 0,1,2,.., \) with probability

\[
P(N_i = n | \lambda_i) = \frac{e^{-\lambda_i} \lambda_i^n}{n!}.
\]

(9)
Here, $N_t$ represents the frequency of CDHW events (in days) at each grid during any given season for year $t$ and $\lambda_t$ is the expected value of Poisson distribution that can be modeled as a function of predictors under the GLM framework as following [53]:

$$\log(\lambda_t) = \beta_0 + \sum_{i=1}^{k} \beta_i x_i(t),$$

(10)

$x_i(t)$ represents selected large-scale drivers, $k$ is the number of predictors, and $\beta_i$ represents the coefficients for the respective predictors. In order to account for areal variations at different latitudes, we modify Eq. (10) as following:

$$\log(\lambda_t) = \beta_0 + \sum_{i=1}^{k} \beta_i x_i(t) + \log(\cos \phi).$$

(11)

where $\phi$ is the latitude and $\log \cos \phi$ serves as an offset term with coefficient 1.

We obtain coefficients for the respective predictors in Eq. (11) by maximizing the log-likelihood function for the Poisson GLM defined as,

$$L = \sum_{t=1}^{T} N_t \log \lambda_t - \lambda_t - \log(N_t!),$$

(12)

The maximized log-likelihood ($L$) provides an indirect measure of how well the model fits the data. However, a positive bias in the estimates of $L$ amplifies as the number of predictors increase, resulting in overfitting of the model. Such a situation can be avoided by using Akaike Information Criteria (AIC) that accounts for these biases and
removes redundant predictors to prevent overfitting. The initial predictors are subjected to a stepwise regression approach based on the AIC estimated as,

$$AIC = 2p - 2\ln L$$

(13)

where \( p \) is the number of parameters in the model.

The Poisson GLM that has the smallest AIC value is considered as the best model for that location. However, the AIC score does not provide any evidence about the absolute quality of the model. Therefore, the chosen predictors of the best AIC model are further tested for statistical significance based on the Wald test. Only predictors, significant at 95% confidence level, are included in the final Poisson GLM to compute the regression coefficients at each grid location.

In order to determine the significance of relationship between the estimated regression coefficients \( \beta_i \) and the random variable \( N_t \), we test the null hypothesis that \( \beta_i = 0 \) is true based on the test statistic \( Z \). In other words, we assume that the expectation of the fitted regression coefficient \( \hat{\beta} \) is 0. Upon standardizing the regression coefficient of individual predictors, we obtain the test statistic \( Z \), which follows a normal distribution as following [54]:

$$Z = \frac{\hat{\beta} - 0}{\hat{\sigma}} = \frac{\hat{\beta}}{\hat{\sigma}}.$$

(14)
The test statistic $Z$ allows test against a two-sided alternative hypothesis that $\beta_i \neq 0$ at a significance level of $\alpha = 0.05$ in which the critical value is the upper $\alpha/2$ percentage point of the standard normal distribution, $z_{\alpha/2}$. The regression coefficient is considered to be statistically significant if $|Z| \geq z_{\alpha/2}$, thereby, rejecting the null hypothesis that $\beta_i = 0$.

3. Results and Discussion

3.1. Potential regional drivers at seasonal scale

The spatial maps of seasonal partial correlations, corresponding to ENSO, PDO, and NAO, are shown in Figures S5 to S7 for the observed precipitation and temperature and in Figures S12 to S14 for the reanalysis precipitation and temperature. These correlations become the basis of our predictor selection in the Poisson GLM model. Correspondingly, the percent of total grid points within each region where the correlations are statistically significant is shown as bar plots in Figures S8 to S11 for observations, and in Figure S15 to S18 for reanalysis. The use of 10% of the total grid points within a region as a minimum criterion for the selection of RDs as predictor leads to a total seven combinations (Figure 2). The geographical distribution of these 7 RDs is shown in Figure 3.

If we consider similarities between the left column (observations) and the right column (reanalysis) RDs, then precipitation and temperature variations are influenced by PDO and NAO (RD6) over Northern Europe and Asia (Figure 3a-3b) in DJF, Eastern Asia and Sahara and Central America in JJA, and Eastern America, Canada and Northern
Figure 3 Regional Map showing combinations of potential regional drivers (RD1 to RD7; filled circles) for the 26 AR5-Climate regions based on the GPCC and CPC (left panel), and the ERA5 (right panel). The list of large-scale climate variabilities corresponding to each of the selected combinations of regional drivers (RD1 to RD7) is provided in Figure 2.
Europe in SON. Similarly, ENSO exhibits a substantial footprint in the monsoon regions, including Southern South America, Northern Australia and South Africa in DJF, and most of South America, South Asia and Northern Australia in SON, consistent with the earlier finding [55]. PDO (RD2) also exhibits influence over many regions of Asia in SON while all of them (ENSO, PDO, NAO; RD7) seem to have an influence over South Asia, Southeast Asia, Northern Australia and Western Asia, which is consistent with previous findings [56–58]. The interaction of PDO with NAO, and ENSO in influencing the precipitation and temperature variability has been reported in the previous studies [59,60].

The existence of several differences between the left and right columns in Figure 3 highlight the fact that the identification of regional drivers of precipitation and temperature variability depends on the source of data. For instance, RDs based on observations show the influence of PDO, and NAO (RD6) over Canada and Mediterranean in DJF (Figure 3a) while use of ERA5 reanalysis replaces PDO with ENSO (RD5) over Canada and only shows the influence of NAO (RD3) over Mediterranean. Similarly, unlike observations, reanalysis doesn’t exhibit influence of PDO over Western and Southern Africa in SON, and over Northern Asia in MAM. Inconsistencies also exists in many other regions and seasons (Figure 3).

The best AIC models that are also statistically significant at the 5% significance level, which are obtained by fitting the RDs (Figure 3) as explanatory variables in the Poisson GLM, are presented geographically in Figure 4. It should be noted that the final models over a region can be identical to the RDs, a subset of RDs or none, which is
Figure 4 Spatial map of the best models (M1 to M7; filled circles) explaining the interannual variability of CDHW events from 1982 to 2016 based on the GPCC and CPC (left panel), and the ERA5 (right panel), which are identified by the Poisson GLM. The list of large-scale natural climate forcing corresponding to each of the selected models (M1 to M7) is provided in Figure 2. The bold boundaries indicate regions exhibiting the influence of a single large-scale driver, which is also consistent in both datasets.
explained in Figure 2. For instance, Northern Asia that has RD5 (ENSO and NAO) as the selected combination of regional drivers (Figure 3e, 3f) in JJA shows M1 (ENSO), M3 (NAO), and M5 (ENSO, NAO) as the best fitted models (Figure 4e, 4f) after the application of the AIC and Wald tests. Similarly, Eastern Asia that has RD2 (PDO) as the selected regional driver (Figure 3 g, h) in SON has no best fitted model (Figure 4g, 4h) after the application of the AIC and Wald tests. Moreover, as previously noted, choice of data source also influences the RDs and hence the final best fitted models (Figure 4, left column versus right column). We identify such regions and hereafter highlight them with a bold boundary.

3.2. Relationship between CDHW events and large-scale climate variability

The formulation of Poisson GLM indicates that the estimated coefficients can be interpreted as a direct measure of sensitivities [37–39]. In addition, the logarithmic function applied in the Poisson GLM is useful to determine the relative importance of every selected large-scale driver, such that 1 unit increase in the regression coefficient \( \beta_i \) can lead to a \( \exp(\beta_i + 1) \) times impact on the predictand. Thus, the higher magnitude of the regression coefficient corresponds to the higher relative impact of the associated driver. Positive (negative) values of the regression coefficients imply that the warm (cold) phase of the associated large-scale drivers has a positive (negative) association with the CDHW events. In this section, we discuss the impacts of warm and cold phases of the associated large-scale drivers across different climate regions of the globe.
3.2.1 El-Nino Southern Oscillation (ENSO)

Figure 5 shows grid points with statistically significant Poisson GLM regression coefficients corresponding to ENSO and CDHW events. Given the dependence of RDs and resulting final models on the choice of data, we only focus on those regions where the results are consistent between observations (left column) and reanalysis (right column). A significant positive influence of the warm phase of ENSO (El Niño) exists in the Southern Hemisphere during the austral summer months (DJF; Figure 5a) over the Amazon, Southern Africa and Northern Australia, and during SON over the Southern Australia. Some of these associations are consistent with earlier studies (Hao et. al., 2018; Hirons & Klingaman, 2015; Min et al., 2013). The negative phase of ENSO (La Niña) exhibits a significant positive influence over parts of Asia in JJA and over the parts of Central America and Western Africa in MAM. Central Europe also exhibits positive influence of El Niño in JJA. However, ENSO is not the only driver over these regions during JJA. Unfortunately, several regions exhibit inconsistencies between the observations and reanalysis, including Southeast Asia during all the four seasons; Alaska, Eastern and Western Africa, and parts of Southern South America in MAM; Canada, and West Asia in JJA, and East Africa in SON.
Figure 5 The statistically significant Poisson GLM regression coefficients for ENSO (Nino3.4) based on the GPCC and CPC dataset (left column), and the ERA5 (right column), explaining the interannual variability of CDHW events during (a, b) DJF, (c, d) MAM, (e, f) JJA, and (g, h) SON. The bold boundaries indicate regions exhibiting the influence of a single large-scale driver, which is also consistent in both datasets.
3.2.2 Pacific Decadal Oscillation (PDO)

The Poisson GLM regression coefficients for PDO explaining the interannual variability of CDHW events during all four seasons is presented in Figure 6. With the exception of JJA, all other seasons exhibit inconsistency in the influence of PDO. In JJA, negative phase of PDO is the sole driver over Western North America (highlighted by bold boundary in Figure 6) and is one of the drivers over central North America, Sahara, Mediterranean, East Asia and Tibet. The role of PDO in modulating the CDHW events over the eastern part of Asia has been previously reported [63]. Previous studies also suggest that the anomalies over the northeastern and tropical Pacific have a significant impact on the occurrence of CDHW events in the contiguous US [64–67].

3.2.3 North Atlantic Oscillation (NAO)

Figure 7 presents the significant Poisson GLM regression coefficients for NAO during the four seasons. The strength of regression coefficients is relatively weak in the case of NAO. However, unlike ENSO and PDO, the results are generally independent of the data source. Both datasets show NAO as the sole significant influencer over Northern Europe and Eastern North America (highlighted by bold boundaries in Figure 7) in JJA. Apart from that, in combination with other modes, role of NAO is seen in JJA over most of Asia, Europe, Northern Africa and Eastern North America. Similarly, negative phase of NAO exhibits more influence over Northern Asia and West Africa while influence of the positive phase of NAO is limited to Northern Europe and parts of Central Asia. Many earlier studies support the influence of NAO on the occurrence of CDHW events over Europe [68] Asia [69] and North America [70,71].
Figure 6 Same as in Figure 5 but for PDO.
3.3. Atmospheric anomalies associated with warm and cold phases

In order to understand atmospheric anomalies associated with the large-scale climate drivers that lead to their regional associations with CHDW events, shown in
Figures 5 to 7, we analyze seasonal anomalies in upper-level (200 mb) velocity potential and divergent winds during warm and cold phases of ENSO, PDO, and NAO. Dry and hot conditions are generally associated with lower level divergence (or upper level convergence) anomalies in the atmosphere [72,73]. The seasonal analyses of divergent winds illustrate the variations in the Walker and Hadley circulation across seasons (Figure S19). In DJF, upper-level divergent wind centers are in the Southern Hemisphere over the western Pacific, the eastern Indian Ocean and the Amazon, which correspond to lower-level monsoonal flow over these regions. In JJA, divergent winds centers shift to the Northern Hemisphere over the Asian and North American monsoon regions. Anomalies in divergent winds can occur due to sea surface temperature anomalies in various oceanic basins [74], which act as a remote connection between natural modes of climate variability and climates over many terrestrial regions.

It should be noted that those regions where more than one large-scale driver influences the occurrence of CDHW events (Figure 3; RD4 to RD7), atmospheric anomalies based on the warm and cold phases of a single driver may not be able to provide a mechanistic explanation of overlying dynamic causes. Therefore, our explanation of atmospheric anomalies is mostly limited to those regions where a single driver is shown to have an influence (M1 to M3; bold boundaries Figure 4 to 7) in the observations and reanalysis.
Figure 8 El Niño (left panel) and La Niña (right panel) based anomalies in divergent winds vectors (m/s) and velocity potential (color shading, unit:m²/s, scaled by 10⁶) contours at 200 mb with respect to the climatology during (a-b) DJF, (c-d) MAM, and (e-f) SON.

Figure 8 illustrates seasonal anomaly in the velocity potential and upper-level divergent winds for the El Niño (left panel) and La Niña (right panel) in DJF, MAM, and SON. The robust positive association between El Niño and the CDHW events during DJF
(Figure 5a, 5b) is explained by the anomalous upper-level convergence over Amazon, Southern Africa, and Australia (Figure 8a). Similarly, upper-level anomalous convergence is present over Australian continent in SON during El Niño, consistent with the positive association of El Niño and the CDHW events during that season (Figure 5g, 5h)). Given that upper-level convergence is an indicator of lower-level divergence or clear weather conditions, these anomalies favor reduction of moist flow particularly during the austral summer over these regions. Circulation anomalies in JJA during El Niño (La Niña) are also consistent with its association with the JJA CDHW events over Central Europe (Central Asia and parts of Northern Asia) (Figure S20). However, ENSO is not the only driver over these regions.

Similarly, Figure 9 depicts the upper-level circulation anomalies during the warm and cold phases of PDO and NAO in JJA. Western North America is the only region where PDO is the sole driver associated with the occurrence of CDHW events in JJA (Figure 4). The sign of regression coefficient suggests that it is the cold phase of PDO that exhibits positive correlation with CDHW (Figure 6e, 6f). This association is consistent with anomalies in upper-level divergent winds in JJA during the cold phase of PDO, which exhibit anomalous upper-level convergence centered over the eastern Pacific. Moreover, the anomalies in upper-level divergent winds over Central America explain the positive association of the CDHW (Figure 6b) events with the cold phase of PDO in MAM (Figure S21). However, this association is not present in the case of observations based CHDW events (Figure 6).
As previously noted, magnitudes of regression coefficients are substantially weaker for NAO compared to ENSO and PDO (Figure 5-7), which suggest relatively weak influence of NAO in the occurrence of CDHW events. During JJA, when NAO influence is spatially more visible (Figure 7e, 7f), upper-level circulation anomalies are relatively small (Figure 9c, 9d), which are consistent with weak magnitudes of regression coefficients.

**Figure 9** Same as in Figure 8 but based on warm (left panel) and cold (right panel) phase of (a, b) PDO, and (c, d) NAO During the JJA season.

It is also important to highlight that regions where a single large-scale driver only exists in the observations-based analysis, reanalysis-based circulations anomalies do not support such associations (Figures S21, S22). For instance, in SON, only in the analyses
based on the observed precipitation and temperature, PDO appears as the single large-scale driver (Figure 3g, 3h) over Western and Southern Africa and positive phase of PDO exhibits association with CDHW events (Figure 4g, 4h). However, anomalies in the reanalysis-based upper-level circulations over Africa exhibit anomalous divergence, which correspond to lower-level convergence or conditions opposite to the ones needed for CDHW events (Figure S21). Likewise, parts of North America and Northern Europe in JJA exhibit NAO as the sole driver in the analyses based on the GPCC and CPC datasets, but ERA5-based circulation anomalies do not support such association of NAO with the occurrence of CDHW events over these regions. These inconsistencies suggest that use of reanalysis to explain anomalies in the observations, particularly those related with fine temporal variations, may not be a robust strategy.

3.4. Surface energy budget anomalies associated with the warm and cold phases

The role of land-atmospheric feedbacks in modulating the characteristics of CDHW events is well recognized over many regions of the globe [4,17,75,76,76]. Under dry and warm conditions, the positive land-atmospheric feedback processes, which are also referred to as soil moisture-temperature coupling, are one of the local-scale controlling factors governing surface energy budget. Characterized by the cessation of cooling from evaporation under dry conditions, such positive feedback loops lead to a decrease in the latent heat fluxes (LHF) and an increase in the sensible heat fluxes (SHF). It is well established that natural climate forcing influence terrestrial water availability, soil moisture deficits and evaporation rates through their controls on precipitation
generating mechanisms [77–79]. These variations in the terrestrial water cycle directly impact the land-surface energy feedbacks and hence influence the occurrence of CDHW events [80].

We investigate the influence of ENSO, PDO, and NAO on these land-surface and atmosphere feedbacks by analyzing seasonal anomalies in the SHF and LHF during the warm and cold phases of ENSO, PDO, and NAO with respect to their seasonal climatology (Figure S23). We use a two-sample Student’s t-test at a 5% significance level to test the significance of anomalies [81]. For any given season, a positive (negative) anomaly in the SHF and a corresponding negative (positive) anomaly in the LHF characterizes a positive land-surface and atmospheric feedback. Thus, the results from the composite analysis are discussed hereafter based on the positive feedback loops with an aim to provide a localized mechanistic explanation behind the association between the large-scale climate forcing and the occurrence of CDHW events, as shown earlier in Figures 5, 6, and 7. As noted earlier, the design of these analyses limits their use only to those cases where a single natural climate forcing is consistently identified as the regional driver in observations and reanalysis.

A significant simultaneous increase in the SHF and decrease in LHF is witnessed over the parts of the Southern Hemisphere in DJF where El Niño exhibits a positive association with the occurrence of CDHW events (Figures 10a, S24) and where upper-level circulation anomalies favor dry conditions on the surface (Figure 8a), such as over Southern Africa and western parts of Northern Australia. Similarly, a significant increase in the SHF and decrease in the LHF is present over major parts of Australia in SON.
(Figures 10e, S24) that exhibit a positive association between El Niño and CDHW, and upper-level circulation anomalies that favor drying.

In case of the cold phase of PDO, a delayed effect of the positive land-atmospheric feedbacks can be noticed over the Western North America during the JJA season (Figure 10d). Given that JJA is the dry season over the Western North America, the surface drying in JJA may be in part influenced by the significant positive high SHF and negative LHF anomalies in the preceding season (MAM; Figures 11b, 11d, S25), which is consistent with its positive association with the CDHW events (Figure 6f). However, results are not very clear in the case of NAO over Eastern North America and Northern Europe in JJA, where it is the only large-scale driver in both datasets. This inconclusiveness in the case of NAO is because both phases of NAO appear to have association with CDHW events over these regions in JJA, which is also relatively weak (Figure 7e, 7f).
**Figure 10** El Niño (left panel) La Niña (right panel) based anomalies in sensible heat flux with respect to the climatology during (a, b) DJF, (c, d) MAM, and (e, f) SON. Stippling represents statistically significant at 95% confidence level. All units are in W/m². The sign convention implemented for the fluxes is positive upwards.
Overall, the anomalies in circulation patterns and surface energy fluxes can explain influence of ENSO, NAO, and PDO in the occurrence of CDHW events when any one of them is the only driver in both datasets (indicated by bold boundaries in

**Figure 11** Same as in Figure 11 but based on the warm (left panel) and cold (right panel) phase of PDO. All units are in W/m². The sign convention implemented for the fluxes is positive upwards.
Figure 5-7). Exceptions in the case of NAO exist, which are understandable given its relatively weak, mixed and spatially sparse relationship with CDHW events (Figure 7). Furthermore, as it is in the case of circulation anomalies, reanalysis-based surface energy fluxes variations are inconsistent when association of ENSO, NAO, and PDO with the CDHW events is only limited to the observations-based analyses (Figure S27).

4. **Summary and Conclusions**

Using weekly se_PDSI and daily Tmax, we estimate the yearly frequency of CDHW events for four seasons from 1982 to 2016 period over the 26 AR5 climate regions of the globe. Moreover, an association between three major large-scale climate forcing (ENSO, PDO, and NAO), and the yearly occurrences of CDHW events is investigated based on the regression coefficients estimated using Poisson GLM for each season. Furthermore, we examine the robustness of such associations by using precipitation and temperature from two sources (observations and reanalysis).

Based on the estimates of Poisson regression coefficients, we find that the warm phase of ENSO has a relatively more dominant and robust footprint, specifically over the Southern Hemisphere. ENSO alone has a significant influence on the yearly occurrences of CDHW events over Northern Australia in DJF and SON, over Southern Africa, Southern parts of South America in DJF, and over Amazon in SON. These regions exhibit consistent associations of CHDW events with ENSO in the observation and reanalysis, which are supported by the anomalies in the reanalysis-based circulations and surface energy fluxes during the respective seasons. Similarly, the cold phase of PDO
exhibits positive influence over Western North America in JJA, which is associated with anomalous upper-level convergence and strong surface drying in JJA and preceding season. Our results indicate a weak influence of NAO in the occurrence of the CHDW events.

In this study, we have provided a preliminary mechanistic understanding regarding the association between the major modes of natural climate variability and the occurrence of CDHW events across the globe. However, there are a number of caveats in our methods and analyses that require improvements in future studies. For instance, the frequency of CDHW events depends on the choice of indices for heatwave and drought identification. The identification of heatwaves, in particular, can vary if a different percentile threshold or consecutive days criteria is used. Therefore, further research is needed for more robust definition of compound drought and heatwave events. Similarly, this study only focuses on three modes of natural climate variability and potentially ignores other large-scale forcing such as SST anomalies in the Indian Ocean (Indian Ocean Dipole), Tropical North Atlantic and Equatorial Atlantic (Atlantic Niño) that can independently impact the occurrences of CDHW events or their co-occurrence with ENSO, PDO and NAO can potentially enhance or dampen their impact [82,83]. Moreover, our analysis also highlights that weekly to sub-monthly scale anomalies in temperature and precipitation, which in some cases give rise to CHDW events, may not be always aligned between the observations and reanalysis across the globe. Therefore, caution must be exercised in the mechanistic explanation of observed anomalies in the precipitation and temperature on the basis of anomalies in the reanalysis-based
circulations and surface energy budget. Nonetheless, our analysis provides a new insight into the mechanistic understanding towards concurrent extremes and should help foster research efforts in this area, especially in improving the seasonal predictability of such extremes.

5. References


10. Mishra V, Thirumalai K, Singh D, Aadhar S. Future exacerbation of hot and dry summer monsoon extremes in India. npj Clim Atmos Sci [Internet]. Nature


63. Yu E, King MP, Sobolowski S, Otterå OH, Gao Y. Asian droughts in the last millennium: a search for robust impacts of Pacific Ocean surface temperature


73. Trenberth KE, Stepaniak DP, Caron JM. The Global Monsoon as Seen through the Divergent Atmospheric Circulation. J Climate [Internet]. American


CHAPTER FOUR

COMPOUND DROUGHT AND HEATWAVE IN A WARMING WORLD

1. Introduction

Compound drought and heatwave (CDHW) events have garnered much attention in recent decades [1–5]. These CDHW events often impact socio-ecological systems [3], which includes wildfires [4,6], massive heat-related deaths [7,8], and loss of crop yield [9,10]. Many regions around the world experienced CDHW events during the summer months of 2003, 2010, 2015, and 2018 in Europe and western Russia [1,6,11,12]; 2012-2014 in the USA [13–15]; 2013 in Australia [16,17]; 2006, 2009–2010, and 2014 in Southwestern and Northern China [18–21].

The evolution of drought and heatwaves are controlled by a variety of land surface fluxes [22], and their spatial patterns are heterogeneous due to the regional differences in precipitation, temperature anomalies, and other hydrological changes [22–26]. Besides, these land surface fluxes are influenced by the background aridity, anthropogenic factors, and large-scale changes in climate patterns [3,27–30].

Numerous studies have investigated CDHW events [3] based on different combinations, such as monthly precipitation and temperature anomalies (Hao et al., 2018a); Standardized Precipitation Index and temperature anomalies [32,33]; precipitation deficit and potential evapotranspiration [34]; and temperature anomalies and precipitation deficit [35]. Nevertheless, the CDHW event estimation framework applied in these studies can be further improved in the context of the timescale as well as the type of indicators used to quantify drought and HWs [3,36,37].
While many studies investigated compound extremes, this study further complements previous studies in three different ways. Firstly, we applied a framework to estimate the frequency, duration, and severity of CDHW events by integrating weekly drought and daily HW information, which was neglected in earlier studies. This framework can overcome the limitations arising from the difference in the temporal evolution of heatwaves (i.e., daily to weekly scale) and droughts (i.e., weekly to monthly scale)[3]. Secondly, the spatial asymmetry of the temporal evolution of CDHW events is investigated for the first time on a global scale. Thirdly, although background aridity plays a vital role in the development of drought and heatwaves, its potential influence on the evolution of CDHW events in different climate regimes (arid, transitional, and humid) remains largely unexplored.

The overall objective of this study is to quantify the spatiotemporal changes in the CDHW events at a global scale based on the following research questions: (a) how does the recent increase in temperature influence the CDHW events, (b) Is there a significant increase in the proportion of HW days coinciding with the extreme drought conditions?; (c) to identify the regions that show a greater increase in the characteristics of CDHW events, and examine if there is spatial asymmetry associated with such growth at the global and hemispherical scale?; (d) whether the background aridity controls the increase in CDHW characteristics at regional to the continental scale; and (e) whether these results are consistent across different datasets. This paper is structured as follows. Section 2 describes the data and methodology used in the study. Section 3 provides results and discussion, followed by the summary and conclusions in Section 4.
2. Methods

2.1. Data

We use daily precipitation (P) dataset obtained from the Global Precipitation Climatology Center (GPCC) (http://gpcc.dwd.de/), available at 1° spatial resolution for the period 1983-2016. GPCC precipitation dataset is relatively consistent across the globe [38], therefore, it is expected to provide spatially robust estimates of CDHW characteristics. Available water content (AWC) is obtained from a soil texture–based global water-holding-capacity map produced by Webb et al., 2000, and extracted at 1° spatial resolution. Daily maximum and minimum 2 m air temperature (T_{max} and T_{min}) dataset is retrieved from the Climate Prediction Center (CPC) (https://www.esrl.noaa.gov/psd/; available at 0.25° spatial resolution) for the period, 1983-2016. Humid, transitional, and arid regions are identified based on the potential evaporation (E_{p}) data obtained from the European Centre for Medium-Range Weather Forecasts Reanalysis 5 (ERA5; https://cds.climate.copernicus.eu/cdsapp#!/home; available at 0.25° spatial resolution) at hourly time-steps for the same period. The datasets from the ERA5 potentially supersede other reanalysis products by its high spatial and temporal resolution. Additionally, we also use daily precipitation, T_{max} and T_{min} data from the ERA5 and T_{max} and T_{min} data from Berkeley Earth Surface Temperature Project (BE; http://berkeleyearth.org/data-new/) to investigate data related inconsistencies. All datasets are regridded at 1° spatial resolution by applying the bilinear interpolation scheme to make it consistent with the GPCC Precipitation dataset.
2.2. Estimation of CDHW Event Characteristics

The summer CDHW events are calculated by integrating heatwave and drought information. In this study, the drought events are derived based on the weekly self-calibrated Palmer Drought Severity Index (sc_PDSI), and HW events are derived based on the abnormally high-temperature anomalies observed for three or more consecutive days. Besides, considering possible epidemiological significance [40], two successive HW events are considered independent if separated by a minimum of four days, otherwise clustered into a single event. The sc_PDSI is calculated based on the methodology proposed by Wells et al., 2004. This procedure includes weekly total precipitation, weekly mean temperature (average of Tmax, and Tmin), and available water capacity (AWC) in the energy-budget framework as inputs for deriving sc_PDSI. This methodology employs the potential evapotranspiration estimated using Thornthwaite's method [42]. In our analysis, the sc_PDSI is derived utilizing the climate characteristics calibrated based on the whole period, 1983-2016.

The CDHW events are estimated based on the time periods when a HW coincides with the extreme drought weeks [3]. An extreme drought week is identified when the scPDSI magnitude falls below the 10th percentile of the weekly values during the study period [3]. The daily temperature anomalies are calculated for each location, separately, based on the corresponding 90th percentile of daily maximum temperature observed in the extended summer seasons (May-October in the northern hemisphere, and November-April in the southern hemisphere) from 1983 to 2016. The concept of identifying HW event is analogous to previous studies conducted for different parts of the Globe [43–46].
In this study, CDHW events are characterized by three metrics: (a) CDHW frequency ($CDHW_f$), defined as the average number of annual CDHW events during the study period; (b) CDHW days ($CDHW_d$), defined as the total number of event days observed for a given year; and (c) CDHW severity ($CDHW_s$), which is calculated based on the cumulative sum of the daily severity values obtained over the consecutive days of the CDHW events. The daily severity is estimated as the product of the daily standardized values of maximum temperatures and the scPDSI value observed in the coinciding extreme drought week. A description of the methodology used in the estimation of CDHWs is provided in A.1. of the Supplementary.

3. Results

3.1. Changes in the CDHW events

We investigate the annual changes in proportion of CDHW events and global land area witnessing these events during the study period, 1983-2016. The proportion of CDHW events for a specific year is calculated based on the ratio (\% change) between the total number of CDHW and HW events occurring at any given location. Hereafter, this ratio (\%) is referred to as $CDHW_r$. The $CDHW_r$ estimation allows us to investigate interannual changes in the proportion of HW events that coincide with the extreme drought weeks observed during the period 1983-2016.

Based on this concept, the time series of $CDHW_r$ are constructed for each grid location and subsequently averaged over the globe. The interannual trends in the global $CDHW_r$ are estimated and compared between the whole period (1983-2016), past period (1983-1999), and the recent warmer period (2000-2016) to investigate the potential contribution
of recent warming. The criteria for selecting the past and recent warmer period is discussed in A.2. of the Supplementary. This criterion is evaluated based on the summer mean of daily Tmax anomalies in the CPC, BE, and ERA5 datasets, as illustrated in Fig S1. It can be observed that the period from 2001 to 2016 was considerably warmer than the preceding period, 1983-2000. Notably, the temporal pattern exhibited by the magnitude of these anomalies shows a close agreement among all the three datasets (Fig S1). Such consistencies among the three (potentially) different temperature data sources can minimize the possibility of uncertainties in the analysis.

In this study, the trends are determined by employing the Sens’s slope estimator, and the statistical significance of the trends is tested based on the MK test [47,48] at 95% confidence level (see A.3 of the Supplementary Information). Similarly, we also examine the interannual changes in the percentage of global land area (estimated as the percentage of land area within the grid cells weighted by the cosine of their latitudes) that witnessed a CDHW event during these three periods. Even across different quality-controlled datasets, trends can often vary substantially, even if year to year variability is congruent and especially for variables with larger uncertainties like precipitation. Specifically, PDSI estimates of drought trends depend critically on the precipitation datasets being used (Mukherjee et al., 2018; Trenberth et al., 2014). To investigate data related inconsistencies, these changes are analyzed based on three different combinations of datasets: (1) precipitation from GPCC and temperature from CPC (GPCC-CPC), (2) precipitation from GPCC and temperature provided by Berkeley Earth Surface
Temperature Project (GPCC-BE), and (3) precipitation and temperature both from ERA5.

These changes are illustrated in Fig1.

Figure 1. (a-c) Linear and scatter plots show the annual timeseries of globally averaged CDHW$_r$ (%) based on (a) GPCC-CPC dataset, (b) GPCC-BE dataset, (c) ERA5 dataset, and (d-f) bar-plots illustrate the percentage of global land area affected by the CDHW events each year during the whole study period, 1983-2016 (%) based on (d) GPCC-CPC
dataset, (e) GPCC-BE dataset, (f) ERA5 dataset. Linear annual trends in CDHW$_r$ and in the global land area are estimated based on the Sen’s slope estimator, and statistical significance in trends are determined based on the MK test for the whole study period (1983-2016), past period (1983-1999), and recent warm period (2000-2016). The numbers in the bracket indicate the estimated slope, and the asterisks denote statistically significant trends (at 95% confidence level).

The global averages of CDHW$_r$ depict a significant (at 95% confidence interval) increasing trend during the whole period, 1983-2016, and in the recent warmer period, 2000-2016 across all three datasets (Fig 1(a-c)). The magnitude of the estimated slope suggests a 0.54-0.55% increase in CDHW$_r$ per year during the recent warmer period in all three datasets, a rate of increment almost double of that (0.29% per year) observed during the whole period in the GPCC-CPC dataset. The relative cooler past period, 1983-1999, on the other hand, shows a decreasing trend in the CDHW$_r$ values in the GPCC-CPC dataset and a negligibly low rate of increase in the GPCC-BE and ERA5 dataset. These results indicate a consistent and significant influence of warming on the recent increases in the drought-related HW events across the globe. The percentage of land area that witnessed a CDHW event (shown in bar plots) has also increased simultaneously in the recent warmer period in all three datasets (Fig 1(d-f)). Both CDHW$_r$ magnitude and spatial extent of CDHW events are found to be the maximum during the year 2012 across all three datasets. Almost 28% of the total HW events occurred during extreme drought weeks in the year 2012 (Fig 1(a-c)), affecting 27-29% of the global land area (Fig 1(d-f)).
3.2. Spatiotemporal changes in CDHW event characteristics

3.2.1. Global and Hemispherical Changes

The spatiotemporal patterns of the CDHW characteristics are quantified based on three metrics, such as the CDHW\(_f\), CDHW\(_d\), and CDHW\(_s\) for the recent warmer period, 2000-2016 with respect to the relatively past period, 1983-1999. The spatial maps for the CDHW\(_f\), CDHW\(_d\), and CDHW\(_s\) for the two periods are presented in Figure S2, Figure S3, and Figure S4, respectively, corresponding to the three combinations of datasets chosen for the study. The corresponding changes between the two time periods are quantified by subtracting the magnitude of CDHW characteristics between recent and past periods (Figure 2).

In most of the regions, a relatively higher frequency (CDHW\(_f\)) of CDHW events (1 to 3 events per year) are observed during the recent period (Figure 2(a-c)). These increases are consistent in all three datasets (GPCC-CPC, GPCC-BE, and ERA5) over the United States, the Amazon basin, central and northern Europe, west and central Asia, and most parts of Australia. Similarly, these regions witness a relative increase in the magnitude of two other CDHW metrics, CDHW\(_d\), and CDHW\(_s\) during the recent period. The CDHW events in these locations are characterized by 2-10 days/year increase in CDHW\(_d\) (Figure 2f) and 6-30 per event/per year increase in CDHW\(_s\) (Figure 2i) during the recent period (2000-2016). Interestingly, the results corresponding to the GPCC-BE and ERA5 datasets show more consistent spatial patterns that differ in the GPCC-CPC dataset over Western South America, North of Northern America, India, China, and most of Africa.
Figure 2. Spatial map showing the changes in (a-c) CDHW$_f$ between two time periods (recent, 2000-2016 minus past, 1983-1999) based on (a) GPCC-CPC dataset, (b) GPCC-BE dataset, (c) ERA5 dataset, (d-f) same as in (a-c) but for the CDHW$_d$ (days/year), and
(g-i) same as in (a-c) but for the CDHWs (per event/year). The numbers illustrated in (a-i) represent the percentage of area affected by an increase in the corresponding CDHW metrics for the NH (red), globe (blue), and SH (black).

Our result suggests that 58-73%, 60-75%, and 56-65% of land area in the NH show an increase in CDHWᵢ, CDHWᵣ, and CDHWₛ, respectively, whereas 51-74%, 54-76%, and 49-65% of the land area witnessed such increase over the SH. These results indicate the presence of a spatial asymmetry, with a higher percentage of area in the NH being affected by the more extended duration and more severe CDHW events as compared to the SH during the recent period as a result of warming. Importantly, these results are robust across the three different combinations of datasets employed in the analysis.

3.2.2. Latitudinal Variations

The latitudinal variations of CDHW metrics are examined to further investigate their spatially asymmetric behavior due to warming at the global scale (Figure 3, and S5). In this study, the latitudinal variations (represented by the heatmaps) are captured at 1° intervals between -54.5°N and 65.5°N. Figure S5 depicts these variations for CDHWᵢ, CDHWᵣ, and CDHWₛ during individual years within the whole study period based on each of the three different combinations of datasets. The latitudinal variation of the CDHW metrics are also compared among the recent and past period based on their corresponding 17-year averages and associated interannual trends (illustrated by the line plots in Fig 3. The latitudes with statistically significant (at 95% confidence level) trends
are indicated by the “square” symbols. The magnitude of the trends and significant
directions (increasing/decreasing) are estimated based on the Sens’s slope estimator and
the MK test [47,48], respectively (see A.3 in the Supplementary Information). This
analysis can reveal the temporal shifts and rate of change of CDHW metrics across the
latitudes as an impact of warming.

The heatmaps show considerable spatial asymmetry for each of the CDHW
metrics across the latitudes that seem to be characterized by a skewed temporal pattern in
response to the warming climate (Fig S5) in all three datasets. This asymmetric behavior
is prominent across the mid-latitude regions (tropical and extratropical regions) of both
hemispheres, with relatively more consistent patterns observed in the northern (southern)
hemisphere in the GPCC-CPC (GPCC-BE) and ERA5 dataset. Peak magnitudes of
CDHWf, CDHWd, and CDHWs in the past period are only limited to a short time period
(1987 to 1989) clustered over the extratropical region (between -30°N to -54.5°N) of the
SH. However, such momentary amplifications can be attributed to sudden increases in
hot and dry events across South America influenced by natural climate variability [50].
On the other hand, the CDHW metrics have amplified considerably during the (relatively
warmer) recent period, 2000-2016, in both hemispheres.

In all three datasets, the CDHW metrics are observed to vary substantially between NH
and SH, and the maximum difference is observed during 2012 (Fig S5). During this year,
the peak magnitudes of CDHWf (1-2 events), CDHWd (7-12 days), and CDHWs (20-25
per event) were observed mostly across the mid-latitude regions (21.5° N-65.5°N) located
in the NH. This particular latitude range encompasses the extratropical regions of Europe,
Asia, and North America. These specific regions witnessed unprecedented heat events accompanied by persistent droughts mostly triggered by anthropogenic warming [8,14,15]. The magnitude of the 2012 European heatwave is found even to exceed that of the 2003 summer heatwave events considered to be one of the most deadly event in the continent [8]. On the other hand, the higher amplification in the SH is clustered over the years 2013-2016 across the latitudes -4.5°N to -45.5°N. The amplified signals of the 2013 CDHWs emerging over the SH can be associated with the unprecedented Australian summer heat (considered as the “angry summer”) attributed to anthropogenic warming [17].

The latitudinal variation of the average CDHW metrics during the past and the recent period provides a more detailed narrative on the association between the spatially asymmetric amplification and the warming climate (Fig 3). The amplification in the recent warmer period is found to be higher over the NH as compared to the SH in all three datasets. More importantly, the broad range in which the minimum and maximum magnitude of the CDHW metrics vary across the latitudes has increased substantially in the recent period in both NH and SH. On the other hand, in all three datasets, this range is found to be relatively smaller and consistent in the past period, which amplifies uniformly across the tropical and extratropical regions. This indicates a higher spatial asymmetry in the recent warmer period relative to the cooler past period in the respective hemispheres. However, the amplifications noted across the equatorial region are overwhelming in the GPCC-BE, and ERA5 dataset that substantially differs from that
observed in the GPCC-CPC dataset. This highlights the potential uncertainties associated with the use of different temperature dataset in the region.

Figure 3 (a-c) Line plots showing the latitudinal variation in the past period (1983-1999; in black) average and recent warm period (2000-2016; in red) average of CDHWf (events/year) based on (a) GPCC-CPC dataset, (b) GPCC-BE dataset, (c) ERA5 dataset, (d-f) interannual trends in CDHW characteristics for the past (in black) and recent warm
period (in red) based on the three individual datasets, (g-l) same as in (a-f) but for CDHWd (days/year), and (m-r) same as in (a-f) but for the CDHWs (per event/year). The square symbols mark the latitudes showing statistically significant trends in the interannual variation of the CDHW metrics.

We further investigated the rate of increase in the spatial variation of CDHW characteristics based on the past and recent periods (Fig3). Although a stable behavior can be observed in the past period, more dynamics in terms of latitudinal symmetry is distinctly visible during the recent period. During the past period, a greater number of latitudes located in the NH witnessed significantly higher positive interannual trends; however, the pattern reversed diametrically during the recent warmer period as the considerably higher trends are observed over a continuous stretch of the SH, particularly between -1°N and -30°N latitudes across all the three datasets. Overall, these results indicate major spatiotemporal asymmetry and heterogeneity associated with the rate of change in CDHW characteristics. However, both the mean magnitudes and trends associated with the recent warmer period are significantly higher in the ERA5 dataset as compared to both GPCC-CPC, and GPCC-BE datasets. This may be linked to the uncertainties associated with the use of different datasets that emerge in the reanalysis-based estimates of CDHW metrics under warmer conditions.
3.3. Effect of background Aridity at Continental Scale

We investigated the possible role of aridity on the asymmetric (heterogeneous) behavior of CDHW metrics. The background aridity has a potential influence on the evaporation and land-atmospheric interactions \([3,27,51,52]\), that directly controls the drought and heatwaves. The influence on evaporation is primarily defined based on the concept of water-limited and energy-limited regimes \([53,54]\). Water-limited evaporation is dominant in the arid regimes, and it is mainly controlled by the variation in soil moisture. On the other hand, the energy-limited evaporation is common in humid regimes and regulated by the availability of surface energy. To investigate the potential influence of such type of control (water-limited and energy-limited) over the spatiotemporal changes in the CDHW events in a warming climate, we examine the CDHW metrics in the arid, transitional, and humid regimes located in six different continents across the globe.

In this study, the global arid, transitional, and humid regimes are identified based on the aridity index \((AI)\) \([55]\). The \(AI\) is calculated based on the ratio between the climatological annual-average potential evaporation and precipitation \((E_p/P)\), which represents the dryness/desertification characteristic over a region \([56,57]\). The global map of the arid \((AI > 2.25)\), transitional \((0.9 < AI \leq 2.25)\), and humid \((AI \leq 0.9)\) regimes derived based on the climatological period of 1983-2016 \([58]\) is shown in Figure 4(a). To examine if the warming climate has a significant influence on the spatiotemporal changes in the CDHW events over these regimes, we investigate the statistically significant change in the spatial distribution and shift in the median value of the CDHW metrics. To
Figure 4 (a) Spatial map showing the arid, transitional, and humid regions identified based on the Aridity Index for the climatological period, 1983-2016, (b-d) Probability density of Past, 1983-1999 period (black), and recent warm, 2000-2016 period mean CDHWs (per event per year) for the arid, transitional, and humid regions of the Globe based on (b) GPCC-CPC, (c) GPCC-BE, and (d) ERA5 datasets.
investigate if these changes are dependent on the choice of data sources, the analysis is performed based on the three different combinations of datasets separately. Firstly, the 17-year mean magnitudes of the CDHW metrics (CDHWf, CDHWd, and CDHWs) are calculated for the past (1983-1999) and recent (2000-2016) periods, separately. Finally, the spatiotemporal changes between the past and relatively warmer recent periods are investigated by comparing the probability density function of mean CDHW metrics (CDHWf, CDHWd, and CDHWs) for three different AI regimes. The probability density functions are generated by using appropriate non-parametric kernel density estimators. Figure 4 (b-d), Fig S6 (b-g), Fig S7 (b-h), Fig S8(b-h) present the probability density functions for the CDHWs, CDHWf, and CDHWd obtained for each of the three different regimes based on the globe and six continents.

The statistically significant change in the spatial distribution and shift in the median value of CDHW metrics are investigated for the recent period relative to the past period by using the two-sample Kolmogorov-Smirnov test, and Wilcoxon rank-sum test (see A.4. in the Supplementary), respectively. In all three datasets, the spatial pattern of the mean CDHW metrics across the globe exhibit a significant (at 95% confidence level) positive shift in the median and the overall distribution during the recent period in all three regimes (arid, transitional, and humid), however a relatively more prominent positive shift in the median can be observed across the arid regimes (Figure 4(b-d)).

Furthermore, a consistent increase in the CDHW metrics in the recent period is notable for the arid and transitional regimes of North America, Europe, Australia, and arid regimes of Asia (Fig S6, S7, and S8). However, considerable spatially heterogeneity
among the climate regimes as well as across the continents can be noted across the datasets for the two time periods. The inter-continental heterogeneities are notable among Africa and South America relative to the rest of the continents in the GPCC-CPC datasets. Even changes in CDHW metrics in the GPCC-CPC dataset across the humid regimes of Europe, North America, and Asia are diametrically opposite to those in the ERA5, and GPCC-BE datasets. Such notable inconsistencies across the humid regions may arise due to the uncertainties in temperature, specifically between the GPCC-CPC and GPCC-BE datasets. This can be explained by the higher sensitivity of evaporation in energy-limited conditions to changes in temperature [36,52]. As such, more significant variability in the results is noted across the humid equatorial regions such as in South America and Africa.

Overall, our results are independent of the choice of datasets across significant regions and indicate an increasing influence (control) of aridity on CDHW events during the recent period (2000-2016) compared to the past (1982-1999). However, the potential influence of arid, transitional, and humid regions on CDHW characteristics varies between the continents in a warming climate. Such inter-regime and inter-continental heterogeneities may arise due to the non-uniformity in land surface fluxes (hydrological changes) over the water-limited and energy-limited regions under global warming [25]. Besides, an increase in the spatiotemporal changes in the magnitude and variability of seasonal precipitation and evaporation [59] can amplify the CDHW event characteristics in the recent warmer period. Furthermore, these heterogeneous impacts can also be attributed to the spatiotemporal variation in soil moisture and wind speed under a
changing climate that plays a significant role in modulating the land-atmospheric interactions leading to increases in frequency, duration, and severity of CDHW events [51,52,60].

4. Discussion and Concluding Remarks

The impact of compound drought and heatwave events results in more significant socio-economic damage compared to individual events. This study investigated compound drought and heatwave characteristics, such as frequency, duration, severity, and spatial extent based on a relatively warmer recent (2000-2016) and a past period (1983-1999). The analysis is performed using three different combinations of observed and reanalysis-based datasets to examine if the results are independent of the choice of data source. In all three datasets, a significant increase in drought-related heat waves and the corresponding spatial extent is noted in the (warmer) recent period. Besides, frequency, duration, and severity of CDHW events have increased substantially during the recent period. Our results highlight the regional amplifications of drought and heatwave characteristics, which are often attributed to anthropogenic climate change [14,15,61–63]. The asymmetric influence of warming was also revealed by the latitudinal variations in CDHW event metrics and the associated temporal skewness in such variations with the NH witnessing more amplification in the recent (2000-2016) warmer climate. The positive influence of warming on the CDHW event metrics are more prominent in arid regions of the globe, and most importantly, the arid regions located in the continents of Asia, Europe, Australia, and North America. However, the CDHW characteristics over the arid regions located in Africa did not witness notable changes,
which may be due to relatively greater (10-20%) amount of precipitation in the recent decade (2000-2009) as compared to the past decade (1990-1999) across some stations located in the arid regimes within the continent [64]. A significant shift in CDHW metrics over global arid lands highlights the potential influence of background aridity and the dominant role of incoming energy and soil moisture variations in modulating the evaporation and intensification of CDHW events in a warmer climate [52]. Data-related inconsistencies in the results are mainly noted across the humid regions, especially in the equatorial belt.

The broad implication of our study is relevant to stakeholders associated with water resources (Mukherjee et al., 2020), agriculture [10,35], and health science-related disciplines. The results from this study are particularly useful for CDHW impact assessment on health risk, crop loss, and energy demands that are often neglected by the decision-makers and stakeholders. Besides, adequate measures can be implemented by developing suitable forecasting tools to provide early warnings related to CDHW events.

The quantification of the local to regional scale CDHW events are seldom challenging due to uncertainty associated with available data [3,49], and differences in the evolution of drought and heatwaves at different spatial scales and temporal resolutions [3]. For example, monthly temperature anomalies compromise the daily variation of HW events; on the other hand, the precipitation-based drought indicators potentially ignore the influence of temperature on droughts under warming scenarios. Furthermore, spatiotemporal changes and heterogeneity in the CDHW event characteristics are often associated with large scale teleconnections [3,65], variations in
soil moisture [60], magnitude and variability of seasonal precipitation and evaporation [59], as well as anthropogenic warming [5,66,67]. Therefore, integrating relative contributions from these influencing factors in a more comprehensive manner can advance our understanding of the evolution of compound events.

5. References


24. Konapala G, Mishra A. Quantifying Climate and Catchment Control on Hydrological Drought in the Continental United States. Water Resources Research [Internet].


1. Introduction

Concurrent Heat wave and drought (CHWD) events have had multiple societal and eco-hydrological impacts including loss of crop yield [1,2], increased wildfires and tree mortality [3], and health hazards[4]. It is well known that CHWD events are typically triggered by the land-atmosphere feedback processes and modulated by the surface energy budget. Climate change triggered by rise in anthropogenic activities has already accelerated such processes leading to increased frequency of CHWD events across many parts of the globe [5–8]. Other major factors that play a crucial role behind triggering such processes are the large-scale natural variabilities. For example, the interannual and interdecadal variations of natural climate are found to be instrumental in influencing the overland atmospheric circulation mainly by triggering a shift from the synoptic weather systems towards the anticyclonic regimes[9]. In addition, observational evidence suggest a poleward expansion of these regimes in the northern as well as the southern hemisphere [9]. These anticyclones are accompanied with clear skies or lack of moisture in the lower atmosphere thereby facilitating drought conditions. The lack of moisture leads to the limited availability of latent energy and the sensible heat flux increases, thereby causing the surface air temperature to rise above the normal conditions for a prolonged duration (also known as HW)[10] resulting in CHWD events. In addition to that, the rise in
surface air temperature initiates a feedback loop with the soil moisture by increasing the atmospheric demand (leading to increased evapotranspiration), therefore, end up further intensifying the existing drought condition. This feedback process is very common in the anticyclonic weather regimes and is generally referred as the soil-temperature coupling [11–13].

Relevant to the understanding of the physical processes that leads to the co-occurrence of drought and HW, drought quantification using only precipitation may lead to underestimation of drying[14], that finally multiplies the uncertainties in the estimation of CHWD characteristics. Hence, soil moisture, and surface temperature anomalies should also be incorporated in CDHW estimation using the energy budget framework. Palmer Drought Severity Index (PDSI)[15] is a comprehensive drought index that incorporates hydroclimatic variables relevant to the estimation of drought under the changing climate[16]. However, most of the previous studies do not consider the effect of soil moisture and temperature while estimating CHWD[7,17,18]. Furthermore, the large scale teleconnection patterns are more instrumental in the formation of anticyclonic regimes rather than the local climate variability[9]. These large-scale patterns should be identified, and a comprehensive analysis that incorporates both the effect of such patterns, as well as the role of anthropogenic global warming (ANT) should be implemented in the investigation of the evolution of CHWD events in the past observations. Previous studies on CHWD characteristics do not consider such effects [5,19].
For the first time, we present a global, and comprehensive analysis on the CHWD events, by integrating the hydroclimatic variables such as, precipitation, soil water, and temperature. Our primary aim will be to identify the attributable climate forcings that has had a significant role in the number of CHWD events in specific climate regions across the globe during the observed period, 1982-2016. Furthermore, we find the relationships that these forcings had with the odds of occurrence of CHWD in those climate regions in the past, and finally predict the increase in the odds of such events in the 1.5°C and 2°C warming limits with respect to the current level of warming.

2. Materials and Methods

2.1. Data

Daily global maximum, and minimum 2 m air temperature (Tmax, and Tmin) dataset provided by Climate Prediction Center (CPC) (from *CPC Global Temperature data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at* https://www.esrl.noaa.gov/psd/) was obtained at 0.5° spatial resolution for the period 1979-2017. We downloaded daily global precipitation (Pr) dataset from the Global Precipitation Climatology Center (GPCC), available at 1° spatial resolution for the temporal period 1982-2016 [20]. All these dataset were regridded and extracted at 2.5° spatial resolution. Available water content (AWC) was obtained from the global database of texture derived AWC by [21], and extracted at 2.5° spatial resolution to make it consistent with the other climate variables (Tmax, Tmin, and Pr). For the estimation of global mean temeprature change, global gridded monthly temperature anomaly data
provided by HadCRUT4[22] was obtained (from https://crudata.uea.ac.uk/cru/data/temperature/). However, these anomalies are estimated with respect to the absolute temperature for the base period, 1961-1990[23], hence, do not actually represent the rise in surface temperature due to anthropogenic activities. Therefore, we re-calculated the anomalies over the globe using the same data, but using the climatological period as the pre-industrial era (1861-1890), and then obtained the global mean temperature change (represented as “ANT” in this study). Furthermore, for estimating the warming level in the current world, global averaged monthly temperature data from 35-General Circulation models participating under the Coupled Model Intercomparison Project Phase-5 (CMIP5-GCM; Table S1) based on the Representative Concentration Pathways 8.5 (RCP8.5) were obtained (from https://climexp.knmi.nl/CMIP5/Tglobal/index.cgi?email=someone@somewhere) for the period 2008-2028. We chose the RCP8.5 scenario as it is found to match most closely with the observed emission scenario [24] as compared to the other RCPs (RCP2.6, RCP4.5, and RCP6).

In addition to that, 6 global large scale indices were used in the study as listed in Table S2. Southern Oscillation Index (SOI) and Dipole Mode Index (DMI/IOD) are the two ENSO index used in the study. The SOI is available from the Bureau of Meteorology (http://www.bom.gov.au/climate/current/soihtm1.shtml), and IOD was obtained from the NOAA Climate Prediction Centre (NOAA CPC; http://www.cpc.ncep.noaa.gov/). The monthly values of Southern Annular Mode (SAM), Arctic Oscillation (AO), North
Atlantic Oscillation (NAO), and Pacific Decadal Oscillation (PDO) were also retrieved from NOAA CPC.

2.2. Estimation of Compound Drought and Heatwave (CHWD) events:

It is well known that, drought estimation at a weekly time scale can help to retain the memory of soil temperature and moisture inherited within a short time scale. This is not only relevant for capturing the diurnal feedback loop but also produces a considerable sample size required in the statistical analysis of rare events such as the co-occurrence of HW and drought events over the observational record. In this study, we define a CHWD event as a HW event that occurred during the drought weeks over a given location and time period.

A threshold-based approach was used to obtain the CHWD events for the period 1982-2016. For each grid point location, the 10\textsuperscript{th} percentile of weekly PDSI_sc (wPDSI_sc) for the reference period, 1981-2010 were obtained as a threshold, and any wPDSI_sc value below that threshold was estimated as a drought week for the period, 1982-2016. Finally, CHWD events were selected such that the daily Tmax value exceeded the TX90pct [25–28] value for 3 or more consecutive days during these drought weeks. The TX90pct was estimated as the calender day 90\textsuperscript{th} percentile of daily Tmax over each 31-day window during the 30 years climatological period, 1981-2010[29].

2.3. Measurement of degree of Susceptibility of HW (DSHW) towards drought:

To get a measure of the degree to which it is more likely that HW and drought will co-occur in a particular grid point location than not, we estimated the DSHW based on observational record. The DSHW was estimated based on conditional formulation of
CHWD events followed by a statistical test for significance. First, probability ($pe$, and $pc$) of occurrence of two mutually exclusive extreme events, HW events with, and without an already existent drought (that influences the background state of the climate) were estimated based on the observational record across the globe. Finally, statistically significant (at 5% significance level) $pe/pc$ ratio greater than 1 were obtained using the two-proportion $z$-test over 1000 realizations of the current sample. The $pe/pc$ ratio showing a significantly greater value than 1 was obtained for each grid point locations and defined as the DSHW in this study. The detailed formulation of DSHW is provided in Appendix C in the supplemental information.

2.4. Measurement of Odds of occurrence of CHWD events

Previous studies have confirmed the link between the odds of occurrence of extreme events and other climate variables using logistic regression model[30,31]. In this study we looked at the relationship between the odds of occurrence of CHWD events in any month with the interannual variability of large-scale climate indices and changes in the global mean temperature during the period, 1982-2016.

2.4.1. Logistic Regression model

We applied a multiple-predictor based Firth logistic regression model that is a special form of Generalized linear model [32] to estimate the penalized regression coefficients corresponding to natural variability of the climate, and ANT. The Firth’s model applies penalized likelihood estimation rather than performing the conventional maximum likelihood estimation to obtain the penalized regression coefficients. The
penalization allows for convergence of the likelihood to finite estimates in conditions of separation and also with sparse data and therefore, may alleviate overfitting [33].

In our analysis, we used the following logistic regression model:

\[
\log \text{it}(Y) = \ln \left( \frac{\pi}{1-\pi} \right) = (\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n) + \beta_{\text{ANT}} X_{\text{ANT}} + \alpha,
\]

(1) where, \( \left( \frac{\pi}{1-\pi} \right) \) is the odds of having more than two CHWD events per year; \( X_1, X_1, \ldots, X_n \) are the large scale climate indices used in the model and \( X_{\text{ANT}} \) is the change in global mean temperature with respect to the pre-industrial period, 1861-1900; \( \alpha, \beta_1, \beta_2, \ldots, \beta_n \), and \( \beta_{\text{ANT}} \) are the corresponding penalized regression coefficients (or scaling factors). Once the model was fitted for the observational distribution the penalized regression coefficients were obtained that we refer as the scaling factors in this study.

2.4.2. Estimating Odd ratio for 1.5°C, and 2°C global warming

In addition to that, we tried to answer the science question- “How much more likely it will be to have a CHWD day (in a month) at 1.5°C, and 2°C global warming limits than it is at the current level of anthropogenic warming?” Analyzing the independent effect of the rise in ANT over the odds of occurrence of more than two CHWD events should be a reasonable motivation towards this objective. This was performed by changing the anthropogenic component to the different warming levels (Current, 1.5°C, and 2°C), keeping the natural component as constant in the regression model. We estimated the current level of warming based on the average of monthly
temperature anomalies (estimated with respect to the pre-industrial period, 1861-1890) for the current world. We define the current world as a 21-year window with the year-2018 at the center (2008-2028) and estimated the monthly anomalies using data from 35 CMIP5-GCMs (Table S1) based on RCP8.5 scenario. Finally, the ratio of odds (or odd ratio (OR)) of monthly occurrence CHWD day for the future warming limits (1.5 °C, and 2 °C) to that for the current warming level was estimated as,

\[
OR_{T} = \frac{\left(\frac{\pi}{1-\pi}\right)_{\text{Current}}}{\left(\frac{\pi}{1-\pi}\right)_{T}} = \exp\left(\frac{\alpha + (\beta_1X_1 + \beta_2X_2 + \ldots + \beta_nX_n) + (\beta_{\text{INT}}T)}{\text{Natural Component}}\right) = \frac{\exp(\beta_{\text{INT}}T)}{\exp(\beta_{\text{INT}}X_{\text{Current}})}
\]  

(2)

where, \( T \) is the selected warming limit of 1.5°C, and 2°C, and \( X_{\text{Current}} \) refers to the current warming level with respect to the pre-industrial period.

3. Results

3.1. Annual increase in the number of CHWD events

We show that, the number of CHWD events has increased annually during the 21st century (Post-2000) compared to that observed during the last two decades of the 20th century (Pre-2000) (Figure 1a). Figure S1a, S1b and Figure 1(a) show the spatial distribution of the average number of events during the Pre-2000 and Post-2000 period, and the corresponding changes in the same between the two periods, respectively. Figure 1(b) show the nonparametric probability density for the average number of CHWD events during the Pre-, and Post-2000 period for the globe. We also performed the Kolmogorov-Smirnov, and the Wilcoxon rank sum test to show that there is a statistically significant (at 5% significance level) difference between the distributions and medians of
Figure 1 (a) Difference between the average number of CHWD events during the Pre-2000 period (1983-1999) and Post-2000 period (2000-2016) (d) Ks density plot for average number of CHWD events during the two periods for the globe, (c) spatial distribution of the ratio of the probabilities where probability of heat wave day conditioned on drought (pe) is significantly (at 5% significance level) greater than probability of heat wave day conditioned on drought (pe), and (d) percentage area of each climate region showing significantly (at 95% confidence level) greater probability of heat wave day conditioned on drought (pe) than that conditioned on no drought (pc).

the CHWD events, respectively, between these two periods. Our analysis suggests an overall annual range of 1-5 number of events during the Post-2000 period (Figure S1b)
with major portions included in most of the climate regions showing an increase of 1-3 number of events per year (Figure 1a). Those regions include the Southern parts of WNA and CAN, Eastern NAU, eastern and south eastern SAF, North-eastern SAS, eastern ENA, Northern MED, central NEU, and almost all over WAS, CEU and NEB. In addition to that, regions such as the Southern EAS, Eastern ALA, Western CGI, and central AMZ show an increase of as high as 5 annual events during the Post-2000 period. However, we do not consider the climate regions, CGI, ALA in our further analysis due to poor quality of available data over these regions.

3.2. Degree of Susceptibility of HW (DSHW) towards drought:

We focus on finding the locations where it is significantly more likely to have HW and drought co-occur in a particular day than not based on the observational record. (Figure S2) We find that majority of grid cells show higher DSHW towards a persistent drought week \((pe/pc > 1; \text{Figure 1c})\). However, the percentage of total area showing such DSHW vary across the different climate regions (Figure 1d). For instance, the climate regions, CAN, CEU, EAS, WNA, WAS, NEB, ENA, NAU, SAS, MED, CAS, ALA, NEU, SAF, and SAU show that more than 2/3rd of the area show statistically significant degree of susceptibility of HW under an ongoing drought condition with majority of the portion showing \(pe/pc\) ratio as high as more than 5 (Figure 1c). On the other hand, climate regions such as the SEA, WAF, EAF, CAM, NAS, TIB, and CGI show that just more than half of the area show statistically significant \(pe/pc\) ratio greater than 1. In addition to that, climate regions such as SAH, WSA, SSA, and AMZ show less than even half of the total area with \(pe/pc\) ratio significantly greater than 1. Therefore, out of the 26
climate regions considered in this study, only 10 climate regions show a significant DSHW over more than 2/3rd of the total area (Figure 1d). Interestingly, all of these regions also show an increase in the number of CHWD events during the Post-2000 period as seen in Figure 1a. Consequently, we do the rest of the analysis based on these 10 climate regions.

3.2.1. Possible Drivers (Natural and Anthropogenic)

Previous studies suggest possible link between the large-scale global circulation patterns or oceanic variabilities and anticyclonic regimes in both the Northern and Southern Hemisphere [9,34–36]. Therefore, understanding and exploring such relationship is key to identify the attributable factors behind the occurrence of rare events such as the CHWD events for the climate regions that show significant DSHW towards drought over more than 2/3rd of the total area. More precisely, we explore any possible link between the monthly total number of CHWD (MT-CHWD) days and the interannual variability in the natural climate[35] as well as the influence of rise in ANT on such extremes during observational record.

We selected six (Table S2) large-scale oscillation indices that have major influence on the global climate variability in an interannual or decadal time scale. To represent the interannual variability, monthly anomalies of these large-scale climate patterns and global mean temperature were smoothed by applying a 12-month running mean filter. Since, different region has different areal extant, area weighted MT-CHWD days were estimated for all of the 10 climate regions. Figure 2a show only the statistically significant (at 5% significance level) Spearman’s partial correlation between the region-
wise area weighted MT-CHWD days and the interannual variability of the large-scale climate indices and ANT for the period, 1982-2016 for the 10 climate regions.

The results suggest that ANT has a strong influence on the observed MT-CHWD days over all of the 10 climate regions during the period 1982-2016 (Figure 2a). This agrees with previous studies that have suggested an increasing global warming footprint and its substantial impact over the increase in heat wave [37], drought [38], and CHWD events [39] at regional scale.

Figure 2 (a) Correlogram showing the significant (at 5% significance level) partial correlation between the number of monthly CHWD days and the interannual variability of large-scale climate indices during the period 1982-2016 based on non-parametric Spearman’s rho. (b) Chord diagram showing the large-scale indices chosen based on the mechanistic explanation.
In addition to that, the interannual natural variability (represented by the six large scale climate indices) also show a significant but weak correlation with the occurrences of CHWD events during the observed period.

(a) \textit{El-Nino Southern Oscillation (ENSO) Index:}

Out of the two ENSO indices used in the partial correlation analysis (SOI, and IOD), interannual variability of SOI show statistically significant positive correlation with the area weighted MT-CHWD days for the regions CNA (0.3), EAS (0.273), ENA (0.27), MED (0.13), WAS (0.29) and WNA (0.15), and negative correlation for the NAU (-0.2) (Figure 2a). It is well known that in the interannual scale, ENSO act as a dominant mode and can trigger significant changes in the weather patterns leading to temperature anomalies and shift in tropical as well as extra-tropical rainfall regimes [40]. In addition to that, strong influence of SOI in the expansion of global anticyclonic regimes has also been reported in previous studies suggesting a poleward expansion of the Hadley cell during November-April [9].

IOD show significant positive correlation for the climate regions such as CEU (0.17), CNA (0.1), EAS (0.23), MED (0.24), SAS (0.14), and WAS (0.32) (Figure 2a). The role of IOD has been suggested in the formation of anticyclonic circulation over the Eastern Asia leading to unusual summer temperature in the year 1961, and 1994 [41]. The IOD-induced divergent flow and diabetic heating originating from the Bay of Bengal around India excites the Rossby wave train propagation during summer towards the EAS climate region [42]. Impact of IOD is also linked to the circulation changes over the Europe and North America [41,43], and negative rainfall anomaly over the WAS climate
region [44]. A significant warming trend and a 10-20% reduction in rainfall is reported over the Indian subcontinent (included in the SAS climate region) over the period 1901-2012 due to rapid warming of the Indian Ocean (positive IOD phase) [45].

(b) North Atlantic Oscillation (NAO)

Strong influence of NAO over European heat wave and drought is evidenced through observational studies that suggest excitation of stationary wave train propagation facilitating anticyclonic weather regimes over the region [46]. Moreover, NAO can be associated with the North Atlantic Jet variability that has strong influence over temperature and precipitation variability over the US and Europe [31,47]. This is also evident in our correlation analysis that show statistically significant Spearman’s correlation coefficient over MED (-0.16) (Figure 2a). Besides MED, three more climate regions (EAS (-0.33), NEB (-0.16), and WAS (-0.28)) also show a significant correlation with the MT-CHWD days (Figure 2a). Except for NEB, where the Atlantic Multidecadal Oscillation (AMO) is the major driver [48], the NAO show marked influence over rainfall and temperature variations over the other climate regions, WAS [49], and EAS [50]. It is to be noted that due to the short span of the temporal period 1982-2016, we did not include the AMO in our analysis.

(c) Pacific Decadal Oscillation (PDO)

PDO show relatively strong negative correlation for the climate regions, CNA (-0.38), ENA (-0.31), and WNA (-0.11) (Figure 2a). This agrees with the previous studies that report a significant influence of anomalies over the north-eastern and tropical Pacific (PDO) on drought and heat wave events in the conterminous US [51–53]. Moreover, we
find negative correlation with the interannual variability of PDO (Figure 2a) and the number of MT-CHWD days over the EAS (-0.33), and WAS (-0.33) (Figure 2a) climate regions which is also evidenced through previous observational studies [54]. However, significant correlations between variability in PDO and climate regions such as NAU (0.11), CEU (-0.11), and MED (-0.24) indicate a possible indirect influence on the CHWD events over these regions. Therefore, we exclude such influences in the further analysis of CHWD events over these regions.

(d) Arctic Oscillation (AO) and Southern Annular mode (SAM)

SAM show positive correlation for the climate regions in the northern hemisphere such as, CEU (0.11), EAS (0.2), ENA (0.12), MED (0.22), SAS (0.19), WAS (0.4), and WNA (0.1) (Figure 2a). On the other hand, significant correlation is found for climate regions, NAU (-0.14), and NEB (0.34) in the southern hemisphere (Figure 2a). It is evidenced that positive SAM has a strong influence on the frequency and poleward expansion of anticyclones in the southern hemisphere [9,55,56] with intensification of Rossby wave in the eastern Australia. However, except for EAS [57], there is no such evidence of SAM index in the northern hemisphere therefore the impact of SAM is not considered in the further analysis of CHWD events over the climate regions, CEU, ENA, MED, SAS, WAS, and WNA. On the other hand, AO that has significant influence over the increased frequency and expansion of anticyclones in the northern hemisphere [9] also show significantly weak correlation for climate regions, CEU (0.14), CNA (0.17), NEB (0.19), and WAS (0.1). In our further analysis, we exclude the effect of AO over the climate regions such as WAS, and NEB.
Finally, based on the physical evidence provided in this section, a chord diagram is presented (Figure 2b) to show the selected large-scale climate indices along with the ANT that has a significant impact on the occurrence of CHWD events for these 10 climate regions.

3.3. Scaling factors associated with the Odds of having CHWD-day

The selected large-scale meteorological perturbations, and ANT (Figure 2b) were used as independent variables to fit the FLM (refer Method section) for the 10 climate regions. Our aim is to find the possible relationship between the odds of having at least one CHWD-day in a month and the combined effect of large-scale climate perturbations and ANT based on the observational record. The odds of having at least one CHWD-day in a month indicate the minimum possible risk associated with the increasing anomaly in these global climate patterns and ANT.

Therefore, monthly binary outcomes (0 and 1) of occurrence, and non-occurrence of CHWD day was used as the dependent variable into the FLM. To account for the anthropogenic component into the FLM, changes in the monthly global mean temperature with respect to the pre-industrial period, 1861-1890 was also added as one of the independent variables. It is to be noted that all the independent monthly variables (natural and ANT) were first smoothed by applying a 12-month running mean and then regressed against the monthly time series of the binary variable. This was done in order to preserve the information in the FLM as obtained from the partial correlation analysis. Finally, the scaling factors and their 5-95% confidence intervals (CI) obtained after fitting the FLM for each of the climate regions are shown in Figure 3. These scaling factors and their 5-
95% CI suggest the multiplicative increase ($\beta > 1$) or decrease ($\beta < 1$) in the monthly odds of a CHWD day for per unit increase in the large-scale climate indices, and ANT. In addition to that, we consider a signal from these large-scale climate variables and ANT to have been detected when the 5-95% CI do not cross zero and consider only the detected signals in our further discussion.

Figure 3 Scaling factors (coefficient of regression) and their corresponding 5-95% CI indicating the sensitivity of odds of occurrence of monthly CHWD days against the interannual variability of large-scale climate variables and ANT obtained from the FLM for the 10 climate regions. The red color indicates the scaling factors for the ANT, and the blue color indicate the same for the large-scale climate indices. The green circles with a blue cross indicate the scaling factors that are not statistically significant (at 5% significant level).

The results (scaling factor, 5% to 95% CI) from the sensitivity analysis suggest that the rise in ANT has had a statistically significant positive impact on the odds of
occurrence of CHWD days for all of the 10 climate regions, CEU (4.1, 2.9 to 5.2), CAN (3.8, 2.7 to 4.9), EAS (5.6, 4.1 to 7.3), ENA (2.9, 2 to 3.9), MED (4.2, 3 to 5.6), NAU (3.7, 2.7 to 4.9), NEB (2.2, 1.3 to 2.3), SAS (4.2, 3 to 5.4), WAS (3.2, 2 to 4.4), WNA (3.3, 2.2 to 4.5) (Figure3). These findings agree with previous studies that report a substantial increase in dry and hot spells in various regions across the globe due to rise in global warming [5–8,39]. However, depending on the climate regions, the large-scale climate oscillations show either positive or negative signals against the odds of occurrence of CHWD day. For instance, increase in SOI show a statistically significant positive relationship (scaling factor, 5% to 95% CI) for the climate regions, ENA (0.29, 0.05 to 0.53), MED (0.29, 0.07 to 0.53), WAS (0.51, 0.11 to 0.62), and WNA (0.36, 0.11 to 0.62), while a negative relationship for NAU (-0.25, -0.44 to -0.07). Similarly, significant effect of SAM can be seen for the climate regions, EAS (-0.6, -1.08 to -0.14), NAU (-0.87, -1.26 to -0.5), and NEB (0.47, 0.14 to 0.82). Increase in positive AO show a significantly increasing relationship with the odds of CHWD day for the climate regions, CEU (0.79, 0.29 to 1.3), CAN (0.72, 0.18 to 1.2), and increase in positive PDO showed a statistically significant decreasing relationship for the climate regions, CAN (-0.51, -0.88 to -0.14), and ENA (-0.379, -0.69 to -0.07). On the other hand, NAO and IOD showed significantly decreasing, and increasing relationship with the odds of CHWD day for the climate regions, EAS (-0.92, -1.8 to -0.05), and WAS (2.96, 1.53 to 4.45), respectively. However, for the climate region, SAS no statistically significant signal is found from the natural variability of the climate.
From the results it is clear that occurrence of CHWD is strongly attributable to the ANT (Figure 3), while natural variability has a very weak or no significant (in case of SAS) influence over the odds of CHWD events for the 10 climate regions. Furthermore, the overall relationship of the large-scale indices and ANT with the odds of occurrence of CHWD day (Figure 3) are found to be consistent with that obtained from the correlation analysis with the MT-CHWD events (Figure 2a) over the same climate regions. Thus, the mechanistic explanations provided based on the correlation analysis, also holds true for the sensitivity analysis (Figure 3) provided in this section.

3.4. Effect of 1.5°C and 2°C rise in Global Warming

Form the sensitivity analysis, it is clear that based on the observational record the monthly odds of occurrence of CHWD day can be attributed to the rise in ANT in almost all of the climate regions. In addition to that, the magnitude of the scaling factors for all the climate regions indicate a substantial increase in the odds is very likely with per unit rise in the ANT forcing in the future climate. Given the continuous rise in warming, it is very likely that the rise in global warming will exceed the 1.5°C and 2°C warming limits by the mid-21st century[58], thereby indicating a possibility of higher odds in the future compared to the present climate scenario. In order to see the likely level of increase, we estimated the ORs for these climate regions as the ratio of monthly odds of occurrence of CHWD day in the 1.5°C, and 2°C warming limits to that in the current warming level. Figure 4 present the two-dimensional CI plot showing the OR and the corresponding 5-95% CI for the 10 climate regions that show significant DSHW towards drought based on
the observational record (Figure 1d). We find OR (5 to 95% CI) as high as 3.5 (2.5 to 5.2), 2.6 (1.98 to 3.5), 2.5 (1.9 to 3.4), 2.5 (1.9 to 3.2), 2.4 (1.8 to 3), 2.3 (1.8 to 3), 2.1 (1.6 to 2.8), 2 (1.6 to 2.7), 1.9 (1.6 to 2.4), 1.7 (1.3 to 2) for climate regions, EAS, MED, SAS, CEU, CAN, NAU, WNA, WAS, ENA, and NEB, respectively (Figure 4). These results suggest that, more than 1.7-fold increase in the odds of CHWD day is very likely under the 1.5°C warming world than it is in the current world in all of the 10 climate regions, with 3.5-fold increase in climate region, EAS.

On the other hand, under the 2 °C warming limit, the climate regions, EAS, MED, SAS, CEU, CNA, NAU, WNA, WAS, ENA, and NEB, is most likely to show ORs of 60.8 (20.4 to 209.18), 22.1 (9.1 to 58), 20.9 (9 to 51.3), 19.5 (8.7 to 45.9), 16.1 (7.4 to 37.2), 15.7 (7.2 to 35.5), 11.36 (5 to 27.4), 10.26 (4.5 to 24.7), 8.6 (4.4 to 17.5), and 5.2 (2.6 to 10.7), respectively (Figure 4). Therefore, climate regions such as, MED, and SAS show about 20-fold increase; CEU, CNA, and NAU show more than 15-fold increase; WNA, and WAS show more than 10-fold increase, and ENA, and NEB show 5 to 8-fold increase in the 2°C warming world. More alarmingly, climate region, EAS show that the odds of having CHWD day in a month will increase by about 60-fold under the 2°C warming world. Therefore, limiting to 1.5 °C warming can mitigate more than 17-fold increase in case of climate region, EAS, 5 to 8-fold increase for climate regions, WNA, NAU, CAN, CEU, SAS, and MED, and 3 to 4-fold increase for climate regions, NEB, ENA, WAS compared to the 2°C warming limit. These results suggest pursuing active efforts to keep the warming levels well below the 2 °C limit [59].
Figure 4 Ratio of odds (OR) for 1.5°C and 2°C warming limits with respect to the current level of warming.

4. Discussion and Conclusion

The precipitation and temperature variability affected by the large-scale climate perturbations often leads to the formation of anticyclonic weather regimes. Under such circumstances, the net radiation received during the daytime becomes the primary component in the surface energy budget that heats up the land surface [11]. The heating process has been accelerated and further intensified by the increased emission of heat trapping gases due to anthropogenic activities [60] leading to increased probability of co-occurrence of HW, and drought events. In this study, observational evidence has been provided that suggest a substantial increase in the number of CHWD events per year (1-5
events per year) across various parts of the globe in the beginning of 21\textsuperscript{st} century (2000-2016).

Furthermore, the HW events were found to be unequally susceptible to the existing drought conditions in the different global climate regions. For example, out of all the climate regions, 10 regions showed a significant DSHW to the existing drought conditions over more than 2/3\textsuperscript{rd} of their corresponding total area. Monthly total number of CHWD days showed significant positive and negative correlation with the interannual variability of few large-scale climate patterns in some of these climate regions, while anthropogenic warming showed significant positive correlation over all the climate regions during the observational period, 1982-2016. Keeping in mind the various shortcomings of the correlation coefficients, such as the susceptibility to outliers and errors arising from linearization, we selected the potential large-scale climate indices based on the mechanistic explanations to avoid any statistical artifact in the results.

Furthermore, we performed an attribution study on the odds of occurrence of these CHWD days in a month during the same period over the 10 climate regions. We found a significantly positive, and multiplicative effect on the odds from the anthropogenic global warming over all of the 10 climate regions. Finally, odd ratios were estimated for these climate regions that were found to be in the range of 1.7 to 3.5, and as high as 5 to 60 at 1.5°C, and 2°C warming levels, respectively, with respect to the current world. Moreover, these odd ratios suggest about 17-fold reduction in the odds in case of climate region, EAS, 5 to 8-fold reduction for climate regions, WNA, NAU, CAN, CEU,
SAS, and MED, and 3 to 4-fold reduction for climate regions, NEB, ENA, WAS in the 1.5°C rise in global warming limit, compared to the 2°C global warming limit.

Our findings suggested that among all the climate regions, EAS is the most affected region due to the rise in anthropogenic warming. This study a preliminary understanding on the combined effects of natural variability of climate and anthropogenic warming on the CHWD events in the observed period. However, the future period may see changes in the large-scale perturbations that may compensate for the increasing anthropogenic warming. Hence, the projected results in the two future warming limits are likely to be over-estimated and therefore, the anomaly in large-scale climate patterns also need to be incorporated in analysis on the CHWD events in the future period. It is also well known that the anticyclonic weather regimes are accompanied with slow-moving jet or stationary blocking zones (caused by the relatively high-pressure ridges) that amplifies the intensity and duration of CHWD events[61], therefore, should also be investigated. In addition to that, a detailed study is necessary that looks into the multiple components of human influences, such as the land-use practices [62], increased effect of dust aerosol [63] that may have a strong influence in the co-occurrence of HW and drought at a global scale.

5. References


18. Sharma S, Mujumdar P. Increasing frequency and spatial extent of concurrent meteorological droughts and heatwaves in India. Scientific Reports [Internet].


42. Qiu Y, Cai W, Guo X, Ng B. The asymmetric influence of the positive and negative IOD events on China’s rainfall. Scientific Reports [Internet]. 2014 [cited 2018 Oct 29];4:4943. Available from: https://www.nature.com/articles/srep04943


63. Huang JP, Liu JJ, Chen B, Nasiri SL. Detection of anthropogenic dust using CALIPSO lidar measurements. Atmospheric Chemistry and Physics [Internet].
1. Introduction

Climate change has increased extreme events (e.g., drought, heatwave, and heat stress) in a warming world [1–3]. Heat stress has led to massive human morbidity and mortality in recent years [2]. The impact of high temperatures is often compounded by high atmospheric relative humidity, which slows heat dissipation from the human body, thereby adding to heat stress (HS) [4]. Numerous recent studies have reported the adverse effect of heat stress on public health and labor efficiency in the United States [6], much of which has been exacerbated by human-caused climate change [2]. Severe recent HS events in the West and the Southeast in summer 2020 are representative of the types of conditions expected to become more common in the CONUS in future [8,9].

Extreme heat events in the CONUS in recent decades have trended toward earlier springtime occurrence and increased severity, frequency, and areal extent [10]. These trends appear connected to anthropogenic warming [11–13]. Critically, increasing global temperatures bring with them increasing moisture, leading to projections of heat stress that are both larger and higher-confidence than those for temperature alone [14,15]. The impact of climate change on atmospheric dynamics such as mid-latitude planetary-wave behavior may lead to more persistent weather extremes including extreme heat events, although this is a topic of active research [16–18].
The US National Weather Service (NWS) characterizes HS using a “feels-like” apparent temperature that combines the effect of heat and humidity, called the heat index (HI) [19]. Throughout the study, HS is used as a qualitative term which describes the effect of extreme heat on the human body, whereas the HI is applied as a metric to quantify the HS. While ultimately an arbitrary definition, the HI usefully communicates risk associated with extreme HS events [2,20]. Many prior studies have employed NWS’s operational definition of excessive heat exposure as corresponding to HI in excess of 105°F (or 40.6°C) [2,21] persisting for 2 hours or more (https://www.weather.gov/bgm/heat). Many consecutive days of heat exposure lead to an increase in mortality risk even in a heat acclimated population [22,23]. In the light of such evidence, in the current study, we extend the definition of HI by examining durations ranging from a day to a week.

Climate change risk assessment has various interpretations; however, its framework remains consistent at the component levels of vulnerability, such as PI and adaptive capacity [24–28]. As such, the assessment of HS-related vulnerability can be usefully assessed with a risk-centered assessment framework applied at any one of these component levels (or a combination of them) [14,25,27]. Drawing upon this concept, in this study, we perform a climate change risk assessment on the extreme heat events in the CONUS by focusing on the potential impact (PI) of the summer (JJA) annual most severe HS in the present and several future climate scenarios.

The potential importance of changes in variance and higher-order moments of the temperature distribution, particularly given its interactions with humidity, motivate the
use of nonstationary frequency analysis in empirically assessing HS. Generalized extreme value (GEV) approaches have gained increasing popularity in this context [14,29]. Common approaches to quantify risk of exposure are based on counting the number of people exposed to certain heat events under stationarity assumptions [2,30–33]. This type of risk assessment potentially ignores the effects of non-stationarity as expressed in the form of statistical distributions for the relevant climate variables with time-dependent heavy tailedness.

Furthermore, daily fluctuations in HS severity, for values in the extreme and impactful range, often impose a considerable health burden with implications on heat-related mortality [2,14,27,34]. Even heat at moderate levels accompanied by large within-season variability is known to cause illness and death, and more broadly is challenging for heat acclimation and long-term adaptability [35,36]. Recent increases in HS have coincided with the locations of greatest socio-economic vulnerability across many parts of the US [35]. To gain a fuller sense of how HS impacts may change, we use measures of HS extremeness and temporal variability to represent severity as well as possible acclimation [2,35].

We apply these characteristics of HS to CMIP5 climate projections, in combination with projected changes in population exposure under the shared socioeconomic pathways [SSPs], to create a composite indicator quantifying the PI of HS (Cardona et al., 2012; Estoque et al., 2020). While various aspects of the risk from extreme heat events are often and productively investigated independently [31,33,37,38], we argue that from a net-impacts point of view the PI of extreme heat events is best
informed through a metric that integrates HS characteristics and the total population potentially impacted (“population exposure”) [24,28].

Building on previous work [2,25,30,31,33], we here compare PI for multiple scenarios and time periods at a grid-cell level across CONUS, allowing us to determine the changes in risk for specified levels of severity, and for short-duration variations which pose a major acclimation and adaptation challenge. Furthermore, we are able to robustly characterize the interaction between the climate-change and population-change effects [31,39] in each CONUS region. With its national scope and multi-part heat-stress representation, our study’s decomposition of the relative importance of climate and population changes has important implications for identifying targeted strategies for limiting total health risks in each region (O’Neill et al., 2017).

2. Data and Methods

2.1. Climate data and Calculation of daily Heat Index

In this study, HS is quantified based on the U.S. National Weather Service recommended heat index (HI) (https://www.wpc.ncep.noaa.gov/html/heatindex_equation.shtml). The HI was calculated based on the algorithm discussed in Anderson et al., 2013 using daily maximum air temperature ($T_{max}$) and daily mean relative humidity ($RH$) data. A brief description of the algorithm is provided in the A.1 of the Supplementary Information. For the historical analysis (observed period, 1979-2019), the Tmax was obtained directly from the Climate Prediction Center (CPC) (available at ftp://ftp.cdc.noaa.gov/Datasets/cpc_global_temp/), whereas RH was evaluated using daily Tmax data from CPC and daily-mean dew point
temperature \((T_d)\). The daily \(T_d\) was derived from the 3-hourly \(T_d\), retrieved from the high-resolution European Centre for Medium-Range Weather Forecasts Reanalysis 5 (ERA5).

The \(RH\) was estimated by using the Magnus approximation as [40],

\[
RH = \frac{\exp \left( \frac{17.625 \times T_d}{243.04 + T_d} \right)}{\exp \left( \frac{17.625 \times T_{\text{max}}}{243.04 + T_{\text{max}}} \right)} \times 100
\]

where, \(T_{\text{max}}\) and \(T_d\) are in °C.

Analysis based on the model simulations was performed using outputs from 9 global climate models (GCMs; as listed in Table S1) participating in the CMIP5 experiment (available from the Earth System Grid Federation, https://esgf-node.llnl.gov/search/cmip5/). Following common practice, we use a single realization (r1i1p1) of each constituent model [41]. The r1i1p1 stands for a single realization or ensemble member that is named based on the initial conditions out of the control simulation. Changes in heat-stress impact were investigated using bias-corrected model output for the RCP4.5 and RCP8.5 scenarios. To perform the bias correction, model-calculated HI for the historical period (1979-2005) was adjusted to match that calculated from reanalysis (CPC and ERA5) over the same period. These corrections were then applied to projected HI under RCP4.5 and RCP8.5 from 2006 to 2100. A brief description of the bias correction technique [42] applied in this study is provided in A.2. of the supplementary information. The daily \(T_{\text{max}}\) data was directly retrieved from the CMIP5 archive, whereas RH was derived from specific humidity \((q)\) and surface pressure \((P_s)\) as [2],
where $\epsilon$ = ratio of gas constants for water vapor and dry air (=0.622), $e_{s0}$ = saturation vapor pressure at $T_0$ (Pa), $L/R = 5,423$ K (latent heat of vaporization divided by the gas constant for water vapor), $T_0 = 273.15$ K and $T_{\text{max}}$ is the maximum air temperature in K. To maintain spatial consistency across the observations and the GCM simulations, all datasets were regridded and extracted at 2° spatial resolution within the CONUS region. Initially, the CMIP5 HI was evaluated for the entire period, 1979-2100, which was subsequently bias-corrected [42] based on the corresponding historical model runs.

### 2.2. Population Estimates

We retrieved observed population estimates from the GPWv4 data at a 1-km resolution [43]. GPWv4 data is only available for 2000 to 2020 at a five-year interval. Therefore, we used the population of the year 2015 to represent the current CONUS population. Gridded total population estimates for the future period are retrieved from the Shared Socioeconomic Pathways (SSPs) [44] at 1/8th degree resolution (available from https://sedac.ciesin.columbia.edu/). We obtain this dataset corresponding to the SSP2 and SSP5 scenarios available for each decade between 2000 and 2100. The SSP2 and SSP5 scenarios were selected because they reflect the moderate and highest socio-economic challenges that extreme heat may pose for mitigation and adaptation. Present and future population estimates are both aggregated to a common 2° grid resolution to match the
resolution of the climate dataset. Annual population estimates at each grid point were obtained by *linear interpolation* in time.

**2.3. Non-stationary GEV Framework to estimate High-end HS Severity**

Estimation of return values quantifies the risk associated with a given climate extreme [29]. In this study, we define the High-end HS severity as the magnitude of the 40-year return value of the HI (hereafter referred to as 40YHI), with an annual probability of exceedance of only 0.25. In our analysis, the 40YHI was estimated for every grid cell, separately, based on a non-stationary GEV framework [45]. The 40YHI is estimated for the 1d-, 3d-, and 7d-HS events, individually, by using a non-stationary GEV framework. This framework is applied to the annual values of summer extreme HS severity corresponding to the five different 40-year climate scenarios, present, RCP4.5 (near-, and far-future period), and RCP8.5 (near-, and far-future period), separately.

**2.3.1. Inclusion of Non-stationarity in GEV modeling**

We employed the Block Maxima (BM) method to restrict attention to the yearly maxima of summer HI (AM-HI) over a given time period. This method is generally applied to study the upper tail characteristics of climate extremes [45]. Our analysis is based on three time periods: historical (1980-2019), near-future (2020-2059), and far-future (2060-2099). For each time period-scenario combination, the BM method was applied to generate 40-year samples of AM-HI at each grid point location of the CONUS. The GEV is a three-parameter distribution comprising of the location (μ), scale (σ), and shape (ξ) parameter, and the theoretical cumulative distribution function can be written as [45]
The non-stationary climate information in the GEV distribution (equation 4) can be incorporated by revising the model parameters to capture the shift in the mean and change in variability within the distribution, without influencing its form. This was achieved by considering linear and non-linear time trends as covariate in the \( \mu \) parameter, and a linear time trend in the \( \sigma \) parameter of the GEV model. In our analysis, \( \xi \) parameter was kept constant, as it can be unrealistic to vary the \( \xi \) parameter as a smooth function of time [45]. Furthermore, it is assumed that by varying the scale of the distribution, it is possible to deal with situations when the tail changes modestly relative to the mean of the distribution [46], as has been previously applied to investigate the non-stationary frequency analysis of extreme events [47,48].

The time-varying GEV model adopted in the study can be denoted as

\[ x_t \mid \mu_t, \sigma_t \sim GEV(\mu_t, \sigma_t, \xi) \]

and the cumulative distribution function can be generalized as,

\[
F_{x_t}(x_t \mid \mu_t, \sigma_t, \xi_0) = \exp \left\{ - \left[ 1 + \frac{\xi_0}{\sigma_t} \left( \frac{x_t - \mu_t}{\sigma_t} \right) \right]^{-\frac{1}{\xi_0}} \right\}.
\] (5)

The location parameters, \( \mu_t \) in equation (5) can be stationary or can vary linearly or quadratically with time, \( t \) (in years) as,

\[
\mu_t = \begin{cases} 
\mu_0 \\
\mu_0 + \mu_1 t \\
\mu_0 + \mu_1 t + \mu_2 t^2 
\end{cases}
\] (6)
while the scale parameter, $\sigma$, can be stationary or can have a log-linear time (in years) dependence given as

$$\sigma_t = \begin{cases} \sigma_0 \\ \exp(\sigma_0 + \sigma_t t) \end{cases}$$

(7)

It should be noted that the scale parameter $\sigma$ is modeled with an exponential function in order to ensure that it only considers positive values. Thus, based on all possible combinations of the regression models in equation (6), and equation (7), along with the assumption of a constant shape parameter, a total of five candidate GEV models (Model-1, Model-2, Model-3, Model-4, and Model-5) were identified (as illustrated in Table S2). Among the five candidates, the best-fit GEV models are selected at each grid location based on the Bayesian Information Criteria (BIC) and Likelihood Ratio Test (LRT). The model parameters associated with the best GEV models are subsequently used to calculate the 40-year return values of annual extreme summer HS events. Procedures used in the non-stationary frequency analysis are discussed in Appendix section A.1 and A.2 in the Supplementary Information. These procedures include selection (based on BIC) and test for significance (LRT) of the GEV models; estimation of the GEV parameters; and calculation of return values and exceedance probabilities.

2.4. Estimation of Potential Impact of summer extreme HS

Our methodology for the estimation of PI of summer extreme HS builds on the IPCC’s conceptual framework of impacts, adaptations, and vulnerability assessment under climate change [24,26,28]. More specifically, we applied an indicator-based assessment technique to determine a composite indicator [25] for PI of HS. This
indicator-based assessment technique is implemented in previous related studies (Inostroza et al., 2016; Wolf & McGregor, 2013; W. Zhang et al., 2019). Moreover, the most severe extreme heat events cause a disproportionate fraction of the societal impacts, motivating a special focus on them here [52]. Here, the PI was estimated for the CONUS grid points by combining the indicator of population exposure with that of HS characteristics (40YHI and decadal trends in the spread of the warm tail), for five different climate scenarios and three different HS durations (1, 3, and 7 days). The composite indicator of PI was estimated by aggregation of these indicators. The underlying formulations are described in three main steps, as follows:

*Step 1 Defining indicator variables*

Exposure is an essential component in vulnerability and is defined as the degree of contact between a person and one or more stressors [24]. In this study, we calculated population exposure for each duration category (short (1-day), medium (3-day), and long (7-day)) of HS events based on the mean of total population estimates — for each of the five climate scenarios, separately.

HS characteristics were selected based on relevance to heat acclimatization and heat-related health adaptations [2,35]. These criteria led to our usage of (1) 40YHI, and (2) decadal trends in warm tail spread (WTS; Spangler & Wellenius, 2020). The 40YHI represents extreme HS severity which is primarily important in the context of accumulated heat burdens [2] that are challenging for heat acclimatization and adaptability. Using a non-stationary framework aids in assessing the health risk from extreme heat events whose statistics change substantially with mean warming (Batibeniz
et al., 2020; Bryan Jones et al., 2015; Mishra et al., 2017). The decadal trends in WTS refers to the trends in intra-seasonal (here, summer (JJA)) range between the maximum and median of the daily HI during each of the 40-year climate scenario-time period combinations. Recent observed WTS trends are greatest in areas of the US with large socioeconomic vulnerability, underscoring the importance of comprehensively understanding future changes [35].

Step 2: Defining a decision matrix:

A decision matrix was formed based on the three selected indicator variables for each grid location of the CONUS and each climate scenario. The general form of the decision matrix for any duration (here, 1-day, 3-day, and 7-day) event can be given as follows:

$\begin{array}{ccccccc}
\text{Columns (}C_j; n \text{ numbers)} & C_1 & C_2 & C_3 & \ldots & C_{j-1} & C_n \\
\text{Rows (}R_i; S \times m \text{ numbers)} & \downarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow \\
R_{i1} & x_{i11} & x_{i12} & x_{i13} & \ldots & x_{i1j} & \ldots & x_{i1n} \\
R_{i2} & x_{i21} & x_{i22} & x_{i23} & \ldots & x_{i2j} & \ldots & x_{i2n} \\
R_{i3} & x_{i31} & x_{i32} & x_{i33} & \ldots & x_{i3j} & \ldots & x_{i3n} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
R_{im} & x_{im1} & x_{im2} & x_{im3} & \ldots & x_{imj} & \ldots & x_{imn} \\
\end{array}$
where $x_{sij}$ is the value of the $j^{th}$ indicator variable corresponding to $s^{th}$ scenario at the $i^{th}$ grid location of the CONUS. In this study, there are $n = 3$ indicator variables; $S = 5$ climate scenarios; and $m = 210$ (resolution 2x2) total grid locations in the CONUS.

Step 3: Estimating Composite Indicator for PI

The composite indicator is estimated by a type of combining of indicators known as geometric aggregation of normalized variables based on the decision matrix. Relative to averaging, geometric aggregation helps ensure that risk is not underestimated, especially for the high-end scenarios. Using equal weights ($=1/3$) assigned to each of the indicators of the decision matrix, we finally estimate the composite indicator for a grid location, $i$ and climate scenario, $s$ as,

$$PI_{si} = \left[ r_{s1i}^{\text{pop}} \times r_{s2i}^{\text{YHS}} \times r_{s3i}^{\text{WTS}} \right]^{1/3} \quad (8)$$

where $r_{sij}$ represent the data values of the $j^{th}$ indicator variable normalized to a common 0 to 1 value range using the min-max normalization method [25].

2.5. Estimating Changes in Risk and Relative Effect of Indicators

To quantify the changes in PI of extreme HS, we calculate the risk ratio (RR) for the composite indicator of PI, given as,

$$RR = \frac{PI_{\text{future}}}{PI_{\text{present}}} \quad (1)$$

where $PI_{\text{future}}$ and $PI_{\text{present}}$ compare future and present climate scenarios such that any possible future increase is denoted by a RR $> 1$.

We use the RR concept to investigate the changes in PI of heat stress (i.e., heat-stress exposure) due to population change; climate change; and their interaction. Because
exposure is function of both climate stressor and population, it would also be interesting to look at their relative impacts when allowing one of the indicators to change while keeping the other fixed [31]. Here, the effect of climate change (population change) is investigated by fixing the population (climate change scenario) at 2015 levels (to the present climate) while allowing the magnitude of 40YHI, trends in WTS, and the population counts to change under the five different climate change scenario-time period combinations. This allows us to examine the contribution of climate change in a framework that integrates the influence of non-linearity in each effect.

3. Results

3.1. Historical Analysis and Future Projections of extreme HS severity

For this study, we use the daily-maximum HI (see Methods) to estimate the annual maximum magnitude of summer extreme HS severity for short, medium and long duration events. The severity for medium and longer duration events is calculated as a 3- and 7-day rolling average of the daily values (referred to hereafter as 3d-HS and 7d-HS). Several recent studies have emphasized these temporal aspects of heat extremes [53–55]. The gridded values of AM-HI (or extreme annual HS severity) for CONUS are derived for the historical period using a reanalysis dataset (see Methods). Future extreme HS severity projections are based on the model-ensemble average from 9 CMIP5 global-climate models (GCMs; Table S1) [see Methods]. Model-computed HI values are additionally bias-corrected [42], as described above.
Figure 1: Historical and CMIP5-model projections of extreme heat stress. (a-c) spatial average of historical observations and CMIP5 modeled projections (period, 2020-2100; shading represents the interquartile range of inter-model variability) of AM-HI representing the extreme HS severity for RCP8.5 (solid brown line) and RCP4.5 (solid red line) and corresponding Tmax (blue and light blue) for (a) 1d-HS, (b) 3d-HS, and (c)
7d-HS, (d-l) spatial maps showing the temporal mean of AM-HI for the CONUS for (d) present (1980-2019), (e) near-future (2020-2059), and (f) far-future (2060-2099) periods corresponding to 1d-HS for the RCP8.5 emission scenario; (g-i) same as in (d-f) but for 3d-HS events; (j-l) same as in (g-i) but for 7d-HS events.

Like other heat-stress metrics, HI exhibits a nonlinear sensitivity towards changes in temperature and absolute humidity, governed by the rapid decrease in physiological latent heat cooling capacity with a rise in vapor pressure [2,56]. As such, increases in temperature for a fixed relative humidity result in even larger increases in the HI. Fig. 1(a-c) demonstrates this relationship for annual extreme HS severity and corresponding maximum air temperature (Tmax) over 1979-2100. Beyond mid-century, the CONUS-mean 1d-HS (7d-HS) severity from climate-model projections based on RCP8.5 depict a steady increase resulting in HI values up to 3.6°C (2.5°C), and up to 6°C (4.6°C) warmer than the measured air temperature. The projected increases in HS severity in excess of 40.6°C (indicated as HS40.6, Fig 1(a-c)) give cause for alarm due to the high mortality rates associated with conditions this severe (Matthews et al., 2017; Wehner et al., 2016).

The maps in Fig 1(d-l), depicting the mean of AM-HI for the present (1980-2019) and RCP8.5 emission scenarios (for near-future, and far-future period), provide a more detailed illustration of the increases. In comparison to the present climate, AM-HI exceeding 40.6°C are not only expected to intensify further over the Gulf of Mexico and Southeast Atlantic coastal regions, but also to spread over the entire East, Midwest, and Great Plains under the RCP8.5 emission scenario. Such spatial patterns of increase are in
close agreement with historical analyses and with climate-model projections of increases in maximum temperature (Fig S1) based on RCP8.5 (Hogan et al., 2019; Raymond et al., 2017). These spatial patterns additionally demonstrate that the relatively slower recent rates of increases in extreme temperatures observed in the Midwest and South partially disappear when the variable of interest is instead heat stress. Note that the process of averaging multiple models smooths out spatial structures, such that the future projections appear much smoother than the present.

A significant increasing trend in the annual HS severity is also noteworthy over much of the western, south-eastern, and north-eastern CONUS during the recent past. In the future climate scenarios, however, such trends occur over the whole CONUS and indeed the globe [2]. The HS severity corresponding to the moderate (RCP4.5) emission scenario depicts a steady increase until the end of the near-future period, and thereafter remains stable throughout the far-future period (Fig 1(a-c), Fig S2, S3).

3.2. Projected Extreme Summer HS Characteristics

We investigate the extreme summer HS characteristics, such as 40YHI and trends in WTS in the five different climate scenario time period combinations. The GEV parameters and, model selection associated with the estimation of 40YHI are illustrated in (Table S2; Fig S4, Fig S5). The geographical distribution of the 40YHI magnitudes is illustrated in Fig S6. Similar to the spatial patterns exhibited by the mean annual 1d-, 3d-, and 7d-HS severity (Fig 1, and Fig S2), the corresponding 40YHI values show a projected amplification in severity over the Gulf of Mexico and Southeast Atlantic coastal plains, Appalachians, Midwest, and Great Plains regions. To further examine how
much of this spatial pattern is due to regional differences in warming versus local differences in humidity, we also estimated the 40-year return levels for the annual maximum temperature (40YTx) for the five scenario-time period combinations. In contrast to the Southeast, which is generally humid, the spatial pattern of the 40YHI matches closely to that of the 40YTx in the western CONUS, where the impact of humidity is negligible (Fig S7).

We next examine the changes in the WTS (see Methods) by investigating the spatial patterns of its decadal trends (Fig S8a-o). These trends show a relatively higher rate of increase in WTS across the Great Lakes in the RCP8.5 (far-future) scenario, apparent even for individual models (Fig S9, S10). These changes indicate a greater potential for ‘heat waves’ as opposed to general summer warming, and suggest particular challenges for heat acclimation and adaptation in this region [2,35]. On the other hand, decrease in the trends in WTS are likely in the far-future (relative to the near-future) in the RCP4.5 scenario. Such decreases may be connected to the representation of radiative forcings in the RCP4.5 scenario that increases until the mid-21st century and becomes stable thereafter [58].

3.3. Changes in Population Exposure

To examine changes in population exposure, we used mean population estimates for the near- and far-future periods under two socioeconomic pathways (SSP2, and SSP5; see methods for data sources). These scenarios were selected based on the interaction of the moderate (RCP4.5) and high (RCP8.5) levels of greenhouse gas emissions with the moderate (SSP2) and the most unconstrained (SSP5) socio-economic pathways. The
changes are investigated based on the baseline fixed to the 2015 CONUS (Present) population. The gridded mean population estimates for all selected climate scenarios are presented in Fig S11.

We investigate the changes in the likely exposed population counts to each of the indicators for HS characteristics (40YHI, and WTS) in the future climate scenarios. The distribution of the 40YHI and the CONUS population estimates is illustrated by contour plots in Fig 2(a-c) for the present and for the four future RCP-SSP time period combinations that we consider. These contour plots are indicated by the data ellipses (or envelopes) that outlines the total CONUS population exposed to the different 40YHI thresholds. A considerably larger population is likely to be exposed to more severe 40YHI in future decades, particularly under the RCP8.5-SSP5 scenario (Fig 2a-c).

The other HS indicator we use, the trends in WTS, also shows considerable increases in total population exposed to more variable summer extremes (Fig 2d-f). These simultaneous increases in HS characteristics and population exposure indicate a likely amplification in the PI of HS events over the CONUS in the RCP8.5-SSP5 scenario due to significant implications on heat acclimation and health adaptability [2,35].
Figure 2: Changes in Population Exposure and Extreme Summer HS Characteristics (a-c)
Data ellipses (or envelops) representing the probability density function evaluated from the bivariate distribution of 40YHI and population estimates for the CONUS in the five climate scenarios, (d-f) probability density for spatial distribution of decadal trends in WTS of daily summer HS severity (summer annual maximum minus median) and bar
plots indicating the total population exposure to specific trend magnitudes. Style of (a-c) is inspired by [59]. Note that the ellipses are approximations intended to show inter-scenario comparisons, rather than precisely calculated population counts.

3.4. Future Changes in Risk of Potential Impact of High-End Summer HS

Future changes in risk of PI are investigated based on the RR. Maps showing the RR for the 1d-HS (3d-HS, and 7d-HS) events are presented in Fig 3(a-l), (Fig S12) which are further summarized by boxplots in Fig 3(m-o) based on the census regions (Fig 3(q)) within the CONUS. These indicate more than 3x likely increase in risk of PI due to the total effect of climate and population changes in the northern Midwest, coastal Pacific Northwest, central California, and northern Utah under the RCP8.5-SSP5 emission scenario, and a 2.5x increase in the Southeast Piedmont, northern Texas, and portions of the Southwest. The climate change effect is greatest in central California, coastal Pacific Northwest, and the Great Lakes and commensurate with the considerable increase in 40YHI and decadal trends in WTS (Fig S6-S8). The population-only effects contribute substantially to the other regions mentioned. Overall, the effect of population dynamics (Fig S11), and trends in WTS (Fig S8-S10) is also reflected in the changes in the risk of PI of HS over the CONUS in the four RCP-SSP combinations (Fig 3 and Fig S12).
Figure 3 Future Changes in Risk of Potential Impact of High-End Summer HS (a-l) Spatial distribution of RR showing the changes in risk of PI of HS in the RCP4.5-SSP2 and RCP8.5-SSP5 scenario in the near-future and far-future period due to the effect of (a-d) total effect from the interaction between climate change and population, (e-h) climate change only, (i-l) population only for 1d-HS events, and (m-o) boxplots showing the spatial distribution of these changes for the census regions (Mid-west, Northeast, South, and West of CONUS shown in (p)) based on these effects (indicated in a-l) for the 1d-, 3d-, and 7d-HS events.

The box and whisker plots in Fig 3(m-o) provide a more vivid comparison between the RRIs in the four future RCP-SSP combinations based on the regions of the CONUS and also highlight the contribution from the total and individual effect of climate and population changes. The likely increases in RR are greatest in the RCP8.5-SSP5 far-future scenario for 1d- and 3d-HS events, with a relatively greater contribution from climate change only (population change only) in the Midwest and West (South). Interestingly, in the Northeast, a greater amplification in RR from climate change alone is noted for the longer duration (3d-HS, and 7d-HS) events, perhaps reflective of model support for slower-moving Rossby waves in this region, a projection which is highly uncertain [17]. On the other hand, for the 7d-HS events, the climate and population changes exhibit a fairly equal contribution in the south and west of CONUS.

There are important caveats that should be kept in mind concerning our results. Recent-generation climate models such as those used in the CMIP5 multi-model
simulations appear unable to faithfully capture some atmospheric mechanisms such as planetary wave resonance, that are implicated in the observed increase in both the magnitude and persistence of key recent extreme heat events (Hogan et al., 2019; Mann et al., 2017b, 2018b). Such limitations may lead to a systematic underestimate of increases in extreme events.

4. Summary and Conclusion

Our findings indicate a substantial increase in the risk of PI from both short and long duration extreme HS in several parts of the CONUS. The possible increases noted over the Pacific Northwest, central California, and a major portion of Mid-west will most likely be dominated by global-mean warming. On the contrary, the risk of PI is likely to increase by more than two times due to population growth alone in the Northeast, Southeastern Piedmont, coastal Pacific Northwest, and parts of Texas, Utah, and California.

Being based on severity and temporal variability relative to local climate, as well as changing population patterns, the integrated indicator for PI that we derive applies equally well across the CONUS, adding a new element to a growing literature on extreme-heat exposure [2,25,30,31,33]. The flexible of the GEV approach allows it to be applied across a range of event types and durations. The resultant projected changes in distribution positions and shapes that we find aid in quantifying increases in risks from extreme heat exposure according to best-available knowledge of the controlling factors of acclimation and adaptability. While our current analysis focuses entirely on the potential
impact of HS, additional determinants of vulnerability such as population density by age group, underlying health conditions, and socioeconomic status are crucial [24,25,28] and would help to translate these findings into tangible adaptation and mitigation agendas.

5. References


27. Im E-S, Pal JS, Eltahir EAB. Deadly heat waves projected in the densely populated agricultural regions of South Asia. Science Advances [Internet]. American Association for the Advancement of Science; 2017 [cited 2020 Sep 6];3:e1603322. Available from: https://advances.sciencemag.org/content/3/8/e1603322


33. Mishra V, Mukherjee S, Kumar R, Stone DA. Heat wave exposure in India in current, 1.5\(\pm\)0.167em\(^{\circ}\)C, and 2.0\(\pm\)0.167em\(^{\circ}\)C worlds. Environ Res Lett [Internet]. IOP Publishing; 2017 [cited 2020 Mar 8];12:124012. Available from: https://doi.org/10.1088%2F1748-9326%2Faa9388


41. Taylor KE, Balaji V, Hankin S, Juckes M, Lawrence B, Pascoe S. CMIP5 data reference syntax (DRS) and controlled vocabularies. PCMDI: San Francisco Bay Area, CA, USA. 2011;


CHAPTER SEVEN

INVESTIGATING THE CASCADING IMPACT OF METEOROLOGICAL
FORCINGS ON EXTREME PRECIPITATION DRIVEN BY ATMOSPHERIC
RIVERS

1. Introduction

Recently, extreme precipitation events leading to massive floods have intensified in intensity, duration, and frequency in different parts of the world [1–4]. Much of this intensification is attributed to the acceleration of the hydrologic cycle caused by climate change [4–7], a considerable proportion of which is associated with large-scale atmospheric features, such as atmospheric rivers (ARs) [8]. Extreme precipitation and flooding associated with ARs have been reported in various parts of the globe [9–14], including the US [15–20].

ARs are long, narrow corridors of strong poleward water vapor transport across the mid-latitudes [21–23]. They can be associated with the warm conveyor belt of extratropical cyclones and are often characterized by strong low-level winds. They are responsible for approximately 90% of all pole-ward atmospheric water vapor transport across the mid-latitudes [23,24].

ARs have a varied range of societal and eco-hydrological impacts, which range from beneficial to devastating. For example, ARs are one of the major contributors to the snowpack and freshwater resources in the arid portions of the southwestern US and responsible for 20–50% of the region’s rainfall accumulation [25]. ARs have been associated with 30% to 60% of the annual variability of total runoff, thereby impacting
the seasonal availability of water resources in the west coast of the US [26]. Landfalling AR storms are also recognized as potential drought busters in the Pacific Northwest by overcoming nearly 60-70% of all persistent droughts in the region [27].

ARs are even known for causing megafloods, often dubbed as “ARkstorms”. These events impact thousands of square miles of urban and agricultural land across the US, disrupting millions of lives and causing massive economic damages [28]. So far, more than 99% of all reported flood damages in the western US have been attributed to severe landfalling AR storms [29]. However, the impact of ARs is not limited to the west coast of the US. In the southeastern US, the major historical 1000-year flood event that occurred on May 1-2, 2010, has been attributed to ARs [18], which affected significant parts of Tennessee, Kentucky, and Mississippi. This specific series of ARs resulted in sustained precipitation (for 48 hours) as high as 344.7 mm over Nashville, Tennessee, leading to flash flooding that resulted in 26 fatalities and caused $2–$3 billion of property damages [18].

Recognizing the potential socio-economic damages (e.g., agriculture and devastating flood damages) caused by the ARs in the SEUS, there has been a growing concern on the AR activities along the coastal southeast US (SEUS) [30]. Many studies highlighted the possible association between major flood events in the central and southeastern US with strong water vapor transport into the SEUS from multiple moisture sources, such as the Caribbean Sea, the Gulf of Mexico, and the Atlantic Ocean [18,24,31–34]. An enhanced source of moisture originating in the Caribbean and the Gulf of Mexico can trigger large-scale flood events in the mid-western US during May-July
The study detected a narrow fetch of moisture steering along the coast of Central America, extending from the Yucatan Peninsula and moving all the way upwards to the western Gulf of Mexico, where it connects with the Great Plains low-level jet (GPLLLJ), a phenomena dubbed “Maya Express” [32]. Further, [33] highlighted that much of the annual maxima floods from 1105 basins across the central US and reported that in 42.5% of the basins, more than half of the flood events are associated with ARs interacting with the SEUS. The major flooding event in the Ohio River Valley (May 1–2, 2010) was triggered by the ARs that originated from the eastern tropical Pacific and the Caribbean Sea (Moore et al., 2011) These studies highlighted the significant role of ARs in causing major flood events.

More recently, global [24] and regional studies [16,17,35,36] have provided crucial information on the climatological and synoptic aspects of ARs affecting the SEUS. More than 10% of AR landfalls along the Gulf of Mexico and Gulf stream contributed to more than 30% of the annual precipitation over most parts of the southeastern US [24]. During the period, 1980-2010, nearly 41% of the heavy precipitation events (> 100 mm/day), and 52% of the spatially widespread heavy precipitation events (affecting 7000 km² area) in the southeastern US were associated with ARs [17]. Focusing on persistent AR events with a duration of at least 48 hours, Debbage et al., 2017 reported that total ~40 persistent AR events per year occurred along the SEUS coastline locations between 1979 and 2014. The study also highlights the potential role of Great Plains Jet (GPJ) and Caribbean Low-Level Jet (CLLJ) on the SEUS ARs, which happens to interact with the western Gulf of Mexico. Miller et al.,
2017 investigated the synoptic patterns associated with minor-major flood events in the Pigeon River basin located downstream of the southern Appalachian Mountains over 5 years period. The synoptic patterns were found to reveal the role of a highly amplified slow-moving positively tilted trough formation that favored the AR-EP events over the region. Rabinowitz et al., (2019) selected 15 AR events between 2000 and 2015 across the Mississippi river valley and categorized their associated synoptic patterns as progressive troughs.

Multiple studies have highlighted the potential role of persistent atmospheric anomalies and structure of moisture transport for a better understanding of the behavior of AR-EP events [18,23,31,36,37]. More specifically, based on the available moisture sources and transport mechanism, it is possible to generate information related to the atmospheric conditions that are particularly moist and unstable at the same time. These synoptic conditions progressively intensify with time, triggering a sequence of favorable conditions for moisture transport and advection until they finally lead to the occurrence of heavy precipitation events. In such a context, “cascading effects” have been increasingly used to understand the sequence of interconnected processes that ultimately lead to extreme precipitation events. The cascading effects can be examined based on the evolution of meteorological features and moisture influx associated with the days before the AR-EP events. While previous studies explored the role of ARs in driving EP events over the SEUS, the potential influence of atmospheric anomalies on such processes was partly explored [16,17]. Furthermore, investigating the influence of atmospheric
anomalies only for coastal interactions [16] limits our understanding of such processes driving EP events across the SEUS region.

This study begins with determining the variation of EP events with respect to AR events. In the next step, the synoptic feature’s space-time evolution is investigated to understand its cascading effect on the moisture transport that influences the 100 most severe AR-EP events during the cold and warm season separately. More specifically, the primary objective of this study is to address the following questions:

(1) How are the AR events and AR-EP events distributed spatially and across different seasons in eight SEUS states (Florida, Alabama, Mississippi, Louisiana, Georgia, North Carolina, South Carolina, and Texas)?

(2) Is there a cascading effect of the synoptic-scale meteorological forcing on the moisture availability and mode of advection associated with observed major AR-EP events in the region between 1979 and 2019?

The rest of the manuscript is organized as follows: Section 2 focuses on the data and methodology applied in the study; the results and relevant discussions are provided in Section 3; and finally, the summary of major findings and concluding remarks are provided in Section 4.
2. Data and Methodology

2.1. Data

AR events are identified within the region bounded by latitudes 15°N to 45°N and longitudes -110°W to -55°W (hereafter referred to as the AR-detection region). Detection of ARs is performed by using specific humidity and zonal and meridional wind fields for 20 pressure levels (between 1,000 and 300 hPa inclusive), retrieved at 6-hourly time-steps for the period 1979-2019 from the high-resolution European Centre for Medium-Range Weather Forecasts Reanalysis 5 (ERA5) (0.25°x 0.25° grid resolution). The hurricane track data from 1979 to 2019 is obtained from the Atlantic Hurricane database (HURDAT2) to segregate the ARs from the Tropical Cyclones (TCs). The HURDAT2 is available at 6-hourly intervals provided by the National Hurricane Center (available at https://www.nhc.noaa.gov/data/#tcr). Mean sea level pressure (MSLP) and 850 mbar geopotential height (Z850) are selected to study the synoptic patterns and their cascading effect on the moisture transport mechanisms. As such, 6-hourly Z850, TCWV, and MSLP data were obtained from the ERA-5 at the native spatial resolution for the 1979-2019 period. While MSLP and Z850 are extensively used to define the large scale meteorological patterns that have a primary influence on extreme precipitation events [38], TCWV is an essential component to assess the moisture availability associated with AR-EPs [39,40].

The impact of ARs on the precipitation and EP events are studied for the SEUS region that includes the eight coastal states, namely, Alabama (AL), Georgia (GA), Louisiana (LA), Mississippi (MS), North Carolina (NC), South Carolina (SC), Texas
(TX), and Florida (FL). The EP events are derived based on the daily observed precipitation dataset (1979-2019) obtained from the Climate Prediction Center (CPC) ([https://www.esrl.noaa.gov/psd/](https://www.esrl.noaa.gov/psd/)), available at 0.5° horizontal grid resolution. The daily precipitation dataset provided by NOAA’s CPC is produced from Global Unified Gauge-Based Analysis (from over 30,000 stations) with a daily time step [41]. The CPC data has been validated against historical records and measurements at nearby stations, concurrent radar/satellite observations, and numerical model forecasts. To maintain consistency in spatial resolution across all datasets used in the study, we apply the Synergraphic Mapping System (SYMAP) (Maurer et al., 2002) to regrid the reanalysis dataset at the 0.5° CPC precipitation grid resolution. Reanalysis data available at 0.5°x0.5° is reasonable to capture the ARs and their potential influence on the EP events [8], and multiple studies investigated regional AR behavior at the same or coarser resolution [11,12,16,17,24,33].

2.2. AR-Detection Methodology

ARs are derived based on an AR-detection algorithm [24], that incorporates the characteristics of vertically integrated vapor transport (IVT) within the lower troposphere at 6-hourly time-steps. The IVT measures the integrated moisture flux based on specific humidity and zonal and meridional wind characteristics. Considering the dual drivers of AR behavior (Integrated water vapor and wind), the IVT magnitude (or intensity) is increasingly used to diagnose ARs across various parts of the globe [14,24,36,43,44]. The IVT intensity is calculated in a Eulerian framework by performing a mass-weighted vertical integration within the pressure levels ranging from 1000 to 300 hPa as follows:
$$IVT = \sqrt{\left( \frac{1}{g} \int_{1000}^{300} qudp \right)^2 + \left( \frac{1}{g} \int_{1000}^{300} qvdp \right)^2}$$

(1),

where $g$ is the acceleration due to gravity (= 9.81 m s$^{-2}$), $q$ is specific humidity (kg kg$^{-1}$), $u$ is zonal wind (m/s), $v$ is meridional wind (m/s), and $dp$ is the pressure difference (Pa) between adjacent pressure levels. The algorithm has been applied in many recent studies related to ARs [14,16,45,46]. The reader is referred to Guan and Waliser (2015) for a detailed discussion on the AR-detection algorithm, the underlying motivation, sensitivity analysis, and its application. A brief discussion on the AR-detection procedure is summarized in Appendix D1 of the Supplementary Material.

It is important to note that, at any time-step, one or more AR landfalls may be detected provided all the requirements (see Appendix D1 of the Supplementary Material) are fulfilled. The algorithm detected a total of 14992 landfalling ARs in the AR-detection region, which is 25.03% of the total 59900 six-hourly time-steps during the study period, 1979-2019.

### 2.3. AR Event and AR-EP Event Identification

The impact of the ARs on the EP events is investigated for the grid cells (0.5°x0.5°) located within the eight SEUS states. Tropical cyclones are seldom identified as a common source of heavy rainfall in the southeastern United States [17,47]. To assess the contribution from ARs alone, we have removed the effect of Tropical cyclones (TC) from the analysis of AR-EP events. This is achieved by removing the AR days that coincide with the TC dates. Out of the detected landfalling ARs, we found 12645 ARs occurring over at least one grid cell over the 1979-2019 period within or on the SEUS.
region boundary. AR duration is an important variable that controls precipitation magnitudes [48]. An AR event is identified based on the duration of ARs that occurred for three or more time steps (18 hours or more)[11,48], and these AR events can generate extreme precipitation events [11,48,49]. A similar methodology was used to examine the association between AR and extreme precipitation storms for the US’s west coast (Lamjiri et al., 2017 and Ralph et al., 2019) and Europe [11]. The AR-events are identified based on all the 12645 ARs. Two ARs were considered independent (distinct) events if they were separated by more than 1 day.

Extreme precipitation events are determined by applying the peak-over-threshold method [2,51] for the SEUS grid locations. The daily extreme precipitation events are first identified based on the 95th (EP95p) and 99th (EP99p) percentile threshold, separately. These percentiles are calculated for all non-zero daily precipitation totals observed during the whole time-period, 1979-2019. Finally, an AR-EP event is defined as the extreme precipitation event observed during or one day after an AR-event. This concept of AR events facilitates the association with the daily precipitation data and have been extensively used in previous studies (e.g., Guan et al., 2010; Huning et al., 2019; Lavers & Villarini, 2013a, 2013b; Neiman, Ralph, Wick, Kuo, et al., 2008). It should be noted that throughout the manuscript, the AR-EP events evaluated based on the 95th (99th) percentile threshold are referred to as AR-EP95p (AR-EP99p).

2.4. Composite Analysis

An event-centered composite analysis is performed to examine the evolution of the synoptic meteorological features and moisture availability and advection. So far,
event-centered composite analysis has been extensively used in numerous studies to explore the relevant synoptic-scale meteorological features associated with a specific meteorological event [38].

The event-centered composite analysis is performed based on the top 100 severe AR-EP_{99p} events, selected separately for the extended warm (May-October) and cold (November-April) seasons. Firstly, the AR-EP_{99p} time-series for individual grids (across the SEUS region) for a given season (warm and cold) are selected to form a matrix, with the dates and the corresponding AR-EP magnitude representing the columns of the matrix. The daily AR-EP magnitudes are then arranged in a decreasing order to identify the top 100 AR-EP events with unique dates for the composite analysis. The selected event days are used to obtain the seasonal anomalies in composites of meteorological variables at zero-day, 2-day, and 5-day lag periods. This approach is useful for exploring the sequence of interconnected processes (cascading effect) driving the top 100 severe AR-EP events.

More specifically, seasonal anomalies in the daily mean composites of the meteorological variables (MSLP, Z850, and TCWV) are calculated at every grid location by subtracting the mean of daily values coinciding with the time of occurrence of the selected top 100 AR-EP events from the long-term mean for the 1979-2019 period. Furthermore, a two-sample Kolmogorov-Smirnov Test (KS-Test) is performed to determine the statistical significance (at 95% confidence level) of the anomalies at every grid location. To determine the moisture advection into the region, we calculate the composite mean of the daily IVT field during which the top 100 AR-EP events are
observed for every grid location. Finally, to assess their temporal evolution over time, each of these composite and composite anomalies is further calculated for 5 days and 2 days prior to the occurrence of these top 100 AR-EP events. A similar approach has been adopted by [54] to investigate the temporal evolution of atmospheric conditions associated with AR-influenced EP along the Dutch coast.

3. Results and Discussion

3.1. Spatial distribution of EP and AR events in SEUS

EP events generally affect a varied range of areas (smaller to larger) across the SEUS throughout the year [55]. To understand the spatial patterns associated with EP and AR events in the SEUS, we constructed spatial maps showing the number of EP$_{95p}$, EP$_{99p}$, and only AR (after removing the TCs) events. Figure 1(a, c, and e) depicts the spatial distribution of the annual event counts for the EP$_{95p}$, EP$_{99p}$, and AR, respectively, for the 1979-2019 period. The corresponding results are further summarized as boxplots based on the eight SEUS states shown in Figure 1 (b, d, and f).

Figure 1(a-d) shows that except for Texas, all other states are affected by 11-16 (2-5) EP$_{95p}$ (EP$_{99p}$) events per year. A homogenous spatial pattern is noted for the EP$_{95p}$ event counts across most parts of the SEUS region, eastward of Texas. The EP$_{99p}$ event counts, however, vary considerably across the SEUS, with a higher number of events clustered in the eastern Gulf coastal plains and the western part of the Florida peninsula. This is also indicated by the boxplots in Figure 1(d) that show a comparatively greater (3-
5) number of EP99p events per year observed over Georgia, Florida, Alabama, and South Carolina.

Figure 1(e-f) shows that the rate of AR activities is found to be relatively higher along the Atlantic coastal plain of the SEUS. More than 12 AR events are observed per year over the southeastern Appalachian range extending southwards along the Atlantic coastal plain up to the southern tip of the Florida peninsula. This spatial distribution is mirrored by the boxplots depicted for the eight SEUS states, as shown in Figure 4(f). A relatively greater (12-15) number of AR events per year affects the easternmost states, including larger portions of South Carolina, Florida, North Carolina, and Georgia. Furthermore, the spatial distribution of AR event counts closely matches the spatial pattern exhibited by the EP95p events, especially along the eastern coast of NC, SC, Georgia, and Florida located along the east side of the Appalachian range.

Interestingly, the regions showing a higher frequency of AR events are found to be located along the trajectory of the Sub-tropical Low-level Jet streams (SLLJ) attributed to the Atlantic warm pool [56]. The SLLJ represents a persistent air mass influx into the subtropics and possibly supplies a significant share of moisture into the region combined with ARs that typically act along the CLLJ or westerlies in the region. This favors the influx of moisture into this region resulting in most AR activities along the coast of Atlantic coastal plains.
Figure 1 (a-b) Total EP95p event counts per year, (c-d) Total EP99p event counts per year, (e-f) AR-event counts per year for the period, 1979-2019 (after removing the TCs).

3.2. AR-EP Characteristics in SEUS

ARs have been typically associated with EP events across different parts of the globe [17,18,30,57–61]. The pronounced temporal and location-dependent behavior exhibited by the number of AR event occurrences (as seen earlier) is likely to be reflected in the EP occurrences driven by the ARs, which is the focus of this analysis. As such, using daily observations of precipitation accumulation from CPC, the AR-EP events are
identified at every grid point based on a peak-over-threshold approach, as discussed in section 2.4. Figure S1 shows the geographical distribution of 95th and 99th percentile thresholds based on which the peak-over-threshold methodology is applied to the daily precipitation totals.

Figure 2 (a-b) Spatial map showing the total number of (a) AR-EP$_{95p}$ and (b) AR-EP$_{99p}$ events observed per year during the 1979-2019 period over the SEUS region, (c-d)
boxplots showing the spatial distribution of (c) AR-EP$_{95\%}$ and (d) AR-EP$_{99\%}$ event counts per year for the period 1979-2019 for the eight states of the SEUS region, and (e-f) spatial map showing the total precipitation magnitude (mm) and the IVT vectors for the ARs identified during the 3$^{rd}$ and 4$^{th}$ May 2010. The intensity of the IVT is represented by the length of the IVT vectors.

In this section, we explore mainly two characteristics of AR-EP events, the spatial distribution of the frequency (per year) of the AR-EP events and the seasonal variation of the frequency and the average magnitude of these events across the SEUS region. The annual frequency of AR-EP events is examined at each grid point to identify any location-dependent behavior revealed by the observed AR-EP events during the selected period. Figure 2(a-d) shows the spatial distribution of the annual frequency of AR-EP95 and AR-EP99 events for the SEUS region and eight states for the 1979-2019 period. In addition to that, we also examine the precipitation characteristics (spatial distribution and magnitude) across the SEUS region for selected ARs observed during 3$^{rd}$ and 4$^{th}$ May 2010 due to their association with large scale flooding [17], as shown in Figure 2(e-f). The seasonal variations of the total AR-EP event count and average magnitudes for the 1979-2019 period are depicted by heatmaps as shown in Figure 3. The seasonal variations are supposed to reveal any temporal dependence exhibited by AR-driven EP events and their severity in the SEUS and including states.

As shown in Figure 2(a-b), the spatial patterns depicted by the annual frequency of AR-EP events are also in close agreement with that obtained for the EP and AR-events
(Figure 1). About 3-5 AR-EP_{99p} and at least 1-2 AR-EP_{99p} events per year are found to occur in the eastern part of the SEUS region, with the greatest number of events occurring primarily in the sub-tropical region along the northwestern coast of the Florida peninsula and south of Georgia. The higher rate of AR-EP events in these regions is also reminiscent of the fact that about 90% of the total water vapor flux is attributed to ARs located within the sub-tropics [23,62]. These results are further reflected by the boxplots that show the greatest number of AR-EP events occurred in Florida and Georgia (Fig 2(c-d)). These results are concordant with the spatial pattern exhibited by the EP_{99p} event counts for these regions (Figure 1(d)).

Figure 2(e-f) shows the spatial pattern and severity of the AR-EP events that cause extreme socio-economic damages in the past. We selected two consecutive ARs observed on 3rd and 4th May 2010. It can be noted that these ARs are associated with a daily precipitation total of more than 100 mm, observed across the junction between the Gulf and Atlantic coastal plain of the SEUS (Figure 2(f)). Previous studies have reported these ARs were associated with major historic flooding in the SEUS [17,18].

To explore the temporal dynamics of the AR-EP event characteristics, we examine the seasonal variation of the total event counts and average precipitation magnitude for the SEUS region and the eight including states for the 1979-2019 period. The seasonal variation of total AR-EP event counts and average precipitation magnitude are examined based on the seasons defined by a 3-month overlapping windows for all 12 months. The 3-month overlap allows for scaling the seasons to represent uniform 90-day seasons. Based on this criterion, we obtained the regional averages of seasonal counts of
AR-EP⁹⁵ and AR-EP⁹⁹ events and the corresponding average precipitation magnitudes as illustrated by the heatmaps in Figure 3(a-b) and Figure 3(c-d), respectively.

Figure 3. The heat maps (a) show the total number and (b) average magnitude of AR-EP⁹⁵ events for the overlapping seasons during the period 1979-2019, (c, and d) same as in (a, and b) but for AR-EP⁹⁹ events.
A marked cold season dependence is noticeable in the total AR-EP event occurrences. A greater number of events is noted over Georgia, Florida, Alabama, and South Carolina during the October-November-December (OND), November-December-January (NDJ), December-January-February (DJF), January-February-March (JFM), February-March-April (FMA), and March-April-May (MAM) season (Figure 3(a, and c)). These regions show a total of more than 50 AR-EP$_{95p}$ and 12-15 AR-EP$_{99p}$ events in each overlapping season between November and April during the 41-year period (1979-2019). Only a few AR-EP events (less than 5-10 events) are found to occur during the extended warm season months (April-May-June (AMJ) to September-October-November (SON)) over the entire SEUS region. These findings are in close agreement with previous studies that suggest the possible contribution of ARs to more frequent rainfall events in the SEUS with a greater influence observed predominantly during the cold season [16,17].

The anomalously high AR-EP counts, markedly noticeable during the months of the extended cold season (November-April), is generally attributed to the large-scale moisture availability and synoptically driven strong transport mechanisms acting simultaneously during the winter/transition months in the region [17,18]. In addition to that, a considerable decrease in the AR-EP event count is notable from spring to summer seasons for most of the regions. As reported by previous studies, this is due to the spring to summer decrease in the large scale and persistent mesoscale convective systems in the
west of the Appalachian Mountains, as well as due to the smaller number of ARs interacting with the CSE-US region during the warm season [16,63,64].

In contrast, the average precipitation magnitude of the AR-EP$_{95p}$ and AR-EP$_{99p}$ events is found to be more than 55-90 mm/day in Louisiana and Mississippi in July-August-September (JAS) and August-September-October (ASO) seasons. The cold season AR-EP$_{95p}$ and AR-EP$_{99p}$ events, on the other hand, are found to be characterized by only 35-60 mm daily total precipitation in the SEUS region. These results indicate that the AR-EP events in the SEUS region are more severe during the warm season than those observed in the cold season. The association between ARs and more severe EP events in the warm season could be related to higher atmospheric instability and moisture availability [55].

Overall, this seasonal-count cycle seems to be equivalent to the annual IVT cycle, reported for the West Coast of the US (Mahoney et al., 2016). The largest vapor transport occurring during the winter and markedly decreasing during the summer can be attributed to the weak transient baroclinic instability during the warm season [17]. This is also in agreement with the previous findings that associate ARs with the winter planetary wave number 4-5, a wavenumber typically associated with the synoptic-scale Extratropical Cyclones in the Northern Hemisphere [23]. Besides, the Extratropical Cyclones that are most prevalent during the winter, are climatologically associated with more number of ARs from October to March. [17].
3.3. Cascading effect of Synoptic Patterns on Moisture Availability and Transport

This section explores the synoptic-scale evolution of atmospheric patterns, moisture availability, and transport mechanisms associated with AR-EP events during the extended warm (May to October) and extended cold (November to April) season. Composites of Z850, MSLP, IVT, and TCWV (refer to Section 2.4 for methodology) are generated 2 and 5 days prior and during the day of the top 100 severe AR-EP events observed during each season. It is expected that such spatiotemporal evolution can help to explore the sequence of events, thereby revealing any cascading effect (if present) of these synoptic-scale meteorological patterns on the moisture availability and transport associated with the top 100 severe AR-EP events. Moreover, the extended warm and cold season months (considered in the composite analysis) are chosen objectively to meet the criteria of maximum seasonal dependence of AR-EP event occurrence and severity in the SEUS region (Figure 3). While such seasonal dependence of the severity of the AR-EP events is noted in the May-October months of the extended warm season for Louisiana and Mississippi, for most of the other states a greater number of AR-EP events are observed during the extended cold season months (November-April) (Figure 3).

The spatial distribution of the maximum daily precipitation totals during the selected top 100 AR-EP event days is illustrated in Figure 4. Interestingly, for the top 100 severe cold (warm) season, AR-EP events are found to exceed a severity of 150 (200) mm/day in the SEUS region, mostly clustered over the eastern Gulf coastal plains (western Gulf and Atlantic coastal plains). The most severe cold season AR-EP events affected Mississippi, Georgia, and Alabama, whereas Louisiana and Mississippi
experienced the most severe warm-season AR-EP events (Figure 4 (b, and d). These results are in close agreement with the spatiotemporal characteristics of the AR-EP events during the warm and cold seasons depicted earlier in Figure 3. We use the top 100 severe AR-EP event days observed for the warm and cold seasons in the subsequent composite analysis.

Figure 4 (a) Spatial map and (b) box plots showing the geographical distribution of maximum AR-EP magnitude for the top 100 severe AR-EP days during the extended cold (Nov-April) season for the 1979-2019 period, and (c), and (d) same as in (a), and (b), respectively, but for the extended warm (May-Oct) season.
Firstly, we examine the spatiotemporal evolution of synoptic-scale meteorological patterns relevant to the occurrence of the top 100 severe AR-EP event days by using daily composite anomaly maps of Z850 and MSLP for the warm and cold seasons. Finally, we investigate if the sequence of spatiotemporal behavior exhibited by such atmospheric patterns shows a cascading effect on the spatiotemporal evolution of moisture advection and source identification associated with the top 100 AR-EP events in the region. For that, daily mean composite maps of the IVT field and daily mean composite anomaly map of TCWV are constructed for both seasons. The obtained results are discussed in the following sub-sections.

(a) Synoptic-Scale Meteorological Patterns

Figure 5 (a-c) and Figure 5 (d-f) depict the spatiotemporal evolution of anomalies in composites of MSLP (shading) and Z850 (contour) for the cold and warm season, respectively, 5-days (indicated by “-5days”) and 2-days (indicated by “-2 days”) prior, and on the day (indicated by “0 days”) the events are observed. It can be noticed that the composite anomalies indicate a synoptic-scale meteorological pattern that exhibits a marked cascading behavior over the region during 5 to 0 days from the time of occurrence of the warm and cold season AR-EP events.

The Z850 anomaly contours for the cold season AR-EP suggest several trough and ridge (shown by positive anomalies) formation from 5 and 2 days before the AR-EP event occurrences. The trough formation (shown by negative anomalies) occurs along the south of the Rockies to the Great Plains region, while the ridge formation occurs along
the east of the Appalachian along the Atlantic coastal plains. Much of the Central Plains and the Gulf Coastal plains are found to correspond to the transition region from the trough to the ridge, which is more pronounced on the day of the events (0-days; Figure 5(c)). Such transitions are typically characterized as a zone of enhanced geostrophic winds (winds associated with the middle latitudes aloft in the troposphere), making the region more vulnerable to ARs [18,65].

MSLP anomalies suggest an eastward extension of the Bermuda high stretching southwards as the day progresses from 2-days before the day, the events are observed. This phenomenon is typical of the North Atlantic region and is mainly observed during the winter months. The extension of Bermuda High acts as a block that prevents the frontal systems from curving out to the Atlantic Ocean, which instead steers into the Gulf of Mexico. This also allows the SLLJ streams to dip further south into the SEUS region, sometimes bringing wintry storms with it, thus marking the importance of the confluence of SLLJ and ARs in relation to the AR-EP events in SEUS region. The region is characterized by a weak pressure gradient 5 days before the occurrence of the events. This pressure gradient, however, becomes stronger 2 days before the onset of the observed events. Interestingly, on the days of the events (0 days), the low-pressure anomaly can be seen to intensify and shift southeast, covering the locations where the EP events are mainly observed (defined by the rectangular bounded region in Figure 5). In addition to that, the cold season synoptic patterns depict a strong pressure dipole with a strong low-pressure center sandwiched between the extended Bermuda High and
abnormal high pressure existing over the western US. This low-pressure system in the region is likely to favor the advection of moisture from the southerly winds.

**Figure 5** Spatial map showing spatiotemporal evolution of (a-c) anomalies in composites of MSLP (shading; units hPa), and Z850 (contours; units: m) for cold (NDJFMA) season average centered on (a) 5 days before, and (b) 2 days before, and (c) on the day the top
100 severe AR-EP events are observed during the cold season; the location of the local maxima (minima) in Z850 anomaly are marked by the letter, “H” (“L”), (d-f) same as in (a-c) but for warm (MJJASO) season; striplings mark the grid points for which the anomalies in MSLP are found to be statistically significant (at 95% confidence level) based on the two-sample KS-test. The region corresponding to the top 100 severe AR-EP events is defined by the rectangle (black).

Similarly, the Z850 anomaly contours for the warm season AR-EP depict an intense low-pressure center, particularly over much of the SEUS. More importantly, the low-pressure center can be seen to intensify over the western Gulf-coastal plains and stretch slightly eastwards between 5 and 2 days before the events. This low-pressure center remains stationary over the region and intensifies until the day of the events, covering much of the Gulf coastal plain. This particular structure is indicative of cut-off lows that are often characterized by the presence of a height cap limiting the vertical movement of an air parcel, which, when moist, increases the chances of heavy rainfall [66], a phenomenon reported by [18] in connection with the severe flooding across the Ohio river basin on May 1-2, 2010. In addition to that, the MSLP pattern exhibits an intense low-pressure anomaly in the Gulf coastal plains, persistent from 2-days before the day of the events. This type of persistent low-pressure structure is indicative of slowly varying background circulations that favor deep moist convection leading to more severe warm-season EPs (Figure 3-4) in the region [55,67].
(b) Moisture Availability and Transport:

To examine the cascading effect of the synoptic-scale meteorological patterns on the moisture availability and transport mechanism associated with such AR-EP events, we calculated the composite mean of the IVT field and anomalies in the composite mean of TCWV for the same event days for the cold and warm season. The composite maps for assessing the moisture availability and transport are presented in Figure 6.

For both cold and warm season, intense IVT field over the strong low-pressure region in the Gulf Coastal Plains (where the AR-EP events are observed as depicted previously in Figure 4) can be observed to develop 2 days before the event occurrences, which amplifies substantially (IVT intensity > 400 kg/m/s) on the days when these events are observed. A dominant role of CLLJ and strong westerlies associated with the SLLJ is also notable from the IVT field for the cold season ARs. The CLLJ can be seen to cross the Yucatan Peninsula that phases with the strong SLLJ steering along the west coast of the Gulf of Mexico, bringing moisture to the region, particularly over the lower Mississippi river basin. For both seasons, the IVT field indicates that the ARs show a tendency to curve out to the center of the extended Bermuda-Azores High towards the eastern coast of Europe. This also explains the relatively higher number of AR landfalls along the Atlantic coastal plains, as indicated by the spatial distribution of AR events in Figure 1(e). Such moisture advection process conforms with the evolution of synoptic meteorological patterns exhibited by the MSLP composite anomalies presented in Figure 5 (a-c). Along the east coast of the Gulf of Mexico, the CLLJ alone can be seen to dominate the advection of moisture into the region, with the Caribbean Sea and the
Figure 6 Spatial map showing composites of IVT field based on (a) 5 days before, (b) 2 days before, and (c) on the day the top 100 AR-EP events are observed during the cold (Nov-April) season, (d-f) same as in (a-c) but for composite anomalies in TCWV, (g-i) same as in (a-c) but for warm (May-Oct) season, and (j-l) same as in (d-f) but for the
warm season. The anomalies are estimated based on the cold or warm season average, and the striplings indicate the grid points where statistically significant (95% confidence level) anomalies are obtained. The region corresponding to the top 100 severe AR-EP events is defined by the rectangle (black).

Atlantic Ocean being the primary sources of moisture. Moreover, the temporal evolution of the availability of moisture (represented by the significantly positive anomalies in TCWV) can be seen to mimic that of the IVT magnitude over the region of interest and the surrounding area of the Southeast coastal plains.

For the warm season, the IVT field associated with the ARs is relatively weaker but spread across a wider area, including both the Gulf and the Atlantic coastal plains (Figure 6(i)). The advection of moisture into the region can be primarily associated with the CLLJ bringing plumes of moisture from the Caribbean Sea and the Atlantic Ocean. On the other hand, the role of CLLJ coupled with the Pacific jet aloft is notable in the moisture transport to the Gulf coastal plains. The CLLJ merges with the strong westerlies associated with the Pacific jet aloft, and steers along the east coast of the Gulf of Mexico, transporting moisture to the lower Mississippi basin where it connects with the GPLLJ. This phenomenon, dubbed the “Maya Express”, has been previously reported in connection with the warm season flooding in the Midwest US [31,32] and closely matches with the moisture advection mechanisms reported in [16]. This is also indicative of a higher moisture flux distributed throughout the lower tropospheric column that highly increases the chance of very heavy rainfall if sufficient convective available
potential energy (CAPE) is present [36]. The transport of moisture to these regions are dominated by two distinct branches of LLJ propagation that can be seen to become stronger during 2-0 days from the event occurrences, thereby, confirming a cascading effect of the synoptic patterns on the moisture transportation by ARs into the region (Figure 5(d-f)).

Furthermore, the southeastern coastal plains are shown to exhibit an intense stationary low-pressure system associated with the warm season AR-EPs (Figure 5(d-f)). The persistent low-pressure structure favors deep moist convection leading to more severe warm-season AR-EPs in the region [55,67]. This is further reinforced by the relatively greater moisture availability associated with the warm-season AR-EPs as indicated by the significantly (13-16 mm) higher positive anomalies of TCWV on the days of the events (Figure 6(i)). The linkages between the spatial evolution of the moisture advection mechanisms and AR-related precipitation events are in close agreement with previous studies that identify the Gulf of Mexico as the primary source of moisture. This highlights the role of CLLJ and the Pacific jet aloft in the advection of moisture into the region [16,18,31–33,36].

Overall, the evolution of the source of moisture and its advection associated with the warm and cold season AR-EP events suggest the presence of a cascading effect from the synoptic-scale meteorological patterns (exhibited by the evolution of MSLP and Z850 composite anomalies in Figure 5). Our results produce new information related to the sequential development of available moisture and advection processes associated with ARs that finally lead to the most severe EP events in the SEUS region. The cold season
AR-EPs are more frequent and driven by relatively stronger dynamical systems indicated by greater IVT intensity. On the other hand, the warm season AR-EPs are more severe and caused by higher atmospheric instability and more moist conditions characterized by stationary low-pressure structures and higher TCWV anomalies.

4. Synthesis and Conclusion

We investigated the cascading impact of synoptic-scale atmospheric patterns on the advection mechanisms linked to the Atmospheric River-related Extreme Precipitation Events in the Coastal South-East US (SEUS). Spatiotemporal variation of the AR-EP event suggests a strong seasonal dependence across most states in the SEUS region. While a greater number of AR-EP events are observed during the cold season, the events are found to be more severe in the warm season. The spatiotemporal evolution of relevant synoptic features suggested a strong cascading effect on the key mechanisms (moisture availability and mode of advection) that drive the occurrence of AR-EP events during the warm and cold seasons in the SEUS region. The following conclusions can be drawn from this study.

(a) Overall, 14992 landfalling ARs were detected (after removing the TCs) from 1979 to 2019. Compared to the warm seasons, a relatively greater number of AR-activities were observed during the cold seasons. The maximum (12-15) number of AR events per year was reported for the SEUS region's easternmost states. These regions experience 11-16 EP_{95p}, and 3-5 EP_{99p} events per year.
(b) The higher number of AR-EP events were observed along the northwestern coast of the Florida peninsula and south of Georgia. In addition to that, about 3-5 AR-EP$_{99p}$ and at least 1-2 AR-EP$_{99p}$ events per year are found to occur in the entire eastern part of the SEUS region. Seasonal counts of AR-EP event occurrences and average precipitation magnitudes showed a pronounced seasonal dependence. Fifty AR-EP$_{95p}$ and twelve AR-EP$_{99p}$ events were observed over Georgia, Florida, Alabama, and South Carolina during the OND, NDJ, DJF, JFM, FMA, and MAM seasons. Only a few AR-EP events (less than 5-10 events) are found to occur during the extended warm season months (AMJ to SON) over the entire SEUS region. On the other hand, the average precipitation magnitude of the AR-EP$_{95p}$ and AR-EP$_{99p}$ events is found to be relatively greater (55-90 mm/day) in the warm seasons.

(c) The evolution of Z850 anomalies associated with the cold season AR-EP events suggested several troughs and ridge formations across much of the Central Plains and the Gulf Coastal plains. These regions are characterized by a zone of enhanced geostrophic winds indicating highly anomalous weather conditions. On the other hand, the warm season anomalies indicated a gradual strengthening of a persistent low-pressure system (2 days before the event timings) resembling that of cut-off lows in the region (covering much of the Gulf coastal plains) that might trigger enhancement of heavy rainfall events in the presence of ARs. The spatiotemporal evolution of the composite anomalies in MSLP, associated with the cold season top 100 AR-EP events, exhibited strengthening of a pressure dipole with an eastward extension of the Bermuda high- and a high-pressure center over the western US.
The moisture transport depicted by the composites of the IVT field associated with both the cold and warm season AR-EP events are found to mimic the spatiotemporal evolution of the MSLP and Z850 composites for the respective seasons. During the cold (warm) season, IVT intensity gradually strengthens 2 days (5-days, 2-days) prior to the events and steer over the SEUS region, finally moving up to the Atlantic Ocean. However, compared to the cold season, the warm season AR-EPs are associated with relatively weaker IVT strength. In addition to that, relatively moister conditions were found to be associated with the warm season AR-EPs, which is indicated by significantly higher TCWV anomalies over the affected region. The marked increase in moisture availability is attributed to the slowly varying background circulations that favor deep moist convection leading to more severe warm-season EPs in the region [55,67]. Such spatiotemporal evolution confirmed the cascading effect of synoptic patterns on the moisture availability and mode of advection associated with AR-EP events in the SEUS region.

The results from this study are expected to provide adequate understanding and background that will help predict future AR-EP events in the SEUS region. The seasonal AR characteristics presented here may be extended further to assess the seasonal water cycle influenced by ARs in the region, which can significantly aid in the advanced planning and management of water resources related to ARs. The synoptic meteorological characteristics discussed in this study can be used as important priors to design early warning and management systems related to AR-driven floods in the SEUS. Future research will aim to address the uncertainties associated with the detection of ARs
in the region [68,69], the impact of other modes of climate variabilities, or if global warming plays a significant role in modulating the statistical behavior or time of emergence of ARs in the SEUS. The possible impact of monsoon ridges on the precipitation variability and its coupling with the large scale climate variabilities [70], or the combined impact of occurrence, intensity, and duration of ARs [71] can be examined. There is also a need to explore the interplay between the ARs and other physical mechanisms such as CAPE [36] and the presence of cut-off lows [60] in a more direct fashion that could bridge specific knowledge gaps.

5. References


Strong, David, Sue Wing, Ian, Tang, Alex, Thomas, Pete, Topping, Ken, and Wills, Chris. 2011;2010–1312.


43. Mundhenk BD, Barnes EA, Maloney ED. All-Season Climatology and Variability of Atmospheric River Frequencies over the North Pacific. J Climate [Internet]. 2016


48. Ralph FM, Coleman T, Neiman PJ, Zamora RJ, Dettinger MD. Observed Impacts of Duration and Seasonality of Atmospheric-River Landfalls on Soil Moisture and


CHAPTER VIII

INVESTIGATING CLIMATE CONTROLS OF GLOBAL FLASH DROUGHTS

1. Introduction

Drought is a creeping phenomenon that slowly evolves over time and gradually spreads over large geographical areas resulting in significant environmental and socio-economic impacts [1,2]. However, a substantial number of drought events are characterized by rapid intensification over a short period, sustaining over a few days to weeks. This rapid intensification is triggered by a complex interaction between sharply elevated atmospheric demand, limited soil moisture, and abruptly higher than usual temperatures, a phenomenon termed "flash drought" [3–6]. Flash drought events are generally unforeseen and difficult to predict, which is why they end up causing devastating socio-economic impacts [7–10]. The co-evolution of highly anomalous meteorological (e.g., precipitation and temperature) conditions relative to the climatological mean and the background state (e.g., aridity) of the region serve as key precursors for the onset and propagation of flash droughts [5,11–13]. Besides, the abrupt transition from an energy-limited to a water-stressed state [14] can create favorable conditions for the onset of these events.

Recent flash drought-related studies have investigated how the onset and propagation of these events are associated with natural climate variability [13,15] or anthropogenic warming [16]. In addition to that, few regional studies have also reported increased flash drought occurrences and intensification across the humid and transitional
(sub-humid, and semi-arid) evaporation regimes [4,10,17–20]. Importantly, these studies either rely on the evaporative stress ratio (ESR) [5,21] or the soil moisture estimates [15,16] as the key variable to quantify flash droughts. Each of these indicators has its limitations and advantages [22,23]. While the evaporative stress ratio directly incorporates the near-surface state variables, the soil moisture-based indicators are more relevant for monitoring the direct impact on vegetation [5,12,22,24]. In the light of such diversification and the need for robust assessment, it is essential to investigate global flash drought characteristics based on these two unique indicators simultaneously.

Furthermore, climate anomalies and surface energy fluxes significantly control the hydrological cycle and global water-availability [25–29]. Nevertheless, so far, the response of global flash droughts to such controls remains unclear. Moreover, previous studies focused on a subset of flash drought characteristics over confined locations. As such, very limited information about the effect of different climatic conditions on the evolution of flash drought characteristics is so far available. This provides a unique opportunity to investigate the development of global flash drought characteristics (e.g., frequency and intensification) in different climate regions (e.g., arid, semi-arid, sub-humid, humid) in response to the climate controls and land-climate interactions, neither of which is explicitly nor sufficiently reported in the literature.

The above discussion highlights the limitations in flash drought research regarding the robustness of flash drought indices, climate controls, and the lack of quantitative assessments for the rapid intensification of these events. Overall, we aim to
answer the following questions to further enhance the scientific understanding of the climate vs. flash drought dynamics globally.

1. What are the global hotspots of most frequent flash droughts that are consistent across two diverse indicators applied simultaneously?

2. What is the effect of surface energy partitioning on the evolution of FD characteristics (frequency and intensity) and how ESR and root-zone soil moisture (RZSM) vary across different evaporation regimes (arid, semi-arid, sub-humid, and humid)?

3. How do the climate controls that trigger FD evolution, vary globally?

4. What are the most dominant climate precursors associated with Flash drought intensity and how does the intensity change in response to these climate anomalies in different evaporation regimes of the globe?

The results are compared based on two commonly used flash drought indices derived separately from the pentad mean evaporative stress ratio (ESR;[18]), and the root-zone soil moisture (RZSM; [16]) using data from the Global Land and Evaporation Amsterdam Model (GLEAM; [30]), European Centre for Medium-Range Weather Forecasts Reanalysis 5 (ERA5), and Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA2). The simultaneous implementation of two diverse indicators is expected to highlight consistency in results and demonstrate robust associations between flash drought characteristics and climatic conditions. This study performed a quantitative analysis to explore the potential linkages between climate controls, surface energy partitioning, and soil moisture-temperature coupling with the onset, evolution, and speed (i.e., intensity) of flash drought events globally. Our findings
suggest varying and contrasting climate controls on the development of flash drought characteristics. This can be attributed to varying response of rapidly intensifying atmospheric moisture demand and soil moisture depletion rates to the transition between energy-limited and water-limited conditions.

2. Methodology

2.1. Flash Drought Identification

Flash droughts are characterized by the rapid intensification of drought conditions over a short period. In this study, flash drought events are identified based on two different methodologies, (1) based on ESR (FD\textsubscript{ESR}) proposed by [18] and (2) based on RZSM (FD\textsubscript{RZSM}) proposed by [22] as discussed below.

(1) \textit{FD\textsubscript{ESR} Detection Methodology}

This methodology relies upon the concept of evaporative stress ratio (ESR), which is calculated based on the ratio between actual evaporation (AE) and potential evaporation (PE) as,

\[
ESR = \frac{AE}{PE} \tag{1}
\]

where \textit{ESR} ranges from zero to approximately one, such that \textit{ESR} approaching zero indicates a very high atmospheric demand for evaporation that is hardly met by the available soil moisture, thus, implying the presence of very high evaporative stress on the environment and vice versa. Daily \textit{ESR} was calculated for the period 1980 to 2018 using global gridded daily \textit{AE} and \textit{PE} dataset obtained from the three different data sources,
third version of the Global Land and Evaporation Amsterdam Model (GLEAM v3.3a; [30]) available at https://www.gleam.eu/, European Centre for Medium-Range Weather Forecasts Reanalysis 5 (ERA5), and Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA2). The GLEAM v3.3a dataset spans between 1980 and 2018 and is available daily for every 0.25° x 0.25° pixels globally. This daily evaporation dataset, provided by GLEAM, is generated based on the Priestley and Taylor (PT) evaporation model. The GLEAM uses various satellite-sensor products to provide relatively more accurate land surface evaporation estimates compared to other satellite- and model-based evaporation models [31,32]. On the other hand, the ERA5 (MERRA2) datasets are available at 0.25° x 0.25° (0.5° × 0.625°) pixels. To maintain consistency, all datasets were regridded to a common 0.5x0.5° grid resolution using a bilinear interpolation scheme. Note that the daily PE data is not available directly from MERRA2. Therefore, PE based on MERRA2 data was calculated separately using the Priestley and Taylor (PT) evaporation model (see A3. of Supplementary) to maintain consistency with the PE data provided by GLEAM.

The standardized ESR (SESR) values are used to identify flash droughts at the pentad scale. Mean pentad values of ESR were calculated and then standardized for each grid point as [18],

$$SESR_{ip} = \frac{ESR_{ip} - \overline{ESR}_{ip}}{\sigma_{ESR_{ip}}}$$  \hspace{1cm} (2)

where \(SESR_{ip}\) (hereafter referred to as SESR) is the z score of the ESR at a specific grid point \((i, j)\) for a specific pentad \(p\), \(\overline{ESR}_{ip}\) is the mean ESR at a particular grid point \((i, j)\)
for a specific pentad $p$ for all years available in the gridded dataset (here, from 1980 to 2018), and $\sigma_{ESR_{ip}}$ is the standard deviation.

Subsequently, the temporal change in SESR was calculated and standardized as,

\[
(\Delta SESR_{ip}) = \frac{\bar{\Delta SESR}_{ip} - \Delta SESR_{ip}}{\sigma_{\Delta SESR_{ip}}}
\]

(3)

where $(\Delta SESR_{ip})_z$ (hereafter referred to as $\Delta SESR$) is the z score of the change in SESR at a specific grid point $(i, j)$ for a particular pentad $p$ for all years available in the gridded dataset, and $\sigma_{\Delta SESR_{ip}}$ is the standard deviation. The SESR and $\Delta SESR$ magnitudes are finally applied to identify the flash drought events following a set of criteria and estimate the intensity of those events worldwide. A detailed explanation of the applied flash drought identification criteria and calculation of flash drought intensity is provided in Appendix F.1., and F.2., respectively, of the supplementary.

The standardized values can easily compare the evaporative stress between regions as well as evaporation regimes. Besides, SESR can be useful for robust comparison over multiple years and during the growing season for agricultural applications. It is further necessary to note that flash droughts can be identified in different ways [5,16]. The ESR based analysis has a key advantage as it directly incorporates the near-surface state variables (e.g., air temperature, wind speed, vapor pressure deficit, latent and sensible heat fluxes, soil moisture, precipitation, and shortwave radiation), which are crucial for capturing the onset, intensification, and end of flash drought [5,18].
(2) \( FD_{RZSM} \) Detection Methodology

The \( FD_{RZSM} \) events are derived based on existing methodology [16] that combines the criteria of rapid decline in RZSM and dry persistency. In this methodology, the detection of \( FD_{RZSM} \) is employed based on the following three criteria [16]:

(i) The pentad mean RZSM decreases from above 40th percentile to 20th percentile, with an average decline rate of not more than 5% in RZSM percentiles for each pentad.

(ii) The FD is considered to have terminated if the declined RZSM rises up to 20th percentile again. These two criteria determine the onset and termination stages of a flash drought event.

(iii) The drought should last for at least 4 pentads (20 days), a slight modification of the existing methodology [16] (3 pentads) to match with the FD persistency criteria applied for the detection of \( FD_{ESR} \) events.

The key advantage of this methodology lies in its ability to capture rapid changes in drying with direct relevance to vegetation health and high sensitivity towards the termination of drought events from rain. However, the existing methodology does not provide a metric for calculating the FD intensity. To calculate the \( FD_{RZSM} \) intensity, we applied a new criterion with slight modification to the existing methodology:

(iv) The mean change in RZSM percentiles during the entire duration of the flash drought must be less than 25th percentile of the climatological changes in RZSM for that grid point and time of year. The \( FD_{RZSM} \) intensity is then calculated as the mean change in RZSM percentiles during the entire FD duration.
This methodology of calculating the FD\textsubscript{RZSM} intensity is analogous to that used for calculating the FD\textsubscript{ESR} intensity (see A2 of Supplementary). Therefore, a unique classification scheme could be applied for both metrics, as illustrated in Table S1. To assess data-related inconsistencies, the FD\textsubscript{RZSM} detection methodology\cite{16} is applied to RZSM data obtained from three different datasets, GLEAM, ERA5, and MERRA2, separately.

2.2. Estimation of Aridity Index

We used the classification adopted by UNEP & Thomas, 1992, which uses the aridity index (AI) to classify the globe into different evaporation regimes: hyper-arid, arid, semi-arid, sub-humid and humid regimes [34,35]. The ranges of aridity values [34] for these five climatic regimes are illustrated in Table S1. The AI was estimated as a ratio between the annual mean precipitation and \textit{PE} for 1980-2018. The mean annual precipitation was calculated using daily total precipitation from the ERA5, whereas the annual mean \textit{PE} was calculated using the daily PE obtained from the GLEAM v3.3a. These regimes are classified based on the Aridity Index (AI) (Methods), which is derived based on the ratio between climatological mean annual precipitation and potential evaporation across the globe. The AI classification used in the study is demonstrated in Table S2, and the locations of the regimes are illustrated in the spatial map shown in Fig S2(c).
2.3. Estimation of Vapor Pressure Deficit

Two reanalysis datasets, ERA5 and MERRA2 were used separately to calculate the vapor pressure deficit (VPD) for the globe. The VPD estimation based on different variables is given as[36],

\[ AVP = 6.112 \times f_w \times e^{\frac{17.67T_d}{T_w+243.5}} \]  

(4)

where, AVP is the actual vapor pressure and \( T_d \) is the dew point temperature (ºC),

\[ SVP = 6.112 \times f_w \times e^{\frac{17.67T_a}{T_w+243.5}} \]

(5)

where, SVP is the saturation vapor pressure (kPa), and \( T_a \) is the daily mean 2m air temperature (ºC), and \( f_w \) is given as,

\[ f_w = 1 + 7 \times 10^{-4} + 3.46 \times 10^{-6} \times P_{surf} \]

(6)

where, \( P_{surf} \) is the surface pressure in (hPa).

Finally, VPD is calculated as,

\[ VPD = AVP - SVP \]

(7)

2.4. Estimation of Soil Moisture-Temperature Coupling Strength

To assess the role of land-atmosphere feedback loops on flash drought onset and intensity, we calculated the strength of soil moisture-temperature coupling during the event days using the \( \pi \) diagnostic proposed by [37] To assess the role of land-atmospheric feedback on flash drought onset and intensity we estimate the strength of soil moisture-temperature coupling by using the diagnostic \( \pi \) proposed by [37]. The diagnostic \( \pi \) is calculated as follows:
\[ pi = e' \times T'' \]  
(8)  

\[ e' = (R_n - \lambda \times AE)' - (R_n - \lambda \times PE)' \]  
(9)  

where \( R_n, AE, PE, \) and \( \lambda \) are the surface net radiation, actual evaporation, potential evaporation, and latent heat of vaporization, respectively. \( AE, PE, and R_n \) are obtained from ERA5, and MERRA2 dataset. Note that \( R_n \) data is not available directly from the MERRA2 dataset, and therefore, calculated separately as the sum of shortwave net radiation and longwave net radiation\[38\] data obtained from MERRA2. \( T \) is the daily temperature anomaly derived by using gridded daily maximum temperature data obtained from the ERA5, and MERRA2, separately. The primes in Equations (4) and (5) represent the z-score for each variable expressed as the number of standard deviations relative to the climatological mean. This framework captures the effect of soil moisture deficits (denoted by \( e \)) on the energy balance. When soil moisture is not a limiting factor (as in the case of humid regions), the actual evaporation equals the potential evaporation, and \( e \) takes up a value equal to zero. In general, the values of \( pi \) range from 0 to 1; as such, a higher value indicates a stronger positive coupling between soil-moisture and temperature. This framework is applied in the previous study to investigate the response of land atmospheric feedback on heatwaves and drought \[39,40\].
2.5. Determining the Variable Importance using Random Forest Algorithm

Random Forest (RF) is a powerful Machine Learning algorithm that works on ensemble learning method for classification and regression trees[41]. RF models have been extensively used for drought monitoring and forecasting studies[42–46], and have similar performance as some of the best-supervised learning algorithms. We used the RF model to evaluate the order of variable importance due to the high accuracy of the algorithm and because it avoids overfitting and efficiently deals with multicollinearity.

The order of variable importance in a RF model is determined based on the percentage increase in mean squared error (%IncMSE) of prediction corresponding to each predictor variable. The %IncMSE is considered as the most robust and informative measure for feature selection in a RF model. The %IncMSE in error of prediction is estimated with out-of-bag-cross validation as a result of a given variable being randomly shuffled. Therefore, the higher the %IncMSE, the higher is the importance of the predictor variable. The number of independent trees (in our case 500) are selected based on trace plots (not shown) of the %MSE. In our analysis, the RF model is employed for each evaporation regime, separately, with the yearly mean FD intensity as the decision variable, and corresponding lagged (0, 1, 2 pentads) yearly mean of standardized anomalies of Pr, Tmax, VPD, and pi as the predictors.

2.6. Estimating Sensitivity of FD Intensity to Climate Variables

Sensitivity of FD intensity was determined based on a multivariate regression model employed for each evaporation regime, separately. The same dependent variable
(FD intensity in percentiles) and lagged (by 0, 1, and 2 pentads) predictor variables, used in the RF model, are used in the regression model given as,

\[
\text{Intensity} = \sum_{n=0}^{2} \alpha_{Ln} \text{Pr}'_{Ln} + \sum_{n=0}^{2} \beta_{Ln} T_{\text{max}}'_{Ln} + \sum_{n=0}^{2} \gamma_{Ln} VPD'_{Ln} + \sum_{n=0}^{2} \phi_{Ln} p_i'_{Ln} + c
\]

(10)

where, \( \text{Pr}' \), \( T_{\text{max}}' \), \( VPD' \), and \( p_i' \) are the yearly mean of pentad standardized anomalies calculated for \( n = 0, 1, \) and 2 pentad lags (denoted as L0, L1, and L2 in Eq 10); \( \alpha, \beta, \gamma, \) and \( \phi \) are the regression coefficients, which define the sensitivity of yearly mean of FD intensity to changes in each predictor variable, and \( c \) is the intercept term.

Note that the regression model is employed for each evaporation regime, separately. The regression model's yearly predictor and dependent variables are first calculated for independent grid locations and then selected from across all grid cells within a given evaporation regime. To handle multicollinearity issues, a penalized regression technique, least absolute shrinkage and selection operator (LASSO) regression[47], was applied. LASSO uses the sum of the absolute values that is called taxicab or Manhattan distance[47,48] and deals with multicollinearity among the predictors by shrinking some coefficients and setting others to 0.

3. Results

3.1. Hotspots of FD Frequency and Effect of Surface Energy Partitioning

We investigated global hotspots of FD frequency and influence of available surface energy partitioning on the FD events and their intensities (Fig1). Two distinct types of flash drought events were analyzed in this study based on existing
methodologies[16,18]. While one method uses pentad mean ESR (referred to as FDESR)[18], the other uses pentad mean RZSM (referred to as FDRZSM)[16] to detect FD events. A detailed explanation of the FD detection procedure is provided in the Methods section. FD intensity was calculated and subsequently classified into moderate to exceptional category[18] (Table S1) as discussed in Methods and Appendix F.2 (Supplementary file). The flash drought events that exhibit moderate to exceptional intensity, capable of causing significant socio-economic losses, are considered in the analysis. The influence of surface energy partitioning was investigated based on the aridity index (AI) calculated using the mean annual Pr and PET derived from the ERA5 and GLEAM datasets, respectively (see Methods).

3.1.1. Hotspots of FD frequency

Hotspots of FD frequency were determined by calculating the percentage of years that experienced at least one flash drought event in the 1980-2018 period. Any location with more than 40-50% of years witnessing an event in both indicators (FDESR, and FDRZSM) is considered a hotspot of FD frequency. The FD frequency hotspots were analyzed and compared between multiple data sources, GLEAM (Fig 1(a, b)), ERA5, and MERRA2 (Fig S1), to improve the robustness of the analysis. Using the GLEAM data sets, both FDESR, and FDRZSM show consistent hotspot locations of high flash drought frequency across many global regions (Fig 1(a-b)). For example, more than 40-50% of flash drought years are apparent in many parts of Asia, including peninsular India, Tibet, and China, west and east Africa, central and central North America, the Mediterranean,
and east-central parts of the Amazon basin and Brazil (Fig 1(a-b)). The higher flash drought frequency, particularly in the Great Lakes Region, western North America, and Southeastern China, matches well with the regional patterns reported in previous studies [13,18,21,49]. However, the $FD_{ESR}$, and $FD_{RZSM}$ frequency hotspots vary substantially in the ERA5 and MERRA2 datasets, mainly across the mid-latitude regions (Fig S1). Such uncertainties can result from disparities in land-surface models and data assimilation techniques applied in the MERRA2 and ERA5[50–52] datasets.

3.1.2. Effect of Surface Energy Partitioning

The association between the number of flash drought occurrences and intensity with the surface energy partitioning is examined for different evaporation regimes. The FD events derived from the GLEAM dataset are used in this analysis, as they show similar spatial patterns between $FD_{ESR}$ and $FD_{RZSM}$ frequency (Fig 1(a-b)). The effect of surface energy portioning on the number of FD occurrences is examined based on contour plots (Fig 1(c-d)) showing the total number of $FD_{ESR}$, and $FD_{RZSM}$ events binned as a function of the climatological mean of annual Pr (Fig S2(a)) and PET (Fig S2(b)) for the 1980–2018 period across the globe (0.5°x0.5° pixels). These contour plots provide a meaningful association between them. Both the FD metrics ($FD_{ESR}$ and $FD_{RZSM}$) indicate a relatively higher number of events for regions with climatological daily precipitation of less than 1 mm and PET more than 0.5-3 mm (Fig 1(c-d), and Fig S2(a-b)). About 2000-5000 $FD_{ESR}$, and $FD_{RZSM}$ events occurred in those regions implying strong influence of available surface energy partitioning on rapid changes in RZSM and evaporative stress.
Naturally, precipitation variability[29] and atmospheric evaporative demand play an important role in the evolution of drought intensity[53–55]. However, these associations are usually complex due to the varying control of surface energy partitioning on terrestrial evaporation across the different evaporation regimes (arid, semi-arid, sub-humid, and humid; Fig S2(c))[55]. While evaporation in water-limited regimes (arid and semi-arid) is more sensitive to precipitation and soil moisture changes, a more prominent effect associated with changes in temperature (or incoming surface radiation) is found for energy-limited regimes (humid and sub-humid). Furthermore, investigating how RZSM, and ESR covary in the two FD definitions across these regimes is necessary to establish a more robust understanding of the association between surface energy partitioning and FD intensity. We investigate these associations by dividing the global regions into four different evaporation regimes, arid, semi-arid, sub-humid, and humid, and then examine the FD intensities across these regimes. For a given regime, we calculated the bivariate density between $FD_{RZSM}$ ($FD_{ESR}$) intensity and mean of pentad-to-pentad changes in SESR (% change in RZSM percentiles) for each FD event, detected within the regime, using the GLEAM dataset. The bivariate densities are illustrated by shaded contours in Fig 1(e-f). A higher density represented by the shaded contours indicates a greater number of flash drought occurrences within an intensity range (y-axis) conditioned on relative change in RZSM or SESR (represented in the x-axis).
Figure 1 Hotspots of FD Frequency and Effect of Available Surface Energy Partitioning

(a-b) spatial map showing the percent of FD_{ESR}, and FD_{RZSM} years between 1980 and 2018 in the GLEAM dataset, (c-d) contour plots showing the total number of (c) FD_{ESR} and (d) FD_{RZSM} events binned as a function of the climatological mean of yearly average precipitation (in mm/day/year) and PET (in mm/day/year) for the global 0.5x0.5º pixels, (e) contour plots showing the bivariate probability densities of FD intensity and (%)
change in RZSM (δRZSM (%)) calculated specific to all \( \text{FD}_\text{ESR} \) events between 1980 and 2018 over the arid, semi-arid, sub-humid, and humid regions, and (f) same as in (e) but for FD intensity and change in SESR calculated specific to all \( \text{FD}_\text{RZSM} \) events.

\( \text{FD}_\text{ESR} \) intensity shows the strongest coupling with relatively higher RZSM depletion rates in the arid regions (Fig 1e). The contour of the strongest coupling shifts gradually from arid to humid conditions, where a relatively slower rate of RZSM depletion is coupled with the \( \text{FD}_\text{ESR} \) intensity. A very similar dependence structure is noted between \( \text{FD}_\text{RZSM} \) event intensity and changes in SESR as they shift from dry to wet regimes (Fig 1f). However, almost no change in SESR (δSESR) is noticed for humid regions corresponding to the \( \text{FD}_\text{RZSM} \) events, suggesting a de-coupling between the atmospheric evaporative stress and RZSM anomalies within the FD duration. This de-coupling behavior can result from asynchronous changes in RZSM and atmospheric demand in humid regions. Such out-of-phase associations are linked to higher initial RZSM conditions and extended memory of soil-moisture in wet conditions[56].

3.3. Climate Controls Associated with Onset and Evolution of Flash droughts

Droughts are primarily triggered by climatic perturbations such as abnormal precipitation and temperature variations[28,29,57], atmospheric evaporative demand[54], and positive land-atmospheric feedback loops[37,39,58]. These climate fluctuations are a sub-set of anticyclonic conditions that inhibit low-level moisture convergences[59] at different spatial and temporal scales. The climate anomalies often have a cascading effect
on the rapid intensification of evaporative stress and soil moisture depletion leading to more favorable conditions for the onset and propagation of flash droughts. We investigate these associations in the context of onset and evolution of each FD event for every 0.5°x0.5° grid cell.

We selected pentad mean of daily total precipitation (Pr), maximum 2m air temperature (Tmax), vapour pressure deficit (VPD), and soil-temperature coupling strength (pi)[14,37] up to four lagged pentads, and three pentads after the FD onset. The standardized anomalies of each of these variables are derived with respect to their climatological pentad mean (for the 1980-2018 period). The procedures followed for calculating VPD and pi are discussed in Methods. These climate variables are obtained from two different reanalysis datasets, ERA5 and MERRA2, to improve the robustness of the results. The total counts of FD$_{ESR}$, and FD$_{RZSM}$ events are calculated at each pixel (0.5°x0.5°) and subsequently binned as a function of the standardized anomalies of each variable corresponding to the selected pentads as illustrated by 2D-contour plots in Fig 2 and Fig S3. Spatial maps in Fig S4-S7 illustrate the spatial distribution of mean of these standardized anomalies corresponding to the selected pentads.

The 2D-contour plots suggest a distinct types of association between the climate anomalies and FD onset and propagation in the two definitions. For example, the 2D-contour plots in Fig 2(a-d) and Fig S3(a-d) suggest that for the majority of FD$_{ESR}$ events, the climate anomalies and ESR anomalies intensify at the same time as the FD evolves. The magnitude of the anomalies corresponding to the maximum number of (about 5000) FD$_{ESR}$ events is observed to increase (or decrease in case of Pr) continuously from one
pentad before the FD onset and reaches the peak (lowest) magnitude after 3 pentads from the FD onset. In contrast, the climate anomalies corresponding to the majority of the FDRZSM events increase (or decrease in case of Pr) continuously two pentads before the FD onset, and reaches the peak (lowest) magnitude after 1 pentad of the FD onset, and thereafter decreases (increases for Pr) again (Fig 2(e-h), and Fig S3(e-h)). This behavior is consistent in all the selected climate variables, Pr, Tmax, VPD, and pi, in ERA5 and MERRA2 datasets, which is also reflected in the spatial maps shown in Fig S4-S5, and Fig S6-S7 for FDES and FDRZSM, respectively.

Figure 2 Climate controls associated with onset and evolution of flash droughts (a) contour plots illustrating the total number of FDES events binned as a function of
standardized Pr anomalies (y-axis) for 1ˢᵗ, 2ⁿᵈ, 3ʳᵈ, and 4ᵗʰ pentad before (denoted by -4 to -1 in the x-axis) and 1ˢᵗ, 2ⁿᵈ, and 3ʳᵈ pentad after (denoted by 1 to 3 in the x-axis) the onset of FD events (denoted by 0 in the x-axis), (b-h) same as in (a) but for standardized anomalies of (b) Tmax, (c) VPD, and (d) pi, (e-h) same as in (a-d) but for the standardized anomalies corresponding to the FD_{RZSM} events. Spatial maps in (a-d) and (e-h) illustrate the mean of standardized anomalies of the climate variables for the 3rd and 1st pentad after the onset of the FD_{ESR} and FD_{RZSM} events.

Overall, these results suggest that while the onset and evolution of FD_{ESR} show simultaneous development with changes in climate anomalies, a delayed or cascading effect of such climate anomalies is noted on the onset and evolution of FD_{RZSM}. Our results are in close agreement with regional studies that suggest a similar spatio-temporal pattern of climatic forcings specific to FD onset and evolution[3,7]. This disparate behavior underscores the effect of soil-moisture memory [56]. Past studies have highlighted the role of soil-moisture memory as a key variable that controls the land-atmospheric feedbacks contributing to the state and temporal variation in droughts[56,60,61]. For most regions, soil moisture tends to have a more extended memory which causes a delayed response of RZSM anomalies to climatic changes.

3.4. Relative Importance of Climate Controls on Flash Drought Intensity

The possible mechanisms responsible for the intensification of drought can be complex and potentially dependent on the background aridity and type of evaporation (energy or water-limited) regime[62–64]. Furthermore, transient evaporation regimes
(sub-humid and semi-arid) have distinct surface energy partitioning that affects soil evaporation in the energy-limited and water-limited regions with varying complexity[65]. Therefore, the most dominating driver of FD intensity is likely to vary significantly over the different evaporation regimes. Consequently, it is important to identify the key climatic precursors that contribute significantly to the variation of FD intensity in a given climate regime. In addition to that, the influence of these precursors often show a dynamic behavior varying from one state to another.

We used a machine learning-based random forest (RF) algorithm[42–46] to quantify the climate controls of FD_{ESR} (FD_{RZSM}) intensity in different climate regions (see Methods). The gridded FD intensities within each evaporation regime are selected in the RF model development. The annual time series of mean FD intensity is derived for all the individual grids, and then pooled together from all the grids within a given evaporation regime to be used as the RF model's predictand. The climate-controlled variables, such as standardized anomalies of Pr, Tmax, VPD, and pi for zero, 1, and 2 pentad lags (hereafter referred to as L0, L1, and L2) specific to each FD event onsets are used as predictors in the RF model. The importance of climate variables is determined based on the percentage increase in mean squared error (%IncMSE) of prediction of the RF model corresponding to each predictor variable (see Methods). A higher magnitude of %IncMSE indicates relatively higher importance of the predictor variable. The %IncMSE for all predictor variables is calculated using both ESR and MERRA2 datasets corresponding to the FD_{ESR} and FD_{RZSM} intensity, as illustrated in Fig 3.
A dominant influence of Pr(L1) anomalies based on both ERA5 and MERRA2 on the interannual variation of $F_{ESR}$ intensity is noted in the humid regimes (Fig 3(a-b)). In contrast, pi(L1) show the most dominant control over interannual variation of $F_{ESR}$ intensity for the sub-humid regimes, which is also consistent in both ERA5 and MERRA2 dataset. This can be explained by the higher transferability of soil temperature memory into atmospheric persistence in transitional regimes, which significantly affects drying rates [66].
Figure 3 Relative importance of climate controls on flash drought intensity (a-d) bar-plots showing the order of relative importance of climate variables influencing the inter-annual variation of (a-b) $FD_{ESR}$, and (c-d) $FD_{RZSM}$ intensity for the different evaporation regimes. Note that both $FD_{ESR}$ and $FD_{RZSM}$ are calculated base on the GLEAM dataset, whereas the climate anomalies are calculated based on the (a, and c) ERA5 and (b, and d) MERRA2 datasets.

For $FD_{RZSM}$, the interannual variation of the FD intensity is dominated by $T_{max}(L1)$ changes in both arid and semi-arid regions based on ERA5 and MERRA2 datasets. On the other hand, $T_{max}(L0)$ shows the most dominant control over the humid regimes, consistent across the two datasets. However, some disparities arising from using different data sources can be noted in the case of $FD_{ESR}$ in the arid and semi-arid regimes and $FD_{RZSM}$ in the sub-humid regimes. For example, based on ERA5, the most dominant control on interannual variation of $FD_{ESR}$ intensity in arid, and semi-arid regions are exhibited by $Pr(L2)$ and $Pr(L1)$, respectively. On the other hand, based on the MERRA2 dataset, $VPD(L2)$ and $pi(L1)$ have the strongest influence in these regimes. Similarly, based on the ERA5 dataset, $Pr(L0)$ has the strongest control in sub-humid regime, whereas, based on MERRA2, $T_{max}(L0)$ shows the most dominant control on interannual variation of $FD_{RZSM}$ intensity.
3.5. Sensitivity of Flash Drought Intensity to Climate Variables

The sensitivity of FD intensity is evaluated by comparing the relative changes of $FD_{ESR}$ and $FD_{RZSM}$ intensity with respect to the changes in the climate variables (Pr, Tmax, VPD, and pi). We applied a penalized regression approach based on the Least Absolute Shrinkage, and Selection Operator (LASSO) technique\textsuperscript{57} to both the data sets (ERA5 and MERRA2). The same dependent and lagged predictors implemented in the RF algorithm are employed in the LASSO regression model for the different evaporation regimes (see Eq. 10 in Methods). The LASSO algorithm efficiently handles multicollinearity issues in lagged variables by shrinking some coefficients and setting others to zero\textsuperscript{[47,48]}. Once the regression coefficients are evaluated for a given regime, changes in FD intensity are calculated relative to changes in the standardized anomalies of the climate variables (see Methods). These changes are represented by linear plots, as shown in Fig 4 and Fig S8.
Figure 4 Sensitivity of flash drought intensity to climate variables (a-d) line plots showing the sensitivity of $F_{D_{ESR}}$ and (e-h) $F_{D_{RZSM}}$ intensity to changes in standardized anomalies of climate variables based on ERA5 dataset and calculated for 0, 1, and 2 pentad lags (denoted as L0, L1, and L2) for the different evaporation regimes.
FD_{ESR} and FD_{RZSM} intensity in most regimes show a higher sensitivity to changes in Pr and Tmax, which supports the relatively greater dominance of these two climatic forcings suggested in the previous analyses (Fig 3). More interestingly, a contrasting response of FD_{ESR} and FD_{RZSM} intensity to changes in the Pr(L0), Pr(L1), and Pr(L2) is notable between the arid and humid regions in both ERA5 (Fig 4(a, d)) and MERRA2 (Fig S8(a, d)) datasets. With the decrease in precipitation, the FD_{ESR} and FD_{RZSM} events are found to speed up more (i.e., FD intensity increases) in the humid regions. In contrast, the FD intensity is found to decrease in the arid regions. A similar response of FD_{ESR} and FD_{RZSM} intensity to changes in Tmax(L1), Tmax(L0), and pi(L0) is observed in both datasets. The sensitivity of FD_{RZSM} intensity to changes in pi(L0), Tmax(L1), and Tmax(L0) are found to increase continuously with the transition from arid to humid regions, as indicated by the positive increase in the slope of the lines in Fig 4 and Fig S8. On the other hand, FD_{ESR} intensity response to pi(L0) changes is diametrically opposite for the arid and humid regimes. Such contrasting climate controls are strongly linked to the difference in the variation of ecosystem functions, such as water-use-efficiency, that are primarily governed by distinct surface energy partitioning in the arid (water-limited) and humid (energy-limited) ecosystems [67,68].

4. Discussion and Conclusion

Our analyses provide evidence that climate variables, such as precipitation, temperature, vapor pressure deficit, and soil-moisture temperature coupling have varying and contrasting control over flash drought onset, evolution, and intensity. The varying
nature of climate control appears to be plausibly connected to the surface energy partitioning that defines the background state of climate aridity[37], which is found to be consistent across different data sources.

We found asynchronous changes in RZSM and evaporative stress during the FD durations, primarily driven by the difference in initial RZSM conditions and more extended memory of soil moisture in wetter regimes[56,60,61]. While the effect of climate variables on rapid intensification of evaporative stress occurs simultaneously, a cascading (time-delayed) climatic impact on the RZSM depletion is observed during the evolution of FDs.

Using a quantitative assessment framework, we conclude that precipitation and temperature anomalies have dominant and contrasting control over FD intensity across the arid and humid regimes. This assessment further reinforces the trade-off between water availability and energy supply as a limiting factor for evaporation across different ecosystems. Nevertheless, the role of (precipitation) temperature is more dominant in regulating (increase in evaporative stress) RZSM depletion rates within the FD duration. Land-atmospheric feedback strength is found to have relatively more influence on FD intensity in the sub-humid regions.

Overall, the results will strengthen our perspective on flash droughts by improving our understanding of the underlying physical processes and essential predictors across different ecosystems. The findings can be further extended to explore the potential influence of coupled changes in magnitude and seasonality of climate variables[69] and develop suitable forecasting tools to provide early warnings [12] across
the flash drought hotspots. By comparing two different flash drought definitions and using multiple datasets, the results from the study are expected to provide a broad and robust understanding of flash drought mechanisms and drivers globally. This is particularly important to highlight the advantages and limitations of the available flash drought definitions used by researchers and stakeholders [4,5,11,18,70]. The new information gained in this study can be further extended to investigate the causal linkages of soil moisture memory length, vegetation fluxes, and water use efficiency with FD characteristics across different ecosystems.

5. References


323


54. Vicente-Serrano SM, McVicar TR, Miralles DG, Yang Y, Tomas-Burguera M. Unraveling the influence of atmospheric evaporative demand on drought and its response to climate change. WIREs Climate Change [Internet]. 2020 [cited 2020


67. Roderick ML, Farquhar GD. A simple framework for relating variations in runoff to variations in climatic conditions and catchment properties. Water Resources Research. Wiley Online Library; 2011;47.


1. Introduction

Compound dry and hot events have received much attention in recent years due to their increasing impacts on agriculture, ecosystem, health, and energy [1–8]. For example, the 2012 summer in the central U.S., one of the unprecedented dry and hot year, caused huge economic losses of about $30 billion [9]. The compound dry and hot extremes that affected Europe and Russia during the summer of 2003 and 2010 are among the most hazardous compound events[10,11] that led to massive socio-economic impacts, including around 40,000 deaths[12], 25% loss of annual crop-yield[13], and extensive forest fires[4,14,15]. In order to reduce the associated potential impacts, it is important to understand the dependencies among dry-hot events using network and dynamic system tools, that can aid in accurate prediction and early-warning systems.

Dry-hot event is generally characterized by precipitation deficit that translate into soil-moisture depletion and positive temperature anomaly across a wide range of spatial and temporal scales. Like any other interconnected hazards[16], the physical processes that cause dry-hot events are interrelated[2], and therefore, have a cascading influence on each other. Although, some retrospective[17] and empirical approaches[2,18,19] have been applied to understand how these events will propagate as cascades across the ecosystem, the extent, scope, and temporal scale of their causal interactions are not yet
well understood. Furthermore, so far most studies do not account for the effect of other dependent variables while measuring the associations between inter-connected events, which may lead to spurious relationships and endogeneity[20].

We use this unique opportunity to assess the casual interaction between dry-hot events across the globe using a probabilistic framework motivated by a system dynamics approach[21–25]. We design two independent cross-scale (temporal) interaction networks[21,24] of compound and cascading dry-hot events based on daily root-zone-soil-moisture (RZSM) and maximum 2m air temperature (Tmax) anomalies for each location across the globe. The daily dry and hot events are considered as nodes of the network. These nodes are cross-linked at intervals of time lags ranging from zero to a week to investigate the non-linear dynamic causal effect of drying on heating and vice-versa. The compound and cascading effects are then quantified for each temporal network as (marginal) casual measures of association between exposure (dry or hot event) and outcome (dry or hot event) conditioned on a set of confounders (meteorological anomalies). This methodology is implemented by applying a standardized logistic regression approach[26] centered on the concept of counterfactual probabilities (see Methods), one of the cornerstones in the modern theory of causal inference[22,23,26,27]. The main advantage of this methodology lies in its ability to robustly estimate the marginal measures of associations, by isolating the main effect of exposure on the outcome variable, accounting for all other dependent variables as confounders[26]. Figure 1 illustrates the directed acyclic graphs representing the dynamic
system, the adopted cross-scale temporal interaction network for the dry-hot events, and the selected exposure, outcome, and confounder variables.

In this study, using daily RZSM data from GLEAM, and daily meteorological variables derived from the European Centre for Medium-Range Weather Forecasts Reanalysis 5 (ERA5), we aim to answer the following questions.
(i) What are the hotspots of compound and cascading dry-hot and hot-dry events across the globe?
(ii) How does the meteorological variables (or confounders) influence compound and cascading effects of drying on heating, and heating on drying in the hot-dry and dry-hot event network, respectively?
(iii) What is the role of background aridity in influencing these compound and cascading effects in the dry-hot and hot-dry event network?

The results from the study highlight the global hotspots where the compound and cascading effects of drying on heating and vice versa, is alarmingly higher, and characterize their scale of interaction which is likely to aid in risk reduction from crop-yield losses, wildfires, and water scarcity across the globe. The underlying mechanisms are also investigated through exploring the potential influence of meteorological anomalies and the role of surface energy partitioning on the causal linkages between terrestrial drying and heating. Our findings suggest more prominent causal effect of drying on heating as compared to that of heating on drying, which is relatively more amplified in the transitional regimes. Overall, the study provides meaningful priors for further research on the role of soil-moisture memory as a key contributor in determining
the nature and scale of interaction between dry and hot events across different ecosystems.

Figure 1 Compound and cascade model framework: (a) directed acyclic graph representing the dynamical system, and XY event network depicting the exposure and outcome variables (denoted by the binary sequency, $X_iY_{i+T}$) and the association with the...
confounder, \( Z \), (c) timeseries depicting association between global mean daily climatology of soil moisture (SM), maximum 2m air temperature (Tmax), and the confounder variables, total precipitation (Pr), total potential evaporation (PE), total actual evaporation (E), vapor pressure deficit (VPD), and surface net radiation (Rn) for 1980-2018 period.

2. Methodology

2.1. Data

In this study, global gridded daily root-zone soil moisture (RZSM), actual evaporation (E), and potential evaporation (PET) is obtained for the period, 1980-2018 from the third version of the Global Land and Evaporation Amsterdam Model (GLEAM v3.3a; [28]) available at https://www.gleam.eu/. Daily total precipitation (Pr), maximum 2m air temperature (Tmax), vapor pressure deficit (VPD), and surface net radiation (Rn) are derived from the European Centre for Medium-Range Weather Forecasts Reanalysis 5 (ERA5;https://cds.climate.copernicus.eu/ cdsapp#!/home). Vapor pressure deficit (VPD) is calculated using daily dew point temperature, daily mean 2m air-temperature, and daily surface pressure obtained from the ERA5 data archives. The GLEAM v3.3a combines various satellite-sensor products and ERA5 net radiation, and air temperature to provide relatively more accurate land surface estimates compared to other satellite- and model-based evaporation models [29,30]. While the GLEAM v3.3a dataset is available directly at daily timescale for every 0.25° x 0.25° pixels worldwide, ERA5 provides data at the same spatial resolution but for hourly timesteps.
2.2. Determining Compound and Cascading Dry and Hot Event Network

Dry events are identified using daily soil moisture by applying a threshold-based approach. A dry event is identified when the daily soil moisture falls below the daily climatological (1980-2018 period) threshold of 30th percentile (U.S. Dry Monitor (USDM)[31]). Soil moisture below the 30th percentile threshold is considered to have harmful effect on crop yield (USDM[31]). Hot events are defined as events during which the daily Tmax exceeds its daily climatological 80th percentile threshold for the 1980-2018 period[32,33]. Both dry and hot events were identified for the 1980-2018 period and the respective daily climatological thresholds were calculated using the whole 41-year time series.

Compound events are referred to as multiple events occurring simultaneously at the same time and location[1,2,34–36]. In this study, compound dry and hot events are defined by the co-occurrence of dry and hot event days at the same location. Cascading events are generally referred to the sequential occurrence of events in a dynamical system with 1-day[2] or multiple time intervals between the occurrences[18,19,37,38]. Here, a dry-hot (hot-dry) cascade is defined as the sequential occurrence of a dry (hot) day followed by a hot (dry) day at pre-defined temporal intervals of 1 to 7 days. These time-intervals are selected for a window of up to 7 days because of the potential increases in difficulty to cope with the socio-ecological impacts of such events as the time window shrinks to a sub-weekly scale[19]. Finally, compound and cascading dry and hot event network are constructed based on the cross-scale interaction[21,24] between
simultaneous and sequential (or lagged) dry and hot event days for the 1980-2018 period, as illustrated in Fig 1(a-b).

2.3. Model-based Estimation of Compound and Cascading Effects

The focus of the study is to explore the compound and cascading effect (CCE) of drying on heating and vice-versa, embedded in the dry-hot and hot-dry event networks. For a given network (dry-hot or hot-dry), the CCE determined based on the causal interaction between drying and heating. In this study, the CCE are quantified based on a metric called attributable fraction (AF)[39]. The AF is a population-specific measure of the proportion of preventable outcomes, e.g., hot (dry) day occurrences, had all days in the time-period been unexposed to dry (hot) event. It is a robust technique popularly used in modern epidemiology and public health and can be implemented to measure exposure-outcome relationship by taking into account necessary confounding measures [40,41].

In the following sections, we provide a discussion on the design of the dry-hot (hot-dry) compound and cascading model and how it is implemented within a logistic regression framework to measure the causal interactions (or AF) between drying and heating.

(i) Model Framework

In a dynamical system, causal interactions can take place through direct or indirect propagation of information within a network that consists of the exposure, outcome and confounders[22]. In most dynamical systems, the confounders have causal association with both the outcome and the exposure variable. If confounding (or independent) effects are not accounted for, it may lead to spurious relationships and
endogeneity[20]. Consequently, to obtain a robust measure of the effect of exposure variable on the outcome, it is important to isolate the main effect of the exposure variable on the outcome variable by accounting for all other variables as confounders[26].

The pathway of information propagation in a dynamical system can be demonstrated by using directed acyclic graphs (DAGs). DAGs present a graphical representation of the problem of confounding[22]. Let Z denote a set of confounders that control both the outcome and the exposure variable. DAG can describe the confounding by Z of the causal relationship between the exposure variable, X and outcome variable, Y, as shown in Fig 1(a). In case of both dry-hot and hot-dry event network, five confounding variables (z ⊆ Z) are used, such as VPD, Pr, daily surface net radiation, actual evaporation, and potential evaporation, that are known to have significant control on both dry (X) and hot (Y) events[42]. Same set of confounders (Z) are used in case of analyzing the causal interaction in a hot (X) and dry (Y) cascade.

(ii) Estimation of AF based on Logistics Regression

We first identify the dry and hot day occurrences based on the definition of compound and cascading event discussed above. The occurrences and non-occurrences of dry, and hot days are subsequently transformed into binary (0/1) time-series, such that, an occurrence is denoted by 1 and non-occurrence denoted by 0. AF was then calculated by fitting a logistic regression model to the binary exposure, X and the binary outcome, Y with model adjustments specifically made to include the confounders (Z) as covariates. The logit regression is implemented for the selected time-intervals (T = 0, 1, 2, 3, 4, 5, 6, and 7) separately, such that T = 0 refers to the compound and T > 0 refers to the cascade
model framework. Hereafter in this study, this framework is referred to as \( X_t Y_{t+T} \), denoting the casual effect of \( X \) at time-step, \( t \) on \( Y \) at time-step, \( t+T \), being measured for \( X \) and \( Z \) lagged by \( T \) days. This framework is illustrated in Figure 1b.

For binary outcomes and exposure, AF can be defined as in Eq (1)[39],

\[
AF = 1 - \frac{P(Y_0 = 1)}{P(Y = 1)},
\]

where the \( P(Y_0 = 1) \) is the counterfactual probability of outcome if the exposure \( X \) is eliminated (i.e., \( X \) set to 0), and \( P(Y = 1) \) is the factual probability of outcome in the population. The AF thus measures the proportion of outcome events (e.g., hot days) that would be prevented if the exposure events (e.g., dry days) were eliminated from the population. Unlike the linear, and log-linear models, logistic regression is a standard choice for estimating AF due to its ability to yield probabilities between 0 and 1. The logistic regression model is defined as,

\[
\text{logit}\{Pr(Y = 1|X,Z)\} = g(X,Z; \beta)
\]

Here \( g(.) \) is an additive function of the variables \( X \) and \( Z \) and could be specified as \( \beta_0 + \beta_1 X + \beta_2 Z \), where \( \beta \) is the parameter vector of the logit model. A regression standardization[27] is then implemented to the fitted model, \( X_t Y_{t+T} \) to estimate marginal measures of association (see A.1. in Supplementary Information). This method uses the logistic regression model to predict the risk ratios of the outcome (\( Y \)), for unexposed (\( X=0 \)) at every pre-defined level of the measured confounders. Finally, these predictions are averaged over the sampling distribution of \( Z \) to produce a standardized risk, for unexposed. Thus, if \( Z \) is sufficient for confounding control, then,
\[ P(Y_0 = 1) = E[P(Y = 1|X = 0,Z)], \]

(3)

where \( P(Y|X,Z) \) is the conditional distribution of \( Y \), given \( X \) and \( Z \).

Once, \( P(Y_0 = 1) \) is estimated following equation 3, \( P(Y = 1) \) is calculated directly from the observational data and implemented in equation 1 to calculate AF. It is important to note that AF is estimated for each model framework \((XY_T)\) separately and denoted by \( AF_T \).

3. Results

3.1. Hotspots of Compound and Cascading Dry-Hot and Hot-Dry Events

The compound and cascading effect (CCE) of dry (hot) on hot (dry) event occurrence is quantified based on the magnitude of attributable fraction (AF)[39]. A data driven compound and cascade model framework is implemented using daily dry-hot and hot-dry event networks for the 1980-2018 period considering multiple time-lags, \( T = 0, 1, 2, 3, 4, 5, 6, \) and \( 7 \) and confounders. AF is calculated for the global grid cells based on the dry-hot (and hot-dry) event networks by applying a regression standardization technique[27] for each of these time-lags, \( T \) (hereafter referred to as \( AF_T \)), separately. The regression standardization is helpful for obtaining robust marginal measures of association between the exposure and outcome variable by accounting for all other variables as confounders. Thus, \( AF_T \) (%) for a dry-hot (hot-dry) event network represents the standardized risk of having a hot (dry) event caused by a dry (hot) event that occurred \( T \) days ago. In other words, \( AF_T \) determines the strength of the CCE of drying on heating
and vice versa. The cascade model framework, implementation, and association with selected confounders (meteorological anomalies) are presented in Fig 1.

Figure 2 Hotspots of Compound and Cascading Effect in Dry-Hot and Hot-Dry Event Network shown by spatial map of (a) maximum AF, (b) corresponding lags for the dry-hot cascade network, and (c-d) same as in (a-b) but for the hot-dry cascade network calculated for the period, 1980-2018. Note that only regions with a positive AF are shown in these spatial maps.

The $AFT$ magnitudes are classified into four different categories, from moderate (0-5%) to exceptional (>15%), as illustrated in Table S1. As such, $AFT$ in the exceptional category indicates that more than 15% of the hot events are caused by dry events that occurred at a lag of $T$ days during the 1980-2018 period. The time lags suggests that the
causal effects in a network can vary from being compounding (simultaneous; lag = 0 days) to cascading (lags > 0 days) in nature. To illustrate the hotspots of CCE of drying on heating (or heating on drying), we calculated the maximum $AF_T$ for the dry-hot (hot-dry) event network across each (0.5°x0.5°) pixel of the globe. We define a given pixel as a hotspot if the corresponding value of maximum AF falls within the extreme (10-15%) to exceptional category range. Figure 2a, and 2c demonstrate the spatial distribution of the maximum $AF_T$ values and the hotspots of CCE for the dry-hot, and hot-dry event network over the globe. The corresponding number of lags for which the AF is found to be maximum is noted for each and every pixel, which is presented in Fig 2(b, and d). The estimates, lower and upper bounds (at 95% confidence level) of AF (%) for each of the lag timings ($T = 0, 1, 2, 3, 4, 5, 6$, and 7 days) are presented in Fig S1-S6.

Fig 2a suggests that extreme to exceptionally (AF > 10%) strong CCE of dry event on hot event occurrence is prominent in major part of North and South American continent, central and southern Africa, southern Europe, southern Asia, and northern and eastern Australia. Fig 2b show that the time lag, for which the maximum CCE of dry event on hot event occurrences are observed, vary substantially across the globe. In North America, the magnitude of AF is found to be particularly higher across the southern and central US over the Great Plains and north-central parts of western US. The AF over majority of the western US is greatest for 0-day lag indicating that a compounding effect of dry day on hot day occurrence is more dominant in the region. On the contrary, the AF magnitude over the Great Plains show an increase for greater time lags (4 to 7 days) indicating the dominance of cascading (sequential) effect of dry day on hot day
occurrence in this region. Similar cascading (compounding) effect of dry day on hot day occurrence is dominant in the hotspot regions located in South America, Central Africa, peninsular India, and northern and eastern Australia (southern Europe, and central Asia).

The hotspots of CCE for the hot-dry event network is however limited to very few regions, such as in northeastern part of South America, some parts of western US, southern Europe, central Africa, and northern parts of India and middle east (Fig 2c). Most of the globe exhibit moderate to severe (0-10%) CCE of hot event on the occurrence of dry event. The hotspot in northeastern South America, central Africa, and northern India indicate that hot events have a cascading effect (lag = 1 to 7 days) on the occurrence of dry days in the region (Fig 2d). On the other hand, both extreme-exceptional compound (lag = 0 days) and cascading (lag = 1 day) effects are notable in the hotspot regions of southern Europe, western US, and middle east.

Nevertheless, the higher effect of dry events on hot events may be associated with the longer SM memory[43,44]. The longer memory of SM is also one of the primary reasons for SM depletion (dry event) being more resilient to changes in hydroclimatic anomalies (confounders; see Methods), leading to weaker CCE of hot events on dry events[45–47].

3.3. Influence of Confounders

The effect of confounders on the CCE of dry-hot and hot-dry event network is investigated based on odd ratios. Odd ratio is calculated by fitting a logistic regression between outcome and exposure with the confounders as covariates. The logit model was
embedded within the compound and cascade model framework (see Methods). Odd ratio is given as \( \exp(\beta) \), where \( \beta \) is the regression coefficient of the logit model, such that \( \exp(\beta) > 1 \) (\( \exp(\beta) < 1 \)) indicates a multiplicative increase (decrease) in the daily odds of an outcome for a given exposure and per unit increase in the confounders (here, standardized anomalies of Pr, ESR, VPD, and Rn), measured for specific time lags. For example, an odd ratio of 1.2 for lag, \( T = 0 \) days in a dry-hot event network implies that for per unit increase in the confounder variable, the odds of occurrence of a hot event are likely to increase by 1.2 times given a dry event has occurred on the same day. Such associations suggest a simultaneous or compounding effect of confounders. Similarly, the odd ratios for lag, \( T > 0 \) days is an indicator of a cascading (lagged) influence of the confounders on the occurrence of a hot day given a dry day is witnessed \( T \) days before. Fig S9, S10, S11 and 3a-h (Fig 3i-p, S12, S13, and S14) illustrate the spatial distribution of statistically significant (at 95% confidence level) odd ratio corresponding to ESR, Rn, Pr, and VPD for the dry-hot (hot-dry) cascade, respectively.

Among the selected confounders, VPD show the greatest positive effect on the odds of occurrence of a hot event conditioned on a dry event, which is indicated by the statistically significant odd ratios greater than 1 consistently across the whole globe (Fig 3(a-h)). The effect of VPD is strongest for a lag of 0 days, and its influence weakens for increasing number of lags. These results suggest that simultaneous increases in VPD is more likely to increase the odds of compounding effect of dry event on hot event occurrences as compared to triggering a sequential dry and hot event cascade. On contrary, a considerable amount of spatial heterogeneity is noted for the odds ratio
Figure 3 Influence of Confounders (a-h) Spatial distribution of statistically significant (at 5% significance level) odd ratios ($\exp(\beta)$) corresponding to standardized anomalies of VPD calculated by fitting the logistic regression model for the dry-hot event network for a lag of (a) 0 days, (b) 1 day, (c) 2 days, (d) 3 days, (e) 4 days, (f) 5 days, (g) 6 days, and
(h) 7 days, and (i-p) same as in (a-h) but corresponding to standardized anomalies of ESR for the hot-dry event network.

corresponding to ESR, Pr, and Rn. For example, simultaneous increases in evaporative stress (indicated by decrease in ESR) have the strongest positive influence on the odds of compounding effect of dry event on hot event (T = 0 days) in major part of South America, southern Africa, and eastern Europe (Fig S9). On the other hand, for southeast Asia, Australia, and southern North America, evaporative stress anomalies show the strongest positive (lagged) association with the odds of witnessing a hot event given a dry event has occurred 7 days, 3-4 days, and 4 days before, respectively (Fig S9). A simultaneous increase in Rn and the odds of occurrence of a hot event in the same day as a dry event can be noted in Northern Russia, Japan, northwest of China, and central Africa (Fig S10). Pr anomalies show a strong lagged (T = 1 day) association in central and southeast Asia including majority of China, central Europe, northwest of North America, some part of central Africa, and South America (Fig S11).

Evaporative stress has the strongest positive influence on the odds of CCE of a hot event on a dry event which is consistent across the globe. This is indicated by the spatial map of odd ratios less than 1 in Fig 3(i-p), that suggest a decrease in the odds of CCE of hot event on dry event occurrence for per unit increase in ESR (or increase in evaporative stress). The compounding effect (lag, T = 0) of ESR is found to be more dominant as compared to its cascading effect (for lag, T > 0). For example, the compounding effect of ESR is strongest in central US, and Australia, and its effect can be
noted to decrease for increasing number of lags across the whole globe. On the other hand, the influence of Rn, Pr, and VPD on the CCE of hot event on dry event is spatially sparse with very little changes across the different lags (T = 0 to 7 days) as shown in Fig S12-S14. These results suggest that the influence of the daily changes in some climate forcings have a spatially disproportionate influence on the dry-hot and hot-dry event cascade across the terrestrial surface. Such spatial heterogeneity may arise from variation in surface energy partitioning which is mainly controlled by background aridity of a region[1].

3.4. Role of Background Aridity

The background aridity of a region play a critical role in controlling the water use efficiency[48], sensitivity of evaporation to changes in temperature and precipitation[1,49], and causal interactions between precipitation, evaporation and soil-moisture[50]. We investigate the control of background aridity on the CCE in dry-hot and hot-dry event network separately to understand the implication of surface energy partitioning for each of these cascades.

Background aridity is quantified based on aridity index (AI)[51] which is calculated as a ratio between annual climatological mean precipitation and annual climatological mean potential evaporation for the 1980-2018 period. The global regions are then subdivided into hyper-arid, arid, semi-arid, sub-humid and humid regimes following an AI based classification system proposed by United Nations Environment Program[51]. The five climate regimes and the corresponding range of AI is illustrated in
We considered the AI range (0.05-1) starting from the arid to humid regimes and divided the globe into 96 sub-regimes at intervals of 0.01. The corresponding pixels within the sub-regions were extracted, and the magnitude of AF (%) was averaged over those pixels for different time lags to investigate the role of background aridity on the CCEs in the dry and hot event network.

Figure 4 Role of Background Aridity

(a) Spatial map showing the classification of global evaporation regimes based on aridity index (AI) calculated as a ratio of mean annual precipitation and potential evaporation for 1980-2018 period, (b) Mean AF (%) binned as
a function of AI and time-lags (0 to 7 days) for the dry-hot event network, and (c) same as in (b) but for the hot-dry event network.

Figure 4b-c displays the variation of AF (represented by shading) from arid to humid regimes for different time lags (represented in the y-axis) corresponding to the dry-hot and hot-dry event network. A very distinct control of background aridity on CCE is noted for the dry-hot and hot-dry event network. The CCE corresponding to the dry-hot event network is found to be relatively stronger (mean AF > 8%) in semi-arid part of the globe (with AI ranging between 0.25 and 0.4) and relatively weaker effect (mean AF < 4%) is noted for the arid (for AI < 0.20) and humid regimes (for AI > 0.8). On contrary, the CCE corresponding to hot-dry network is found to be relatively stronger (mean AF > 4%) for sub-regions falling in the transition zone from sub-humid to humid regime (denoted by AI between 0.63 and 0.8). These results underscore non-linear control of surface energy portioning on causal interaction between drying and heating of the terrestrial surface, which is notably stronger in the transitional (semi-arid, and sub-humid) regimes, generally characterized by stronger land-atmosphere coupling [52–54].

4. Discussion and Conclusion

Compound and cascading dry and hot events have significant impact on ecosystem functioning[1,2,34,37]. Understanding the causal interactions between these events all over the globe, therefore becomes extremely important and requires new methods to account for the measures and scale of such interactions. Here using a
standardized logistic regression approach[26], for the first time, we defined a cross-scale interaction-based compound and cascade model framework for measuring the casual effects in global drying and heating in dry-hot and hot-dry event networks. Two distinct event networks, dry-hot and hot-dry, are constructed using daily root-zone soil-moisture and maximum air temperature anomalies by embedding time lags ranging from 0 to 7 days. The compound and cascading effects of drying and heating are subsequently analyzed in these event networks based on attributable fraction[26], that measures the causal effects in exposure-outcome (dry-hot or hot-dry) relationship conditioned on multiple confounders (meteorological anomalies).

The results from the study reveal crucial aspects of the causal interactions between dry and hot events, including their global hotspots, meteorological drivers, and effect of background aridity. A relatively stronger compounding and cascading (causal) effect of drying on heating is noted (as compared to that of heating on drying) across the globe which may be linked to the persistency of soil moisture memory[43,50,55]. Extreme to exceptionally strong CCE (indicated by AF > 10%) of drying on heating was found in major part of western and Great Plains of North America, and tropical South America, central and southern Africa, southern Europe, southern Asia, and northern and eastern Australia. The time lags of such associations exhibited substantial spatial heterogeneity suggesting the varying nature of interactions between drying and heating from one place to another, from being compounding (simultaneous; lag = 0 days) to cascading (lags > 0 days). Vapor pressure deficit is found to have the strongest positive effect on the CCE of drying on heating for different time lags uniformly across the globe
with a more dominant influence on the compounding effect relative to cascading effect. The influence of precipitation, evaporative stress and surface net radiation is found to be spatially sparse. On the other hand, evaporative stress exhibits the most dominant positive influence on the CCE of heating on drying across the globe, whereas a considerable spatial heterogeneity in the influence of precipitation, vapor pressure deficit and surface net radiation is noted for different time lags. Background aridity seems to have a non-linear association with both CCE of drying on heating and heating on drying, with a relatively stronger interaction between dry and hot events noted for the transitional regimes.

Soil moisture drying has significant implication on heating [56] that propagates as cascades across the physical and human systems affecting agriculture and human-health[8,57]. Our study has a broader implication in bridging the gap between disaster risk reduction and climate change adaptation[16,35–37] with potential to provide a more nuanced framework for assessing interconnected and cascading risks. The results from the study can be usefully transformed to assess the changes in risk of exposure to interconnected hazards[58–60] and forecasting skill[50,61]. Our findings can be further channelized to provide more in-depth understanding on the association of dry and hot cascades with length of soil-moisture memory[43,50,55], anticyclonic circulations and blocking[62,63], land-atmosphere coupling[52–54], regional moisture transport[34,64], vegetation fluxes[65], water use efficiency[48,66,67], compound changes in climate variability[68], and large-scale teleconnections[34].
5. References


112. Available from: https://journals.ametsoc.org/view/journals/hydr/7/5/jhm533_1.xml


364
CHAPTER X
CONCLUSION AND RECOMMENDATION

There is widespread evidence of intensification in hydroclimatic extremes due to acceleration of hydrological cycle caused by natural and anthropogenic changes in climate. In future, climate change impacts will emerge in various ways at the same time or propagate as cascades across the physical-human systems. Furthermore, non-stationarity in climate will most likely have a non-linear and complex effect on the statistical behavior or distribution of hydroclimatic extremes. Therefore, more nuanced techniques are required to quantify these simultaneous or cascading impacts occurring concurrently and in different combinations of extremes. This thesis specifically focused on the detection, investigation of underlying mechanisms, and impact assessment of four different compound and cascading hydroclimatic extremes, (i) compound drought and heatwaves, (ii) combined heat extremes and humidity, (iii) meteorological forcing associated with extreme precipitation events caused by atmospheric rivers, and (iv) flash droughts. These hydroclimatic extremes are analyzed across different spatiotemporal scale using concept of system dynamics and causality, robust statistical methods, and multiple datasets. The specific research objectives that this thesis focused on are:

(a) Highlighting the major challenges associated with the application of different drought indices in climate change studies.

(b) Robust detection of compound drought and heatwaves, and quantifying their frequency, duration, and severity on a global scale.
(c) Investigate and quantify the association of compound drought and heatwave occurrences with large scale climate variability, moisture convergence/divergence, and land surface energy fluxes across different parts of the globe.

(d) Quantifying the influence of anthropogenic warming on changes in compound drought and heatwave characteristics across the globe and their association with the background aridity of the global regions.

(e) Unraveling and quantifying the relative contribution of natural climate variability and anthropogenic warming on the occurrence of compound drought and heatwaves.

(f) Evaluating the possible changes on combined and potential impact of heat extremes and humidity on human population under the non-stationary climate scenarios in future.

(g) Cascading influence of meteorological forcing on the progression of moisture fluxes leading to extreme precipitation events associated with atmospheric rivers.

(h) Investigating the climate controls on global flash droughts across the different evaporation regimes.

(i) Proposing a cascade model framework to quantify the compound and cascading effects in a dry and hot event network.

Based on the above objectives, the following conclusion are drawn from the study.

(a) Accurate formulation of robust drought indices is important to investigate drought characteristics under the warming climate. Because different drought indices
show different degrees of sensitivity to the same level of continental warming, robustness of drought indices against change in temperature and other variables should be prioritized. A formulation of drought indices without considering the factors that govern the background state may lead to drought artifacts under a warming climate. Consideration of downscaling techniques, availability of climate data, estimation of potential evapotranspiration (PET), baseline period, non-stationary climate information, and anthropogenic forcing can be additional challenges for a reliable drought assessment under climate change. As one formulation of PET based on temperatures can lead to overestimation of future drying, estimation of PET based on the energy budget framework can be a better approach compared to only temperature-based equations.

(b) El Niño Southern Oscillation (ENSO) exhibits robust association with compound drought and heatwaves over the Southern Hemisphere during the austral summer and fall, while Pacific Decadal Oscillation (PDO) influences their occurrences over the Western North America in the Northern Hemisphere during the boreal summer, which is supported by the composites of anomalies in the atmospheric circulations and surface energy budget. North Atlantic Oscillation (NAO) association with compound drought and heatwave events is relatively weak. The compound drought and heatwave occurrence over other regions is driven by a combination of these large-scale natural forcing. The analyses also highlight that the cooccurrence of weekly to sub-monthly scale anomalies in the observed temperature and precipitation may not be always aligned between the observations and the reanalysis. Therefore, caution must be exercised while
explaining such observed anomalies on the basis of reanalysis-based circulations and surface energy budget.

(c) Significant increases in drought-related heatwaves and affected global land area is observed in recent (warmer) periods, 2000-2018. Several regions across the globe witnessed rise in compound drought and heatwave frequency (one to three events/year), duration (2–10 days/year), and severity. This increasing pattern is spatially asymmetric, and greater amplification is observed across the Northern hemisphere due to recent warming. Furthermore, the background aridity influences the spatiotemporal evolution of compound drought and heatwave frequency events.

(d) An attribution study performed to quantify the odds of occurrence of these compound drought and heatwave days in a month over 10 climate regions revealed a significantly positive, and multiplicative effect on the odds from the anthropogenic global warming. Odd ratios estimated for these climate regions were found to be in the range of 1.7 to 3.5, and as high as 5 to 60 at 1.5°C, and 2°C warming levels, respectively. Among all the climate regions, East Asia is the most affected region due to the rise in anthropogenic warming.

(e) Potential impact of short-to-medium duration (1–7 days) heat stress events is likely to increase more than three-fold across densely populated regions of the U.S. including the Northeast, Southeast Piedmont, Midwest, and parts of the Desert Southwest by late this century (2060–2099) under the highest emissions scenario. The contribution from climate change alone more than doubles the impact in the coastal Pacific Northwest,
central California, and the Great Lakes region, implying a substantial increase in heat stress risk without aggressive mitigation efforts.

(f) The seasonal frequency of EP events associated with ARs suggests that more frequent AR-EP events occur during the colder months (November to April). In contrast, the AR-EP events are less frequent but more severe in the warmer months (May to October). A total of 12-15 AR-EP events, with severity exceeding the 99th percentile precipitation threshold, were observed during the 3-month overlapping seasons between November and April from 1979 to 2019 that affected Georgia, Florida, Alabama, and South Carolina. On the other hand, the average precipitation magnitude of the AR-EP events is relatively higher (55-90 mm/day) in the warmer months (May to October).

To explore the cascading nature of relevant meteorological forcing on the physical processes that favor such events, an event-centered composite analysis was performed based on the top 100 severe AR-EP events observed during the extended cold and warm season, separately. It was observed that during the progression of the AR-EP events, the anomalies associated with composite mean sea level pressure (MSLP) and 850mb geopotential height (Z850) make a transition from the trough to ridge formation along with a south-eastward extension of Bermuda High in the cold season. The spatiotemporal evolution of these meteorological variables is found to have a cascading effect on the mode of moisture transport indicated by integrated vapor transport (IVT) and moisture availability shown by total column water vapor associated with the major AR-EP events. The warm season IVT field gets stronger 2-days before the AR-EP event occurrences indicating a continuous increase in moisture influx into the Gulf and Atlantic
Coastal Plains. Similar strengthening of IVT is noted over the Gulf Coastal Plains regions during the cold season. A cascading effect is also noted for the moisture availability indicated by a significant increase in total column water vapor over the Gulf of Mexico 2 days before the events. Overall, the cold season AR-EPs are driven by relatively stronger dynamical systems indicated by greater IVT intensity. In contrast, the warm season AR-EPs are associated with a weaker IVT field, higher atmospheric instability, and more moist conditions.

(g) Varying as well as contrasting spatial drivers of FD frequency, intensity, and their evolution were revealed. Changes in precipitation, temperature, vapor pressure deficit and soil-temperature coupling play an important role with a cascading (concurrent) impact on the evolution of FDs. Precipitation and temperature (soil-temperature coupling) have the most dominant and contrasting control over FD intensity across the arid and humid (sub-humid) regions, underscoring the importance of water and energy supply as limiting factors in regulating FD intensification rates.

(h) Relatively stronger compound and cascading effect (CCE) of drying on heating than that of heating on drying with extreme to exceptionally strong (AF>10%) CCE were noted in major part of North and South American continent, central and southern Africa, southern Europe, southern Asia, and northern and eastern Australia. The impact of meteorological variables on the CCEs is quantified based on odd ratios that suggest, vapor pressure deficit (evaporative stress) has the strongest influence on the CCE of drying on heating (heating on drying) across the globe. The role of background
aridity is also investigated which suggest relatively more amplified CCEs in the transitional evaporation regimes.

**Recommendation for future study:** Modern research has started focusing on the usage of real-time as well as big data for impact forecasting of extreme hydroclimatic events. Machine learning tools, and the concept of artificial intelligence can also be implemented using the prior knowledge of the physical system to evaluate and investigate the crucial aspect of interconnected events and their impacts. More direct approaches linking the impact of extreme events with the socio-economic indicators can also provide a better understanding of their compounding and cascading impacts across the physical and human systems. More robust techniques complementing the non-linear and bidirectional causations among the inter-connected extremes may prove to be valuable resources to understand these real-world problems and possible impacts under a system dynamics framework. More emphasis can be given on policy driven research that can directly connect the quantification, assessment, and research outputs to the stakeholder’s needs. Research projects that puts the scientists, stakeholders, policy makers, and communicators under a single banner is likely to enhance and aid in better mitigation efforts for reducing the compound and cascading impacts of hydroclimatic extremes under the changing climate scenario.
APPENDICES

Appendix A

Supplementary Information for Chapter 3

Figure S1. 26 climate regions used in the study as defined in the AR5-SREX (IPCC SREX 2012).
Figure S2 Spatial distribution of yearly frequency of extreme drought events based on (left panel) GPCC and CPC dataset, and (right panel) ERA5 dataset during the (a, and b) DJF (c, and d) MAM, (e, and f) JJA, and (g, and h) SON.
Figure S3 Spatial distribution of yearly frequency of heatwave events based on (left panel) GPCC and CPC dataset, and (right panel) ERA5 dataset during the (a, and b) DJF (c, and d) MAM, (e, and f) JJA, and (g, and h) SON.
Figure S4 Spatial distribution of yearly frequency of CDHW events based on (left panel) GPCC and CPC dataset, and (right panel) ERA5 dataset during the (a, and b) DJF (c, and d) MAM, (e, and f) JJA, and (g, and h) SON.
Figure S5 Spatial distribution of correlation between ENSO (Nino3.4) and (left panel) CPC based seasonal mean of average temperature; and (right panel) GPCC based seasonal mean of total monthly precipitation during (a, and b) DJF, (c, and d) MAM, (e, and f) JJA, and (g, and h) SON. Grid cell where the correlations are found to be statistically significant (at 5% significance level) are marked by striplings.
Figure S6 Spatial distribution of correlation between PDO and (left panel) CPC based seasonal mean of average temperature; and (right panel) GPCC based seasonal mean of total monthly precipitation during (a, and b) DJF, (c, and d) MAM, (e, and f) JJA, and (g, and h) SON. Grid cell where the correlations are found to be statistically significant (at 5% significance level) are marked by striplings.
Figure S7 Spatial distribution of correlation between NAO and (left panel) CPC based seasonal mean of monthly average temperature; and (right panel) GPCC based seasonal mean of total monthly precipitation during (a, and b) DJF, (c, and d) MAM, (e, and f) JJA, and (g, and h) SON. Grid cell where the correlations are found to be statistically significant (at 5% significance level) are marked by striplings.
Figure S8 Bar plot showing area (or percentage of grid cells) within each of the 26 climate regions showing statistically significant (at 5% significance level) partial correlation estimates (as shown in Figure S5 to Figure S7) between the large scale climate oscillations (ENSO, PDO, and NAO) and the CPC based seasonal mean of monthly average temperature (red), and the GPCC based seasonal mean of total monthly precipitation (blue) during DJF.
Figure S9 Bar plot showing area (or percentage of grid cells) within each of the 26 climate regions showing statistically significant (at 5% significance level) partial correlation estimates (as shown in Figure S5 to Figure S7) between the large scale climate oscillations (ENSO, PDO, and NAO) and the CPC based seasonal mean of monthly average temperature (red), and the GPCC based seasonal mean of total monthly precipitation (blue) during MAM.
Figure S10 Bar plot showing area (or percentage of grid cells) within each of the 26 climate regions showing statistically significant (at 5% significance level) partial correlation estimates (as shown in Figure S5 to Figure S7) between the large scale climate oscillations (ENSO, PDO, and NAO) and the CPC based seasonal mean of monthly average temperature (red), and the GPCC based seasonal mean of total monthly precipitation (blue) during JJA.
Figure S11 Bar plot showing area (or percentage of grid cells) within each of the 26 climate regions showing statistically significant (at 5% significance level) partial correlation estimates (as shown in Figure S5 to Figure S7) between the large scale climate oscillations (ENSO, PDO, and NAO) and the CPC based seasonal mean of monthly average temperature (red), and the GPCC based seasonal mean of total monthly precipitation (blue) during SON.
Figure S12 Spatial distribution of correlation between ENSO (Nino3.4) and (left panel) seasonal mean of average temperature; and (right panel) seasonal mean of total monthly precipitation from ERA5 during (a, and b) DJF, (c, and d) MAM, (e, and f) JJA, and (g, and h) SON. Grid cell where the correlations are found to be statistically significant (at 5% significance level) are marked by striplings.
Figure S13 Spatial distribution of correlation between PDO and (left panel) seasonal mean of average temperature; and (right panel) seasonal mean of total monthly precipitation from ERA5 during (a, and b) DJF, (c, and d) MAM, (e, and f) JJA, and (g, and h) SON. Grid cell where the correlations are found to be statistically significant (at 5% significance level) are marked by striplings.
Figure S14 Spatial distribution of correlation between NAO and (left panel) seasonal mean of average temperature; and (right panel) seasonal mean of total monthly precipitation from ERA5 during (a, and b) DJF, (c, and d) MAM, (e, and f) JJA, and (g, and h) SON. Grid cell where the correlations are found to be statistically significant (at 5% significance level) are marked by striplings.
Figure S15 Bar plot showing area (or percentage of grid cells) within each of the 26 climate regions showing statistically significant (at 5% significance level) partial correlation estimates (as shown in Figure S12 to Figure S14) between the large scale climate oscillations (ENSO, PDO, and NAO) and the seasonal mean of monthly average temperature (red), and the seasonal mean of total monthly precipitation (blue) from EAR5 during DJF.
Figure S16 Bar plot showing area (or percentage of grid cells) within each of the 26 climate regions showing statistically significant (at 5% significance level) partial correlation estimates (as shown in Figure S12 to Figure S14) between the large scale climate oscillations (ENSO, PDO, and NAO) and the seasonal mean of monthly average temperature (red), and the seasonal mean of total monthly precipitation (blue) from EAR5 during MAM.

Figure S17 Bar plot showing area (or percentage of grid cells) within each of the 26 climate regions showing statistically significant (at 5% significance level) partial correlation estimates (as shown in Figure S12 to Figure S14) between the large scale climate oscillations (ENSO, PDO, and NAO) and the seasonal mean of monthly average temperature (red), and the seasonal mean of total monthly precipitation (blue) from EAR5 during JJA.
Figure S18 Bar plot showing area (or percentage of grid cells) within each of the 26 climate regions showing statistically significant (at 5% significance level) partial correlation estimates (as shown in Figure S12 to Figure S14) between the large scale climate oscillations (ENSO, PDO, and NAO) and the seasonal mean of monthly average temperature (red), and the seasonal mean of total monthly precipitation (blue) from EAR5 during SON.
Fig S19 Seasonal climatology of divergent winds vectors (m/s) and velocity potential (color shading, unit: m²/s, scaled by 10⁶) contours at 200 mb.
Fig S20 (a) El Nino (b) La Nina based anomalies in divergent winds vectors (m/s) and velocity potential (color shading, unit: m²/s, scaled by 10⁶) contours at 200 mb with respect to the climatology during JJA.
Figure S21 PDO warm phase (left panel) and cold phase (right panel) based anomalies in divergent winds vectors (m/s) and velocity potential (color shading, unit: m$^2$/s, scaled by 10$^6$) contours at 200 mb with respect to the climatology during (a-b) DJF, (c-d) MAM, and (e-f) SON.
Figure S22 NAO warm phase (left panel) and cold phase (right panel) based anomalies in divergent winds vectors (m/s) and velocity potential (color shading, unit: m²/s, scaled by 10⁶) contours at 200 mb with respect to the climatology during (a-b) DJF, (c-d) MAM, and (e-f) SON.
Figure S23 Composites of surface sensible heat flux (left panel), and latent heat flux (right panel) for (a-b) DJF, (c-d) MAM, (e-f) JJA, and (g-h) SON. All units are in W/m$^2$. 
Figure S24 El Nino (left panel) La Nina (right panel) based anomalies in surface latent heat flux with respect to the climatology during (a-b) DJF, (c-d) MAM, (e-f) JJA, and (g-h) SON. The grid cells where the difference is statistically significant at 95% confidence level are marked with stippling. The sign convention implemented for the fluxes is positive upwards.
Figure S25 Same as in Figure S25 but based on the warm (left panel) and cold (right panel) phase of PDO.
Figure S26 Same as in Figure S25 but based on the warm (left panel) and cold (right panel) phase of NAO.
Figure S27 Surface sensible heat flux anomalies for the warm and cold phase of (a) ENSO, (c-f) PDO, and (g-j) NAO.
Appendix B

Supplementary Information for Chapter 4

B.1. Estimation of CDHW Severity

The severity of a CDHW (CDHWs) event is calculated based on the cumulative sum of the daily severity values obtained over the consecutive days of the event. The daily severity is estimated as the product of the daily standardized values of maximum

\[\text{SCPDSI}_w\]

and the \(\text{SCPDSI}_w\) value observed in the coinciding extreme drought week, \(w\)

Thus, the severity for an event, \(i\) is given as,

\[CDHW_{i} = \sum_{d=1}^{D_i} (-1 \times \text{SCPDSI}_{w,d}) \times \left( \frac{T_{\text{max},d} - T_{25p}}{T_{75p} - T_{25p}} \right) ; D_i \geq 3, d \in w\]

where \(D_i\) = duration of the \(i^{th}\) CDHW event; \(d\) indicates the heatwave day coinciding with drought week, \(w\); and \(T_{25p}\) and \(T_{75p}\) are the 25\(^{th}\) and 75\(^{th}\) percentile of the daily maximum temperature during the summer season, respectively. Finally, the \(\text{CDHW}_s\) for a given year is calculated as the average of the \(\text{CDHW}_{si}\) magnitudes observed within the year, which is estimated as,

\[\frac{\sum_{i=1}^{N} \text{CDHW}_{si}}{N} ; \text{if } N \neq 0\]

\[\text{CDHW}_s = 0 \quad \text{; if } N = 0\]

where \(N\) represents the total number of events observed in a given year.
B2. Selection of Recent Warmer Period

To investigate the role of warming, we compare the CDHW characteristics based on two equal periods between 1983 and 2016, past period (1983-1999), and a recent warmer period (2000-2016). The recent warmer period is determined based on the anomalies of summer mean daily maximum temperature (Tmax) estimated for each year between 1983 and 2016 for the global land areas, as shown in Fig S1. At first, the anomalies are evaluated for each grid location with respect to the long-term, 1983-2016 average, and finally, spatially averaged across the global land surface. Additionally, to investigate if the results are independent of the data sources, the anomalies are estimated based on three different datasets, daily gridded Tmax from (1) CPC, (2) Berkeley Earth (BE), and (3) ERA5. The results illustrated in Fig S1 suggest that corresponding to each of the three datasets, the 2001 to 2016 period is found to be considerably warmer than the preceding period, 1983-2000. This is indicated by the positive anomalies in temperature observed since the year 2001 that continues to increase almost uninterruptedly (except for the year 2004) until the year 2016 (up to 0.5-0.75°C across all three datasets). Importantly, the temporal pattern exhibited by the magnitude of these anomalies shows a close agreement among all the three datasets.

B.3. Trend Analysis:

Interannual trends in the CDHW$_f$, CDHW$_d$, and CDHW$_s$ are estimated by using the Sen’s slope estimator (Sen, 1968), and tested for significance based on the Mann-Kendall (MK) trend test (Kendall, 1948; Mann, 1945) under the null hypothesis of no
trend. A detailed explanation on the MK test and Sen slope estimator is provided in the following sub-sections.

**B.3.1. Mann-Kendall test**

The MK test is a non-parametric test that considers the test statistic, $S$ to have zero mean and variance estimated as (Kendall, 1948; Mann, 1945)

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{sgn}(x_j - x_i)$$

(1)

$$\text{sgn}(x_j - x_i) = \begin{cases} 
+1, & \text{if } x_j - x_i > 0 \\
0, & \text{if } x_j - x_i = 0 \\
-1, & \text{if } x_j - x_i < 0 
\end{cases}$$

(2)

$$\text{Var}(S) = \frac{n(n-1)(2n+5) - \sum_{i=1}^{m} t_i(t_i - 1)(2t_i + 5)}{18}$$

(3)

where $n$ is the number of data points, $m$ is the number of tied groups; $t$ denotes the number of ties of extent $i$, $x_i$ and $x_j$ are the data values in the time series at $i^{th}$ and $j^{th}$ data point (such that $j > i$), respectively, and $\text{sgn}(x_j - x_i)$ is the sign function. In cases where the sample size $n > 10$, the standard normal variable $Z$ is computed as,
Positive values of $Z$ indicate increasing trends, while negative values imply decreasing trends in the time series. The null hypothesis of no trend is rejected for an absolute value of $Z$ greater than $Z_{1-\alpha/2}$, obtained from the standard normal cumulative distribution tables. In this study, significance level of $\alpha = 0.05$ is used.

### B.3.2. Sen’s slope estimator

Sen, 1968 developed a non-parametric procedure for estimating the true slope (change per unit time) given a linear trend is present in a time series. The slope of trend in the sample of $N$ pairs of data are first computed as,

$$Q_i = \frac{x_k - x_j}{k - j}, \text{ for } i = 1, \ldots, N$$

(5)

where $x_j$ and $x_k$ are data values in the time series at data points $j$ and $k$ ($k > j$), respectively. The Sen’s estimator of slope is then computed as the median of these $N$ values of $Q_i$ such that,

$$Q_{med} = \begin{cases} Q \lfloor (N+1)/2 \rfloor & \text{if } N \text{ is odd} \\ \frac{1}{2} (Q_{\lfloor N/2 \rfloor} + Q_{\lfloor (N+2)/2 \rfloor}) & \text{if } N \text{ is even} \end{cases}$$

(6)
Finally, the confidence interval of $Q_{med}$ can be computed at specific confidence levels (in this work, $\alpha = 0.05$ is used) as follows:

$$C_\alpha = Z_{1-\alpha/2}\sqrt{Var(S)}$$

(7)

where $Var(S)$ has been computed as in Eq. (3), and $Z_{1-\alpha/2}$ is obtained from the standard normal cumulative distribution tables.

Then, $M_1 = \frac{N - C_\alpha}{2}$ and $M_2 = \frac{N + C_\alpha}{2}$ are computed, where the lower and upper limits of the confidence interval, $Q_{min}$ and $Q_{max}$ are the $M_{th}$ largest and the $(M2+1)th$ largest of the $N$ ordered slope estimates $Q_i$ (Gilbert, 1987). The slope $Q_{med}$ is considered to be significantly different from zero if the two limits ($Q_{min}$ and $Q_{max}$) have similar sign.

**B.4. Statistical Analysis**

**B.4.1. Two-sample Kolmogorov-Smirnov Test**

The two-sample Kolmogorov-Smirnov (KS) test is used to test whether two samples come from the same or different distribution family. KS is a nonparametric test that can evaluate two distribution functions (for two different samples) based on the distance between their empirical distribution functions. The null hypothesis is that the two distribution functions are drawn from the same distribution at a given significance level (here, $\alpha = 0.05$). Here, the two-sample KS test is employed to assesses differences between the spatial distribution of the CDHW event characteristics evaluated for 1982-
1999, and 2000-2016 for different regimes (arid, transitional, and humid) across the globe and the six different continents. The test indicates whether the data from the two periods come from the same distribution at a 0.05 significance level.

For two given samples, one with a sample size of \( m \) and a cumulative distribution function (CDF) of \( F(x) \) and the other with a sample size of \( n \) and a CDF of \( G(x) \). The maximum distance between the two CDFs can be given as,

\[
D_{m,n} = \max_x | F(x) - G(x) |
\]

As for the Kolmogorov-Smirnov test for normality, the null hypothesis of same distribution (at significance level \( \alpha \)) is rejected if \( D_{m,n} > D_{m,n,\alpha} \) where \( D_{m,n,\alpha} \) is the critical value with \( m \) and \( n \) being sufficiently large and \( D_{m,n,\alpha} \) given by

\[
D_{m,n,\alpha} = c(\alpha) \sqrt{\frac{m+n}{mn}}
\]

where \( c(\alpha) = \) the inverse of the Kolmogorov distribution at \( \alpha \)

**B.4.2. Wilcoxon Rank-sum Test**

The Wilkoxon Rank-sum, sometimes called the Mann Whitney U test is a nonparametric test to compare outcomes between two independent populations. It is extensively used to compare the medians between the two populations. In contrast, the null and two-sided research hypotheses for the nonparametric test are stated as follows:

\( H_0 \): The two populations are equal versus

\( H_1 \): The two populations are not equal.
Here, this test is performed as a two-sided test with the null hypothesis that the medians of the populations are equal at a given significance level (here, $\alpha = 0.05$). The observations from the two samples (with say sample sizes $n$, and $m$) are pooled into one combined sample (sample size of $n + m$) while keeping a track of which sample each observation comes from, and then ranking lowest to highest from 1 to $n+m$, respectively.

Here, the test is employed to assess differences between the median of spatial distribution of the CDHW event characteristics evaluated for 1982-1999, and 2000-2016 for different regimes (arid, transitional, and humid) across the globe and the six different continents.
Figures

(a) CPC

(b) BE

(c) ERA5
Figure S1  Anomalies in summer mean of daily maximum temperature (Tmax) for each year between 1983 and 2016 with respect to the long-term (1983-2016) average based on (a) CPC, (b) Berkeley earth surface temperature project, and (c) ERA5 dataset.

Figure S2. (a-f) Spatial map showing the CDHWf (events/year) observed during (left panel) past period, 1983-1999, and (right panel) recent warmer period, 2000-2016, based on (a-b) GPCC-CPC, (c-d) GPCC-BE, and (e-f) ERA5 dataset.
Figure S3. (a-f) Spatial map showing the CDHWd (days/year) observed during (left panel) past period, 1983-1999, and (right panel) recent warmer period, 2000-2016, based on (a-b) GPCC-CPC, (c-d) GPCC-BE, and (e-f) ERA5 dataset.
Figure S4. (a-f) Spatial map showing the CDHWs (per event/year) observed during (left panel) past period, 1983-1999, and (right panel) recent warmer period, 2000-2016, based on (a-b) GPCC-CPC, (c-d) GPCC-BE, and (e-f) ERA5 dataset.
**Figure S5** (a-c) Heatmaps showing the latitudinal variation of CDHWf (events/year) for each year within the period, 1983-2016 based on (a) GPCC-CPC dataset, (b) GPCC-BE dataset, (c) ERA5 dataset, (d-f) same as in (a-c) but for the CDHWd (days/year), and (g-i) same as in (a-c) but for the CDHWs (per event/year).
Figure S6 Figure 4 (a) Spatial map showing the arid, transitional, and humid regions identified based on the Aridity Index for the climatological period, 1983-2016, (b-h) Probability density of Past, 1983-1999 period (black), and recent warm, 2000-2016
period mean CDHWs (per event per year) for the arid, transitional, and humid regions of
the (b) Europe, (c) North America, (d) Asia, (e) South America, (f) Australia, and (g)
Africa based on GPCC-CPC, GPCC-BE, and ERA5 datasets.
Figure S7 Same as in Figure 4 but for CDHWf (events/year)
Figure S8 Same as in Figure 4 but for CDHWd (days/year)
Appendix C

Supplementary Information for Chapter 5

Table S1 List of CMIP5 Models used for the calculation of global mean temperature change

<table>
<thead>
<tr>
<th>Sl. No.</th>
<th>Model</th>
<th>Sl. No.</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ACCESS1-0</td>
<td>18</td>
<td>GISS-E2-H</td>
</tr>
<tr>
<td>2</td>
<td>ACCESS1-3</td>
<td>19</td>
<td>GISS-E2-R</td>
</tr>
<tr>
<td>3</td>
<td>BNU-ESM</td>
<td>20</td>
<td>HadGEM2-AO</td>
</tr>
<tr>
<td>4</td>
<td>CCSM4</td>
<td>21</td>
<td>HadGEM2-CC</td>
</tr>
<tr>
<td>5</td>
<td>CESM1-BGC</td>
<td>22</td>
<td>HadGEM2-ES</td>
</tr>
<tr>
<td>6</td>
<td>CESM1-CAM5</td>
<td>23</td>
<td>IPSL-CM5A-LR</td>
</tr>
<tr>
<td>7</td>
<td>CMCC-CM</td>
<td>24</td>
<td>IPSL-CM5A-MR</td>
</tr>
<tr>
<td>8</td>
<td>CMCC-CMS</td>
<td>25</td>
<td>IPSL-CM5B-LR</td>
</tr>
<tr>
<td>9</td>
<td>CNRM-CM5</td>
<td>26</td>
<td>MIROC-ESM-CHEM</td>
</tr>
<tr>
<td>10</td>
<td>CSIRO-Mk3-6-0</td>
<td>27</td>
<td>MIROC-ESM</td>
</tr>
<tr>
<td>11</td>
<td>CanESM2</td>
<td>28</td>
<td>MIROC5</td>
</tr>
<tr>
<td>12</td>
<td>EC-EARTH</td>
<td>29</td>
<td>MPI-ESM-LR</td>
</tr>
<tr>
<td>13</td>
<td>FGOALS_g2</td>
<td>30</td>
<td>MPI-ESM-MR</td>
</tr>
<tr>
<td>14</td>
<td>FIO-ESM</td>
<td>31</td>
<td>MRI-CGCM3</td>
</tr>
<tr>
<td>15</td>
<td>GFDL-CM3</td>
<td>32</td>
<td>NorESM1-M</td>
</tr>
<tr>
<td>16</td>
<td>GFDL-ESM2G</td>
<td>33</td>
<td>NorESM1-ME</td>
</tr>
<tr>
<td>17</td>
<td>GFDL-ESM2M</td>
<td>34</td>
<td>bcc-csm1-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>35</td>
<td>inmcm4</td>
</tr>
</tbody>
</table>
Table S2 Possible drivers considered in the correlation analysis and in the FLM.

<table>
<thead>
<tr>
<th>Acronyms</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOI</td>
<td>Southern Oscillation Index</td>
</tr>
<tr>
<td>DMI/IOD</td>
<td>Indian Ocean Dipole</td>
</tr>
<tr>
<td>SAM</td>
<td>Southern Annular Mode</td>
</tr>
<tr>
<td>AO</td>
<td>Arctic Oscillation</td>
</tr>
<tr>
<td>NAO</td>
<td>North Atlantic Oscillation</td>
</tr>
<tr>
<td>PDO</td>
<td>Pacific Decadal Oscillation</td>
</tr>
</tbody>
</table>

**Calculation of Degree of Susceptibility of Heatwave towards Drought**

A Bayesian approach was adopted and the probability of having a HW day conditioned on drought weeks \((pe)\) were estimated at each grid-point location and compared with that conditioned on no drought weeks \((pc)\) for the observational period. In other words, we selected the HW days conditioned on drought as our experiment, and HW days conditioned on no drought as the control data set. The following equations illustrate the methodology for the estimation of \(pc\) and \(pe\) using relative frequencies of daily CHWD events at each grid cell locations across the globe during the observational period, 1982-2016.

\[
pe = P(HW\ day \mid Drought) \\
(1)
\]

\[
pc = P(HW\ day \mid No\ Drought) \\
(2)
\]

Therefore, applying the Baye’s theorem we estimated \(pe\) as,
\[ pe = \frac{P(HW \text{ day } \cap \text{ Drought})}{P(Drought)} \]\\
(3)\\
\[ pc = \frac{P(HW \text{ day } \cap \text{ No Drought})}{P(\text{No Drought})} \]\\
(4)\\
\[ \therefore pe / pc = \frac{P(HW \text{ day } \cap \text{Drought}) \times P(\text{No Drought})}{P(HW \text{ day } \cap \text{No Drought}) \times P(Drought)} \]

where,

\[ P(HW \text{ day } \cap \text{Drought}) = \sum_{i=1}^{N} \frac{n_{e_i}}{N} \left\{ \begin{array}{ll} n_{e_i} = 1, & \text{if } HWT_{max,i} \geq TX 90\text{ pct, } & PDSI_{sc_i} \leq PDSI_{10p} \\ n_{e_i} = 0, & \text{otherwise.} \end{array} \right\} \]

(5)\\
\[ P(HW \text{ day } \cap \text{No Drought}) = \sum_{i=1}^{N} \frac{n_{c_i}}{N} \left\{ \begin{array}{ll} n_{c_i} = 1, & \text{if } HWT_{max,i} \geq TX 90\text{ pct, } & PDSI_{sc_i} > 0 \\ n_{c_i} = 0, & \text{otherwise} \end{array} \right\} \]

(6)\\
\[ P(Drought) = \sum_{i=1}^{N} \frac{n_{e_i}}{N} \left\{ \begin{array}{ll} n_{e_i} = 1, & \text{if } PDSI_{sc} \leq PDSI_{10p} \\ n_{e_i} = 0, & \text{otherwise.} \end{array} \right\} \]

(7)
\[
P(\text{No Drought}) = \frac{\sum_{i=1}^{N} n_{c_i}}{N} \begin{cases} 
nc_i = 1, & \text{if } PDSI_{sc} > 0 \\
nc_i = 0, & \text{otherwise}
\end{cases}
\]

(8)

where, \(i\) is the daily time-step, and \(N\) is the total number of daily time steps for the period, 1982-2016; \(HWT_{max_i}\) is the daily \(T_{max}\) corresponding to the \(i^{th}\) time step that constitute a HW day.

Next, we performed a two-proportion z-test (or Chi-square test) to select the \(pe/pc\) values that have statistically significant (at 5\% significance level) higher values than 1. This refers to a one-sided hypothesis testing for which the statistical hypothesis can be formulated as,

\[
H_0: \frac{pe}{pc} \leq 1 \quad \text{versus} \quad H_1: \frac{pe}{pc} > 1 ; \quad \text{Upper tailed test.}
\]

However, the z-statistic is based on a standard normal distribution. Therefore, to remove the normality assumption, the results obtained for the two mutually exclusive events were resampled, producing 1000 realizations each with replacement. The z-statistic was estimated for each of these samples as well as for the observed distribution as,

\[
z = \frac{pe - pc}{\sigma}
\]

(9)

where, \(\sigma\) is the standard error from the pooled sample and can be given as,

\[
\sigma = \sqrt{p(1-p)\left(\frac{1}{ne} + \frac{1}{nc}\right)}
\]

(10)
The DSHW towards an existing drought condition was measured by the $pe/pc$ ratio that showed statistically significant value greater than 1.
Figure S1 Average number of CHWD events during the (a) Pre-2000 period (1983-1999) and (b) Post-2000 period (2000-2016)
Figure S2 Degree of susceptibility of heatwave towards drought.
Appendix D

Supplementary Information for Chapter 6

D.1. Calculation of Heat Index

The computation of heat index (HI) is performed using a multi-stage algorithm that applies a simple formula proposed by Steadman, 1979b, 1979a and a regression equation along with some adjustments (Rothfusz, 1990). The equations consider the relationship between temperature and relative humidity. The procedure for the calculation of HI is available at https://www.wpc.ncep.noaa.gov/html/heatindex_equation.shtml and can be summarized in the following four steps:

Step 1: Calculate the HI using the simple formula provided by Steadman, 1979b, 1979a given as,

\[
HI = 0.5 \times \left\{ T + 61.0 + \left[ (T - 68.0) \times 1.2 \right] + \left( RH \times 0.094 \right) \right\}
\]

(D.1.1)

where \( T \) is the temperature in °F and \( RH \) is relative humidity in percent. HI is the heat index expressed as an apparent temperature in °F.

Step 2: Average the HI with the T

Step 3: If the result in step 2 is equal to or greater than 80°F, then apply the regression equation (Rothfusz, 1990) given as,

\[
HI = -42.34 + 2.05 \times T + 10.14 \times RH - 0.22 \times T \times RH - 0.0068 \times T \times T
- 0.05481717 \times RH \times RH + 0.0012 \times T \times T \times RH + 0.00085 \times T \times RH \times RH
- 0.00000199 \times T \times T \times RH \times RH
\]

(D.1.2)
Step 4: If the RH is less than 13% and the $T$ is between $80^\circ F$ and $112^\circ F$, then the following adjustment is subtracted from the HI calculated in step 3,

$$HI_{adj(-)} = \frac{(13-RH)}{4} \times \sqrt{\frac{17-[T-95]}{17}}$$

(D.1.3)

Step 5: If the RH is greater than 85% and the $T$ is between $80^\circ F$ and $87^\circ F$, then the following adjustment is added to HI calculated in step 3:

$$HI_{adj(+)} = \left(\frac{RH-85}{10}\right) \times \left(\frac{87-T}{5}\right)$$

(D.1.4)

Note that we used daily maximum temperature ($T_{max}$) to calculate HI in this study.

D.2. Bias Correction of daily HI projections:

We applied a non-stationary bias correction technique, following the Equidistant Cumulative Distribution Function Matching (EDCDFm) procedure proposed by (Li et al., 2010). The EDCDFm is an improvement over the traditional Quantile Mapping (QM) approach that has been shown to produce reliable bias corrected GCM outputs for temperature projections under non-stationary conditions (Li et al., 2010; Miao et al., 2016; Wang & Chen, 2014). Unlike the traditional QM approach, it incorporates the change in variance and skewness in addition to considering shift in the mean of the distribution. A non-parametric approach to estimate cumulative distribution functions (CDFs) of daily HI from observations and GCMs is applied in the study. The quantiles obtained from the non-parametric CDF for the projections of a specific GCM (period,
2006-2100) was subsequently mapped onto the same quantile of the observed distribution, in the reference period, 1979-2005. Similarly, the quantiles obtained from the CDF of future projections of daily HI was subsequently mapped onto the same quantile of the distribution of the historical model, in the reference period, 1979-2005.

This improved quantile mapping method can be mathematically written as

\[
T_{\text{model\_adjusted}} = T_{\text{model\_proj}} + F_{\text{obs\_ref}}^{-1}\left(F_{\text{model\_proj}}(T_{\text{model\_proj}})\right) - F_{\text{model\_ref}}^{-1}\left(F_{\text{model\_proj}}(T_{\text{model\_proj}})\right)
\]

\[(\text{D.1.5})\]

where \(F_{\text{model\_proj}}\) is the CDF of the model for a future projection (RCP4.5, and RCP8.5 scenario) period, and \(F_{\text{obs\_ref}}^{-1}\), and \(F_{\text{model\_ref}}^{-1}\) are quantile functions for observations and model (historical scenario) in the reference period, respectively.

**D.3. Parameter Estimation, Model Selection, and Significance Testing**

The parameter estimation and model selection was performed by using the log-likelihood function which can be derived from the cumulative distribution function provided in equation (3) of Chapter 6. For a sample of \(n\) years of annual maxima with values \(x_i = (x_1, \ldots, x_n)\), the log-likelihood function can be derived for \(\xi \neq 0\) as (Coles, 2001a)

\[
l = \log L(x_i | \mu, \sigma, \xi) = -n \log \sigma - (1 + 1/\xi_0) \sum_{i=1}^{n} \log \left[1 + \xi_0 \left(\frac{x_i - \mu}{\sigma_i}\right)\right] - \sum_{i=1}^{n} \left[1 + \xi_0 \left(\frac{x_i - \mu}{\sigma_i}\right)\right]^{-1/\xi_0}
\]

\[(\text{C.1.6})\]

and for \(\xi = 0\) as
\[ l = \log L(x_i | \mu_i, \sigma_i) = -n \log \sigma_i - \sum_{i=1}^{n} \left( \frac{x_i - \mu_i}{\sigma_i} \right) - \sum_{i=1}^{n} \exp \left[ -\left( \frac{x_i - \mu_i}{\sigma_i} \right) \right] \]

\[ \text{(D.1.7)} \]

The stationary and non-stationary models (Table S2) are fitted to the time series of 1d-HS, 3d-HS, and 7d-HS of the summer season, separately, by maximizing the log-likelihood function. After fitting the candidate models (Table S2), the best models were selected based on the Bayesian information criterion (BIC; (Schwarz, 1978)). The BIC introduces a penalty term in order to compensate for the increasing number of parameters in the model as,

\[ BIC = -2\hat{l} + p \log n \]

\[ \text{(D.1.8)} \]

where, \( \hat{l} \) is the maximized log-likelihood of the model containing \( p \) parameters and \( n \) samples. The model with the smallest BIC value is chosen as the best-fit model.

There is a high possibility that the selected models based on the BIC values are obtained by chance, therefore, we performed a likelihood ratio test (LRT) to find if the inclusion of the time trends as covariates in the parameters of the selected model contributes significantly to the model performance(Katz, 2013). The LRT is expressed as the difference in the deviance statistics of the nested models, mathematically represented as,
\[ LRT = -2 \log_e \left( \frac{L_s}{L_g} \right) \]
\[ = -2 \left( \log_e(L_s) - \log_e(L_g) \right) \]
\[ = \text{deviance}_s - \text{deviance}_g \]

(D.1.9)

where \( L_s \) is the likelihood function for the simpler (s) model, and \( L_g \) represents the same but for the general (g) model. The simpler model has fewer parameters than the general model. It is assumed that the difference between the deviance statistics as derived from the equation (D.1.9.) approximately follows the \( \chi^2 \) distribution with 1 degree of freedom under the null hypothesis. Finally, for any given level of significance testing (\( \alpha=5\% \)), the null hypothesis of \( LRT=0 \) is rejected in favor of the general model if the value of \( LRT \) is found to be greater than the \( (1-\alpha)^{th} \) quantile of the \( \chi^2 \) distribution with degrees of freedom equal to the difference in the number of parameters between the nest models.

**A.4. Estimation of Return Levels**

The return level for a given time period are estimated at CONUS gridpoints by using the best-fit stationary and non-stationary GEV models. We estimate effective 40-year return levels (referred to as 40YHS) and the \( \Pr(\text{HS}>40.6) \) that fulfill our objective to examine the high-end HS severity and exceedances above a critical limit of 40.6°C to assess the heat burden(Matthews et al., 2017). Subsequently, the 40YHS value is calculated by using the estimated GEV parameters as(Coles, 2001b):
where \( p (=1/40) \) is the exceedance probability, and \( \hat{\mu}, \hat{\sigma}, \) and \( \hat{\xi} \) are the GEV parameters for the best-fit GEV model at any gridpoint. Return levels for the stationary GEV model (Model-1; Table S2) were estimated based on setting the parameters \( \hat{\mu} = \mu_0, \hat{\sigma} = \sigma_0, \) and \( \hat{\xi} = \xi_0. \) On the other hand, the effective 40-yr return level for the non-stationary GEV models (Model-2, and Model-4; Table S2), with the time trend as a covariate only in the location parameter (\( \mu_t \)), is estimated by using the 95th percentile of \( \mu_{t1}, \mu_{t2}, \mu_{t3}, \ldots, \mu_{tn} \) values such that \( \hat{\mu} = Q_{95}\left(\mu_{t1}, \mu_{t2}, \ldots, \mu_{tn}\right), \hat{\sigma} = \sigma_0, \) and \( \hat{\xi} = \xi_0. \) Similarly, in the case of the non-stationary GEV models (Model-3, and Model-5; Table S2), where both location and scale parameters vary with time, the effective 40-yr return levels were derived by estimating the 95th percentile of both \( \mu_t \) and \( \sigma_t \) as,

\[
\hat{\mu} = Q_{95}\left(\mu_{t1}, \mu_{t2}, \ldots, \mu_{tn}\right)
\]
\[
\hat{\sigma} = Q_{95}\left(\sigma_{t1}, \sigma_{t2}, \ldots, \sigma_{tn}\right)
\]
\[
\hat{\xi} = \xi_0
\]

A similar approach for estimation of return levels has been adopted in previous studies related to the non-stationary analysis of extreme temperature events (Ouarda & Charron, 2018)
Tables:

Table S1: List of 9 CMIP5-GCMs based on the historical and RCP (RCP4.5, and RCP8.5) scenarios used in the study. The data are analyzed over historical (1979-2005) and future (RCP4.5 and RCP8.5, 2006-2100) periods.

<table>
<thead>
<tr>
<th>Sl. No.</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CanEMS2</td>
</tr>
<tr>
<td>2</td>
<td>CSIRO-Mk3-6-0</td>
</tr>
<tr>
<td>3</td>
<td>GFDL-ESM2G</td>
</tr>
<tr>
<td>4</td>
<td>IPSL-CM5A-LR</td>
</tr>
<tr>
<td>5</td>
<td>IPSL-CM5A-MR</td>
</tr>
<tr>
<td>6</td>
<td>MIROC5</td>
</tr>
<tr>
<td>7</td>
<td>MIROC-ESM</td>
</tr>
<tr>
<td>8</td>
<td>MIROC-ESM-CHEM</td>
</tr>
<tr>
<td>9</td>
<td>MRI-CGCM3</td>
</tr>
</tbody>
</table>

Table S2: The five competing stationary and non-stationary GEV models used in the study.

<table>
<thead>
<tr>
<th>Models</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Location ((\mu))</td>
</tr>
<tr>
<td>Model-1</td>
<td>(\mu_0)</td>
</tr>
<tr>
<td>Model-2</td>
<td>(\mu_0 + \mu_1 t)</td>
</tr>
<tr>
<td>Model-3</td>
<td>(\mu_0 + \mu_1 t)</td>
</tr>
<tr>
<td>Model-4</td>
<td>(\mu_0 + \mu_1 t + \mu_2 t^2)</td>
</tr>
<tr>
<td>Model-5</td>
<td>(\mu_0 + \mu_1 t + \mu_2 t^2)</td>
</tr>
</tbody>
</table>
Figures:

**Figure S1** (a-i) Spatial map of CONUS showing the mean of annual maximum daily maximum temperature (Tx) for the (a) present (1980-2019), (b) near-future (2020-2059),
and (c) far-future (2060-2099) climate scenario corresponding to 1day event (1d-Tx) for the RCP4.5 emission scenario, (d-f) same as in (a-c) but for 3d-Tx event, and (g-i) same as in (a-c) but for 7d-Tx event.

**Figure S2** Change in annual extreme HS severity. (a-o) spatial maps of change in the severity of (a) 1d-HS, (b) 3d-HS, and (c) 7d-HS events in the present climate, (d-f) same
as in (a-c) but during the near-future corresponding to RCP4.5 scenario, (g-i) same as in (a-c) but during near-future corresponding to RCP8.5 scenario, (j-l) same as in (a-c) but during far-future corresponding to RCP4.5 scenario, and (m-o) same as in (a-c) but during far-future corresponding to RCP8.5 scenario. Changes in severity of the HS events are estimated based on Sens's slope methodology (Sen, 1968) for each of the 40-year scenarios and gridpoints and the significant trends (tested at 5% significance level based on the pre-whitened Mann Kendall Trend test (Kendall, 1948; Mann, 1945)) are marked by stippling.

**Figure S3** (a-i) Map of mean of extreme summer HS severity for the CONUS in the (a) present (1980-2019), (b) near-future (2020-2059), and (c) far-future (2060-2099) climate scenario corresponding to 1d-HS event for the RCP4.5 emission scenario, (d-f) same as in (a-c) but for 3d-HS event, and (g-i) same as in (a-c) but for 7d-HS event.
Figure S4 (a-c) Maps of best-fit GEV-models obtained for the (a) historical, (b) near-future based on RCP4.5, (c) near-future based on RCP8.5, (d) far-future based on RCP4.5, and (e) far-future based on RCP8.5 climate scenario for the 1d-HS events, (f-j) same as in (a-e) but for 3d-HS events, and (k-o) same as in (a-e) but for 7d-HS events. Model-1 represents the stationary GEV model, while Model-2 to Model-5 represents the nonstationary GEV models. Spatial maps showing the range of p-values corresponding to 1d-HS events obtained from LR-test considering Model-1 as the simpler (s) model, and the models, Model-2 to Model-5 as the general (g) model, (f-j) same as in (a-e) but for 3d-HS events, and (k-o) same as in (a-e) but for 7d-HS events.
**Figure S5** (a-c) Spatial distribution of (a-e) location, (f-j) scale and (k-o) shape parameters corresponding to the AM-HI for the 1d-HS events corresponding to the five scenario-time period combinations used in the study. The estimation of the parameters is performed following equation A.1.11.
Figure S6 (a-c) Spatial distribution of 40YHI (°C) severity for the (a) present, (b) near-future based on RCP8.5, and (c) far-future based on RCP8.5 climate scenario corresponding to 1d-HS event, (d-f) same as in (a-c) but for 3d-HS event, (g-i) same as in (a-c) but for 7d-HS event.
Figure S7 same as in Figure S6 but for 40YTx (°C).
Figure S8 Spatial maps showing the decadal trends in Warm Tail Spread (WTS)(Spangler & Wellenius, 2020) for (a-e) 1d-HS, (f-j) 3d-HS, and (k-o) 7d-HS events.
Figure S9 Spatial maps showing the trends in WTS for each CMIP5 model for the 1d-HS events corresponding to the five scenario-time period combinations used in the study.
Figure S10 Kernel density plots for the trends in WTS for the sub-regions in the nine individual CMIP5 models for the 1d-HS events corresponding to the five scenario-time period combinations used in the study.
Figure S11 Gridded population count for (a) Present (2015), (b) near-future based on SSP2, (c) near-future based on SSP5, (d) far-future based on SSP2, and (e) far-future based on SSP5.
Appendix E

Supplementary Information for Chapter 7

AR Detection Criteria

A brief discussion on the AR-detection procedure and criteria (Guan and Waliser, 2015) is summarized in the following five steps.

(i) Thresholding of IVT strength: ARs are defined based on the IVT intensity exceeding a fixed lower limit (100 kg m$^{-1}$ s$^{-1}$) or a threshold value estimated at every grid cell and for each of the 12 months. This lower limit of 100 kg m$^{-1}$ s$^{-1}$ is selected based on manual examination, and corroboration with previous AR detection results (Gorodetskaya et al., 2014; Guan and Waliser, 2015). While, the fixed lower limit is applied to all grid cells, it is observed that varying the lower limit between 50 and 150 kg m$^{-1}$ s$^{-1}$ brings no notable impact on the AR detection (Guan and Waliser, 2015). The selected threshold employed in the AR detection methodology is therefore robust to picking up noise. The IVT threshold for a month is calculated as the 85th percentile of IVT intensity over all time steps during an overlapping 5-month window centered around that month between 1979 and 2019. The 85th percentile threshold is found to skillfully capture the ARs across the Central United States (Nayak et al., 2014), and it has been applied in other regional studies including SEUS and Britain (Debbage et al., 2017; Lavers et al., 2012).

(ii) Determining the AR-objects: The grid cells where the IVT intensity exceeds either one of the corresponding threshold or 100 kg m$^{-1}$ s$^{-1}$ (whichever is greater) are extracted and retained for identifying a set of AR-objects corresponding to each time-
step. The set of AR objects determined from these grid cells are such that any object artificially cut off at the circular longitudinal boundary is considered as one object. The retained objects are further checked for direction, geometry, and landfall criteria in the subsequent steps.

(iii) Checking for IVT direction: The retained AR objects for each time-step are checked for (a) coherence in IVT direction such that an object is discarded if more than half of the grid cells have IVT deviating more than 45° from the mean IVT of the object; (b) an appreciable poleward component with a minimum IVT intensity of 50 kg m$^{-1}$ s$^{-1}$, and (c) consistency between the mean IVT direction and overall orientation such that an object is discarded if the direction of the mean IVT deviates from the overall orientation by more than 45°.

(iv) Checking for Geometry: The AR objects that fulfill the requirements of IVT direction are further subjected to a check for the object’s geometry. For an object, the axis is first identified, based on which its corresponding object-length, and -width are determined. The objects having a length greater than or equal to 2,000 km and a length-to-width ratio greater than or equal to 2 are retained as ARs. The ARs, retained hereafter, are checked for landfall requirements in the subsequent step. While employing this algorithm, we considered a broad spatial domain extending up to 55°W latitude over the Atlantic Ocean, and the Caribbean and neighboring islands less than 2.5x2.5° are not considered in the analysis (Guan and Waliser, 2015). The small islands are considered as the ocean surface following the original algorithm (Guan and Waliser, 2015).
Furthermore, following the same algorithm, the 2000 km length criteria have been employed by Debbage et al., 2017 to analyze AR interactions along the coastal SEUS.

(v) Checking for landfall: If the AR intersects the coastline such that the mean AR IVT is directed onshore, the intersecting coastal grid cells are identified. The landfall characteristics are then determined based on the identified intersecting coastal grids such that the grid cell with the maximum onshore IVT intensity is selected as the location of landfall. Finally, an AR is considered to have made a landfall if the remaining length of the AR located over the ocean is found to be greater than or equal to 1000 km.

Figure S1 Geographical distribution of (a) 95th and (b) 99th percentile thresholds (in mm) based on which the EP events are calculated in the study. The threshold is estimated
based on all non-zero daily precipitation totals at each precipitation grid for the period 1979-2019.
Appendix F

Supplementary Information for Chapter 8

F.1. Flash drought Identification Criteria based on ESR

The flash drought events are identified by following the criteria proposed by \(^1\) that emphasizes both longevity and impact of rapidly intensifying drought events. The criteria are summarized as

(a) Flash drought events are required to have a minimum length of five SESR changes \(\Delta \text{SESR} \), which is equivalent to a duration of six pentads (30 days).

(b) A final SESR value below the 20th percentile of SESR values.

(c) The \(\Delta \text{SESR} \) value must be at or below the 40th percentile between individual pentads, and no more than one \(\Delta \text{SESR} \) above the 40th percentile following a \(\Delta \text{SESR} \) that meets the former criterion.

(d) The mean change in SESR (\(\Delta \text{SESR} \)) during the entire duration of the flash drought must be less than the 25th percentile of the climatological changes in SESR for that grid point and time of year.

F.2. Calculation of Flash Drought Intensity based on ESR

Following \(^1\), flash droughts were categorized by four drought intensity indices (FD1, FD2, FD3, and FD4) based on the moderate (FD1) to exceptional (FD4) rate of intensification calculated by using the mean SESR change (\(\Delta \text{SESR} \)) during the entire event. The percentile thresholds used for each category are illustrated in Table S2. It is important to note that at a given grid point all percentile values used in criteria (b) and (c) in A.1., were taken from the pentad climatology of all SESR available between 1980 and
2018, whereas, for criterion (d), percentiles were calculated from the distribution of $\Delta\text{SESR}$ for all pentads that were encompassed within the flash drought event. The Flash drought intensity is calculated as the mean SESR change ($\Delta\text{SESR}$) during the flash drought period \(^1\). A more detailed explanation on the flash drought identification methodology is available in \(^1\).

F.3. Calculation of Potential Evaporation using Priestley-Taylor Equation

The Priestly-Taylor (PT) equation \(^2\) is extensively used in the calculation of potential evaporation (PE), applied to drought quantification. It is a simplified version of the Penman-Monteith method, where the vapor pressure deficit and convection terms are reduced to an empirical constant, \(\alpha(=1.26)\). The PT equation is given as,

\[
PE = \alpha \times \frac{\Delta \times (R_s - G)}{\lambda_v \times (\Delta + \gamma)} \times 1000,
\]

where, PE is in mm/day, \(\lambda_v\) is volumetric latent heat of vaporization, 2453 MJ m\(^{-3}\), \(G\) is soil heat flux density (MJ m\(^{-2}\) day\(^{-1}\)), \(\gamma(=0.000665 \times P)\) is the psychrometric constant (kPa °C\(^{-1}\)), \(P\) is the atmospheric pressure (kPa), \(R_n\) is the surface net radiation (MJ m\(^{-2}\) day\(^{-1}\)) calculated as, \(R_n = R_s - R_L\), where, \(R_s\) and \(R_L\) are the surface net incoming shortwave radiation, and surface net outgoing longwave radiation, respectively. \(\Delta\) is slope of the saturation vapor pressure-temperature curve (kPa °C\(^{-1}\)) calculated as,

\[
\Delta = \frac{4098 \times e_s}{(T_a + 237.3)^2}
\]

where, \(e_s\) is the saturation vapor pressure (kPa) given as,
\[ e_s = 6.112 \times \left(1 + 7 \times 10^{-4} + 3.46 \times 10^{-6} \times P_{surf}\right) \times e^{\frac{17.67 T_a}{T_a + 243.5}} \]  

(3),

\(T_a\) is the daily mean 2m air temperature (°C), and \(P_{surf}\) is the surface pressure (hPa).

Tables

Table S1 Thresholds used for flash drought classification.

<table>
<thead>
<tr>
<th>Flash drought intensity index</th>
<th>Flash drought intensity</th>
<th>Mean change in SESR or mean 1m RZSM (percentile)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FD1</td>
<td>Moderate flash drought</td>
<td>20th - 25th</td>
</tr>
<tr>
<td>FD2</td>
<td>Severe flash drought</td>
<td>15th - 20th</td>
</tr>
<tr>
<td>FD3</td>
<td>Extreme flash drought</td>
<td>10th - 15th</td>
</tr>
<tr>
<td>FD4</td>
<td>Exceptional flash drought</td>
<td>&lt;10th</td>
</tr>
</tbody>
</table>

Table S2 Aridity index (AI) classification system based on \(^3\).

<table>
<thead>
<tr>
<th>Zones</th>
<th>Range of Aridity Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyper-arid</td>
<td>AI &lt; 0.05</td>
</tr>
<tr>
<td>Arid</td>
<td>0.05 &lt; AI &lt; 0.20</td>
</tr>
<tr>
<td>Semi-arid</td>
<td>0.20 &lt; AI &lt; 0.50</td>
</tr>
<tr>
<td>Sub-humid</td>
<td>0.50 &lt; AI &lt; 0.65</td>
</tr>
<tr>
<td>Humid</td>
<td>0.65 &lt; AI</td>
</tr>
</tbody>
</table>
Figures

Figure S1 Spatial map showing the percent of (a) $F_{D_{ESR}}$, and (b) $F_{D_{RZSM}}$ years between 1980 and 2018 based on ERA5 dataset, and (c-d) same as in (a-b) but based on MERRA2 dataset.
Figure S2 Spatial map showing the climatological mean (a) Pr (mm/day/year), (b) climatological mean PET (mm/day/year), and (c) AI calculated for the period, 1980-2018. The spatial map in (c) illustrates the humid, sub-humid, semi-arid, arid, and hyper-arid climate regimes classified across the globe based on the range of AI (Table S2).
**Figure S3** Same as in Fig 2 but based on MERRA2 dataset.
Figure S4 Spatial maps showing the mean standardized anomalies of (left) Pr, (2nd column) Tmax, (3rd column) VPD, and (right) pi based on the ERA5 dataset calculated at one pentad intervals starting from 4 pentads before (denoted by -4) to 3 pentads after (denoted by 3) the onset timings of $F_{DESR}$. 
Figure S5 same as in Fig S4 but for mean standardized anomalies based on the MERRA2 dataset.
**Figure S6** same as in Fig S4 but for mean standardized anomalies based on the ERA5 dataset calculated at one pentad intervals starting from 4 pentads before (denoted by -4) to 3 pentads after (denoted by 3) the onset timings of $\text{FD}_{\text{RZSM}}$. 
Figure S7 same as in Fig S4 but for mean standardized anomalies based on the MERRA2 dataset calculated at one pentad intervals starting from 4 pentads before (denoted by -4) to 3 pentads after (denoted by 3) the onset timings of $F_{D_{RZSM}}$. 
Figure S8 same as in Figure 4 but for climate anomalies calculated based on MERRA2 dataset. Note that the FD_{ESR} and FD_{RZSM} intensity are calculated based on the GLEAM dataset.
Appendix G

Supplementary Information for Chapter 9

G.1. Estimation of Vapor Pressure Deficit

In this study, daily vapor pressure deficit (VPD) is calculated by using daily mean dew point temperature, daily mean 2m air temperature, and surface pressure obtained from the ERA5 dataset. The VPD estimation based on these variables is given as\(^\text{68}\),

\[
AVP = 6.112 \times f_w \times e^{\frac{17.67T_d}{T_d + 243.5}}
\]

where, AVP is the actual vapor pressure and \(T_d\) is the dew point temperature (ºC),

\[
SVP = 6.112 \times f_w \times e^{\frac{17.67T_a}{T_a + 243.5}}
\]

where, SVP is the saturation vapor pressure (kPa), and \(T_a\) is the daily mean 2m air temperature (ºC), and \(f_w\) is given as,

\[
f_w = 1 + 7 \times 10^{-4} + 3.46 \times 10^{-6} \times P_{\text{surf}}
\]

where, \(P_{\text{surf}}\) is the surface pressure in (hPa).

Finally, VPD is calculated as,

\[
VPD = AVP - SVP
\]
Figures

**Figure S1** Spatial maps showing the estimates of AF (%) of dry days on hot day occurrence in a dry-hot cascade network considering (a) 0 days, (b) 1 day, (c) 2 days, (d) 3 days, (e) 4 days, (f) 5 days, (g) 6 days, and (h) 7 days lag.

**Figure S2** Spatial maps showing the lower bound of (5-95%) confidence interval for AF (%) of dry days on hot day occurrence in a dry-hot cascade network considering (a) 0
days, (b) 1 day, (c) 2 days, (d) 3 days, (e) 4 days, (f) 5 days, (g) 6 days, and (h) 7 days lag.

**Figure S3** Spatial maps showing the upper bound of (5-95%) confidence interval for AF (%) of dry days on hot day occurrence in a dry-hot cascade network considering (a) 0 days, (b) 1 day, (c) 2 days, (d) 3 days, (e) 4 days, (f) 5 days, (g) 6 days, and (h) 7 days lag.
**Figure S4** Spatial maps showing the estimates of AF (%) of hot days on dry day occurrence in a hot-dry cascade network considering (a) 0 days, (b) 1 day, (c) 2 days, (d) 3 days, (e) 4 days, (f) 5 days, (g) 6 days, and (h) 7 days lag.

**Figure S5** Spatial maps showing the lower bound of (5-95%) confidence interval for AF (%) of hot days on dry day occurrence in a hot-dry cascade network considering (a) 0 days, (b) 1 day, (c) 2 days, (d) 3 days, (e) 4 days, (f) 5 days, (g) 6 days, and (h) 7 days lag.
Figure S6: Spatial maps showing the upper bound of (5-95%) confidence interval for AF (%) of hot days on dry day occurrence in a hot-dry cascade network considering (a) 0 days, (b) 1 day, (c) 2 days, (d) 3 days, (e) 4 days, (f) 5 days, (g) 6 days, and (h) 7 days lag.

Figure S7: Spatial distribution of lagged correlation coefficients between daily SM (lagged) and Tmax for (a) 0 days, (b) 1 day, (c) 2 days, (d) 3 days, (e) 4 days, (f) 5 days, (g) 6 days, and (h) 7 days lag.
Figure S8 Spatial distribution of lagged correlation coefficients between daily Tmax (lagged) and SM for (a) 0 days, (b) 1 day, (c) 2 days, (d) 3 days, (e) 4 days, (f) 5 days, (g) 6 days, and (h) 7 days lag.

Figure S9 Spatial distribution of statistically significant (at 5% significance level) odd ratios \( \exp(\beta) \) corresponding to ESR anomalies calculated by fitting the logistic regression model for the dry-hot cascade network for a lag of (a) 0 days, (b) 1 day, (c) 2
days, (d) 3 days, (e) 4 days, (f) 5 days, (g) 6 days, and (h) 7 days. Note that, a decrease in ESR value indicates an increase in evaporative stress.

Figure S10 same as in figure S9 but corresponding to surface net radiation (Rn) anomalies.

Figure S11 same as in figure S9 but corresponding to precipitation (Pr) anomalies.
Figure S12 Spatial distribution of statistically significant (at 5% significance level) odd ratios ($\exp(\beta)$) corresponding to surface net-radiation (Rn) anomalies calculated by fitting the logistic regression model for the hot-dry cascade network for a lag of (a) 0 days, (b) 1 day, (c) 2 days, (d) 3 days, (e) 4 days, (f) 5 days, (g) 6 days, and (h) 7 days.

Figure S13 same as in figure S12 but corresponding to precipitation (Pr) anomalies.
**Figure S14** same as in figure S12 but corresponding to VPD anomalies.