Best Practices for Incorporating Non-stationarity in Extreme Precipitation and Flooding Design Values

SERDP Project RC-2517

KENNETH E. KUNKEL
North Carolina State University

KATHERINE SCHLEF
Western New England University

CASEY BROWN AND BAPTISTE FRANÇOIS
University of Massachusetts

YONAS DEMISSIE AND EUGENE YAN
Washington State University

DENNIS P. LETTENMAIER AND KIMBERLY J. WANG
University of California, Los Angeles

ANNA WAGNER
U.S. Army Engineer Research and Development Center
Cold Regions Research and Engineering Laboratory

MARK S. WIGMOSTA
Pacific Northwest National Laboratory

THOMAS R. KARL
Climate and Weather, LLC

DAVID R. EASTERLING
National Oceanic and Atmospheric Administration

Distribution Statement A
This document has been cleared for public release

SERDP
DOD • EPA • DOE
This report was prepared under contract to the Department of Defense Strategic Environmental Research and Development Program (SERDP). The publication of this report does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official policy or position of the Department of Defense. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the Department of Defense.
**REPORT DOCUMENTATION PAGE**

**Title:** Best Practices for Incorporating Non-stationarity in Extreme Precipitation and Flooding Design Values

**Authors:** Kenneth E. Kunkel, North Carolina State University
Katherine Schlef, Western New England University
Casey Brown and Baptiste François, University of Massachusetts
Yonas Demissie and Eugene Yan, Washington State University
Anna Wagner, U.S. Army Engineer Research and Development Center
Cold Regions Research and Engineering Laboratory
Mark S. Wigmosta, Pacific Northwest National Laboratory
Thomas R. Karl, Climate and Weather, LLC
David R. Easterling, National Oceanic and Atmospheric Administration

**Performing Organization:**
North Carolina State University
151 Patton Avenue
Asheville, NC 28801

**Sponsoring/Monitoring Agency:**
Strategic Environmental Research and Development Program (SERDP)
4800 Mark Center Drive, Suite 16F16
Alexandria, VA 22350-3605

**Description:**
This technical report provides guidance to developers of extreme precipitation and flooding design values. There is now strong evidence, both theoretically and observationally, that extreme precipitation is increasing and will continue to increase in the foreseeable future. As a result, new approaches to design values are required. A number of approaches are reviewed and assessed for both extreme precipitation and riverine flooding. Traditional approaches that have relied on a stationary climate are no longer adequate for planning beyond a decade or two. Global Climate Model simulations, new nonstationary statistical models, next-generation Interval-Duration-Frequency curves, emerging methods for regional aggregation of extreme events, adaptive management of infrastructure design, and accommodation of the additional risk related to imperfect accuracy of projections of conditions are all important tools to address the changes that lie ahead. Each of these is reviewed and assessed in this “Best Practices” technical report.

**Subject Terms:**
Non-stationarity, Extreme Precipitation, Flooding Design Values, Anthropogenically-Forced Climate Change
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Executive Summary</td>
<td>2</td>
</tr>
<tr>
<td>1.0</td>
<td>INTRODUCTION</td>
<td>3</td>
</tr>
<tr>
<td>2.0</td>
<td>Observed Meteorological Changes and Linkages to Flooding</td>
<td>5</td>
</tr>
<tr>
<td>2.1</td>
<td>Large-Area Meteorological Changes</td>
<td>5</td>
</tr>
<tr>
<td>2.2</td>
<td>Regional Meteorological Changes</td>
<td>6</td>
</tr>
<tr>
<td>2.3</td>
<td>Regional Hydrologic Flooding</td>
<td>8</td>
</tr>
<tr>
<td>3.0</td>
<td>Assessing and Understanding Atmospheric Drivers of Change</td>
<td>13</td>
</tr>
<tr>
<td>3.1</td>
<td>Climate Model Projections</td>
<td>13</td>
</tr>
<tr>
<td>3.2</td>
<td>Assessing the Meteorology of Extreme Precipitation Within Climate Models</td>
<td>17</td>
</tr>
<tr>
<td>4.0</td>
<td>Frameworks for Developing IDF Values</td>
<td>19</td>
</tr>
<tr>
<td>4.1</td>
<td>Exclusive Use of Recent Historical Data</td>
<td>19</td>
</tr>
<tr>
<td>4.2</td>
<td>Historical Observations Calibrated to GCMs for a Simulated Future</td>
<td>19</td>
</tr>
<tr>
<td>4.3</td>
<td>Historical Observations Calibrated to the Best GCMs</td>
<td>20</td>
</tr>
<tr>
<td>5.0</td>
<td>Estimation of IDF Relationships in a Changing Climate</td>
<td>23</td>
</tr>
<tr>
<td>5.1</td>
<td>Detection of Non-stationarity over Other Deterministic and Stochastic Changes</td>
<td>24</td>
</tr>
<tr>
<td>5.2</td>
<td>Nonstationary Frequency Analysis</td>
<td>26</td>
</tr>
<tr>
<td>5.3</td>
<td>Covariates for Characterizing Changes in Extreme Precipitations</td>
<td>27</td>
</tr>
<tr>
<td>5.4</td>
<td>Functions to Relate Parameters and Covariates</td>
<td>29</td>
</tr>
<tr>
<td>5.5</td>
<td>Non-stationary IDF Curves</td>
<td>30</td>
</tr>
<tr>
<td>6.0</td>
<td>Non-stationary Regional Frequency Analysis and Non-stationary IDAF Relationships for Extreme Precipitation</td>
<td>33</td>
</tr>
<tr>
<td>6.1</td>
<td>Non-stationary Regional Frequency Analysis</td>
<td>33</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Homogeneous Regions</td>
<td>33</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Regional Flood Curves</td>
<td>34</td>
</tr>
<tr>
<td>6.1.3</td>
<td>Non-stationary Trend Detection at the Regional Level</td>
<td>35</td>
</tr>
<tr>
<td>6.1.4</td>
<td>Application to Station Data</td>
<td>36</td>
</tr>
<tr>
<td>6.1.5</td>
<td>Application to Gridded Data</td>
<td>36</td>
</tr>
<tr>
<td>6.2</td>
<td>Updated IDAF Curves</td>
<td>36</td>
</tr>
<tr>
<td>6.2.1</td>
<td>IDF Curve</td>
<td>37</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Areal Reduction Factors</td>
<td>37</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>6.3 Path Forward</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>7.0 Strategies to Address Non-stationarity in Hydrologic Extremes</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>7.1 Estimates from First Principles</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>7.2 Projection by the Hydrologic Simulation Approach</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>7.3 Projection by the Climate-Informed Approach (Statistical Methods)</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>7.4 Calculation of Non-stationary Design Flood</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>7.5 Recommended Practices for Riverine Flooding</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>8.0 Rain-on-Snow/Snowmelt Events and Their Estimation</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>8.1 Current Practice</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>8.2 Physics-Based Models</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>8.3 IDF-Based Hydrologic Design</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>8.3.1 Limitations of Precipitation-Based IDF Curves</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>8.3.2 NG-IDF Based Design</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>9.0 Effective Application of Non-stationary and Uncertain Design Values</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>9.1 Addressing Uncertainty with Design Storms and Floods</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>9.2 Climate Factors</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>9.3 Prudent Approach</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>9.4 Robust-Based Decision Methods</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>9.5 The Way Forward</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>10.0 Recommendations</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>10.1 Short-term Planning (20 years or less)</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>10.2 Long-term Planning (more than 20 years)</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>10.3 Global Climate Models</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>10.4 Special Considerations Related to Riverine Flooding</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>10.5 Rain-on-Snow Riverine Flooding</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>10.6 Regional Averaging</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>10.7 Designing for a Changing Climate</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>References</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>IDF curves for a rain gauge in Lafayette, Louisiana (Latitude: 30.2050, Longitude: -91.9875) (Source: NOAA ATLAS 14, Precipitation Frequency Data).</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Observed changes (%) in heavy precipitation defined different ways for two different time periods (Source: Easterling et al. 2017).</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Longitudinally averaged, multimodel median extreme precipitation (99.9th percentile daily precipitation) for the past (blue) and future (red) under the Special Report on Emission Scenarios (SRES) A1B scenario (a balanced emphasis on all energy sources). The shaded region represents the 50% interquartile range of the climate models’ predictions for the past period (Source: O’Gorman and Schneider 2009).</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Projected change (%) in maximum daily precipitable water (PWmax) by the late 21st century relative to late 20th century under the high (RCP8.5) emissions scenario. This is an average of 13 CMIP5 models (Source: updated from Kunkel et al. 2013).</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Correlation coefficients between the magnitude of the annual daily maximum precipitation values and the simultaneous value of the 3-hour maximum precipitable water for (a) NCEP/NCAR reanalysis for 1949-2017, (b) NCEP/NCAR reanalysis for 1980-2017, and (c) MERRA-2 reanalysis for 1980-2017). Statistically significant (0.05 level) positive slopes are depicted with triangles, while dots represent non-significant slopes (Source: Kunkel, Stevens, et al. 2020).</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Steps to incorporate non-stationarity into IDF curves.</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Schematic representation of the return levels and periods (probability of exceedances) for stationary and non-stationary frequency analysis (Source: Salas and Obeysekera 2014).</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Schematics of the two primary approaches for flood projections. These are intended to be examples and not prescriptive for every analysis.</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Comparison of the measured and modeled snow water equivalent (SWE) in the (A) open, (B) forest, and (C) gap sites. Note the different y-axis scales. In (C), the simulated SWE at the forest site is replotted on the same scale for comparison (Source: Sun et al. 2018).</td>
<td></td>
</tr>
</tbody>
</table>
| 12     | Relative percentage of differences ($q_{NG} - q_{PREC} / q_{PREC} \times 100$) of the 10-, 25-, 50, and 100-year events between the peak design floods derived from NG-IDF ($q_{NG}$) and PREC-IDF ($q_{PREC}$). Stations with red circles indicate underdesign ($q_{NG} > q_{PREC}$), stations with blue circles suggest overdesign ($q_{NG} < q_{PREC}$), and black dots indicate proper design.
for the stations. The diameter of the circles indicates the magnitude of the relative difference (Source: Yan et al. 2018). ............................................................................................................. 48

Figure 13. Top row: the relative peak design flood error \((q_{\text{PREC}} - q_{\text{DHSVM}})/ q_{\text{DHSVM}} \times 100)\) for (a) 10-year event, (b) 25-year event, and (c) 50-year event between the peak design floods derived from PREC-IDF \(q_{\text{PREC}}\) and DHSVM continuous simulations \(q_{\text{DHSVM}}\). Bottom row: the relative peak design flood error \((q_{\text{NG}} - q_{\text{DHSVM}})/ q_{\text{DHSVM}} \times 100)\) for (d) 10-year event, (e) 25-year event, and (c) 50-year event between the peak design floods derived from NG-IDF \(q_{\text{NG}}\) and continuous simulations \(q_{\text{DHSVM}}\). For all panels, stations with warm colors indicate overpredictions; stations with blue color suggest underpredictions (Source: Yan et al. 2018). ............................................................................................................. 49

Figure 14. Top row: scatterplots of the (a) 10-, (b) 25-, and (c) 50-year normalized peak design floods derived from PREC-IDF curves versus normalized peak design floods from DHSVM continuous simulations for the 246 SNOTEL stations across the western United States. Bottom row: scatterplots of the (d) 10-, (e) 25-, and (f) 50-year normalized peak design floods derived from NG-IDF curves versus normalized peak design floods from DHSVM continuous simulations. For all panels, the black line indicates the 1:1 line. For each return period, all peak floods were normalized by the maximum DHSVM peak flood across all 246 stations (Source: Yan et al. 2018). ............................................................................................................. 50
## ACRONYMS AND ABBREVIATIONS

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEP</td>
<td>Annual Exceedance Probability</td>
</tr>
<tr>
<td>AMO</td>
<td>Atlantic Multidecadal Oscillation</td>
</tr>
<tr>
<td>AMP</td>
<td>Annual Maximum Precipitations</td>
</tr>
<tr>
<td>AR</td>
<td>Atmospheric Rivers</td>
</tr>
<tr>
<td>ARF</td>
<td>areal reduction factor</td>
</tr>
<tr>
<td>ARI</td>
<td>Annual Return Intervals</td>
</tr>
<tr>
<td>AWR</td>
<td>available water for runoff</td>
</tr>
<tr>
<td>CAUSES</td>
<td>ClimAtological Effects Under Synoptic Extreme States</td>
</tr>
<tr>
<td>CCSM</td>
<td>Community Climate System Model</td>
</tr>
<tr>
<td>CDF</td>
<td>Cumulative Probability Function</td>
</tr>
<tr>
<td>CMIP</td>
<td>Coupled Model Intercomparison Project</td>
</tr>
<tr>
<td>CONUS</td>
<td>continental United States</td>
</tr>
<tr>
<td>DEFRA</td>
<td>Department for Environment, Food and Rural Affairs</td>
</tr>
<tr>
<td>DHSVM</td>
<td>Distributed Hydrology Soil Vegetation Model</td>
</tr>
<tr>
<td>e.g.</td>
<td>for example</td>
</tr>
<tr>
<td>ENSO</td>
<td>El Niño Sothern Oscillation</td>
</tr>
<tr>
<td>GAMLSS</td>
<td>Generalized Model for Location, Scale, and Shape</td>
</tr>
<tr>
<td>GCM</td>
<td>Global Climate Model</td>
</tr>
<tr>
<td>GEV</td>
<td>Generalized Extreme Value</td>
</tr>
<tr>
<td>GW-PCA</td>
<td>Geographically Weighted Principal Component Analyses</td>
</tr>
<tr>
<td>i.e.</td>
<td>that is</td>
</tr>
<tr>
<td>IAM</td>
<td>Integrated Assessment Modeling</td>
</tr>
<tr>
<td>IDAF</td>
<td>Intensity-Duration-Area-Frequency</td>
</tr>
<tr>
<td>IDF</td>
<td>Intensity Duration Frequency</td>
</tr>
<tr>
<td>IOD</td>
<td>India Ocean Dipole</td>
</tr>
<tr>
<td>IOP</td>
<td>Institute of Physics</td>
</tr>
<tr>
<td>IPCC</td>
<td>Intergovernmental Panel on Climate Change</td>
</tr>
<tr>
<td>km</td>
<td>kilometers</td>
</tr>
<tr>
<td>MC</td>
<td>Monsoon Climate</td>
</tr>
<tr>
<td>MK</td>
<td>Mann–Kendall</td>
</tr>
<tr>
<td>mm</td>
<td>millimeters</td>
</tr>
<tr>
<td>n</td>
<td>years</td>
</tr>
<tr>
<td>NAO</td>
<td>North Atlantic Oscillation</td>
</tr>
<tr>
<td>NG-IDF</td>
<td>next-generation IDF</td>
</tr>
<tr>
<td>NOAA</td>
<td>National Oceanic and Atmosphere Administration</td>
</tr>
<tr>
<td>NWS</td>
<td>National Weather Service</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>P</td>
<td>probability</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>PDO</td>
<td>Pacific Decadal Oscillation</td>
</tr>
<tr>
<td>PNW</td>
<td>Pacific Northwest</td>
</tr>
<tr>
<td>POT</td>
<td>Peak Over Threshold</td>
</tr>
<tr>
<td>ppm</td>
<td>parts per million</td>
</tr>
<tr>
<td>PREC-IDF</td>
<td>precipitation-based Intensity-Duration-Frequency</td>
</tr>
<tr>
<td>PWmax</td>
<td>maximum daily Precipitable Water</td>
</tr>
<tr>
<td>RCPs</td>
<td>Representative Concentration Pathways</td>
</tr>
<tr>
<td>RFA</td>
<td>Regional Frequency Analysis</td>
</tr>
<tr>
<td>RMSE</td>
<td>root mean square error</td>
</tr>
<tr>
<td>ROS</td>
<td>Rain-On-Snow</td>
</tr>
<tr>
<td>SNOTEL</td>
<td>Snowpack Telemetry</td>
</tr>
<tr>
<td>SWE</td>
<td>snow water equivalent</td>
</tr>
<tr>
<td>TC</td>
<td>Tropical Cyclone</td>
</tr>
<tr>
<td>TR</td>
<td>Technical Release</td>
</tr>
<tr>
<td>USGCRP</td>
<td>Climate Science Special Report</td>
</tr>
<tr>
<td>°C</td>
<td>degrees Celsius</td>
</tr>
<tr>
<td>°F</td>
<td>degrees Fahrenheit</td>
</tr>
<tr>
<td>μ</td>
<td>location</td>
</tr>
<tr>
<td>yr</td>
<td>year</td>
</tr>
<tr>
<td>α</td>
<td>significance</td>
</tr>
<tr>
<td>ξ</td>
<td>shape</td>
</tr>
<tr>
<td>σ</td>
<td>scale</td>
</tr>
</tbody>
</table>
Executive Summary

This technical report provides guidance to developers of extreme precipitation and flooding design values. There is now strong evidence, both theoretically and observationally, that extreme precipitation is increasing and will continue to increase in the foreseeable future. As a result, new approaches to design values are required. A number of approaches are reviewed and assessed for both extreme precipitation and riverine flooding. Traditional approaches that have relied on a stationary climate are no longer adequate for planning beyond a decade or two. Global Climate Model simulations, new non-stationary statistical models, next-generation Interval-Duration-Frequency curves, emerging methods for regional aggregation of extreme events, adaptive management of infrastructure design, and accommodation of the additional risk related to imperfect accuracy of projections of conditions are all important tools to address the changes that lie ahead. Each of these is reviewed and assessed in this “Best Practices” technical report.
1 Introduction

Public- and private-sector interests have increasingly become aware of the deleterious effects of suboptimal and ineffective flood mitigation for investment decisions, planning, and operations (e.g., Cusick 2019). Mounting evidence for recent changes in extreme precipitation risks and a growing scientific consensus around future increases in such risks have motivated increased research around approaches to address this issue. Yet, there is no community consensus around best practice methods for flood design risk. The purpose of this technical guidance report is to summarize ongoing work and the state of knowledge on this topic and to provide recommendations for best practices.

A basic tool that has served engineers and planners is the Intensity-Duration-Frequency (IDF) set of curves, which provide a basic foundation for estimating current and future flood risk. IDF curves can provide return period estimates (or probability of exceedances) for various amounts of precipitation (or streamflow) over different durations. A brief review of the construction of these curves is useful as it is used throughout this guide. A variety of commercial and publicly available websites can calculate IDF curves with the input of precipitation (or streamflow) data, and they provide a variety of options for IDF calculations.

For the purposes of illustrating an IDF curve, it is assumed that the data can be well described by probabilities and statistics that do not change with time (i.e., stationarity). The IDF curves can be based purely on empirical data or, more often, with the aid of a probability distribution applied to the observed data. To begin constructing an IDF curve, a set of data is required of some number of years ($n$). The data can be transformed into a series of annual maxima accumulated precipitation for various duration times depending on the resolution of the original data set; e.g., 1 minute, 10 minutes, 30 minutes, 60 minutes, 6 hours, 24 hours, 48 hours, 5 days, etc. The quantity of interest can either be precipitation or streamflow. For example, if there is a 30-year record of precipitation with 1-minute resolution, empirical plotting of the annual maxima for each of the various durations noted above can be ranked across the years from highest (1) to lowest (30). Each of these annual maxima is associated with an empirical exceedance probability ($p$) that can be calculated by:

$$p = \text{rank}/(n + 1)$$  \hspace{1cm} (1)

The value of $p$ can be transformed into empirical annual return intervals (ARIs) with the equation:

$$\text{ARI} = (n + 1)/\text{rank}$$  \hspace{1cm} (2)

The precipitation (or streamflow) for each ranked year can be plotted against its associated return interval for each of the selected durations to form a set of empirical curves. By transforming all the values of precipitation (or streamflow) across various durations to the same units of total accumulated precipitation, a set of IDF curves can be plotted. Figure 1 provides an example of a set of IDF curves that have been developed by fitting an appropriately selected theoretical probability distribution to the dataset of annual maxima (in contrast to plotting the discrete empirically derived values given in equation [2]).
These IDF curves represent the best estimate for extreme precipitation amounts based on the observed dataset, but with the assumption of stationarity, or no change in the parameters governing the probability distribution over time. As will be described in subsequent chapters, this assumption is now increasingly being violated in a number of regions, and this trend is expected to continue into the foreseeable future as the climate changes. New methods are needed to address this problem, as it compounds errors of identifying both the appropriate probability distribution and its associated parameters. Subsequent chapters will present the case for why climate is and will continue to change and what can be done to address this non-stationarity when calculating flood potential related to extreme precipitation events.
2 Observed Meteorological Changes and Linkages to Flooding

2.1 Large-Area Meteorological Changes

The global surface temperature (land and ocean) has warmed almost 1°C since the start of the 20th century, and surface temperatures continue to warm, with 2015, 2016, 2017, 2018, and 2019 being the 3rd, 1st, 4th, 5th, and 2nd warmest on record, respectively (Figure 2). Spatially, the greatest warming has occurred at the higher latitudes, with some areas of higher-latitude Canada, Alaska, and parts of eastern Russia warming over 2°C, more than twice the change in the global average. However, not all parts of the globe have warmed over this period. A few small areas, including the southeastern United States, and areas of the ocean southeast of Greenland show little change or a slight cooling since the start of the 20th century.


The global increase of temperature has led to other changes in the climate system. In particular, warming has increased atmospheric water vapor concentrations. The Clausius–Clapeyron equation describes the relationship between air temperature and the saturated value of atmospheric water vapor. Saturation vapor pressure increases by ~7% for every 1°C increase in global mean surface temperature. Global observations of near-surface atmospheric water vapor show that it has increased along with the observed warming of the climate (Seneviratne et al. 2012) and is well manifested in the United States (Kunkel, Karl, et al. 2020).
Recent analyses of past extreme precipitation events in several regions of the world, including the United States, have shown the possible presence of non-stationarity resulting from either anthropogenic climate change or natural climate variability (Fischer and Knutti 2015; Janssen et al. 2014; Kunkel et al. 2010; Kunkel, Karl, et al. 2020; Min et al. 2011; Wasco and Sharma 2015; Westra et al. 2013; Zhang et al. 2013). In some cases, the extremes have exhibited significant positive trends despite decreases or no trends in mean and total annual precipitation (Kunkel, Karl, et al. 2020; Easterling et al. 2017; Goswami et al. 2006; Groisman et al. 2005). Statistically significant increases in the percentage area experiencing precipitation extremes are found for the Northern Hemisphere regions over the past half-century (Dittus et al. 2015). Also, over the past century, the change in total precipitation is less pronounced compared to the increase in extreme precipitation (Donat et al. 2016; Kunkel, Karl, et al. 2020).

It is important to distinguish between low-frequency climate variability and trend. Identification is compounded by limited available precipitation records (Matalas 2012). For example, the low-frequency components of climate variability such as El Niño–Southern Oscillation (ENSO) and decadal and multidecadal oscillations such as the Pacific Decadal Oscillation (PDO) and Atlantic Multidecadal Oscillation (AMO) can obscure longer-term climate forcings affecting the precipitation records (Salas and Obeysekera 2014). Consequently, identification of truly non-stationary precipitation patterns and their significance for the engineering analyses have been the primary challenge in adopting the concept of non-stationarity in the IDF curves. Evaluation of data with long periods of record is preferred for determining whether, for example, the observed changes in extreme precipitations are results of low-frequency climate variability with recoverable changes or climate change with relatively irreversible changes (Kumar et al. 2013; Stedinger and Griffis 2011; Villarini, Serinaldi, et al. 2009). Other than the recommendation to use data with long periods of record, limited guidance is available in frequency analysis literature on how to detect and differentiate long-term monotonic trends from the shorter quasi-cyclic climate variabilities. On the other hand, statistical and data-mining approaches are available to help quantify serial correlation at various frequencies in the time series (Harrouni and Guessoum 2009).

2.2 Regional Meteorological Changes
The non-stationarity commonly observed in extreme precipitation time series leads to an upward shift in the IDF curves, with the extent of the shift varying for different durations and return periods of the precipitation. Particularly, short-duration storms associated with convective precipitation lasting less than a day have shown the strongest non-stationarity across different parts of the world (Ban et al. 2015; Berg et al. 2013; Feng et al. 2016; Landsea et al. 2010; Lenderink and Meijgaard 2008; Loriaux et al. 2013; Prein et al. 2016). A number of studies demonstrate the extent of the shifts in the IDF curves based on past precipitation records. Rosenberg et al. (2010) evaluated the changes in annual precipitation maxima between 1956–1980 and 1981–2005 for storm duration ranging from 1 hour to 10 days in Seattle and other major cities in the region. Their results showed that the annual maximum precipitation in Seattle increased for all durations, causing the return period of a 50-year storm to decrease to approximately 8 years. Groisman et al. (2012) showed a 40% increase in the frequency of extreme precipitation events during the recent decades over the central United States. DeGaetano (2009) noted positive trends in the generalized extreme value (GEV) distribution location parameter (analogous to the mean) for the majority of the U.S. precipitation stations. Demaria et al. (2017) studied the effect of non-stationary climate on IDF curves and risk of failure of hydraulic infrastructures for two
climatologically different watersheds in the United States and found that stationary climate assumptions led to much lower estimated sub-daily intensities than those under non-stationary assumptions. Similar trends are noted by Buonomo et al. (2007) for the European region, where the greatest shift is likely to occur for precipitation with longer return periods and shorter durations.

In many areas, the increasing trends in heavy precipitation are expected to continue in the future as warmer temperatures lead to more atmospheric water vapor, resulting in the intensification of heavy precipitation for most parts of the world (Groisman et al. 2012; Kharin et al. 2013; Lehmann et al. 2015; Pfahl et al. 2017; Sillmann et al. 2013). In the United States, a clear signal of increases in heavy precipitation events has emerged (Kunkel, Karl, et al. 2020; Easterling et al. 2017). Figure 3 shows regional changes in different measures of heavy precipitation for two different time periods: 1901–2016 and 1958–2016. Averaged across the contiguous United States, numerous measures show an increase over their respective time periods (Kunkel, Karl, et al. 2020). Regionally, the largest increases have occurred in the northeastern and midwestern United States. However, a few other regions show decreases over the last several decades. Trends of extreme precipitation have mostly decreased in the southwestern United States (Kunkel, Karl, et al. 2020) for various return intervals and durations. Additionally, Hawai’i and Puerto Rico both show declines for the time periods for which they have enough data to analyze. So, the general increases in water vapor are not the sole factor in extreme precipitation trends, as changes in storm-producing weather dynamics are also important (Kunkel et al. 2012).

Precipitation is ultimately caused by specific meteorological and climatological phenomena operating at a variety of space–time scales, as classically described by Hirschboeck (1988). The resulting large-scale, regional, and local atmospheric patterns, which result in various levels of precipitation intensity and duration, have been substantiated in countless studies. For example, in addition to regional and local synoptic weather phenomena, simple global- or continental-scale analyses have demonstrated correlations between flood events and broad patterns such as the El Niño–Southern Oscillation (ENSO), the North Atlantic Oscillation, and the Pacific Decadal Oscillation, among others (D. Lee et al. 2018; Ward et al. 2010, 2014). The extent to which these patterns are likely to change in a warmer world also will affect the future intensity and frequency of extreme precipitation events.

In California and the Pacific Northwest (PNW), extreme precipitation is often associated with atmospheric rivers (ARs), which transport warm, moist air of tropical origin in narrow bands that produce heavy precipitation over the coastal mountains. Leung and Qian (2009) compared two AR events that have similar atmospheric structures but produced noticeable differences in flooding in California. They found that antecedent soil moisture and existing snowpack are important factors, so changes in these land-surface conditions, besides changes in extreme precipitation, must be considered in estimating non-stationarity effects, including changes in future flooding on the ground. For example, a particularly warm, AR-driven rain-on-snow (ROS) event in the PNW melted 35%–100% of the snowpack on the western slope of the northern and central Cascades and contributed to a flood event comparable in magnitude to the record set in the PNW (Marks et al. 1998).
2.3 Regional Hydrologic Flooding

While precipitation type is naturally an important distinguishing characteristic of different types of floods, another distinguishing characteristic is useful—specifically, pluvial (rain) versus fluvial (river) flooding. Fluvial floods are strongly dependent upon the river basin characteristics and are referred to as riverine, or streamflow, flooding. The precipitation effects on streamflow depend on multiple factors, including snow storage, the coupling of the water–energy balance, and the humidity index in the basin (Sankarasubramanian et al. 2001).

Figure 3. Observed changes (%) in heavy precipitation defined different ways for two different time periods (Source: Easterling et al. 2017).
Flood events can be broadly classified into categories based on the proximate cause of the event. As defined by R. Merz and Blöschl (2003), there are five classifications: long-rain, short-rain, flash, rain-on-snow, and snowmelt floods. In a similar, but alternate classification, Berghuijs et al. (2016) group floods into those caused by the single largest precipitation event, the single largest series of precipitation events, the single largest precipitation excess event in relation to soil moisture storage capacity, and the single largest snowmelt or rain-on-snow event.

Antecedent conditions and runoff response dynamics are important proximate causes of floods and are used to distinguish different flood types in the classification schemes. Arguably, the strongest forcing of antecedent conditions related to soil moisture and streamflow is the annual cycle. The effect of comparable amounts of precipitation on streamflow during seasons of low soil moisture appreciably differs from seasons of high soil moisture. Additionally, antecedent soil moisture conditions not only influence streamflow and runoff, but they have been shown to also affect extreme precipitation in regions where soil moisture and precipitation are strongly coupled (Koster et al. 2004), such as the Mississippi River basin. In contrast, floods can also occur during non-extreme rainfall falling on saturated soil, as in the Mississippi River basin, or by the melting of accumulated snow (see Chapter 8 and Berghuijs et al. 2016). Similarly, while topography influences rainfall patterns, such as orographic rainfall (Roe 2005), flooding is determined both by topography (e.g., channel slope) as well as land-surface cover (e.g., forested compared to urban; Jacobson 2011; Liu et al. 2006). The complexity of floods as a temporal and spatial aggregation of water over a landscape, in comparison to extreme precipitation, necessitates additional special considerations. For example, Waylen and Woo (1982) indicated that large floods occurred because of either heavy winter rainfall or spring snowmelt in the Cascade Range of the Pacific Northwest. Jarrett (1990) and Kampf and Lefsky (2016) found that the magnitude of extreme rainfall events decreased substantially and that large floods were dominated by snowmelt at high altitudes in the Colorado Front Range. England et al. (2010) also reported that the dominant flood-generating mechanism transitioned from rainfall in lower elevations to snowmelt in higher elevations in Colorado. Berghuijs et al. (2016) investigated the dominant flood-generating mechanisms across the continental United States and suggested that rain-on-snow and snowmelt events were more robust predictors of the flooding response than rainfall only over the western United States.

Non-stationarity in flood events may be generated by non-stationarity in any of the proximate or ultimate causes of floods. Changes in land-use and impervious cover are driven by human population growth and development (Lambin et al. 2001) and can also be caused by natural variability and anthropogenic climate change (Collins et al. 2013; IPCC 2014). Given the many different sources of non-stationarity, the mathematical representation of change in distribution parameters over time can take innumerable forms. For example, gradually changing phenomena, such as global warming or urbanization, could introduce a monotonic trend; low-frequency cycles from natural climate variability, such as the Atlantic Oscillation, could result in quasi-periodic behavior; and abrupt alteration from infrastructure development or catastrophic natural events, such as volcanos or hurricanes, could cause outliers or permanently change the hydrologic regime.

The many possible causes and mathematical forms of non-stationarity make identification and attribution of non-stationarity in a given streamflow record a challenge. One salient example is the Red River of the North in North Dakota, which has been extensively studied by the United States Army Corps of Engineers and academics. The presence of non-stationarity in records at this river gauge station is
strongly debated—if it does exist, it could be, but has not been definitively, attributed to any combination of the following factors: regulation, changes in land use and cover, climate change, and varying hydroclimatic regimes (Serinaldi and Kilsby, 2015). This example for only one gauge highlights the difficulty of identifying, attributing, and ultimately modeling non-stationarity for fluvial floods more broadly.

In the paleo record, non-stationarity in flood magnitude and frequency is clearly evident (Paasche and Støren 2014). For example, paleo studies of floods in the southwestern United States show that large flood events exhibit multi-century periodicity and manifest in distinct epochs. For example, Schimmelmann et al. (2003) provide evidence that California floods and other palaeoclimate records suggest that expressions of regional climates are modulated by solar variability with a ~200-year periodicity. Other natural variations occur during periods of cool and moist climate with frequent El Niño events or with global reorganizations of atmospheric circulation such as the transition from cold to warm climates (Ely et al. 1993; Schimmelmann et al. 2003). The timescale of paleo records implies that the observed non-stationarity is likely caused by significant changes in climate, indicating that anthropogenically induced climate change can have the same result, even if not yet fully evident in the historical record. Model simulations of future climate indicate that the influence of anthropogenic climate change on extreme precipitation and streamflow in some regions will not be distinguishable from natural climate variability until at least 2050, if not later (Hulme et al. 1999; Martel et al. 2018; Schleß, Steinschneider, et al. 2018). This raises the concern that by the time a significant signal from anthropogenic climate change is detected in the historical flood record and acted upon for design purposes, significant damages may have already been incurred.

Analysis of data from stream gauges without significant upstream human diversions has been completed across the United States. Since the late 19th or early 20th century, results indicate considerable spatial variability and, in some large regions, a lack of spatial coherence (Archfield et al. 2016). Nonetheless, there is a notable tendency for increasing peak flows in the northern and central Mississippi valley region and the Northeast to be consistent with increases in mean annual precipitation (Figure 4). In the United States, the increase in annual total precipitation and the extremes of daily precipitation (Kunkel, Karl, et al. 2020) are most often due to warm-season increases, with the exception of the Northeast and the Mid-Mississippi valley region. The warm season (May–Oct) is generally when soil moisture and river levels are at seasonally low levels, thus likely mitigating the effects of the increases in extreme precipitation on fluvial floods but not necessarily pluvial flooding. Interestingly, there is no strong statistical evidence that annual peak streamflow magnitude is broadly increasing in the United States. In fact, the Southwest shows a statistically significant decrease (Hirsch and Ryberg 2012), which is in line with the change in extreme precipitation in that area (Kunkel, Karl, et al. 2020); this is in contrast to the significant increase in global mean temperature, atmospheric greenhouse gases, and extreme precipitation across the rest of the United States. Although Luke et al. (2017) show that the number of rivers showing significant increases during the last 100 years is about twice the number showing decreases (8% versus 4%), most gauges show no trend. The lack of widespread positive trends in streamflow, in contrast to extreme precipitation, is also evident, even when inhomogeneous or impaired river gauges (e.g., rivers with human diversions) are included in the analysis. Villarini, Serinaldi et al. (2009) also show little in the way of significant positive trends in the annual maximum stream discharge for 50 streams with 100 years of record.
The apparent inconsistency and lack of a direct relationship between observed increases in extreme precipitation and changes in floods is likely attributable to several reasons: in many U.S. regions, floods are driven by snowmelt and rain-on-snow events (see Chapter 8) or by the cumulative effect of moderate rainfall over multiple days or weeks on saturated soil; the seasonality of floods can be different than the seasonality of trends of extreme precipitation events; observed positive trends of extreme precipitation in the United States are primarily expressed during the warm season (Kunkel, Karl, et al. 2020), when streamflow tends to be at its seasonal low. All of this suggests that positive streamflow trends are more likely to be detected within smaller basin sizes during the warm season, especially those with fewer upstream diversions where flash floods linked to extreme short-duration precipitation are prominent.

It is also noteworthy that the seasonality of the annual maximum flood has become more (or less) variable depending on changes in antecedent storage (event rainfall) in regions of highly (loosely) synchronized water and energy cycles (Ye et al. 2017). It should be noted that change points are also observed in some gauges, often corresponding to changes in gauging practice or the start of regulation (Villarini, Serinaldi, et al. 2009). Furthermore, positive magnification factors for decadal and centennial floods are observed in many regions, although gauges in pristine watersheds have the smallest magnification factors, followed by unregulated gauges, followed by regulated gauges (Vogel et al. 2011). In particular, the evidence for non-stationarity in streamflow from the historical flood records in the United States is varied and depends on the characteristics (e.g., magnitude and seasonality), type of non-stationarity (e.g., trend or change point) and level of impairment (e.g., pristine, unregulated, or regulated). This change in seasonality is also found across Europe: Blöschl et al. (2017) and Madsen et al. (2014) have found continental scale changes in the seasonality of floods in response to warmer temperatures.

In summary, there is evidence in the historic record that changes in extreme precipitation are affecting the trends in the seasonality and the magnitude of fluvial-caused floods, primarily in areas where increases in extreme precipitation are dominant year-round in the United States (e.g., the Northeast and Mid-Mississippi valley regions), but at this point, linking this to human-induced global climate change is equivocal. By contrast, the more general increase in extreme precipitation events (and pluvial-induced flooding) noted in the previous section has been linked to climate change. On the other hand, regulation or land-use change has been shown to introduce change points and resulting magnification of peak riverine floods, but not in the form of trends.
Figure 4. Relationships between century-scale trends (% per decade) in river-flow annual maximum and total annual precipitation (~1909–2008) (Source: adapted from Peterson et al. 2013).
3 Assessing and Understanding Atmospheric Drivers of Change

3.1 Climate Model Projections
Climate models are a critical tool to understand changes in flood-related variables such as precipitation. As concentrations of greenhouse gases continue to increase unabatedly, many climate model simulations that run to the end of the 21st century rely on various scenarios of greenhouse gas emissions. These scenarios are based on projections of how future greenhouse gas emissions might evolve by considering how global population, technology, and policy could unfold in this century. The scenarios result in different rates of climate change, because the physical laws governing the radiative balance of the planet are sensitive to the amount of greenhouse gases in the atmosphere. The scenarios provide information useful for assessing how these different rates of climate change will impact human and natural systems.

The global scientific community has run a set of experiments called the Coupled Model Intercomparison Project (CMIP). Dozens of different climate models have been used to simulate historical and future climate. CMIP has gone through several phases of evolution over the course of years. Originally conceived as guidance in scientific assessments of climate change to be used by policymakers, it is now a major component of the World Climate Research Programme and is entering phase 6, with some new results appearing in 2019 and 2020 and more expected in subsequent years. CMIP5 results are readily available, and they have been a cornerstone in the policy-related assessments of IPCC (2014) and the U.S. National Climate Assessments. CMIP5 uses four scenarios of emissions called Representative Concentration Pathways (RCPs) to provide scenarios of radiative forcing of the climate to the end of the 21st century (van Vuuren et al. 2011). The four RCPs (2.6, 4.5, 6.0, 8.5) provide scenarios of radiative forcing with different trajectories through this century. RCP8.5 provides a high emissions pathway, where radiative forcing reaches and exceeds 8.5 W/m² by 2100. RCP6.0 and RCP4.5 are termed “stabilization” scenarios, where radiative forcing is stabilized at 6.0 and 4.5 W/m², respectively. Finally, RCP2.6 was developed to provide a scenario where radiative forcing peaks at 2.6 W/m² and then declines before the year 2100. CMIP6 builds on these scenarios and adds new scenarios, filling in critical gaps. This includes the effect of peak and declining radiative forcing, scenarios that limit global surface warming to 2°C, and additional forcings such as land-use changes and short-lived atmospheric constituents.

The RCPs were developed using integrated assessment modeling (IAM) simulations. Each RCP (e.g., RCP6.0) can be produced using different assumptions of the trajectories of technological, socioeconomic, and policy processes that lead to various radiative forcing scenarios. IAM groups test different mixes of these processes. Since the future path of greenhouse gas emissions is strongly influenced by national policies, the forcing trajectories are heavily dependent on both today’s and unknown future policies. The uncertainty associated with which, if any, of these future pathways will mimic human actions results in one of the major uncertainties in global climate model (GCM) projections. Nonetheless, they provide insights as to what is likely to occur within any given scenario.

Figure 5, from O’Gorman and Schneider (2009), shows the future and past extreme precipitation obtained from multiple climate model simulations used in the Coupled Model Intercomparison Project phase 3 (CMIP3). The increase in the intensity of extreme precipitation is generally expected to exceed...
that of the annual mean precipitation in most regions of the world (Kharin et al. 2007; Sun et al. 2007; Sillmann et al. 2013; Westra et al. 2014). The warming climate also leads to an increase in the number of extreme precipitation events in most parts of the world (Grum et al. 2006; Mukherjee et al. 2018; Mailhot et al. 2007; Xu et al. 2018). In the United States, the number of extreme events (exceeding a 5-year return period) is projected to increase by two to three times the historical average by the end of the 21st century under the RCP8.5 scenario (Easterling et al. 2017).

Figure 5. Longitudinally averaged, multimodel median extreme precipitation (99.9th percentile daily precipitation) for the past (blue) and future (red) under the Special Report on Emission Scenarios (SRES) A1B scenario (a balanced emphasis on all energy sources). The shaded region represents the 50% interquartile range of the climate models' predictions for the past period (Source: O’Gorman and Schneider 2009).

The many different models within CMIP enable multimodel ensembles of simulations to be analyzed for identical RCPs. Ensembles based on a single model also provide a means for evaluating uncertainty. Ensembles are produced using one model, but each model simulation begins with slightly different starting conditions. These single-model ensembles enable assessment of the climate variability generated by the model.

The uncertainty of the model physics and chemistry can be assessed by comparing different model simulations to produce multimodel ensembles. Differences in model formulation (what processes are explicitly modeled and how they are resolved within the model) lead to differences among models. This provides a means to assess model formulation uncertainties.

Yet another means to assess model formulation uncertainty is to vary model parameterizations within a model. The governing parameters of the model are determined by the model resolution. Model parameterizations arise because of limited spatial and temporal resolution (e.g., tens of km and hourly time steps, respectively) within the model to represent small-scale and very fast processes. The effect of varying parameters is referred to as perturbed physics (Parker 2013). Perturbed physics ensembles can be used to evaluate uncertainty in the choice of parameterization values for sub-grid scale and rapidly evolving processes. The sensitivity of the perturbed physics within any model and the differences among models lead to variations in how much warming a model produces with a doubling of CO$_2$ over pre-industrial levels (270 versus 540 ppm), commonly called the climate sensitivity.
Simulations of the climate faithfully reproduce many important aspects of the climate. Nonetheless, at regional or local scales, simulations from global climate models vary from model to model, and at these scales they can sometimes differ appreciably from what might be expected given the observed regional or local climate. Changes at these smaller space scales are relatively more uncertain, owing to a number of factors, including the choice of RCP, the resolution of the model both in time and space, and differences in model complexity and how that complexity is addressed. When comparing to observations, interannual and decadal natural climate variability can especially confound assessments of model fidelity to the real world at these smaller scales. Nonetheless, climate models continue to be the only tool to examine impacts on the climate system of increased radiative forcing that incorporate a near-comprehensive set of relevant physical and chemical processes in a consistent manner.

It is a complex process to detect and correct for climate model errors in the simulation of heavy and extreme precipitation. This is because water vapor and temperature alone are not the only variables affecting heavy and extreme precipitation. For example, buoyancy-driven convection, which is also very important in many heavy precipitation events, is usually parameterized because the relatively coarse spatial resolution of the global model (generally ~100 or more km) means that convection usually occurs at the sub-grid scale. Even regional climate models, which typically are run at the 10–50 km resolution, do not solve the issues, since convection still needs to be parameterized rather than explicitly modeled. However, there are a few examples where regional models are run at very fine scales, and convection can be directly modeled, eliminating the need for convective parameterizations. Prein et al. (2017) present a convection-permitting simulation for North America at the 4 km grid resolution that is able to explicitly simulate deep convection and captures the main characteristics of observed mesoscale-convective systems, including their size, rain rate, forward speed, and lifetime. Even then, however, other processes are still parameterized, which is a source of model biases (Rummukainen et al. 2015). Nonetheless, progress is being made in modeling local-scale precipitation events that produce very heavy rainfall.

Since temperature changes can be directly tied to changes in radiative forcing and the saturation vapor pressure is directly linked to temperature, projections of these two quantities are among the most reliable. In other words, with continued increases in human-caused forcing of the climate system, there is high confidence that both temperature and water vapor will continue to increase as well. Based on an ensemble of climate model projections and the basic physical relationship between atmospheric water vapor and storms, a warming climate is expected to further increase the intensity and frequency of extreme precipitation events (Bador et al. 2018). The amount of atmospheric water vapor is a limiting factor in heavy precipitation events.

Climate model projections based on different scenarios of continued increases in greenhouse gases all show warming. Depending on the greenhouse gas increase scenario, by the end of this century, warming ranges from +1.5°C to 4.5°C for the globally averaged surface temperature (IPCC 2014). Warming is virtually certain if human-caused forcing of the climate continues. Virtually all model simulations show the greatest warming occurring at the highest latitudes in both hemispheres. Interestingly, the models generally show that the change in total precipitation is more spatially variable than the expected widespread increase in heavy precipitation events.
Figure 6 shows the changes in the daily maximum precipitable water (total water vapor in a vertical column of air) under a high emissions scenario (RCP8.5). The increases are very large, at least 20% and 30% over most of the mid- and high latitudes.

**Figure 6.** Projected change (%) in maximum daily precipitable water (PW\textsubscript{max}) by the late 21st century relative to late 20th century under the high (RCP8.5) emissions scenario. This is an average of 13 CMIP5 models (Source: updated from Kunkel et al. 2013).

Although direct use of model projections of heavy precipitation remains problematic owing to the issues that models have in simulating precipitation rates (Jones and Randall 2011; O’Gorman 2015), models universally show large increases in heavy precipitation events in both low and high emissions simulations for the 21st century. For example, Janssen et al. (2014) looked at the extreme precipitation from CMIP5 projections to find an overall increasing trend over the continental United States (CONUS) for the RCP4.5 and RCP8.5 projections. They divided CONUS into seven subregions, where both RCP4.5 and RCP8.5 showed increased extreme precipitation across all subregions, with RCP8.5 having a larger magnitude of increase. Other studies also show that observed increasing trends in heavy precipitation are expected to continue in the future, with likely intensification in most parts of the world (Groisman et al. 2012; Kharin et al. 2013; Lehmann et al. 2015; Pfahl et al. 2017; Sillmann et al. 2013). Similar to the observed trends, the changes in future extreme precipitations are expected to be relatively larger than the change in mean precipitation (Berg et al. 2013). These changes have significant implications on the estimation of the Intensity-Duration-Frequency (IDF) curves and their applications for engineering designs and flood management, which both traditionally relied on the stationary patterns of extreme precipitation (DeGaetano and Castellano 2017; Madsen et al. 2009; Willems et al. 2012).
3.2 Assessing the Meteorology of Extreme Precipitation Within Climate Models

As important as atmospheric water vapor is with respect to extreme precipitation, it can only be regarded as providing the potential for extreme precipitation events. The atmosphere can have very high concentrations of water vapor on many days and in many parts of the globe without any precipitation. Weather and climate systems are enablers of extreme precipitation events when ample water vapor is present. Regionally, there are various types of weather and climate types that draw upon water vapor to produce extreme precipitation events. For example, in the United States, tropical cyclones are an important weather type in the South, Southeast and Northeast, but in the West, winter storms and their associated frontal systems are key (Kunkel et al. 2012). An important challenge for effectively assessing future extreme precipitation events is to understand how different weather and climate types differentially operate on ambient atmospheric water vapor and to see how climate models simulate these effects.

To demonstrate the impact of meteorological (and climatological) systems, Kunkel, Stevens, et al. (2020) calculated the relationship between extreme precipitation events and water vapor across the United States using the annual maximum daily precipitation data from approximately 3,000 stations. The results (Figure 7) show a positive relationship between annual daily maximum precipitation amounts and the precipitable water magnitude, but there is considerable regional variability in the slope of the relationship. Kunkel, Stevens, et al. (2020) attribute the regional variability to different orography, nearness of water sources, and the types, frequency, intensity, and speed of weather and climate events drawing upon ambient atmospheric water vapor. Understanding the causes of these differing relationships is ongoing, but they point to the importance of regional differences in future projections of extreme precipitation events.
Figure 7. Correlation coefficients between the magnitude of the annual daily maximum precipitation values and the simultaneous value of the 3-hour maximum precipitable water for (a) NCEP/NCAR reanalysis for 1949-2017, (b) NCEP/NCAR reanalysis for 1980-2017, and (c) MERRA-2 reanalysis for 1980-2017. Statistically significant (0.05 level) positive slopes are depicted with triangles, while dots represent non-significant slopes (Source: Kunkel, Stevens, et al. 2020).
4 Frameworks for Developing IDF Values

The use of GCM simulations is playing an increasing role for estimating future values of IDF extreme precipitation. In fact, research and applications for estimating extreme precipitation over approximately the last 5 years can be categorized into three approaches. First, there are analyses that rely exclusively on historical observations (e.g., Ganguli and Coulibaly 2017; Cheng and Aghakouchack 2014; and the eleven volumes of NOAA Atlas 14, e.g., Bonnin et al. 2004). Second, other analyses use historical observations to calibrate or downscale GCM projections of precipitation with the use of non-stationary statistics (e.g., DeGaetano and Castellano 2017; Srivastav et al. 2014 and Lima et al. 2016). Finally, there are analyses that also use historical observations to calibrate or downscale the output from the “best” GCMs and also use non-stationary statistics (e.g., Simonovic et al. 2016; Chandra et al. 2015; and Agilan and Umamahesh 2016). This brief review will highlight the implications of each approach, as related to their ability to provide effective future IDF values. All these approaches have provided useful tools for assessing the likelihood of various types of extreme precipitation probabilities, but understanding their limitations and when to apply which technique is paramount to effective design criteria.

4.1 Exclusive Use of Recent Historical Data
Evidence suggests that the use of recent historical observations and non-stationary statistics can provide useful estimates of IDF out to about a decade or so (Ganguli and Coulibaly 2017). The advantage of using this approach is tied to at least three issues. First, at short time-horizons of a decade or less, natural climate variability can readily swamp longer-term forced changes (Easterling and Wehner 2009). Second, by focusing only on historical observations, high-time-resolution sub-daily (minutes, hours) extreme event data, which sometimes is of key importance to users, is often directly available without resorting to any loss of information through downscaling techniques. Finally, the differences between IDF statistics using stationary or non-stationary statistics for relatively short future time horizons are competing with errors related to estimating the statistical model parameters derived from the precipitation time series and the measurement errors of precipitation at high intensity, among other kinds of errors. The former has been recognized, but the latter has rarely been included in the error estimates of the precipitation IDF. The error characteristics of the measurement of precipitation data vary with precipitation rate, instrument type, precipitation type, and ambient weather conditions. In a field comparison of various precipitation gauges, Lanza et al. (2007) show that depending on the gauge used, errors and systematic biases in the range of 5% to 15% are common.

There is some evidence to suggest that the use of non-stationary versus stationary statistics has the largest impact at the sub-daily hourly durations versus the longer multi-day durations (Cheng and Aghakouchack 2014), but this needs more evaluation prior to generalizing. Additionally, caution is required for users of stationary IDF for long time horizons (e.g., NOAA Atlas 14), because differences between stationary and non-stationary methods can grow substantially as shown by Cheng and Aghakouchack (2014) and others as described in the next two subsections.

4.2 Historical Observations Calibrated to GCMs for a Simulated Future
The use of GCMs greatly enhances consideration of various types of non-stationarity that can manifest as future climate. Although these methods have added considerable value in our ability to define future IDF, the use of GCMs for future IDF does come with a few caveats. Bias corrections as well as spatial
and temporal downscaling are all required in the use of the climate model output. These are contingent on the assumption that the relationships developed to remove biases and downscale the data are time invariant, and this is questionable. An extreme case in point includes regions where surface boundary conditions change dramatically, such as a transition from lake- or sea-ice cover to no lake- or sea-ice. This change influences regional vertical profiles of temperature and humidity, which may not be well reflected in the historical record. Lanzante et al. (2018) demonstrate that discrepancies can arise in the derived relationships of model-historical observations which vary geographically, seasonally, across weather conditions, and by the types of downscaling methods used. They conclude that “whether a particular pitfall may be a serious concern depends on the details of a study’s climate data needs and sensitivities — a factor that can preclude simple one-size-fits-all guidance.” Other concerns include evidence to suggest that the relationships that are required between daily total precipitation (generally available from the models) and hourly and minute precipitation intensities (often requested by users) appear to depend on the rate of change in temperature or atmospheric water vapor (Westra et al., 2014).

The use of GCMs in the development of IDF’s has now been applied across a variety of regions. In one of the larger applications, DeGaetano and Castellano (2017) developed IDF’s using well over a hundred stations across the state of New York, dozens of climate models, and a variety of downscaling methods. Their results showed that biases were reduced most by pooling together the output from all the GCMs, but it was not readily apparent what downscaling method delivered the best results, although they did not include a weather generator downscaling method and some analyses have used this approach (Chandra et al. 2015). Other analyses (e.g., Srivastav et al. 2014, following Li et al. 2010) have focused on downscaling methods that ensure preservation of any changes in the distribution of annual maximum precipitation produced by the GCMs from current to future climate. Nonetheless, it remains unclear whether or not there is a universally best downscaling method.

In terms of identifying the best extreme value distribution, often necessary when using the output of GCMs to IDF’s, there is also ambiguity, but Lima et al. (2016) argue that the Bayesian beta distribution provides a unique advantage of being able to identify the error bounds associated with the parameters of the statistical extreme value distribution. The Bayes formulation is more complex than classical generalized extreme value (GEV) methods, so the merits would have to be considered, along with the ability to address other sources of errors (e.g. spatial downscaling, temporal downscaling, measurement, future sources of atmospheric composition). Some (e.g., Chandra et al. 2015) have argued that statistical model parameter uncertainty is more important than GCM model uncertainty, although this was based on an analysis of a single station, highlighting the quandary of assessing uncertainties.

4.3 Historical Observations Calibrated to the Best GCMs
Given the relatively large variability of future values of maximum precipitation arising from many different GCMs, a number of studies have sought to identify the “best” GCM in terms of its ability to simulate climate. One of the most popular approaches is the method of reliability ensemble averaging (REA) proposed by Giorgi and Mearns (2003), where the similarity of the GCM-simulated climate to the observed climate for selected variables is weighted along with the similarity of the future climate to those simulated by other models. The variables to be assessed are left to the analyst. For IDF applications, daily precipitation amounts have often been used (Agilan and Umamahesh 2016;
Simonovic et al. 2016). Chandra et al. (2015) provide an example for one station showing that the use of the REA method also favorably influences the uncertainty related to the selection of the statistical extreme value distribution and its parameters.

How well the GCMs produce the synoptic and mesoscale weather associated with extreme precipitation events, as opposed to specific weather or climate variables, has been an omitted aspect of identifying the best GCMs. Kunkel et al. (2012) provide an extensive analysis of the causes of extreme precipitation events in the United States. They find that different types of fronts and various storms are key drivers. The ability of a GCM to simulate these events could provide further insights as to the efficacy of future extreme precipitation as simulated by the GCMs, including appropriate weighting of the GCMs. For example, one approach is to establish relationships between extreme precipitation events and key synoptic meteorological/climatological systems known to directly affect extreme precipitation rates. Then, future estimates of extreme precipitation $P$ ($P_{fut}$) can make use of the relationships between extreme precipitation and key synoptic meteorological and climate systems for a given amount of ambient water vapor. Changes of the key causes of extreme precipitation within each region can be assessed with observations and evaluated within GCMs to estimate future values of $P$. How well current GCMs simulate the historical weather and climate systems provides a measure of confidence of the impacts of future changes. Mathematically, these relationships can be approximated in a ClimAtological effects Under Synoptic Extreme States (CAUSES) equation:

$$P_{d,f}^{fut}(x,y,t) = P_{d,f}^{NA14}(x,y)[1 + \alpha \Delta PW(x,y,t)][1 + G\sum_{s=1}^{S} B(FRT(x,y,s)\Delta FRT(x,y,s,t) + \gamma ETC(x,y,s)\Delta ETC(x,y,s,t) + \delta TC(x,y)\Delta TC(x,y,t) + \epsilon MC(x,y)\Delta MC(x,y,t))]$$  \hfill (3)

where,

$P_{d,f}^{fut}(x,y,t)$ = precipitation design value at future year $t$ for duration $d$ and frequency $f$ at location $(x,y)$;

$P_{d,f}^{NA14}(x,y)$ = current precipitation design value for duration $d$ and frequency $f$ at location $(x,y)$.

For example, for those locations within the U.S., these values can be taken from NOAA Atlas 14, which provides a baseline stationary estimate of extreme precipitation;

$FRT(x,y,s)$ = the fraction of all events at point $(x,y)$ that are caused by fronts in season $s$ for the current climate;

$ETC(x,y,s)$ = the fraction of all events at point $(x,y)$ that are caused by extratropical cyclones in season $s$ for the current climate;

$TC(x,y)$ = the fraction of all events at point $(x,y)$ that are caused by tropical cyclones in the tropical cyclone season for the current climate;

$MC(x,y)$ = the fraction of all events at point $(x,y)$ that are caused by the North American Monsoon in the monsoon season for the current climate;

$\Delta FRT(x,y,s,t)$ = the fractional change in frontal frequency in season $s$ at point $(x,y)$ at future year $t$;

$\Delta ETC(x,y,s,t)$ = the fractional change in frequency of extratropical cyclones in season $s$ at point $(x,y)$ at future year $t$;
\[ \Delta TC(x,y,t) = \text{the fractional change in frequency of landfalling tropical cyclones at point (x,y) at future year } t; \]

\[ \Delta MC(x,y,t) = \text{the fractional change in frequency of North American Monsoon moisture surges at point (x,y) at future year } t; \]

\[ \Delta PW(x,y,t) = \text{fractional change in precipitable water in future year } t \text{ at point (x,y);} \]

\[ G = \text{the function that transforms the weather system frequency changes into quantitative changes in IDF values;} \]

\[ \alpha, \beta, \gamma, \delta, \epsilon = \text{coefficients defined by the empirical relationships between the respective variable and extreme precipitation amounts in the observed data,} \]

and \( FRT(x,y,s), ETC(x,y,s), TC(x,y), \) and \( MS(x,y) \) have been standardized such that:

\[ TC(x,y) + MS(x,y) + \sum_{s=1}^{4}[FRT(x,y,s) + ETC(x,y,s)] = 1. \]

Based on current GCM capability, \( \Delta TC \) and \( \Delta MC \) are set to zero. The rationale for \( \Delta TC = 0 \) is based on a comprehensive assessment of confidence related to changes in TC climatology as the climate warms (Knutson et al. 2020). In that assessment the highest confidence was associated with TC precipitation rates scaling with water vapor increases as the climate warms. This is already included in our CAUSES equation through PW. Lowest and mixed confidence was given to changes in the frequency of TC frequency and thus \( \Delta TC \) is set to zero. Until GCMs increase resolution and reduce parameterization to resolve eye wall characteristics and clouds, it would be risky to project any change in TC frequency at regional and local space scales. Regarding the MC, future changes remain uncertain. Depending on the model, model resolution, and bias corrections, projections range from an intensified circulation with more intense rainfall to reduced early-season circulation and less rainfall (Meyer and Jin 2017; Pascale et al. 2017). As a result, \( \Delta MC \) is also set to zero. With these changes, the CAUSES equation (3) simplifies to:

\[ p_{d,f}^{\text{fut}}(x,y,t) = p_{d,f}^{\text{NA14}}(x,y)[1 + \alpha \Delta PW(x,y,t)] \left[ 1 + G[\sum_{s=1}^{4} \beta (FRT(x,y,s) \Delta FRT(x,y,s,t) + \gamma ETC(x,y,s) \Delta ETC(x,y,s,t))] \right] \]
5 Estimation of IDF Relationships in a Changing Climate

The IDF curves, which are used to characterize extreme storms for various engineering applications, will be directly impacted by the expected change in extreme precipitation. Traditionally, the curves are developed assuming no significant difference in extreme precipitation throughout the lifetime of a structure or infrastructure that may be expected to last 100 years or more. Not accounting for the expected precipitation change (or non-stationarity) in the IDF curves may pose a significant flooding risk on existing stormwater infrastructure designed to handle storms with longer return periods. For example, much of the drainage system installed in the last several decades use IDF curves that are outdated by as much as half a century, making the system potentially inadequate and vulnerable for flooding under strongly non-stationary precipitation conditions (Kessler 2011; Mailhot and Duchesne 2010; Nie et al. 2009; Rosenberg et al. 2010). Cheng and AghaKouchack (2014) and Yilmaz and Perera (2014) showed that under a non-stationarity condition, the current IDF curves can substantially underestimate precipitation extremes and thus, they may not be suitable for infrastructure design in a changing climate. Demaria et al. (2017) found that the non-stationarity in extreme precipitation may increase the risk of failure of a hydraulic structure by 25% for a 100-year return period and a project life of 100 years. Sarhadi and Soulis (2017) incorporated the impact of different nonstationary conditions on the occurrence of extreme precipitation in the Great Lakes area and confirmed the underestimation of extreme precipitation under the stationary assumption. It is thus critical for keeping the IDF curves up-to-date by incorporating recent observational and climate projections data to better represent the changes in rainfall extremes in engineering standards and codes.

Unlike the stationary frequency analysis, which is mostly governed by the extreme value theory, so far there has been no theoretical framework to characterize the non-stationary extreme events. Instead, different techniques were proposed to extend the stationary frequency analysis to nonstationary frequency analysis (Coles 2001). These techniques mostly involve 1) detecting the non-stationarity using trend analysis, 2) representing the trend in the extreme value probability distributions using covariate-dependent parameters (including time), and 3) determining average return period from non-stationary return periods to develop the IDF curves. Figure 8 shows the steps to incorporate precipitation non-stationarity into the IDF curves. An extensive review of the procedure was provided by Coles (2001) for general extreme value analyses and by Katz et al. (2002), Khaliq et al. (2006), and AghaKouchak et al. (2013) for hydrological analyses. Despite the usefulness of the above approach to address changes in extreme precipitations, users often face challenges in deciding how to detect and distinguish the different changes observed within extremes, how to incorporate the non-stationarity into the frequency analysis, and how to apply and interpret the return period concept under nonstationary conditions. In this chapter, we will review methodologies used for identifying and incorporating non-stationarity into the frequency analysis to develop or update the IDF curves and highlight some of their main practical limitations.
5.1 Detection of Non-stationarity over Other Deterministic and Stochastic Changes

In general, three types of trend analysis techniques are used to detect changes in the IDF curves. In the first approach, trend analysis is directly conducted on the extreme precipitation time series using either parametric or non-parametric (i.e., distribution-free) trend tests. The parametric test fits a linear regression line to the time series data and uses the slope of the line, with the student-t statistic for the significant test, to determine trends. The regression analysis requires that the residuals from the fitted regression line be normally distributed, an assumption not required by the non-parametric test such as the Mann–Kendall (MK). The MK test analyzes the sign of the difference between successive data points and compares the resulted sum of signs (or test statistic S) with the standard normal deviate Z-value $\left(Z_{1-a/2}\right)$ corresponding to a desired level of significance ($\alpha$). The positive and negative S values indicate increasing and decreasing trends, respectively, with the trend being significant if $|S|$ is greater than the Z-value for a given significance level. The second approach directly uses the non-stationary probability distributions, often with changing parameters, and evaluates the significance of the changes in the parameters values to determine trends. Badjana et al. (2017) applied a Bayesian trend analysis for different probability distributions with time-varying location and scale parameters to identify long-term trends in annual mean rainfall, duration, and maximum rainfall. The third approach compares the performance (or likelihood ratio) of the non-stationary distribution against the stationary distribution and evaluates the significance of the performance difference to identify the presence of a trend in the dataset. For example, Katz (2013) used the likelihood ratio to test the null hypothesis of no-trend in a
parameter value by comparing the negative log likelihood of the stationary distribution and the non-stationary distribution.

The MK approach captures monotonically increasing or decreasing trends without being affected by outliers, while the likelihood-ratio and the change-in-parameters approaches can capture other types of non-stationarities but also be affected by outliers. The MK approach is more sensitive to data gaps, which can lead to underestimating the trends. Furthermore, the likelihood-ratio and the change-in-parameters approaches have the advantage of evaluating the potential relations between the changes in extreme precipitation and large-scale climate variabilities (discussed further in section 5.3). Other less common approaches in hydrology that can be used to test the stationarity of a time series include the Kwiatkowski–Phillips–Schmidt–Shin test, Phillips–Perron test, Spearman’s Rho test, and Augmented Dickey–Fuller test. A comprehensive review of trend analysis methods used to assess hydroclimatic variables in a changing climate was provided by Teegavarapu (2018). In addition to the linear trend, it is essential to examine the precipitation time series for a potential shift or abrupt change using the change-detection or homogeneous analysis. The Pettit test and von Neumann ratio test are commonly used for identifying the change points in hydroclimatological time series (e.g., Jaiswal et al. 2015; Ganguli and Coulibaly 2017). Perreault et al. (2000) also introduced a Bayesian change-point method to investigate the sudden change in annual energy inflows for hydropower that can easily be adapted for extreme precipitation time series. Another useful statistical test for non-stationarity is the two-sample Kolmogorov–Smirnov test that can be used to assess the change in probability distributions between the past and future precipitation or any other two periods.

The trend analysis results can vary depending on what approach and metric are used and whether the analysis is done using a station’s data or regional data, highlighting the particular care that should be taken when analyzing trends. Zhang et al. (2004), for example, compared the different methods for detecting trends in the magnitude of extreme values and found that the trend analysis based on the parameters of extreme value distributions identifies more statistically significant trends. Agilan and Umapahesh (2016) observed that even if most stations do not exhibit significant trends based on the MK test, they have shown a significant trend when likelihood ratio between the non-stationary and stationary distributions was used. Depending on how extreme precipitation is defined (i.e., annual maximum precipitations [AMP] versus peaks over threshold [POT]), the trend analysis may also lead to different results. Wi et al. (2016) compared the trend analysis results using the MK test for AMP and POT time series and found different results: the POT series showed a significantly reduced number of stations with an increasing trend, and even some stations showed a decreasing trend, as compared to that of the AMP series. Determining the significance of the observed trends is another source of ambiguity in the trend analysis as this involves some level of subjectivity in setting a pre-defined probability threshold, such as the p-value. Furthermore, Serinaldi et al. (2018) showed that even if the trend analysis results indicate statistically significant trends in the data, they are mostly uninformative and do not support nonstationary frequency analysis and modeling without assuming a priori additional information on the underlying stochastic process. Finally, the data sources, either from individual stations or regionally aggregated, can affect the trend analysis results. The regional MK test (Hirsch and Slack 1984) can be used to assess trends at a regional scale. Aamir and Hassan (2018) found that even though most of the stations in a region showed no significant trends in precipitation, the regional level MK analysis showed significant trends.
In order to address the above limitations and improve the reliability of trend analysis, we recommend: 1) using long-record precipitation data that is supplemented, where possible, by well-calibrated proxy data to reconstruct the longest records with minimal time-dependent biases relative to any observed changes in precipitation; 2) using a regional trend analysis to minimize the effect of noisy records from point sources (i.e., individual stations); 3) examining the potential changes associated with the underlying causes of the trends, and 4) evaluating and choosing a method for trend analysis to ensure adequate representation of the non-stationarity that is relevant for the IDF curve development or update. The 2017 Climate Science Special Report (USGCRP 2017) also recommends distinguishing regional and station-based trends depending on the types of frequency analysis to be employed (e.g., regional or at-site frequency analyses). The ongoing progress in big data science and advances in data mining and pattern recognition methodologies, such as deep machine learning, have the potential to capture complex trends and their causal relationships with other variables. However, their applications for the IDF curves and follow-up engineering designs and management need to be carefully evaluated and justified given the difficulty and associated uncertainty in predicting such trends and functional relationships, as well as the difficulty of interpreting the resulted return periods and levels for engineering applications.

5.2 Nonstationary Frequency Analysis
A stationary extreme precipitation series is described using a fixed cumulative probability function (CDF) with constant parameters. Following the extreme value theory, the CDF is typically considered as a generalized extreme value (GEV) distribution for AMP series and generalized Pareto distributions for POT series (Coles 2001; Katz et al. 2002), although in some cases other distributions such as Log-Pearson Type III (e.g., Griffis and Stedinger 2007) have been applied. For AMP series ($X$) under a no-trend assumption, the GEV model ($M_0$) can be formalized as follows:

$$M_0: X \sim GEV(\mu, \sigma, \xi)$$

where the location ($\mu$), scale ($\sigma$), and shape ($\xi$) are considered to be constant over time.

For the non-stationary case (i.e., independent but non-identically distributed random values), the type of CDF is commonly assumed to remain the same over time, while the distribution parameters vary as a function of time or other explanatory variables (also known as covariates), which also may vary with time. Consequently, the stationary distributions can be easily extended to non-stationary distributions ($M_1$) whose parameters are not constant but change as a function of a covariate (e.g., Cheng et al. 2014; Coles 2001; Katz et al. 2002; Ouarda and El-Adlouni 2011; Renard et al. 2006, 2013; Strupczewski and Kaczmarek 2001; Sugahara et al. 2009). In most cases, time is used as a covariate to allow the distribution parameters to change as a function of time. For annual AMP series ($X$), the nonstationary GEV model ($M_1$) can be formalized as follows:

$$M_1: X \sim GEV(\mu_t, \sigma_t, \xi_t)$$

$$\mu_t = \mu_0 + \mu_1 t$$

$$\sigma_t = \exp(\sigma_0 + \sigma_1 t)$$

$$\xi_t = \xi_0 + \xi_1 t$$

Under these typical formulations, the slope parameters ($\mu_1$, $\sigma_1$, $\xi_1$) represent the trend or the level of non-stationarity. The exponent or the log function in the scale parameter equation is needed to ensure
positive values. If the non-stationary is dominantly a linear trend, using time-varying location parameter is sufficient, while considering constant values for the shape and scale parameters (e.g., Cheng and AghaKouchack 2014; Demaria et al. 2017; Liuzzo and Freni 2015; Wi et al. 2016). If the spread or variance of the AMP or POT series changes with time, the scale parameter is allowed to change as an exponential or log function of time (El-Adlouni et al. 2007; Katz 2013; Menéndez and Woodworth 2010; Salas et al. 2018; Salas and Obeysekera 2014; Sarhadi and Soulis 2017; Yilmaz and Perera 2014).

In most regions, the location and scale parameters play major roles in capturing the trends in precipitation extremes, while the shape parameter is relatively insensitive to the trends, providing evidence of a bounded upper tail in the extreme value distribution (DeGaetano 2009; Hosking et al. 1985). Khairin and Zwiers (2005) used linear time-dependent trends for the location and shape parameters and a log-linear trend for the scale parameter of the GEV distributions of global temperature and precipitation extremes. Their results showed no statistically significant difference between the GEV distributions with and without time-varying shape parameters, indicating that a linear change in the location and scale parameters were sufficient to capture the observed trends in extreme temperature and precipitation. In a similar study, Kyselý et al. (2010) have shown that making both the shape and scale parameters time-dependent did not improve the frequency analysis of temperature. Consequently, Koutsoyiannis (2004) suggested using a spatially constant shape parameter for GEV, while varying the location and scale parameters according to climatological regions and variability of the yearly extremes to characterize global extreme precipitation. This is supported by Ragulina and Reitan (2017), who have found a global average shape parameter value around 0.14, with a 95% confidence interval between 0.13 and 0.15 using daily gridded precipitation extremes. The estimation of the shape parameter is also found to be difficult, particularly when considering covariates and the relative short-record length (Coles 2001; Papalexiou and Koutsoyiannis 2013; Renard et al. 2006).

Similar to the stationary distributions, the parameters of the nonstationary distributions can be estimated using the maximum-likelihood method (Smith 1985), the generalized-maximum-likelihood method (El-Adlouni et al. 2007), and the Bayesian approach (Cheng et al. 2014; Cheng and AghaKouchack 2014). In general, the maximum-likelihood-based methods are efficient when the sample size is sufficiently large, with the generalized-maximum-likelihood method performing better than the standard maximum-likelihood method (El-Adlouni et al. 2007; Martins and Stedinger 2000). Because of the complexity of the likelihood function, the likelihood-based parameter estimations can only be obtained through numerical methods, which may lead to estimation bias when the sample size is small. To resolve the bias in the estimated location and shape parameters, Martins and Stedinger (2000) suggest fixing or using a prior distribution for the shape parameter of the GEV based on the most probable values.

5.3 Covariates for Characterizing Changes in Extreme Precipitations

Many recent studies performing the nonstationary frequency analysis allow the parameters of extreme value distributions to vary as functions of not only time but also exogenous variables or covariates. As described in Section 5.4, such analyses, in addition to estimating changes in extreme precipitation with time, can also be used to analyze whether the changes are associated with atmospheric and geographical conditions. Some also argue that incorporating climatic covariates instead of time allows for evaluation of the risk of extreme precipitation on a monthly and seasonal basis (as opposed to annually) and better incorporation of climate model projections in the frequency analysis (e.g., Tramblay
et al. 2013). Agilan and Umamahesh (2018) suggest that the use of time as a sole covariate does not necessarily lead to a better model. Instead, potential covariates relevant to a study area should be identified and evaluated in terms of their influence on extreme precipitation and its change over time. As previously discussed, extreme precipitation in a particular region is often governed by large-scale climatological conditions such as the El Niño–Southern Oscillation (ENSO), Pacific Decadal Oscillation (PDO) and North Atlantic Oscillation (NAO); alternatively, local factors such as topography, temperature and wind condition can be used as covariates. Agilan and Umamahesh (2018) provided a more comprehensive list of potential covariates and data sources. These covariates are related to the change in the parameters of non-stationary extreme value distributions (e.g., Cannon 2010; Coles 2001; El-Adlouni et al. 2007; Katz et al. 2002; Kharin and Zwiers 2005; Mailhot and Duchesne 2010; Maraun et al. 2010; Sugahara et al. 2009; Tramblay et al. 2013; Vasiliades et al. 2015).

The influence of the different covariates on extreme precipitation is expected to vary from region to region. For example, Kunkel et al. (2012) show that extreme precipitation events are linked to varying synoptic weather systems across the United States. This includes fronts, extratropical and tropical storms, mesoscale convective complexes, and other phenomena. Additionally, their impacts vary appreciably by region and season. Thus, identifying those covariates that contribute to the non-stationary patterns of the extreme precipitations at a given location is an important part of the non-stationary frequency analysis (B. Merz et al. 2012). ENSO and NAO are the two major modes of climate variability in North America that influence both the mean and extreme precipitation, primarily in winter and spring seasons (Kenyon and Hegerl 2010; Schubert et al. 2008; Whan and Zwiers 2017). Zhang et al. (2010) analyzed the influence of large-scale climate variability on winter maximum daily precipitation over North America using indices representing the ENSO, PDO, and NAO as covariates to the location and scale parameters of the GEV distribution. They have found that the ENSO condition often leads to a statistically significant increase in the winter extreme precipitations over southern North America and to a decrease in the winter extreme precipitations in the Great Plains, Great Lakes, and Ohio River valley. This finding is consistent with the one reported by Meehl et al. (2007), who found that El Niño in winter intensifies precipitation extremes along the southern United States, the eastern coast, and the Southwest. In contrast, precipitation intensity decreases in the Pacific Northwest and the Midwest during the El Niño season. Groisman et al. (2012) found that the increases in both the frequency and intensity of precipitation over the Northern Hemisphere are associated with changes in mean annual temperature, which is linked to water vapor changes.

For other parts of the world, for example, the NAO and, to a lesser extent, ENSO indices have been recognized as the leading covariates for the winter extreme precipitation anomaly in Europe (Tabari and Willems 2018). Westra et al. (2013) investigated the presence of trends in the annual maximum one-day precipitation over the globe using a non-stationary GEV distribution and found a positive correlation between the intensity of extreme precipitation and annual global temperature anomaly for about two-thirds of the stations they have analyzed. Agilan and Umamahesh (2017) used urbanization, temperature anomaly, global warming, ENSO cycle, and Indian Ocean Dipole (IOD) and their combinations to develop 62 non-stationary GEV models for characterizing the extreme precipitation in Hyderabad, India. The non-stationary GEV distribution, which uses the temperature anomaly as a covariate, outperformed the other covariates for short-duration annual maximum rainfall series, while the non-stationary GEV distribution with ENSO covariate performed better for long-duration series. In a
similar study, Tramblay et al. (2011) reviewed the potential covariates for characterizing the change in extreme storm frequencies and magnitudes in the Mediterranean region. Their results indicate that covariates related to the frequency of southern synoptic circulation patterns, air temperature, and sea level pressure provide significant improvements in estimating the occurrence and the magnitude of storm events in the region. Tramblay et al. (2013) further demonstrated that the southern circulation pattern serves as an effective covariate for the Poisson distribution parameter to estimate storm frequencies, while the monthly air temperature is effective for the generalized Pareto distribution scale parameter to estimate storm magnitudes.

Despite distinct regional differences in the response of extreme precipitation to atmospheric circulation, the effects of ENSO are widely seen throughout the world, including India, Africa, South America, North America, and, weakly, Europe (Kenyon and Hegerl 2010). Temperature and annual precipitation also showed a higher correlation with extreme precipitation in most regions. Finally, it should be noted that the choice of covariates, without a thorough evaluation of the contributions of different covariates a priori, can introduce considerable uncertainty in estimating non-stationary return levels and may compromise the advantage of using a non-stationary distribution over a stationary one (Agilan and Umamahesh 2018). Particular care is required to select the large-scale climate phenomena as covariates as some of them (e.g., ENSO and NAO) are more relevant for inter-annual changes of precipitation extremes, while others (e.g., the Atlantic Multidecadal Oscillation and Interdecadal Pacific Oscillation) are better related to decadal variability of precipitation extremes.

5.4 Functions to Relate Parameters and Covariates

As discussed in the previous section, the non-stationarity in extreme precipitations is typically represented by considering the parameters of the extreme value probability distribution to vary as a certain function of the relevant covariates. Theoretically, the functions should reflect the nature of observed trends (e.g., linear, nonlinear, or periodic) and vary from region to region to represent potential differences in the trends and associated covariates. In addition, Mentaschi et al. (2016) have shown that the choice of the function depends on the extreme value distribution chosen to describe a given precipitation series. Since there is a limitation in theoretical or process-based information to determine suitable functional relationships in most regions, rigorous analysis of several alternative functions is needed using statistical criteria that balance model complexity and fitting error to either select the best model or weighted-average suitable models. This, however, is not a standard practice in the extreme precipitation frequency analysis, where a linear function is widely used without providing a clear rationale for the selection of the function (e.g., Cheng et al. 2014; Cheng and AghaKouchack 2014; Katz et al. 2002; Yilmaz and Perera 2014). The few exceptions are Um et al. (2017), who compared 30 different nonlinear functions using statistical and information-based criteria to characterize extreme precipitation in eight U.S. cities using a non-stationary GEV. Vasilides et al. (2015) argue that a priori specification of the functional relationships between the distributions parameters and covariates are often neither appropriate nor realistic. Agilan and Umamahesh (2017) further demonstrated this limitation and have shown that without detailed examination of the nature of the underlying trend, the direct application of linear trends to the parameters may increase the bias in the non-stationary IDF curve compared to the stationary IDF curve.

Several studies have suggested the use of flexible nonlinear functions to overcome some of the limitations associated with the linear function–based non-stationary GEV (e.g., Agilan and Umamahesh
Rigby and Stasinopoulos (2005) introduced the generalized model for location, scale, and shape (GAMLSS) framework to relate parameters of general distributions with covariates using linear, non-linear, parametric, or non-parametric functions. Villarini, Smith, et al. (2009) applied the GAMLSS with cubic splines function for flood frequency analysis and obtained a better description of the variability in the mean and variance of the annual maximum peak discharge. Cannon (2010) introduced a probabilistic neural network to model the parameters of a non-stationary GEV distribution as a function of covariates. Vasiliades et al. (2015) demonstrated the flexibility of the probabilistic neural network for a non-stationary extreme value analysis of precipitation data in Greece. Agilan and Umamahesh (2017) have developed a non-stationary GEV distribution using a multi-objective genetic algorithm, which allows specifying a nonlinear relationship between the distribution parameter and covariates. Bayesian techniques were also applied to evaluate the different functional relationship between distribution parameters and time and Southern Oscillation Index covariates and their influence on non-stationary IDF curves (Sarhadi and Soulis 2017). An alternative and probably more effective approach was suggested by Mentaschi et al. (2016), which involves first transforming the non-stationary data to stationary data and modeling them using a stationary distribution and then transforming the results back into a non-stationary extreme value distribution. It is thus recommended to consider different non-stationary models through combinations of time and other covariates of the location and the scale parameters. Then, apply either the Bayesian model averaging to combine the different models or select the appropriate model based on its performance and parsimony.

5.5 Non-stationary IDF Curves

For stationary events, the probability of a particular storm event (or return level) to be exceeded is the same for all years, leading to a constant average return period, which is calculated as the inverse of the probability of exceedance. For the stationary events, there is thus a one-to-one relationship between return period and return level (Cooley 2013), while for the non-stationary events, the return period of a given storm event and the corresponding probability of exceedance vary as a function of time and/or covariates. Consequently, the return period is not only dependent on the return level but is also affected by the time used for the analysis. For non-stationary events with an increasing trend, the probability of exceedance is initially low but increases over the years. For example, Villarini, Smith, et al. (2009) have found that the return period of an annual peak discharge can range from a maximum value of more than 5,000 years in 1957 (the first year of the annual peak discharge data) to a minimum value of almost 8 years in 2007. Determining the return periods for practical engineering design is, in fact, the main challenge in adopting the non-stationary frequency analysis, as both the return periods and levels change depending on the time of the analysis; it is even more challenging if the return periods vary with other covariates.

In order to get a one-to-one relationship between return periods and return levels under a non-stationary condition for constructing the IDF curves, Cooley (2013) and Salas and Obeysekera (2014) extended the average return period concept from stationary cases to non-stationary cases. The return period is typically defined by either the average recurrent interval (ARI) for peaks over threshold extreme events or by the annual exceedance probability (AEP) for annual maximum extreme events (Cooley 2013; NOAA NWS n.d.). The AEP concept is used here to describe the return period and level under non-stationary extreme precipitations. Also, a return period is defined as the waiting time for a
given storm to be exceeded for the first time (Olsen et al. 1998). An alternative definition of the return period is provided by Parey et al. (2007) and Read and Vogel (2015).

As shown in Figure 9, the exceedance probability \( p \) for any particular (or design) storm \( z_p \) remains constant year to year for the stationary case. Assuming that the annual maximum precipitation values \( X \) are independent, the waiting time required for the storm event \( z_p \) to be exceeded for the first time is a random variable that can be described using a geometric distribution. The expected value of the waiting time, which is also known as the return period \( T \), is computed using:

\[
T = E(X) = \sum_{x=1}^{\infty} xp(1-p)^{x-1} = \frac{1}{p}
\]

(7)

Furthermore, the variance of the return period is \( Var(X) = \frac{1-p}{p^2} \). These are well-known statistics to characterize return periods for a stationary and independent annual maximum series.

![Figure 9. Schematic representation of the return levels and periods (probability of exceedances) for stationary and non-stationary frequency analysis (Source: Salas and Obeysekera 2014).](image)

Cooley (2013) and Salas and Obeysekera (2014) have extended the above relationships for non-stationary cases where, unlike for the stationary case, the probability of exceedance \( p_t \) varies with time. For an increasing trend, the probability of exceedance \( p_t \) also increases over time. Similar to the stationary case, under independent extreme precipitation time series, the waiting time required for the storm event \( z_p \) to be exceeded for the first time is a random variable that can be described using a geometric distribution. The expected time for the first exceedance (or return period \( T \)) is computed using:

\[
T = E(X) = 1 + \sum_{x=1}^{x_{\text{max}}} \prod_{t=1}^{x} (1 - p_t)
\]

(8)

The above equation can be used to estimate the return period \( T \) for non-stationary conditions with increasing trend. If the trend is negative, the return period can be calculated using:

\[
T = E(X) = \sum_{x=1}^{\infty} xp_x \prod_{t=1}^{x-1} (1 - p_t)
\]

(9)

Where \( p_t \) is exceedance probability at \( t = x \), \( x_{\text{max}} \) is the year at which \( p_x = 1 \), or the year that the design storm will definitely be exceeded. This condition is possible when we have increasing trend in annual maximum precipitation that leads to increasing probability of exceedance with a potential to reach one. 
The return periods from stationary (equation [7]) and non-stationary (equations [8] and [9]) can be directly compared to assess the impact of non-stationarity on IDF curves and associated engineering design and flooding risk. Alternatively, the average return period is computed using different quantiles to estimate the corresponding distribution parameter, $\mu(t)$ and $\sigma(t)$. For example, Cheng et al. (2014) used the median of $\mu(t)$ to determine the “effective” return level for the year corresponding to the midpoint of the historical time series. For the low-risk design, the 95 percentiles of the $\mu(t)$ values is used. The effective return period can also be used if covariates other than time are used to represent the non-stationary in the IDF curves.
6 Non-stationary Regional Frequency Analysis and Non-stationary IDAF Relationships for Extreme Precipitation

In this chapter, we review the uses of regional frequency analysis and applications of Intensity-Duration-Area-Frequency (IDAF) curves. Current methodology has assumptions of stationary behavior in the rainfall data. We will point out how impacts of non-stationary behavior will impact the methodology as standing and then suggest new practices that have been published recently that incorporate non-stationarity in some way. We conclude the discussion with recommendations on when or where these areal-based analyses may prove useful.

6.1 Non-stationary Regional Frequency Analysis

Regional frequency analysis (RFA) is based on the assumption that within a homogeneous region, extreme rainfall has an identical frequency distribution apart from a scaling factor (an index flood) (Hosking and Wallis 1997). Being able to use multiple observations is advantageous, as it allows for more accurate estimates of quantiles than at-site observations alone (e.g., Hosking and Wallis 1988; Lettenmaier et al. 1987; Lettenmaier and Potter 1985). The general outline of the regional frequency analysis includes the following:

1. Find homogeneous regions by grouping similar stations and then testing for homogeneity.
2. Pool data across the region to find the underlying identical distribution and then estimate flood quantiles.

6.1.1 Homogeneous Regions

Formation of homogeneous regions generally follows an iterative process that includes clustering or grouping stations based on their similar at-site characteristics, applying a homogeneity test, and then regrouping stations based on the results of the homogeneity test such that the stations that are most dissimilar may be put into other existing groups or be used as the basis of a new grouping. This process is iterated until all regions are found to be acceptably homogeneous (which can be defined by the user). The criteria used to create the station clusters often include site characteristics, such as latitude, longitude, elevation, and major rainfall characteristics, such as mean annual precipitation or timing of the wet season. The characteristics that define only the physical characteristics of the site are stationary and will not change with time and will likely be as useful in clustering stations in the future as they are currently. Statistics based on and meant to describe rainfall behavior, however, are subject to change under non-stationary assumptions. The manner in which these statistics change can thus affect the homogeneity of the regions in a different time period. Regions may be changing with time in a non-uniform manner, particularly in areas where human intervention is prominent (Milly et al. 2008).

Principal component analysis (PCA) is a method commonly used to identify spatial patterns. This method takes in a set of intercorrelated variables (stations located near each other have intercorrelated statistics) and returns a reduced set of variables that are linear combinations of the original variables. The reduced set of variables are the principal components and are not linearly correlated to each other, unlike the original variables. Chang and Chen (2016) suggest replacing PCA with geographically weighted principal component analyses (GW-PCA) for a better explanation of total variance in a more
heterogeneous space. This approach may allow for one to get more information from a homogeneous region based on stationary assumptions.

We put forward two suggestions to further research in the face of non-stationary region formation. First, one may consider doing some measurement of sensitivity of at-site characteristics (either with time or temperature) into the initial clustering analysis. This creates a better understanding of which regions are most vulnerable to needing reassessment in the future. A second suggestion is to include in the clustering stations a characteristic that more directly represents the source of or mechanism behind precipitation at each station. This could look like grouping together stations with the same source of atmospheric moisture or with the same dominating meteorological mechanism in creating extreme precipitation (see e.g., Wallis et al. 2007). By including a more direct representation of the source and mechanism behind extreme precipitation, we hope to increase the ability to capture the changes in both across a region as well.

Some authors, including Hosking and Wallis (1997), Guttman (1993), and Schaefer (1990), note that homogeneous regions do not need to be physically contiguous. Enforcing physically contiguous regions can increase the measured heterogeneity by including stations that are nearby but not necessarily homogeneous and excluding stations that would be homogeneous but are not within a certain physical range. However, there is an implicit assumption that stations that have an identical underlying distribution, apart from the flood index, have rainfall patterns that are driven by the same physical mechanisms, and that as the climate shifts, the drivers of change will occur in a physically contiguous manner. Physically contiguous regions are also more commonly seen when regional analysis is applied to a large area, such as CONUS (see, for example, the NOAA Atlas 14 regions). Deciding whether or not to enforce physical continuity in homogeneous regions comes with its own sets of possible advantages and disadvantages.

### 6.1.2 Regional Flood Curves

Once homogeneous regions are established, the data across the region is pooled and used for finding an underlying distribution. There are many ways to fit a probability distribution to the data, estimate the parameters of the distribution, and then test for goodness of fit. Common distributions used in extreme frequency analysis include the three-parameter generalized extreme value (GEV) distribution, which includes the Gumbel distribution when the shape parameter equals zero, the Log Pearson III distribution, and the Pareto distribution. In this chapter, we will use the GEV distribution as our example distribution.

The index flood method is based on the assumption that all sites within a homogeneous region have an identical distribution, save for a site-specific scaling factor or index flood. The index flood is often the mean or median extreme event specific to the station but can be any type of location parameter of the site-specific distribution (Hosking and Wallis 1997). The underlying distribution common to the region is represented by the unitless regional growth curve, where \( Q_i(F) \) is the at-site quantile function, \( \mu_i \) is the index flood, and \( q(F) \) is the regional growth curve. To incorporate the effects of non-stationarity into the regional growth curve, we add a time-index into the location parameter (\( \mu \)) or into the location and scale (\( \mu, \sigma \)) parameters. The result is a non-stationary regional growth curve analogous to a non-stationary IDF curve, where a time index influences the resulting estimated quantiles where \( \mu(t), \sigma(t) \)
are time-dependent location and scale parameters for the region and $K_R$ is the time-independent shape parameter for the region.

$$Q_i(F) = \mu_i \ast q(F)$$

$$Q_i(F) = Q(x; \mu_i, \sigma_i, \kappa_i) = \mu_i + \sigma_i ((-\log(x))^{-\kappa_i} - 1)/\kappa_i$$ \hspace{1cm} (10)

where $\mu_i, \sigma_i, \kappa_i$ are location, scale, and shape parameters at site $i$, and

$$q(F) = q(x; \mu_R, \sigma_R, \kappa_R) = \mu_R + \sigma_R ((-\log(x))^{-\kappa_R} - 1)/\kappa_R$$ \hspace{1cm} (11)

where $\mu_R, \sigma_R, \kappa_R$ are location, scale, and shape parameters for the region.

$$\hat{\theta}_k^R = \sum_{i=1}^{N} n_i \hat{\theta}_k^{(i)} / \sum_{i=1}^{N} n_i$$ \hspace{1cm} (13)

where $\hat{\theta}_k^R$ is the estimated regional parameter, $n_i$ is the length of record for site, $i$, and $\hat{\theta}_k^{(i)}$ is the estimated at-site parameter.

$$\mu(t) = \mu_1 \ast t + \mu_0$$

$$\sigma(t) = \sigma_1 \ast t + \sigma_0$$

$$q(F) = q(x; \mu_R(t), \sigma_R(t), \kappa_R) = \mu_R(t) + \sigma_R(t)((-\log(x))^{-\kappa_R} - 1)/\kappa_R$$ \hspace{1cm} (14)

where $\mu_R(t), \sigma_R(t)$ are time-dependent location and scale parameters for the region and $K_R$ is the time-independent shape parameter for the region.

### 6.1.3 Non-stationary Trend Detection at the Regional Level

It is most likely that some stations will test significantly for non-stationary behavior before other stations within the same region. To accommodate for this, one could allow for mixed behavior, meaning allowing some stations to be modeled as non-stationary and others as stationary depending on some criteria or test for significant non-stationary behavior within a region. However, one of the benefits of regionalization is the ability to reduce the noise around a signal by using regional averaging. Individual stations will need to have a much clearer signal in order to test for significant nonstationary behavior as compared to regionalized data. In this way, allowing for mixed behavior defeats the purpose of using regionalization analysis. Further, allowing mixed behavior is a cautious approach in incorporating non-stationarity, as its effects will be diluted at the regional level. We suggest, instead, testing non-stationarity at the regional level in such a manner that if the region has significant non-stationary behavior, it is assumed that all stations within the region should be modeled as non-stationary as well.

Conversely, if the region does not test for significant non-stationarity, then we proceed to treat all stations within the region as stationary.

Cunderlik and Burn (2003) use a regional bootstrap re-sampling technique to test for the presence of significant non-stationary behavior at the regional level, where non-stationarity is allowed in the first and second moment. They first determine how many stations within a region test for significant non-stationarity (see Chapter 5 for trend detection methods at the station level). They then pair a regional bootstrap re-sampling technique along with Monte Carlo simulations to determine if the number of non-stationary stations is significant enough for the region to be considered non-stationary as a whole.
The results of trend detection at the station level are affected by cross-correlation, which is preserved in the regional bootstrap re-sampling technique. Livezey and Chen (1983) use Monte Carlo simulation to establish field significance for collective set of statistics. Cunderlik and Ouarda (2006) use a regional trend analysis to estimate the non-stationary first and second moments of a region. They use the index flood method but replace three components with a time-dependent analog of the stationary model.

6.1.4 Application to Station Data
Cunderlik and Ouarda (2006) applied the non-stationary index flood method to a study area in the South British Columbia Mountains climate region to fit the data to a regional GEV distribution with a non-stationary location and non-stationary scale parameter. They tested at each station and then across each region in the study area for the presence of significant trends. Regions that were significantly (alpha at 10%) non-stationary then got fitted with a non-stationary GEV distribution, and a new non-stationary regional growth curve was created and compared to the stationary one. They found that the non-stationary distribution reduced model bias and root mean square error (RMSE). Tramblay et al. (2013) used the POT instead of AMS method and fit a non-stationary Poisson distribution to a regional sampling of stations over southern France to find that non-stationary regional quantile estimates provide a better fit than the stationary model. Leclerc and Ouarda (2007) applied non-stationary regional frequency analysis to a study area in the northeast United States. They used a non-stationary GEV distribution to estimate quantiles representing the 5-year and 100-year flood. They found that while the stationary and non-stationary model were both positively biased, the non-stationary model had smaller bias and RMSE values than the stationary model.

6.1.5 Application to Gridded Data
Kharin and Zwiers (2005) fit a non-stationary regional GEV distribution to three different GCM simulations. They allowed for a time-dependent location and scale parameters, as estimated with the method of maximum likelihood. They found that the method of maximum likelihood is preferred over stationary L-moments and over an approximated non-stationary method where the change in time-dependent parameters is estimated via moving windows. Gilroy and McCuen (2012) also applied a regional analysis to GCM simulations following a procedure similar to one set by Kharin and Zwiers (2005). Both papers found an exponential function instead of a linear function for relating location to time as well as for relating scale to time. They compared the GCM distributions to the observed precipitation from the same area to find the reduction factors needed to make the gridded areal precipitation comparable to a point location.

6.2 Updated IDAF Curves
IDAF curves give estimates of areal rainfall. When the area component of the IDAF approaches zero, the rainfall is point rainfall and is represented by an IDF curve instead. Thus, IDAF is the natural areal extension of the IDF concept and depends on both a temporal scale (duration) and spatial scale (area). With properties from simple scaling (scale invariance) applied to space and dynamical scaling applied to time, the IDAF curves can model extreme rainfall distributions across varying durations and areas. The areal reduction factor (ARF) is the ratio that relates point rainfall to areal rainfall. By introducing non-stationarity into the IDAF curve, a time element must be included for both components.

The two components to the IDAF curve are 1) the Intensity-Duration-Frequency curve and 2) the areal reduction factor.
\[ \text{IDAF} = \text{IDF} \times \text{ARF} \]
\[ \text{IDAF}(t) = \text{IDF}(t) \times \text{ARF} \]

### 6.2.1 IDF Curve

See Chapter 5 for information on how non-stationary behavior affects IDF curves, as well as recent work on the topic.

### 6.2.2 Areal Reduction Factors

The ARF is the ratio of areal rainfall to point rainfall. A commonly used formulation was established by the U.S. Weather Bureau in Technical Paper 29 (U.S. Weather Bureau 1957). Allen and DeGaetano (2005) found this equation to be adequate for areas under 1,000 km² and reviewed other ARF formulations for areas over 1,000 km², where the simplistic TP-29 definition may not be the best suited. More contemporary definitions of the ARF take advantage of the multi-scaling properties of rainfall (De Michele et al. 2001) to create a scaling relationship between two different sizes of areas with the same central location. While other definitions based on non-scaling rainfall exist, such work has become less popular since the 2000s. The ARF definition presented by De Michele et al. (2001) utilizes the simple scaling and multi-scaling properties of rainfall in their construction:

\[ \text{ARF}_{TP-29} = \frac{1}{n} \frac{\sum_{i=1}^{n} \tilde{R}_j}{k \sum_{j=1}^{n} \left( \frac{1}{k \sum_{j=1}^{n} R_{ij}} \right) \sum_{j=1}^{n} R_{ij}} \]

where \( n \) is the number of years on record, \( \tilde{R}_j \) is the annual maximum areal rainfall for year \( j \), \( k \) is the number of stations in the area, and \( R_{ij} \) is the annual maximum point rainfall at site \( i \) for year \( j \).

While there is no work currently published on the non-stationary modeling of ARFs, it is likely that research into this area can proceed in a manner similar to constructing stationary ARFs. One can create a rain field across time (duration) and space aggregations at various years and compare the ratio of the areal rainfall to the point rainfall but apply this approach on a year-by-year, or window-by-window, basis to get a general idea of how a particular ARF may change with time. The non-stationary ARF should effectively represent the rate of non-stationarity at the areal level versus point level. In a carefully homogenized region, where the changes in precipitation mechanism are the same across the region, there may be little change to the ARF with time if the change in areal rainfall and the change in point rainfall are similar.

### 6.3 Path Forward

There is still a considerable amount of work to be done on the topic of non-stationary regional frequency analysis and non-stationary IDAF curves, yet decisions are needed today for IDAF estimates. Here, we summarize our final suggestions, comments, and remarks on complications and identify areas of research that need development.

For non-stationary regional frequency analysis, we recommend 1) creating homogeneous regions that remain homogeneous with time by identifying areas undergoing similar precipitation mechanisms and similar changes in precipitation mechanisms; 2) testing for the significance of a non-stationary trend at the regional level; and 3) incorporating a time index into the formulation of the regional growth curve.
For non-stationary IDAF curves, we recommend examining how non-stationarity affects the IDF curve (see Chapter 5) and the ARF. Currently, there is no established method for creating a non-stationary ARF or overall IDAF curve. A non-stationary IDAF curve requires that simple scaling in time and dynamical scaling over space must still hold true, along with the underlying assumptions of the scaling.
7 Strategies to Address Non-stationarity in Hydrologic Extremes

Vexing challenges of incorporating non-stationarity into extreme precipitation IDF's and flood design values is the focus of this chapter. Though related, non-stationarity in coastal flooding is beyond the scope of this Guide; note, too, that this chapter is not intended to be directly applicable to flash floods, urban flooding, or wetland floods that occur from short-duration extreme precipitation-related events that cause landslides, soil erosion, and other harmful impacts. The chapter does include a discussion of the weaknesses and strengths of various methods to guide the user in terms of best practices. The chapter opens with non-stationarity estimates using general first-principles estimates of future floods. Subsequently, the chapter describes the two prominent approaches of flood projections under non-stationarity: hydrologic simulation and climate informed (Figure 10). Where applicable, a summary of expected changes in floods based on each method is also provided. Naturally, there are many hybrid approaches that mix physically based models and statistical methods; however, only widely accepted approaches are described here. Generalized schematics for each approach are shown in Figure 10. Finally, this chapter describes methods for calculating non-stationary design values. Additional reviews and summaries of the topics covered in this chapter can also be found elsewhere (François et al. 2019; Salas et al. 2012, 2018).

7.1 Estimates from First Principles
A simple estimate of future riverine floods can be developed from first principles. According to the Clausius–Clapeyron equation, anthropogenic global warming is expected to lead to increased moisture-holding capacity of the atmosphere, which, without other changes, leads to increased precipitation extremes and, potentially, to greater floods. Additionally, increased temperatures will cause more precipitation to fall as rain rather than snow and will modify the timing of snowmelt, which can change both the seasonality and magnitude of floods. Apart from climate change, natural climate variability
operating at decadal to multidecadal scales leads to cycles of flood magnitude and frequency. Human development, such as impervious surfaces from expanding urbanization or land-clearing for agriculture or industry, will lead to a faster rainfall-runoff response, increasing the potential for flash floods. However, simple first-order estimates are not sufficient for design, due to the complex coupling and feedbacks between atmospheric and land-surface forcing. For example, the first-order effects of increased temperatures on extreme precipitation, and hence floods, are modulated by additional mechanisms present in weather and climate systems and modulated by orography, including moist-adiabatic temperature lapse rate, upward velocity, and localized temperature (O’Gorman and Schneider 2009; Trenberth 2011; Kunkel, Stevens, et al., 2020). Similarly, urbanization not only increases impervious surfaces in the immediate locality but also incentivizes increased consumption, leading to a much larger ecological footprint than simply the urban boundaries (Lambin et al. 2001, 2003). These challenges have led to the development of physically based models and statistical methods for flood projection.

### 7.2 Projection by the Hydrologic Simulation Approach

The key idea of projecting floods using hydrologic simulation is to mathematically represent the mechanisms governing the movement of water. This approach is also commonly referred to as the top-down or model-chain approach and has been widely implemented at various levels of complexity and spatial scales around the globe. Studies using the hydrologic simulation approach that account for both climate change and direct anthropogenic forcing of riverine flood events (e.g., Gilroy and McCuen 2012; Huong and Pathirana 2013) are less prominent in the literature than those focusing solely on climate change impacts. This likely occurs because multiple sources of non-stationarity introduce additional modeling complexity and because anthropogenic forcing is often used when the interest is primarily flood damages (e.g., Winsemius et al. 2016) rather than riverine flood events. Yet, observed non-stationarity is more prominent in impaired gauges than in reference gauges, as discussed previously (Villarini, Serinaldi, et al. 2009; Vogel et al. 2011). Thus, this limited focus on other sources of non-stationarity besides climate change engenders a limited perspective of future riverine flood change.

The more common implementation of the hydrologic simulation approach focuses solely on forcings from climate. In climate-focused studies, projections of climate variables, usually precipitation and temperature, from an ensemble of GCMs are used to force a hydrologic or land-surface model; subsequently, riverine flood events are identified within the resulting streamflow time series. Studies following this approach at the global scale, using the most recent generation of GCM projections, find that large-scale patterns of flood change are generally consistent across models (Dankers et al. 2014; Hirabayashi et al. 2013). In particular, there are projected increases (decreases) in the 30-year, five-day average peak flows and the 100-year flood for regions of Siberia and Southeast Asia (regions of Europe and North America). In contrast, at the individual basin scale, the same studies find large variations in the magnitude and sign of change between models. Studies following this approach for continental, regional, or local scales are too numerous to summarize here; however, no study on climate impacts specific to floods for the entire United States was identified.

The advantage of using the hydrologic simulation approach is that feedbacks between coupled systems can be represented to the fullest extent possible given the current state of knowledge. However, there are many concerns and challenges associated with this approach. In general, even the most complex models do not fully represent the underlying physical processes (e.g., GCMs poorly simulate localized
rainfall, and hydrologic models poorly simulate peak flows), and they often have high computational demands and lead to increasing uncertainty propagation. For studies assessing climate change impacts, the shortcomings of GCM simulation of rainfall, which is a primary driver of flood events, is a prominent concern. At the global scale, most GCMs over- (under-) estimate precipitation over complex topographic (arid) regions (Mehran et al. 2014). Over the contiguous United States, GCMs reproduce large-scale patterns but vary widely at regional scales (Sheffield, Camargo, et al. 2013). This behavior occurs not only because of the limited ability of GCMs to resolve fine-scale features and processes (Kundzewicz and Stakhiv 2010) but also because GCMs often do not preserve the teleconnections that drive precipitation (Langenbrunner and Neelin 2013; Y. Y. Lee and Black 2013; Polade et al. 2013; Sheffield, Barrett, et al. 2013). These issues are aggravated for precipitation extremes; GCM bias is higher (Mehran et al. 2014), and observed trends are underestimated and vary widely among models (Wuebbles et al. 2014). To address these issues, a suite of statistical and dynamical downscaling and bias-correction techniques have been developed that increase the spatial resolution of GCM outputs and decrease bias relative to observations. However, as described in previous chapters, these methods also have their own shortcomings, adding more uncertainty to projections; for example, bias correction can alter the relationship between spatiotemporal variables and violate conservation principles, while downscaling is strongly dependent on the forcing from the GCM (Ehret et al. 2012; Fowler et al. 2007). Thus, while the hydrologic simulation approach should theoretically provide the best projections, in practice, our limited knowledge of and ability to reproduce the complexity of the hydrologic cycle hinders confidence in using projections from this approach to guide design values.

7.3 Projection by the Climate-Informed Approach (Statistical Methods)

The key idea of projecting floods using statistical methods is to condition the parameters of the flood distribution on time-varying predictors using the historic record and then to use projections of the parameter to force changes in the distribution. In its simplest implementation, the location parameter of the extreme value distribution (i.e., the mean) could be regressed on time; if the regression is linear, this could ostensibly represent a monotonic trend forced by climate change or urbanization. However, a simple regression on time is not recommended both because it is abstracted from any physical mechanism driving the change and because the past does not necessarily represent the future (Blöschl and Montanari 2010; Jain and Lall 2001). For example, there is no reason why the forcings that dictated climate change of the past will follow a monotonic linear future trend. Additionally, a linear trend in particular does not capture low-frequency (i.e., multidecadal) variability; in fact, if too short a historical record is analyzed, it may cause such variability to be confounded with longer-term changes. Luke et al. (2017), in a study of stream gauges across the United States with at least 60 years of data, convincingly show that extrapolation of an observed trend in the first half of the record rarely improves projections for the second half of the record. Thus, regression on time alone is not recommended as a statistical method for flood projection.

A better approach is to condition the parameters of the flood distribution on time-varying predictors that represent a physical forcing mechanism. Example physical forcing mechanisms include population, an index that tracks reservoir development, precipitation as rainfall or snow, and large-scale oceanic or atmospheric patterns (e.g., Delgado et al. 2014; Kwon et al. 2008; López and Francés 2013; Villarini, Smith, et al. 2009, among others). To date, such methods have been applied only to specific regions or localities; thus, global- or continental-scale projections of floods are not available. Advantages of this
method are that the simplicity of a statistical model allows easy identification of which sources of non-stationarity contribute the greatest uncertainty. Both anthropogenic and natural forcings can be incorporated as predictors based on informed projections (see Figure 10) of both climate and non-climate changes. Furthermore, only the predictand (i.e., flood events) is simulated, rather than the full streamflow time series. The disadvantages of this method are that 1) important predictors or feedback mechanisms between predictors can be difficult to identify or may be inadvertently left out of the model and 2) the identified relationship between riverine floods and predictors used in the model may themselves be non-stationary. For example, a sufficient change in climate could cause current storm tracks to change their frequency and intensity, modifying the original physical mechanisms used to develop the statistical model.

Recently, the subset of statistical methods that use large-scale oceanic or atmospheric patterns, termed climate-informed approaches, have gained interest in the academic community and have been formalized into a general methodology (Kwon et al. 2008; Schlef, François, et al. 2018). The motivation for a climate-informed approach is that larger-scale patterns are more skillfully simulated by GCMs as compared to localized temperature and precipitation (Fuentes-Franco et al. 2016; Ning and Bradley 2016; Sheffield, Camargo, et al. 2013). Thus, climate-informed approaches attempt to blend statistical methods with the information from physically based GCMs that is both credible in reference to the historic record and robust under climate change. In summary, while statistical methods, and particularly climate-informed approaches, are still relatively new compared to the hydrologic simulation approach, they are proving to be a viable alternative for riverine flood projections under non-stationarity.

7.4 Calculation of Non-stationary Design Flood

The hydrologic simulation approach provides a time series output of streamflow, which can be post-processed to identify flood magnitudes. In contrast, a statistical model conditioned on predictors will provide a flood distribution for each time step. Regardless, given the results from either method, the next challenge is to develop a value suitable for design. Traditionally, under the assumption of stationarity, design is based on the flood magnitude associated with a particular frequency of occurrence (e.g., the 100-year flood that occurs, on average, once every 100 years and has a 1% probability of occurring in a given year), which can be calculated from the flood distribution. However, under non-stationarity, this definition, which is based on the geometric distribution, is no longer valid.

A common solution in most of the non-stationary flood projection literature is to calculate a return period from projected flood magnitudes by assuming stationarity within pre-specified time periods. For projections from the hydrologic simulation approach, this can be calculated directly (e.g., Hirabayashi et al. 2013). In contrast, projections from statistical methods provide a distribution for each time period that can be post-processed (e.g., Delgado et al. 2014; Schlef, François, et al. 2018). The limitations of assuming stationarity within a time period are twofold: 1) it violates the motivating assumption of non-stationarity, and 2) given that most projections extend to 2100 at most, it imposes a relatively short record from which to fit a distribution (e.g., on the order of 30 to 60 years, as in Hirabayashi et al. 2013 or Schlef, François, et al. 2018, respectively), biasing the upper tail of the distribution.

An alternative approach to calculating non-stationary return periods is to use the non-homogeneous geometric distribution, where the non-stationary return period can be defined either as the waiting time until an exceedance occurs (Salas and Obeysekera 2014) or the time length for which the probability of
exceedance is one (Parey et al. 2007). A slightly different definition is given by Rootzén and Katz (2013), who argue for the use of “design life level,” which is the probability of the event occurring within the planning horizon. In contrast, Read and Vogel (2015) advocate that return periods should be replaced with the concept of reliability, defined as the probability that an exceedance event will not occur over the planning period, which provides information about expected performance, rather than expected failure. Thus, calculating a non-stationary metric for design remains an open question, especially given the computational challenges associated with some formulations.

7.5 Recommended Practices for Riverine Flooding
Riverine floods are caused by precipitation, antecedent conditions, and runoff response dynamics. These proximate causes, driven by climate variability, climate change, and human development, are potential sources of non-stationarity. However, evidence for the non-stationarity in the historic streamflow record is difficult to identify, inconclusive, and, where observed, not readily attributable to climate change. In contrast, evidence for non-stationarity in paleo river-related records is clearly evident and, given the long timescales, attributable to climate change. Future changes in riverine flood events can be simply estimated from first principles; for more quantitative projection of floods under non-stationarity, there are two primary approaches: hydrologic simulation and statistical methods (particularly, climate-informed methods). Hydrologic simulation represents the complex processes governing floods to the fullest extent possible given current knowledge and computational techniques. Yet hydrologic simulation has multiple limitations, including poor simulation of local precipitation in GCMs and of peak flows in hydrologic models. Statistical methods, particularly climate-informed ones, directly project future floods based on predictors, which, in the case of large-scale oceanic or atmospheric patterns, have the potential to be more skillfully projected than local forcing mechanisms, such as precipitation. However, climate-informed approaches may omit important predictors or feedback mechanisms and assume that the identified relationship between floods and predictors is stationary. Finally, given future riverine flood projections, there are multiple methods for calculating design values, depending on the type of projection and purpose of design.

Given the mixed evidence for non-stationarity in the recent historic flood record (i.e., the last 100 to 200 years, at most) and the challenges associated with developing design values accounting for non-stationarity, it is readily apparent why the assumption of stationarity, possibly updated to reflect the most recent 30 years, is still recommended by some researchers (England et al. 2018; Luke et al. 2017). In fact, the literature contains a sharp theoretical debate about assuming (or not assuming) stationarity (Koutsoyiannis and Montanari 2015; Lins and Cohn 2011; Matalas 2012; Milly et al. 2008, 2015; Montanari and Koutsoyiannis 2014; Serinaldi and Kilsby 2015). Our belief, founded on theoretical and data-based considerations, is that non-stationarity cannot be ignored. Therefore, we recommend that non-stationary flood projections be developed using statistical methods, in particular those that are climate informed. In cases where adequate predictors cannot be identified, then the hydrologic simulation approach should be used with caution. Furthermore, we recommend calculating a design value based on the assumption of stationarity within a specified time period, realizing that this is not ideal and that further research is required. Finally, in all cases, we recommend a thorough assessment of model credibility (i.e., whether the model adequately simulates observations and incorporates all primary driving mechanisms), robustness (i.e., whether the model formulation and predictors will
remain valid under future change), and the accompanying uncertainty (i.e., the degree of confidence in projections).
8 Rain-on-Snow/Snowmelt Events and Their Estimation

Snow-dominated regions can be particularly sensitive to non-stationary climate due to the impact of changes in both precipitation and temperature. A relatively minor increase in temperature alone can shift the dominant peak runoff mechanism from snowmelt to rain-on-snow, or even to a rain-dominated regime. Furthermore, a systematic and consistent hydrologic design approach explicitly accounting for snow processes is lacking in many snow-dominated locations.

Unlike the relatively even seasonal distribution of precipitation in the eastern United States, most precipitation in the western United States falls in late fall, winter, or early spring, and much of it is temporarily stored as snowpack in high-altitude mountainous regions. Snowmelt from mountain snowpack contributes about 70%–80% of the total annual runoff in this region (Bales et al. 2006; Clark et al. 2001; Fassnacht et al. 2003). Although extreme precipitation over a short time period is often the cause of extreme flood events, a great number of large flood events are associated with snowmelt from deep snowpack, especially during rain-on-snow (ROS) events (Bookhagen and Burbank 2010; Fang et al. 2014; Kattelmann 1997). McCabe et al. (2007) have shown that, historically, with warming in the western United States, the occurrence of ROS events has declined at the lowest sites and has increased at higher sites. The decreasing number of low-altitude ROS events reflects declining periods of time when snow is on the ground and available to be rained upon. The increasing frequency of ROS events at higher altitudes probably reflects increasing tendencies toward storms with unusually high snowlines and unusually broad rainy areas in many parts of the West (e.g., Knowles et al. 2006). This result is further confirmed by Yan et al. (2018), who found that the frequency of ROS events showed a statistically significant increase in the mountainous regions of the northwestern United States. The Pacific Northwest is especially sensitive to a warming climate, because at these sites near the mean freezing level, a moderate change in winter temperature can strongly reduce the fraction of precipitation that falls as snow (Mote 2003; Safeeq et al. 2016).

Based on the discussion above, extreme precipitation alone may not be the best predictor of flood response in many snow-dominated regions. This is particularly true where extreme precipitation falls as winter snow or where it falls as rain on existing snowpack. On the other hand, if peak flows result from summer thunderstorms, extreme precipitation may be a viable predictor even if these locations have consistent winter snowpack. Thus, we suggest that available water for runoff (AWR) at the ground surface may be a more robust predictor. AWR is defined through mass balance for a selected time interval as:

\[
AWR = (\text{PREC} - S) - \Delta \text{SWE}
\]

(17)

where PREC is atmospheric precipitation, S is equal to condensation (negative) or evaporation/sublimation (positive), and ΔSWE is the change in ground snowpack water content (Yan et al. 2018). A time series of AWR can be estimated from observations or through hydrologic simulation and can be subject to the sample types of statistical analysis presented in Chapter 4.

This chapter discusses current approaches and possible improvements for hydrologic design in these regions. We begin with an overview of snowmelt runoff processes, followed by a discussion of current design practice using both physics-based, continuous simulation models and standard IDF-based design. The chapter concludes with design recommendations and a potential path forward.
8.1 Current Practice
A systematic, coordinated, and consistent approach for local surface water design manuals is not available for snow-dominated regions of the United States. Local design approaches vary from the blind approach of simply using precipitation-based Intensity-Duration-Frequency (PREC-IDF) curves (e.g., Snohomish County Washington Government 2016); to the tuning factor approach of adding a snowmelt factor to the PREC-IDF values (e.g., Washington State Department of Ecology 2004); and, for large, high-cost, and high-risk infrastructure design, sophisticated physics-based hydrologic modeling (FAA 2013). Both the blind and tuning factor approaches are based on traditional IDF-based hydrologic design that implicitly assumes precipitation is in the form of rainfall that is immediately subject to rainfall-runoff processes. This assumption has potentially significant implications in regions where snowfall is a major component of precipitation. If an extreme precipitation event is primarily snowfall, much of the precipitation may not be immediately available for the rainfall-to-runoff process, which can lead to overdesign and incur unnecessary cost. On the other hand, under-design will occur if snowmelt or ROS rates exceed extreme precipitation, potentially leading to significant underestimates of flood risk.

Despite scientific advances, a large portion of infrastructure in snow-dominated regions is still designed using traditional IDF technology, including the NOAA PREC-IDF curves (NOAA Atlas 14). It is a common and standard practice for engineers to use a single-event-based model (e.g., Cronshey et al. 1986) for modeling the rainfall-runoff process. The use of an advanced physically based modeling approach can be cost-prohibitive in the design of smaller infrastructure, such as highway culverts or residential storm water systems. Another critical constraint is adherence to local surface water design manuals, which may recommend or require the use of PREC-IDF curves with selected single-event hydrologic models. It has been more than 50 years since the IDF curve technology was first introduced in hydrology. Over the past decades, the use of physically based hydrologic models for predicting hydrologic responses has become more common. However, because of the high cost of the advanced and sophisticated physically based models (i.e., labor, staff expertise, and computational demand) and local regulation as discussed above, IDF-based technology will likely continue to play an important role in hydrologic design through the foreseeable future, especially for the design of small-scale infrastructure. For a more complete discussion of design risk, see Chapter 9.

8.2 Physics-Based Models
The frequency, intensity, and duration of snowmelt and/or ROS events respond to local changes in precipitation, air temperature, incoming short- and long-wave radiation, wind, and vapor pressure, as well as antecedent snowpack conditions including pack-surface albedo, temperature, and snow water equivalent (SWE). Continuous, physics-based distributed hydrologic models use detailed mass and energy-balance algorithms to track these processes through space and time to simulate streamflow directly, instead of relying on an IDF curve to drive an event-based hydrologic model (Hamlet et al. 2005, 2013; Hamlet and Lettenmaier 2007; S. Y. Lee et al. 2016; Mote et al. 2005; Tohver et al. 2014). The following physically based hydrologic models use an energy- and mass-balance approach to track both snow accumulation and melt, as well as soil moisture and runoff generation: the Variable Infiltration Capacity (Liang et al. 1994), the Distributed Hydrology Soil Vegetation Model (DHSVM; Wigmosta et al. 1994), the Precipitation Runoff Modeling System (Leavesley et al. 1983), the Weather Research and Forecasting Model Hydrological modeling extension package (Gochis et al. 2015), the Structure for Unifying Multiple Modeling Alternatives (Clark et al. 2015), the Soil Water Assessment Tool (Arnold et al. 2012), and the Council for Regulatory Environmental Modeling (EPA 2009). Stand-alone snow models, such as the SNOW-17 (Anderson 1976), the Utah Energy Balance Snow Model (Tarboton et al. 1995),
and the iSnobal (Marks et al. 1999), are also available and widely used in research (Hay and Clark 2003; Painter et al. 2016; Raleigh and Lundquist 2012).

Calibration and validation of hydrologic models are often based on streamflow observations only. However, verifying specific hydrologic processes (e.g., snow processes) with additional datasets can help constrain the parameterization (Dickerson-Lange and Mitchell 2013; Du et al. 2014) and avoid over-parameterization, a common problem with complex physics-based models. This is particularly important for snow-dominated river basins where flow regime is closely linked to snowpack dynamics, including the spatial extent and duration of snow cover, peak SWE, and the onset of melt. Snowpack Telemetry (SNOTEL) data provide daily and/or subdaily time series of SWE, snow depth, and, in some cases, other observations such as soil temperature. They are also a source of high-elevation meteorological data, including air temperature and precipitation. It is noted that under-catch of snowfall is a common problem with SNOTEL data.

Yan et al. (2018) produced quantitative estimates of AWR during rain, snowmelt, and ROS events at 376 SNOTEL stations across the western United States that each had at least 30 years of high-quality records. They provide a daily time series of the following parameters at each SNOTEL site:

- Total precipitation
- Rainfall only
- AWR from melt without precipitation
- AWR from melt associated with mixed rainfall and snowfall
- AWR from all melt events including ROS
- AWR from ROS only
- AWR from all events plus rainfall on snow-free ground

Yan et al. (2018) used this information to generate next-generation IDF (NG-IDF) curves for application in snow-dominated locations (see Section 8.3.2). Hamlet (2018) argues the importance of this dataset for the validation and refinement of off-line hydrology models and land-surface schemes embedded in climate models, which can then be used to extend these data in space and time to create more comprehensive products to guide infrastructure design. Typically, performance of a model’s snow and ROS components is through comparison with observed streamflow, measurements of SWE, and less commonly remotely sensed snow extent and freezing elevation products. The new dataset will allow different models’ estimates of P + snowmelt + drainage reaching the land surface to be explicitly compared with a dataset derived from observations. As noted by Hamlet (2018), this approach also avoids confounding effects associated with different infiltration schemes in different models when comparing to streamflow extremes. Likewise, because SNOTEL data also provide coincident SWE data, errors in antecedent conditions can be accounted for in evaluating model simulations. When available, we strongly suggest the use of this type of information when applying physics-based continuous simulation models in snow-dominated regions. Refer to Chapter 7 for a more complete discussion of physics-based continuous simulation models for addressing non-stationarity in riverine flooding.

Local climate and topography directly impact snow processes through mass balance and energy balance. Overstory vegetation also plays a complex and important role in both intercepting snowfall and the energy balance of below-canopy snowpack. Snow accumulation and ablation in mountain forest environments depend critically on forest structure (Broxton et al. 2015; Connaughton 1935; Moore and McCaughey 1997; Pomeroy et al. 1998; Varhola et al. 2010). Many studies found that forested areas, in contrast to open areas, commonly accumulate less snow and thus produce less AWR, due mainly to canopy interception and evapotranspiration of up to 60% of accumulated snow (Cristea et al. 2014;
Hedstrom and Pomeroy 1998; Marks et al. 1998; McCabe et al. 2007; McCabe and Clark 2005; Pomeroy et al. 1998; Regonda et al. 2005; Stednick 1996; Stewart et al. 2005; Troendle and King 1985). Previous research suggests that thinner forest canopy advances melt due to reduced radiation attenuation (Link and Marks 1999; Troendle and King 1985; Varhola et al. 2010). However, the opposite may occur when the enhanced longwave radiation under canopies offsets the reduction in shortwave radiation from canopy attenuation, a situation referred to as the radiative paradox (Ambach 1974; Lundquist et al. 2013; Sicart et al. 2004).

The role of vegetation on snow dynamics remains an important area of active research. For example, Sun et al. (2018) developed a canopy gap component in DHSVM as a unique vegetation cover type that is characterized on the sub-grid scale; the component includes physics-based representation of the processes that allow snowpack in the forest gap to be simulated separately from snowpack in the surrounding forest. The predicted snow water SWE using the enhanced DHSVM showed good agreement with sub-hourly SWE measurements at three experimental sites in Idaho—a completely open site, a densely forested site, and a canopy gap site (Figure 11).

![Figure 11. Comparison of the measured and modeled snow water equivalent (SWE) in the (A) open, (B) forest, and (C) gap sites. Note the different y-axis scales. In (C), the simulated SWE at the forest site is replotted on the same scale for comparison (Source: Sun et al. 2018).](image)

### 8.3 IDF-Based Hydrologic Design

As noted previously, despite scientific advances (i.e., physics-based, distributed, continuous simulation models), a large portion of infrastructure in snow-dominated regions is currently designed using traditional IDF technology—the design storm approach. Although the design storm approach is straightforward to use, it greatly simplifies the hydrologic process and suffers from several problematic requirements and assumptions (Camici et al. 2011; Cheng et al. 2014; Milly et al. 2008; Rogger et al. 2012; Yan et al. 2018; McCuen 2002):

1. The need to select a specific design storm hyetograph (both shape and duration)
2. The need to specify antecedent moisture condition prior to the storm event
3. The assumption that the return period of design storm and resulting flood event are equal
4. The assumption that the occurrence probability of extreme precipitation events does not change significantly over time (i.e., stationarity)
5. The assumption that the design storm is in the form of rainfall that is instantly available for the rainfall-runoff process
We illustrate the implications of some of these assumptions below and offer a potential path forward when conditions dictate the use of this approach in snow-dominated locations.

8.3.1 Limitations of Precipitation-Based IDF Curves

Yan et al. (2018) proposed NG-IDF curves based on annual maximum of the AWR time series (equation [17]) as an alternative to PREC-IDF curves for hydrologic design in both rain- and snow-dominated regions. They compared NG-IDF and PREC-IDF at nearly 400 SNOTEL stations across the western United States that each had at least 30 years of high-quality records (assuming $S = 0$ in equation [17]). PREC-IDF curves at 45% of the stations were subject to under-design, many with substantial underestimation of 10- to 100-year extreme events, for which the PREC-IDF curves can underestimate AWR by as much as 125% due to snowmelt and ROS events. They also found the potential for overdesign at 20% of the SNOTEL locations, primarily in the Middle Rockies and Arizona mountains.

Beyond the IDF assessment, Yan et al. (2018) further compared peak design flood estimates from the National Resource Conservation Service TR-55 hydrologic model (Cronsheey et al. 1986) driven by NG-IDF and PREC-IDF curves ($q_{NG}$ and $q_{PREC}$, respectively) at 399 SNOTEL stations across the western United States. They found that about 72% of the SNOTEL locations in the western United States showed the potential for underdesign, wherein the PREC-IDF curves underestimated peak design floods by as much as 324% (Figure 12).

It is also noted that the relative errors between the PREC-IDF and NG-IDF curves do not necessarily generate errors of similar magnitude in the derived design floods because of highly nonlinear hydrologic processes and the implicit assumptions in the design storm approach. Only about 45% of stations were shown to be underdesigned when comparing IDF curves (Yan et al. 2018); this increased to 72% when considering flood peaks (Yan et al. 2018). The authors also found that the peak flood errors can be either larger or smaller than the associated IDF errors, depending on the local climate condition. This result must be considered in the impact study of non-stationary climate on infrastructure design, because many current non-stationary IDF studies only focus on the IDF assessment and do not consider the derived design flood estimate.
Figure 12. Relative percentage of differences ($q_{NG} - q_{PREC}/q_{PREC} \times 100$) of the 10-, 25-, 50-, and 100-year events between the peak design floods derived from NG-IDF ($q_{NG}$) and PREC-IDF ($q_{PREC}$). Stations with red circles indicate underdesign ($q_{NG} > q_{PREC}$), stations with blue circles suggest overdesign ($q_{NG} < q_{PREC}$), and black dots indicate proper design for the stations. The diameter of the circles indicates the magnitude of the relative difference (Source: Yan et al. 2018).

8.3.2 NG-IDF Based Design

Evaluation of new methods for design flood estimation (e.g., NG-IDF curves) is a fundamental question in hydrology, because there is no absolute reference against which results can be compared (Boughton and Droop 2003; Calver et al. 2009; Yan and Moradkhani 2016). In many cases, the IDF event-based design method is evaluated against a physics-based hydrologic continuous simulation model (Calver et al. 2009; Camici et al. 2011; Grimaldi et al. 2012; Rogger et al. 2012; Boughton et al. 2002). Recently, Yan et al. (2018) used a simple cascade of linear reservoirs as a “representative” hillslope characterized by its length ($L$) and time of concentration ($T_c$), which were used to parameterize the TR-55 single-event hydrologic model. Meteorological data from each SNOTEL station were used to drive the physics-based DHSVM hydrologic model and generate a time series of AWR and streamflow from the representative hillslope. NG-IDF curves were derived from the AWR and used with the TR-55 model to estimate flood frequency. These values were compared at each SNOTEL station to flood frequency estimates from PREC-IDF curves and flood frequency estimates calculated directly from the DHSVM simulated streamflow ($q_{DHSVM}$), with the results presented in Figure 13.
Figure 13. Top row: the relative peak design flood error \((q_{\text{PREC}} - q_{\text{DHSVM}}) / q_{\text{DHSVM}} \times 100\) for (a) 10-year event, (b) 25-year event, and (c) 50-year event between the peak design floods derived from PREC-IDF \((q_{\text{PREC}})\) and DHSVM continuous simulations \((q_{\text{DHSVM}})\). Bottom row: the relative peak design flood error \((q_{\text{NG}} - q_{\text{DHSVM}}) / q_{\text{DHSVM}} \times 100\) for (d) 10-year event, (e) 25-year event, and (c) 50-year event between the peak design floods derived from NG-IDF \((q_{\text{NG}})\) and continuous simulations \((q_{\text{DHSVM}})\). For all panels, stations with warm colors indicate overpredictions; stations with blue color suggest underpredictions (Source: Yan et al. 2018).

In Figure 13, it is noted that in order to estimate the potential peak flood using the TR-55 event-based model, the antecedent moisture condition was set to wet condition. For each IDF return period, the duration (selected from 1–30 days) that produced the largest flood peak was selected to be the critical design duration in the TR-55 model. It is observed that for all three return periods, \(q_{\text{PREC}}\) showed large errors relative to \(q_{\text{DHSVM}}\) at most of the stations, whereas \(q_{\text{NG}}\) showed much better agreement with the DHSVM results. Figure 14 presents these results as scatterplots of the peak design flood estimates between NG-IDF, PREC-IDF, and DHSVM. When using the PREC-IDF curves, about 74%, 72%, and 68% of stations showed relative errors larger than \(\pm20\%\) for the 10-, 25-, and 50-year events, respectively. When using the NG-IDF curves, only 19%, 15%, and 13% of stations showed errors larger than \(\pm20\%\) for the 10-, 25-, and 50-year events, respectively. Due to the aforementioned limitations within the event-based method, differences between continuous simulation and event-based flood frequencies are unavoidable, and Calver et al. (2009) suggest that errors within 20% are very good in practice given this hydrologically challenging context.
Figure 14. Top row: scatterplots of the (a) 10-, (b) 25-, and (c) 50-year normalized peak design floods derived from PREC-IDF curves versus normalized peak design floods from DHSVM continuous simulations for the 246 SNOTEL stations across the western United States. Bottom row: scatterplots of the (d) 10-, (e) 25-, and (f) 50-year normalized peak design floods derived from NG-IDF curves versus normalized peak design floods from DHSVM continuous simulations. For all panels, the black line indicates the 1:1 line. For each return period, all peak floods were normalized by the maximum DHSVM peak flood across all 246 stations (Source: Yan et al. 2018).

Based on these results, it can be concluded that there is an emerging need for updating IDF curves for hydrologic design in snow-dominated regions. NG-IDF curves may provide an alternative to standard PREC-IDF curves. NG-IDF curves provide a consistent design approach that also works in rain-dominated systems (equal to PREC-IDF curves for this case). They can be considered an enhancement to the standard PREC-IDF design approach, thereby minimizing technology transfer to practicing engineers.

The observation-based NG-IDF curves discussed above were limited to SNOTEL stations with bare soil conditions (open areas) across the western United States. The location of infrastructure is unlikely to be coincident with SNOTEL stations in practical design problems. As a result, a desired path forward is to extend these AWR datasets in spatial and temporal scale using a well-calibrated, physics-based hydrologic model to provide more comprehensive NG-IDF products for future infrastructure design. The effects of different vegetation cover on NG-IDF curves should be also included. To fully implement this approach, it would be necessary to develop NG-IDF curves for specific vegetation types, such as those
employed by TR-55. This would likely be accomplished using a physics-based hydrologic model that would account for snow-canopy interaction and provide AWR.
# Effective Application of Non-stationary and Uncertain Design Values

Appropriate decision-making tools given non-stationary design values is the focus of this chapter. Such tools aim to address the uncertainty in design values that stems from non-stationarity.

## 9.1 Addressing Uncertainty with Design Storms and Floods

To begin, we note that under the assumption of stationarity, design has traditionally relied on pre-determined return periods (e.g., 100-year storm or flood); such standards are often more politically than scientifically or economically motivated. More recently, but still in the context of stationarity, design has followed risk-based approaches that seek to minimize total expected costs (i.e., balance the cost of protection with the expected costs associated with a disastrous event). In both cases, uncertainty is often quantified using confidence intervals, which usually do not influence design (i.e., design is usually based on the mean or median value). Rather, uncertainty is typically addressed by applying safety factors to a design value.

Under an assumption of non-stationarity, there are several approaches to incorporating uncertainty into design: climate factors, the prudent approach, and robustness-based decision methods (François et al., 2019). Each is discussed in more detail in the sections below, which are heavily based on François et al. (2019). See Section 3.1, where the sources of uncertainty stemming from the use of models and GCM projections of climate are described.

## 9.2 Climate Factors

In the context of uncertainty associated with climate change, additional safety factors are called climate allowances or climate factors. Through a review of existing guidelines in Europe, Madsen et al. (2014) found that the use of climate factors with the explicit purpose of protecting against climate change is rare; exceptions include Germany, Norway, and the United Kingdom (DEFRA 2006; Environmental Agency 2016; Hennegriff et al. 2006; Lawrence and Hisdal 2011). Although climate factors are easy to apply, such a “simplistic adjustment” is a result of “poorly understood impacts of future climate change” (Kuklicke and Demeritt 2016). Climate factors lack flexibility because they are generally prescribed for a single time horizon (one exception is the recent United Kingdom guidelines) and incorrectly estimate small-scale variability because they are usually defined by basin or political boundaries (Madsen et al. 2014). The calculation of climate factors is generally described in technical reports, which often lack clear descriptions of calculation methods and are based on the difference between historical and projected simulations of streamflow from GCMs. Furthermore, studies that assess climate factor performance are limited but needed, especially given that the resulting climate factors can be highly sensitive to the modeling choices (e.g., the first-generation of United Kingdom climate factors, see DEFRA 2006; Prudhomme et al. 2010; Reynard et al. 2005).

## 9.3 Prudent Approach

Even under the uncertainty of climate change, the direction of change of some key variables can be known with relatively high confidence; for example, temperature (and water vapor) has and will continue to increase under climate change. This information can be used to qualitatively infer the likely direction of change in hydrological extremes (e.g., decreased snowpack would lead to smaller and earlier peak flows, or intensified extreme precipitation will increase peak flows), assuming sufficiently
negligible feedback effects. Given this information, decision-makers might decide to opt for a prudent approach, based on the precautionary principle (Gollier and Treich 2003). In the context of climate change mitigation, Kirkwood (2011) noted that “the prudent path lies somewhere between doing absolutely nothing about climate change and doing everything possible.” In the context of decision-making for design, the prudent approach consists of making design decisions based on expected changes that are known with relatively high confidence. The prudent approach is especially applicable for projects involving discrete choices, such as culvert size or levee height. Choosing the next higher discrete level of design is often justifiable when its marginal cost is small compared to the total infrastructure cost (Hallegatte 2009).

9.4 Robust-Based Decision Methods
To deal with deep uncertainty associated with climate change, Lempert (2002) introduced the idea of robustness. Unlike a risk-based approach, which finds an optimal design for one assumed future state of the world, a robustness-based approach finds a design that will be satisfactory (i.e., perform well; Simon 1956) for a large range of plausible futures. Multiple approaches using the concept of robustness have been developed: Robust Decision Making (Lempert et al. 2003), Info-Gap analysis (Ben-Haim 2006), Scenario-Neutral approach (Prudhomme et al. 2010), and Decision Scaling (Brown et al. 2011, 2012). Stakhiv (2011) provides an extended discussion on the use of robust decision-making methods for water resource management under climate change. Despite the value of robustness-based decision methods, quantitative approaches to their implementation are relatively new and require increased application and improvement. For example, Hall et al. (2012) showed that Info-Gap and Robust Decision Making provide similar but not identical solutions for the same case study and note that the comparison improved understanding of the system and the proposed management options.

9.5 The Way Forward
For addressing the uncertainty associated with non-stationarity from climate change, we recommend robustness-based decision methods, which are the most quantitatively advanced of the available methods, despite being only recently developed. Regardless of the chosen approach, it is worth noting that a static infrastructure design is implied. Instead of static design, adaptive management rests on flexible design decisions that benefit from increasing knowledge on expected changes and uncertainty through time (e.g., Hui et al. 2018); two approaches, which have been applied to floods, are Dynamic Adaptive Policy Pathways (Haasnoot et al. 2013) and real options analysis (e.g., Hino and Hall 2017). Adaptive management can reduce or delay initial and often large investment costs associated with infrastructure construction that may subsequently become unnecessary under future states of the world. We recommend that adaptive approaches be combined with robustness-based approaches to promote design that is both robust and flexible under non-stationarity from climate change.
10 Recommendations

10.1 Short-term Planning (20 years or less)
Updates to IDF values reflecting the current risk should adopt the following approach. First, a period of analysis should be chosen that is long-enough to discern decadal trends. It should be analyzed through the application of non-stationary methods of data analysis and/or risk-based approaches that incorporate the contemporary changes in risk. Different non-stationary statistical models can be considered through combinations of time and other covariates of the location and the scale parameters of the distribution of the extreme being analyzed. Then, apply either the Bayesian model averaging to combine the different models or select the appropriate model based on model’s performance and parsimony. Second, evaluation of observed trends at the local level should recognize the inherent spatial variability in metrics of extremes and the high uncertainty at individual points. Because of the relatively high spatial and temporal variability of extremes, it is difficult to identify underlying trends at the local level using single stations. Regional pooling of station data helps to filter out the spatial variability and more robustly identifies underlying climate signals. Third, some methods, when applied independently across different durations of extreme events, can result in physically impossible crossings of the IDF curves. Such instances need to be identified and corrected.

10.2 Long-term Planning (more than 20 years)
Beyond 20 years, changes in climate are expected to be large. They will violate the assumption of stationarity. Fundamental physical principles associated with anthropogenically forced global warming will lead to increases in atmospheric water vapor content, which will increase the potential for extreme precipitation. As a result, future increases in IDF values are highly likely and should be incorporated into decision-making. Robustness-based decision methods are needed for addressing the uncertainty associated with non-stationarity from climate change using the most quantitatively advanced available methods, despite their being only recently developed. Although static infrastructure design is normally implied, adaptive management resting on flexible design decisions that benefit from increasing knowledge of expected changes and uncertainty through time will be more effective (Hui et al. 2018), including Dynamic Adaptive Policy Pathways (Haasnoot et al. 2013) and real options analysis (e.g., Hino and Hall 2017).

10.3 Global Climate Models
Global and regional climate model simulations are best suited for estimating long-term changes in large-scale climate conditions affecting extreme precipitation, such as atmospheric water vapor concentrations and large-scale weather systems such as fronts and tropical and extratropical cyclones. The direct use of precipitation data from these models is not recommended because of biases that have persisted through several generations of models in simulating the record of observed extreme precipitation. Rather, the application of observed relationships linking extreme precipitation amounts and large-scale atmospheric conditions such as temperature, vertical velocity, stability, and humidity to model simulations is preferred. This capitalizes most directly on the strengths of the vast majority of today’s GCMs. Cloud-resolving climate models avoid the need for convective parameterizations, a major source of extreme precipitation biases in contemporary GCMs, but available simulations from such models are currently insufficient in number and length to provide a robust basis for future IDF values.
10.4 Special Considerations Related to Riverine Flooding

For long-term planning, non-stationary riverine flood projections and design values should be developed using statistical models that are climate informed. Only in situations where it is not possible to identify an adequate set of climate predictors should the hydrologic simulation approach be used (see Chapter 7). Any calculation of a design value based on the assumption of stationarity (e.g., for short-term planning) requires trend analysis to ensure that the specified time period used to develop the statistical model is truly stationary. Consideration of the following will improve the reliability of trend analysis: 1) use long-record data that is supplemented, where possible, by well-calibrated proxy data to reconstruct the longest records with minimal time-dependent biases relative to any observed changes in the extremes; 2) use a regional trend analysis to minimize the effect of noisy records from point sources (i.e., individual stations); 3) examine the potential changes associated with the underlying causes of the trends; and 4) evaluate and choose a method for trend analysis that ensures adequate representation of the non-stationarity that is relevant for the IDF curve development or update. Finally, in all cases, a thorough assessment is required of the 1) model credibility (i.e., whether the model adequately simulates observations and incorporates critical driving mechanisms), 2) model robustness (i.e., whether the model formulation and predictors will remain valid under future change), and 3) accompanying uncertainty (i.e., the degree of confidence in the projections).

10.5 Rain-on-Snow Riverine Flooding

Hydrologic design approach for snow-dominated regions in a non-stationary climate is challenging, especially in areas where snowmelt contributes to total runoff. Traditionally, the design storm approach (IDF technology) based on single-event-based models is standard practice for engineers designing infrastructure in snow-dominated regions. A time series of the available water for runoff (AWR) at the ground surface is more robust for flood predictor than extreme precipitation alone. AWR can be derived from observations or through hydrologic simulation, where IDF curves can be developed from annual maximum of the AWR time series. The set of next-generation IDF curves should be used as an enhancement to the standard precipitation-related IDF design approach historically used in snow-dominated regions.

In areas where the flood-causing mechanism is dominated by antecedent soil moisture and snowmelt, the direct use of rainfall-based IDF curves are not recommended. Instead, hydrological analysis is required to translate the rainfall IDF curves to corresponding runoff IDF curves.

10.6 Regional Averaging

For non-stationary regional frequency analysis, we recommend first creating homogeneous regions that remain homogeneous over time by identifying areas undergoing the same precipitation mechanisms and the same changes in the forcings of the extreme being analyzed. Testing is needed to determine the significance of a non-stationary trend at the regional level, as well as including a time index into the formulation of the regional growth curve.

For non-stationary IDAF curves, we recommend examining how non-stationarity affects the IDF curve (see Chapter 6) and the ARF. Currently, there is no established method for creating a non-stationary ARF or overall IDAF curve. A non-stationary IDAF curve requires that simple scaling in time and space must still hold true for the IDF, along with the underlying assumptions of the scaling.
10.7 Designing for a Changing Climate
Because the design of stormwater infrastructure such as culverts and drainage are dependent on design variables that are likely to change in time and are difficult to accurately predict into the future, a prudent approach for such facilities is to design for increasing extremes using a safety factor. This could include an upper range of confidence intervals based on historical data. In addition, adaptive approaches that include options for the expansion of capacity to manage extremes as they evolve in the future should be explored. While the exact degree of increases in extreme precipitation and related flooding are difficult to specify, the evidence strongly suggests the need to plan for greater risk in the future.
References


Bador, M., M.G. Donat, O. Geoffroy, and L.V. Alexander, 2018: Assessing the robustness of
future extreme precipitation intensification in the CMIP5 ensemble. *Journal of Climate, 31* (16), 6505–6525. [http://dx.doi.org/10.1175/JCLI-D-17-0683.1](http://dx.doi.org/10.1175/JCLI-D-17-0683.1)


Berg, P., C. Moseley, and J.O. Haerter, 2013: Strong increase in convective precipitation in response to higher temperatures. *Nature Geoscience, 6* (3), 181–185. [http://dx.doi.org/10.1038/ngeo1731](http://dx.doi.org/10.1038/ngeo1731)


Cheng, L. and A. AghaKouchak, 2014: Nonstationary precipitation Intensity-Duration-Frequency curves for infrastructure design in a changing climate. *Scientific Reports, 4*(1), 7093. [http://dx.doi.org/10.1038/srep07093](http://dx.doi.org/10.1038/srep07093)


DEFRA, 2006: Flood and Coastal Defence Appraisal Guidance (FCDPAG3), Economic
Appraisal Supplementary Note to Operating Authorities – Climate Change Impacts.
Department for Environment Food and Rural Affairs, 9 pp. http://bailey.persona-
ipi.com/Public-Inquiries/M4-Newport/C%20-
%20Core%20Documents/17.%20Road%20Drainage%20and%20the%20Water%20Environ-
ment/17.2.20.pdf


2012


Goswami, B.N., V. Venugopal, D. Sengupta, M.S. Madhusoodanan, and P.K. Xavier, 2006: Increasing trend of extreme rain events over India in a warming environment. *Science*, **314** (5804), 1442. [http://dx.doi.org/10.1126/science.1132027](http://dx.doi.org/10.1126/science.1132027)


Abfall, 53 (8).


Kattelmann, R., 1997: Flooding from rain-on-snow events in the Sierra Nevada. In *Destructive Water: Water-Caused Natural Disasters, Their Abatement and Control*, Anaheim, CA, June 1996. [https://data.sacriver.org/assets/5ffd2763a1b5d1a3078c4bd991651244/application/pdf/Flooding_from_rain-on-snow_events_in_the_Sierra_Nevada.pdf](https://data.sacriver.org/assets/5ffd2763a1b5d1a3078c4bd991651244/application/pdf/Flooding_from_rain-on-snow_events_in_the_Sierra_Nevada.pdf)


Kharin, V.V. and F.W. Zwiers, 2005: Estimating extremes in transient climate change simulations. *Journal of Climate*, **18** (8), 1156–1173. [http://dx.doi.org/10.1175/JCLI3320.1](http://dx.doi.org/10.1175/JCLI3320.1)


Kunkel, K.E., D.R. Easterling, D.A.R. Kristovich, B. Gleason, L. Stoecker, and R. Smith, 2012:


Lenderink, G. and E. van Meijgaard, 2008: Increase in hourly precipitation extremes beyond expectations from temperature changes. Nature Geoscience, 1 (8), 511–514. http://dx.doi.org/10.1038/ngeo262


Lundquist, J.D., S.E. Dickerson-Lange, J.A. Lutz, and N.C. Cristea, 2013: Lower forest density enhances snow retention in regions with warmer winters: A global framework developed
from plot-scale observations and modeling. *Water Resources Research*, **49** (10), 6356–6370. [http://dx.doi.org/10.1002/wrcr.20504](http://dx.doi.org/10.1002/wrcr.20504)


Mailhot, A. and S. Duchesne, 2010: Design criteria of urban drainage infrastructures under climate change. *Journal of Water Resources Planning and Management*, **136** (2), 201–208. [http://dx.doi.org/10.1061/(ASCE)WR.1943-5452.0000023](http://dx.doi.org/10.1061/(ASCE)WR.1943-5452.0000023)


Martel, J.-L., A. Mailhot, F. Brissette, and D. Caya, 2018: Role of natural climate variability in the detection of anthropogenic climate change signal for mean and extreme precipitation at local and regional scales. *Journal of Climate*, **31** (11), 4241–4263. [http://dx.doi.org/10.1175/JCLI-D-17-0282.1](http://dx.doi.org/10.1175/JCLI-D-17-0282.1)


Matalas Nicholas, C., 2012: Comment on the announced death of stationarity. *Journal of Water Resources Planning and Management*, **138** (4), 311–312. [http://dx.doi.org/10.1061/(ASCE)WR.1943-5452.0000215](http://dx.doi.org/10.1061/(ASCE)WR.1943-5452.0000215)

McCabe, G.J. and M.P. Clark, 2005: Trends and variability in snowmelt runoff in the western
http://dx.doi.org/10.1175/JHM428.1

http://dx.doi.org/10.1175/BAMS-88-3-319

http://dx.doi.org/10.1061/(ASCE)1084-0699(2002)7:1(43)

http://dx.doi.org/10.1029/2007GL031027

http://dx.doi.org/10.1002/2013JD021152

http://dx.doi.org/10.1029/2009JC005997

http://dx.doi.org/10.5194/hess-20-3527-2016

http://dx.doi.org/10.5194/hess-16-1379-2012

http://dx.doi.org/10.1029/2002WR001952

http://dx.doi.org/10.1007/s00382-016-3352-8

http://dx.doi.org/10.1126/science.1151915

http://dx.doi.org/10.1002/2015WR017408


Mukherjee, S., S. Aadhar, D. Stone, and V. Mishra, 2018: Increase in extreme precipitation events under anthropogenic warming in India. *Weather and Climate Extremes*, **20**, 45–53. [http://dx.doi.org/10.1016/j.wace.2018.03.005](http://dx.doi.org/10.1016/j.wace.2018.03.005)


Ning, L. and R.S. Bradley, 2016: NAO and PNA influences on winter temperature and precipitation over the eastern United States in CMIP5 GCMs. *Climate Dynamics*, **46** (3), 1257–1276. [http://dx.doi.org/10.1007/s00382-015-2643-9](http://dx.doi.org/10.1007/s00382-015-2643-9)


Salas Jose, D., B. Rajagopalan, L. Saito, and C. Brown, 2012: Special section on climate change and water resources: Climate nonstationarity and water resources management. *Journal of Water Resources Planning and Management*, **138** (5), 385–388. [http://dx.doi.org/10.1061/(ASCE)WR.1943-5452.0000279](http://dx.doi.org/10.1061/(ASCE)WR.1943-5452.0000279)


Schlef Katherine, E., S. Steinschneider, and M. Brown Casey, 2018: Spatiotemporal impacts of climate and demand on water supply in the Apalachicola-Chattahoochee-Flint basin.
Journal of Water Resources Planning and Management, 144 (2), 05017020.
http://dx.doi.org/10.1061/(ASCE)WR.1943-5452.0000865


http://dx.doi.org/10.1002/jgrd.50203


Tabari, H. and P. Willems, 2018: Seasonally varying footprint of climate change on precipitation in the Middle East. *Scientific Reports*, **8** (1), 4435. [http://dx.doi.org/10.1038/s41598-018-22795-8](http://dx.doi.org/10.1038/s41598-018-22795-8)


Wasko, C. and A. Sharma, 2015: Steeper temporal distribution of rain intensity at higher temperatures within Australian storms. *Nature Geoscience*, 8 (7), 527–529. [http://dx.doi.org/10.1038/ngeo2456](http://dx.doi.org/10.1038/ngeo2456)


Westra, S., L.V. Alexander, and F.W. Zwiers, 2013: Global increasing trends in annual maximum daily precipitation. *Journal of Climate*, 26 (11), 3904–3918. [http://dx.doi.org/10.1175/JCLI-D-12-00502.1](http://dx.doi.org/10.1175/JCLI-D-12-00502.1)


Xu, Y., X. Gao, F. Giorgi, B. Zhou, Y. Shi, J. Wu, and Y. Zhang, 2018: Projected changes in temperature and precipitation extremes over China as measured by 50-yr return values and periods based on a CMIP5 ensemble. *Advances in Atmospheric Sciences*, 35 (4), 376–388. [http://dx.doi.org/10.1007/s00376-017-6269-1](http://dx.doi.org/10.1007/s00376-017-6269-1)


Yilmaz, A.G. and B.J.C. Perera, 2014: Extreme rainfall nonstationarity investigation and

