Experiences of Hydrological Frequency Analysis in Taiwan
- Coping with Extraordinary Extremes and Climate Change

Prof. Ke-Sheng Cheng
Department of Bioenvironmental Systems Engineering, National Taiwan University
Outline

• Introduction

• Improving accuracy of frequency analysis
  – Presence of outliers (extraordinary rainfall extremes)
  – Spatial correlation (non-Gaussian random field simulation)

• Frequency analysis of multi-site rainfall extremes
Introduction

• Frequent occurrences of disasters induced by heavy rainfalls in Taiwan
  – Landslides
  – Debris flows
  – Flooding and urban inundation

• Increasing occurrences of rainfall extremes (some of them are record breaking) in recent years

• Almost all of the long-duration rainfall extremes (longer than 12 hours) were produced by typhoons.
Examples of **annual max events** in Taiwan

<table>
<thead>
<tr>
<th>Design durations</th>
<th>1965-year</th>
<th>1969-year</th>
<th>1974-year</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Duration (hrs)</strong>*</td>
<td>1 2 3 4 6 12 18 24 48 72</td>
<td>1 2 3 4 6 12 18 24 48 72</td>
<td>1 2 3 4 6 12 18 24 48 72</td>
</tr>
<tr>
<td>Duration (hrs)*</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>1983-year</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hosolau</td>
<td>5/23</td>
<td>5/23</td>
<td>5/23</td>
</tr>
<tr>
<td>Wutuh</td>
<td>10/1</td>
<td>10/1</td>
<td>6/3</td>
</tr>
<tr>
<td><strong>1987-year</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hosolau</td>
<td>7/26</td>
<td>10/22</td>
<td>10/22</td>
</tr>
<tr>
<td>Wutuh</td>
<td>10/22</td>
<td>7/26</td>
<td>10/22</td>
</tr>
<tr>
<td><strong>1994-year</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wutuh</td>
<td>6/18</td>
<td>6/18</td>
<td>6/18</td>
</tr>
<tr>
<td>Year</td>
<td>Station</td>
<td>Stations Observing Annual Max Rainfalls</td>
<td>Proportion</td>
</tr>
<tr>
<td>-------</td>
<td>----------</td>
<td>----------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>19840624</td>
<td>WYNNE魏恩</td>
<td>26</td>
<td>0.269231</td>
</tr>
<tr>
<td>19850626</td>
<td>-</td>
<td>28</td>
<td>0.214286</td>
</tr>
<tr>
<td>19860919</td>
<td>ABBY艾貝</td>
<td>28</td>
<td>0.392857</td>
</tr>
<tr>
<td>19870727</td>
<td>ALEX亞力士</td>
<td>28</td>
<td>0.392857</td>
</tr>
<tr>
<td>19880813</td>
<td>-</td>
<td>18</td>
<td>0.555556</td>
</tr>
<tr>
<td>19890912</td>
<td>SARAH莎拉</td>
<td>25</td>
<td>0.8</td>
</tr>
<tr>
<td>19900523</td>
<td>OFELIA歐菲莉</td>
<td>28</td>
<td>0.571429</td>
</tr>
<tr>
<td>19910522</td>
<td>-</td>
<td>26</td>
<td>0.307892</td>
</tr>
<tr>
<td>19920830</td>
<td>POLLY寶莉</td>
<td>27</td>
<td>0.666667</td>
</tr>
<tr>
<td>19930912</td>
<td>ABB亞伯</td>
<td>20</td>
<td>0.5</td>
</tr>
<tr>
<td>19940803</td>
<td>CAILIN凱特琳</td>
<td>28</td>
<td>0.678571</td>
</tr>
<tr>
<td>19950608</td>
<td>DEANNA薇安娜</td>
<td>28</td>
<td>0.464286</td>
</tr>
<tr>
<td>19960731</td>
<td>HERB賀伯</td>
<td>28</td>
<td>0.75</td>
</tr>
<tr>
<td>19970828</td>
<td>AMBER安伯</td>
<td>26</td>
<td>0.5</td>
</tr>
<tr>
<td>19980804</td>
<td>OTTO奧托</td>
<td>27</td>
<td>0.518519</td>
</tr>
<tr>
<td>19990811</td>
<td>-</td>
<td>27</td>
<td>0.222222</td>
</tr>
<tr>
<td>20000822</td>
<td>BILIS碧利斯</td>
<td>27</td>
<td>0.62963</td>
</tr>
<tr>
<td>20010917</td>
<td>NARI納莉</td>
<td>27</td>
<td>0.481481</td>
</tr>
<tr>
<td>20020805</td>
<td>-</td>
<td>28</td>
<td>0.428571</td>
</tr>
<tr>
<td>20030606</td>
<td>-</td>
<td>28</td>
<td>0.464286</td>
</tr>
<tr>
<td>20040702</td>
<td>MINDULLE敏督利</td>
<td>28</td>
<td>0.857143</td>
</tr>
<tr>
<td>20050718</td>
<td>HAITANG海棠</td>
<td>28</td>
<td>0.607143</td>
</tr>
<tr>
<td>20060809</td>
<td>-</td>
<td>28</td>
<td>0.678571</td>
</tr>
<tr>
<td>20071006</td>
<td>KROSA柯羅莎</td>
<td>28</td>
<td>0.321429</td>
</tr>
<tr>
<td>20080717</td>
<td>KALMAEBGI卡玫基</td>
<td>28</td>
<td>0.642857</td>
</tr>
<tr>
<td>20090808</td>
<td>MORAKOT摩拉克</td>
<td>27</td>
<td>0.962963</td>
</tr>
<tr>
<td>20100919</td>
<td>FANAPI凡那比</td>
<td>27</td>
<td>0.666667</td>
</tr>
</tbody>
</table>

Proportion of rainfall stations observing annual max rainfalls.
• A large number of rain gauges maintained by CWB and WRA. Some of them have record length longer than 50 years. Most of them have less than 30 years record length.

• Design rainfalls play a key role in studies related to climate change and disaster mitigation.

• Problems in rainfall frequency analysis
  – Short record length (less than 30 years) (**small sample size**)
  – Record breaking rainfall extremes (**presence of extreme outliers**)

  • Typhoon Morakot
• Typhoon Morakot (2009)

![Graph showing accumulated rainfall over time for Typhoon Morakot. The graph includes data points and a map of rainfall distribution across Taiwan. The data points indicate significant rainfall accumulation over various durations, with a world record of 2467 mm in 48 hours and 2361 mm in 24 hours. The map highlights areas affected by the typhoon, with color gradients indicating varying levels of rainfall.]
• Catastrophic storm rainfalls (or extraordinary rainfalls) often are considered as extreme outliers. Whether or not such rainfalls should be included in site-specific frequency analysis is disputable.

— 24-hr annual max. rainfalls (Morakot) of 2009

• 甲仙 1077 mm
• 泰武 1747 mm
• 大湖山 1329 mm
• 阿禮 1237 mm
• Frequency analysis of 24-hr **annual maximum rainfalls (AMR)** at Jia-Sien station using **50 years** of historical data
  – 1040mm/24 hours (by Morakot) excluding Morakot – **901 years** return period
  – Return period inclusive of Morakot – **171 years**

The same amount (1040mm/24 hours) was found to be associated with a return period of **more than 2000 years** by another study which used 25 years of annual maximum rainfalls.
Morakot rainfalls were not included in the frequency analysis.

<table>
<thead>
<tr>
<th>流域</th>
<th>雨量站</th>
<th>鄉鎮名稱</th>
<th>24 小時</th>
<th>48 小時</th>
<th>72 小時</th>
<th>報總降雨</th>
<th>影響長度</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>高屏溪</td>
<td>甲仙</td>
<td>1077.5</td>
<td>&gt;2000</td>
<td>1601.0</td>
<td>1856.0</td>
<td>1916.0</td>
</tr>
<tr>
<td></td>
<td>高屏（5）屏東市</td>
<td>屏東縣屏東市</td>
<td>667.0</td>
<td>141</td>
<td>886.0</td>
<td>124</td>
<td>947.0</td>
</tr>
<tr>
<td></td>
<td>美濃（2）美濃鎮</td>
<td>高雄縣美濃鎮</td>
<td>507.0</td>
<td>&gt;2000</td>
<td>749.0</td>
<td>&gt;2000</td>
<td>828.0</td>
</tr>
<tr>
<td></td>
<td>屏東</td>
<td>屏東縣屏東市</td>
<td>666.0</td>
<td>140</td>
<td>906.0</td>
<td>143</td>
<td>974.5</td>
</tr>
<tr>
<td></td>
<td>溪埔</td>
<td>高雄縣大樹鄉</td>
<td>729.5</td>
<td>271</td>
<td>994.5</td>
<td>265</td>
<td>1057.5</td>
</tr>
<tr>
<td></td>
<td>旗山</td>
<td>高雄縣旗山鎮</td>
<td>621.0</td>
<td>&gt;2000</td>
<td>813.0</td>
<td>&gt;2000</td>
<td>854.5</td>
</tr>
<tr>
<td></td>
<td>尾寮山</td>
<td>屏東縣三地門</td>
<td>1414.5</td>
<td>&gt;2000</td>
<td>2215.0</td>
<td>&gt;2000</td>
<td>2564.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>高雄縣甲仙鄉</td>
<td>1077.5</td>
<td>&gt;2000</td>
<td>1601.0</td>
<td>&gt;2000</td>
<td>1856.0</td>
</tr>
</tbody>
</table>

資料來源：「莫拉克颱風暴雨量及洪流量分析」，經濟部水利署，民國 98 年 9 月。原資料尚有潮水溪、北港溪、朴子溪、八掌溪、急水溪、曾文溪、鹽水溪、二仁溪、東港溪、四重溪、林邊溪、知本溪等各雨量站不同降雨延時降雨量頻率分析結果。
• Concurrent occurrences of extraordinary rainfalls at different rain gauges
  – Several stations had 24-hr rainfalls exceeding 100-yr return period.
  – Site-specific events of 100-yr return period.
  – What is the return period of the event of multi-site 100-yr return period?
    • \((100)^4 = 100,000,000\) years (4 sites), assuming independence

– Redefining extreme events
  • multi-site extreme events w.r.t. specified durations and rainfall thresholds
  • Spatial covariation of rainfall extremes
• Spatial covariation of rainfall extremes
  – By using site-specific annual maximum rainfall series for frequency analysis implies a significant loss of valuable information.
## 24-hr Event Maximum Rainfalls

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>160</td>
<td>355</td>
<td>164</td>
<td>387</td>
<td>173</td>
<td>518</td>
<td>383</td>
<td>693</td>
<td>142</td>
<td>1242</td>
<td>321</td>
<td>1493</td>
</tr>
<tr>
<td>2</td>
<td>293</td>
<td>410</td>
<td>383</td>
<td>537</td>
<td></td>
<td></td>
<td>535</td>
<td>505</td>
<td>621</td>
<td>1009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>390</td>
<td>358</td>
<td>313</td>
<td>384</td>
<td>412</td>
<td></td>
<td>696</td>
<td>704</td>
<td>289</td>
<td>548</td>
<td>463</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>168</td>
<td>239</td>
<td>359</td>
<td>471</td>
<td>344</td>
<td></td>
<td>683</td>
<td>858</td>
<td>1219</td>
<td>47</td>
<td>400</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>617</td>
<td>356</td>
<td>186</td>
<td></td>
<td></td>
<td></td>
<td>408</td>
<td>560</td>
<td>230</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>567</td>
<td>280</td>
<td>180</td>
<td>231</td>
<td></td>
<td></td>
<td>347</td>
<td>800</td>
<td>111</td>
<td>171</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Probability distribution of Annual Maximum Rainfalls (Small sample size)

**Station A**

**Station B**
Probability distribution of Event-Maximum Rainfalls (Large sample size)
• Coping with outliers
  – Regional Frequency Analysis
  – Stochastic simulation of multi-site event-maximum rainfalls
Fundamental concept of regional frequency analysis

• For areas with short record length or without rainfall or flow measurements, hydrological frequency analysis needs to be conducted using data from sites of similar hydrological characteristics.

• Data observed at different sites within a “homogeneous region” can be combined and used in regional frequency analysis.
General procedures of regional frequency analysis (Hosking and Wallis, 1997)

1. Data screening
   – Correctness check
   – Data should be stationary over time.

2. Identifying homogeneous regions
   – A set of characteristic variables should be chosen and used for delineation of homogeneous regions.
   – Characteristic variables may include geographic and hydrological variables.
   – Homogeneous regions are often determined by cluster analysis.
3. Choice of an appropriate regional frequency distribution

- GOF test using **rescaled samples** from different sites within the same homogeneous region.
- The chosen distribution not only should fit the data well but also yield quantile estimates that are robust to physically plausible deviations of the true frequency distribution from the chosen frequency distribution.
4. Parameter estimation of the regional frequency distribution
   – Estimating parameters of the site-specific frequency distribution
   – Estimating parameters of the regional frequency distribution using record-length weighted average.

5. Calculating various quantiles of the hydrological variable under investigation.
An experimental stochastic simulation

• We generated 7 random samples (sample size n = 40) of a gamma distribution with mean=600 and std dev =346.

• The data set is considered equivalent to annual max rainfalls (record length = 40) at 7 stations.

• Outliers were detected in four of the seven series.

• Return period of the maximum value of each individual AMR series was calculated.

| Max value in AMR | 1797.758 | 1159.976 | 1567.84 | 2066.767 | 1326.539 | 1288.869 | 1624.943 |
Study area and rainfall stations

25 rainfall stations in southern Taiwan. (1951 – 2010) Not all stations have the same record length. Annual maximum rainfalls (AMR) of various durations (1, 2, 6, 12, 18, 24, 48, 72 hours)
Event-maximum rainfalls of various design durations

- Event-max 1, 2, 6, 12, 18, 24, 48, 72-hr typhoon rainfalls at individual sites.
- Approximately 120 events
- Complete series
Delineation of homogeneous regions (K-mean cluster analysis)

• 24 classification features (8 design durations x 3 parameters – mean, std dev, skewness)
  – Normalization of individual features

• Two homogeneous regions (25 stations)
Regional frequency analysis

- Delineating homogeneous regions
颱風強降雨頻率分布

2010 凡那比

2009 莫拉克
Hot spots for occurrences of extreme rainfalls

1992 – 2010
Number of extreme typhoon events
Rescaled variables for regional frequency analysis – Frequency factors

\[ K_{ijk} = \frac{x_{ijk} - \mu_{ik}}{\sigma_{ik}} \]
Goodness-of-fit test and parameters estimation

- L-moment ratio diagrams (LMRD) for goodness-of-fit test
  - Pearson Type 3 distribution
- Parameters estimation
  - Method of L-moments
  - Record-length-weighted regional parameters
L-moment-ratio diagram GOF test

甲仙（2），t = 24hr

- Gaussian
- PT3
- EV-1

L-skewness
嘉義  t = 24hr  L-moment Ratio Diagram for GOF Tests
Covariance structure of the random field

- Covariance matrices are semi-positive definite.
- Experimental covariance matrices often do not satisfy the semi-positive definite condition.
- Modeling the covariance structure of frequency factors (event-max rainfalls) by variogram modeling.
Relationship between semivariogram and covariance function

\[ \gamma(h) = \frac{1}{2} \text{Var}[Z(x + h) - Z(x)] \]
\[ = C(0) - C(h) \]

\[ C(x_i, x_j) = C(h) \]
\[ \text{sill} = C(0) = 1 \]
Semi-variogram modeling (24-hr EMR)

\[ \gamma(h) = \omega \left( 1 - \exp \left( -\frac{h}{a} \right) \right) \]

\[ \omega = C(0) = 1 \]

Variogram model: \( a = 81 \); \( R^2 = 0.88 \)
Semi-variograms of EMRs of other durations
Stochastic Simulation of Multi-site Event-Max Rainfalls

- Pearson type III (Non-Gaussian) random field simulation
- Covariance Transformation Approach
  - Covariance matrix of multivariate PT3 distribution
  - Covariance matrix of multivariate standard Gaussian distribution
  - Multivariate Gaussian simulation
  - Transforming simulated multivariate Gaussian realizations to PT3 realizations
\( \rho_{XY} \sim \rho_{UV} \) Conversion

\[
\rho_{XY} \approx (A_X A_Y - 3A_X C_Y - 3C_X A_Y + 9C_X C_Y) \rho_{UV} + 2B_X B_Y \rho_{UV}^2 + 6C_X C_Y \rho_{UV}^3
\]

\[
A_X = 1 + \left( \frac{\gamma_X}{6} \right)^4 \quad B_X = \frac{\gamma_X}{6} - \left( \frac{\gamma_X}{6} \right)^3 \quad C_X = \frac{1}{3} \left( \frac{\gamma_X}{6} \right)^2
\]

\[
A_Y = 1 + \left( \frac{\gamma_Y}{6} \right)^4 \quad B_Y = \frac{\gamma_Y}{6} - \left( \frac{\gamma_Y}{6} \right)^3 \quad C_Y = \frac{1}{3} \left( \frac{\gamma_Y}{6} \right)^2
\]
• Each simulation run generated one sample of t-hr multi-site event maximum rainfalls.
• Simulated samples preserved the spatial covariation of multi-site EMRs as well as the marginal distributions.
• 10,000 samples were generated.
  – Multi-site t-hr EMRs of 10,000 typhoon events.
• The number of typhoons vary from one year to another.
  – Annual count of typhoons is a random variable.
• **Determination of t-hr rainfall of T-yr return period**
  
  – Average number of typhoons per year, \( m = 2.43 \)
  
  – Return period, \( T=100 \) years
  
  – Exceedance probability of the event-max rainfalls, \( p_E=1/(100*2.43) \)
Return period, $T = 50\text{yr}$
Duration, $t = 24\text{hr}$
Return Period, $T = 200$ yr
Duration, $t = 24$ hr
<table>
<thead>
<tr>
<th></th>
<th>24hr雨量(mm)</th>
<th>24hr重現期(年)</th>
<th>48hr雨量(mm)</th>
<th>48hr重現期(年)</th>
<th>72hr雨量(mm)</th>
<th>72hr重現期(年)</th>
</tr>
</thead>
<tbody>
<tr>
<td>北港(2)</td>
<td>358</td>
<td>11</td>
<td>433</td>
<td>11</td>
<td>460</td>
<td>12</td>
</tr>
<tr>
<td>大埔</td>
<td>720</td>
<td>53</td>
<td>921</td>
<td>57</td>
<td>1069</td>
<td>118</td>
</tr>
<tr>
<td>沙坑</td>
<td>607</td>
<td>32</td>
<td>759</td>
<td>31</td>
<td>888</td>
<td>52</td>
</tr>
<tr>
<td>中坑(3)</td>
<td>497</td>
<td>24</td>
<td>636</td>
<td>26</td>
<td>744</td>
<td>42</td>
</tr>
<tr>
<td>溪口(3)</td>
<td>324</td>
<td>15</td>
<td>403</td>
<td>16</td>
<td>450</td>
<td>24</td>
</tr>
<tr>
<td>六溪</td>
<td>791</td>
<td>43</td>
<td>1024</td>
<td>55</td>
<td>1122</td>
<td>59</td>
</tr>
<tr>
<td>木柵</td>
<td>778</td>
<td>102</td>
<td>1037</td>
<td>54</td>
<td>1188</td>
<td>80</td>
</tr>
<tr>
<td>阿蓮(2)</td>
<td>630</td>
<td>41</td>
<td>868</td>
<td>42</td>
<td>924</td>
<td>41</td>
</tr>
<tr>
<td>旗山(4)</td>
<td>519</td>
<td>17</td>
<td>760</td>
<td>23</td>
<td>819</td>
<td>22</td>
</tr>
<tr>
<td>美濃(2)</td>
<td>338</td>
<td>4</td>
<td>509</td>
<td>6</td>
<td>569</td>
<td>6</td>
</tr>
<tr>
<td>新豐</td>
<td>908</td>
<td>63</td>
<td>1179</td>
<td>51</td>
<td>1195</td>
<td>36</td>
</tr>
<tr>
<td>屏東(5)</td>
<td>676</td>
<td>40</td>
<td>878</td>
<td>45</td>
<td>938</td>
<td>45</td>
</tr>
<tr>
<td>南和</td>
<td>614</td>
<td>15</td>
<td>988</td>
<td>39</td>
<td>1100</td>
<td>62</td>
</tr>
<tr>
<td>嘉義</td>
<td>526</td>
<td>22</td>
<td>646</td>
<td>17</td>
<td>697</td>
<td>24</td>
</tr>
<tr>
<td>台南</td>
<td>535</td>
<td>61</td>
<td>709</td>
<td>53</td>
<td>736</td>
<td>42</td>
</tr>
<tr>
<td>高雄</td>
<td>538</td>
<td>40</td>
<td>755</td>
<td>57</td>
<td>801</td>
<td>53</td>
</tr>
<tr>
<td>恆春</td>
<td>528</td>
<td>27</td>
<td>715</td>
<td>34</td>
<td>730</td>
<td>28</td>
</tr>
<tr>
<td>關子嶺(2)</td>
<td>1098</td>
<td>92</td>
<td>1490</td>
<td>126</td>
<td>1683</td>
<td>221</td>
</tr>
<tr>
<td>甲仙(2)</td>
<td>1040</td>
<td>65</td>
<td>1614</td>
<td>177</td>
<td>1915</td>
<td>204</td>
</tr>
<tr>
<td>紹家</td>
<td>818</td>
<td>110</td>
<td>1176</td>
<td>141</td>
<td>1333</td>
<td>130</td>
</tr>
<tr>
<td>三地門</td>
<td>825</td>
<td>18</td>
<td>1109</td>
<td>25</td>
<td>1235</td>
<td>29</td>
</tr>
<tr>
<td>大湖山</td>
<td>1329</td>
<td>48</td>
<td>1958</td>
<td>89</td>
<td>2533</td>
<td>306</td>
</tr>
<tr>
<td>樟腦寮(2)</td>
<td>868</td>
<td>22</td>
<td>1395</td>
<td>55</td>
<td>1846</td>
<td>141</td>
</tr>
<tr>
<td>泰武(1)</td>
<td>1747</td>
<td>47</td>
<td>2938</td>
<td>121</td>
<td>3417</td>
<td>171</td>
</tr>
<tr>
<td>阿禮</td>
<td>1237</td>
<td>23</td>
<td>1908</td>
<td>38</td>
<td>2335</td>
<td>66</td>
</tr>
</tbody>
</table>
甲仙(2)站 IDF Curve

IDF Curve

降雨強度 (mm/hr)

延時(hr)
Estimating return period of multisite Morakot rainfall extremes

• Four rain gauges within the Kaoping River Basin recorded 24-hr rainfalls close to 1,000mm (908, 1040, 825, 1237 mm).

• Define a multi-site extreme event
  – over 1,000 mm 24-hr rainfalls at all four sites
  – Among the 10,000 simulated events, only 8 events satisfied the above requirement.
  – Average number of typhoons per year, m=2.43
  – Multi-site extreme event return period $T = 514$ years.
• Further look at the simulation results
  – Preserving the spatial pattern of rainfall extremes
Type A
Type B
同時發生超過100年重現期的最大24小時降雨之測站散佈圖
Type C
同時發生超過100年重現期的最大24小時降雨之測站散佈圖
Type D
Summary

• We developed a stochastic approach for simulation of multi-site event-max rainfalls to cope with the problems of outliers and short record length in hydrological frequency analysis.

• By increasing the sample size and considering the spatial covariation of EMRs, the return periods of site-specific and multi-site rainfall extremes can be better estimated.
Rationale of BVG simulation using frequency factor

• From the viewpoint of random number generation, the frequency factor can be considered as a random variable \( K \), and \( K_T \) is a value of \( K \) with exceedance probability \( 1/T \).

• Frequency factor of the Pearson type III distribution can be approximated by

\[
K_T \approx z + \left( z^2 - 1 \right) \frac{\gamma_X}{6} + \frac{1}{3} \left( z^3 - 6z \right) \left( \frac{\gamma_X}{6} \right)^2 \\
- \left( z^2 - 1 \right) \left( \frac{\gamma_X}{6} \right)^3 + z \left( \frac{\gamma_X}{6} \right)^4 - \frac{1}{3} \left( \frac{\gamma_X}{6} \right)^5
\]

[A]
• General equation for hydrological frequency analysis

\[ X_T = \mu_X + K_T \sigma_X \]

Given \( \mu_X, \sigma_X \) and \( \gamma_X \), if we can generate a set of random numbers of \( K \), say \( k_1, k_2, \ldots, k_n \), then a random sample of \( X \), say \( x_1, x_2, \ldots, x_n \), can be obtained by \( x_i = \mu_X + k_i \sigma_X \).

Note that each \( k_i, i = 1, 2, \ldots, n \), corresponds to its own exceedence probability \( 1/T_i \).
• The gamma distribution is a special case of the Pearson type III distribution with a zero location parameter. Therefore, it seems plausible to generate random samples of a bivariate gamma distribution based on two jointly distributed frequency factors.

\[
K_T \approx z + \left( z^2 - 1 \right) \frac{\gamma_X}{6} + \frac{1}{3} \left( z^3 - 6z \right) \left( \frac{\gamma_X}{6} \right)^2
\
- \left( z^2 - 1 \right) \left( \frac{\gamma_X}{6} \right)^3 + z \left( \frac{\gamma_X}{6} \right)^4 - \frac{1}{3} \left( \frac{\gamma_X}{6} \right)^5
\]

[A]
Gamma density

\[ f_X(x; \alpha, \beta) = \frac{1}{\alpha \Gamma(\beta)} \left(\frac{x}{\alpha}\right)^{\beta-1} e^{-x/\alpha}, \quad 0 \leq x < +\infty \]

\[ \alpha = \frac{\sigma}{\sqrt{\beta}} > 0, \quad \beta = \left(\frac{2}{\gamma}\right)^2 > 0, \quad \mu = \alpha \beta = \sigma \sqrt{\beta} > 0 \]

\( \mu, \sigma, \) and \( \gamma \) are the mean, standard deviation, and skewness coefficient of \( X \) (or \( Y \)), respectively, and \( \alpha \) and \( \beta \) are respectively the scale and shape parameters of the gamma density.

\( \sigma = \frac{\mu \gamma}{2} \)
Assume two gamma random variables $X$ and $Y$ are jointly distributed.

The two random variables are respectively associated with their frequency factors $K_X$ and $K_Y$.

Equation (A) indicates that the frequency factor $K_X$ of a random variable $X$ with gamma density is approximated by a function of the standard normal deviate and the coefficient of skewness of the gamma density.
\[ K_T \approx z + \left( z^2 - 1 \right) \frac{\gamma_X}{6} + \frac{1}{3} \left( z^3 - 6z \right) \left( \frac{\gamma_X}{6} \right)^2 - \left( z^2 - 1 \right) \left( \frac{\gamma_X}{6} \right)^3 + z \left( \frac{\gamma_X}{6} \right)^4 - \frac{1}{3} \left( \frac{\gamma_X}{6} \right)^5 \]

Simulation of the frequency factor \( K_X \) can be achieved by generating a random sample of the standard normal deviate \( U \), say \( u_1, u_2, \ldots, u_n \), and then utilizing Eq. (A) to obtain \( k_{x_1}, k_{x_2}, \ldots, k_{x_n} \) from \( u_1, u_2, \ldots, u_n \).

However, for a bivariate gamma density \( f_{XY}(x, y) \), the two frequency factors \( K_X \) and \( K_Y \) are correlated through two correlated standard normal deviates \( U \) and \( V \), with a correlation coefficient \( \rho_{UV} \).
• Thus, random number generation of the second frequency factor $K_Y$ must take into consideration the correlation between $K_X$ and $K_Y$, which stems from the correlation between $U$ and $V$. 
Conditional normal density

- Given a random number of $U$, say $u$, the conditional density of $V$ is expressed by the following conditional normal density

$$
\phi_{V|U}(v \mid U = u) = \frac{1}{\sqrt{2\pi(1 - \rho_{UV}^2)}} \cdot \exp\left\{-\frac{1}{2} \left[ \frac{v - \rho_{UV} u}{\sqrt{1 - \rho_{UV}^2}} \right]^2 \right\}
$$

with mean $\rho_{UV} u$ and variance $1 - \rho_{UV}^2$. 
Thus, based on a random sample \( u_1, u_2, \ldots, u_n \) of \( U \), a random sample of \( V \), say \( v_1, v_2, \ldots, v_n \), can be generated by a normal random number generator with means \( \rho_{UV} u_i \) \( (i = 1, 2, \ldots, n) \) and variance \( 1 - \rho_{UV}^2 \).

\[
\phi_{V|U}(v \mid U = u) = \frac{1}{\sqrt{2\pi(1 - \rho_{UV}^2)}} \cdot \exp \left\{ -\frac{1}{2} \left[ \frac{v - \rho_{UV} u}{\sqrt{1 - \rho_{UV}^2}} \right]^2 \right\}
\]
\[ X = \mu_X + K\sigma_X \]

From the two sets of random samples \( u_1, u_2, \ldots, u_n \) and \( v_1, v_2, \ldots, v_n \), Eq. (A) can be used to obtain random samples of the two frequency factors \( K_X \) and \( K_Y \), i.e., \( k_{x_1}, k_{x_2}, \ldots, k_{x_n} \) and \( k_{y_1}, k_{y_2}, \ldots, k_{y_n} \).

Given the expected values (\( \mu_X \) and \( \mu_Y \)) and standard deviations (\( \sigma_X \) and \( \sigma_Y \)) of random variables \( X \) and \( Y \), random samples of the bivariate gamma distribution can be obtained by transferring \( k_{x_1}, k_{x_2}, \ldots, k_{x_n} \) and \( k_{y_1}, k_{y_2}, \ldots, k_{y_n} \) to \( x_1, x_2, \ldots, x_n \) and \( y_1, y_2, \ldots, y_n \).
Given the means, standard deviations and coefficients of skewness of two correlated gamma random variables \( X \) and \( Y \), i.e.,
\[
(\mu_X, \sigma_X, \gamma_X) \quad \text{and} \quad (\mu_Y, \sigma_Y, \gamma_Y)
\]
[Note: \( \sigma = \mu \gamma / 2 \)]
Also given the correlation coefficient of \( X \) and \( Y \), i.e., \( \rho_{XY} \)

Converting the correlation coefficient \( \rho_{XY} \) to \( \rho_{UV} \) using Eq. (10)

Generating a random sample \( u_1, u_2, \ldots, u_n \) of
\[
U \sim N(0,1)
\]
(Eq. 8)

Generating a random sample \( v_1, v_2, \ldots, v_n \) of
\[
V \sim N(0,1)
\]
(Eq. 2)
Flowchart of BVG simulation (2/2)

1. Generating a random sample $u_1, u_2, \ldots, u_n$ of $U \sim N(0,1)$  
   \[ \text{(Eq. 2)} \]
2. Calculating a random sample of $K_x$  
   \( (k_{x_1}, k_{x_2}, \ldots, k_{x_n}) \)  
   \[ \text{(Eq. 3)} \]
3. Calculating a random sample of $X$  
   \( (x_1, x_2, \ldots, x_n) \)

1. Generating a random sample $v_1, v_2, \ldots, v_n$ of $V \sim N(0,1)$  
   \[ \text{(Eq. 2)} \]
2. Calculating a random sample of $K_y$  
   \( (k_{y_1}, k_{y_2}, \ldots, k_{y_n}) \)  
   \[ \text{(Eq. 3)} \]
3. Calculating a random sample of $Y$  
   \( (y_1, y_2, \ldots, y_n) \)
In practice, stochastic simulation of a bivariate gamma distribution requires the generated random samples to have pre-specified mean, standard deviation, coefficient of skewness \((\mu_X, \sigma_X, \gamma_X)\) and \((\mu_Y, \sigma_Y, \gamma_Y)\), and correlation coefficient \(\rho_{XY}\). In order for the generated samples to meet such requirements, the correlation coefficient \(\rho_{UV}\) must be determined from the pre-specified \(\gamma_X, \gamma_Y, \) and \(\rho_{XY}\) through the following equation:
\( \rho_{XY} \sim \rho_{UV} \) Conversion

\[
\rho_{XY} \approx \left(A_X A_Y - 3A_X C_Y - 3C_X A_Y + 9C_X C_Y\right)\rho_{UV} + 2B_X B_Y \rho_{UV}^2 + 6C_X C_Y \rho_{UV}^3 \tag{B}
\]

\[
A_X = 1 + \left(\frac{\gamma_X}{6}\right)^4 \quad B_X = \frac{\gamma_X}{6} - \left(\frac{\gamma_X}{6}\right)^3 \quad C_X = \frac{1}{3}\left(\frac{\gamma_X}{6}\right)^2
\]

\[
A_Y = 1 + \left(\frac{\gamma_Y}{6}\right)^4 \quad B_Y = \frac{\gamma_Y}{6} - \left(\frac{\gamma_Y}{6}\right)^3 \quad C_Y = \frac{1}{3}\left(\frac{\gamma_Y}{6}\right)^2
\]
Derivation of the $\rho_{XY} \sim \rho_{UV}$ relationship

Suppose that random variables $X$ and $Y$ form a bivariate gamma distribution. Given the means ($\mu_X$ and $\mu_Y$) and standard deviations ($\sigma_X$ and $\sigma_Y$), $X$ and $Y$ can be respectively expressed in terms of their corresponding frequency factors $K_X$ and $K_Y$, i.e.,

$$X = \mu_X + K_X \sigma_X \quad \text{and} \quad Y = \mu_Y + K_Y \sigma_Y.$$  

Note that, with given means $\mu_X$ and $\mu_Y$ and standard deviations $\sigma_X$ and $\sigma_Y$, the coefficients of skewness $\gamma_X$ and $\gamma_Y$ can be readily determined.
From the above equations, it can be easily shown that frequency factors $K_X$ and $K_Y$ are distributed with zero mean and unit standard deviation, and correlation coefficient of $X$ and $Y$ is equivalent to correlation coefficient of $K_X$ and $K_Y$, i.e.,

$$E[K_X] = E[K_Y] = 0,$$

$$Var[K_X] = Var[K_Y] = 1,$$

and

$$\rho_{XY} = \rho_{K_XK_Y}.$$
• Frequency factors $K_X$ and $K_Y$ can be respectively approximated by

$$K_X \approx U + \left(U^2 - 1\right)\frac{\gamma_X}{6} + \frac{1}{3}\left(U^3 - 6U\right)\left(\frac{\gamma_X}{6}\right)^2 - \left(U^2 - 1\right)\left(\frac{\gamma_X}{6}\right)^3 + U\left(\frac{\gamma_X}{6}\right)^4 - \frac{1}{3}\left(\frac{\gamma_X}{6}\right)^5$$

$$K_Y \approx V + \left(V^2 - 1\right)\frac{\gamma_Y}{6} + \frac{1}{3}\left(V^3 - 6V\right)\left(\frac{\gamma_Y}{6}\right)^2 - \left(V^2 - 1\right)\left(\frac{\gamma_Y}{6}\right)^3 + V\left(\frac{\gamma_Y}{6}\right)^4 - \frac{1}{3}\left(\frac{\gamma_Y}{6}\right)^5$$

where $U$ and $V$ both are random variables with standard normal density and are correlated with correlation coefficient $\rho_{UV}$. 
• Correlation coefficient of $K_X$ and $K_Y$ can be derived as follows:

$$\rho_{K_XK_Y} = Cov(K_X, K_Y) = E[K_X K_Y]$$

$$\approx E\left\{ U + \left(U^2 - 1\right) \frac{\gamma_X}{6} + \frac{1}{3} \left(U^3 - 6U\right) \left(\frac{\gamma_X}{6}\right)^2 - (U^2 - 1) \left(\frac{\gamma_X}{6}\right)^3 + U \left(\frac{\gamma_X}{6}\right)^4 - \frac{1}{3} \left(\frac{\gamma_X}{6}\right)^5 \right\} \right.$$
\[ K_x \approx U \left[ 1 + \left(\frac{\gamma_x}{6}\right)^4 \right] + (U^2 - 1) \left[ \frac{\gamma_x}{6} - \left(\frac{\gamma_x}{6}\right)^3 \right] + \frac{1}{3} (U^3 - 6U) \left(\frac{\gamma_x}{6}\right)^2 - \frac{1}{3} \left(\frac{\gamma_x}{6}\right)^5 \]

\[ = A_x U + B_x (U^2 - 1) + C_x (U^3 - 6U) + D_x \]

\[ K_y \approx V \left[ 1 + \left(\frac{\gamma_y}{6}\right)^4 \right] + (V^2 - 1) \left[ \frac{\gamma_y}{6} - \left(\frac{\gamma_y}{6}\right)^3 \right] + \frac{1}{3} (V^3 - 6V) \left(\frac{\gamma_y}{6}\right)^2 - \frac{1}{3} \left(\frac{\gamma_y}{6}\right)^5 \]

\[ = A_y V + B_y (V^2 - 1) + C_y (V^3 - 6V) + D_y \]

\[ A_x = 1 + \left(\frac{\gamma_x}{6}\right)^4, \quad B_x = \frac{\gamma_x}{6} - \left(\frac{\gamma_x}{6}\right)^3, \quad C_x = \frac{1}{3} \left(\frac{\gamma_x}{6}\right)^2, \quad D_x = -\frac{1}{3} \left(\frac{\gamma_x}{6}\right)^5 \]

\[ A_y = 1 + \left(\frac{\gamma_y}{6}\right)^4, \quad B_y = \frac{\gamma_y}{6} - \left(\frac{\gamma_y}{6}\right)^3, \quad C_y = \frac{1}{3} \left(\frac{\gamma_y}{6}\right)^2, \quad D_y = -\frac{1}{3} \left(\frac{\gamma_y}{6}\right)^5 \]
\[ E[K_X K_Y] \approx \begin{bmatrix}
A_X A_Y U V + A_X B_Y U (V^2 - 1) + A_X C_Y U (V^3 - 6V) \\
A_X D_Y U + B_X A_Y V (U^2 - 1) + B_X B_Y (U^2 - 1)(V^2 - 1) \\
B_X C_Y (U^2 - 1)(V^3 - 6V) + B_X D_Y (U^2 - 1) \\
C_X A_Y (U^3 - 6U)V + C_X B_Y (U^3 - 6U)(V^2 - 1) \\
C_X C_Y (U^3 - 6U)(V^3 - 6V) + C_X D_Y (U^3 - 6U) \\
D_X A_Y V + D_X B_Y (V^2 - 1) + D_X C_Y (V^3 - 6V) + D_X D_Y
\end{bmatrix} \]
\[ E[K_X K_Y] \]
\[ \approx E \left[ A_X A_Y U V + A_X C_Y U (V^3 - 6V) + B_X B_Y (U^2 - 1) (V^2 - 1) \right. \\
\[ + C_X A_Y (U^3 - 6U) V + C_X C_Y (U^3 - 6U) (V^3 - 6V) + D_X D_Y \] \]

\[ E[K_X] \approx E \left[ A_X U + B_X (U^2 - 1) + C_X (U^3 - 6U) + D_X \right] = D_X \]

Since \( K_X \) and \( K_Y \) are distributed with zero means, it follows that
\[ D_X = D_Y \approx 0 \]
\[ \rho_{K_XK_Y} = E[K_XK_Y] \]

\[ \approx E \left[ A_XA_YUV + A_XC_YU(V^3 - 6V) + B_XB_Y(U^2 - 1)(V^2 - 1) 
+ C_XA_Y(U^3 - 6U)V + C_XC_Y(U^3 - 6U)(V^3 - 6V) \right] \]

\[ = A_XA_Y\rho_{UV} + A_XC_Y\left[ E(UV^3) - 6\rho_{UV} \right] + B_XB_Y\left[ E(U^2V^2) - 1 \right] \]

\[ + C_XA_Y\left[ E(U^3V) - 6\rho_{UV} \right] + C_XC_Y\left[ E(U^3V^3) - 6E(UV^3) - 6E(U^3V) + 36\rho_{UV} \right] \]
It can also be shown that

\[
E(U^2V^2) = 2\rho_{UV}^2 + 1 \quad E(U^3V^3) = 6\rho_{UV}^3 + 9\rho_{UV}
\]

Thus,

\[
E(U^3V) = E(UV^3) = 3\rho_{UV}
\]

Thus,

\[
\rho_{XY} = \rho_{KxK_Y} \approx \left(A_X A_Y - 3A_X C_Y - 3C_X A_Y + 9C_X C_Y\right)\rho_{UV}
\]

\[
+ 2B_X B_Y \rho_{UV}^2 + 6C_X C_Y \rho_{UV}^3
\]
Equation (B) indicates that $\rho_{XY}$ can be expressed as a third order polynomial of $\rho_{UV}$. It is therefore of practical concern whether there exits a unique $\rho_{UV}$ for a given set of $(\gamma_X, \gamma_Y, \rho_{XY})$. Or equivalently, given a set of $(\gamma_X, \gamma_Y, \rho_{XY})$, does Eq. (B) return a single-value of $\rho_{UV}$?
\( \rho_{XY} \sim \rho_{UV} \)  
Single - Value Relationship ip

We have also proved that Eq. (B) is indeed a single-value function.
Proof of Eq. (B) as a single-value function

Let \( f(\rho_{UV}) = \partial \rho_{XY} / \partial \rho_{UV} \). From Eq. (B) we have

\[
\begin{align*}
    f(\rho_{UV}) &= (A_X A_Y - 3A_X C_Y - 3C_X A_Y + 9C_X C_Y) \\
                  &+ 4B_X B_Y \rho_{UV} + 18C_X C_Y \rho_{UV}^2 \\
    A_X A_Y - 3A_X C_Y - 3C_X A_Y + 9C_X C_Y &= (A_X - 3C_X)(A_Y - 3C_Y)
\end{align*}
\]

\[
A_X - 3C_X = 1 + \left(\frac{\gamma_X}{6}\right)^4 - \left(\frac{\gamma_X}{6}\right)^2 = \left[\left(\frac{\gamma_X}{6}\right)^2 - 1\right]^2 + \left(\frac{\gamma_X}{6}\right)^2 > 0
\]

\[
A_Y - 3C_Y = 1 + \left(\frac{\gamma_Y}{6}\right)^4 - \left(\frac{\gamma_Y}{6}\right)^2 = \left[\left(\frac{\gamma_Y}{6}\right)^2 - 1\right]^2 + \left(\frac{\gamma_Y}{6}\right)^2 > 0
\]
• Therefore,

\[ A_X A_Y - 3A_X C_Y - 3C_X A_Y + 9C_X C_Y \]

\[ = \left\{ \left[ \left( \frac{\gamma_X}{6} \right)^2 - 1 \right]^2 + \left( \frac{\gamma_X}{6} \right)^2 \right\} \left\{ \left[ \left( \frac{\gamma_Y}{6} \right)^2 - 1 \right]^2 + \left( \frac{\gamma_Y}{6} \right)^2 \right\} \]
Let \( g(\rho_{UV}) = 4B_X B_Y \rho_{UV} + 18C_X C_Y \rho_{UV}^2 \)

\[
= \frac{\gamma_X \gamma_Y}{9} \left[ \left( \frac{\gamma_X}{6} \right)^2 - 1 \right] \left[ \left( \frac{\gamma_Y}{6} \right)^2 - 1 \right] \rho_{UV} + \frac{18}{9} \left( \frac{\gamma_X}{6} \right)^2 \left( \frac{\gamma_Y}{6} \right)^2 \rho_{UV}^2
\]

Also, let \( G_X = \frac{\gamma_X}{6} \), \( G_Y = \frac{\gamma_Y}{6} \). We then have

\[
g(\rho_{UV}) = 4G_X G_Y (G_X^2 - 1)(G_Y^2 - 1) \rho_{UV} + 2G_X^2 G_Y^2 \rho_{UV}^2
\]
\[ f(\rho_{UV}) = \left[(G_X^2 - 1)^2 + G_X^2\right]\left[(G_Y^2 - 1)^2 + G_Y^2\right] + 4G_X G_Y (G_X^2 - 1)(G_Y^2 - 1)\rho_{UV} + 2G_X^2 G_Y^2 \rho_{UV}^2 \]

\[ \frac{\partial f}{\partial \rho_{UV}} = 4G_X G_Y (G_X^2 - 1)(G_Y^2 - 1) + 4G_X^2 G_Y^2 \rho_{UV} \]

Let \( \frac{\partial f}{\partial \rho_{UV}} = 0 \), it yields an extreme value of \( f \) at

\[ \rho_{UV}^* = -\frac{(G_X^2 - 1)(G_Y^2 - 1)}{G_X G_Y} \]
$$f(\rho_{UV}^*) = \left[ (G_X^2 - 1)^2 + G_X^2 \right] \left[ (G_Y^2 - 1)^2 + G_Y^2 \right]$$
$$- 4(G_X^2 - 1)^2 (G_Y^2 - 1)^2 + 2(G_X^2 - 1)^2 (G_Y^2 - 1)^2$$
$$= \left[ (G_X^2 - 1)^2 + G_X^2 \right] \left[ (G_Y^2 - 1)^2 + G_Y^2 \right] - 2(G_X^2 - 1)^2 (G_Y^2 - 1)^2$$
$$= G_X^2 (G_Y^2 - 1)^2 + (G_X^2 - 1)^2 G_Y^2 + G_X^2 G_Y^2 - (G_X^2 - 1)^2 (G_Y^2 - 1)^2$$

Since $$-1 \leq \rho_{UV} \leq 1$$ (or equivalently, $$\rho_{UV}^2 \leq 1$$), it yields

$$(G_X^2 - 1)^2 (G_Y^2 - 1)^2 \leq G_X^2 G_Y^2$$

Thus,

$$f(\rho_{UV}^*) \geq G_X^2 (G_Y^2 - 1)^2 + (G_X^2 - 1)^2 G_Y^2 > 0$$
We now check the second derivative of $f(\rho_{UV})$, i.e.,

$$\frac{\partial^2 f}{\partial (\rho_{UV})^2} = 4G_x^2G_y^2 > 0$$

Therefore, $f(\rho_{UV}^*) > 0$ is the minimum of the function $f(\rho_{UV}) = \partial \rho_{XY} / \partial \rho_{UV}$. It follows that $f(\rho_{UV}) = \partial \rho_{XY} / \partial \rho_{UV} > 0$ for all possible values of $\rho_{UV}$. 
• The above equation indicates $\rho_{XY}$ increases with increasing $\rho_{UV}$, and thus Eq. (B) is a single-value function.
**Conceptual description of a gamma random field simulation approach**

Given a pair of bivariate gamma random variables \((X,Y)\) with known properties:

\[ \mu_X, \mu_Y, \gamma_X, \gamma_Y, \rho_{XY} \]

- Converting \(\rho_{XY}\) to \(\rho_{UV}\) where \((U,V)\) represents a pair of bivariate standard normal variables.

Given a homogeneous and isotropic random field \(Z(x)\) with known gamma density and covariance function \(C_Z(h)\) or variogram \(\gamma_Z(h)\):

- Converting \(C_Z(h)\) to \(C_W(h)\) where \(W(x)\) is a random field with standard normal density and covariance function \(C_W(h)\).
Generating a random sample of \((U, V)\) with sample size \(n\), i.e. \(\{(u_i, v_i), i = 1, \ldots, n\}\).

Individually and independently converting \(u_i\) to \(x_i\) and \(v_i\) to \(y_i\). The resultant \(\{(x_i, y_i), i = 1, \ldots, n\}\) is a random sample of the bivariate random variables \((X, Y)\).

Generating a realization of \(W\), i.e. \(\{w(i, j), i = 1, \ldots, n ; j = 1, \ldots, m\}\) where \((i, j)\) represents a spatial location and \(n\) and \(m\) defines the extent of the spatial domain.

Individually and independently converting \(w(i, j)\) to \(z(i, j)\). The resultant \(\{z(i, j), i = 1, \ldots, n ; j = 1, \ldots, m\}\) is a realization of the random field \(Z(x)\).