Developments in deriving Precipitation Frequency and PMP in China

"Key issues of engineering hydrology"

Prof. Bingzhang Lin
lbz@nuist.edu.cn

Applied Hydrometeorological Research Institute (AHMRI)、
College of Hydrometeorology (COHM)
Nanjing University of Information Science & Technology (NUIST), CHINA

Presentation for ESEWG meeting at NWS
15 August, 2013 Silver Spring, MD, USA
Nanjing University of Information Science & Technology

南京信息工程大学 (www.nuist.edu.cn)
What are the key issues in Engineering Hydrologic Studies?

---Hydrologic frequency & PMP estimation in China
100 years of AMS rainfall data at a station in the U.S.

(011901) 28-4229 ANMAX

100

3.46  2.60  5.65  4.90  3.02  7.15  2.97  2.86  2.70  3.18
3.39  2.13  2.34  2.61  4.39  2.04  3.06  3.92  3.00  3.67
1.97  1.45  2.78  2.73  3.98  1.98  2.17  2.00  1.80  2.08
1.72  2.61  3.10  3.73  2.82  2.17  2.10  6.78  6.05  4.55
1.92  3.20  3.50  3.45  2.78  3.33  1.42  2.62  3.23  4.20
2.42  3.23  2.22  4.25  3.21  4.02  1.37  4.16  1.81  5.05
2.44  3.84  2.80  2.29  3.35  3.18  5.06  2.41  3.58  2.10
4.29  3.80  3.78  6.63  2.77  3.42  2.25  2.71  2.84  2.35
2.83  1.90  2.63  2.46  4.68  4.10  2.51  1.95  2.59  4.71
4.94  3.88  2.81  2.51  2.85  3.72  2.91  2.26  4.88  2.63
100 years of AMS rainfall data

(Sorted + Grouped)

(011901) 28-4229 ANMAX

100

<table>
<thead>
<tr>
<th>1.55</th>
<th>1.61</th>
<th>1.64</th>
<th>1.95</th>
<th>2.04</th>
<th>2.05</th>
<th>2.15</th>
<th>2.18</th>
<th>2.21</th>
<th>2.23</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.25</td>
<td>2.27</td>
<td>2.31</td>
<td>2.36</td>
<td>2.38</td>
<td>2.38</td>
<td>2.42</td>
<td>2.46</td>
<td>2.46</td>
<td>2.52</td>
</tr>
<tr>
<td>2.55</td>
<td>2.56</td>
<td>2.60</td>
<td>2.65</td>
<td>2.66</td>
<td>2.73</td>
<td>2.74</td>
<td>2.77</td>
<td>2.79</td>
<td>2.85</td>
</tr>
<tr>
<td>2.85</td>
<td>2.94</td>
<td>2.95</td>
<td>2.96</td>
<td>2.96</td>
<td>2.97</td>
<td>2.98</td>
<td>2.98</td>
<td>3.06</td>
<td>3.07</td>
</tr>
<tr>
<td>3.10</td>
<td>3.14</td>
<td>3.15</td>
<td>3.15</td>
<td>3.18</td>
<td>3.19</td>
<td>3.20</td>
<td>3.21</td>
<td>3.22</td>
<td>3.23</td>
</tr>
<tr>
<td>3.64</td>
<td>3.66</td>
<td>3.66</td>
<td>3.78</td>
<td>3.80</td>
<td>3.84</td>
<td>3.88</td>
<td>3.91</td>
<td>3.92</td>
<td>3.97</td>
</tr>
<tr>
<td>4.06</td>
<td>4.16</td>
<td>4.22</td>
<td>4.23</td>
<td>4.29</td>
<td>4.31</td>
<td>4.35</td>
<td>4.40</td>
<td>4.45</td>
<td>4.51</td>
</tr>
<tr>
<td>4.56</td>
<td>4.65</td>
<td>4.72</td>
<td>4.76</td>
<td>4.82</td>
<td>4.86</td>
<td>4.98</td>
<td>5.16</td>
<td>5.31</td>
<td>5.34</td>
</tr>
<tr>
<td>5.53</td>
<td>5.56</td>
<td>5.60</td>
<td>5.73</td>
<td>5.74</td>
<td>6.41</td>
<td>6.86</td>
<td>7.52</td>
<td>7.69</td>
<td>8.11</td>
</tr>
</tbody>
</table>
Histogram

Empirical frequency curve

AMS Histogram

AMS Cumulative Frequency Plot

Increment = 0.5 in.

Fig. 1 AMS histogram and empirical frequency curve
Fig. 2 Probability density function curve
What is Frequency Analysis?

Frequency Analysis -- is a statistical approach using sample or samples to estimate the population probability distribution.
What are the Essential Issues of FA?

1. Precision
2. Accuracy
Nature of the Frequency Analysis (1)

- Like shooting practice: Precision and Accuracy (e.g. firing 10 shots)

Good precision
Poor accuracy
Nature of the Frequency Analysis (2)

• Like shooting practice: Precision and Accuracy

Poor precision
Moderate accuracy
Nature of the Frequency Analysis (3)

- Like shooting practice: Precision and Accuracy

Good precision
Good accuracy
Two Impossible Things in FA

1. Theoretical true value of frequencies is unknown forever. (100-year ?)

2. There is no analytical way to derive a theoretical distribution to best fit the data. (GLO or GEV or PE3 ?)
Where is the bull’s-eye?

- The true value of frequencies such as 100-year is unknown forever.

Something like bulls-eye unknown while shooting.
Global Climate Change Makes the Issue more Complicated

Let data talk!
Location of Daily Stations in OH Study Area

Fig. 3 Location of raingauges in Ohio River Basin
Only 16% or 1797 sites exhibited linear trend in mean for AMS in Ohio River Basin in past century

Fig. 4 Map of stations that exhibit linear trend in mean for AMP over 20th century
Only 18% or 531 sites exhibited shift in mean for AMS in past century

(compression of prior- to post- 1958 )

Fig. 5 Map of stations that exhibit shift in mean for AMP over the 20th century
Investigation of 1,741 daily AMS in Mainland China

Fig. 6 Location of raingauges in China (partially)
Only 17.5% of 2,436 sites exhibited linear trend in mean of AMS in China

Fig. 7 Spatial distribution of raingauges with linear trend in China
Only 10.8% of 1,649 sites exhibited shift in mean for AMS in China

Fig. 8 Spatial distribution of raingauges with shift in China
Findings

- Generally speaking, there was no obvious linear trend and shift in mean for daily AMS in Ohio of the U.S. and in China in the past century;

- However, there was more than 50% of tested sites that exhibited a clear increase in variance of daily AMS in OH River Basin, with SD increased by 23% for the latter half- to the former half- century. **What does it imply?**
It implies:

We may observe more and more extreme hydrometeorological events (droughts or floods) in the Ohio River Basin area in the near future than before though their mean does not change. So the world.
Facing with the acceleration of climate change,

What is our job?
One of Our Mission

Exploration of a Robust and Reliable Approach to Performing Precipitation Frequency Analysis of Extreme Events.
What we should/can do? (1)

L-Moments Method -- focusing
on the issue of precision
in terms of parameter estimation
Methods of Parameter Estimation

1. Conventional Moments Method (CMM)

\[ M_{p,0,0} = E[X^p] = \int_0^1 x^p dF(x) \]

2. L-Moments Method (LMM)

\[ M_{1,r,0} = E[X\{F(X)^r\}] = \int_0^1 x\{F(x)^r\} dF(x) \]

Advantage – Linearity

Power shift
L-Moments

Definition: L-moments are expectations of certain linear combinations of order statistics (Hosking, 1989)

\[
\lambda_r \equiv r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E[X_r - k: r], \quad r = 1, 2, \ldots
\]

\[
\lambda_1 = EX
\]
\[
\lambda_2 = \frac{1}{2} E(X_{2:2} - X_{1:2})
\]
\[
\lambda_3 = \frac{1}{3} E(X_{3:3} - 2X_{2:3} + X_{1:3})
\]
\[
\lambda_4 = \frac{1}{4} E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4})
\]

(When \( X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n} \))
Application of RLM to Precipitation Frequency Analysis in the U.S.

Fig. 9  Pennsylvania data will be used for comparison
Comparison of CMM to LMM on Biasedness

Data Cs vs. Simulated
(Pennsylvania 228 daily max, fit GE)

Number of simulations: 1000

Mean Simulated Cs vs. Data Cs

Data L-Cs vs. Simulated
(Pennsylvania 228 daily max, fit GE)

Number of simulations: 1000

Mean Simulated L-Cs vs. Data L-Cs

High biasedness of Cs          Much less biasedness of L-Cs

Fig. 10  Comparison for biasedness between CMM and LMM based on PA data
Comparison of CMM to LMM on Robustness

CMM plays poorly to outlier (10.37”/day on 7/22/1947 at #2682, PA)
Cs cannot model this outlier even for N = 500 yr.

Data Cs vs. Simulated Cs
(Pennsylvania 228 daily max, fit GEV)

Number of simulations: 1000
Data N = 500 yr. for all stations

Data L-Cs vs. Simulated L-Cs
(Pennsylvania 228 daily max, fit GEV)

Number of simulations: 1000
Data N = 500 yr. for all stations

CMM: poor modeling to outlier
LMM: very well modeling to outlier

Fig. 11 Comparison for robustness between CMM and LMM based on PA data
The same findings on biasedness and robustness have been drawn for annual AMS precipitation data in China.
Location of 96 raingauges in Taihu Lake (6,134 km²)

Fig. 13 Map of Taihu Lake Basin with distribution of 96 raingauges
(96+45) Sites of Taihu Lake divided into 8 HGS Regions

Fig. 14 Regionalization of Taihu Lake Basin
Comparison for Taihu Lake Data (1) on Biasedness

Fig. 1 Biasedness on skewness of CMM

Fig. 2 Unbiasedness on skewness of LMM

Fig. 15 Comparison for biasedness between CMM and LMM based on Taihu data
Comparison for Taihu Lake Data (2) on Robustness

Fig. 3 Difficult to model the outlier by CMM

Fig. 4 Robustness to outlier by LMM

Fig. 16 Comparison for robustness between CMM and LMM based on Taihu data
More comparisons

• CMM is less sensitive to screen the data than the LMM does in terms of statistical characteristics.

• As a result, the *Pearson Type III* has been officially selected for fitting data (rainfall & streamflow) in China since 1950s because the CMM has widely been adopted in design studies there.
PE3 is the best-fit when CMM used (1-1)

**Relationship of Skewness vs. Kurtosis**
(The Ohio River Basin: 1-day/24-hour)

Based on 4253 stations in the basin

**Fig. 17** Diagram of Cs vs. Ck for daily stations in the OH area (all points)
PE3 is the best-fit when CMM used (1-2)

Relationship of Skewness vs. Kurtosis
(The Ohio River Basin: 1-day/24-hour)

Based on 4253 stations in the area

Mean (Skewness, Kurtosis)
(1.9666, 8.1912)

Fig. 18 Diagram of Cs vs. Ck for daily stations in the OH area (mean point)
GEV is the best-fit when LMM used (2-1)

Fig. 19 Diagram of L-Cs vs. L-Ck for daily stations in the OH area (all points)
GEV is the best-fit when LMM used

![Diagram of L-Cs vs. L-Ck for daily stations in the OH area (mean point)](image)

**RELATIONSHIP OF L-CS vs. L-CK**
(The Ohio River Basin: 24-hour)

Based on 4253 stations in the basin

Mean (L-cs, L-ck) = (0.3549, 0.2399)

Mean (L-cs, L-ck) = (0.3549, 0.2399)

Page 42  Fig. 20  Diagram of L-Cs vs. L-Ck for daily stations in the OH area (mean point)
What we should/can do? (2)

Regional Analysis – focusing on the issue of **accuracy** in terms of uncertainties of quantiles
Assume: A rainfall could be decomposed into the regional component reflecting the common characteristics in the region and the local component reflecting the individual local characteristics.

\[ Q_{T,i,j} = q_{T,i} \times \bar{x}_{i,j} \]

\[ q_{T,i} = \frac{Q_{T,i,j}}{\bar{x}_{i,j}} \]
In Homogeneous Regions
(Assumption & major procedure)

1. A rainfall could be decomposed into common (related to common characteristics) & local (related to locality) components;
2. The common components at all sites in a region are identically distributed to fit a dimensionless probability distribution at length-weighted;
3. The regionally dimensionless probability distribution is then applied to each individual site by “superposition” to get their own probability distribution;
4. Calculate the quantiles at each individual site.
Regionalization → Reduces quantile uncertainty

Reduced uncertainties in terms of confidence interval (Regional analysis provides more stable estimates)

**Fig. 21** Comparison for uncertainties of quantiles between the at-site FA and regional FA

<table>
<thead>
<tr>
<th>Return Periods (years)</th>
<th>Single Station</th>
<th>Regional</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-y</td>
<td>0.0314</td>
<td>0.0353</td>
</tr>
<tr>
<td>5-y</td>
<td>0.0391</td>
<td>0.0363</td>
</tr>
<tr>
<td>10-y</td>
<td>0.0499</td>
<td>0.0376</td>
</tr>
<tr>
<td>25-y</td>
<td>0.0727</td>
<td>0.0411</td>
</tr>
<tr>
<td>50-y</td>
<td>0.0961</td>
<td>0.0449</td>
</tr>
<tr>
<td>100-y</td>
<td>0.1211</td>
<td>0.0497</td>
</tr>
<tr>
<td>200-y</td>
<td>0.1501</td>
<td>0.0554</td>
</tr>
<tr>
<td>500-y</td>
<td>0.1927</td>
<td>0.0642</td>
</tr>
<tr>
<td>1000-y</td>
<td>0.2278</td>
<td>0.0715</td>
</tr>
</tbody>
</table>
Status of FA in China

1. **Current** – *Subjective, Limited info*
   CMM + （single, one duration, fixed P-III）+ G-O-F by eyes

2. **Projected** – *Objective, Full of info*
   LMM + （regionalization, multi-duration, more distributions）+ G-O-F by criteria
Relevant Topics for FA

- Advantages of the RLA (Regional L-moments Analysis)
- Criteria to identify homogeneous regions
- Goodness-of-fit
- Real-data-check
- Consistency adjustments over time & space
- Intersite dependence
- Uncertainties of quantiles – conf. intervals
- Sampling methods – AMS or PDS or AES
Online FA Deliverables (1)

Example

Taihu Lake Basin

Fig. 22 Home page of online FA products for Taihu Lake Basin
Online FA Deliverables (3)

A result-window pop-up

Fig. 24  Online FA products for Taihu Lake Basin (2)
Online FA Deliverables (4)

Fig. 25  Online FA products for Taihu Lake Basin (3)
Precipitation FA for **Guangxi** in SW China (1)

Example

**Gx 1h Distribution**

Legend
- 1h Stations
- Gx Boundary
- **Gx Value**
  - High: 6050
  - Low: -245

Fig. 26  Online FA products for Guangxi in SW China (1)
Precipitation FA for **Guangxi** in SW China (2)

**Example**

Fig. 27 Online FA products for Guangxi in SW China (2)
Two Tricky things in FA

1. "Theoretical is Empirical"
   --Selection of a theoretical probability distribution to fit the data is empirical.

2. "Empirical is theoretical"
   --Position formula \( \frac{m}{n+1} \) to plot an empirical curve in a diagram is theoretical (meaning it can be analytically derived).
Second mission
Another Mission
--Explore a meaningful upper limit of rainfall

In China PMP estimation is required for design studies for large infrastructure (big dams, nuclear power stations) as regulatory standards in terms of flood-control, and flood-mitigation planning for large cities as well.

Probable Maximum Precipitation
-- Hydrometeorologically causally causal approach
Starting in 1930s — The U.S. Weather Bureau first introduced the upper limit to precipitation (Late Prof. Ven Te Chow used it)

Maximum Possible Precipitation (MPP) → Probable Maximum Precipitation (PMP)
Definition of PMP

Probable maximum precipitation (PMP) is defined as the greatest depth of precipitation for a given duration meteorologically possible for a design watershed\(^{(1)}\) or a given storm area\(^{(2)}\) at a particular location at a particular time of Year\(^{(3)}\), with no allowance made for long-term climate trends.


*Comments: (1) and (2) are not equivalent; (3) is irrelevant.
--By Prof. B Lin*
### PMP Estimation Methodology -- International Practice

In general, mainly two types of approaches in design practice of PMP studies: I. Hydrometeorological (HYDROME) & II. Statistical (STAT)

<table>
<thead>
<tr>
<th>(I-a) Moisture maximization</th>
<th>Maximum 12-hr persisting dew point (HYDROME)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(I-b) Storm transposition</td>
<td>Storm Separation + Adjustments (HYDROME)</td>
</tr>
<tr>
<td>(I-c) Use of D-A-D curves</td>
<td>Envelopment (HYDROME)</td>
</tr>
<tr>
<td>(II) Statistical approach</td>
<td>Modified frequency analysis (STAT)</td>
</tr>
</tbody>
</table>
The most popular means – Storm Transposition (ST)

The most difficult job is how to take ST in terrain area – orographic effects?

Key: Storm separation

- Convergence component
- Orographic component
Laminar Model

- $R$ is the rainfall rate $\text{mm/s}$; $v_1$ is the inflow wind speed $\text{m/s}$; $\Delta P$ is the inflow pressure difference $\text{hPa}$; $\overline{q}_1$ and $\overline{q}_2$ are the inlet and outlet average specific humidities $\text{g/kg}$; $g$ is the acceleration due to gravity $\text{cm/s}^2$; $\rho$ is the density of water $\text{g/cm}^3$; $Y$ is the horizontal depth of the hill $\text{m}$.

$$R = \frac{v_1 \Delta P_1 (\overline{q}_1 - \overline{q}_2)}{Y} \cdot \frac{1}{g \rho}$$

(*NWS once applied unsuccessfully the Laminar Model to account the terrain effect on rainfall in a PMP study in 1961 and 1966*)
For a storm rainfall, the rainfall intensity for a given point \( P(x,y) \) in a drainage at any time can be defined by

\[
I(x,y,t) = I_0(x,y,t) \times f(x,y,t)
\]

(Hence, the area-averaged rainfall \( R \) for the whole drainage area of \( A \) during the period of time is given below:

\[
\overline{R}_{A} = \frac{\iint_{A} r_{\Delta}(x,y)\,dxdy}{\iint_{A} dxdy} = \frac{\iint_{A} r_{0,\Delta}(x,y) \times f_{\Delta}(x,y)\,dxdy}{\iint_{A} dxdy} \approx \frac{\sum_{i}^{m} \sum_{j}^{n} r_{0,\Delta}(x_i,y_j) \times f_{\Delta}(x_i,y_j)\,dxdy}{\sum_{i}^{m} \sum_{j}^{n} \Delta x_i \Delta y_j}
\]

(* Lin, Bingzhang, WMO NO-1045, Geneva, 2009)
10 Tasks for the PMP Study in HK

- Inception report
- Historical rainfall data acquisition (SE China)
- Storm survey/selection and analysis
- Transposition analysis of selected storms
- Synoptic analysis + apply the storm separation technique (SDOIFs)
- Statistical estimation
- Orientation + transposition adjustment
- Development of DAD with moisture maximization
- Impact of climate change/Long-term trends in rainfall extremes
- Comprehensive comparisons
- Final report
Procedures

Selection of major storms in south & southeast China region

Determination of geographical coverage

Collection of rainfall data

Stage 1

Storm data

Historical data

Stage 2

Analysis of Target Storm

Transposition Analysis

Isohyet analysis

Applying SDOIF method

Convergence component

Local SDOIF

Embryonic PMP

Storm separation

Statistical analysis

PMP "lower" / "upper" bounds

Example

Fig. 29 Sketch of PMP estimation procedure for HK as example
Moving track of Morakot Typhoon

Fig. 30  Track of Morakot Typhoon, 2009-08-03~11
## World, China mainland and Taiwan rainfall (By 2009)

<table>
<thead>
<tr>
<th>Location</th>
<th>Date</th>
<th>Duration</th>
<th>Amount</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shinliao, Taiwan (north)</td>
<td>October 17, 1967</td>
<td>24-hour</td>
<td>1,672 mm</td>
<td>Taiwan record</td>
</tr>
<tr>
<td>Jiayi Alishan, Taiwan</td>
<td>August 8, 2009</td>
<td>24-hour</td>
<td>1,623 mm</td>
<td>Typhoon Morakot</td>
</tr>
<tr>
<td>Linzhuang, China mainland</td>
<td>August 7, 1975</td>
<td>24-hour</td>
<td>1,060 mm</td>
<td>Sup. Typhoon Nina</td>
</tr>
<tr>
<td>Cilaos, La Reunion Island</td>
<td>March 15, 1952</td>
<td>24-hour</td>
<td>1,870 mm</td>
<td>Tropical Cyclone</td>
</tr>
<tr>
<td>Jiayi Alishan, Taiwan</td>
<td>August 8-9, 2009</td>
<td>48-hour</td>
<td>2,361 mm</td>
<td>Typhoon Morakot</td>
</tr>
<tr>
<td>Linzhuang, China mainland</td>
<td>August 7-8, 1975</td>
<td>48-hour</td>
<td>1,279 mm</td>
<td>Sup. Typhoon Nina</td>
</tr>
<tr>
<td>Cilaos, La Reunion Island</td>
<td>March 15-17, 1952</td>
<td>2-day</td>
<td>2,500 mm</td>
<td>Tropical Cyclone</td>
</tr>
<tr>
<td>Jiayi Alishan, Taiwan</td>
<td>August 8-10, 2009</td>
<td>3-day</td>
<td>2,747 mm</td>
<td>Typhoon Morakot</td>
</tr>
<tr>
<td>Linzhuang, China mainland</td>
<td>August 6-8, 1975</td>
<td>3-day</td>
<td>1,605 mm</td>
<td>Sup. Typhoon Nina</td>
</tr>
<tr>
<td>Cilaos, La Reunion Island</td>
<td>March 15-18, 1952</td>
<td>3-day</td>
<td>3,240 mm</td>
<td>Tropical Cyclone</td>
</tr>
<tr>
<td>Grand-Ilet, La Reunion Island</td>
<td>January 24-27, 1980</td>
<td>3-day</td>
<td>3,241 mm</td>
<td>Cyclone Hyacinthe</td>
</tr>
<tr>
<td>Commerson's Crater, La Reunion Island</td>
<td>February 24-26, 2007</td>
<td>3-day</td>
<td>3,929 mm</td>
<td>Cyclone Gamede</td>
</tr>
</tbody>
</table>
Issue of the Typhoon Morakot (1)
(1,000-year plus event?)

Lessons/findings learnt from the Morakot:

- Total rainfall >> historical records in the Mainland China
- 24-hr rainfall >> 1,000-year estimate in U.S. and PRVI
- 24-hr rainfall ~ 24-hr PMP estimate on Hainan Island
- 2-day rainfall ~ the world record
Possible causes:

- High topography of local terrain (rainfall intensification mechanism)
- Typhoon circulation
- Strong south-westerly monsoon system

However, current typhoon intensity forecast skill is still poor because of the lack of understanding on the complex interactions between ocean and typhoons.
Isohyets of 24-hr for Morakot Typhoon

- 1,583mm / 24-hr
- 2,372mm / 48-hr
- 2,682mm / 72-hr

(based on hourly rainfall observations)
Moisture Flux of Morakot Storm

Development of the SDOIF for the Target Area

- Major Moisture Flux during Morakot (left)

- Power Spectrum of the WNPSM Index (right)

(After Chi–Cherng Hong, Taipei Municipal University of Education, Taipei, Taiwan)

Fig. 33 Model simulation of moisture flux for Morakot Typhoon
Fig. 34 Location of raingauges in Taiwan for analysis of OIF
Gridded 24h OIF for Morakot (10kmx10km)

Orographic Intensification Factor

Fig. 35 Map of 24-hr gridded OIF for Morakot Storm
Before & After Storm Separation

Fig. 36 Before and after decomposition of 24-hr Morakot Storm rainfall
Generalized Convergence Component Pattern

(based on synoptic analysis of 4 major storms)

Fig. 37  Generalized convergence component pattern of Morakot Storm
Convergence Pattern can be Transposed in a Wider Region
Application example: Transposed to HK

Fig. 39 Illustration of synoptic analysis for HK in terms of moisture flux
Development of OIF in Design Area, HK

(5km x 5km)

24 hour - Orographic Intensification Factors

<table>
<thead>
<tr>
<th>113.9</th>
<th>114.0</th>
<th>114.1</th>
<th>114.2</th>
<th>114.3</th>
<th>114.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.34</td>
<td>1.33</td>
<td>1.31</td>
<td>1.28</td>
<td>1.23</td>
<td>1.17</td>
</tr>
<tr>
<td>1.36</td>
<td>1.35</td>
<td>1.32</td>
<td>1.25</td>
<td>1.16</td>
<td>1.09</td>
</tr>
<tr>
<td>1.38</td>
<td>1.41</td>
<td>1.45</td>
<td>1.48</td>
<td>1.37</td>
<td>1.17</td>
</tr>
<tr>
<td>1.36</td>
<td>1.39</td>
<td>1.45</td>
<td>1.46</td>
<td>1.23</td>
<td>1.14</td>
</tr>
<tr>
<td>1.25</td>
<td>1.23</td>
<td>1.20</td>
<td>1.15</td>
<td>1.21</td>
<td>1.30</td>
</tr>
<tr>
<td>1.19</td>
<td>1.17</td>
<td>1.14</td>
<td>1.12</td>
<td>1.13</td>
<td>1.34</td>
</tr>
<tr>
<td>1.15</td>
<td>1.13</td>
<td>1.10</td>
<td>1.07</td>
<td>1.04</td>
<td>1.05</td>
</tr>
<tr>
<td>1.12</td>
<td>1.10</td>
<td>1.07</td>
<td>1.05</td>
<td>1.02</td>
<td>1.06</td>
</tr>
</tbody>
</table>
Convert the Convergence into gridded

• Convert the convergence component of Morakot into a gridded frame like the gridded SDOIF

• Then cut-off the center piece to match the HK area size
E-W orientation with different center points (as example)

(Peak at Lantau)

(Peak at Tai Mo Shan)

Fig. 41 Orientation of transposed convergence pattern (1)
E-W orientation with different center points (as example)

(Peak at Lantau)

(Peak at Tai Mo Shan)

Fig. 42 Orientation of transposed convergence pattern (2)
Calculation of Embryonic PMP

- Distribution of Gridded-average Embryonic PMP in HK

Fig. 43 Gridded embryonic PMP for HK (unfinished; example only)
More orientations
Orientation 22.5° – 24hr
(Embryonic PMP)

Centered at Lantau

Centered at Tai Mo Shan

Fig. 44 Orientation of transposed convergence pattern (3)
NE-SW orientation – 22.5°

(Embryonic PMP)

Centered at Tai Mo Shan

Centered at Lantau

Fig. 45 Embryonic PMP for HK-1 (example only)
Orientation 45° – 24hr
(Embryonic PMP)

Centered at Lantau

Centered at Tai Mo Shan

Fig. 46   Orientation of transposed convergence pattern (4)
NE-SW orientation – 45°

Embryonic PMP

Centered at Tai Mo Shan

24 hour - PMP Estimates

Centered at Lantau

24 hour - PMP Estimates

Fig. 47  Embryonic PMP for HK-2 (example only)
Depth-Area-Duration Curves

Master DAD for Top 20 HK storms after Maximization

Example

Fig. 48  DAD curve-1 (example only)
Depth-Area-Duration Curves

24hr PMP

Example
Moisture Maximization

Ratio of Moisture Maximization:

Ratio of moisture maximization for transposition

\[ r = \frac{W_m}{W_r} = \frac{W_{27.4}}{W_{24}} = \frac{99.6}{74.0} = 1.346 \]

in which, \( W_m \) is the historical maximum consistent dew point which is \( 27.4^\circ C \) in design area while the \( W_r \) is the representative dew point, \( 24.0^\circ C \), for the Morakot storm in target area.
Statistical Approach (1)

PMP, $X_{PMP}$

$$X_{PMP} = \bar{X}_n + K_m \times S_n$$

where

- $\bar{X}_n$ the mean of the n maxima
- $S_n$ the standard deviation of the n maxima

Fig. 50 Sketch for statistical approach to PMP (1)
“Modified” Frequency Analysis

• “standard deviation”, $K_m$, is added to the mean in the frequency equation

\[ X_{PMP} = \bar{X}_n + K_m \times S_n = (1 + K_m \times C_{vn}) \times \bar{X}_n \]

where $\bar{X}_n$ and $S_n$ are the mean and the standard deviation of the $n$ maxima, $C_{vn}$ is the coefficient of variation of the sample with $n$ values

• calculated in a unique way that the maximum observed value ($\bar{X}_n$) from the historical series will be omitted in the computation

\[ X_{pmp} = \bar{X}_n + K_m \times S_n \]

\[ K_m = \frac{(X_m - \bar{X}_{n-1})}{S_{n-1}} \]

\[ X_m = \bar{X}_{n-1} + K_m \times S_{n-1} \]

where $\bar{X}_{n-1}$ and $S_{n-1}$ are the mean and the standard deviation of the rainfall series from which the maximum record rainfall was omitted.
Criteria to Check the Eligibility and Stability of Using the Method

- Beyond the WMO No. 1045
- The criterion of **minimum data size** of $N_m$
  \[ N_m = \phi_m^2 + 2 \]
  where $\phi_m$ is the maximum deviation from mean and is directly computed from the following equation,
  \[ \phi_m = \frac{(X_m - \bar{X}_n)}{S_n} \]
- The criterion of the **stable size** of $N_s$ in terms of 10% relative error (Lin, 1981; Lin & Vogel, 1993)
  \[ N_s \geq 5.76 \times (\phi_m^2 + 2) \]
The Table of $K_m$ (Lin & Vogel, 1993)

<table>
<thead>
<tr>
<th>n</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>132.38</td>
<td>224.90</td>
<td>353.41</td>
<td>522.92</td>
<td>742.43</td>
<td>1014.94</td>
<td>1397.46</td>
</tr>
<tr>
<td>38</td>
<td>7.29</td>
<td>224.90</td>
<td>353.41</td>
<td>522.92</td>
<td>742.43</td>
<td>1014.94</td>
<td>1397.46</td>
</tr>
<tr>
<td>51</td>
<td>6.65</td>
<td>6.60</td>
<td>11.74</td>
<td>14.25</td>
<td>\</td>
<td>\</td>
<td>\</td>
</tr>
<tr>
<td>66</td>
<td>7.58</td>
<td>9.95</td>
<td>13.65</td>
<td>1014.94</td>
<td>1397.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>83</td>
<td>5.50</td>
<td>6.65</td>
<td>6.60</td>
<td>11.74</td>
<td>14.25</td>
<td>\</td>
<td>\</td>
</tr>
<tr>
<td>100</td>
<td>5.07</td>
<td>6.11</td>
<td>7.18</td>
<td>8.27</td>
<td>9.39</td>
<td>10.55</td>
<td>17.05</td>
</tr>
<tr>
<td>102</td>
<td>5.01</td>
<td>6.02</td>
<td>7.05</td>
<td>8.05</td>
<td>9.08</td>
<td>10.10</td>
<td>15.35</td>
</tr>
<tr>
<td>150</td>
<td>5.01</td>
<td>6.02</td>
<td>7.05</td>
<td>8.05</td>
<td>9.08</td>
<td>10.10</td>
<td>15.35</td>
</tr>
<tr>
<td>200</td>
<td>5.01</td>
<td>6.02</td>
<td>7.05</td>
<td>8.05</td>
<td>9.08</td>
<td>10.10</td>
<td>15.35</td>
</tr>
<tr>
<td>219</td>
<td>5.01</td>
<td>6.02</td>
<td>7.05</td>
<td>8.05</td>
<td>9.08</td>
<td>10.10</td>
<td>15.35</td>
</tr>
<tr>
<td>294</td>
<td>7.70</td>
<td>7.79</td>
<td>7.79</td>
<td>7.79</td>
<td>\</td>
<td>\</td>
<td>\</td>
</tr>
</tbody>
</table>

Notes: 1. $N_r$ refers to the minimum number of data length required to make the $K_m$-value method reasonable.
2. $N_s$ refers to the required number of data length to make a more statistically stable results.
Statistical Approach (5)

- Relationship of Variation Coefficient $\phi_m$ with $N_s$

Take some stations out when $N_s > 3.5n$ as it may cause 50% in error in terms of $K_m$.

Example:
220 > 3.5x60 causes 50% in $K_m$ ($3 / 6 = 0.5$)
Statistical Approach (6)

Probable maximum adjustment of sample mean $\overline{X}$

PMP, $X_{PMP}$

$$\overline{X}' = \overline{X} + 3 \times \sigma_{\overline{X}} \approx (1 + \frac{3 \times C_{yn}}{\sqrt{n}}) \times \overline{X}_n$$

where $\sigma_{\overline{X}}$ is the standard deviation of the mean

(Covering 99.7% of the probability area)

Fig. 51 Sketch for statistical approach to PMP (2)
Quality of PMP Estimates

It depends upon:

1) **Availability of data**;

2) **Depth of the study**.
1: How to determine the **lower limit of the integration of the pdf**?

\[ F(x) = \int_{-\infty}^{+\infty} f(x) \, dx = \int_{a}^{b} f(x) \, dx = \int_{??}^{?} f(x) \, dx = 1 \]

\[ P_{Ex} = \int_{a}^{+\infty} f(x) \, dx = 1 \]

\[ P_{Non} = \int_{-\infty}^{a} f(x) \, dx = 0 \]

Denominator = 0, computation crashed!

Fig. 52 Illustration of pdf curve (1)
Data sampling methods

- **Annual Maximum Series (AMS)**
  Annual Maximum Series (AMS) data consist of the largest event in each year, regardless of whether the second largest event in a year exceeds the largest events of other years.

- **Partial Duration Series (PDS)** *
  A partial duration series is a series of data which are selected so that their magnitude is greater than a predefined base value.

- **Annual Exceedance Series (AES)** *
  If the base value of the PDS is selected so that the number of values in the series is equal to the number of years of the record, the series is called an annual exceedance series. The AES may be regarded as a special case of the PDS. In the study, the PDS refers AES.

Exceedance frequencies of data (1)

- Location of the Semiarid study area

Fig. 53 SA data -- used for test of sampling methods
Exceedance frequencies of data (2)

**Table 1** Exceedance probabilities for region 1, SA

<table>
<thead>
<tr>
<th>Station ID</th>
<th>Data years</th>
<th>Return Period (R.P.) / Exceedance Probability (E.P.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2-y</td>
</tr>
<tr>
<td>04-0029</td>
<td>56</td>
<td>0.5</td>
</tr>
<tr>
<td>04-1476</td>
<td>53</td>
<td>0.76</td>
</tr>
<tr>
<td>04-1805</td>
<td>38</td>
<td>0.64</td>
</tr>
<tr>
<td>04-2964</td>
<td>48</td>
<td>0.76</td>
</tr>
<tr>
<td>04-4838</td>
<td>40</td>
<td>0.66</td>
</tr>
<tr>
<td>04-9053</td>
<td>69</td>
<td>0.67</td>
</tr>
<tr>
<td>35-2018</td>
<td>41</td>
<td>0.68</td>
</tr>
<tr>
<td>35-3232</td>
<td>30</td>
<td>0.60</td>
</tr>
<tr>
<td>35-5174</td>
<td>32</td>
<td>0.56</td>
</tr>
<tr>
<td>35-7354</td>
<td>67</td>
<td>0.75</td>
</tr>
<tr>
<td>35-8007</td>
<td>47</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.82</td>
</tr>
<tr>
<td><strong>Average E.P.</strong></td>
<td></td>
<td><strong>0.687</strong></td>
</tr>
<tr>
<td><strong>Corresponding R.P.</strong></td>
<td></td>
<td><strong>1.45-y</strong></td>
</tr>
</tbody>
</table>
Exceedance frequencies of data

Data Exceedance-freq. vs Theoretical Exceedance-proba.
(based on 1,438 stations data over 59 daily regions, SA)

Fig. 54 Illustration for comparison of frequency to probability
Underestimated quantiles based on AMS

• It is clear from Table 1 above:
  – The average exceedance frequencies of the 11 stations in the region 1 are 0.687, 0.223, 0.104, 0.032 and 0.014 for return periods of 2-y, 5-y, 10-y, 25-y, 50-y and 100-y, respectively.
  – The corresponding real return periods are calculated to be 1.45-y, 4.49-y, 9.57-y, 20.75-y, 31.45-y and 73.53-y.
• It is also clear from the Chart above:
  – The area averaged exceedance frequencies are higher than the corresponding analytical exceedance probabilities for 2-y through 100-y, particularly much higher for 2-y to 10-y.

• Conclusion:
  – Under current concept, quantiles were underestimated based on AMS particularly for 2~10-year.
What is wrong?

- Obviously, the quantiles for frequent events, particularly 2-y thru 10-y, have been underestimated for long time under current estimation approach.
- The problem comes from inconsistency between the definition of the return period of quantiles and the data sampling method that creates the AMS data for frequency analysis to get the quantiles.
- The return period is defined for an average time interval in unit of year for $x_T$ to occur over a large time period. It doesn’t mean occurring once per each time interval.
- However, the AMS takes only the largest event in each year, regardless of whether the second largest event in a year exceeds the largest events of other years. Something (some high values) has been missed.
Relation (equation) of PDS-AMS

- Conversion of PDS-AMS:
  \[
  T_{AES} = \left[ \ln\left( \frac{T_{AMS}}{T_{AMS} - 1} \right) \right]^{-1}
  \]
  (Ven Te Chow, 1964)

- or,
  \[
  T_{AMS} = \frac{1}{1 - e^{-\frac{1}{T_{AES}}}}
  \]
How to correct?

• Two ways to correct the underestimation:
  
  – To use the AES data in combination with the use of the exceedance probabilities listed in Table 2, i.e. 0.5, 0.2, 0.1, 0.04 and 0.02 for return periods of 2-y, 5-y, 10-y, 25-y and 50-y.

  – Or, to use the AMS data in combination with the use of the exceedance probabilities listed in Table 3, i.e. 0.3935, 0.1813, 0.0952, 0.0392 and 0.0198 for return periods of 2-y, 5-y, 10-y, 25-y and 50-y.

• The two ways are deemed to be equivalent in quantiles estimation.
Findings / Suggestions

- Quantiles based on AMS are underestimated;
- Concept of PDS or AES is in accordance with the Return Period;
- **Recommend**: It is YES to continue to employ the AMS data but with adjustments of non-exceedance probabilities based on Ven Te Chow’s equation of PDS-AMS.
Relation (the equation) of PDS-AMS

- **Table 2** Return periods based on AMS data

<table>
<thead>
<tr>
<th>$T_{AES}$ (-year)</th>
<th>$T_{AMS}$ (-year)</th>
<th>$P_E = 1 / T_{AMS}$</th>
<th>$P_{NON} = 1 – 1 / T_{AMS}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N/A)</td>
<td>1</td>
<td>1.0</td>
<td>0.0*</td>
</tr>
<tr>
<td>1.44</td>
<td>2</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>4.48</td>
<td>5</td>
<td>0.2</td>
<td>0.8</td>
</tr>
<tr>
<td>9.49</td>
<td>10</td>
<td>0.1</td>
<td>0.9</td>
</tr>
<tr>
<td>24.50</td>
<td>25</td>
<td>0.04</td>
<td>0.96</td>
</tr>
<tr>
<td>49.50</td>
<td>50</td>
<td>0.02</td>
<td>0.98</td>
</tr>
</tbody>
</table>

* Note: It is incomputable for the 1-year event under AMS-based data. The frequency computer program is designed and coded based on the non-exceedance probability.
Thus, 1-year event can be estimated

- Table 3  Return periods based on AES data

<table>
<thead>
<tr>
<th>$T_{AES}$ (-year)</th>
<th>$T_{AMS}$ (-year)</th>
<th>$P_E = 1 / T_{AMS}$</th>
<th>$P_{NON} = 1 - 1 / T_{AMS}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.58</td>
<td>0.6321</td>
<td>0.3679</td>
</tr>
<tr>
<td>2</td>
<td>2.54</td>
<td>0.3935</td>
<td>0.6065</td>
</tr>
<tr>
<td>5</td>
<td>5.52</td>
<td>0.1813</td>
<td>0.8187</td>
</tr>
<tr>
<td>10</td>
<td>10.51</td>
<td>0.0952</td>
<td>0.9048</td>
</tr>
<tr>
<td>25</td>
<td>25.50</td>
<td>0.0392</td>
<td>0.9608</td>
</tr>
<tr>
<td>50</td>
<td>50.50</td>
<td>0.0198</td>
<td>0.9802</td>
</tr>
</tbody>
</table>

$p_E$ stands for exceedance probability; $p_{NON}$ stands for Non-exceedance probability.
2: How to determine the upper limit of the integration of the pdf?

\[ F(x) = \int_{-\infty}^{+\infty} f(x) \, dx = \int_{a}^{b} f(x) \, dx = \int_{?}^{?} f(x) \, dx = 1 \]

\[ P_{Ex} = P(X > b) = \int_{b}^{+\infty} f(x) \, dx = 0 \]

It means: the pdf should converge at \( b \).
Can we do?
In other words:

What is the probability of an estimated PMP?
In current textbooks it assumes:

• All quantiles are normally distributed → leading to divergence of upper tail

 Against reality
However, my studies say:
“No, not the case!”

1. Quantiles vary asymmetrically
2. Around 25-50yr – symmetrical variation
3. Quantiles < 25-50yr – positively skewed
4. Quantiles > 25-50yr – negatively skewed

leading to convergence of upper tail

(My investigations of a great number of AMS precipitation data in the U.S. and China support my findings; see below Figs. 56, 57, 58)
Findings (Results) over 84 Regions, OH

Ratios of Upper vs. Lower Confi-limit
(Mean-ratio over 84 daily regions, OH)

Fig. 56 Ratios of (upper vs lower) for Ohio River Basin
Findings (Results) over 59 Regions, SA

Fig. 57 Rations of (upper vs lower) for SW Semiarid U.S.
Findings (Results) over 8 Regions, Taihu

Fig. 58 Rations of (upper vs lower) for Taihu Lake
These Studies indicate that:

Upper tail of frequency distribution tends to converge

These evidences suggest that

the upper tail of the probability distribution should converge to a certain value by an asymptote.
**Conclusion:** Estimation of the upper limit of integration of the PDF is doable

\[ \int_{-\infty}^{+\infty} f(x) \, dx = 1 \]

\[ \int_{?}^{?} f(x) \, dx = 1 \]

\[ \int_{a}^{b} f(x) \, dx = 1 \]
Thus, **the frequency analysis and the PMP study can be unified**, tested each other, complemented and no longer fighting against each other – **this may change the entire world of the hydrologic design studies. Amazing!**

Wow!
It’s exciting, but not easy!

There is still a long way to go.

However, the direction to go is certain. *It’s doable!*
Acknowledgements

1. NOAA of the U.S.: *Precipitation Frequency Atlases Update project*;
2. MWR of China: *Application of Regional L-Moments Method to Flood-Mitigation Planning* (#201001047, ongoing);
3. MWR of China: *Impact of Climate Change on PMP Estimation and the Countermeasures to Flood-Mitigation* (#201101033, ongoing);
4. HKSAR Government GEO: *PMP Estimation for Hong Kong* (ongoing);
Anything else?
Announcement

Conference on the theme of
“Extreme Hydrometeorological Events and Flood-Control & Disaster-Reduction with Risk Analysis” to be held on 25-27 October 2013 in Xiamen of South China, a charming seaside city, one of the most beautiful cities in China, hosted by the NUIST.
The city of Xiamen

Conference to be held here, 25-27 October 2013
The End

Thank you

Contact person:
Dr. Yehui Zhang, Assistant to Director
Applied Hydrometeorological Research Institute (AHMRI)
Nanjing University of Information Science & Technology (NUIST)
Email: zhangyehui@gmail.com