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

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）

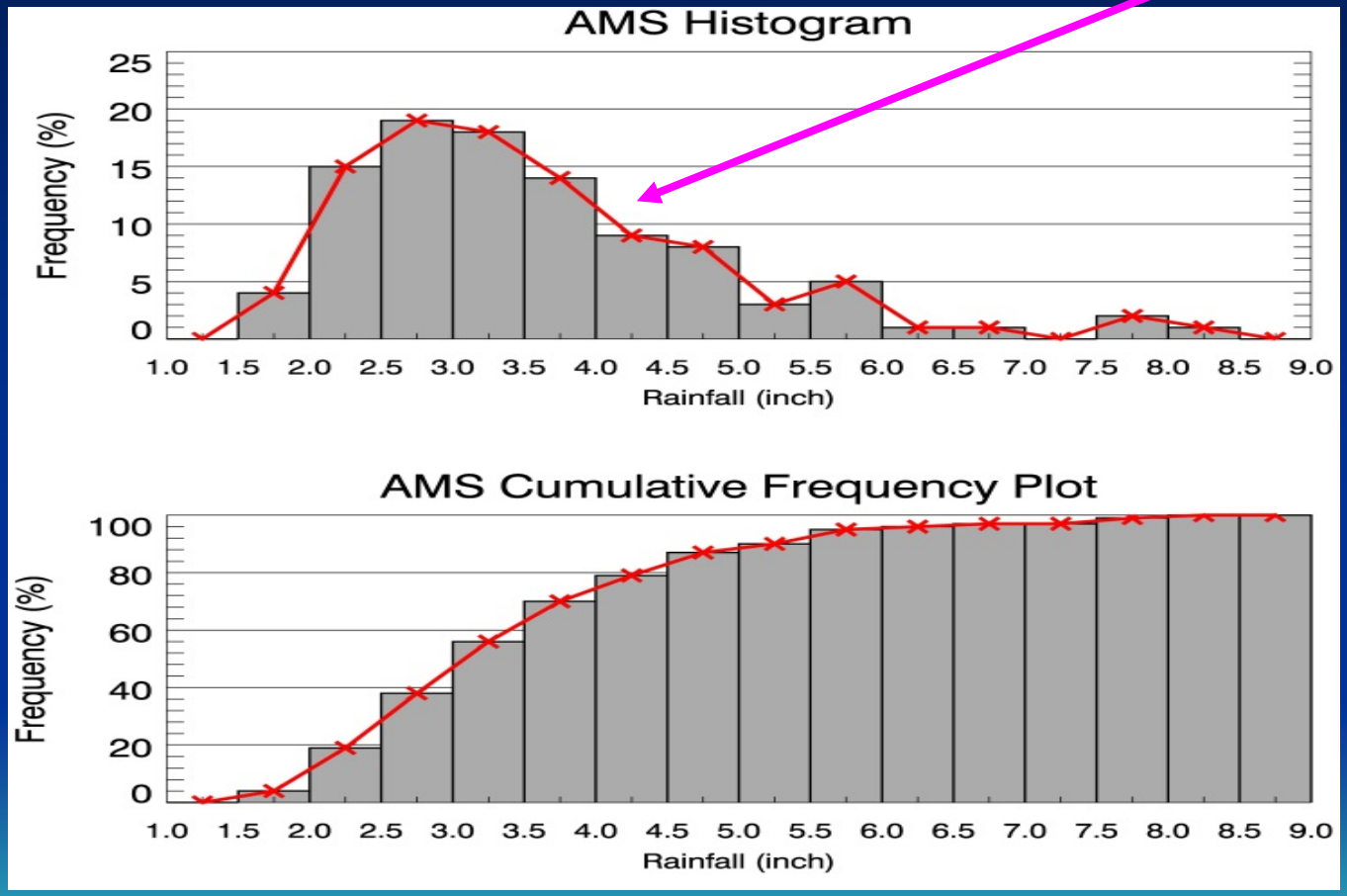
(011901) 28-4229 ANMAX

100

1.55	1.61	1.64	1.95	2.04	2.05	2.15	2.18	2.21	2.23
2.25	2.27	2.31	2.36	2.38	2.38	2.42	2.46	2.46	2.52
2.55	2.56	2.60	2.65	2.66	2.73	2.74	2.77	2.79	2.85
2.85	2.94	2.95	2.96	2.96	2.97	2.98	2.98	3.06	3.07
3.10	3.14	3.15	3.15	3.18	3.19	3.20	3.21	3.22	3.23
3.24	3.30	3.37	3.40	3.42	3.47	3.52	3.61	3.61	3.63
3.64	3.66	3.66	3.78	3.80	3.84	3.88	3.91	3.92	3.97
4.06	4.16	4.22	4.23	4.29	4.31	4.35	4.40	4.45	4.51
4.56	4.65	4.72	4.76	4.82	4.86	4.98	5.16	5.31	5.34
5.53	5.56	5.60	5.73	5.74	6.41	6.86	7.52	7.69	8.11

Histogram

Empirical frequency curve



Increment = 0.5 in.

Fig. 1 AMS histogram and empirical frequency curve

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

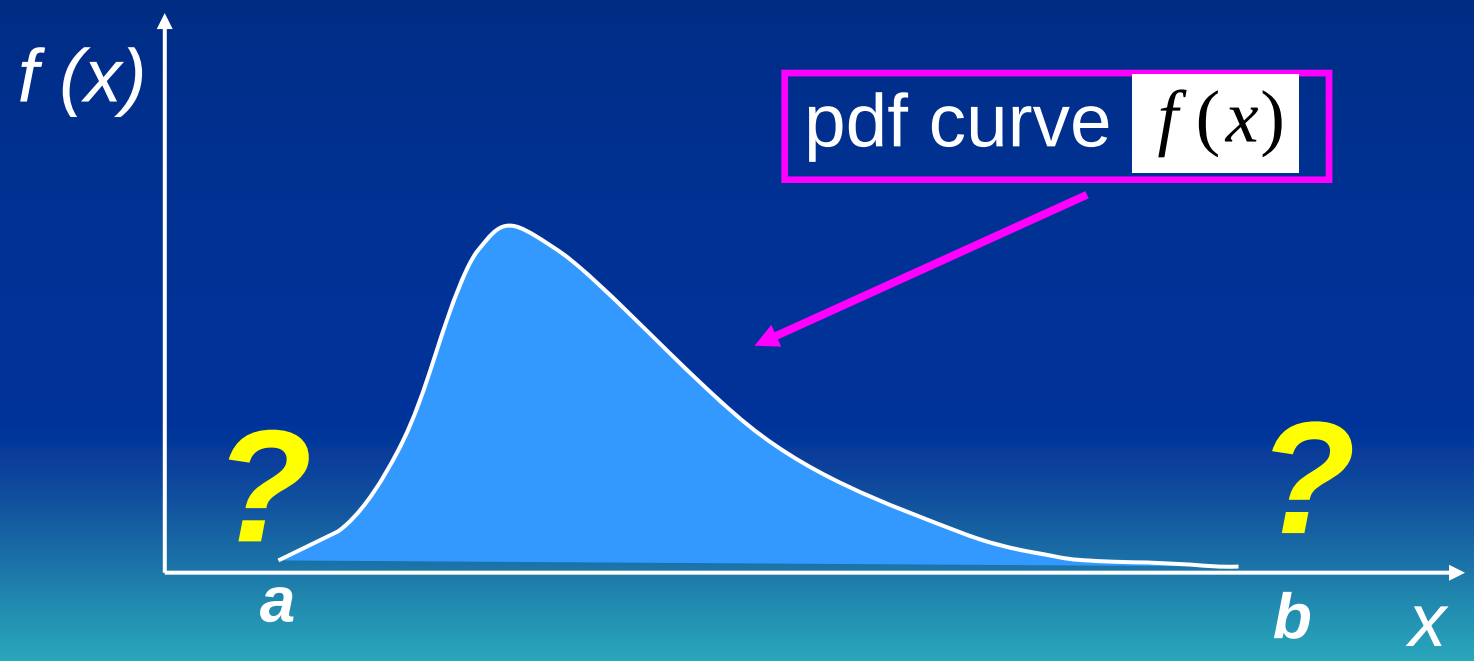


Fig. 2 Probability density function curve

What is Frequency Analysis ?

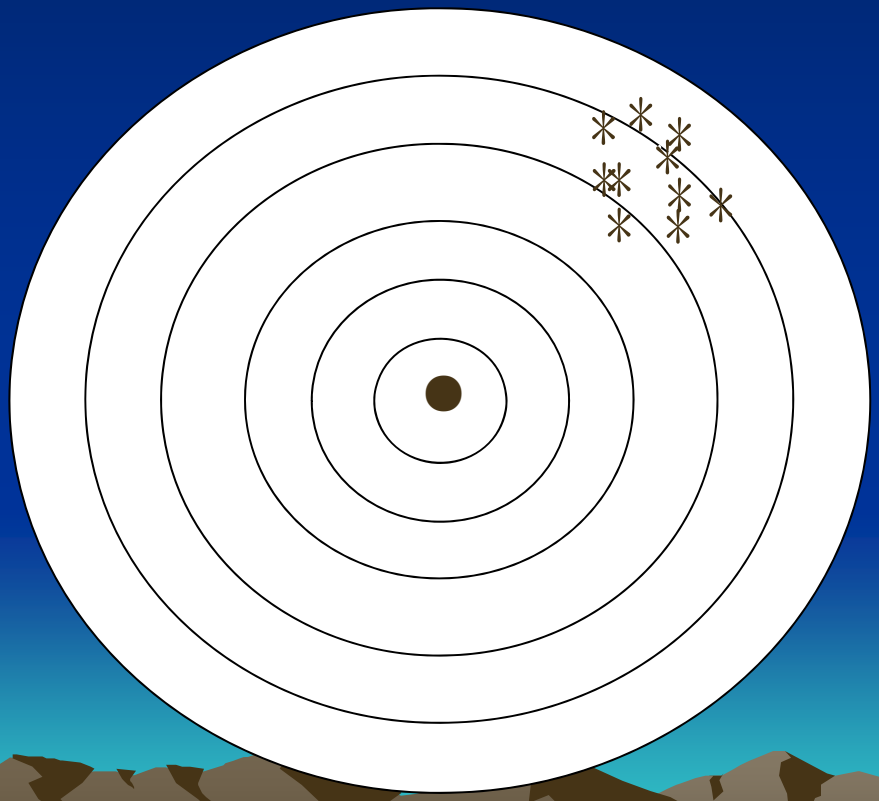
Frequency Analysis -- is a statistical approach using sample or samples to estimate the population probability distribution.

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)

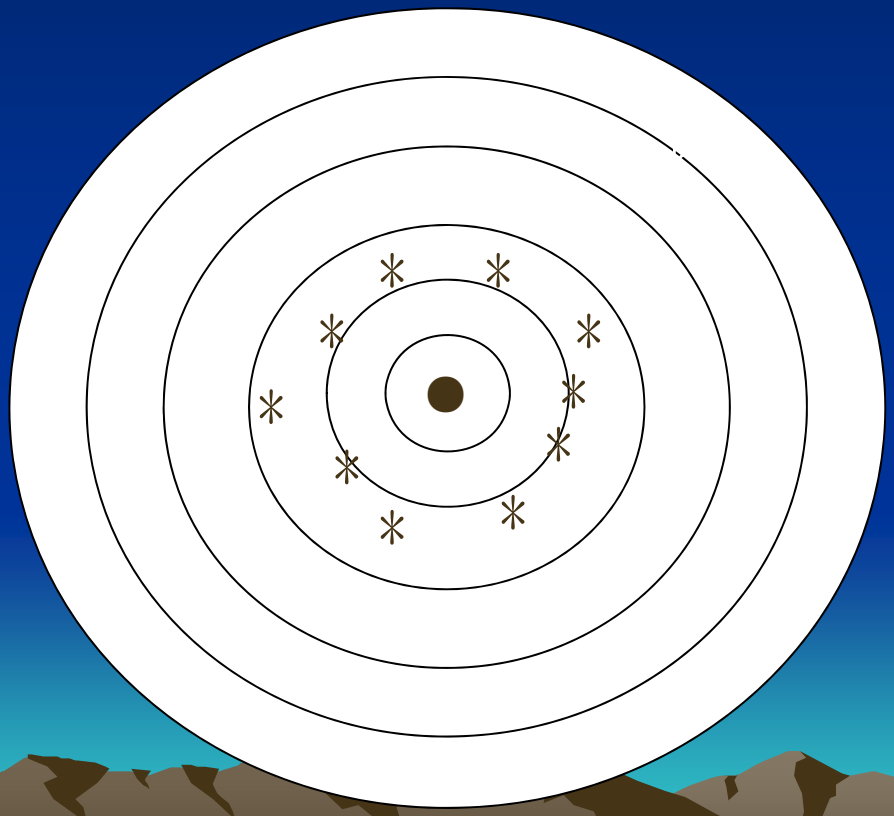


Good precision

Poor accuracy

Nature of the Frequency Analysis (2)

- Like shooting practice: Precision and Accuracy

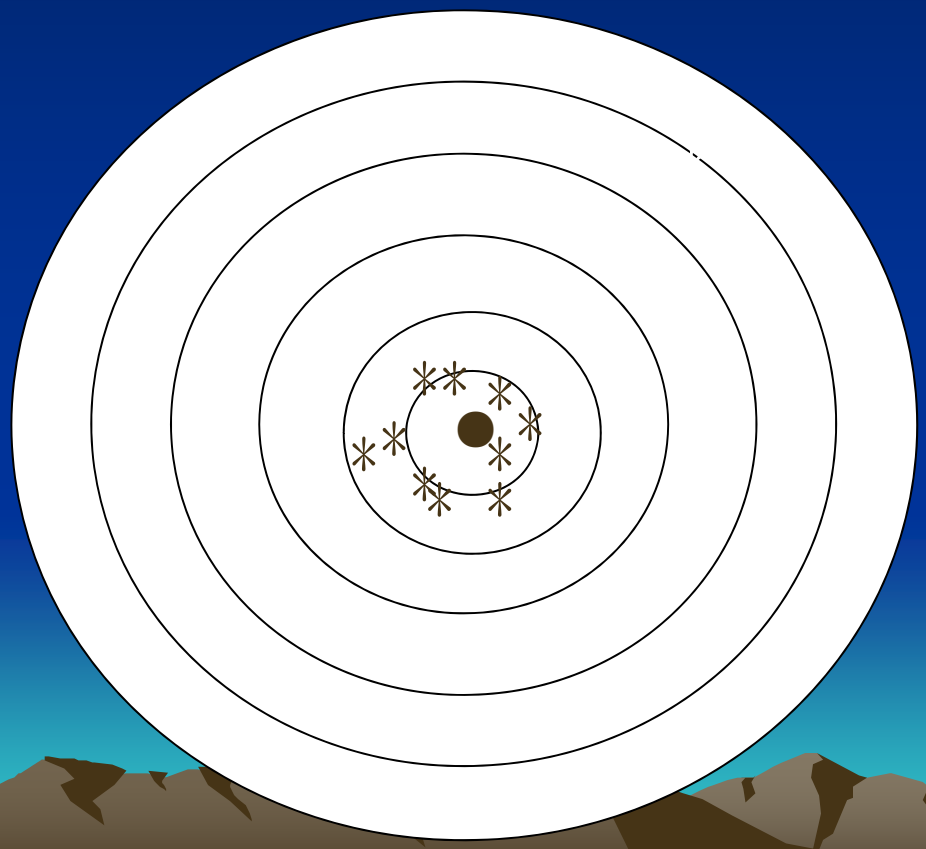


Poor precision

Moderate accuracy

Nature of the Frequency Analysis (3)

- Like shooting practice: Precision and Accuracy



Good precision

Good accuracy

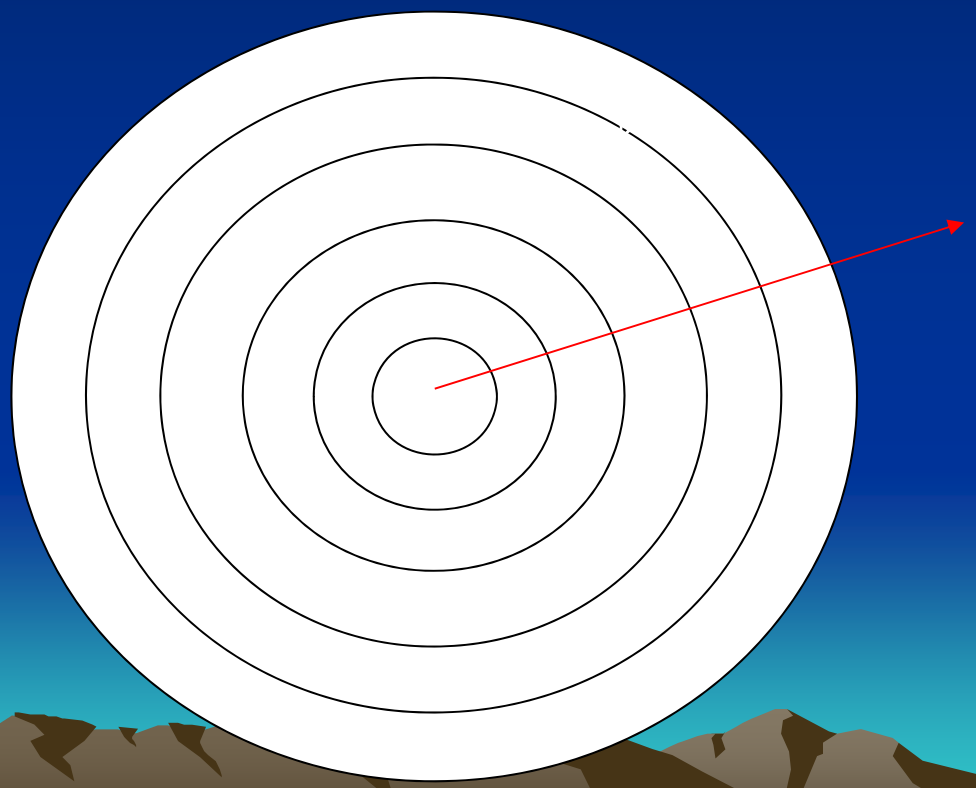
Two Impossible Things in FA

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

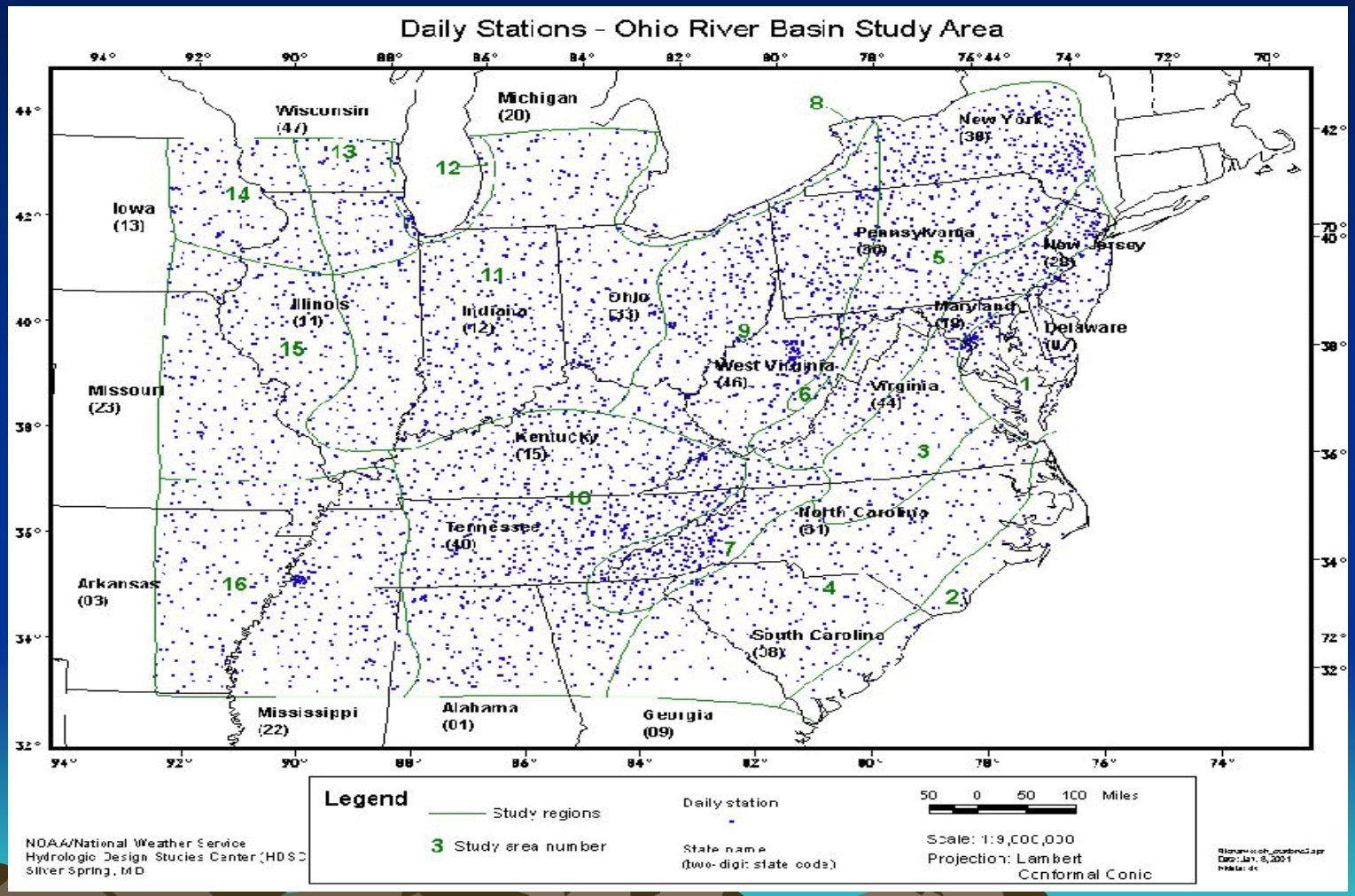


Fig. 3 Location of raingauges in Ohio River Basin

Only 16% or 1797 sites exhibited linear trend in mean for AMS in Ohio River Basin in past century

（1900 – 2000年）

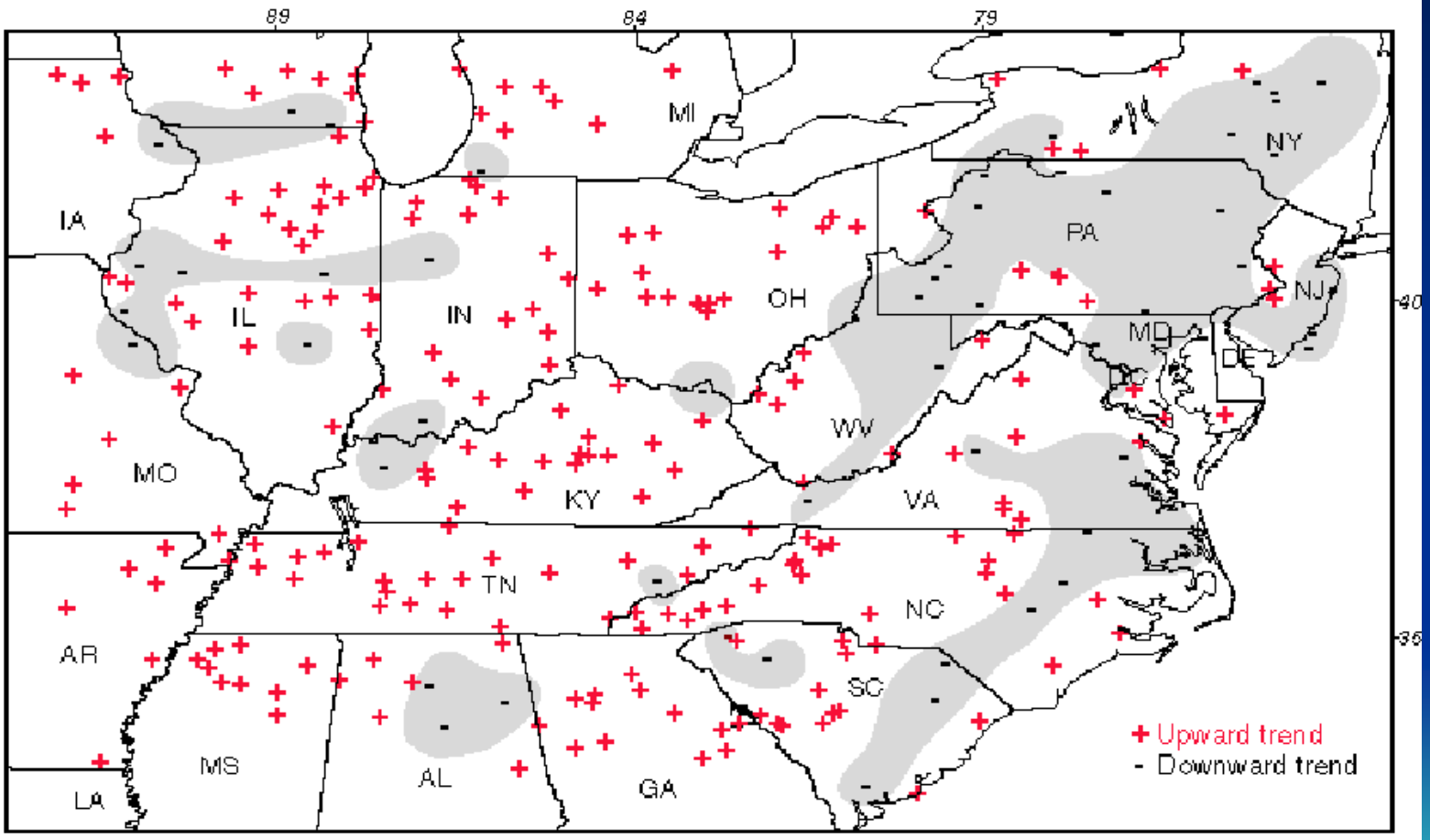


Fig. 4 Map of stations that exhibit linear trend in mean for AMP over 20th century

Only 18% or 531 sites exhibited shift in mean for AMS in past century

（compression of prior- to post- 1958 ）

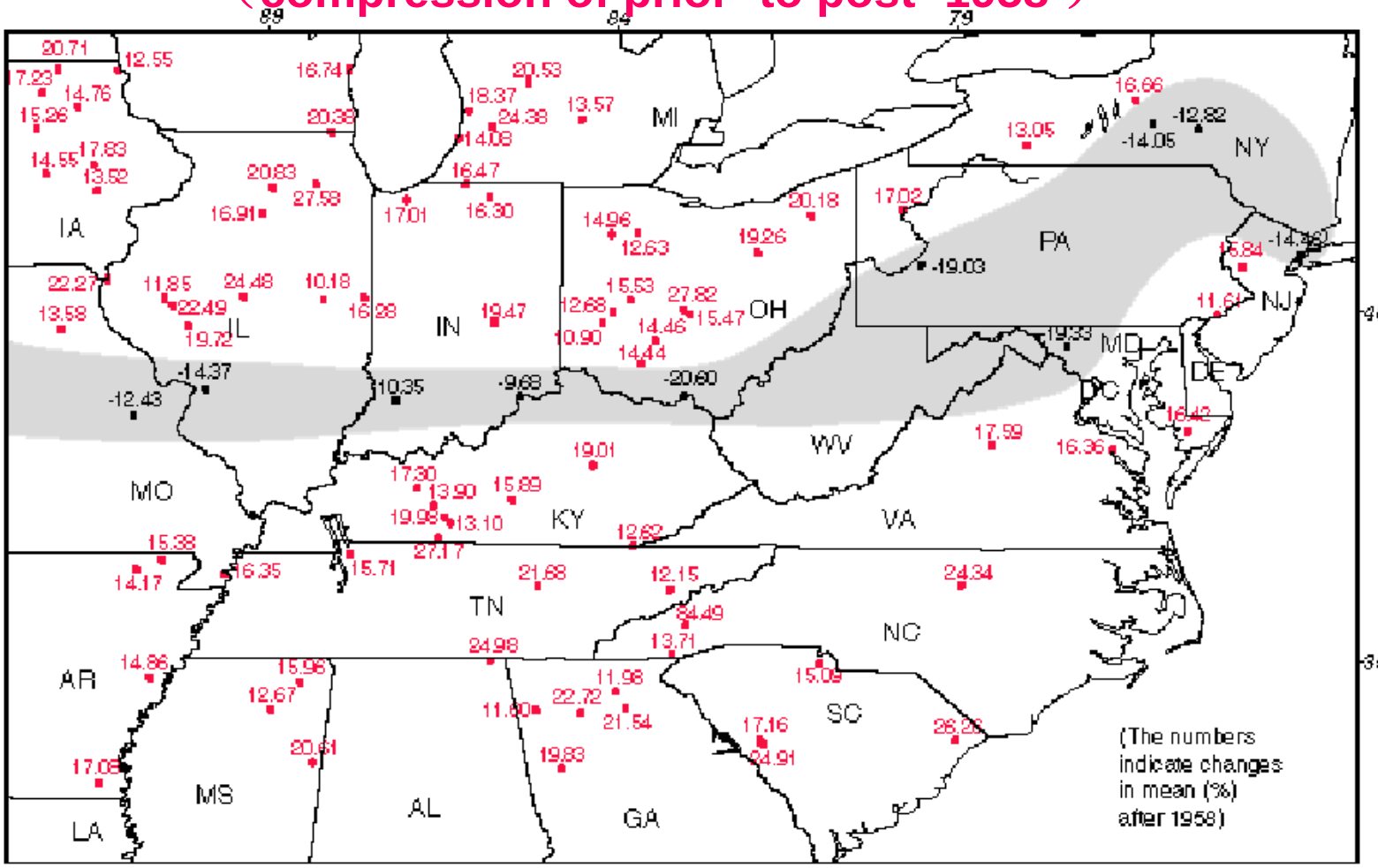


Fig. 5 Map of stations that exhibit shift in mean for AMP over the 20th century

Investigation of 1,741 daily AMS in Mainland China

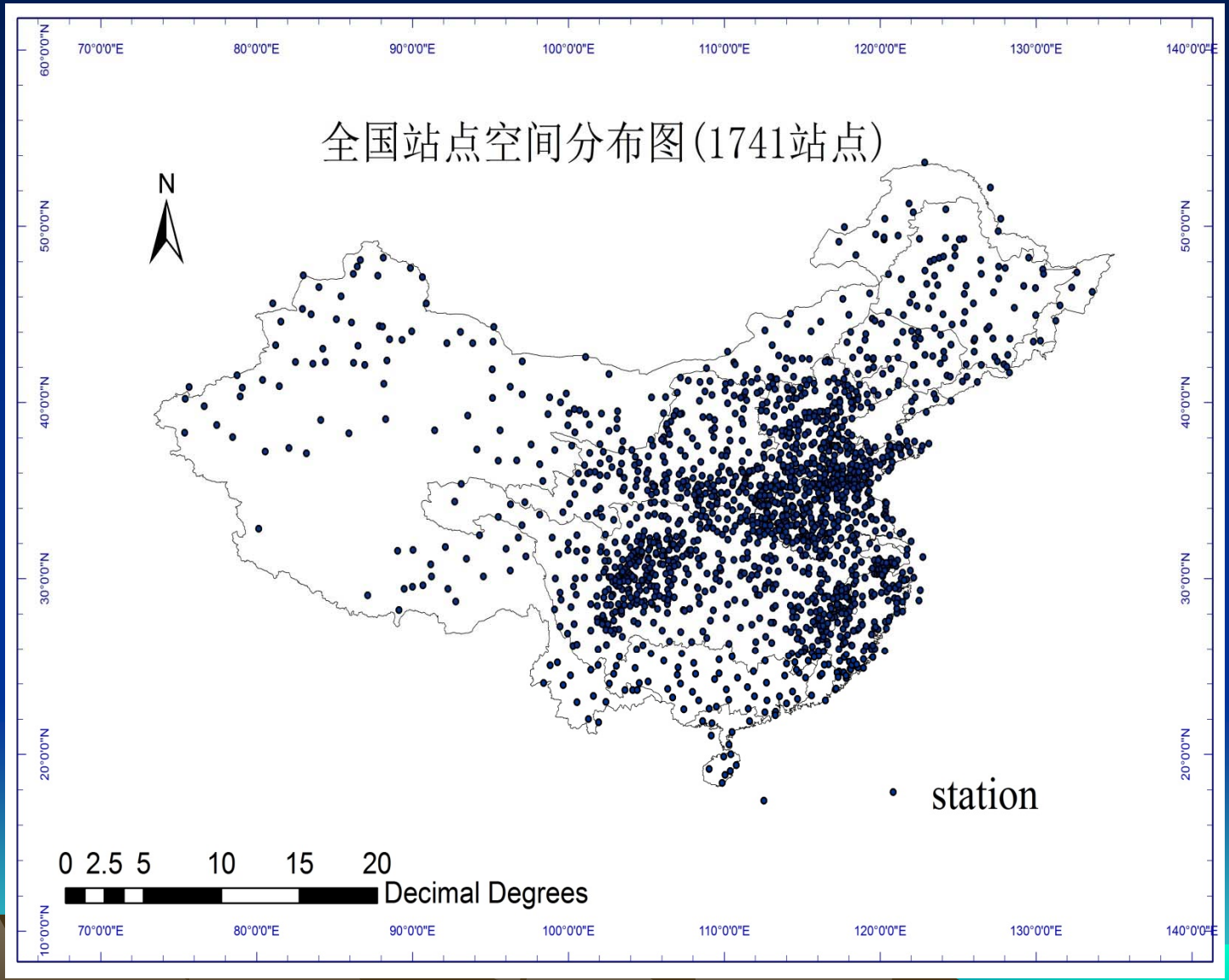


Fig. 6 Location of raingauges in China (partially)

Only 17.5% of 2,436 sites exhibited linear trend in mean of AMS in China

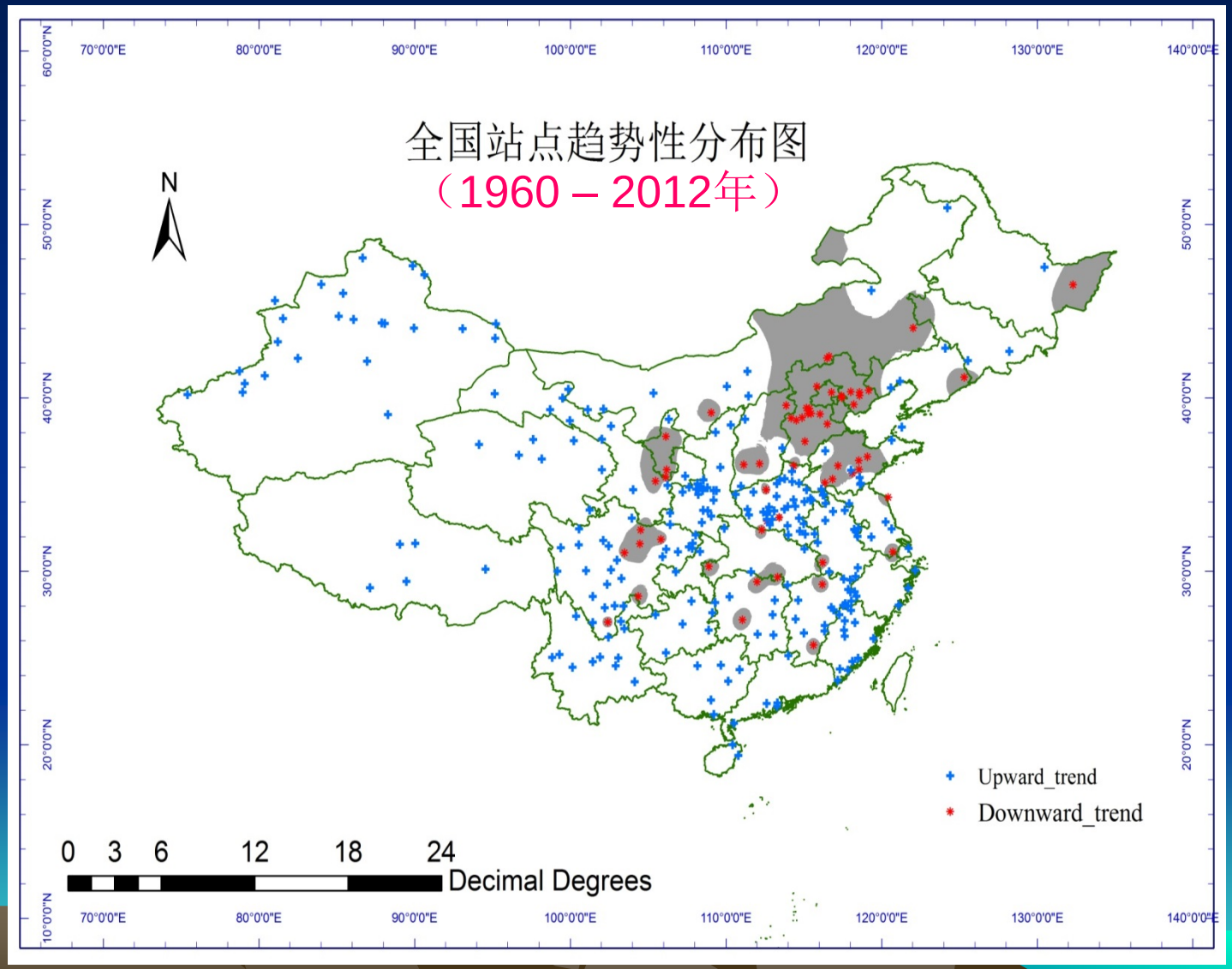


Fig. 7 Spatial distribution of raingauges with linear trend in China

Only 10.8% of 1,649 sites exhibited shift in mean for AMS in China

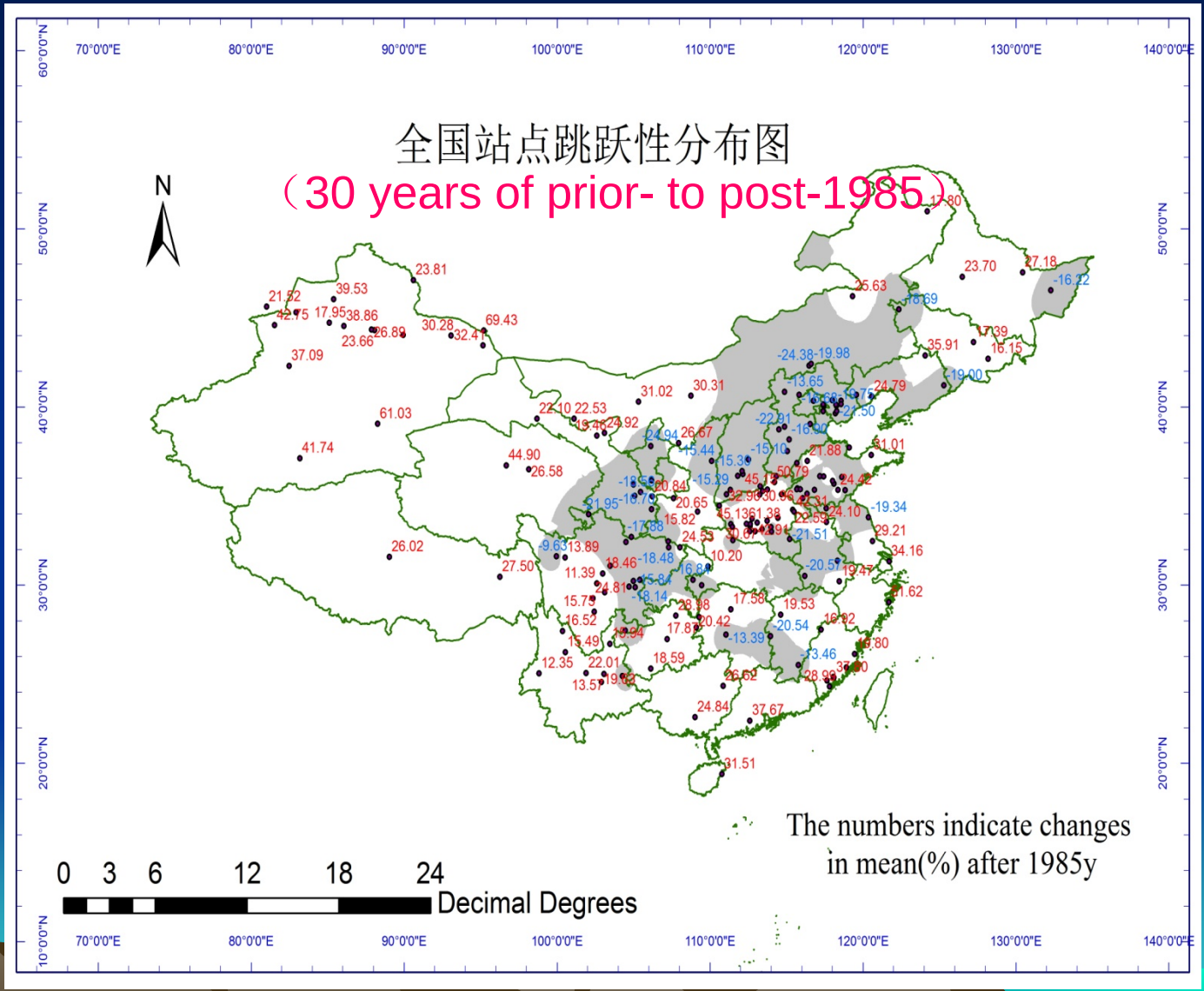


Fig.8 Spatial distribution of raingauges with shift in China

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)

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

Power shift

Advantage – Linearity

2. L-Moments Method (LMM)

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

L-Moments

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

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

$$\lambda_1 = EX$$

$$\lambda_2 = \frac{1}{2}E(X_{2:2} - X_{1:2})$$

$$\lambda_3 = \frac{1}{3}E(X_{3:3} - 2X_{2:3} + X_{1:3})$$

$$\lambda_4 = \frac{1}{4}E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4})$$

$$(When \quad X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n})$$

Application of RLM to Precipitation Frequency Analysis in the U.S.

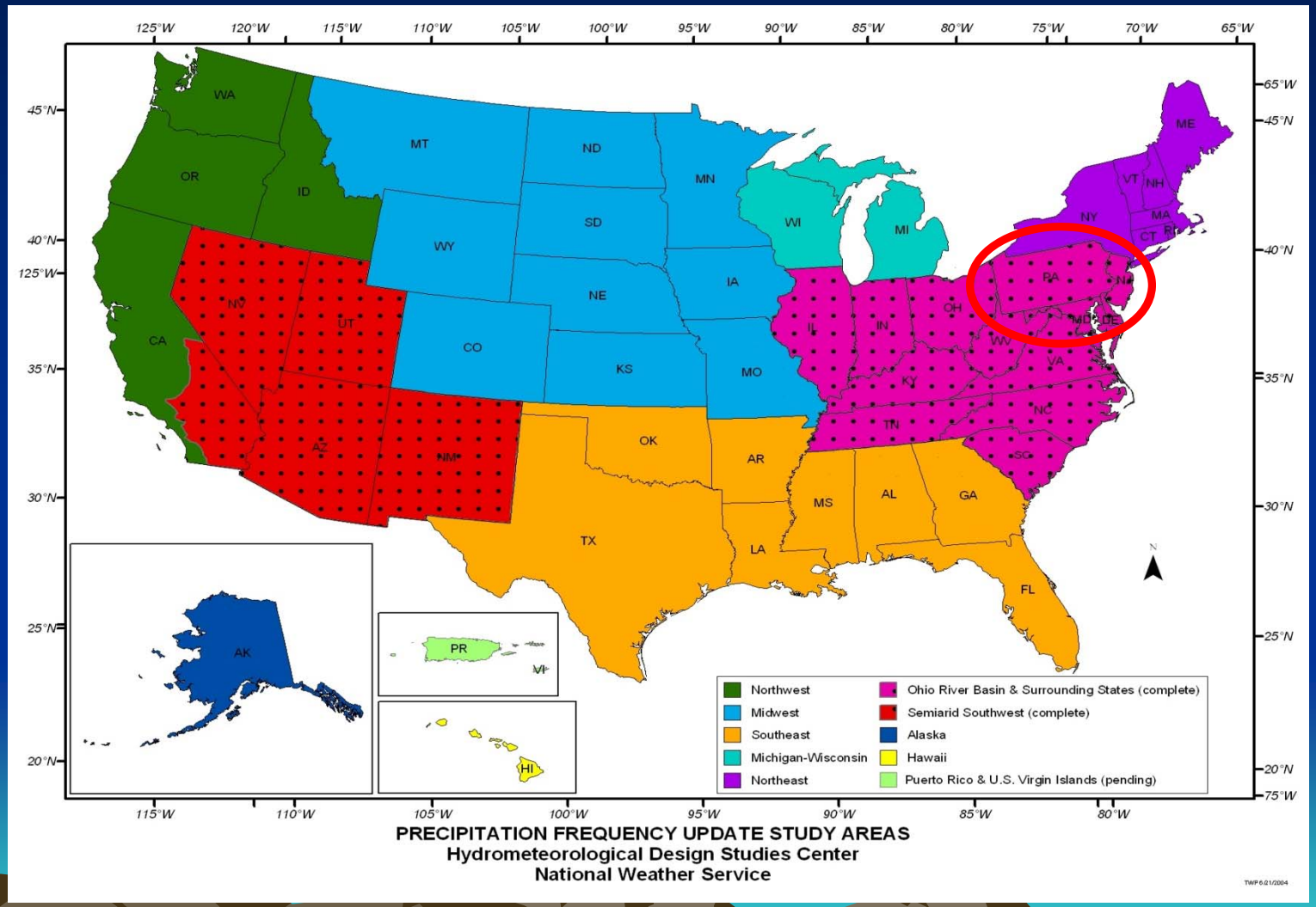
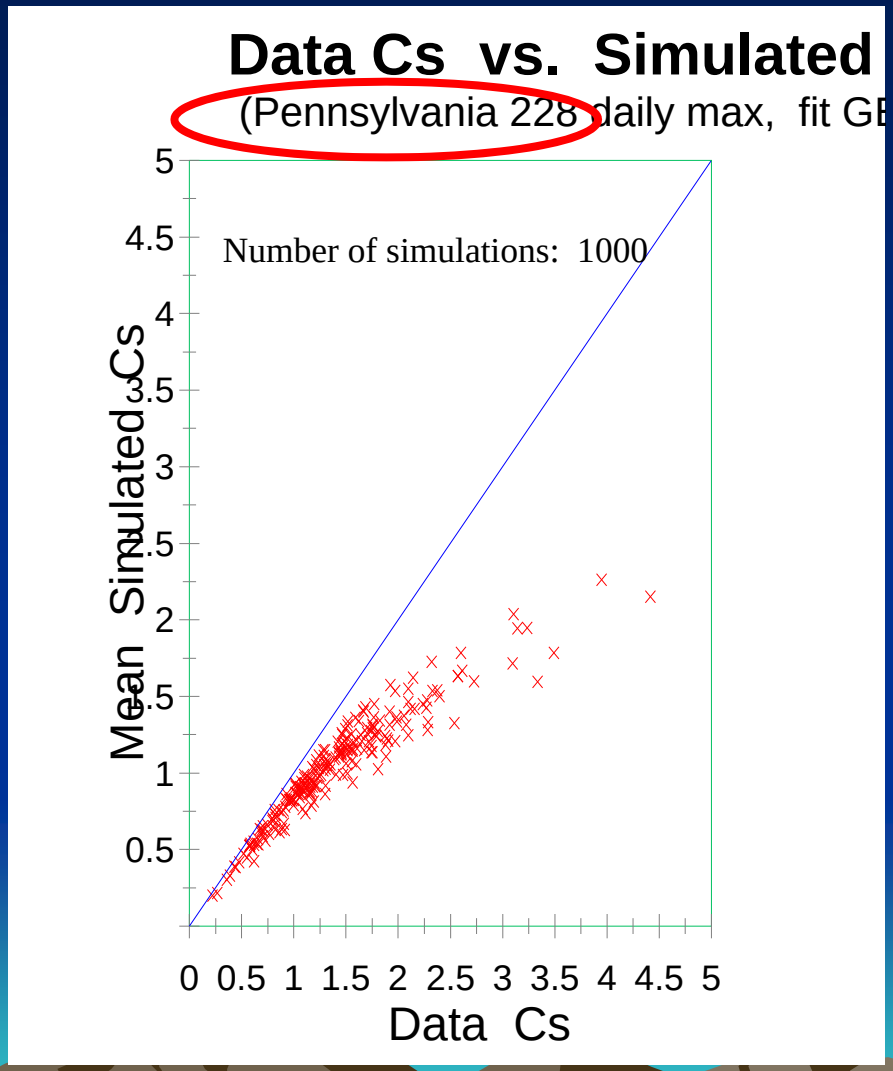
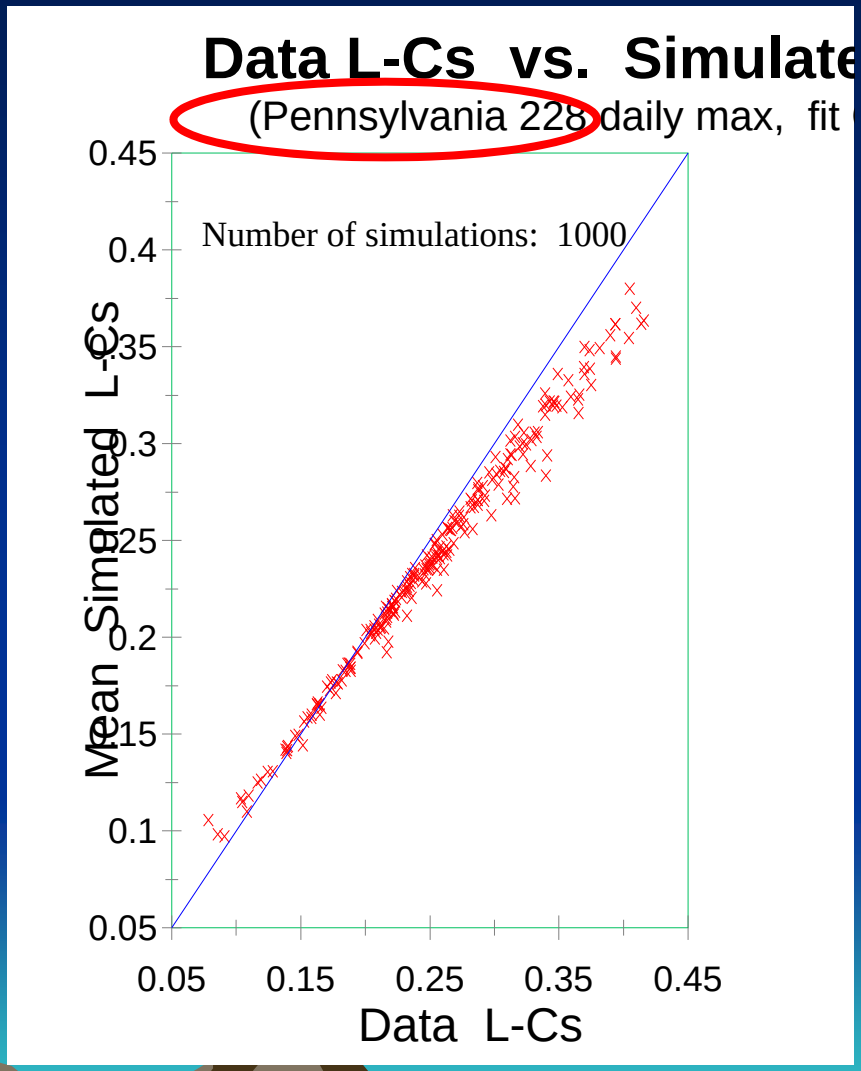


Fig. 9 Pennsylvania data will be used for comparison

Comparison of CMM to LMM on Biasedness



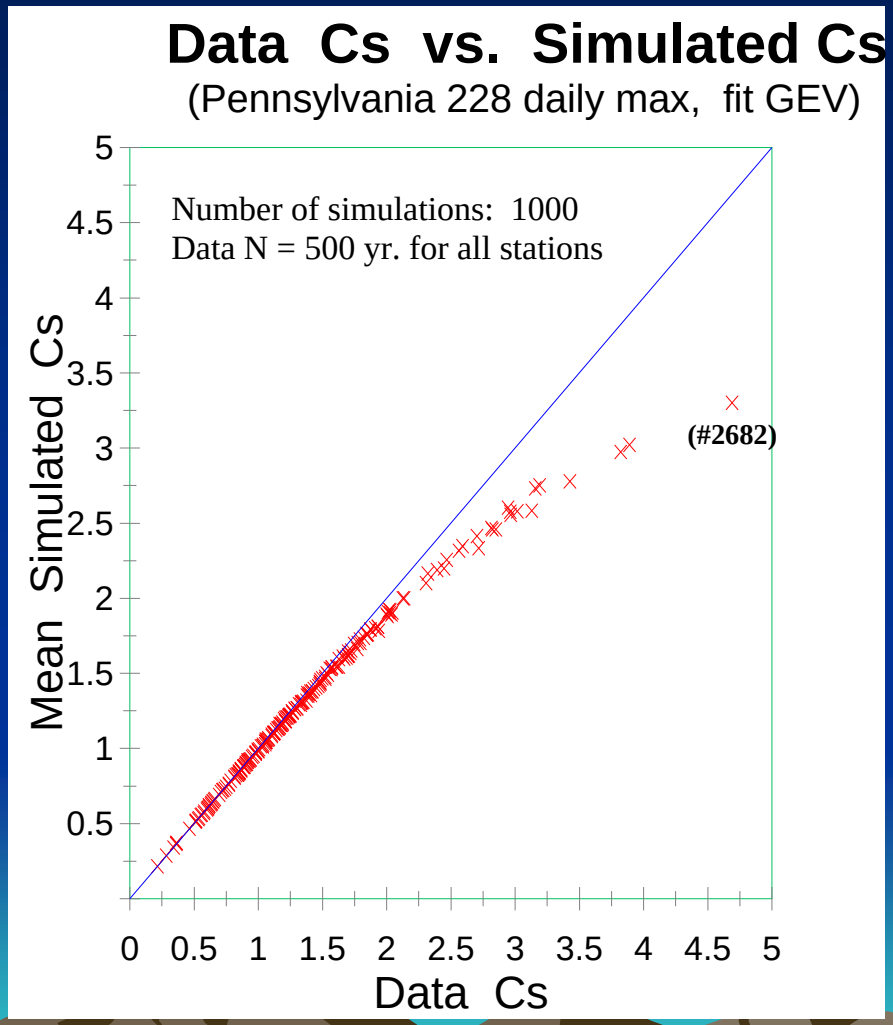


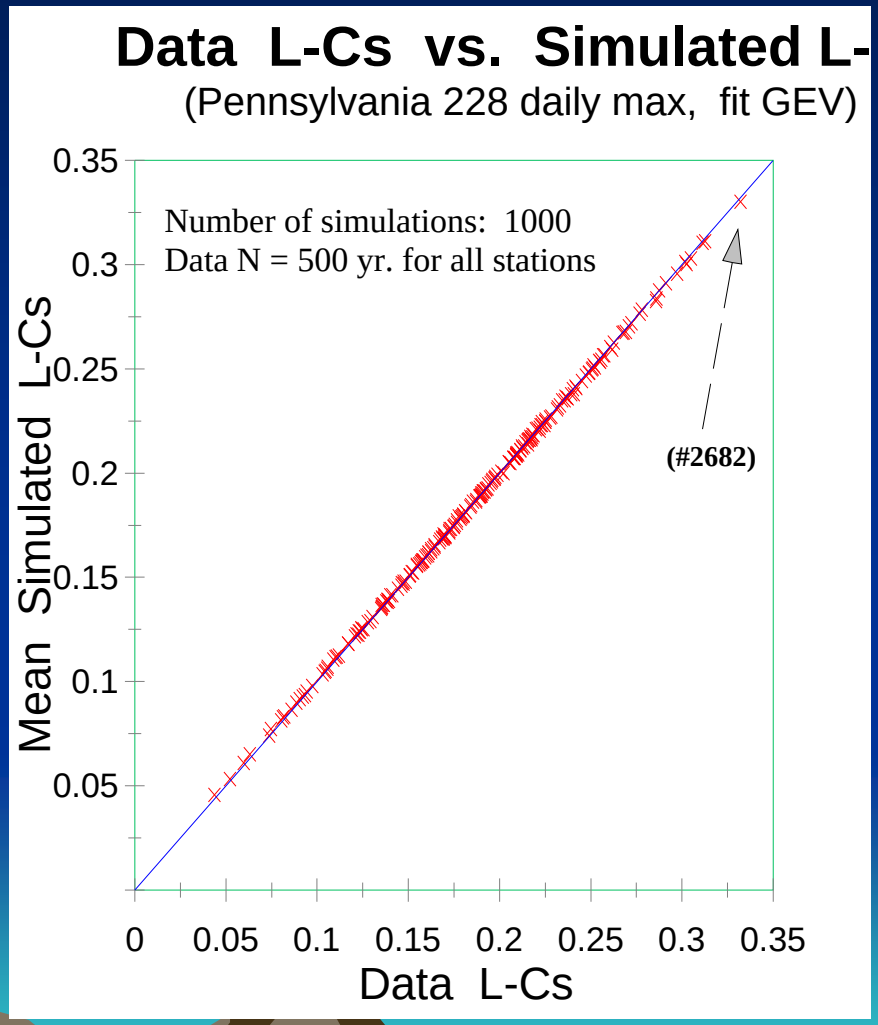
High biasedness of Cs **Much less biasedness of L-Cs**

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.





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

Fig. 11Comparison 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

Regional L-moments Analysis in China

(A pilot study – Application to Taihu Lake Basin)

Fig. 12 Location of Taihu Lake Basin

Location of 96 raingauges in Taihu Lake (6,134 km^2)

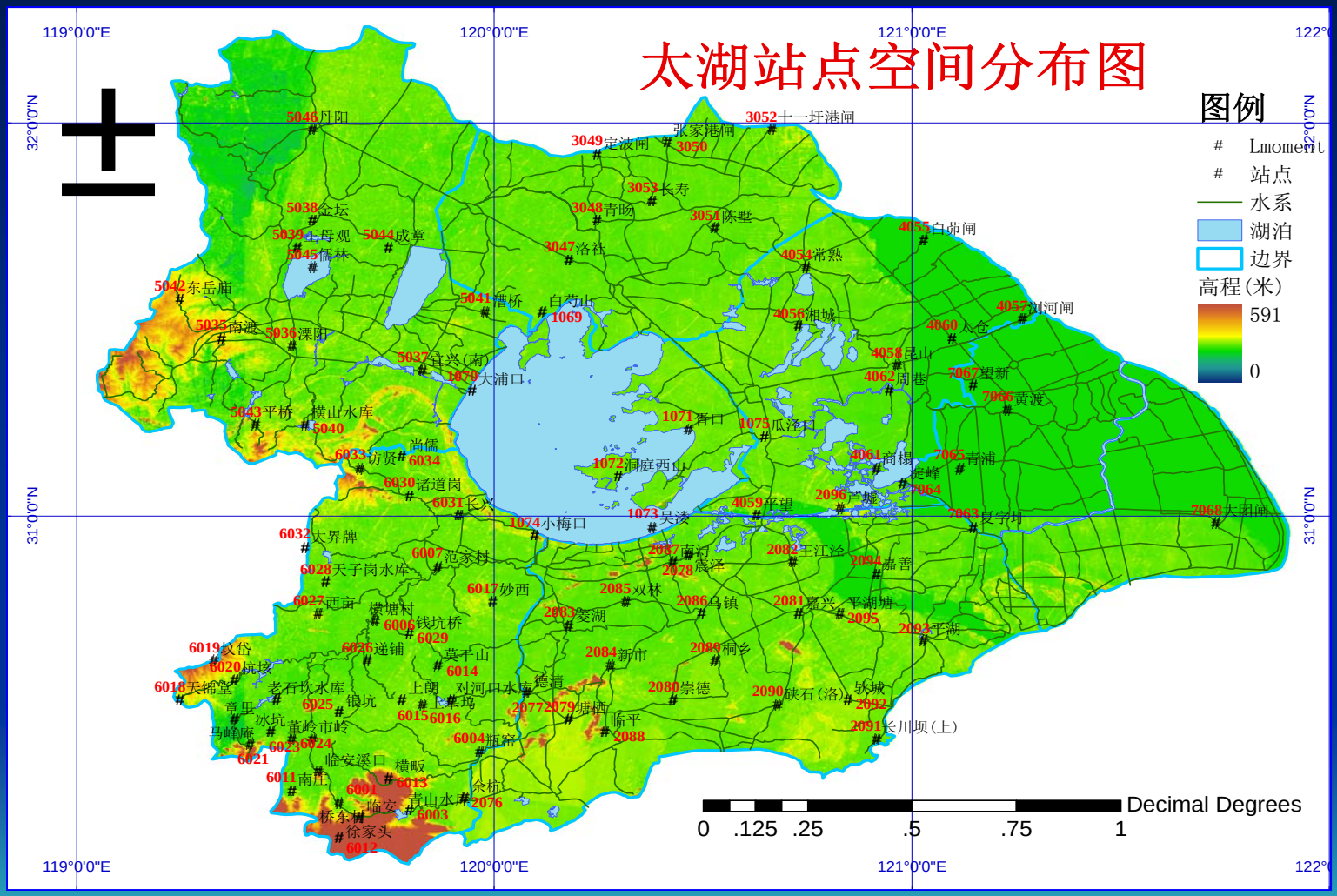


Fig. 13 Map of Taihu Lake Basin with distribution of 96 raingauges

(96+45) Sites of Taihu Lake divided into 8 HGS Regions

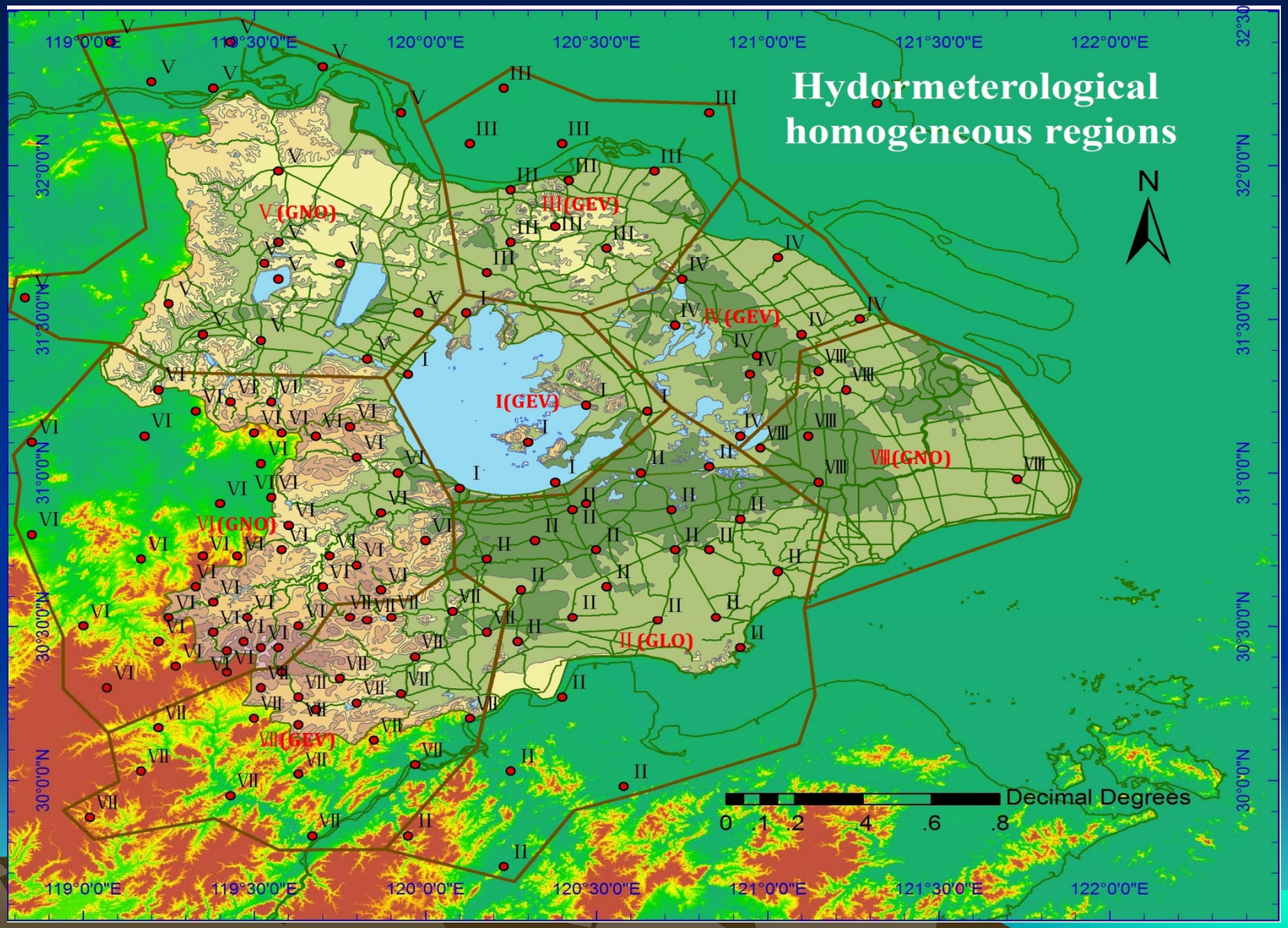


Fig. 14 Regionalization of Taihu Lake Basin

Comparison for Taihu Lake Data (1)
on Biasedness

Taihu 96 daily max, fit GLO

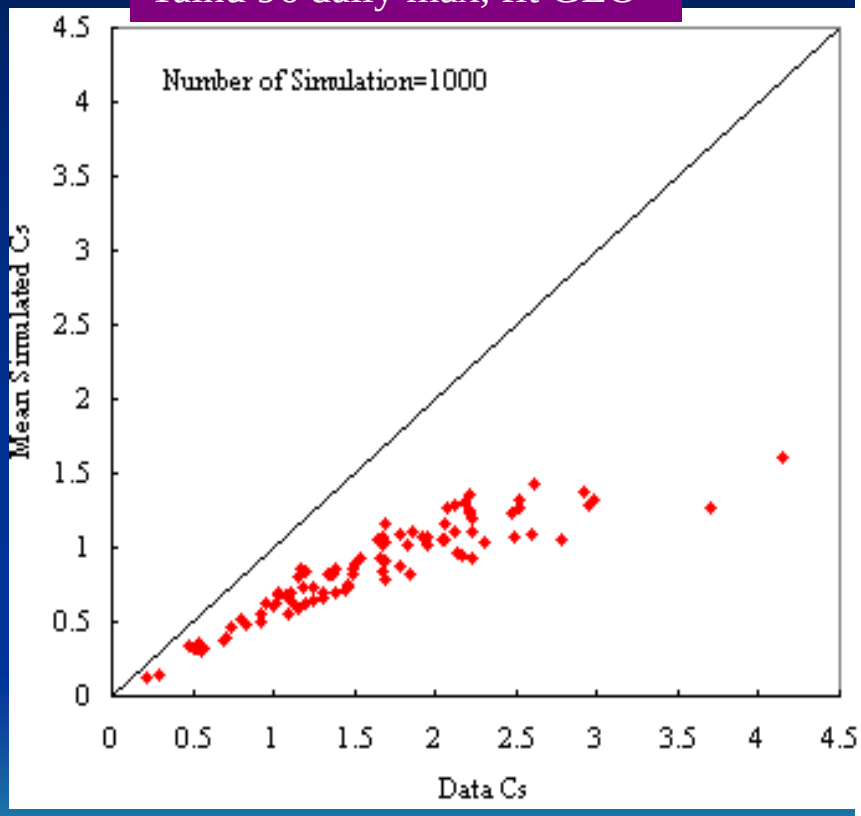


Fig.1 Biasedness on skewness of CMM

Taihu 96 daily max, fit GLO

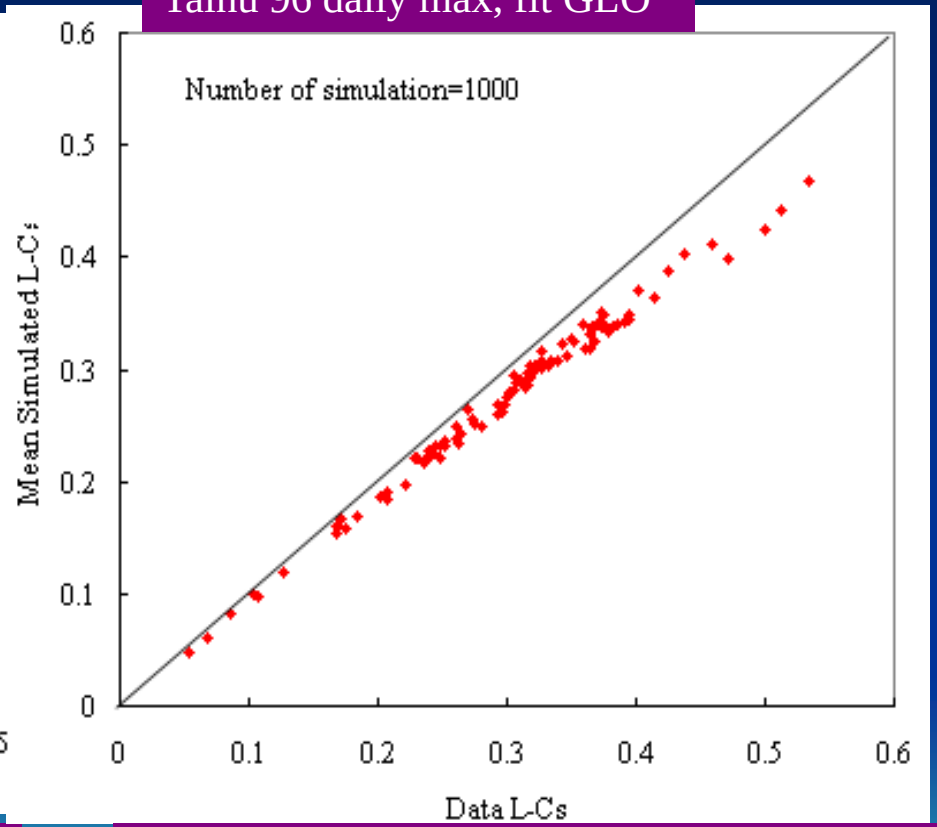


Fig.2 Unbiasedness on skewness of LMM

Fig. 15 Comparison for biasedness between CMM and LMM based on Taihu data

Comparison for Taihu Lake Data (2) on Robustness

Taihu 96 daily max, fit GLO

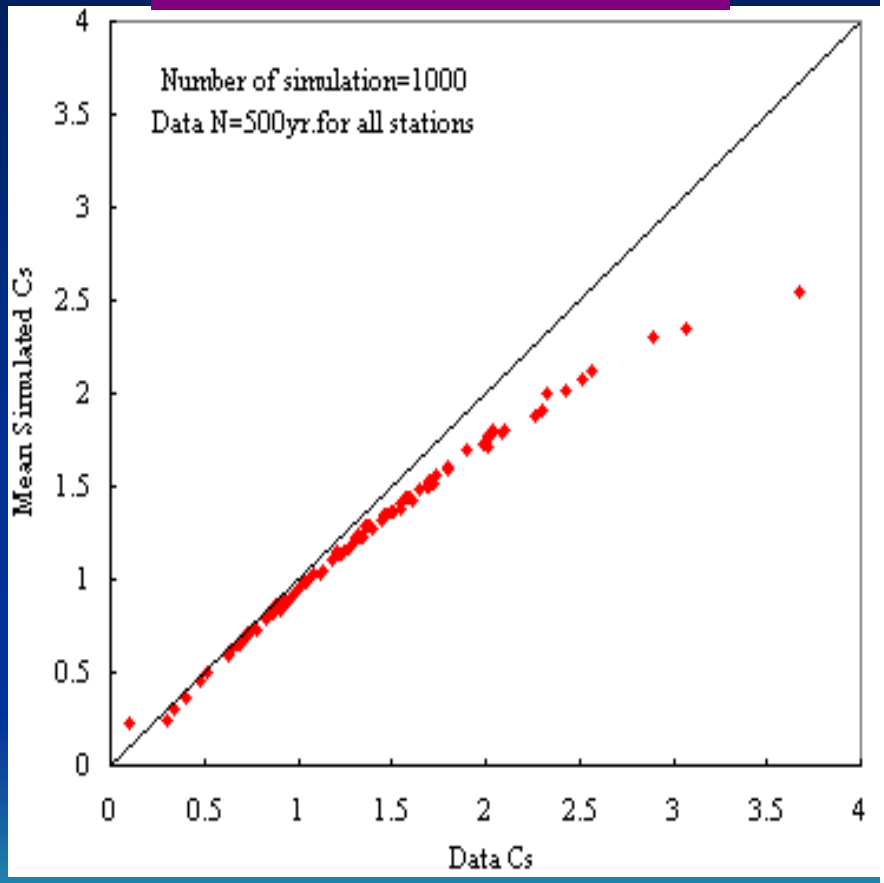


Fig.3 Difficult to model the outlier by CMM

Taihu 96 daily max, fit GLO

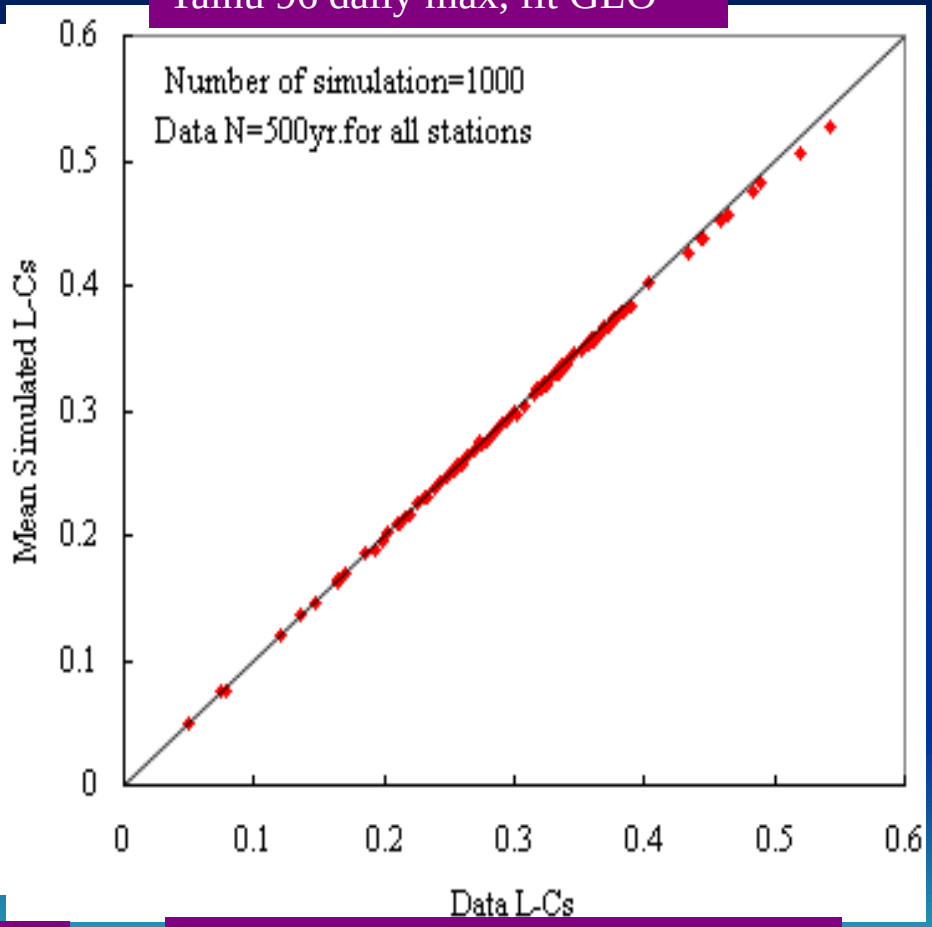


Fig.4 Robustness to outlier by LMM

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)

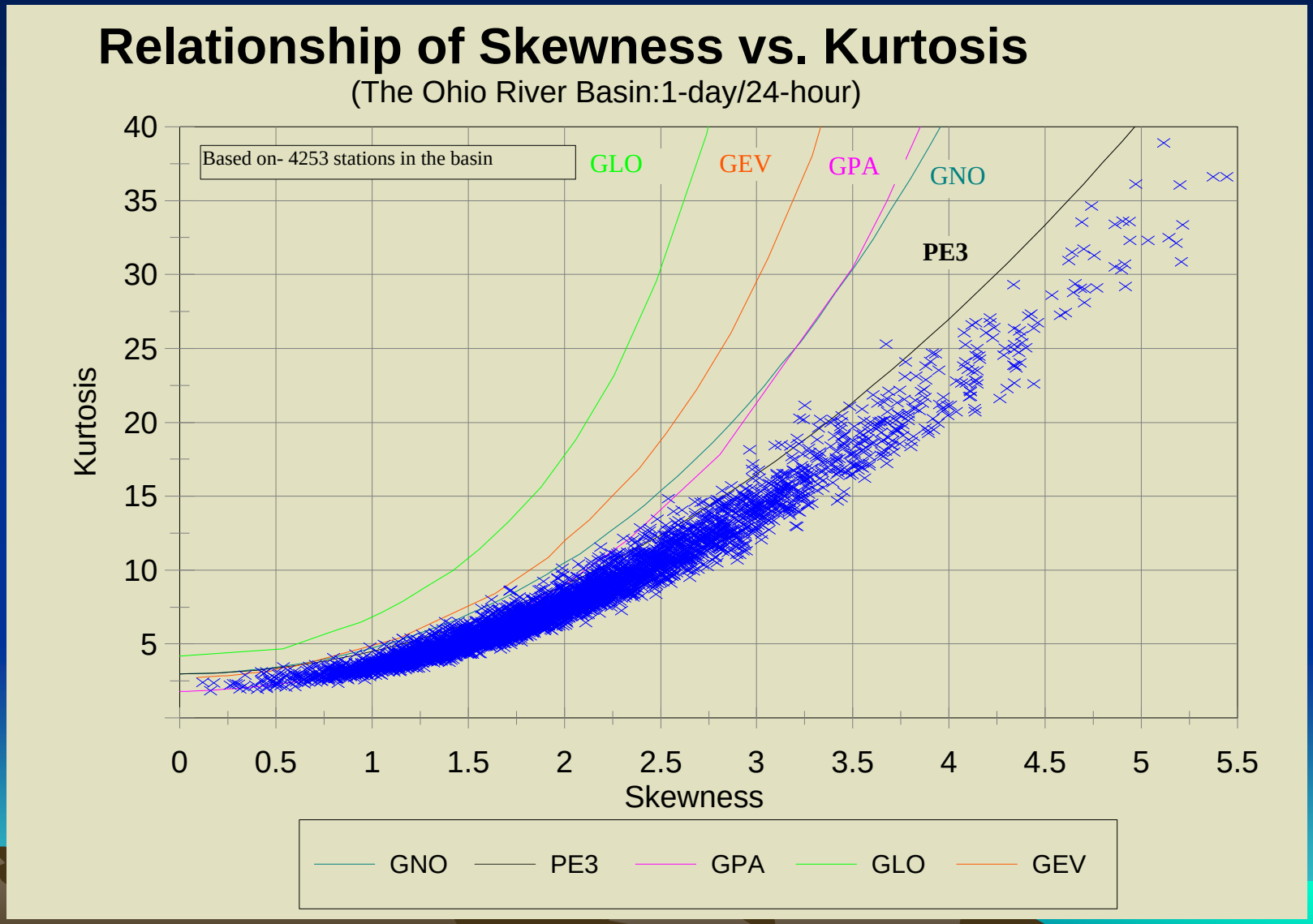


Fig. 17 Diagram of Cs vs. Ck for daily stations in the OH area (all points)

PE3 is the best-fit when CMM used (1-2)

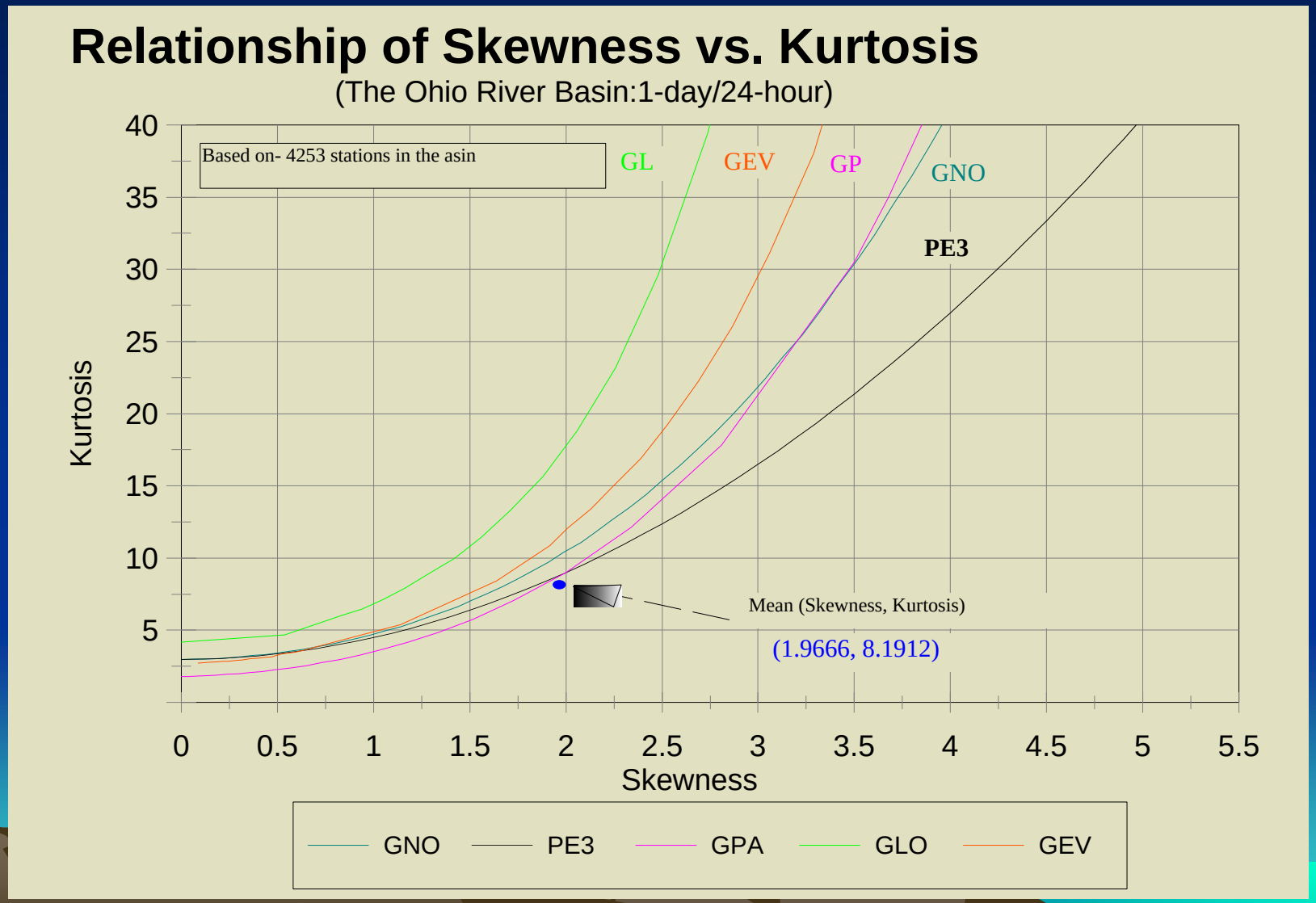


Fig. 18 Diagram of Cs vs. Ck for daily stations in the OH area (mean point)

GEV is the best-fit when LMM used (2-1)

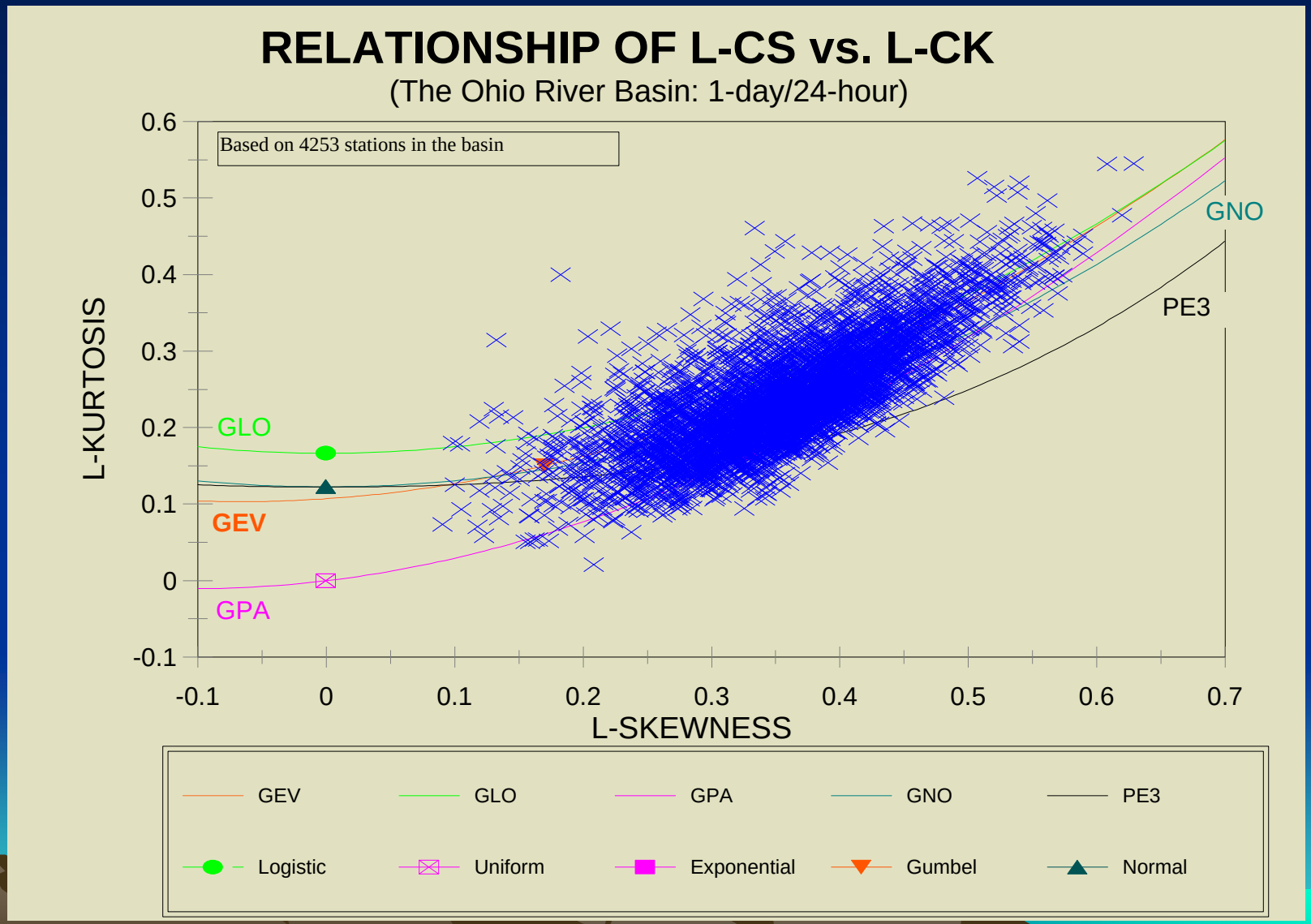


Fig. 19 Diagram of L-Cs vs. L-Ck for daily stations in the OH area (all points)

GEV is the best-fit when LMM used (2-2)

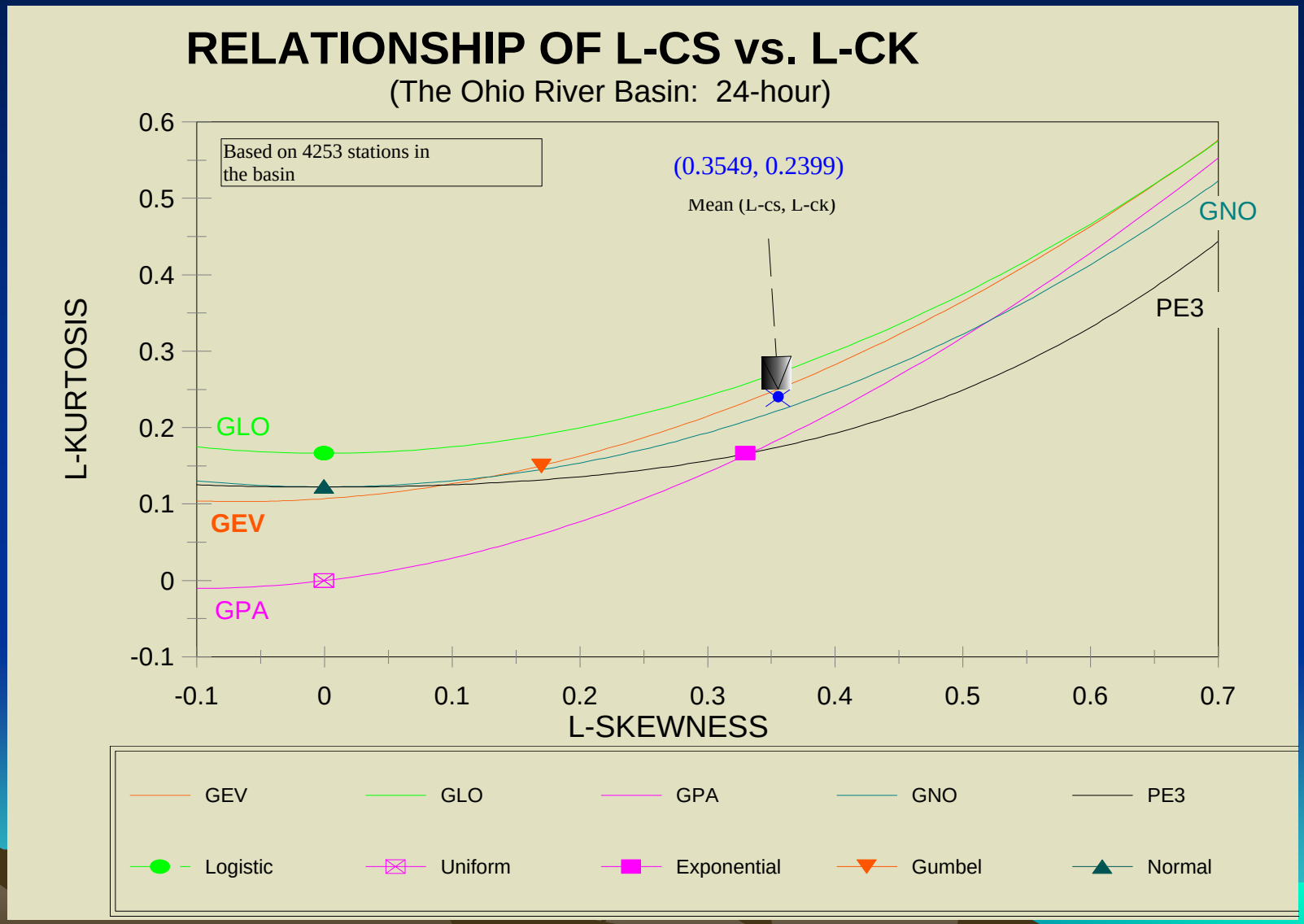


Fig. 20 Diagram of L-Cs vs. L-Ck for daily stations in the OH area (mean point)

What we should/can do? (2)

Regional Analysis – focusing on the issue of ***accuracy*** in terms of uncertainties of quantiles

Regionalization – Homogeneous Regions

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.

$$Q_{T,i,j} = q_{T,i} * \overline{x}_{i,j}$$

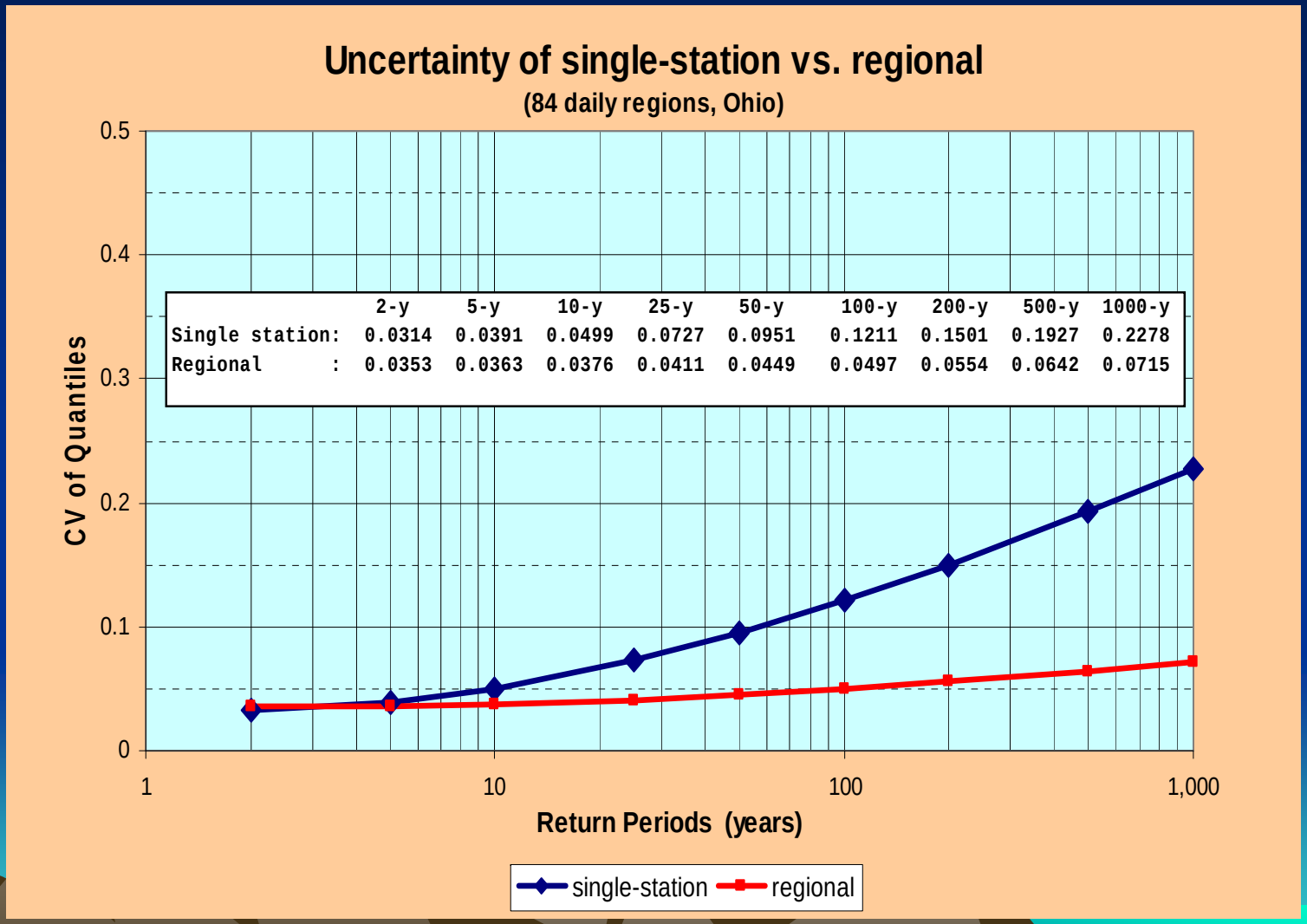
$$q_{T,i} = \frac{Q_{T,i,j}}{\overline{x}_{i,j}}$$

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 sites by “superposition” to get their own probability distribution；
4. Calculate the quantiles at each individual sites.

Regionalization →*Reduces quantile uncertainty*

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



	2-y	5-y	10-y	25-y	50-y	100-y	200-y	500-y	1000-y
Single station:	0.0314	0.0391	0.0499	0.0727	0.0951	0.1211	0.1501	0.1927	0.2278
Regional :	0.0353	0.0363	0.0376	0.0411	0.0449	0.0497	0.0554	0.0642	0.0715

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

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)

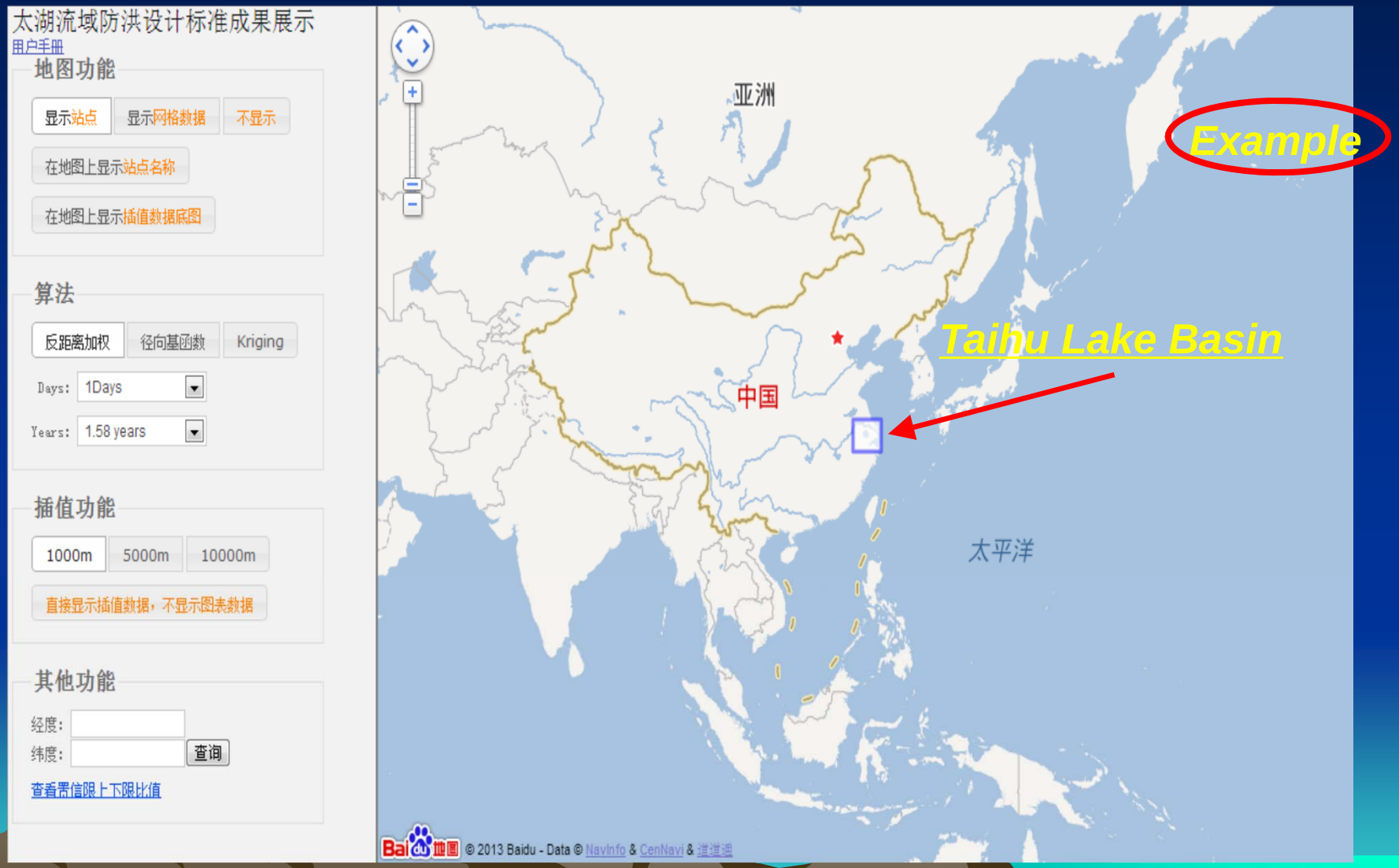
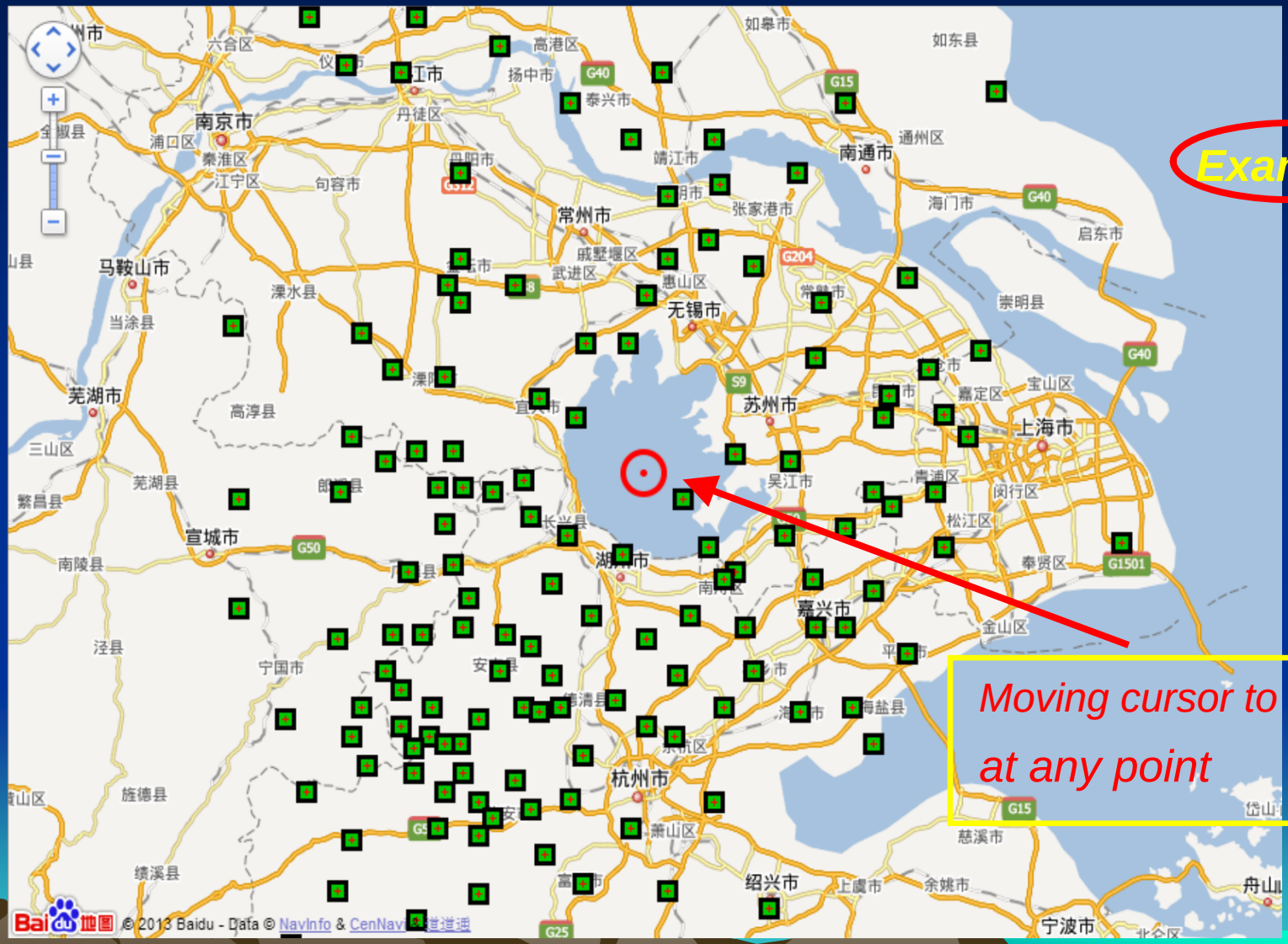


Fig. 22 Home page of online FA products for Taihu Lake Basin

Online FA Deliverables (2)



Example

Moving cursor to click at any point

Fig. 23 Online FA products for Taihu Lake Basin (1)

Online FA Deliverables (3)

大浦口(119.95,31.32)

Precipitation Frequency Estimates(mm)

ARI(Year)	Y1_58	Y2	Y5	Y10	Y25	Y50	Y100	Y200	Y500	Y1000	Y5000	Y10000
1 Day	73.57	92.46	121.48	146.97	186.19	220.61	259.47	304.18	372.78	433	606.75	[illegible]
3 Day	97.62	120.52	155.67	186.52	233.93	275.5	322.41	376.34	459	531.5	740.42	851.45
7 Day	126.77	157.18	205.22	248.57	317.11	378.91	450.37	534.5	667.2	786.88	1146.4	1345.07
10 Day	145.67	180.05	232.03	276.93	344.87	403.52	468.83	542.89	675.58	794.29	1153.6	1349.29

Example

A result-window pop-up

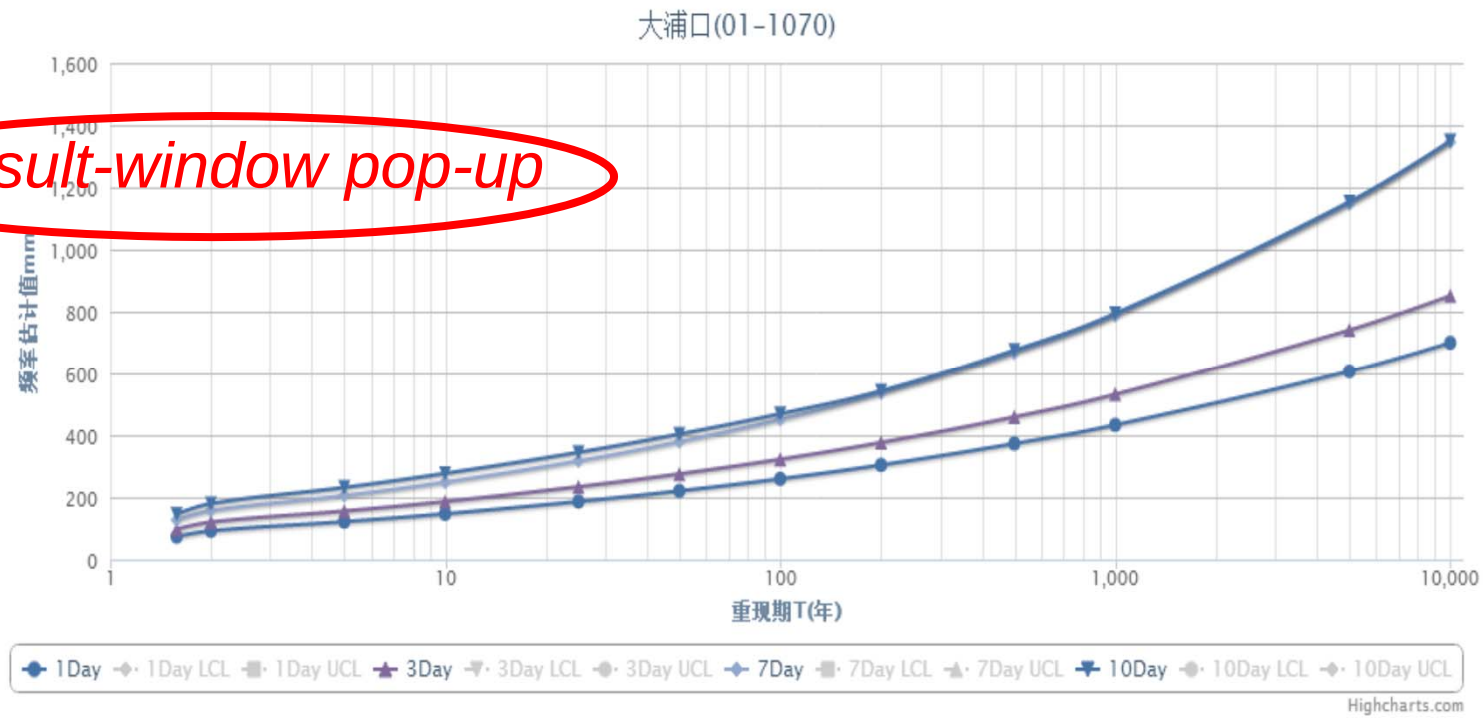


Fig. 24 Online FA products for Taihu Lake Basin (2)

Online FA Deliverables (4)

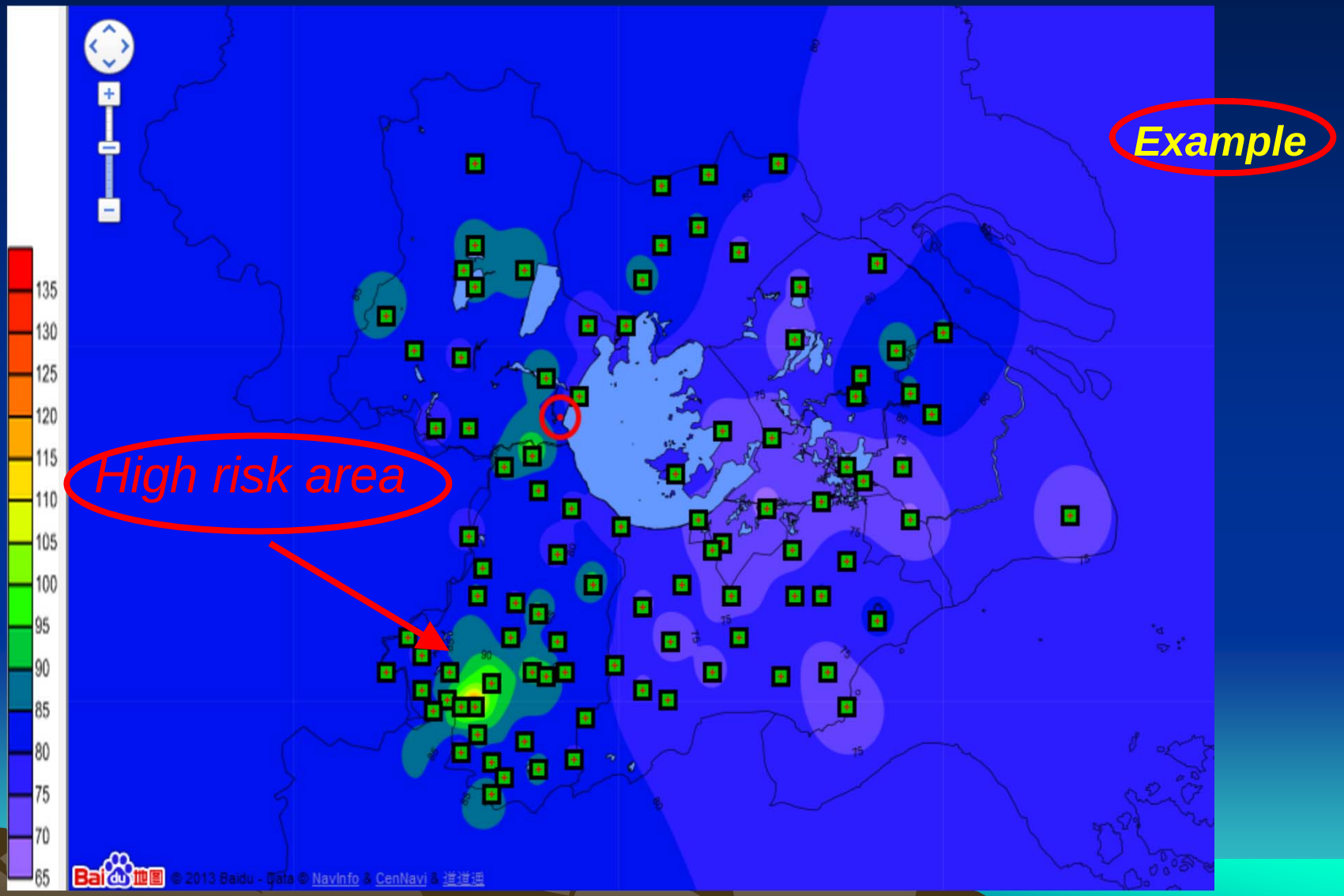


Fig. 25 Online FA products for Taihu Lake Basin (3)

Precipitation FA for *Guangxi* in SW China (1)

Example

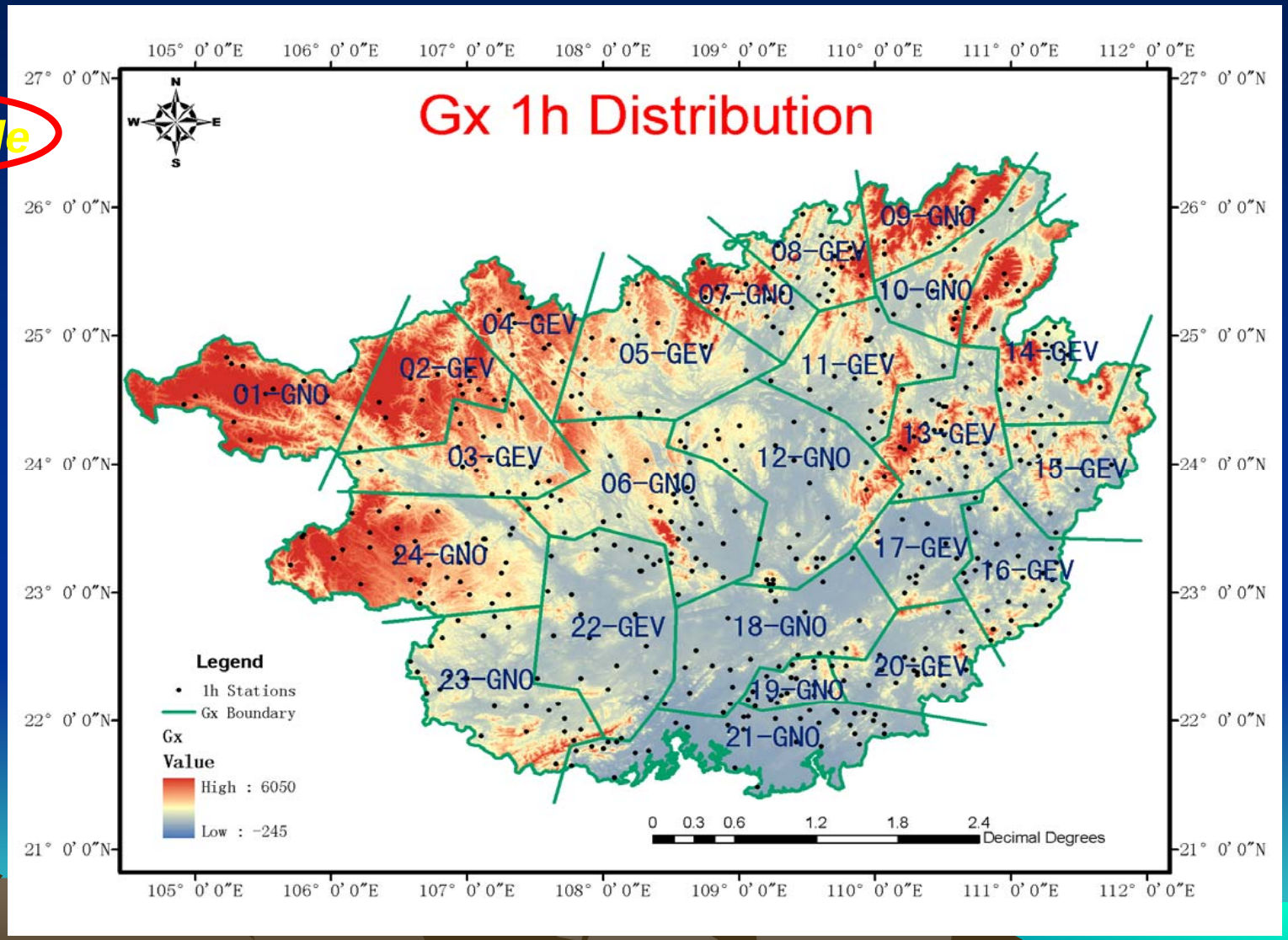


Fig. 26 Online FA products for Guangxi in SW China (1)

Precipitation FA for *Guangxi* in SW China (2)

Example

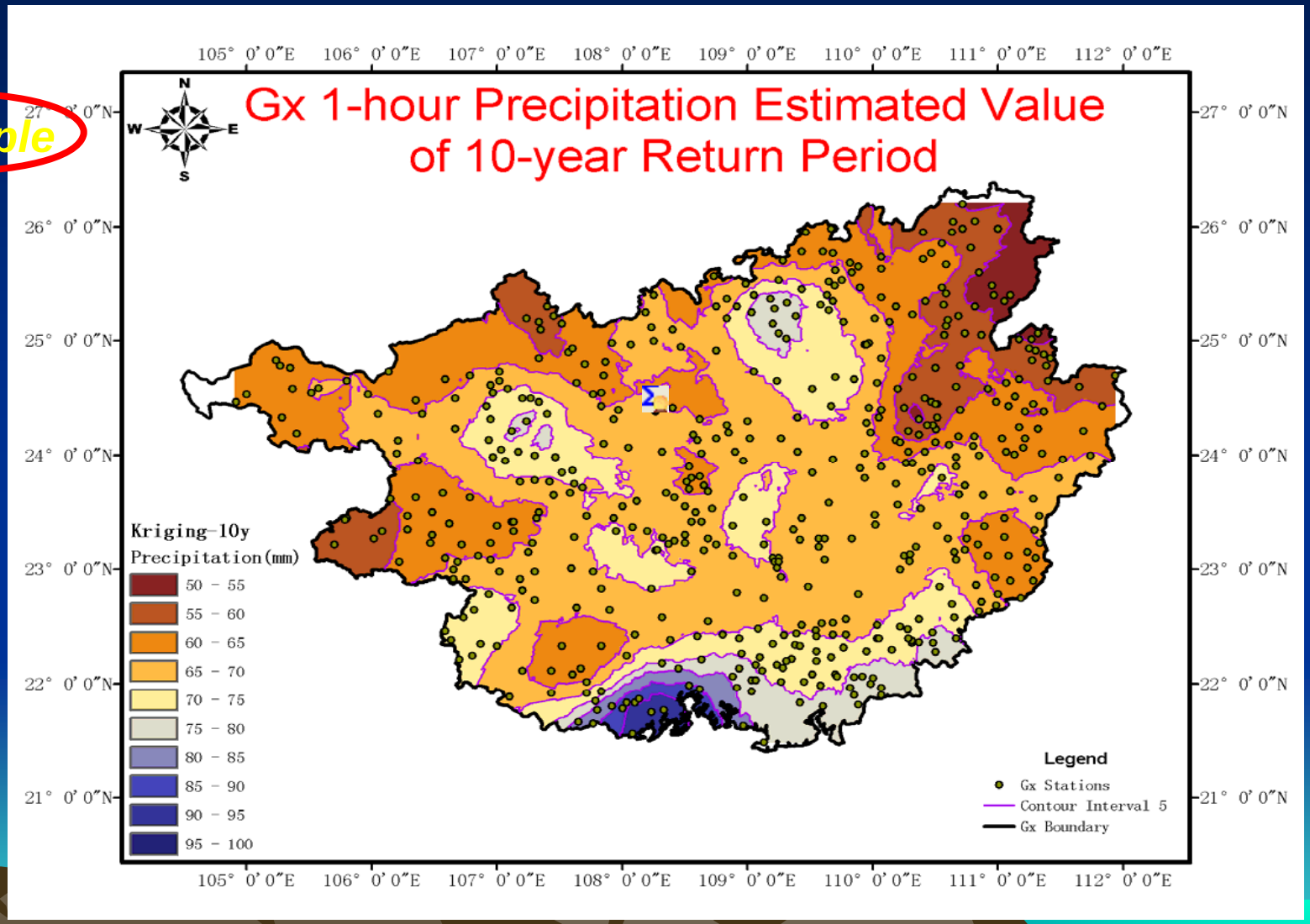


Fig. 27 Online FA products for Guangxi in SW China (2)

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 (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 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.

(WMO No.-1045, *Manual on Estimation of PMP*, Geneva, 2009)

***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)

(I-a) Moisture maximization	Maximum 12-hr persisting dew point (HYDROME)
(I-b) Storm transposition	Storm Separation + Adjustments (HYDROME)
(I-c) Use of D-A-D curves	Envelopment (HYDROME)
(II) Statistical approach	Modified frequency analysis (STAT)

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为降雨率mm/s; v_1为该层入流风速 m/s; △P为该层入流压力差 hPa; $\bar{q}_1$ & $\bar{q}_2$ 为入、出流平均比湿 g/kg；g为重力加速度cm/s²; ρ 为水的密度 g/cm³；Y为该山体的水平深度 m。

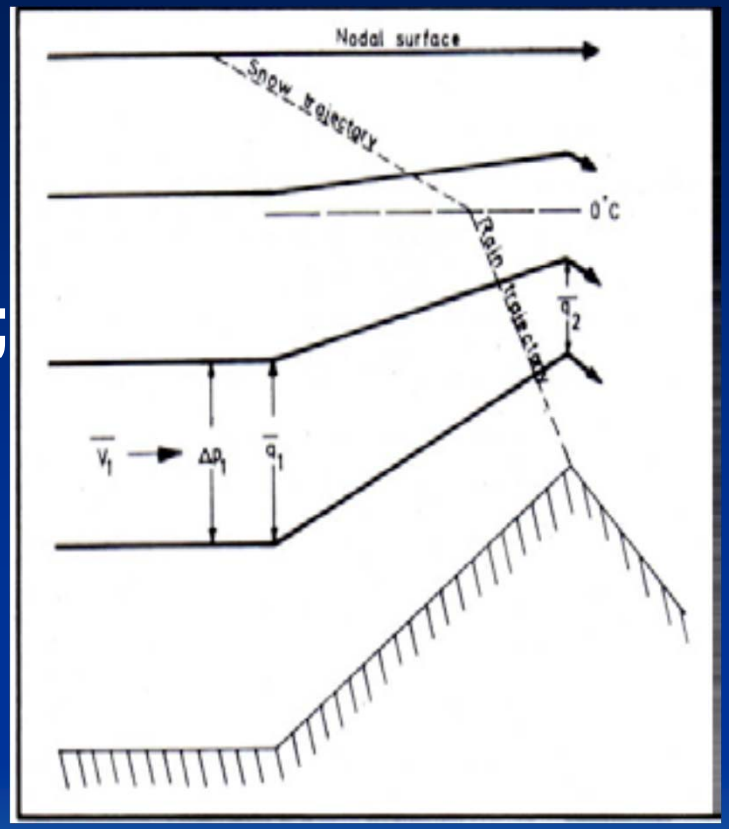


Fig. 28 Sketch of Laminar Model

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

（*NWS once applied unsuccessfully the Laminar Model to account the terrain effect on rainfall in a PMP study in 1961 and 1966)

Synoptic +Statistics+ Orographic

For a storm rainfall, the rainfall intensity for a given point P(x,y) in a drainage at any time can be defined by

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

（SDOIF method*）

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

$$\overline{R}_{\Delta t,A}=\frac{\iint_A r_{\Delta t}(x,y)dxdy}{\iint_A dxdy}=\frac{\iint_A r_{0,\Delta t}(x,y)\times f_{\Delta t}(x,y)dxdy}{\iint_A dxdy}\approx\frac{\sum_i^m\sum_j^n r_{0,\Delta t}(x_i,y_j)\times f_{\Delta t}(x_i,y_j)dxdy}{\sum_i^m\sum_j^n \Delta x_i\Delta y_j}$$

（* Lin，Bingzhang， WMO NO-1045，Geneva，2009）

10 Tasks for the PMP Study in HK

Example

- 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

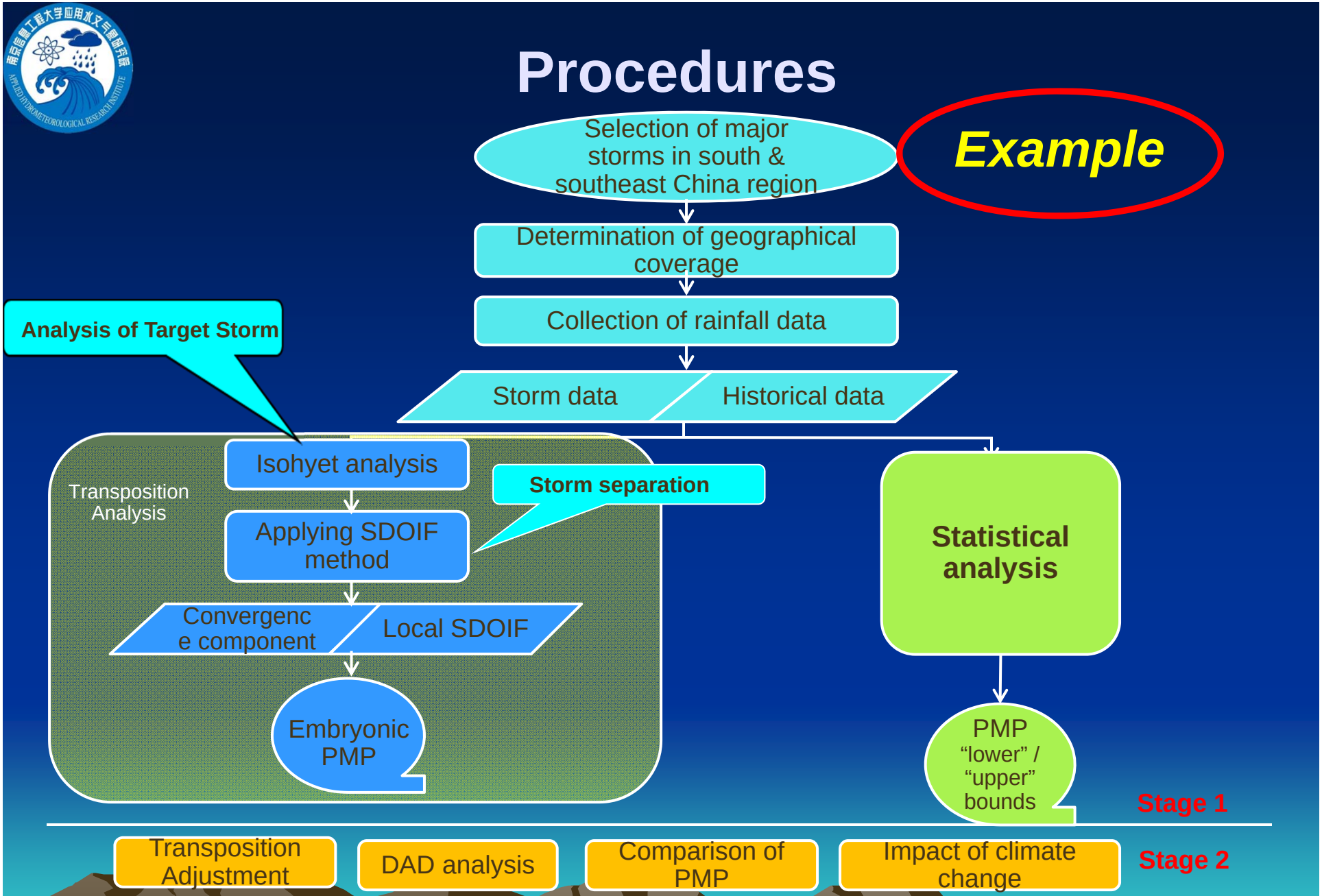


Fig. 29 Sketch of PMP estimation procedure for HK as example

Moving track of Morakot Typhoon

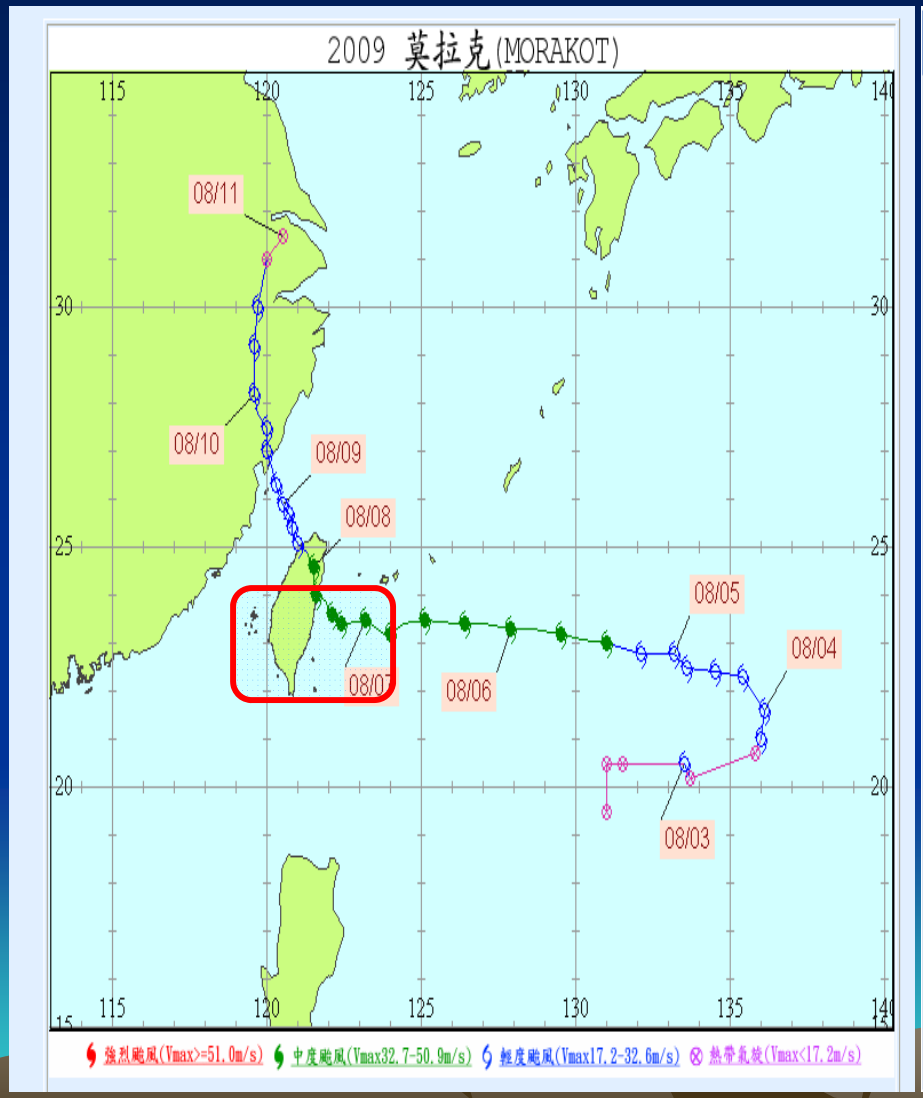


Fig. 30 Track of Morakot Typhoon, 2009-08-03~11

World, China mainland and Taiwan rainfall (By 2009)

Location	Date	Duration	Amount	Note
Shinliao, Taiwan (north)	October 17, 1967	24-hour	**1,672 mm**	**Taiwan record**
Jiayi Alishan, Taiwan	**August 8, 2009**	**24-hour**	**1,623 mm**	Typhoon Morakot
Linzhuang, China mainland	August 7, 1975	24-hour	1,060 mm	Sup. Typhoon Nina
Cilaos，La Reunion Island	March 15, 1952	24-hour	**1,870 mm**	Tropical Cyclone
Jiayi Alishan, Taiwan	**August 8-9, 2009**	**48-hour**	**2,361 mm**	**Typhoon Morakot**
Linzhuang, China mainland	August 7-8, 1975	48-hour	1,279 mm	Sup. Typhoon Nina
Cilaos，La Reunion Island	March 15-17, 1952	2-day	**2,500 mm**	Tropical Cyclone
Jiayi Alishan, Taiwan	**August 8-10, 2009**	**3-day**	**2,747 mm**	**Typhoon Morakot**
Linzhuang, China mainland	August 6-8, 1975	3-day	1,605 mm	Sup. Typhoon Nina
Cilaos，La Reunion Island	March 15-18, 1952	3-day	3,240 mm	Tropical Cyclone
Grand-llet, La Reunion Island	January 24-27, 1980	3-day	3,241 mm	Cyclone Hyacinthe
Commerson's Crater，La Reunion Island	February 24-26, 2007	3-day	**3,929 mm**	**Cyclone Gamede**

Issue of the Typhoon Morakot (1)

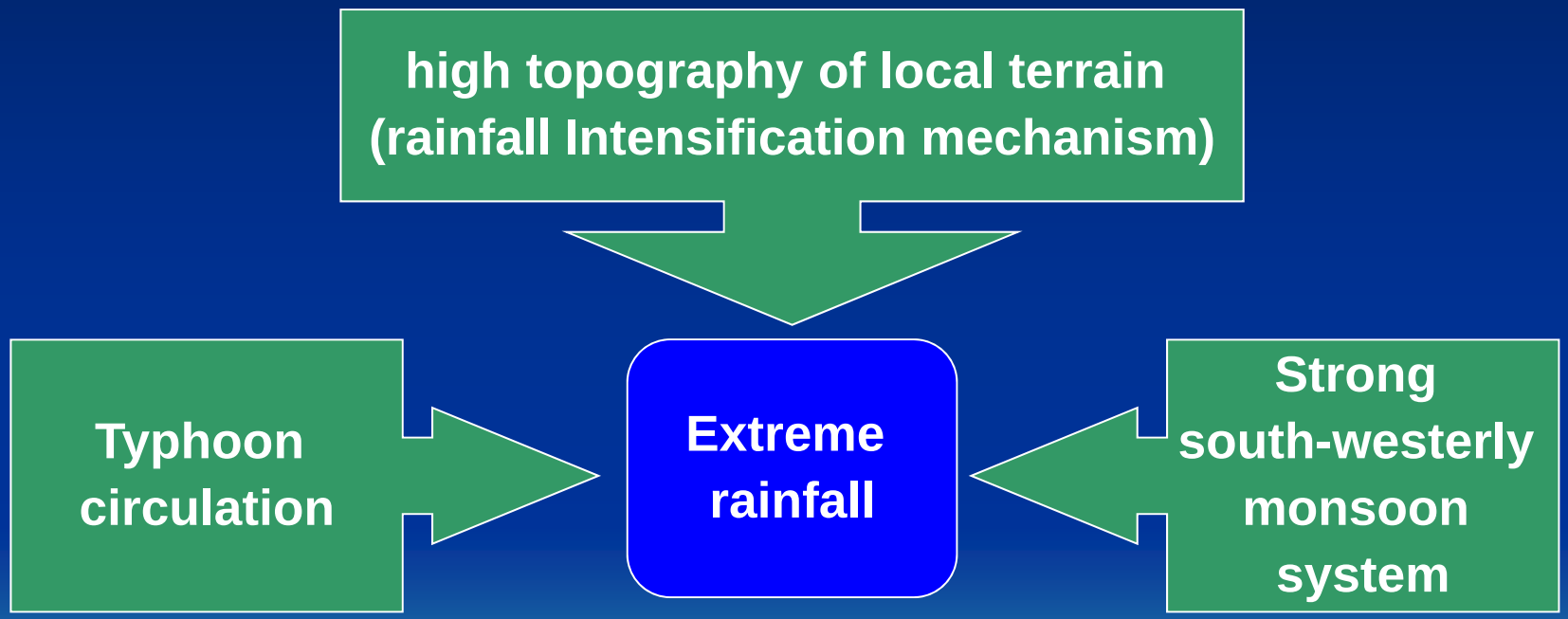
(1,000-year plus event?)

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**

Issue of the Typhoon Morakot (2)

Possible causes:



However, current typhoon intensity forecast skill is still poor because of the lack of understanding on the complex interactions between ocean and typhoons

Fig. 31 Sketch of causes for Morakot Typhoon Storm

Isohyets of 24-hr for Morakot Typhoon

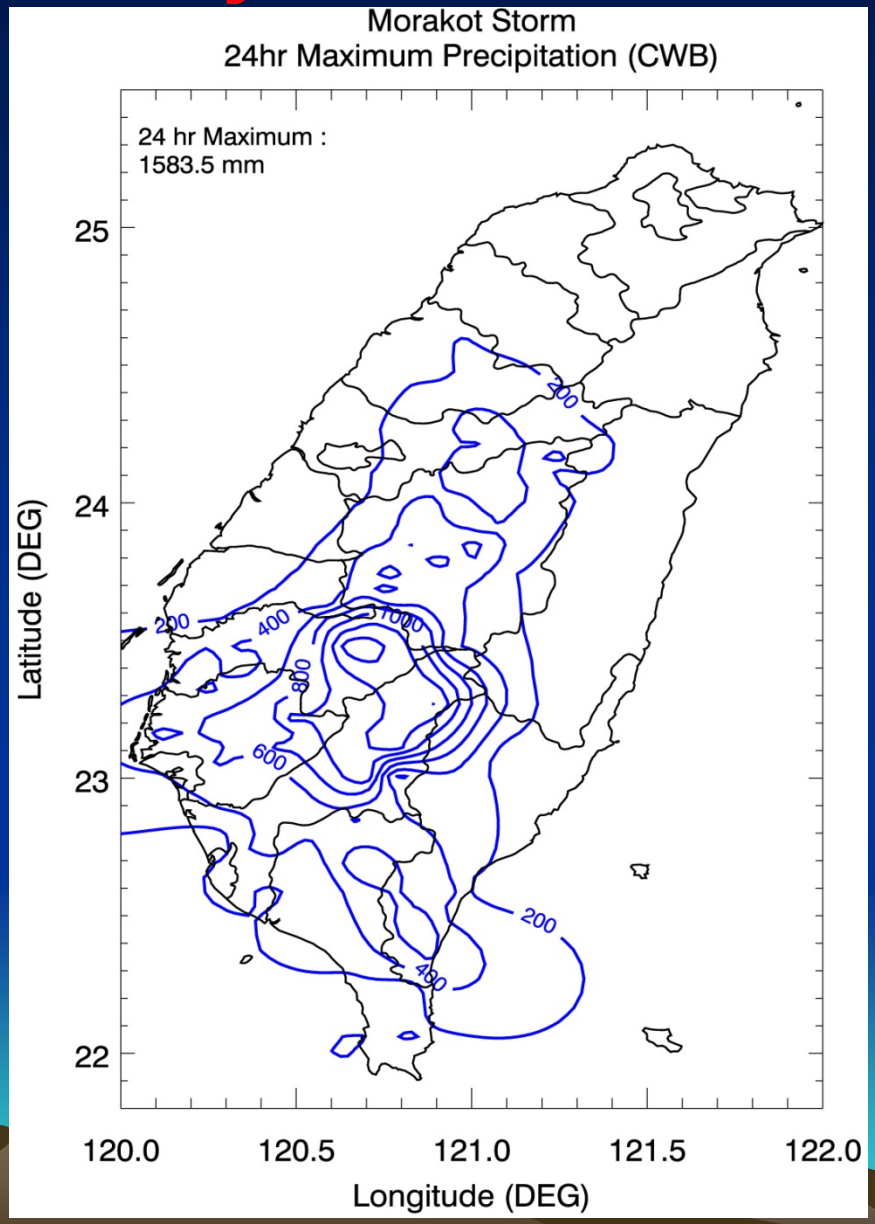


Fig. 32 24-hr isohyets of Morakot Storm

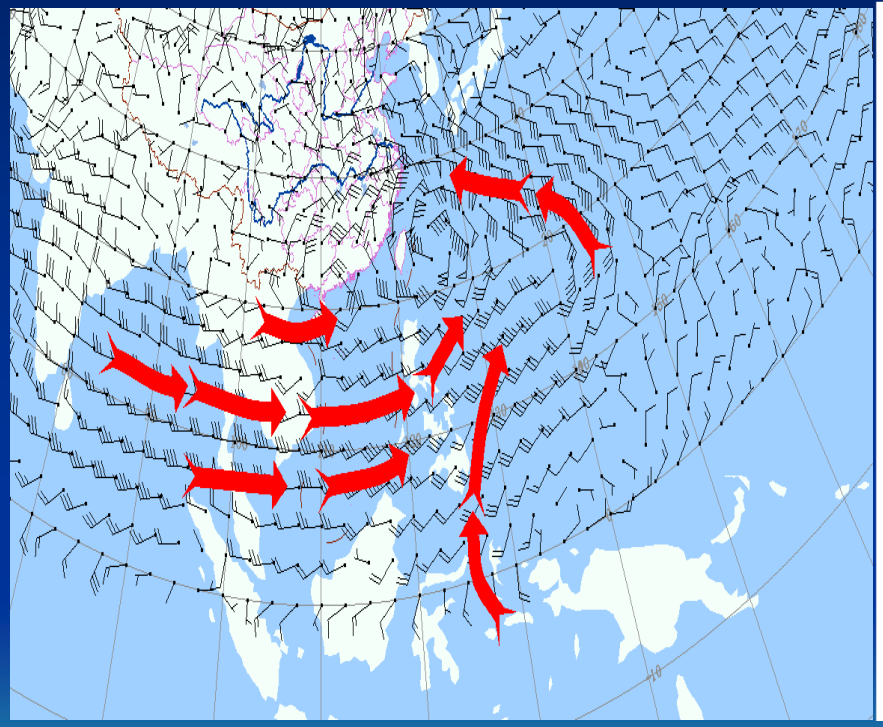
- **1,583mm / 24-hr**
- **2,372mm / 48-hr**
- **2,682mm / 72-hr**

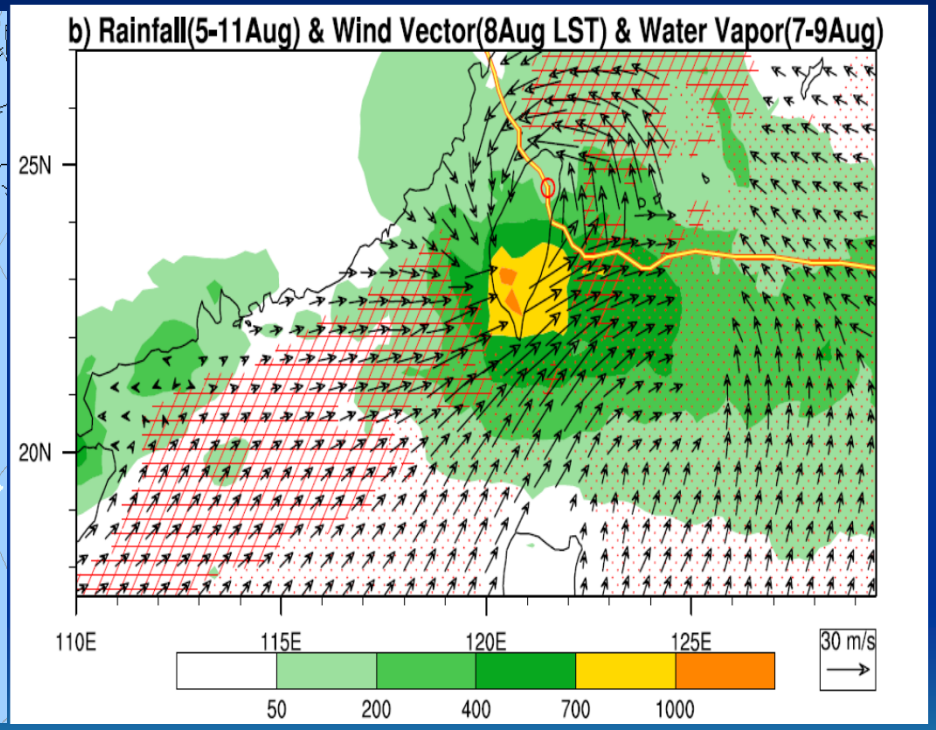
(based on hourly rainfall observations)

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

Locations of Key Raingauges in Taiwan

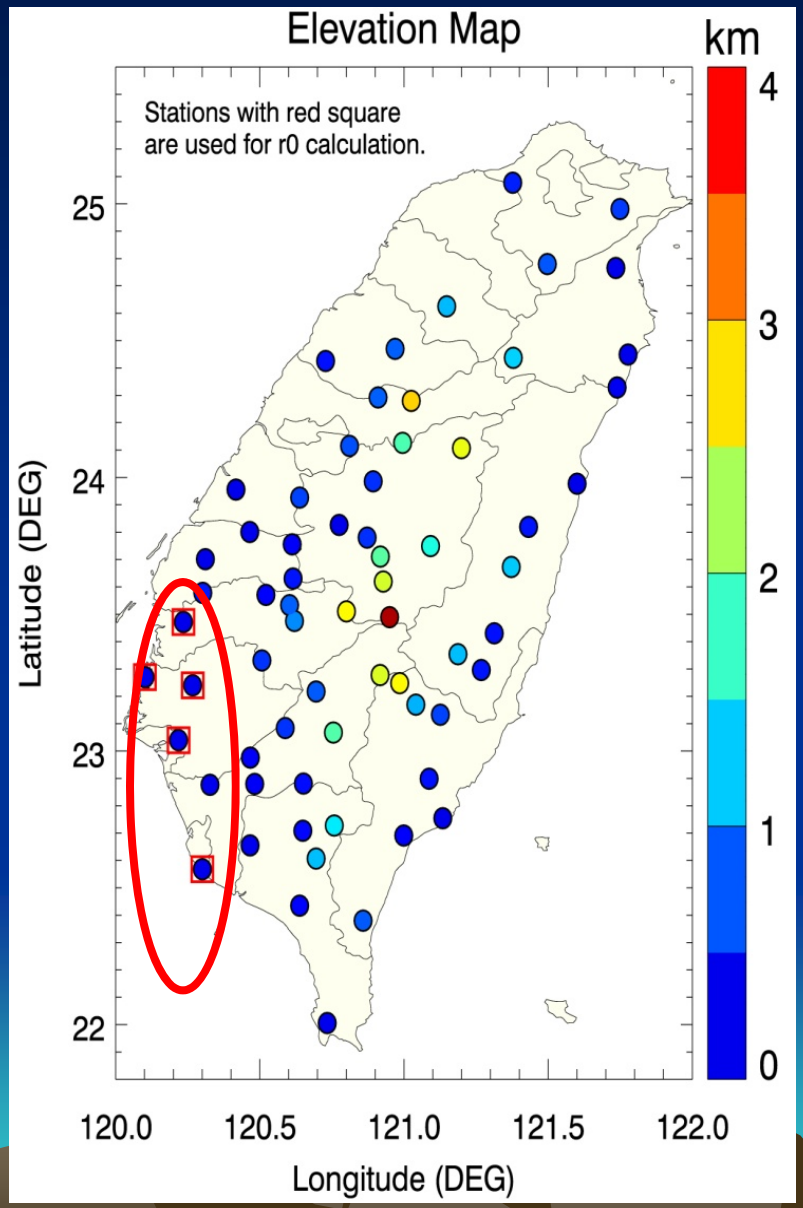
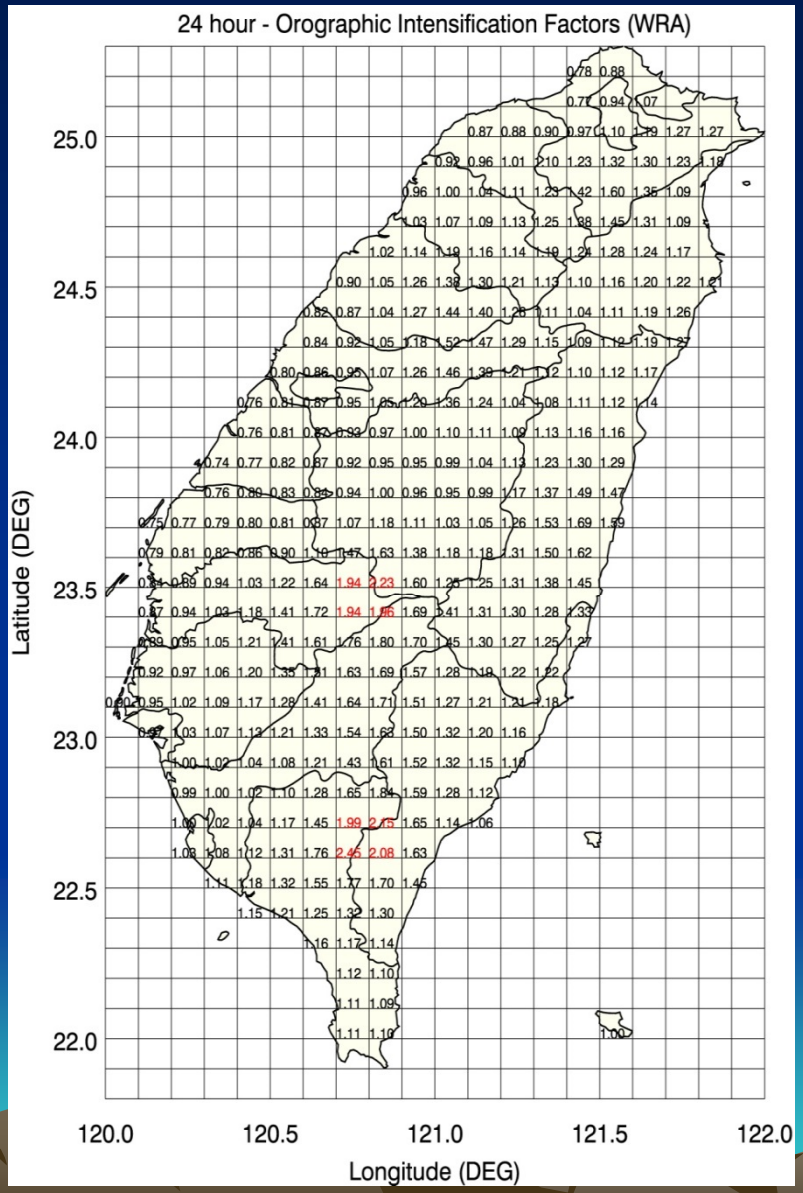


Fig. 34 Location of raingauges in Taiwan for analysis of OIF

Gridded 24h OIF for Morakot (10kmx10km)



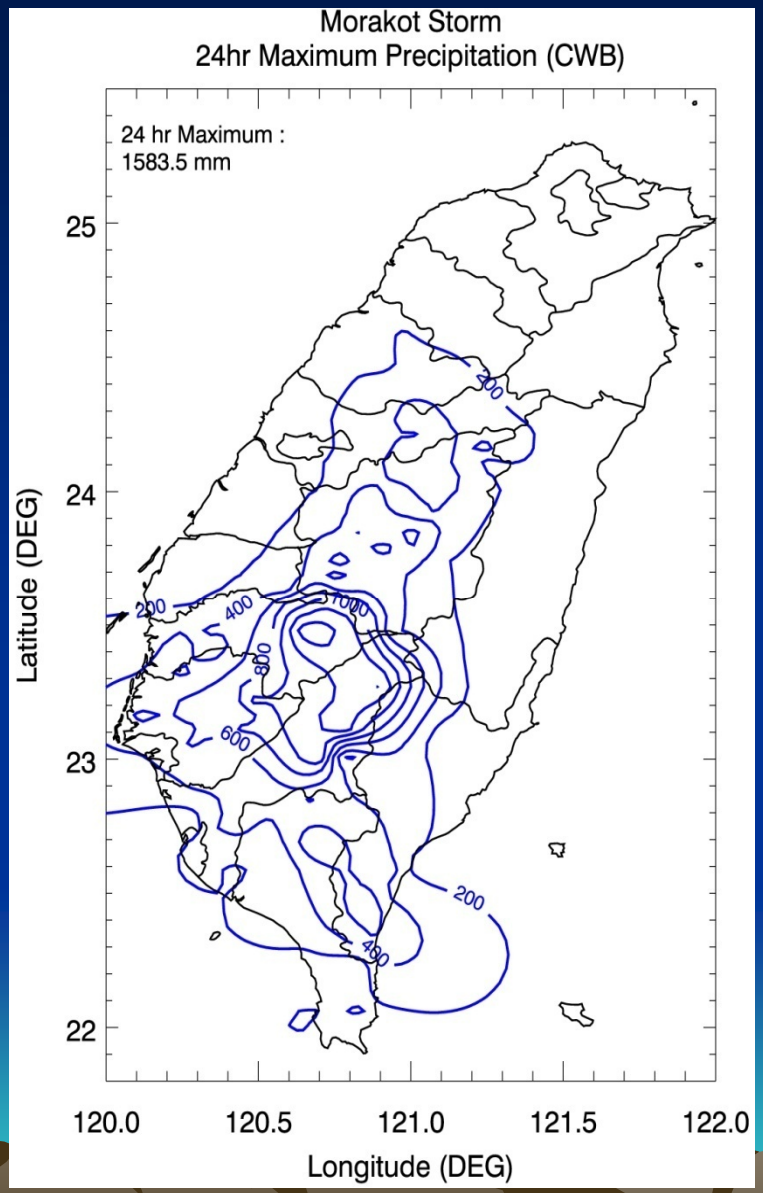
Orographic

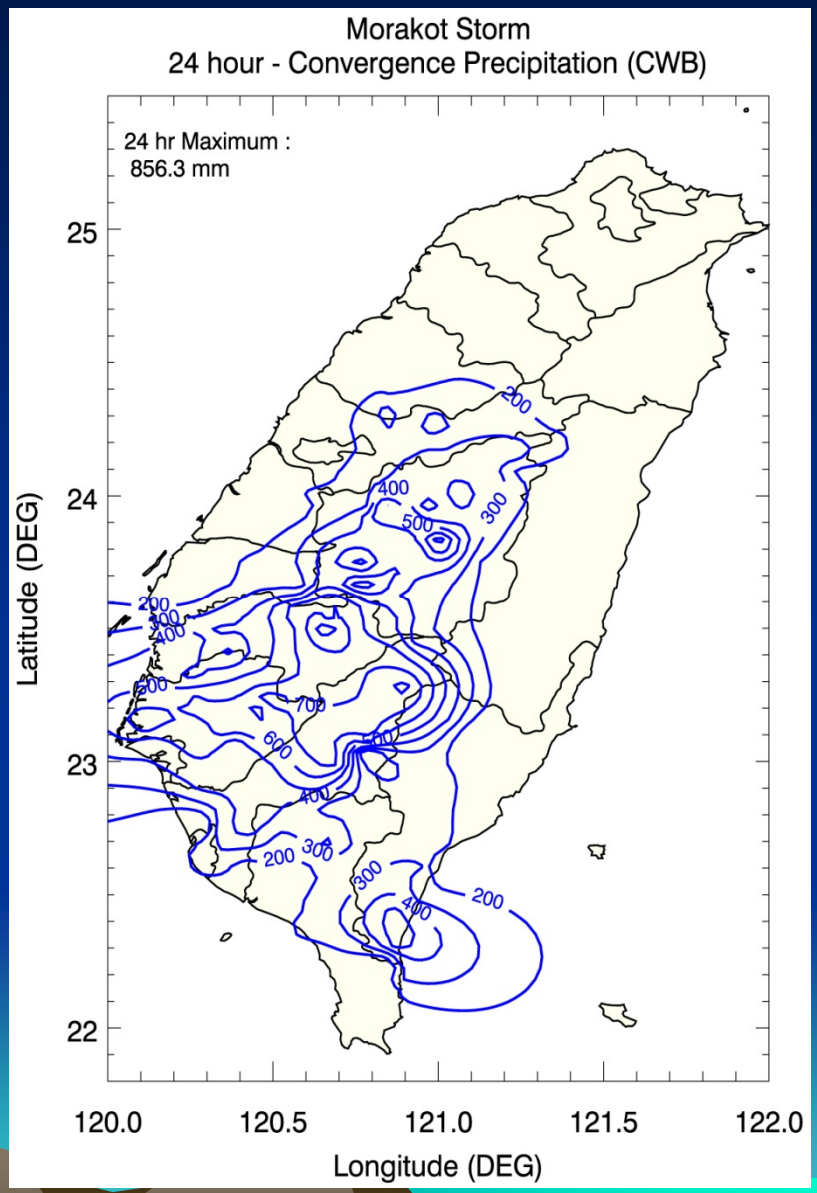
Intensification

Factor

Fig. 35 Map of 24-hr gridded OIF for Morakot Storm

Before & After Storm Separation





Before

After

Fig. 36 Before and after decomposition of 24-hr Morakot Storm rainfall

Generalized Convergence Component Pattern

(based on synoptic analysis of 4 major storms)

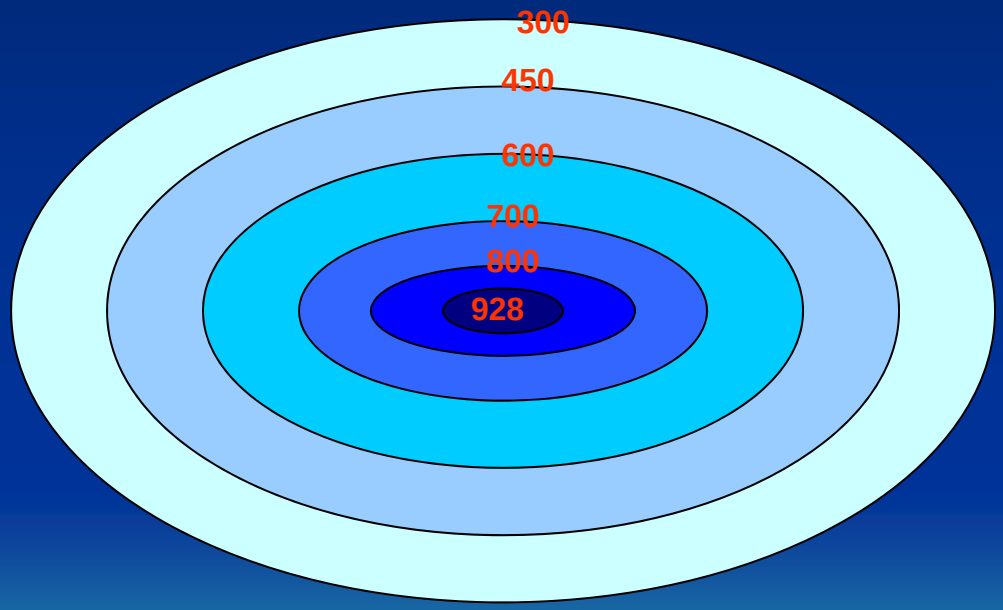


Fig. 37 Generalized convergence component pattern of Morakot Storm

Convergence Pattern can be Transposed in a Wider Region

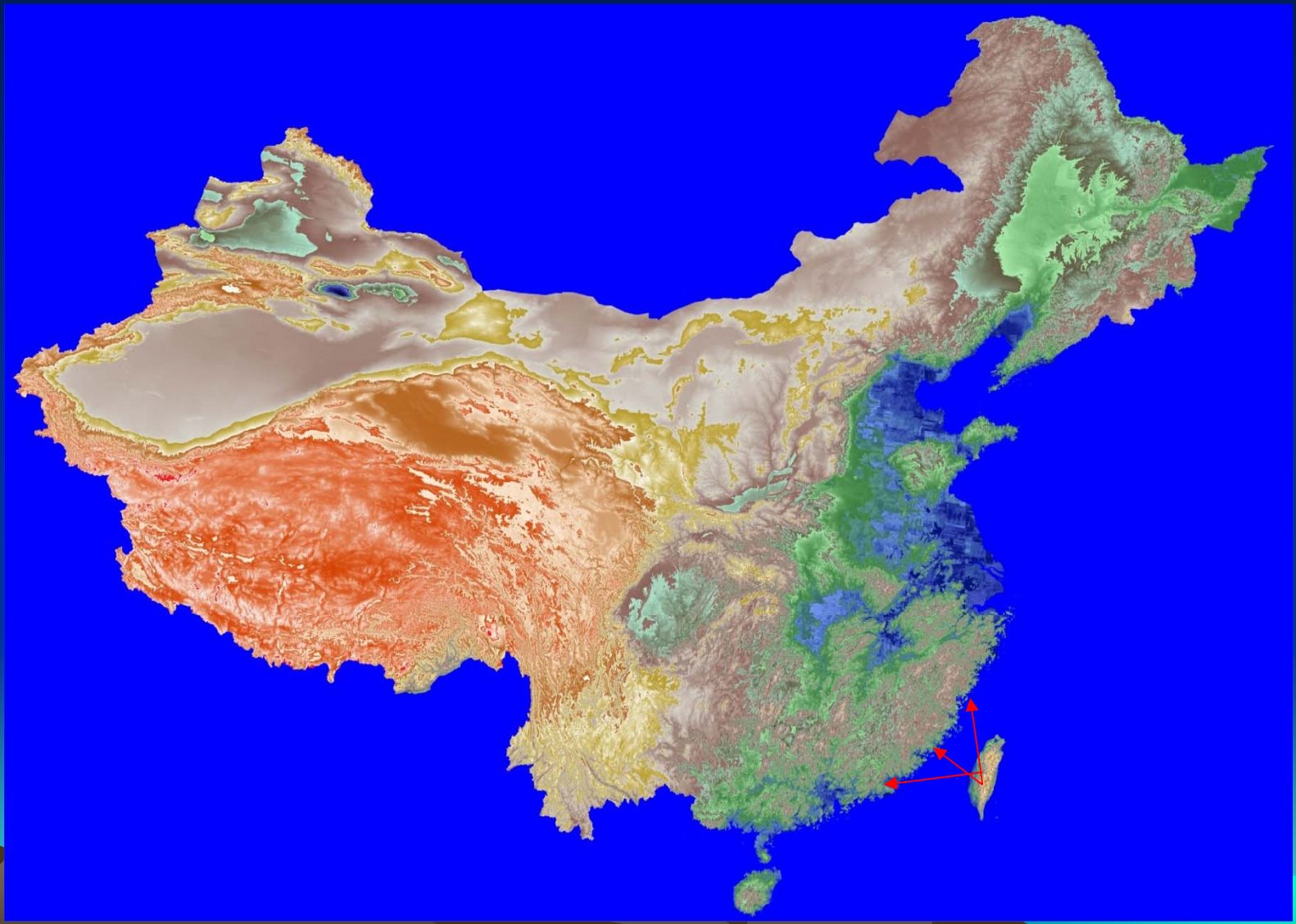


Fig. 38 Illustration of transposition for convergence pattern of Morakot Storm

Application example: Transposed to HK

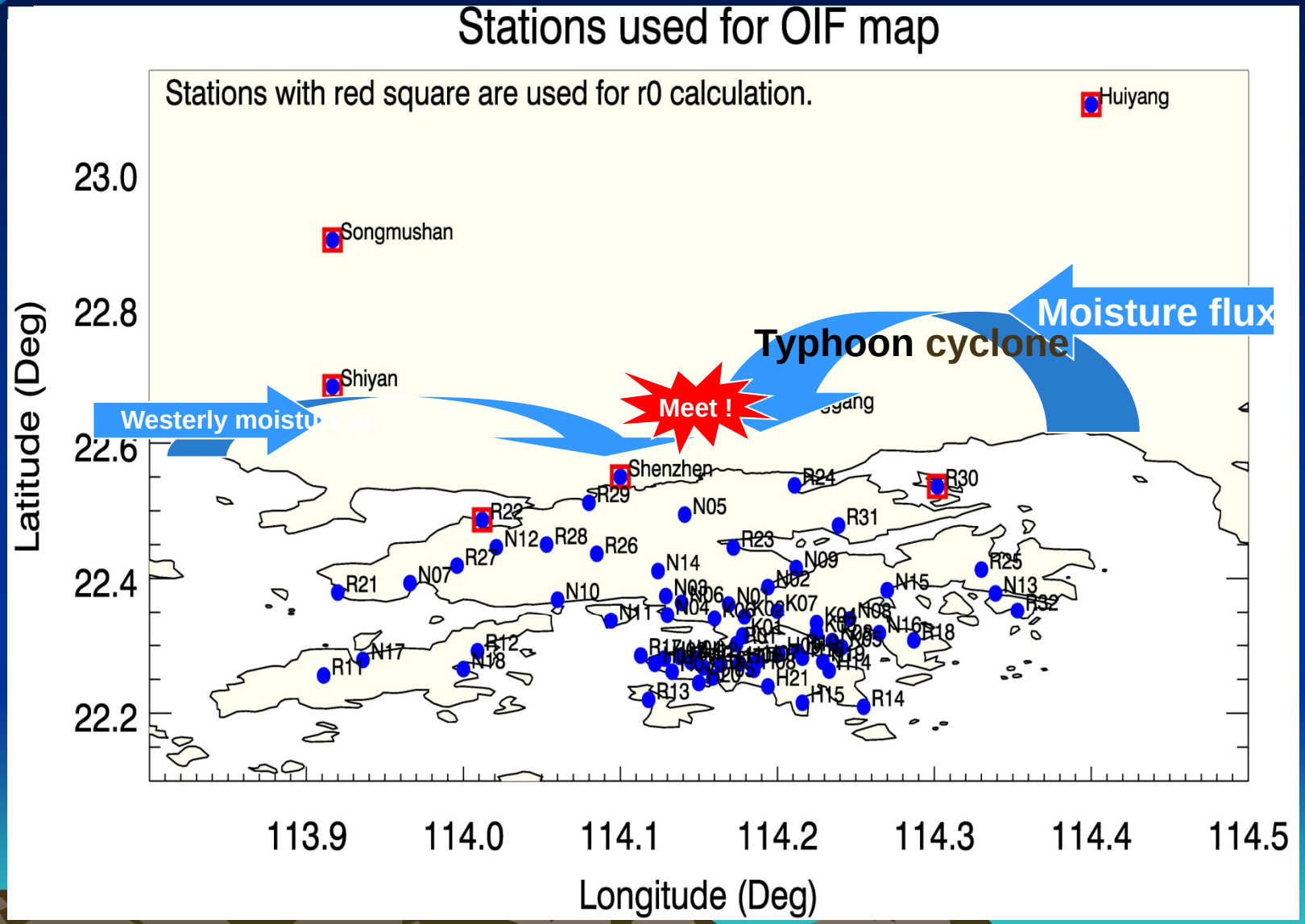
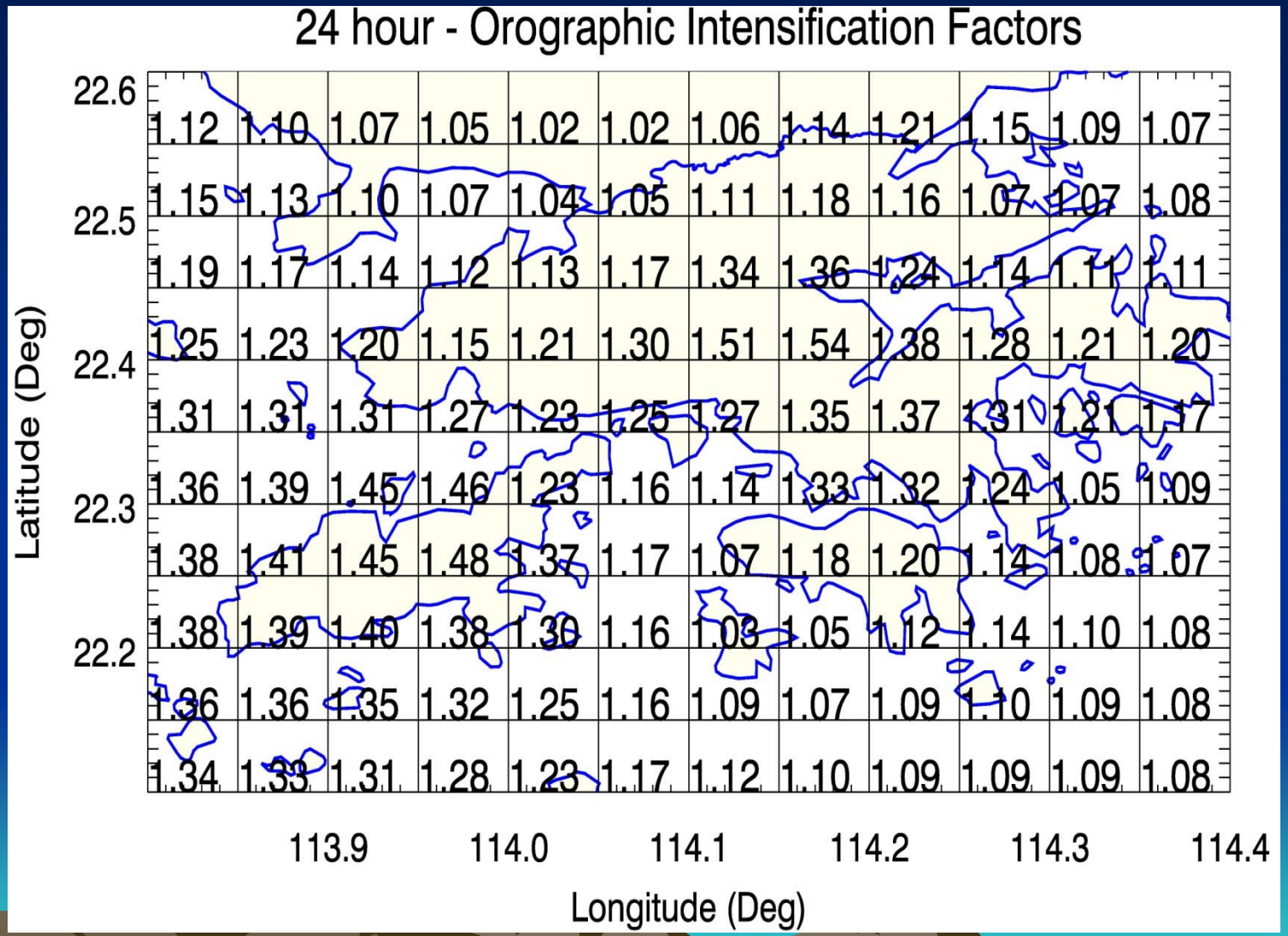


Fig. 39 Illustration of synoptic analysis for HK in terms of moisture flux

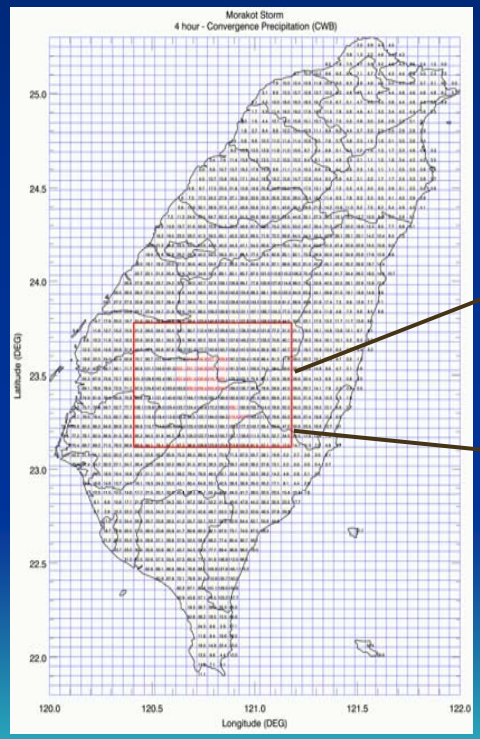
Development of OIF in Design Area, HK

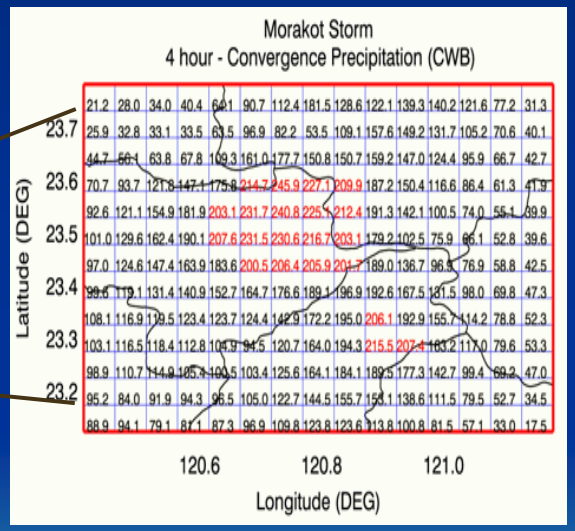
(5kmx5km)



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



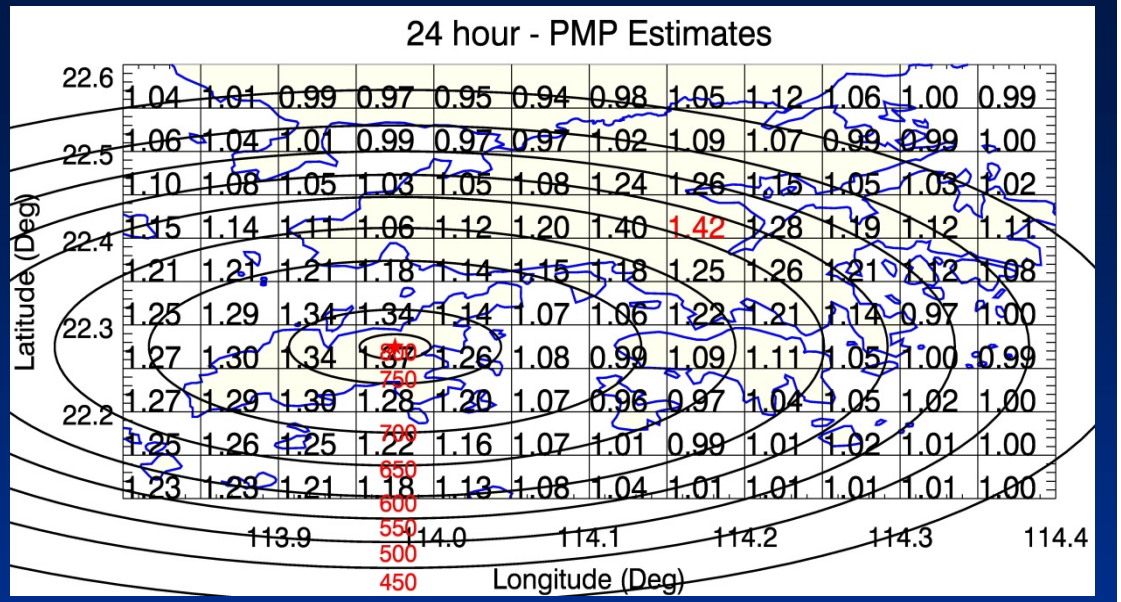


4-hr sub-mesh

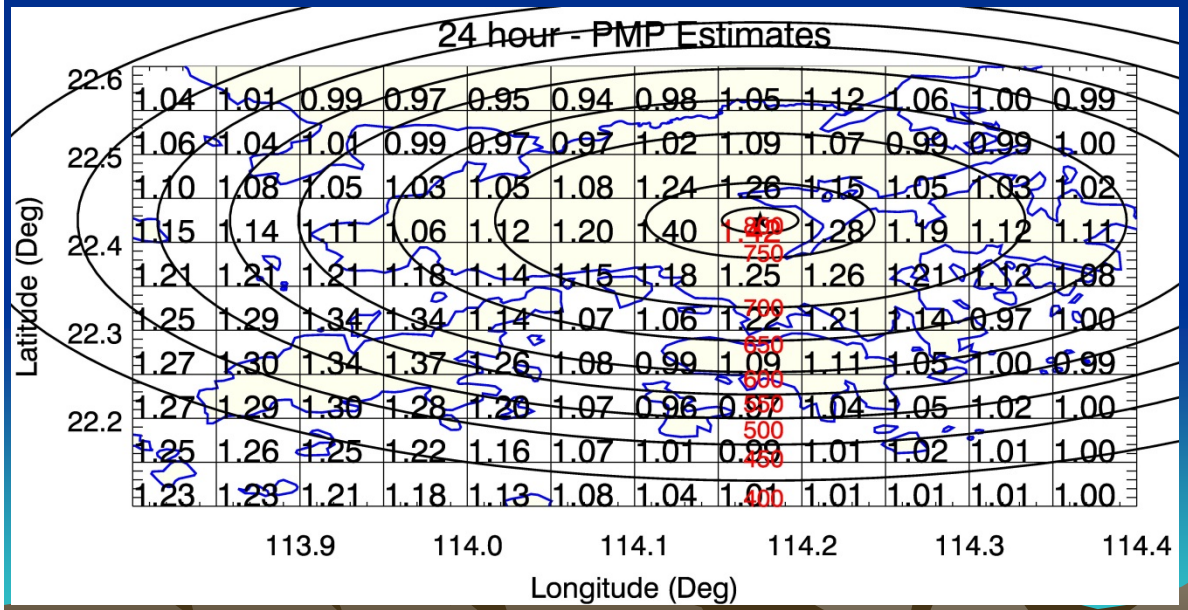
showing here is for the 4-hr Morakot pattern at resolution of 5kmx5km

Fig. 40 Illustration of cut-off sub-mesh of convergence pattern

E-W orientation with different center points （as example)



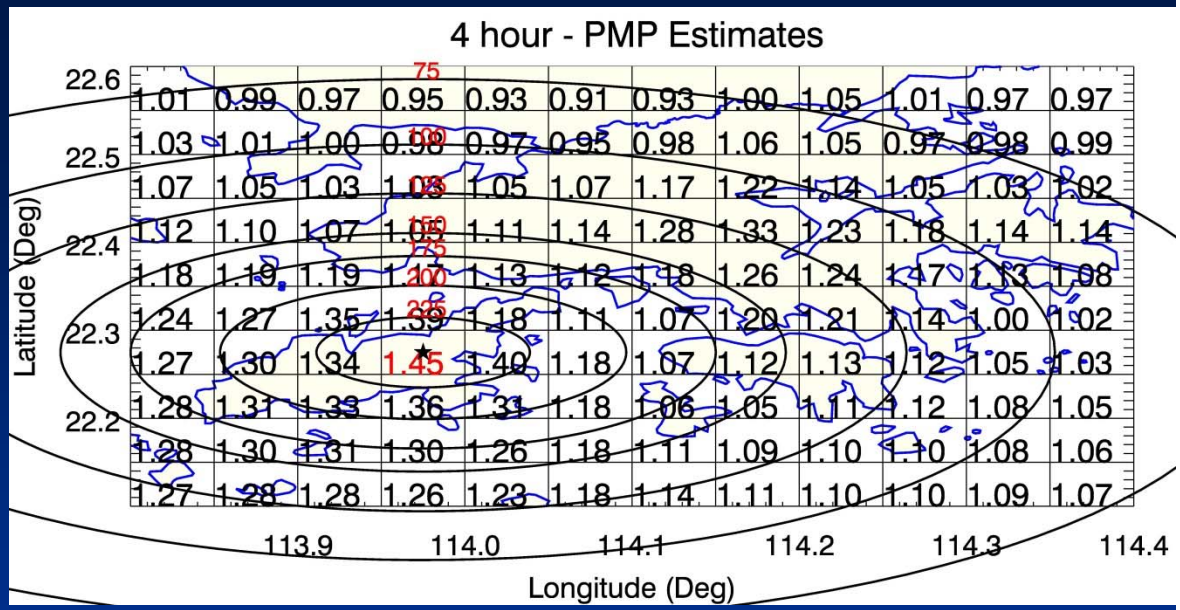
(Peak at Lantau)



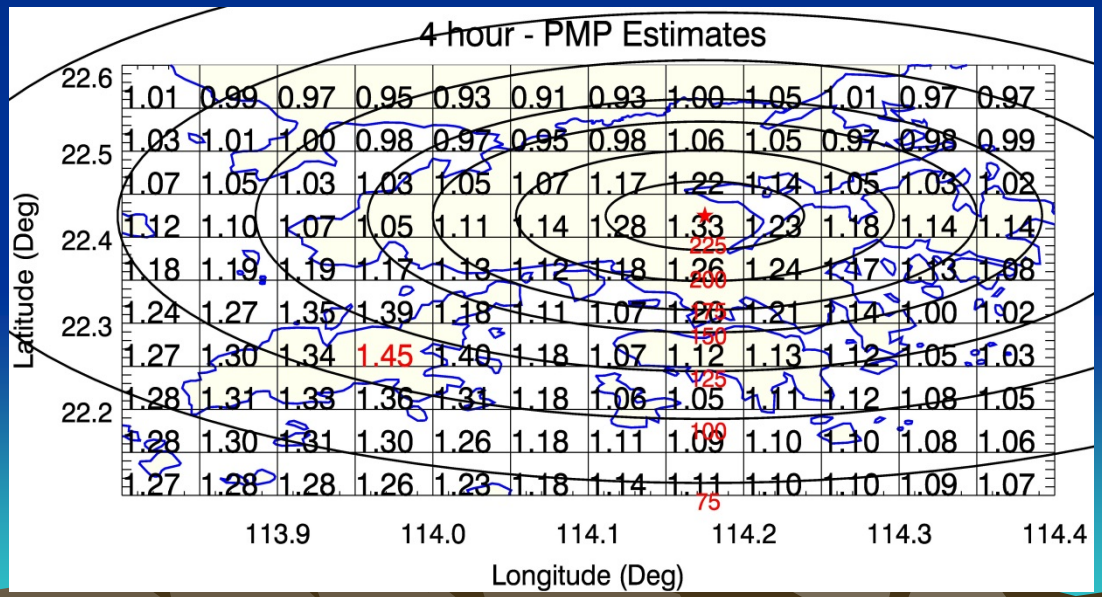
(Peak at Tai Mo Shan)

Fig. 41 Orientation of transposed convergence pattern (1)

E-W orientation with different center points （as example)



(Peak at Lantau)

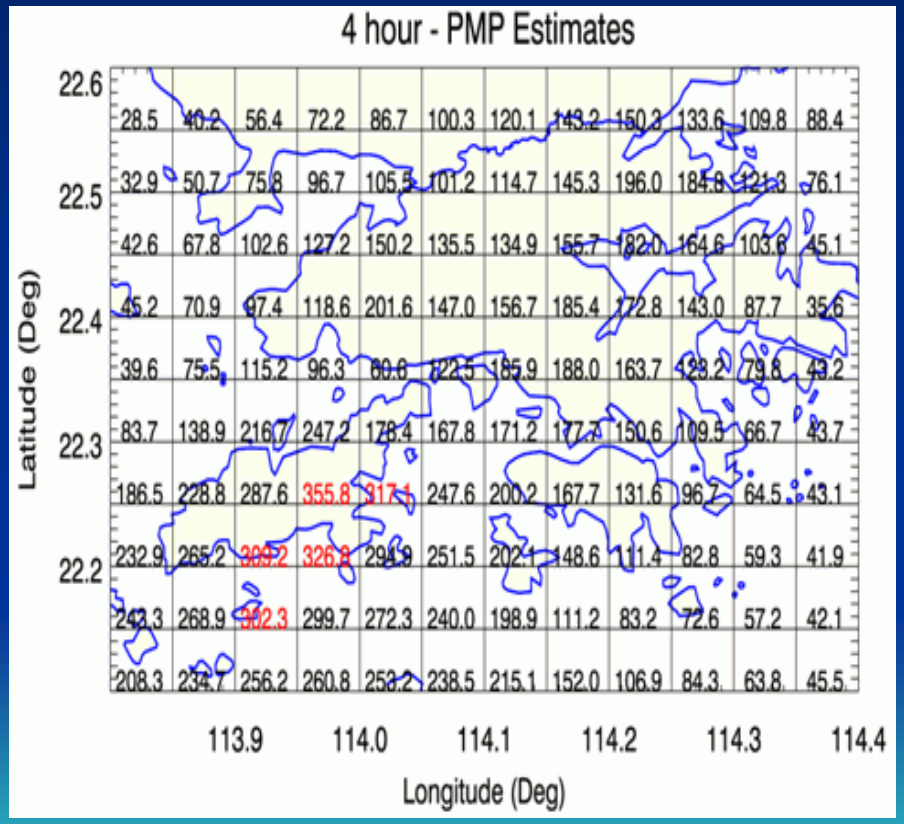


(Peak at Tai Mo Shan)

Fig. 42 Orientation of transposed convergence pattern (2)

Calculation of Embryonic PMP

- Distribution of Gridded-average Embryonic PMP in HK



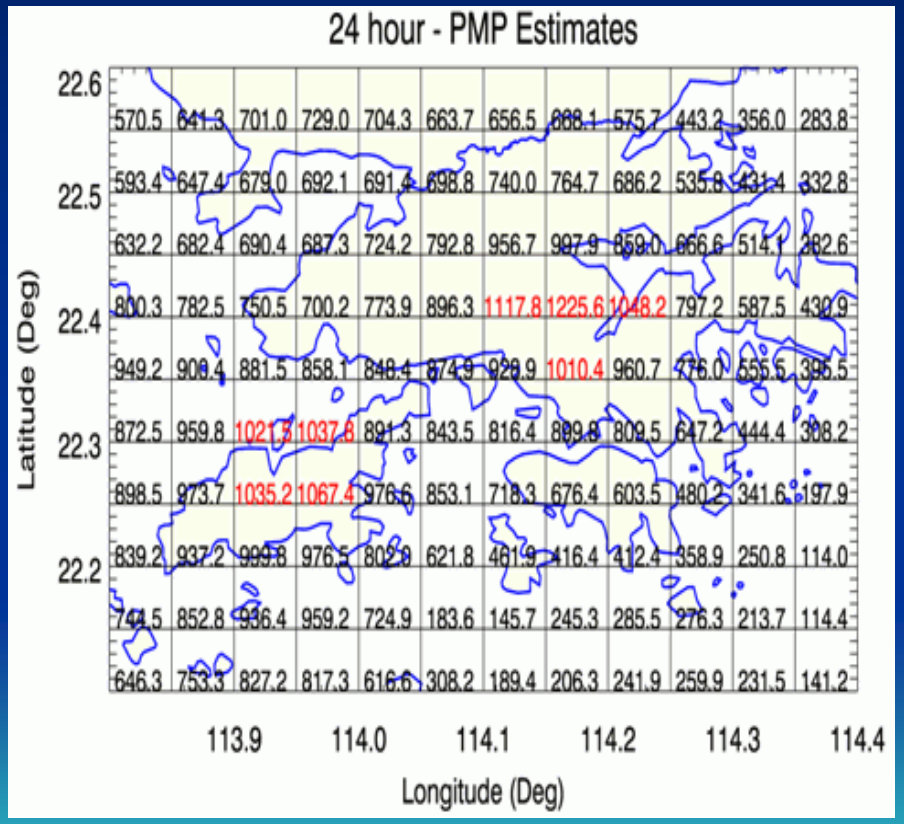
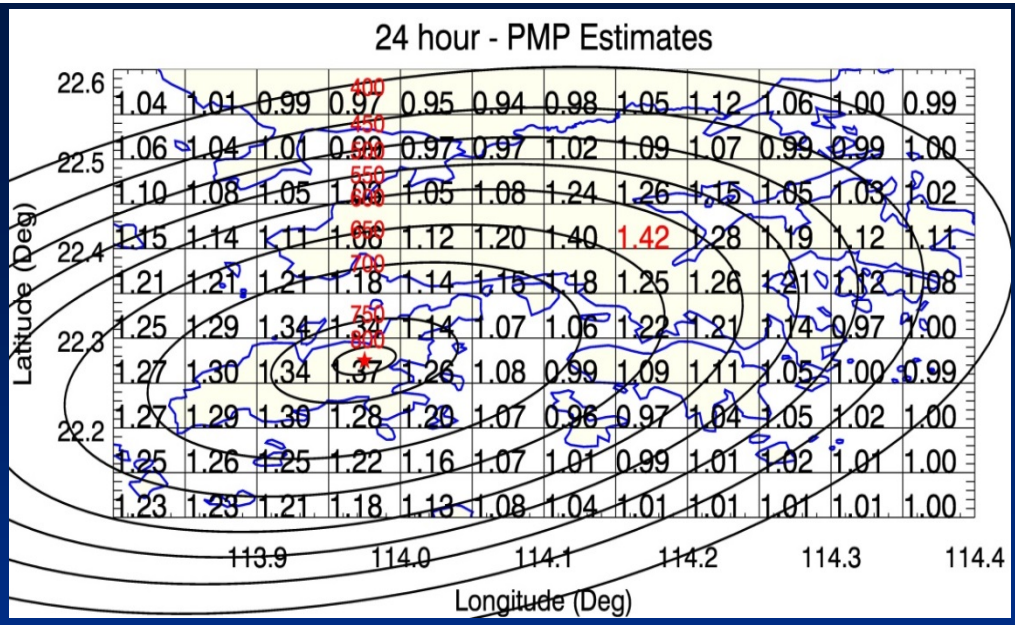


Fig. 43 Gridded embryonic PMP for HK (unfinished; example only)

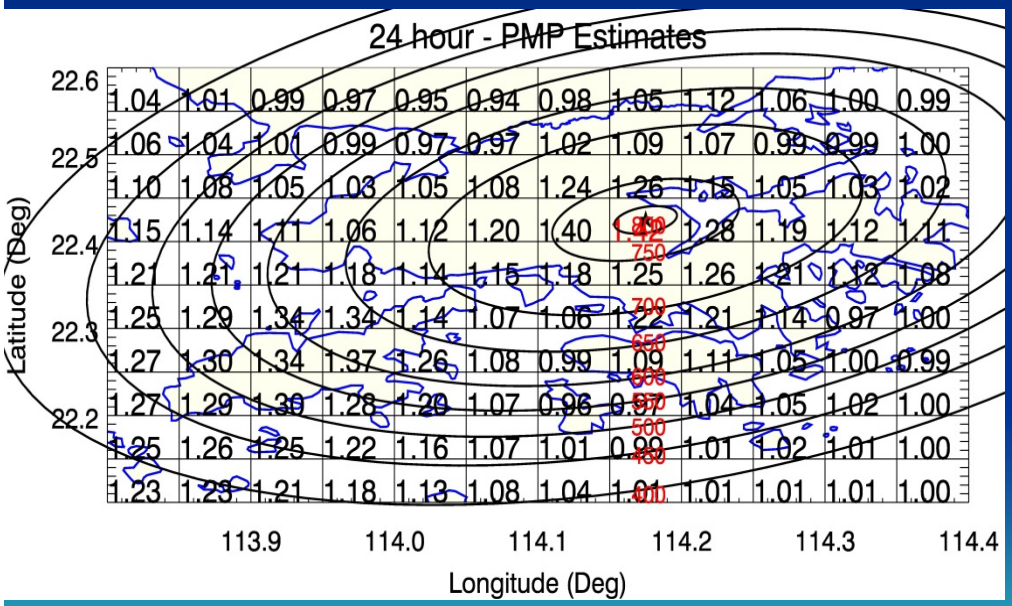
More orientations

Orientation 22.5° – 24hr

(Embryonic PMP)



Centered at Lantau



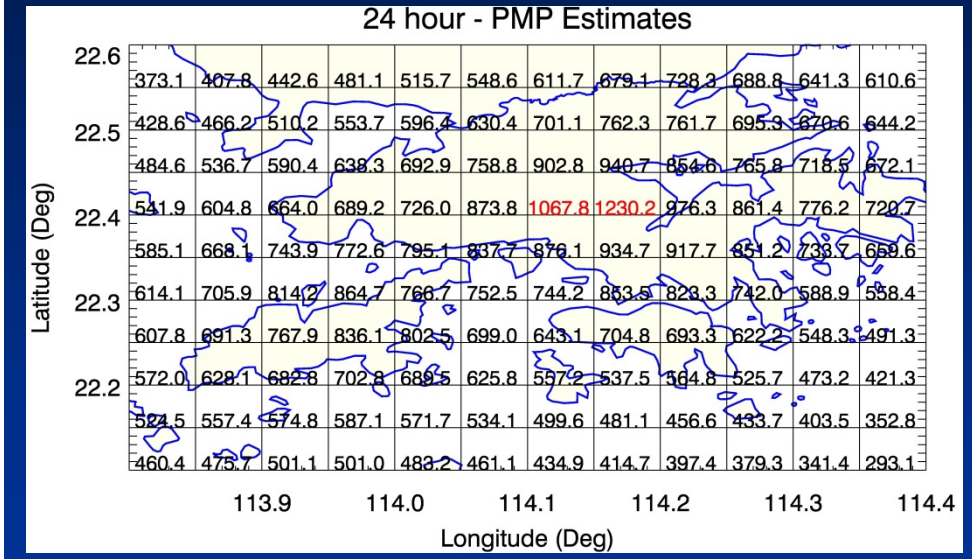
Centered at Tai Mo Shan

Fig. 44 Orientation of transposed convergence pattern (3)

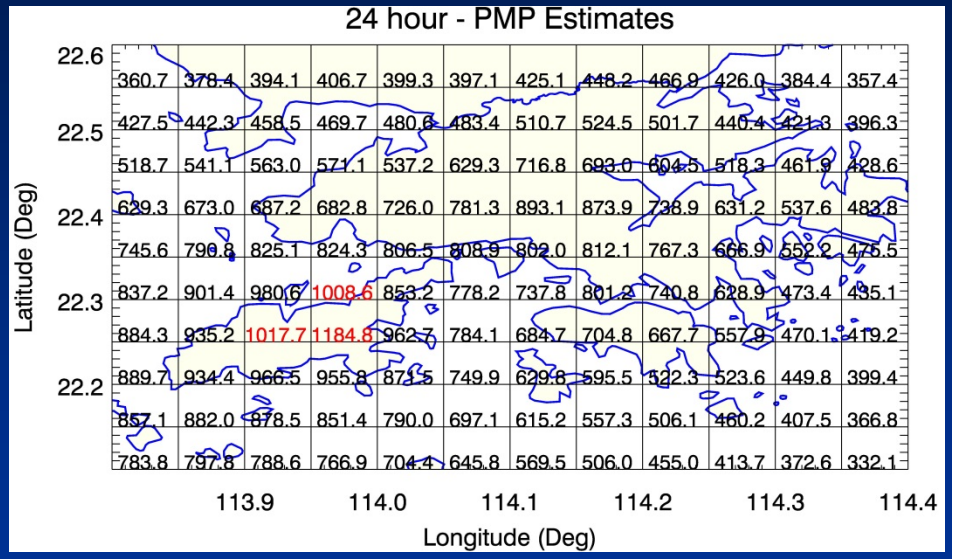
NE-SW orientation – 22.5°

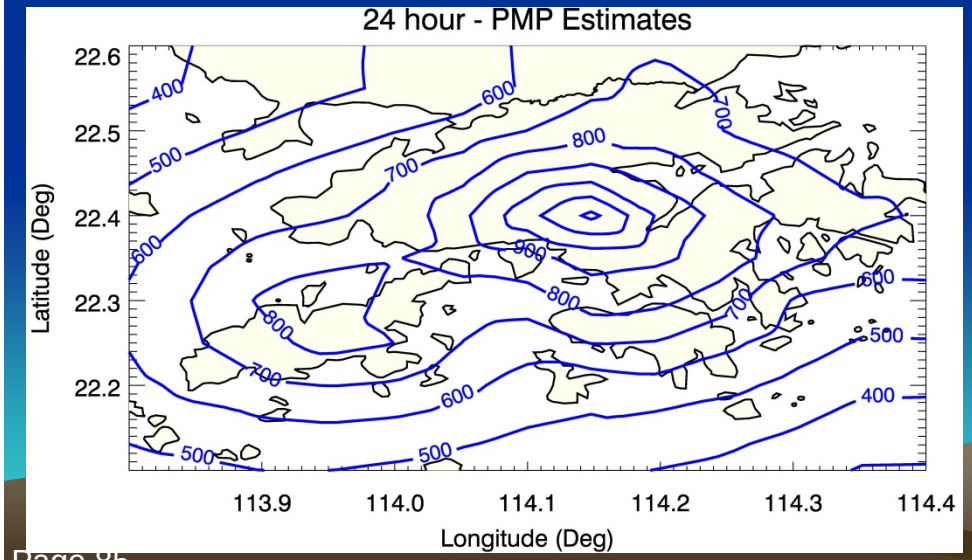
(Embryonic PMP)

Centered at Tai Mo Shan



Centered at Lantau





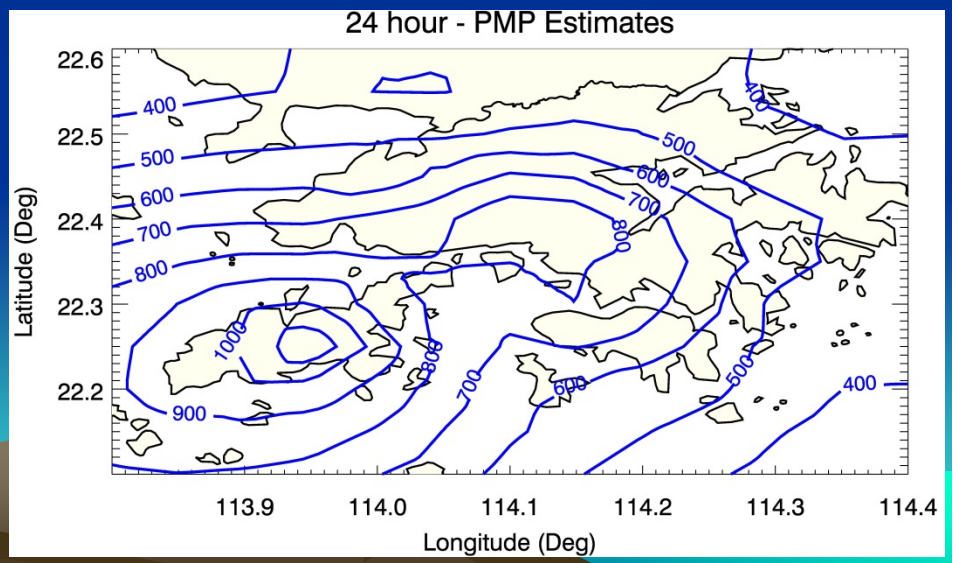
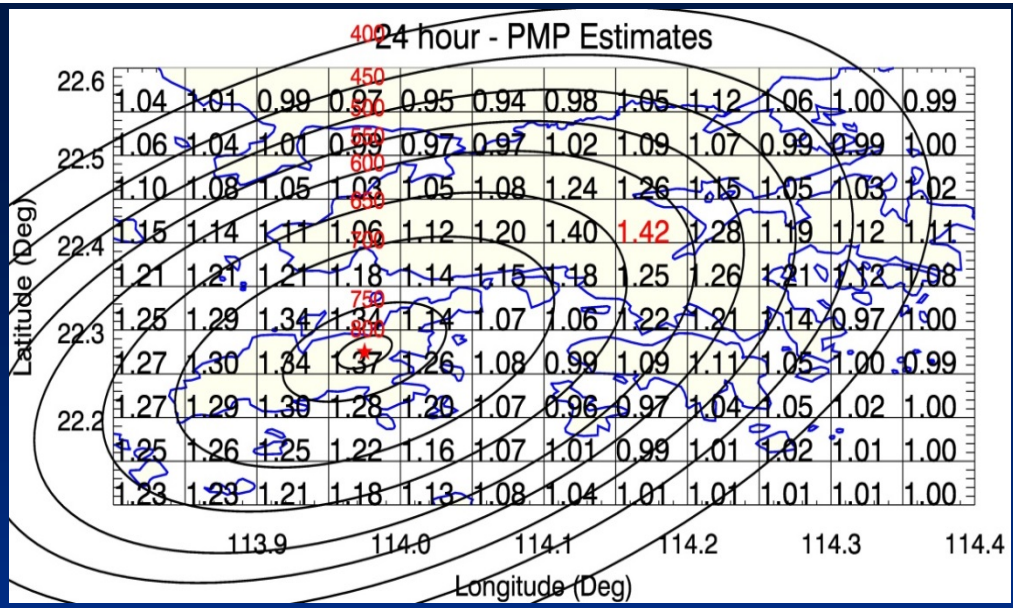


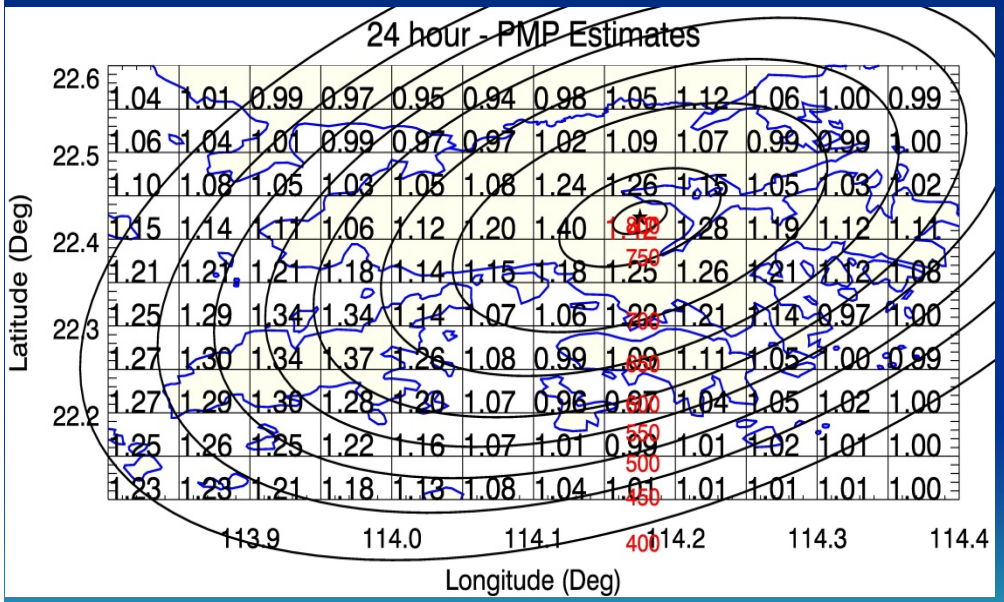
Fig. 45 Embryonic PMP for HK-1 (example only)

Orientation 45° – 24hr

(Embryonic PMP)



Centered at Lantau



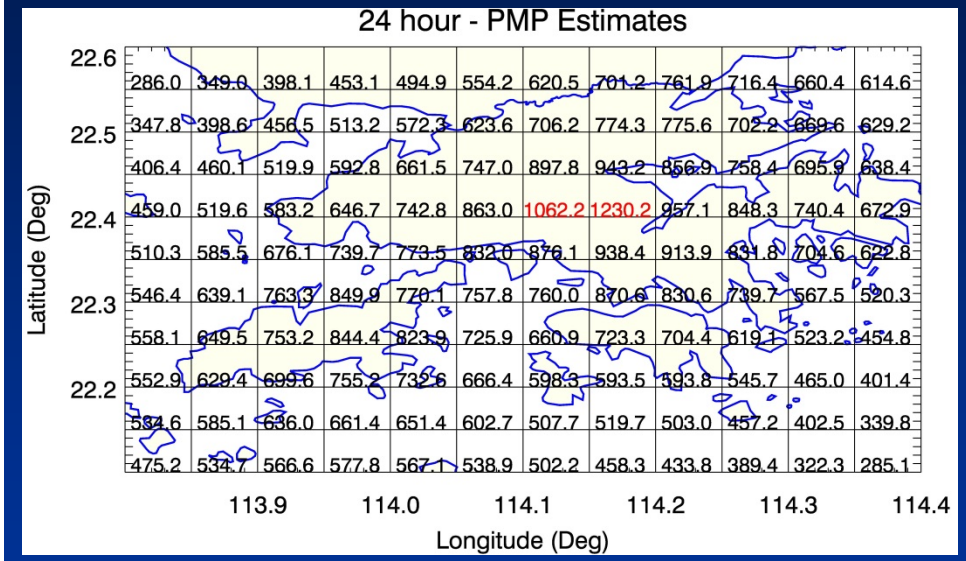
Centered at Tai Mo Shan

Fig. 46 Orientation of transposed convergence pattern (4)

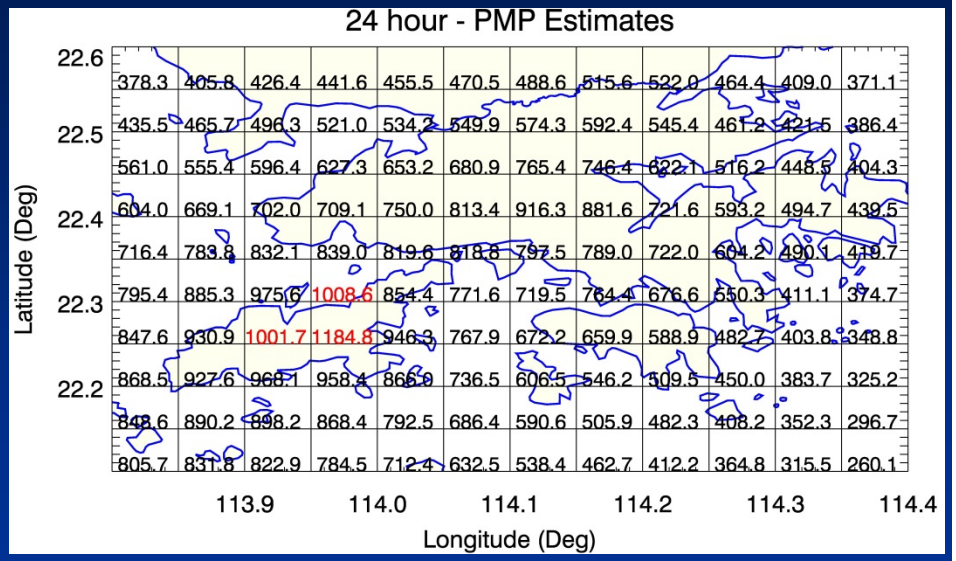
NE-SW orientation – 45º

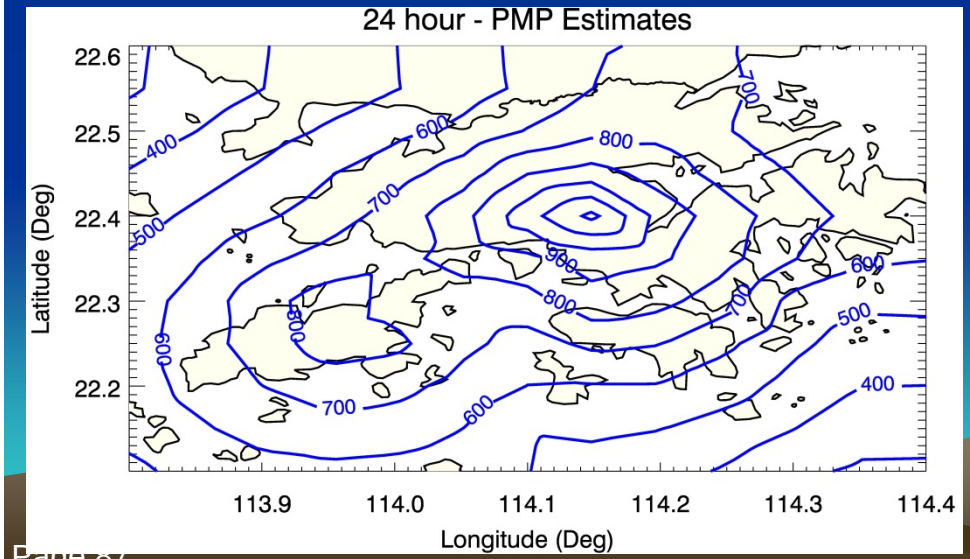
(Embryonic PMP)

Centered at Tai Mo Shan



Centered at Lantau





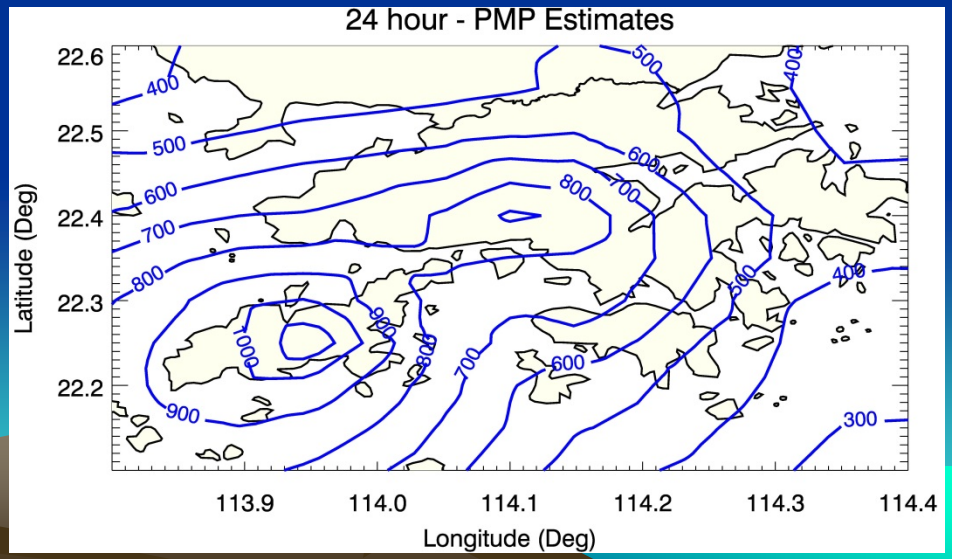


Fig. 47 Embryonic PMP for HK-2 (example only)

Depth-Area-Duration Curves （1）

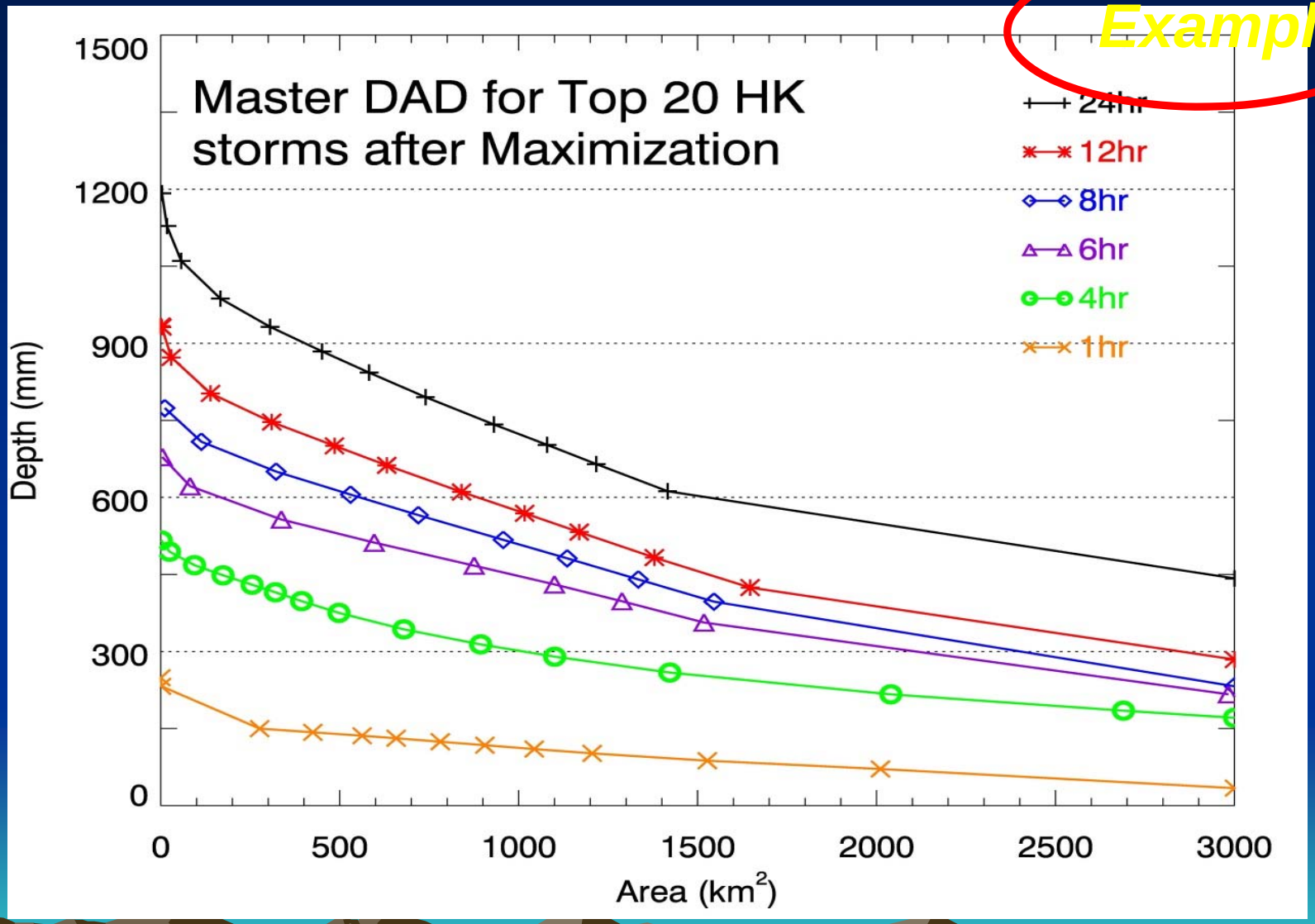


Fig. 48 DAD curve-1 (example only)

Depth-Area-Duration Curves （2）

Example

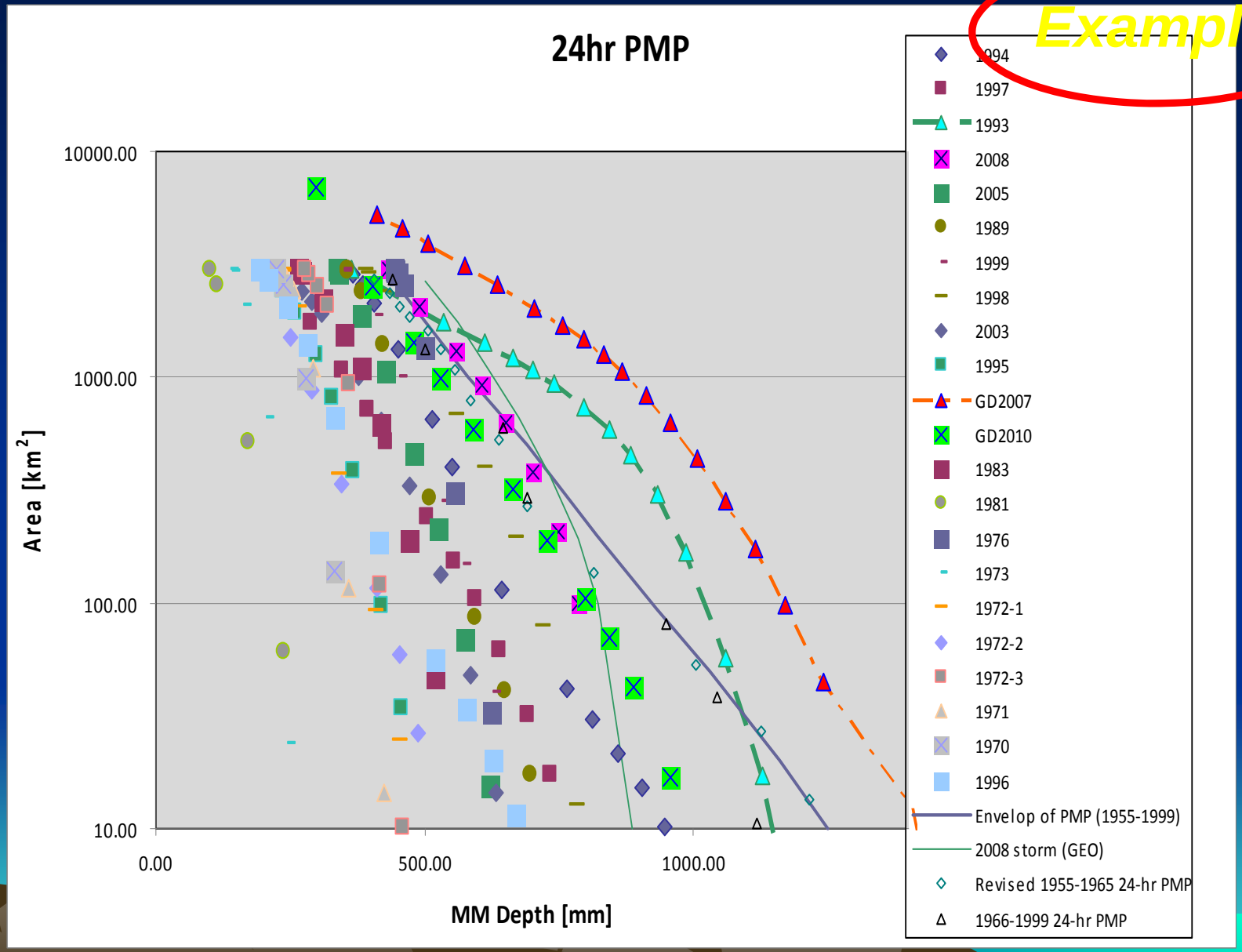


Fig. 49 DAD curve-2 (example only)

Moisture Maximization

Ratio of Moisture Maximization:

Ratio of moisture maximization for transposition

$r = W_m / W_r = W_{27.4} / W_{24} = 99.6 / 74.0 = 1.346$

in which, W_m is the historical maximum consistent dew point which is 27.4°C *in design area* while the W_r is the representative dew point, 24.0°C, for the Morakot storm *in target area*.

Statistical Approach (1)

PMP, X_{PMP}

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

where $\overline{X}_n$ the mean of the n maxima

S_n the standard deviation of the n maxima

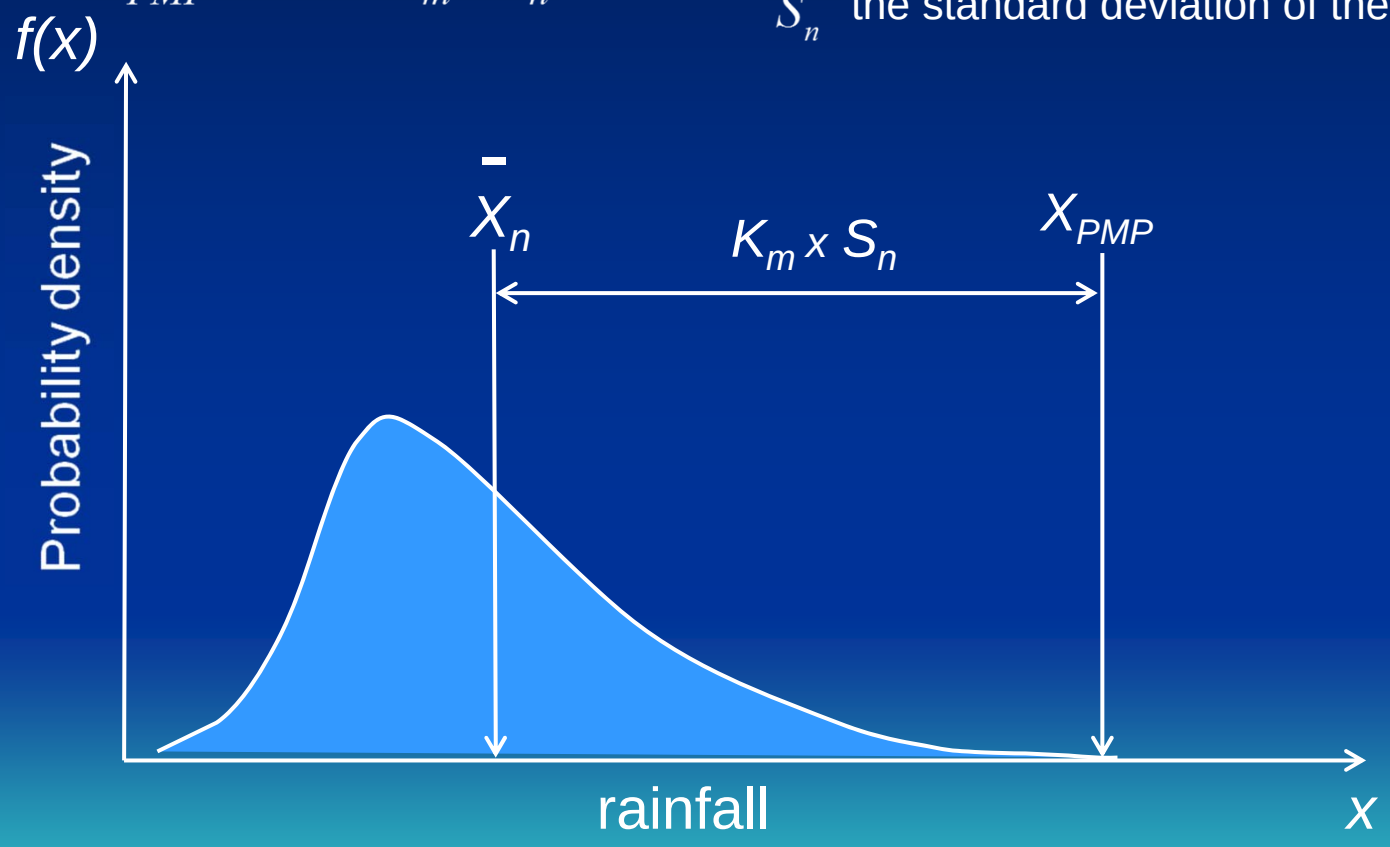


Fig. 50 Sketch for statistical approach to PMP (1)

Statistical Approach (2)

"Modified" Frequency Analysis

- "standard deviation", K_m, is added to the mean in the frequency equation

$$X_{PMP} = \overline{X}_n + K_m \times S_n = (1 + K_m \times C_{vn}) \times \overline{X}_n \longrightarrow \boxed{X_{pmp} = \overline{X}_n + K_m \times S_n}$$

where $\overline{X}_n$ 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 (X_m) from the historical series <u>will be omitted</u> in the computation

$$K_m = \frac{(X_m - \overline{X}_{n-1})}{S_{n-1}} \longrightarrow \boxed{X_m = \overline{X}_{n-1} + K_m \times S_{n-1}}$$

Tricky

where $\overline{X}_{n-1}$ and S_{n-1} are the mean and the standard deviation of the rainfall series **<u>from which the maximum record rainfall was omitted.</u>**

Statistical Approach (3) – New development

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 ϕ_m is the maximum deviation from mean and is directly computed from the following equation,

$$\phi_m = \frac{(X_m - \overline{X}_n)}{S_n}$$

- The criterion of the **stable size** of N_s in terms of 10% relative error (*Lin, 1981; Lin & Vogel, 1993*)

$$N_s \geq 5.76 \times (\phi_m^2 + 2)$$

Statistical Approach (4)

The Table of K_m (Lin & Vogel, 1993)

Appendix 1 List of variations of K_m with variation of n under different values of T_m

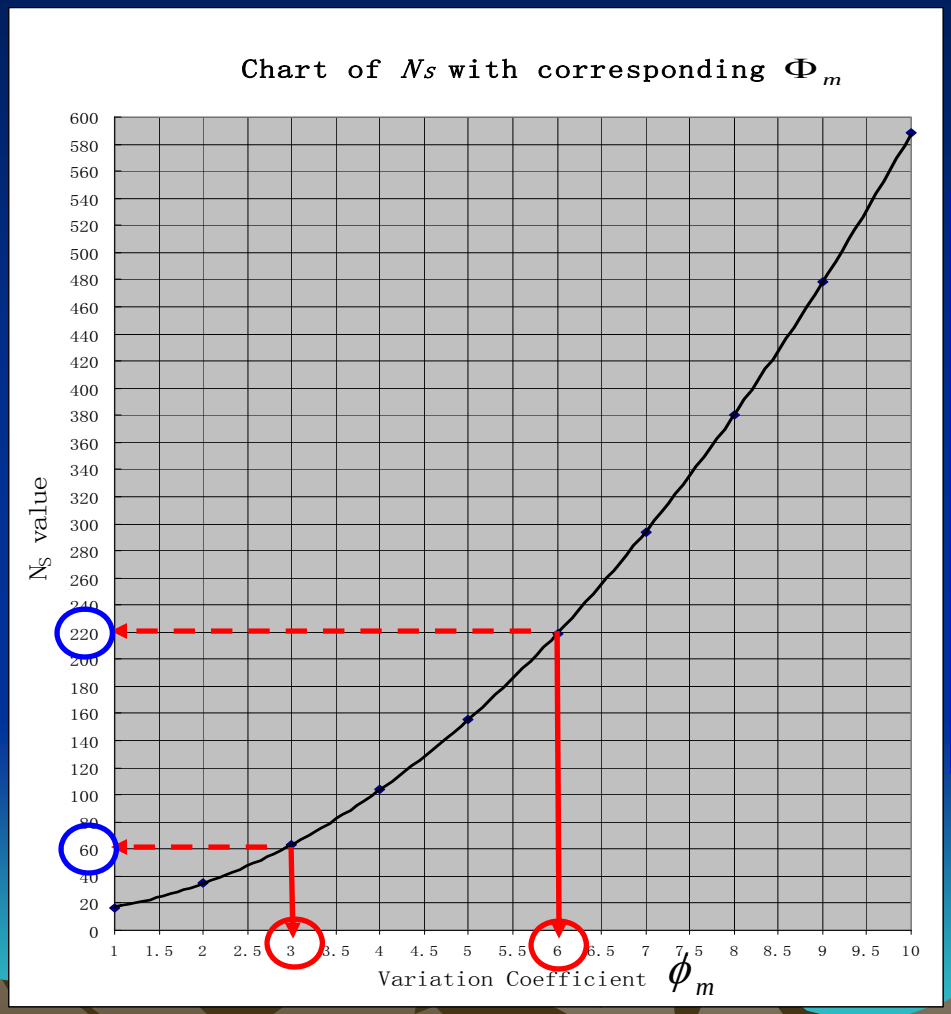
n	Tm 5	6	7	8	9	10	15
27	132.38						
38		224.90					
50	7.29						
51			353.41				
66				523.92			
83					742.43		
100		7.58	9.95	13.65			
102						1014.94	
150	**5.50**						
200		6.65			11.74	14.25	
219		**6.60**					
227							3397.46
294			**7.70**				
300			7.67			12.29	
380				**8.80**			
400				8.74			
478					**9.90**		
500	5.14	6.24	7.38	8.58	9.84		20.28
588						**11.00**	
600						10.97	
1000	5.07	6.11	7.18	8.27	9.39	10.55	17.05
1308							**16.50**
1500							16.28
5000	5.01	6.02	7.05	8.05	9.08	10.10	15.35
10000							15.17
Nm	27	38	51	66	83	102	227
Ns	**150**	**219**	**294**	**380**	**478**	**588**	**1308**

Notes: 1. N_m 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.

10% in error

Statistical Approach (5)

- Relationship of Variation Coefficient ϕ_m with N_s



Take some stations out when N_s > 3.5 n 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)

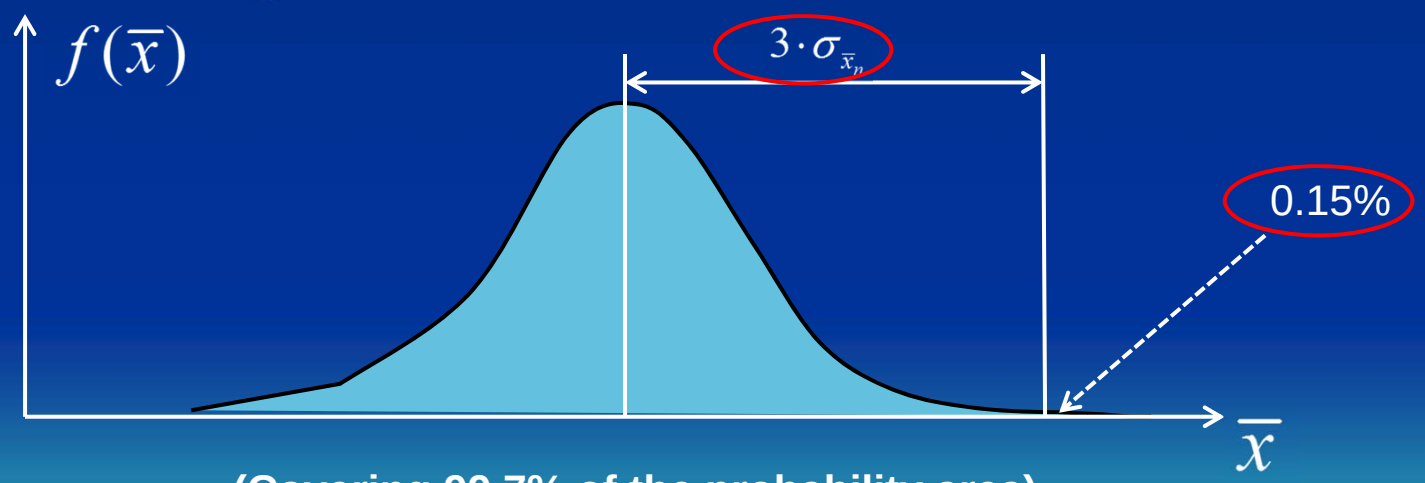
Statistical Approach (6)

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

PMP, X_{PMP}

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

where $\sigma_{\overline{x}_n}$ 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!

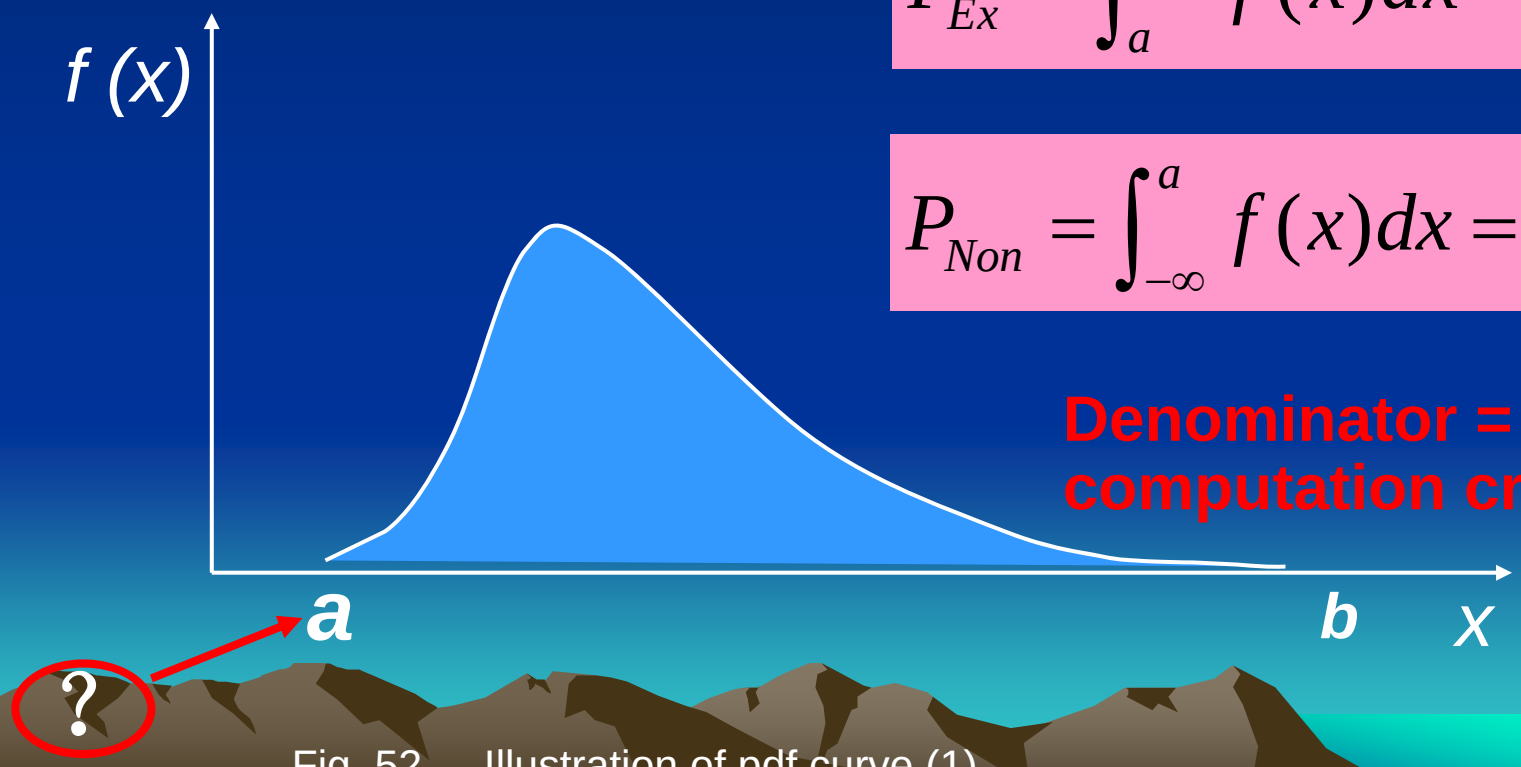


Fig. 52 Illustration of pdf curve (1)

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.

(* Ven Te Chow, *Applied Hydrology,* 1988 edition)

Exceedance frequencies of data (1)

- Location of the Semiarid study area

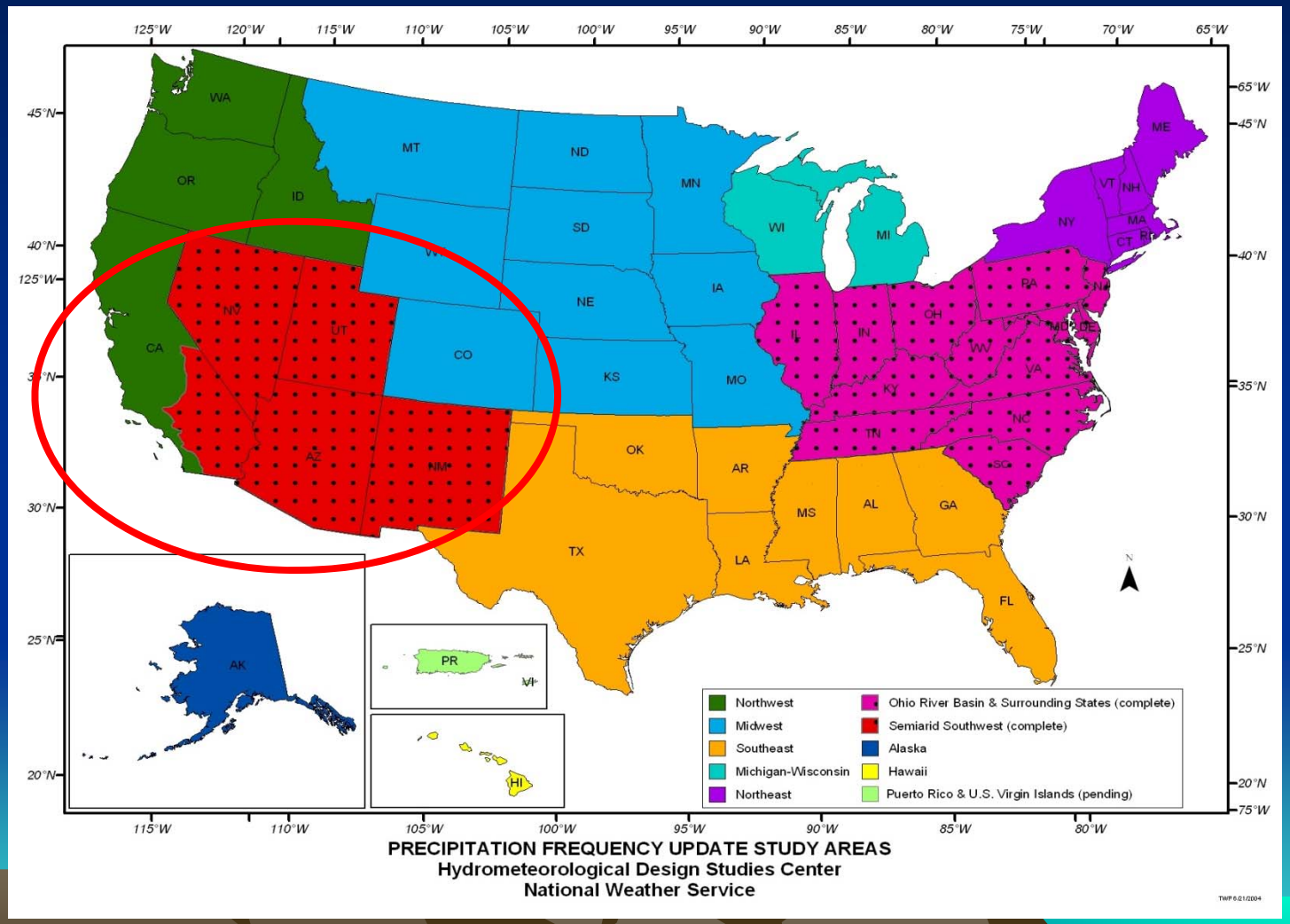


Fig. 53 SA data -- used for test of sampling methods

Exceedance frequencies of data (2)

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

Station ID	Data years	Return Period (R.P.) / Exceedance Probability (E.P.)					
		2-y	5-y	10-y	25-y	50-y	100-y
		0.5	0.2	0.1	0.04	0.02	0.01
04-0029	56	0.76	0.18	0.07	0.04	0.04	0.02
04-1476	53	0.64	0.26	0.13	0.06	0.04	0.04
04-1805	38	0.76	0.22	0.13	0.03	0.03	0
04-2964	48	0.66	0.21	0.04	0.04	0.04	0.02
04-4838	40	0.67	0.26	0.13	0.10	0.05	0.03
04-9053	69	0.68	0.25	0.09	0.03	0.01	0
35-2018	41	0.60	0.24	0.19	0.05	0.02	0.02
35-3232	30	0.56	0.16	0.10	0.07	0.03	0
35-5174	32	0.75	0.22	0.09	0.06	0.06	0
35-7354	67	0.66	0.20	0.11	0.05	0.03	0.02
35-8007	47	0.82	0.25	0.07	0	0	0
Average E. P.		***0.687***	***0.223***	***0.104***	***0.048***	***0.032***	***0.014***
Corresponding R.P.		***1.45-y***	***4.49-y***	***9.57-y***	***20.75-y***	***31.45-y***	***73.53-y***

Exceedance frequencies of data（3）

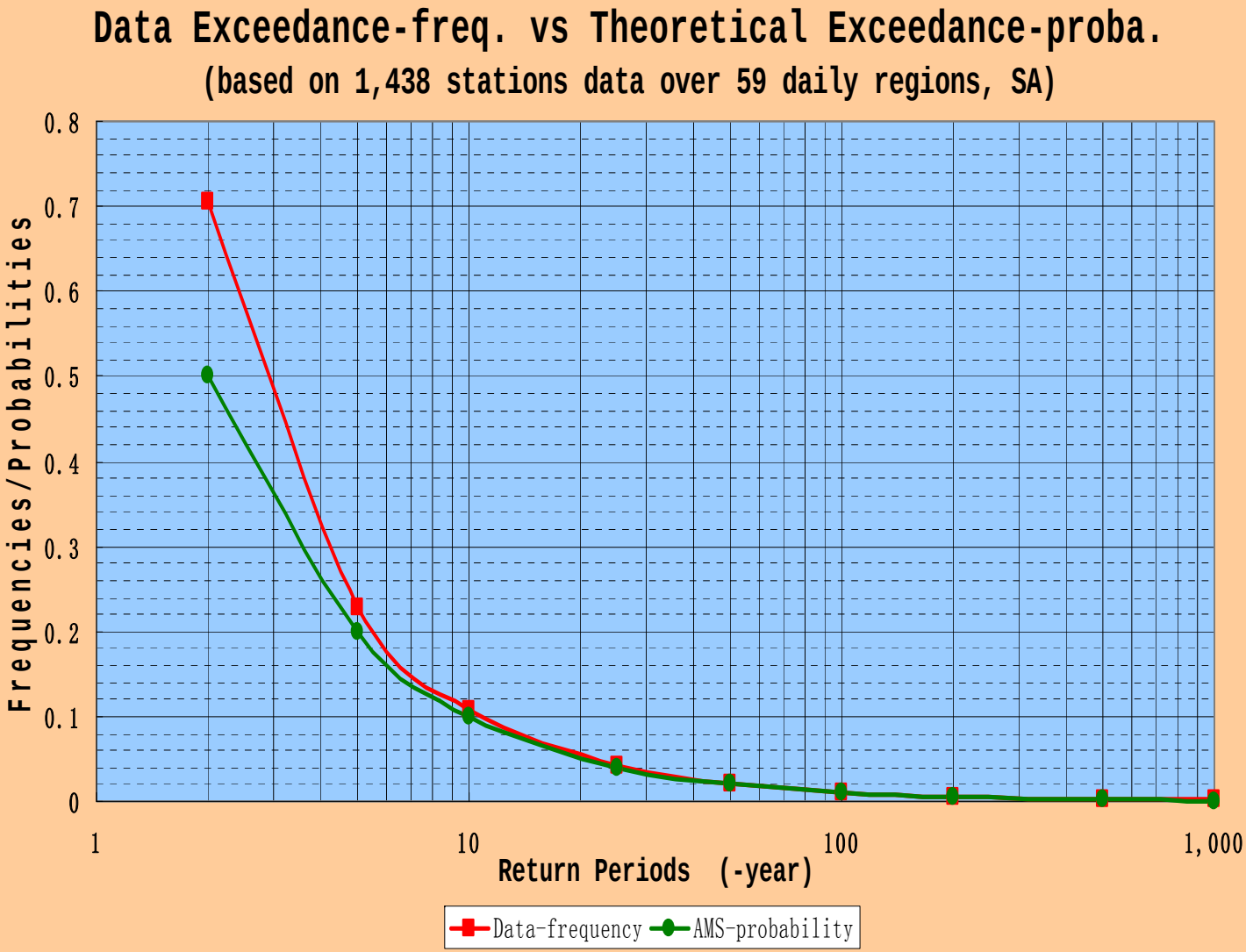


Fig. 54 Illustration for comparison of frequency to probability

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:

$$T_{AES} = [\ln(\frac{T_{AMS}}{T_{AMS} - 1})]^{-1}$$

(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

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

* 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

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

*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.

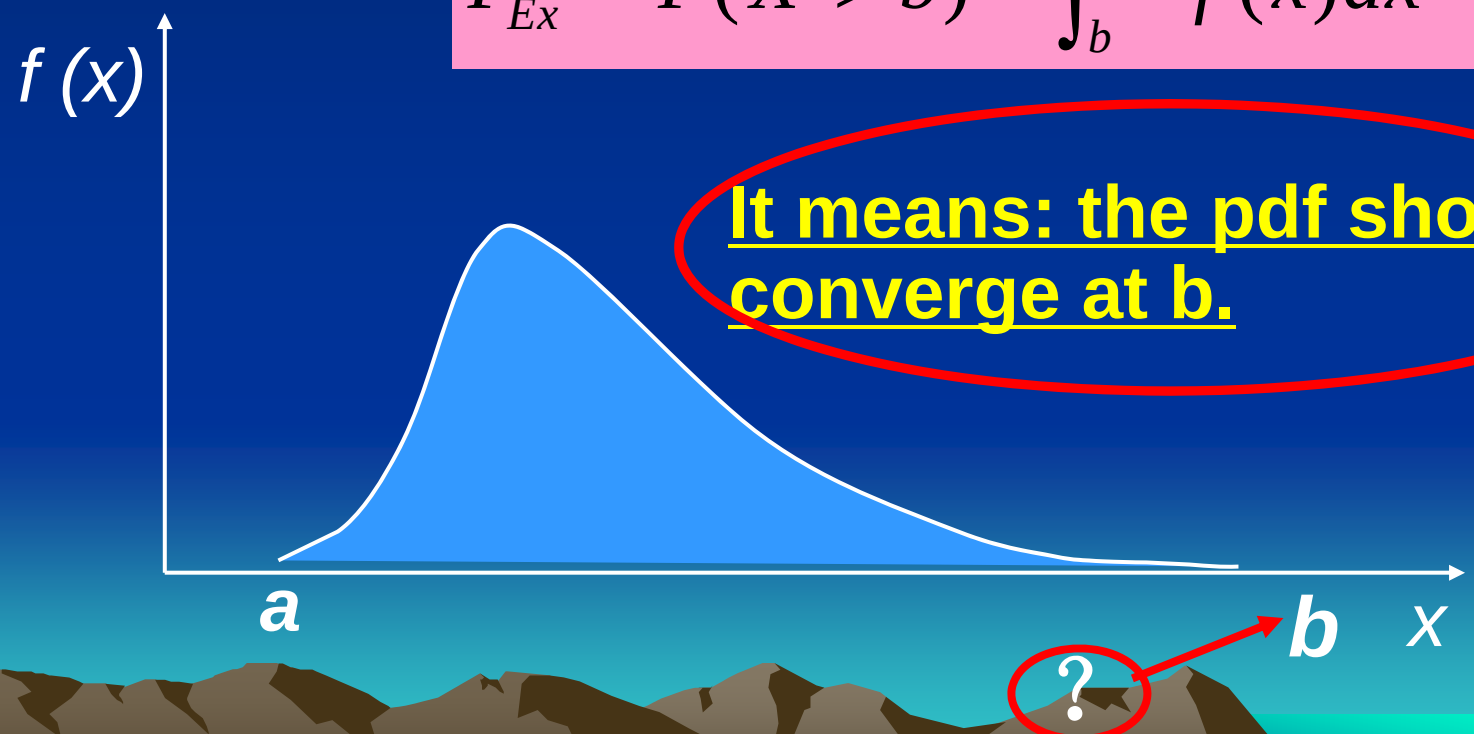


Fig. 55 Illustration of pdf curve (2)

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

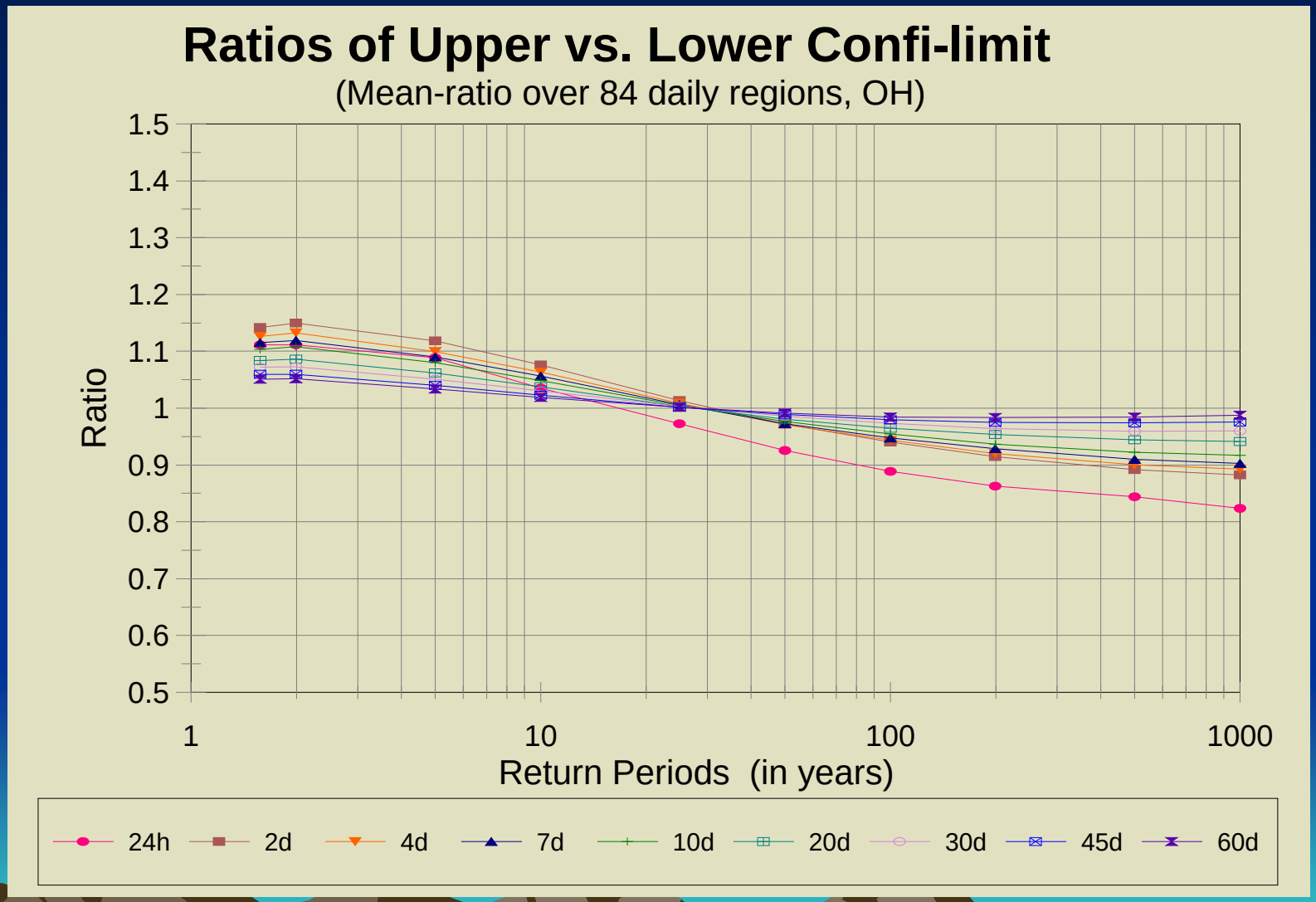


Fig. 56 Rations of (upper vs lower) for Ohio River Basin

Findings (Results) over 59 Regions, SA

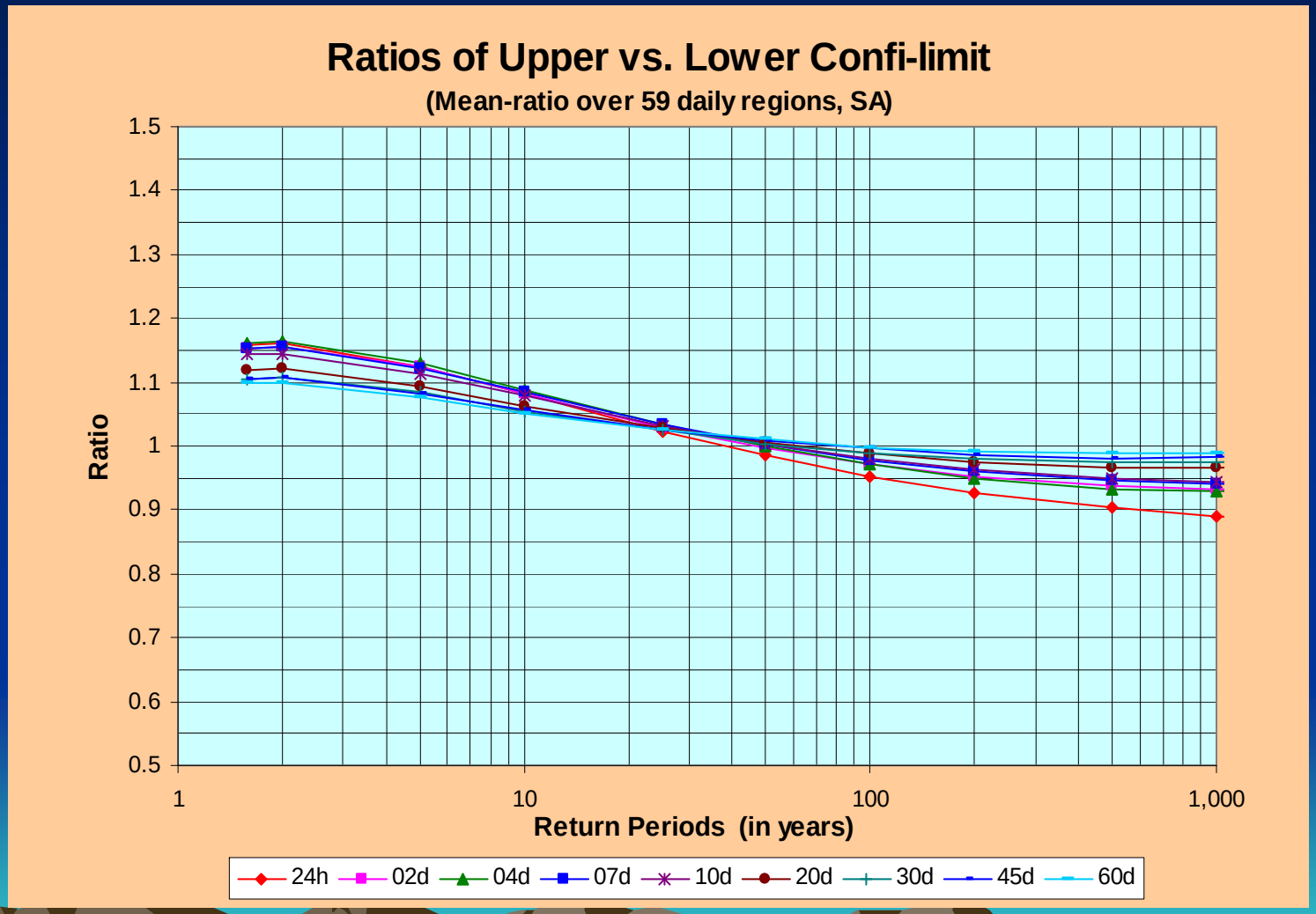
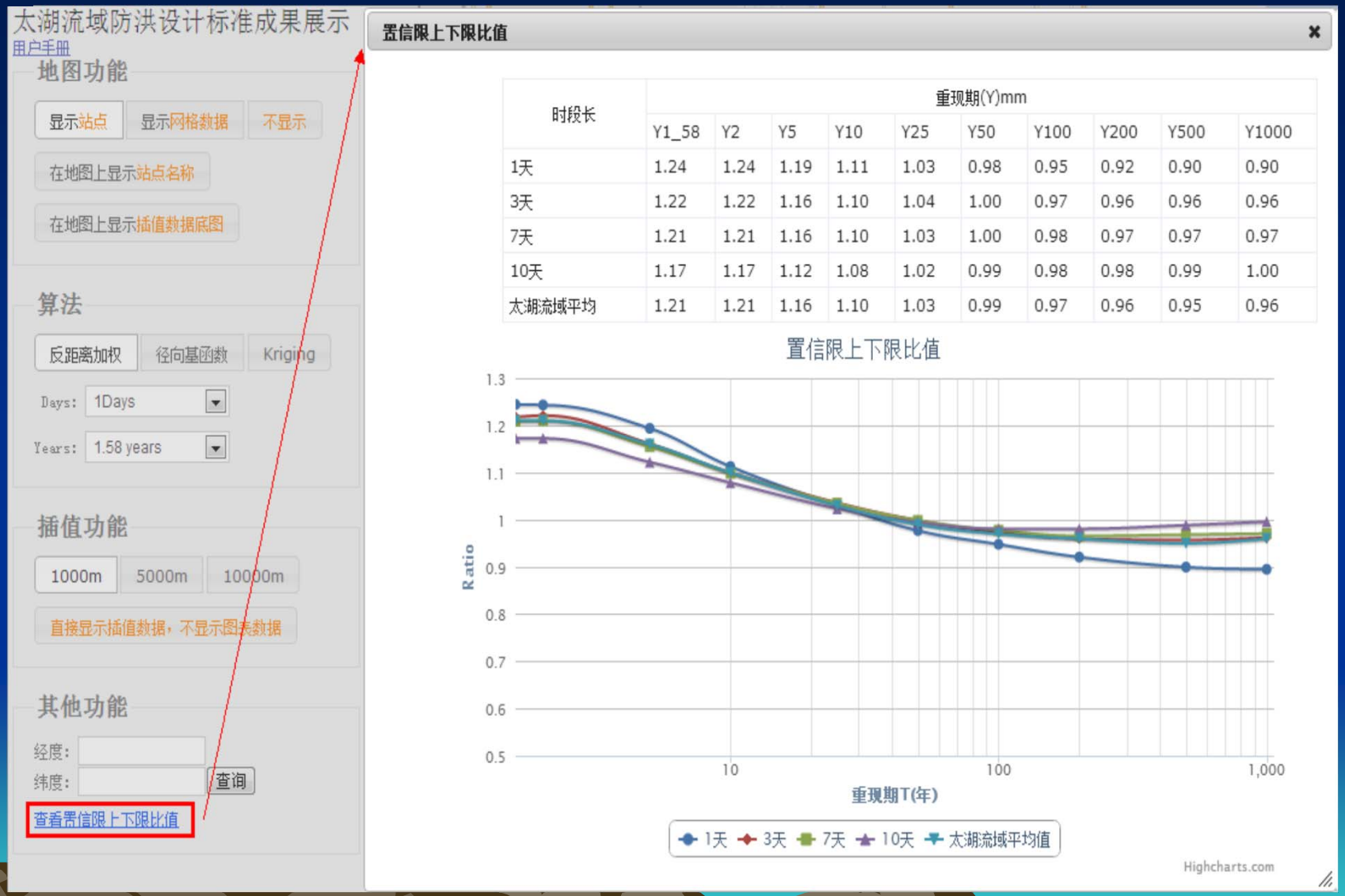


Fig. 57 Rations of (upper vs lower) for SW Semiarid U.S.

Findings (Results) over 8 Regions, Taihu



时段长	重现期(Y)mm									
	Y1_58	Y2	Y5	Y10	Y25	Y50	Y100	Y200	Y500	Y1000
1天	1.24	1.24	1.19	1.11	1.03	0.98	0.95	0.92	0.90	0.90
3天	1.22	1.22	1.16	1.10	1.04	1.00	0.97	0.96	0.96	0.96
7天	1.21	1.21	1.16	1.10	1.03	1.00	0.98	0.97	0.97	0.97
10天	1.17	1.17	1.12	1.08	1.02	0.99	0.98	0.98	0.99	1.00
太湖流域平均	1.21	1.21	1.16	1.10	1.03	0.99	0.97	0.96	0.95	0.96

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***

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

$$\int_{?}^{?} f(x)dx = 1$$

$$\int_{a}^{b} f(x)dx = 1 \longrightarrow \textbf{OK}$$

Uncertainties of quantiles

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!*

南京信息工程大学应用水文气象研究所

APPLIED HYDROMETEOROLOGICAL RESEARCH INSTITUTE

AHMRI

南京信息工程大学

All rights reserved.

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

Welcome!

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*

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