香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AS1.16-D3061

Classification and Diagnosis of Summer Monsoon Rainfall Patterns and their Potential Predictability in Southeast China

Dai Lun & Cheng Tat Fan

Supervisor: Prof. Lu Mengqian

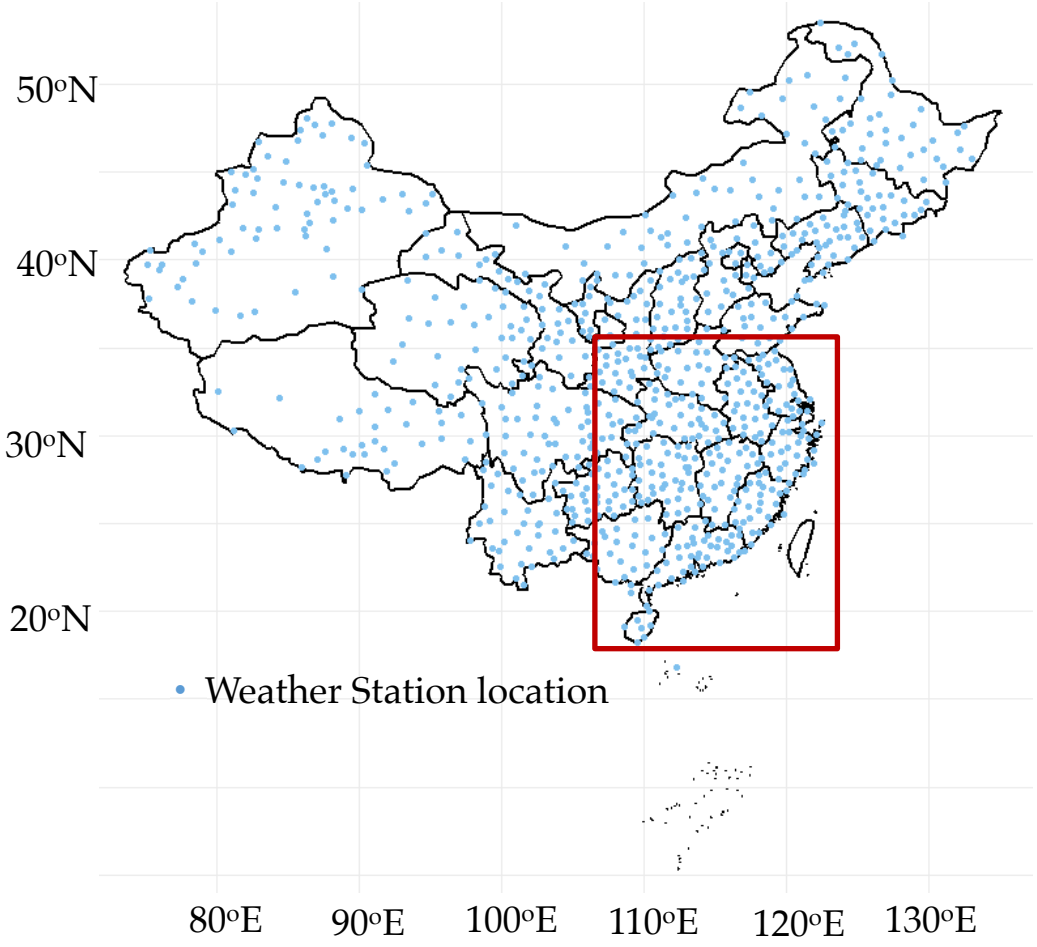
Dai, L., Cheng, T. F., & Lu, M. (2020). Summer Monsoon Rainfall Patterns and Predictability over Southeast China. Water Resources Research, 56(2). https://doi.org/10.1029/2019WR025515

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

- **Questions**
  - *i. Rainfall patterns & their variability ?*
  - *ii. Atmospheric dynamics and teleconnected oceanic signals?*
  - *iii. Any predictability?*
  - *iv. Informative dynamic divers?*
- **Methods**
  - *i. Self- organizing Map (SOM)*
  - *ii. Composite Analysis*
  - *iii. Multiple Linear Regression (MLR)*
  - *iv. Empirical orthogonal Functions (EOF)*
- **Data**
  - *i. Rainfall gauge data from China Meteorological Data Service Center*
  - *ii. Environmental data from ERA-interim*
- **Results**
  - *i. Spatial Rainfall Patterns*
  - *ii. Temporal variabilities & spatial grouping*
  - *iii. Diagnosis of Atmospheric dynamics and teleconnected oceanic signals*
  - *iv. Model building & Validation & Blind prediction*
  - *v. Interpretation of informative predictors*

Conclusions and Some Highlights

Data —— *Rainfall data & environmental variables*



Study period: 1979-2018 (May- August)

Rainfall:

1. 864 stations over mainland China.
2. 272 stations in SEC (108°E-125°E, 18°N-36°N)
3. High quality dataset with less than 1% missing value and low-quality value.

Environmental variables:

1. Reanalysis Datasets from ERA-Interim (2.5°× 2.5°)
2. Including:
   a. Geopotential Height at 200 hPa (GPH 200)
   b. Geopotential Height at 850 hPa (GPH 850)
   c. Sea Surface Temperature (SST)
   d. 2-meter Temperature (T2m)
   e. Vertically Integrated water vapor transport(IVT)
   f. Total precipitation (PP)

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

Results —— *Spatial Rainfall Patterns*

Input Matrix —*SOM*→ Output —*Average within clusters*→ Rainfall Patterns

Station 1 Station 272

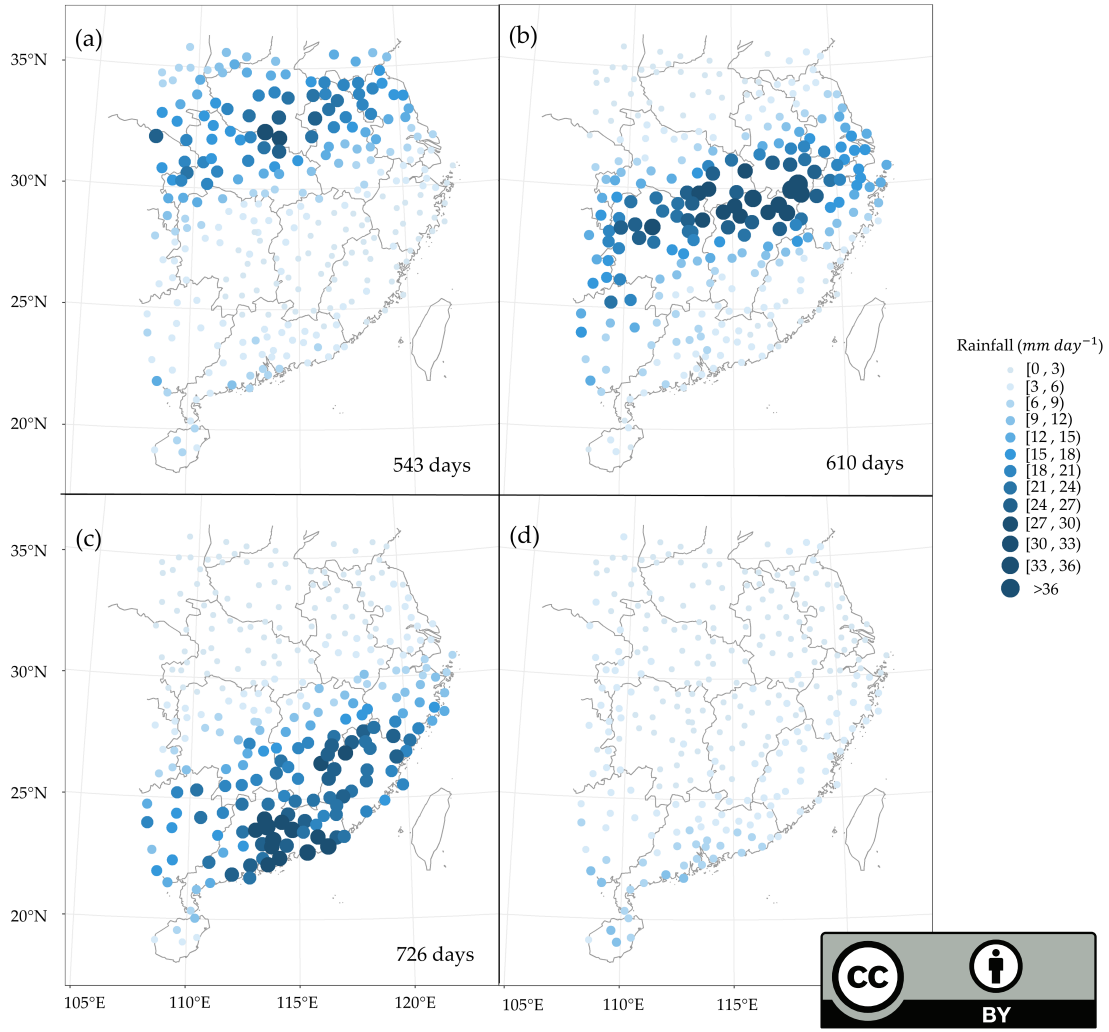
$$\begin{bmatrix} x_{1,1} & \cdots & x_{1,272} \\ \vdots & & \vdots \\ x_{123,1} & \ddots & x_{123,272} \\ \vdots & & \vdots \\ x_{4675,1} & & x_{4675,272} \\ \vdots & & \vdots \\ x_{4797,1} & & x_{4797,272} \end{bmatrix}$$

1st Year MJJA

39th Year MJJA

$x_{i,j}$: rainfall intensity in day i and station j

Four Clusters



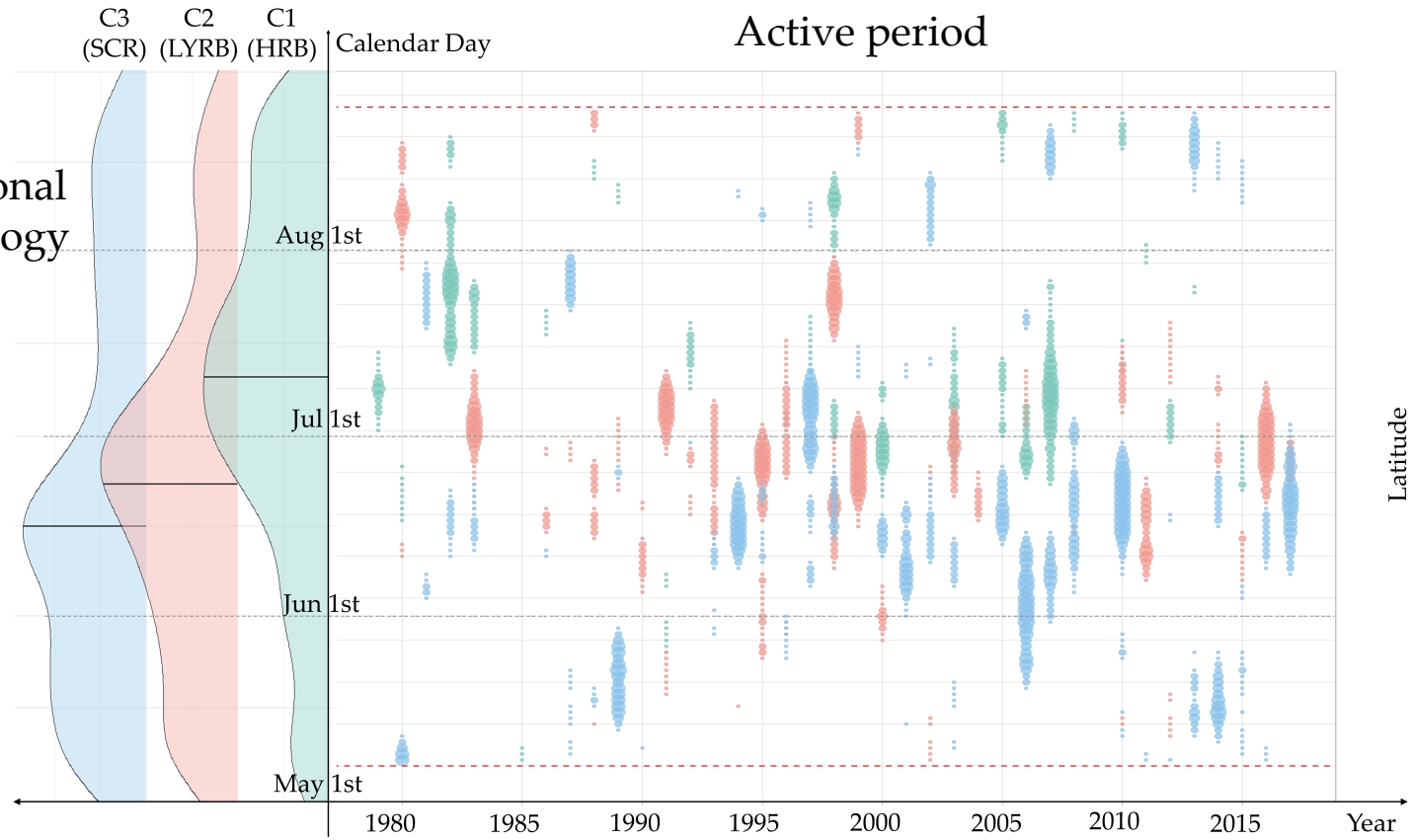
Why choose "four clusters":

- Reasonably capture the rain-belt over SEC
- Significantly distinguishable between each pair of clusters

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

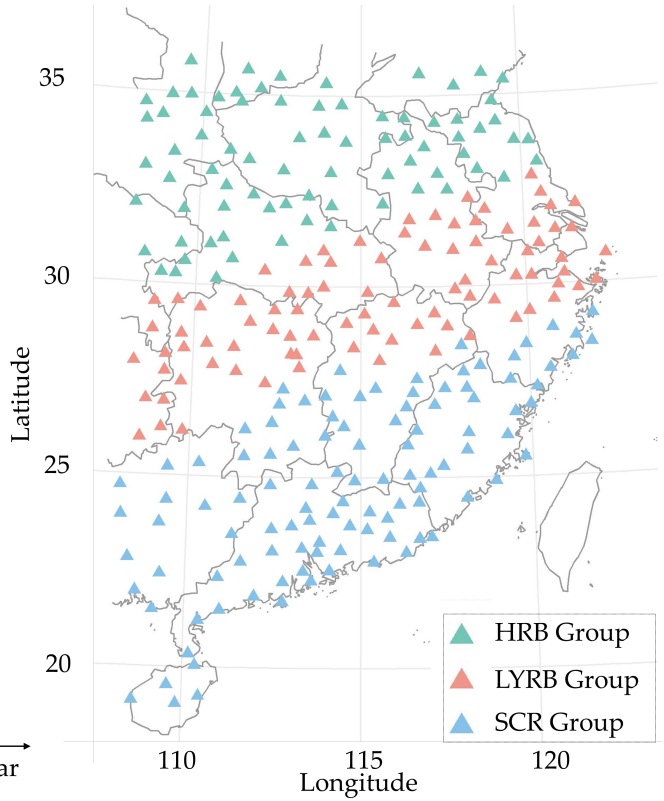
Results —— *Temporal variabilities & spatial grouping*

Subseasonal Climatology

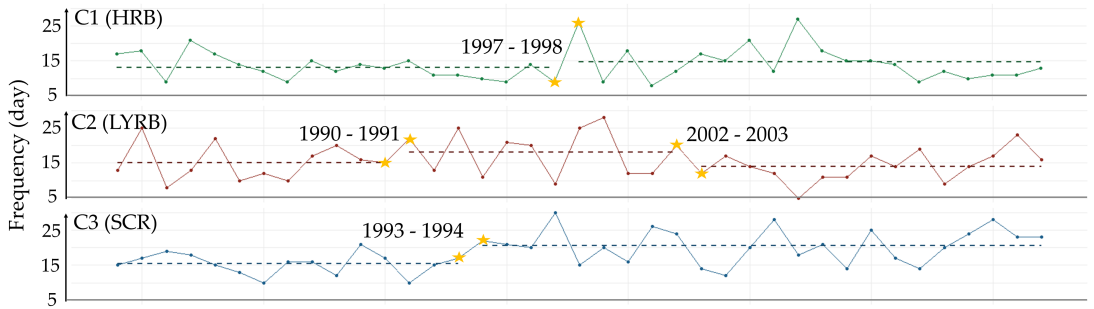


Active period

Spatial Grouping



Annual frequency change points



C1 C2 C3

Change point ★

Occurrence (day) 5 6 7 8 9 >=10

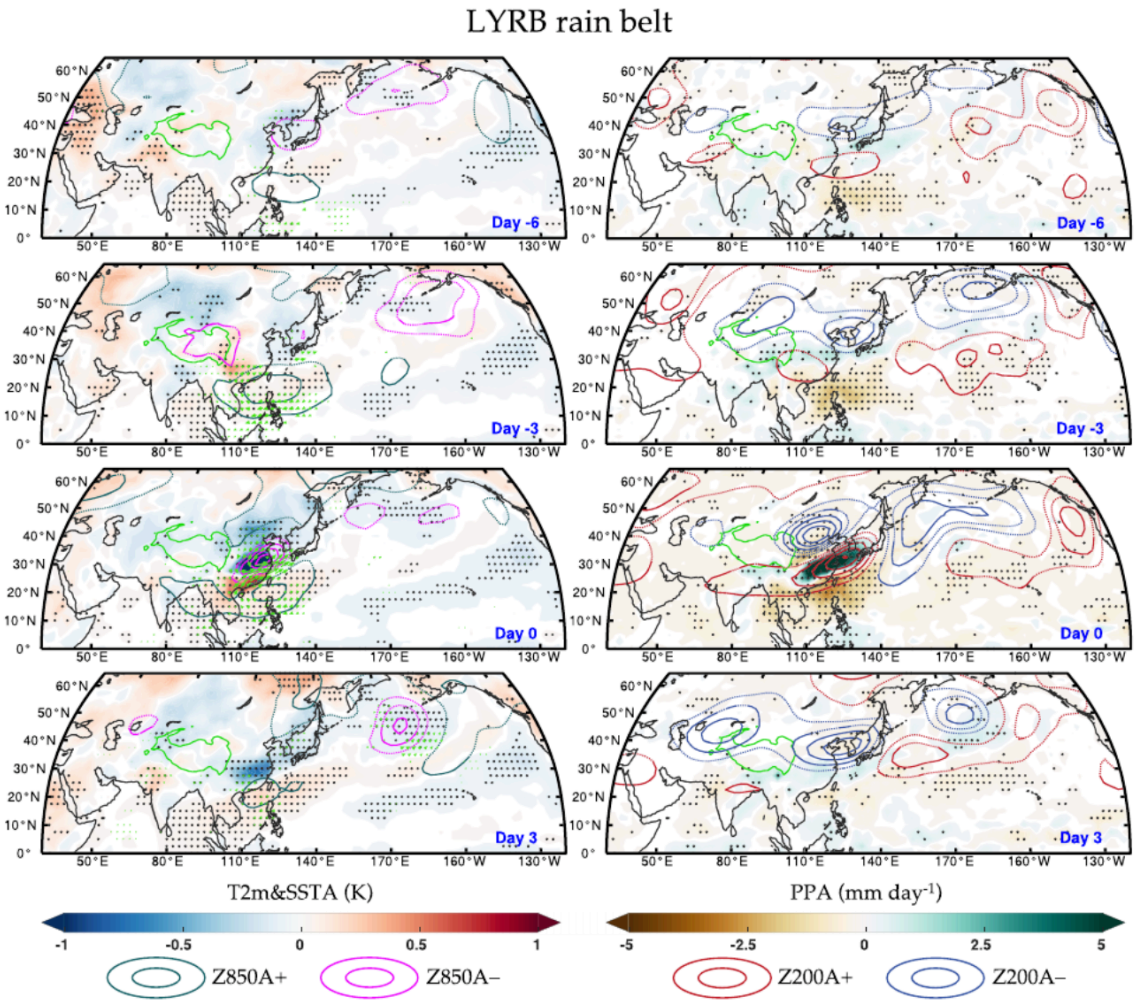
Results ——— *Diagnosis of Atmospheric dynamics and teleconnected oceanic signals*

Composite of Cluster 2

T2mA&SSTA (shaded)

Z850A+ & Z850A-
(contour, interval: 3 m)

IVTA (vector)



PPA (shaded)

Z200A (+) & Z200A (-)
(contour, interval: 5 m)

Results —— *Model building & Validation & Blind prediction (Part I)*

Transformed Daily rainfall

GPHA 850 data

EOF

EOF Spatial patterns

EOF time series

Average within each group

Concurrent & one day lead

Regional Rainfall ⟺ 20 predictors *Full model*

Forward stepwise variable selection based on BIC

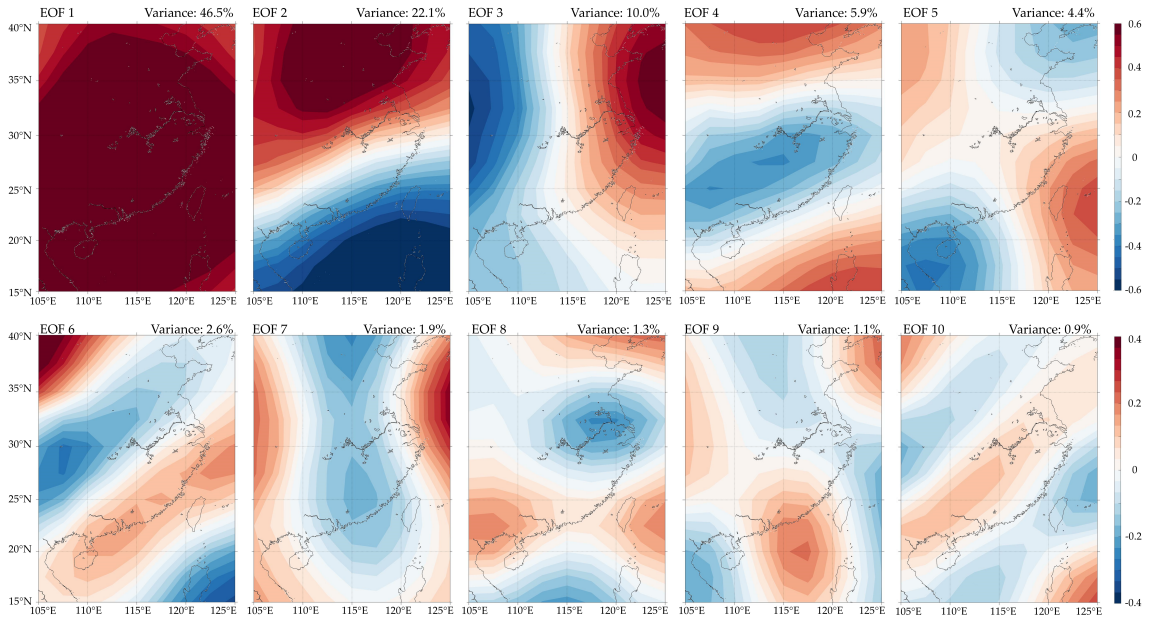
Regional Rainfall ⟺ Selective Predictors *Final model*

Blind prediction

Leave-one-year-rolling-out cross validation

….

….



Results —— *Model building & Validation & Blind prediction (Part II)*

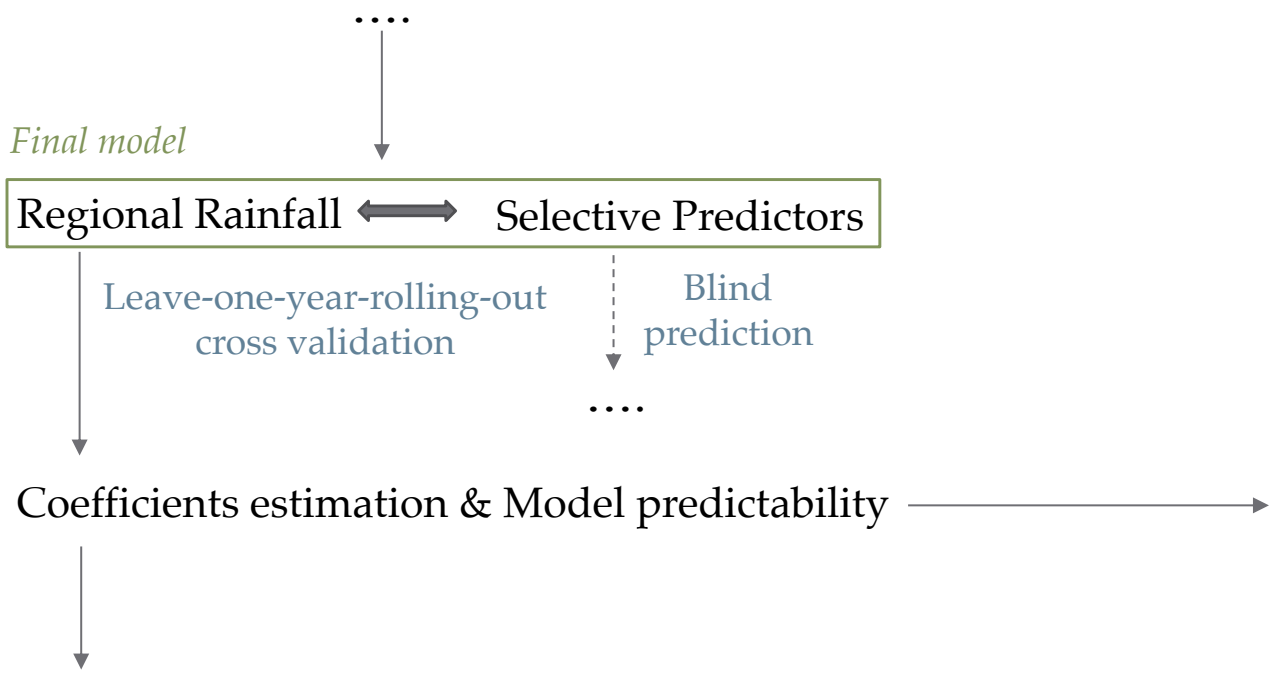


Table 1
The Assessment Metrics for CV and Blind Prediction

Assessment Metrics	Definition	Calculation[a]
R^2_{pred}	Predictive R-squared	$1-\frac{\sum_k (\hat{y}_k - y_{obs_k})^2}{\sum_k (y_{obs_k} - \bar{y}_{obs})^2}$
MSEP	Mean squared errors of prediction	$\frac{\sum_k (\hat{y}_k - y_{obs_k})^2}{K}$
PIR	Ratio of observations within 95% prediction interval (PI)	$\frac{\#\ of\ y_{obs_k}\ within\ PI}{K}$
cor	Pearson correlation between predicted values and observations	$\frac{Cov\left(Y_{obs}, \hat{Y}\right)}{\sigma_Y \sigma_{\hat{Y}}}$

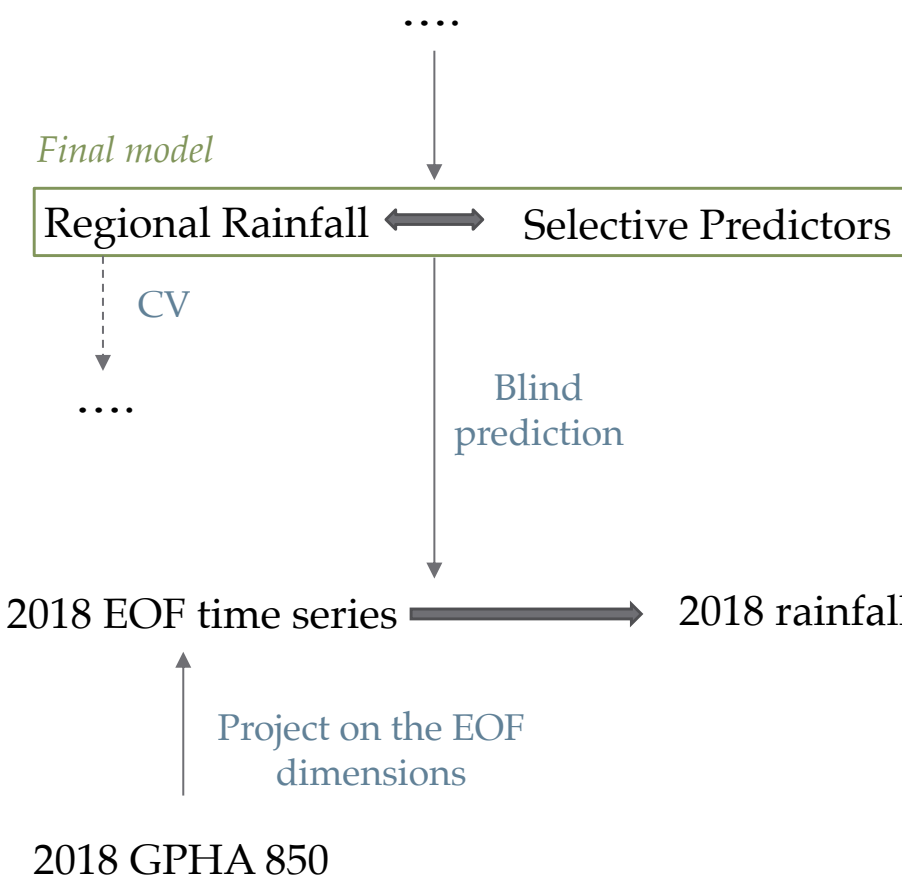
[a]Number of observations: K; Predicted values: $\hat{Y} = \{\hat{y}_k; k = 1, ..., K\}$; Observations: $Y_{obs} = \{y_{obs_k}; k = 1, ..., K\}$; $\bar{y}_{obs} = (\sum_k y_{obs_k})/K$.

Table 3
Selected Predictors and the Assessments of the Final MLR Models

Final MLR Models	Predictors[a, b]	R^2 (39-year)[c]	R^2 (38-year CV)[d]	$\overline{R^2_{pred}}$ (1-year CV)[e]
HRB	EOFs 2, 3.1, 4, 6.1, 8 and 10.1	0.39	[0.38, 0.39]	0.37
LYRB	EOFs 2, 4, 5.1, 6.1, 7, 8, 9 and 10	0.47	[0.46, 0.47]	0.45
SCR	EOFs 1, 4, 5, 6, 7, 8, 9 and 10	0.50	[0.50, 0.51]	0.48

[a]Predictors with (without) an underline denote a negative (positive) sign of their regression coefficients. [b]EOF 3.1 denotes EOF 3 at lead 1 day and a similar convention applies to others. [c]R-squared based on 39-year data (1979 – 2017). [d]R-squared range based on 38-year data in each CV loop. [e]Average predictive R-squared based on the remaining 1-year data in each CV loop.

Results —— *Model building & Validation & Blind prediction (Part III)*



	HRB	LYRB	SCR
R^2_{pred}	0.45	0.50	0.41
MSEP	0.29	0.24	0.18
PIR	0.97	0.98	0.98
cor	0.67	0.73	0.68

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

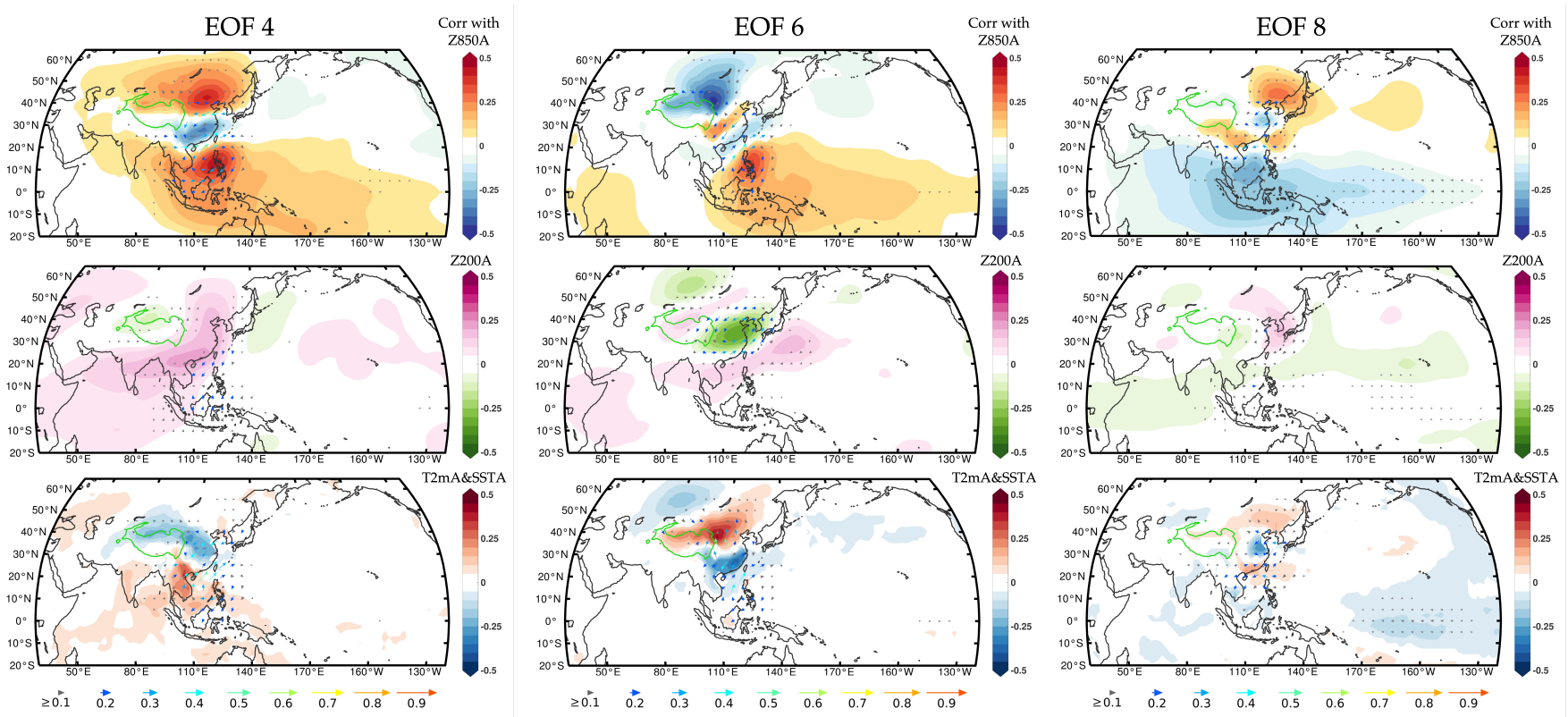
Results ——— *Interpretation of informative predictors*

Correlation maps

Z850A (Shaded)
IVTA (vectors)

Z200A (Shaded)
uv200A (vectors)

T2mA&SSTA
(Shaded)
uv10m (vectors)



香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

Conclusions and Some Highlights

- Distinct rain belt patterns classified by the Self-organizing Map are highly in line with the summer monsoon front in Southeast China.
- Redefined rain belts lead to new discovery of informative regional circulation anomalies and upper-level wave trains for improved rainfall predictability.
- Unprecedented regional daily rainfall prediction skills are achieved over the three rain belts in Southeast China.
- Offered some insights on the source of biases embedded in numerical models, possibly amplified in statistical downscaling of climate model outputs with a deeper understanding of monsoonal dynamics.

Thank you for your attention!