Regional frequency analysis of hydro-meteorological extremes

Non-standard aspects

Dhouha Ouali, Ph.D.

Research associate, MEOPAR-PCIC

Oct. 25th 2017

Introduction…

Flood risk

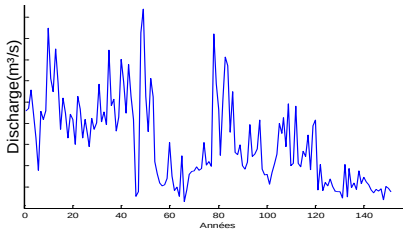
Drought risk

Risk Assessment

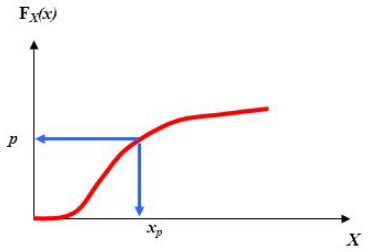
water resource management

Frequency Analysis

General scheme



- Extract a sample of extreme values from the discharges series
- Fit an adequate statistical distribution
- Estimate parameters
- Compute quantities of interest (quantiles)



Frequency Analysis

1. **Available** data series are **too short**

2. **Unavailable** data series

Consequences:

1. Large uncertainties

2. No estimates of flood design

Using data from different sites can reduce the uncertainties and provide estimates at ungauged sites

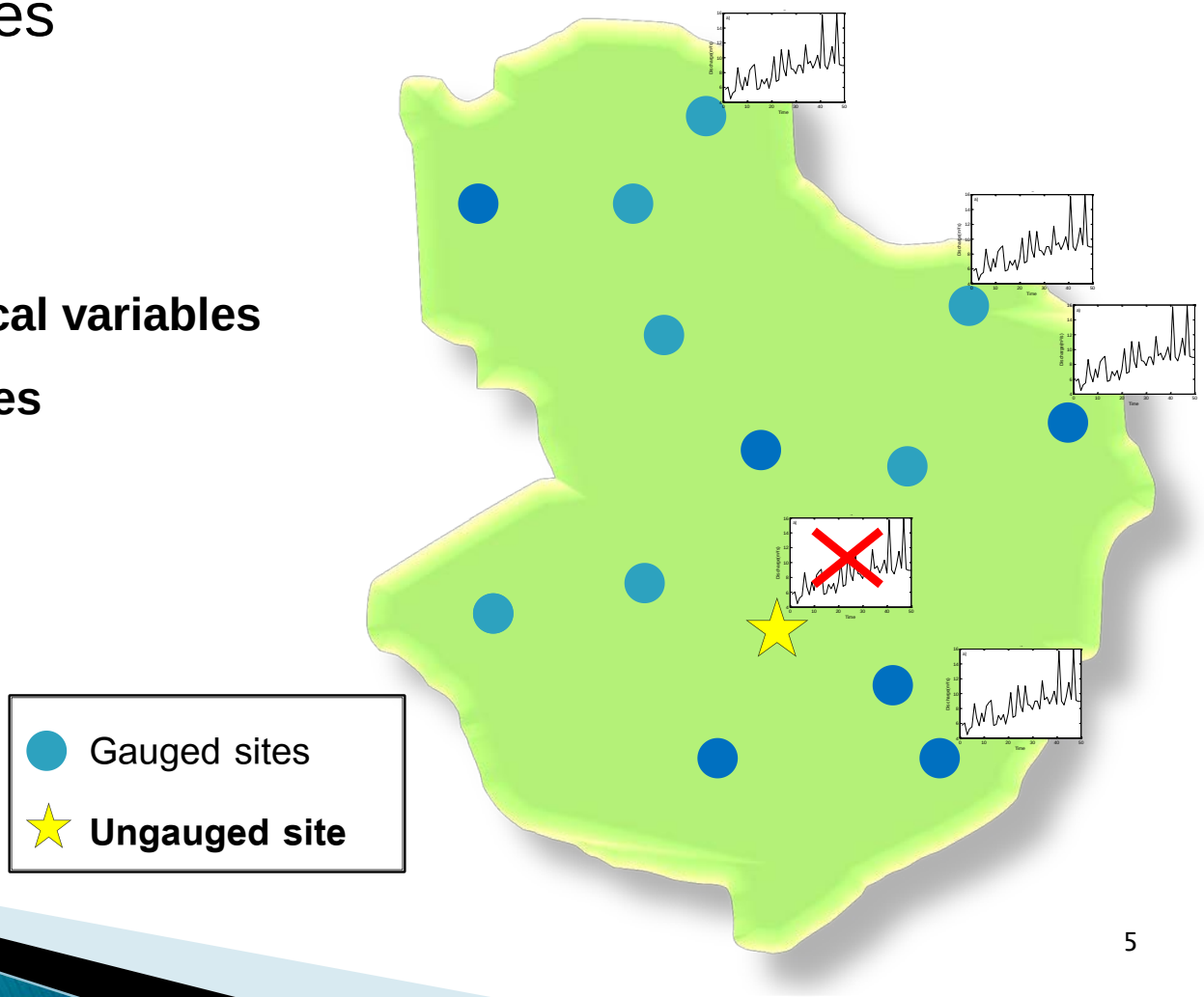
Regional Frequency analysis (RFA)

- Basic principles

Each gauged site:

X: Physio-meteorological variables

Y: Hydrological variables



Regional Frequency analysis (RFA)

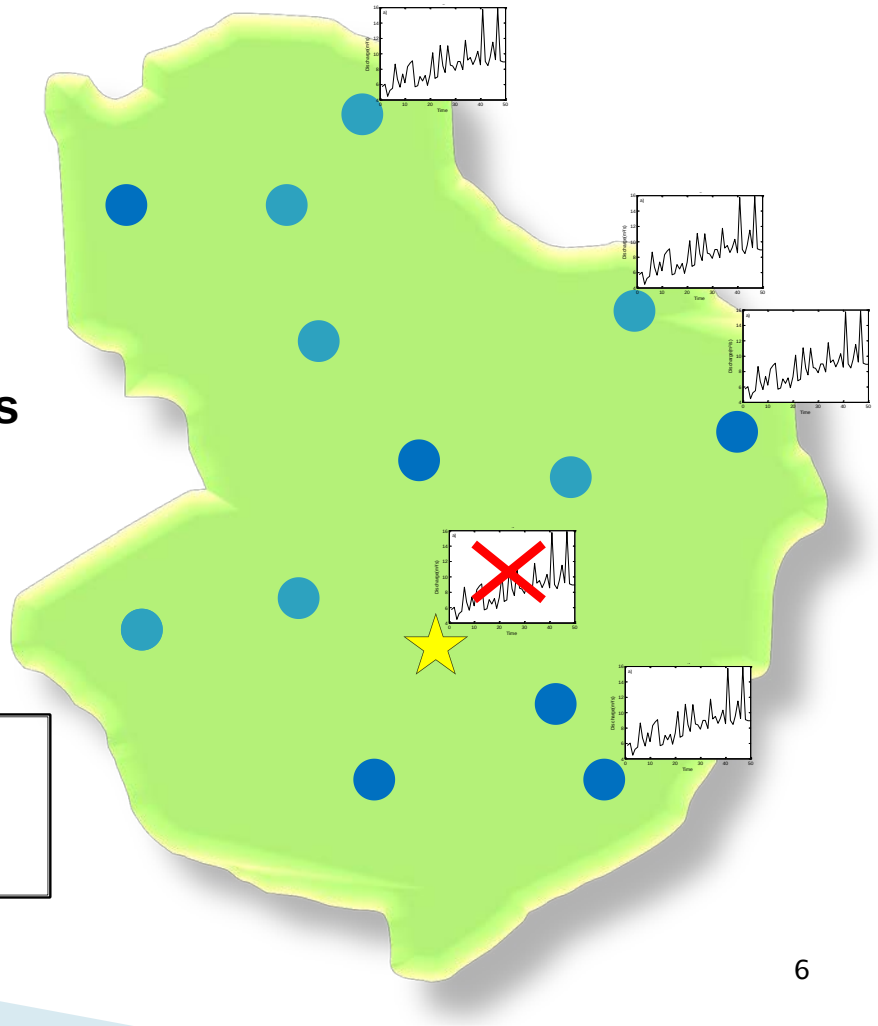
1. At-site frequency Analysis

Each gauged site:

X: Physio-meteorological variables

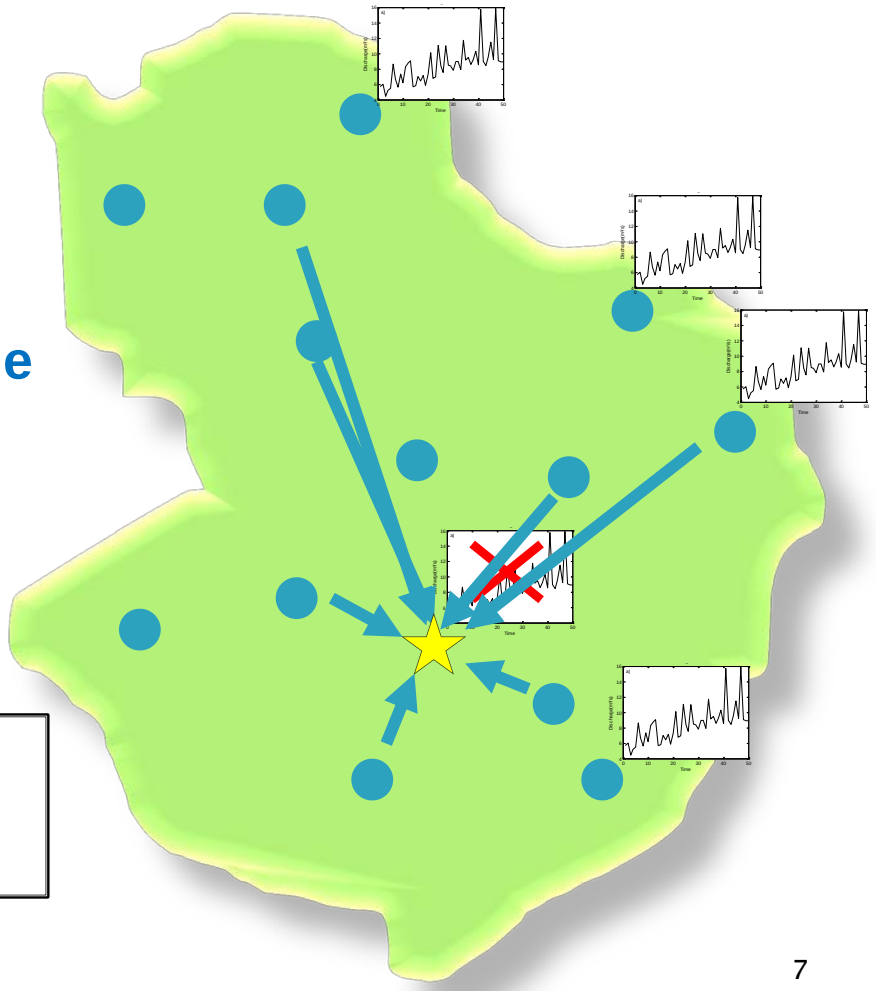
Y: Hydrological variables hydrologiques

Quantile estimate at gauged sites



Regional Frequency analysis (RFA)

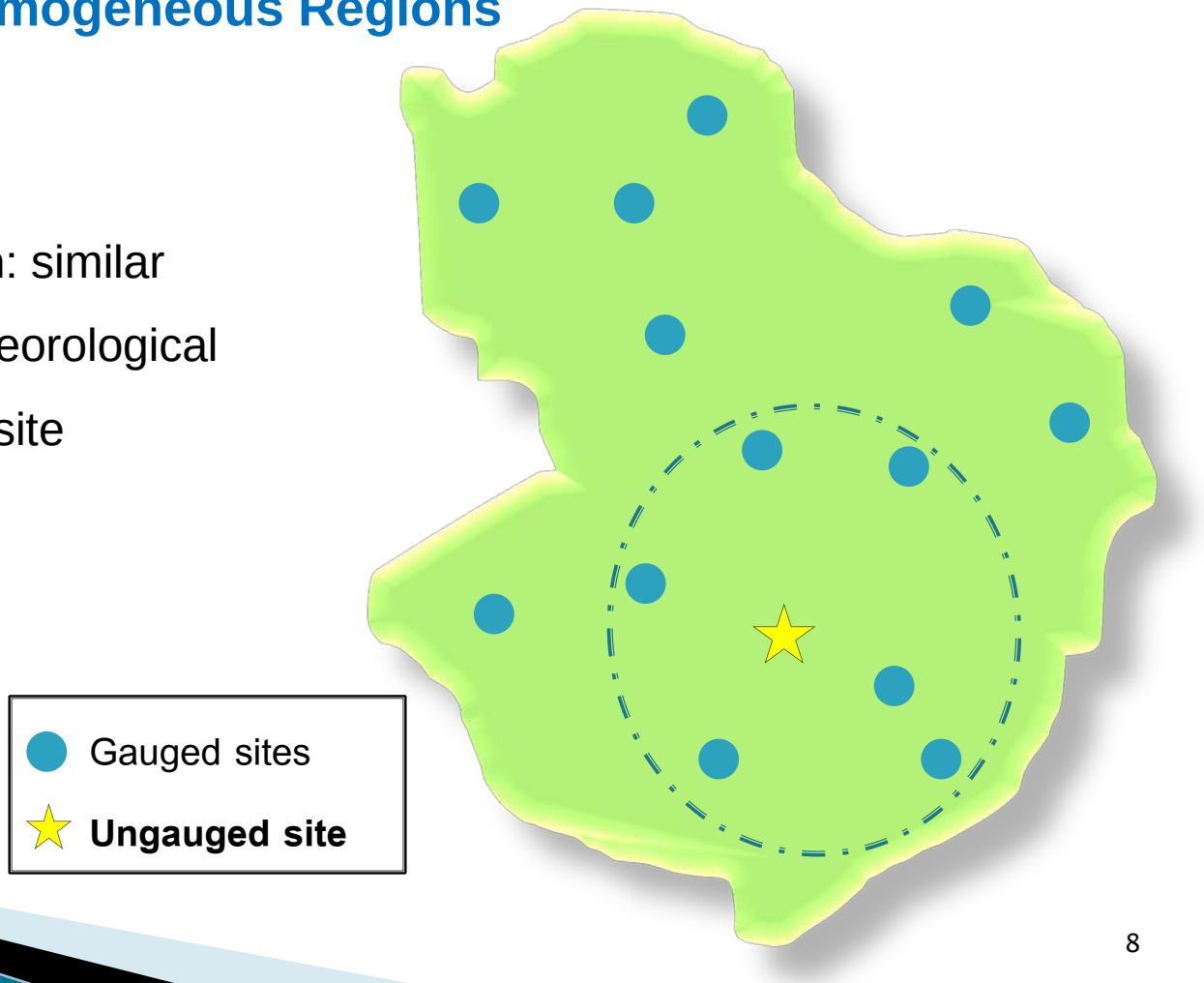
Transfer of the hydrological information from gauged sites to the ungauged site



Regional Frequency analysis (RFA)

2. Delineation of Homogeneous Regions

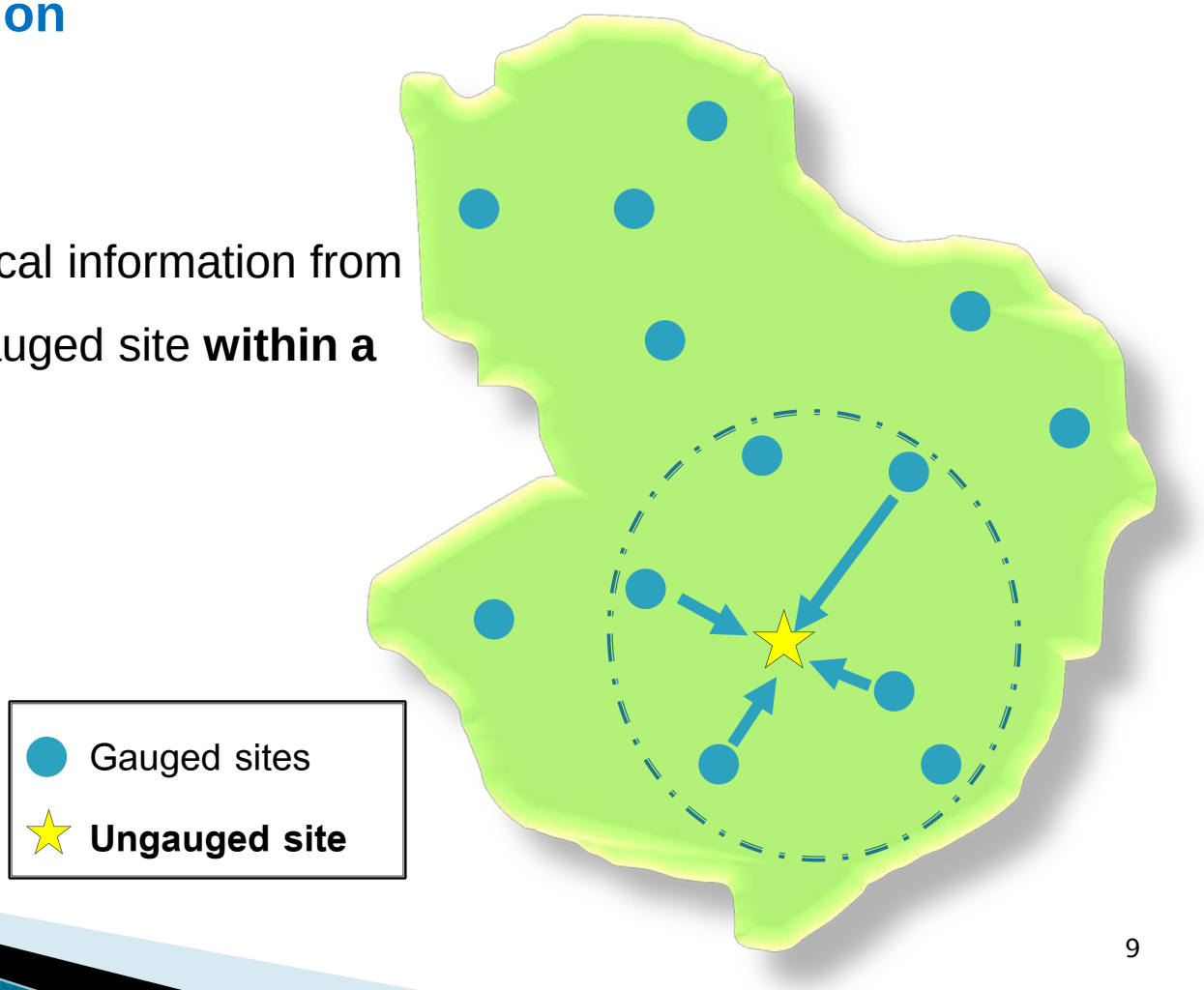
Homogeneous region: similar physiographic and meteorological attributes to the target site



Regional Frequency analysis (RFA)

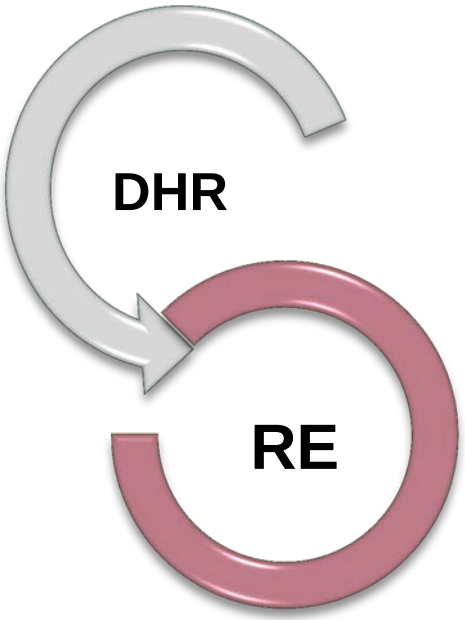
3. Regional Estimation

Transfer of the hydrological information from gauged sites to the ungauged site **within a homogeneous region**

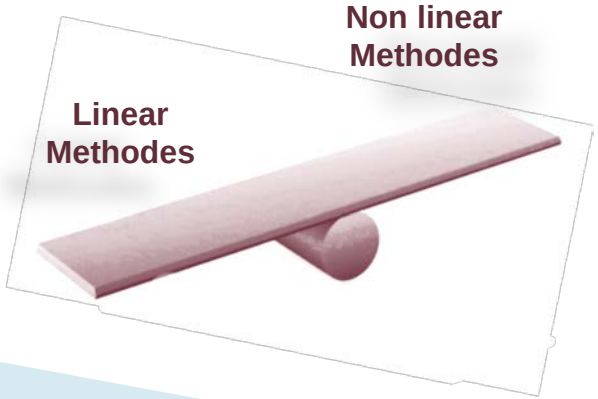


Problems definition

A. Nonlinear modeling of the hydrological processes



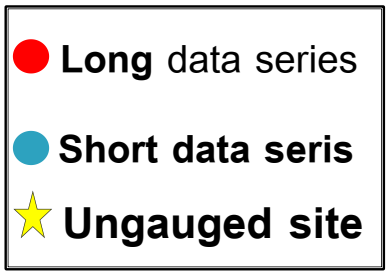
Not exploited

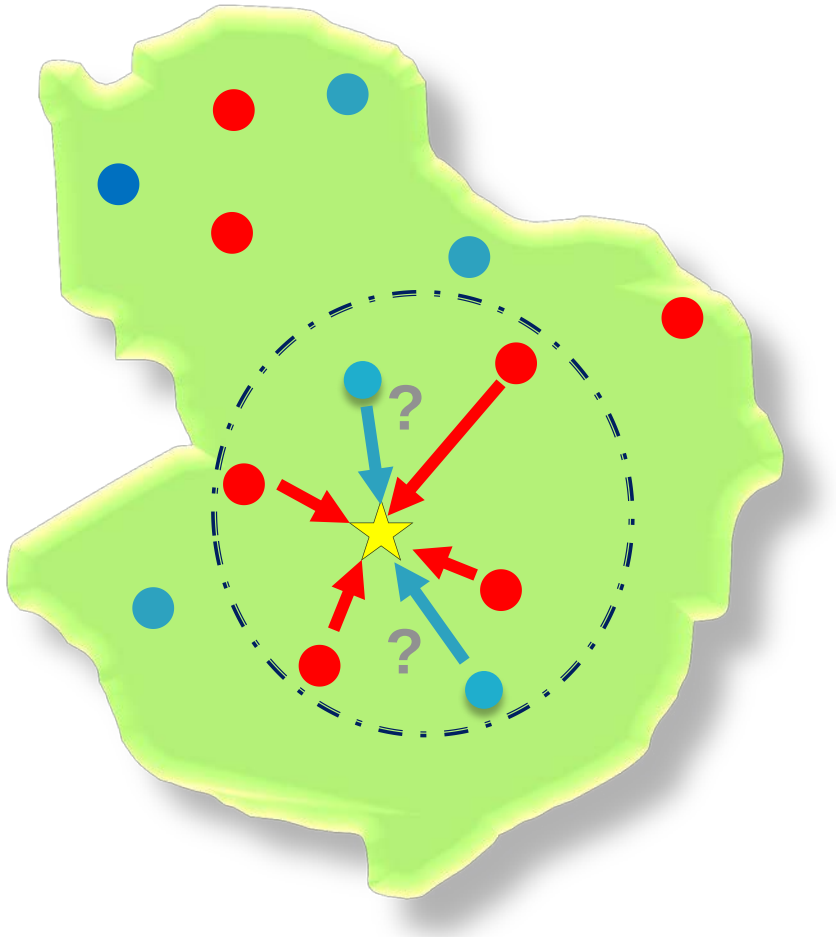


Problems definition

B. Exploitation of the hydrological information

Ignore sites with short data sets





Proposed methodologies

Linear approaches in DHR

- Canonical Correlation Analysis (CCA)

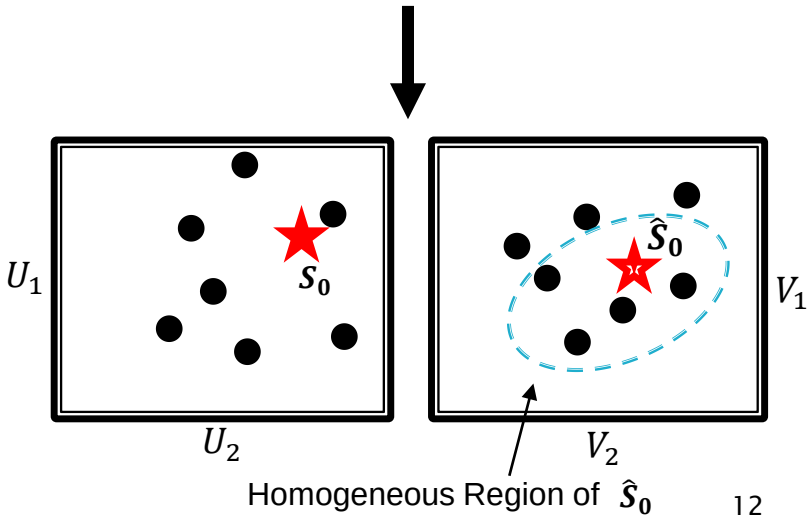
$$X = \begin{Bmatrix} X_1 \\ \vdots \\ \vdots \\ X_q \end{Bmatrix} \xrightarrow{\text{U=a X}} U = \begin{Bmatrix} U_1 \\ \vdots \\ \vdots \\ U_r \end{Bmatrix}$$

Maximize

$$cor(U,V)$$

Linear Relationships

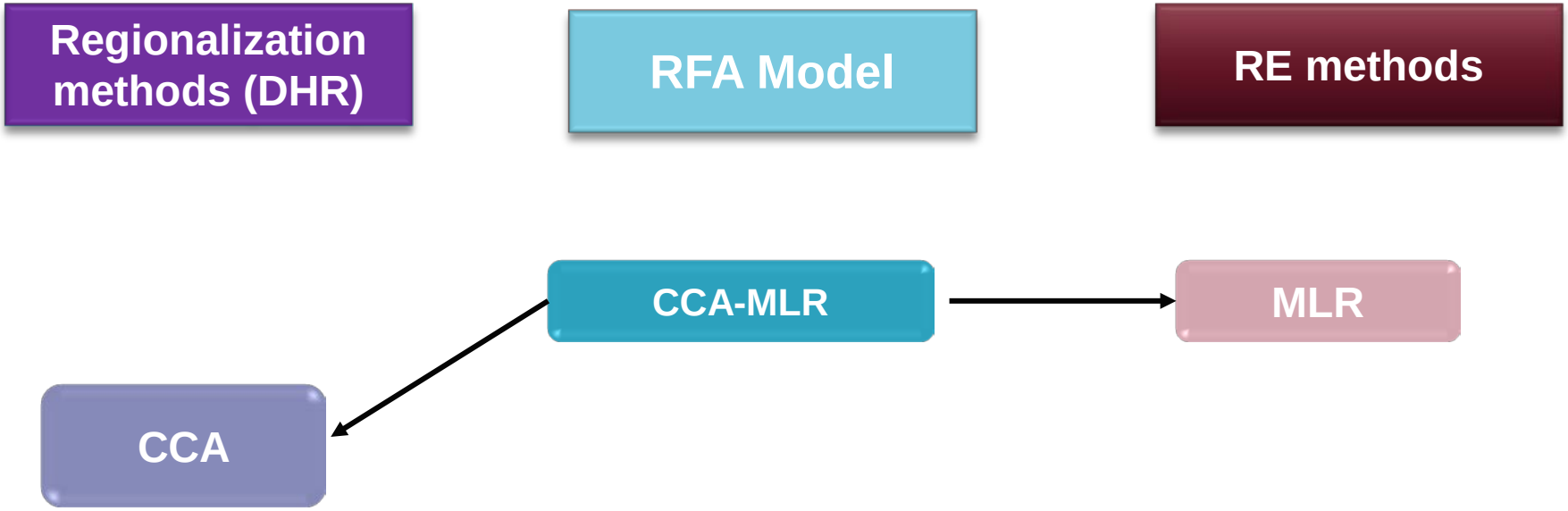
$$Y = \begin{Bmatrix} Y_1 \\ \vdots \\ \vdots \\ Y_r \end{Bmatrix} \xrightarrow{\text{V = b Y}} V = \begin{Bmatrix} V_1 \\ \vdots \\ \vdots \\ V_r \end{Bmatrix}$$



Proposed methodologies

Linear approaches in DHR

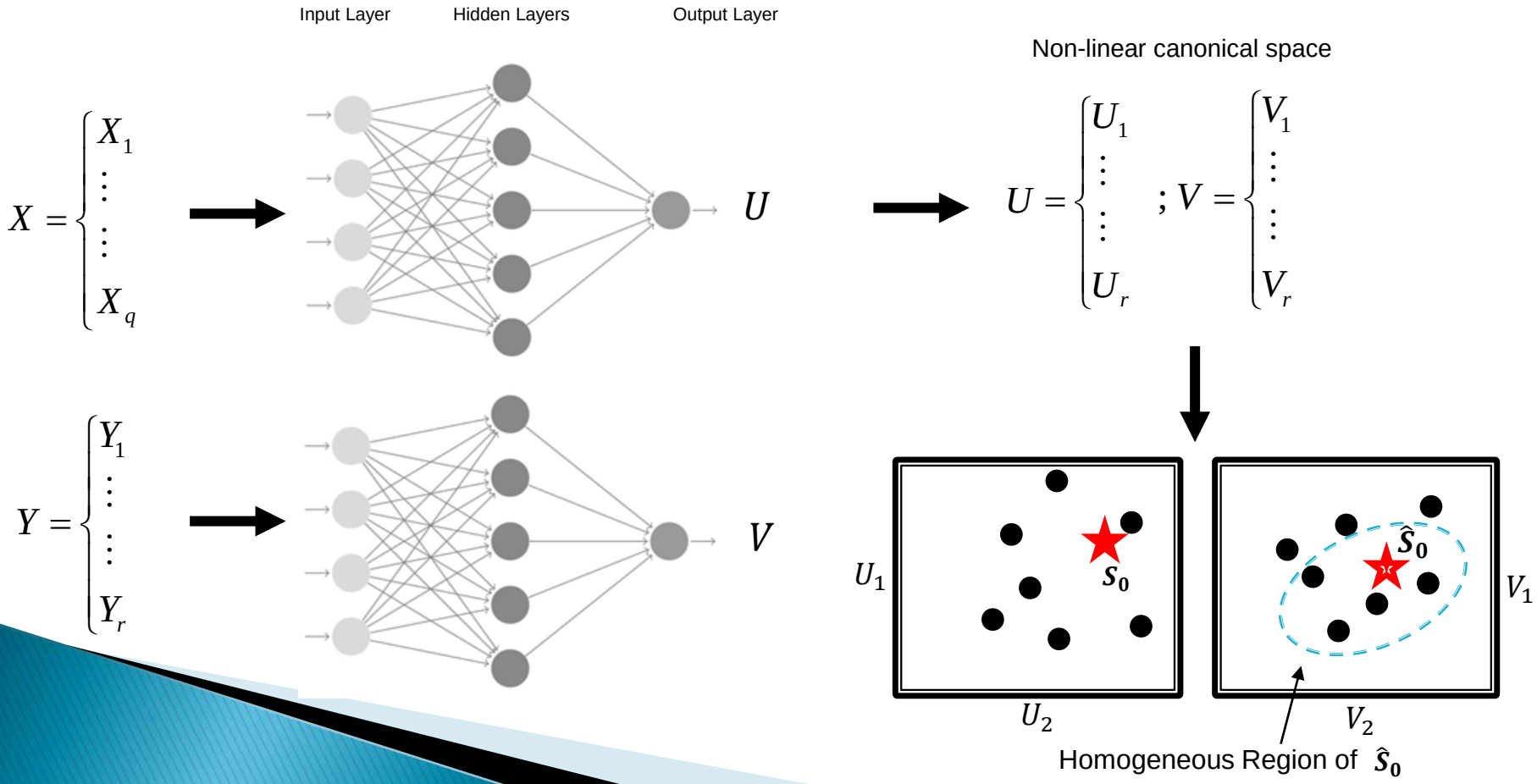
- Canonical Correlation Analysis (CCA)



Proposed methodologies

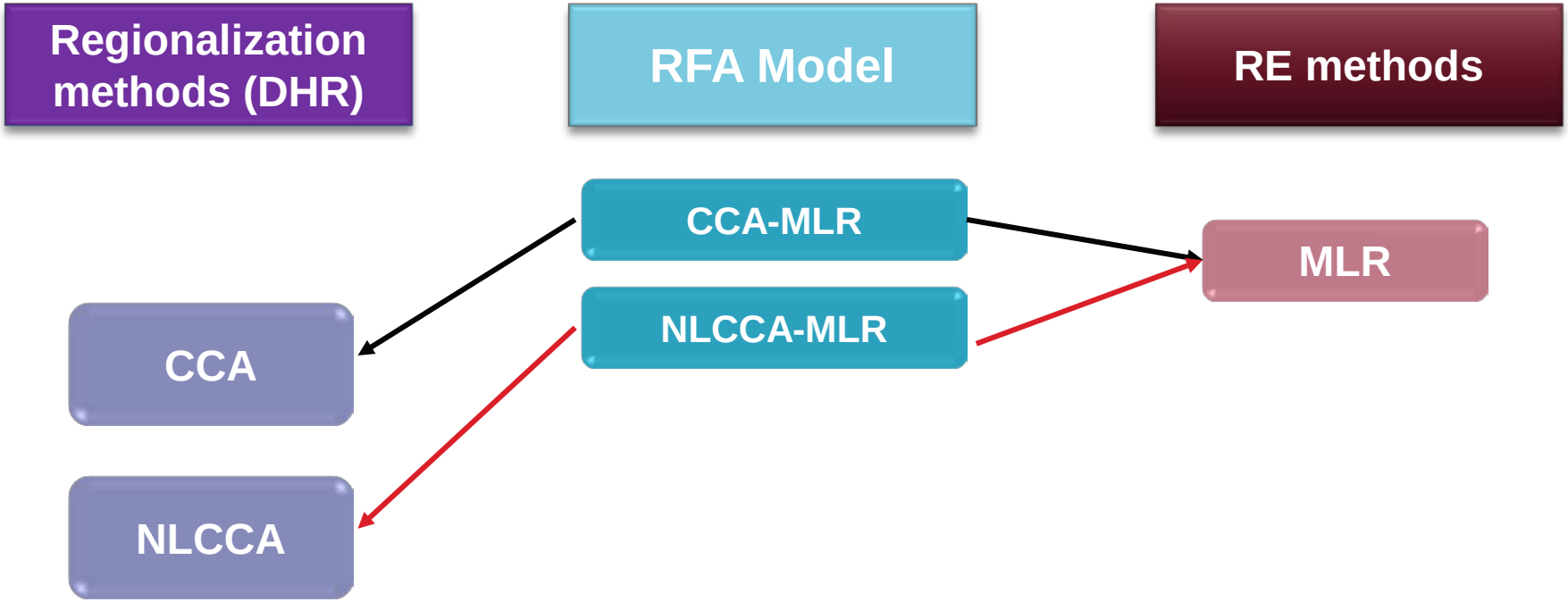
A.1.Non-linear approaches in DHR

- Non-linear Canonical Correlation Analysis (NLCCA)



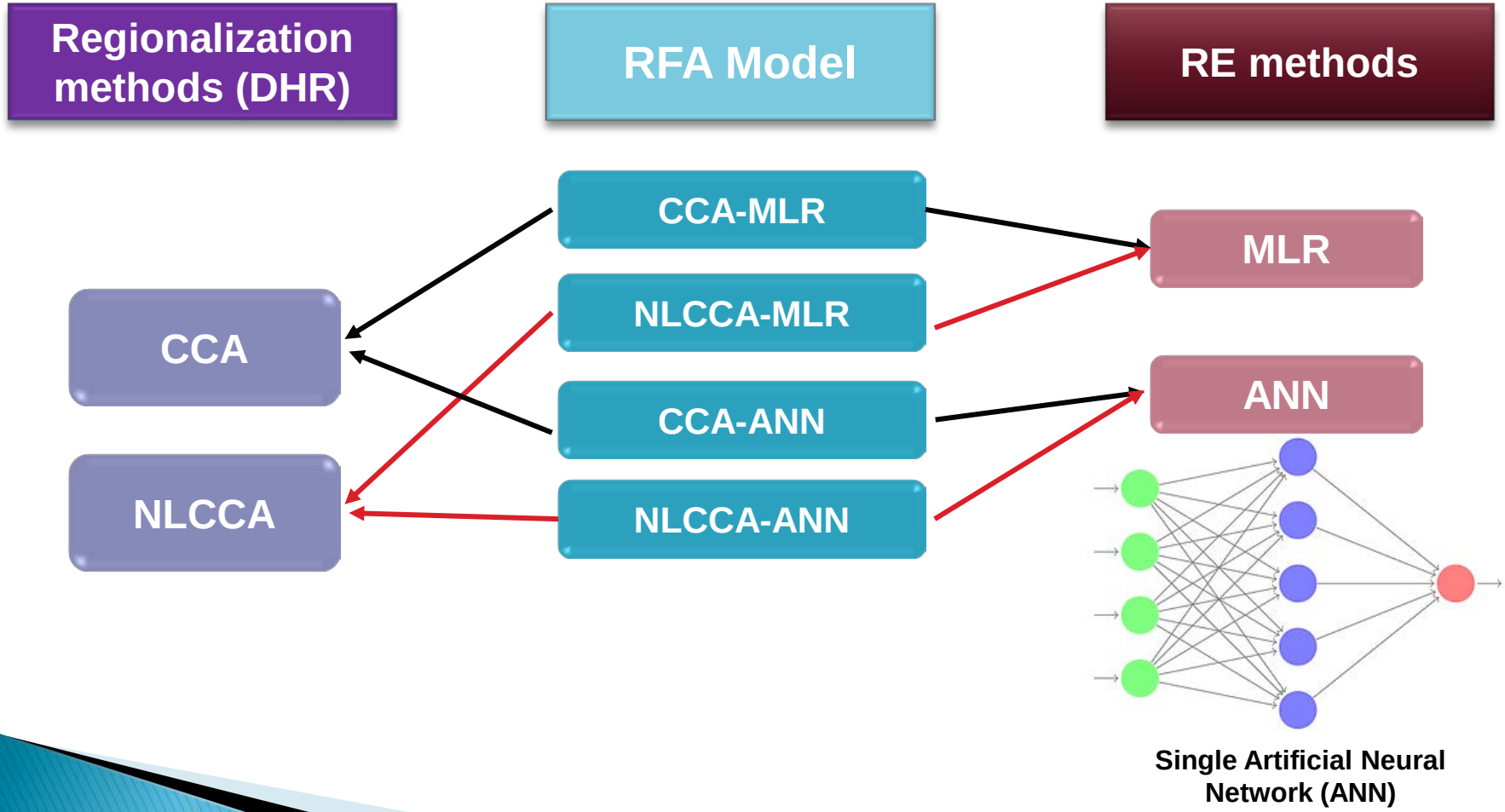
Proposed methodologies

A.1.Non-linear approaches in DHR



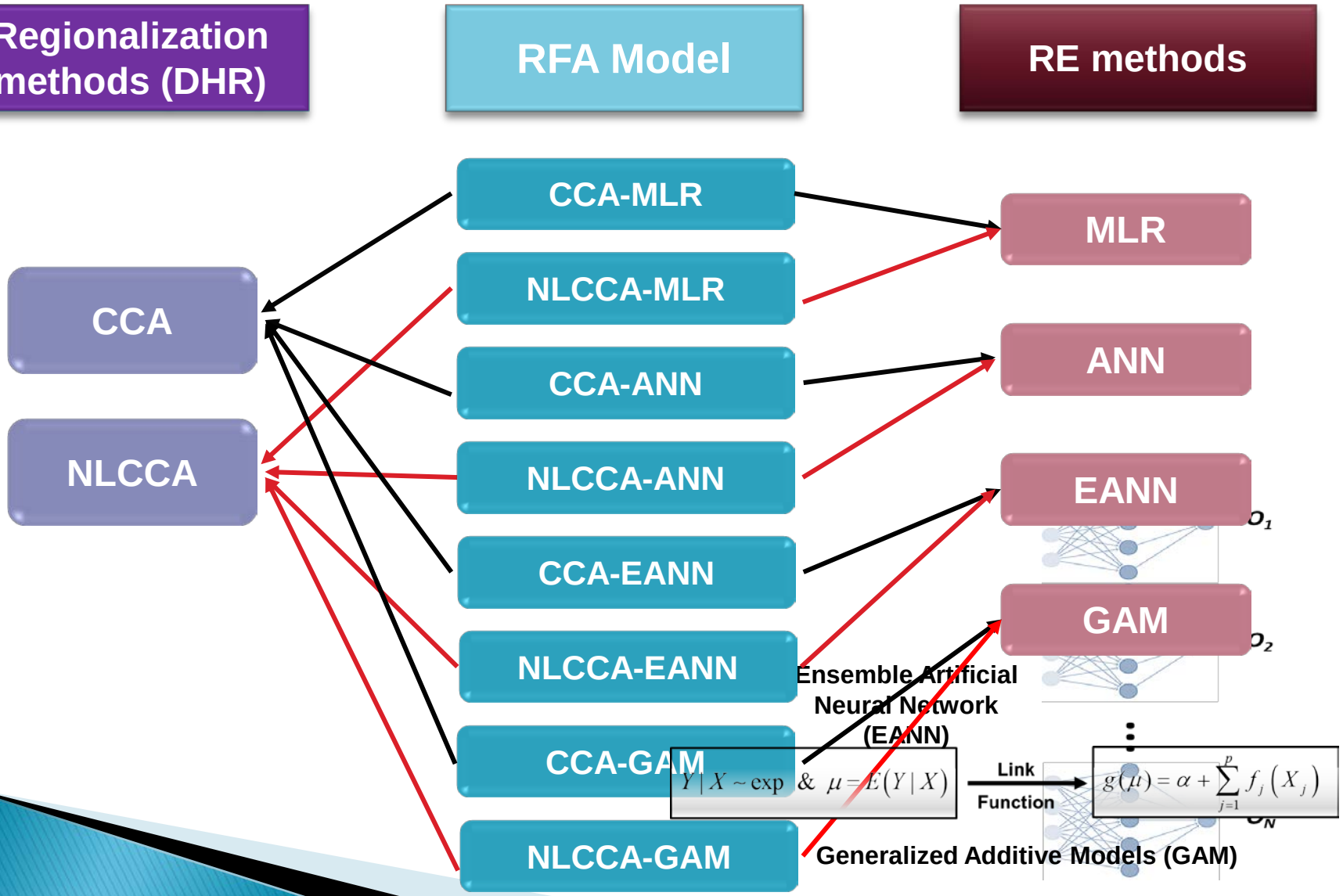
Proposed methodologies

A.2.Non-linear approaches in RE

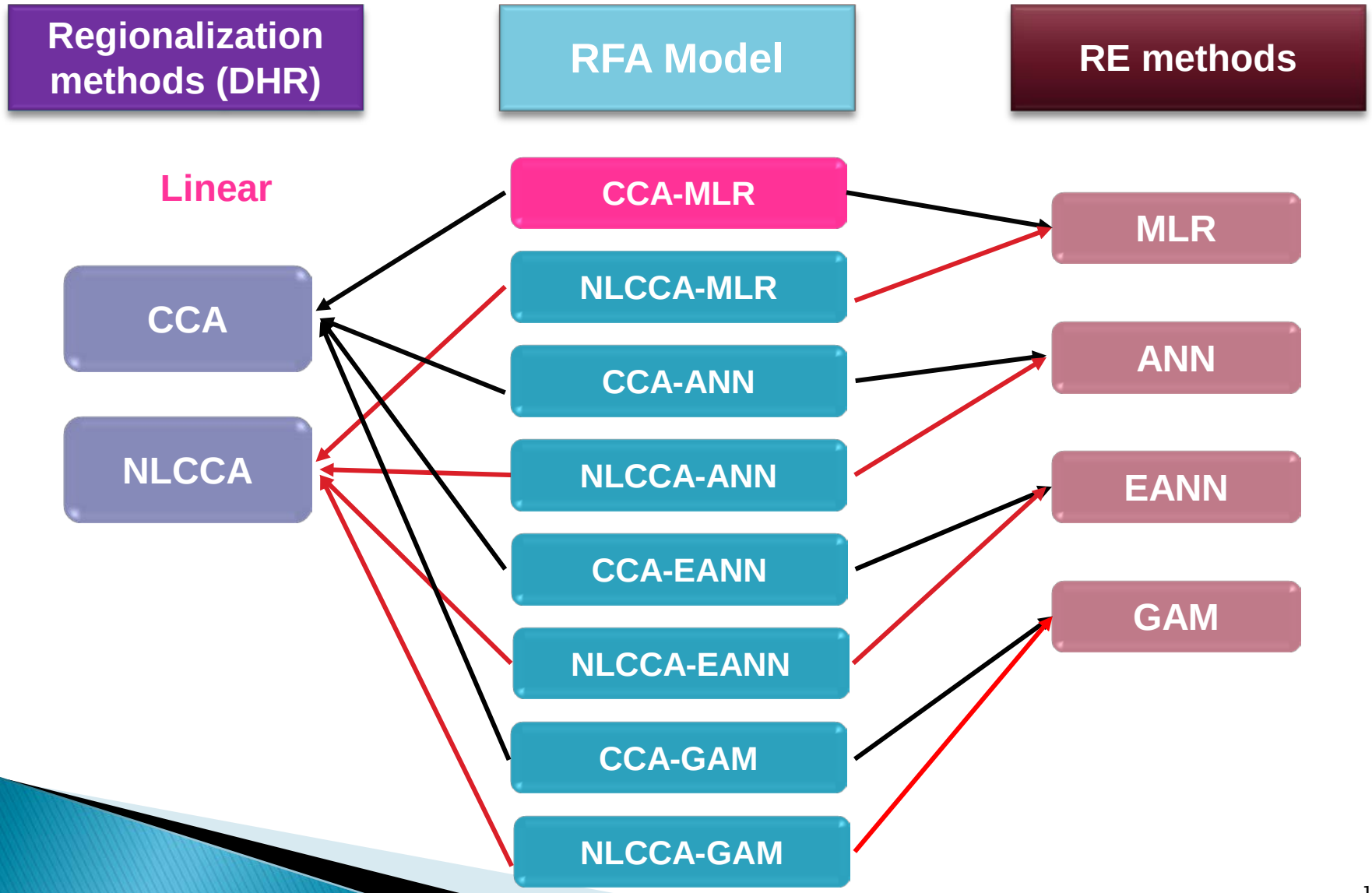


Single Artificial Neural Network (ANN)

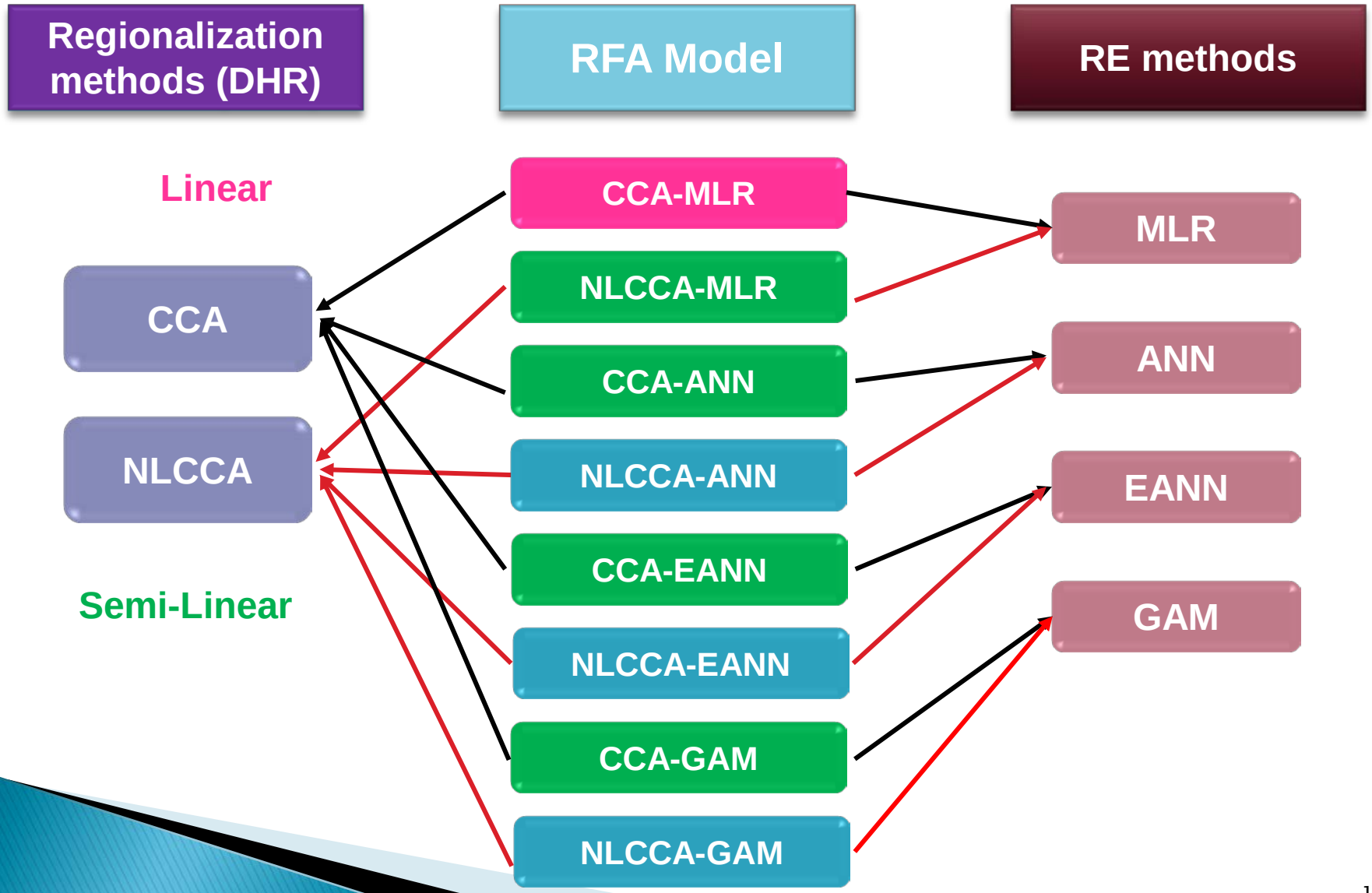
Proposed methodologies



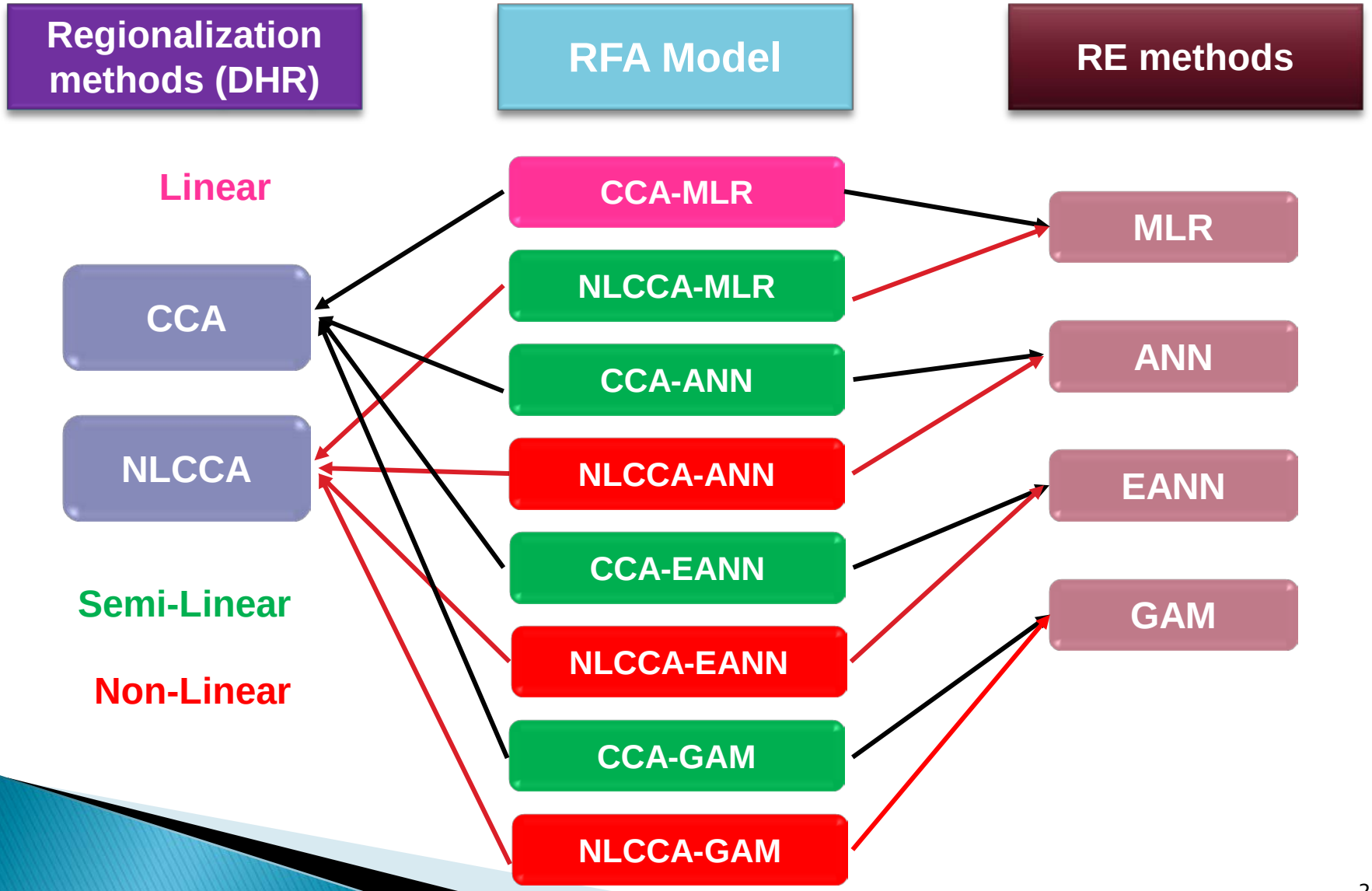
Proposed methodologies



Proposed methodologies



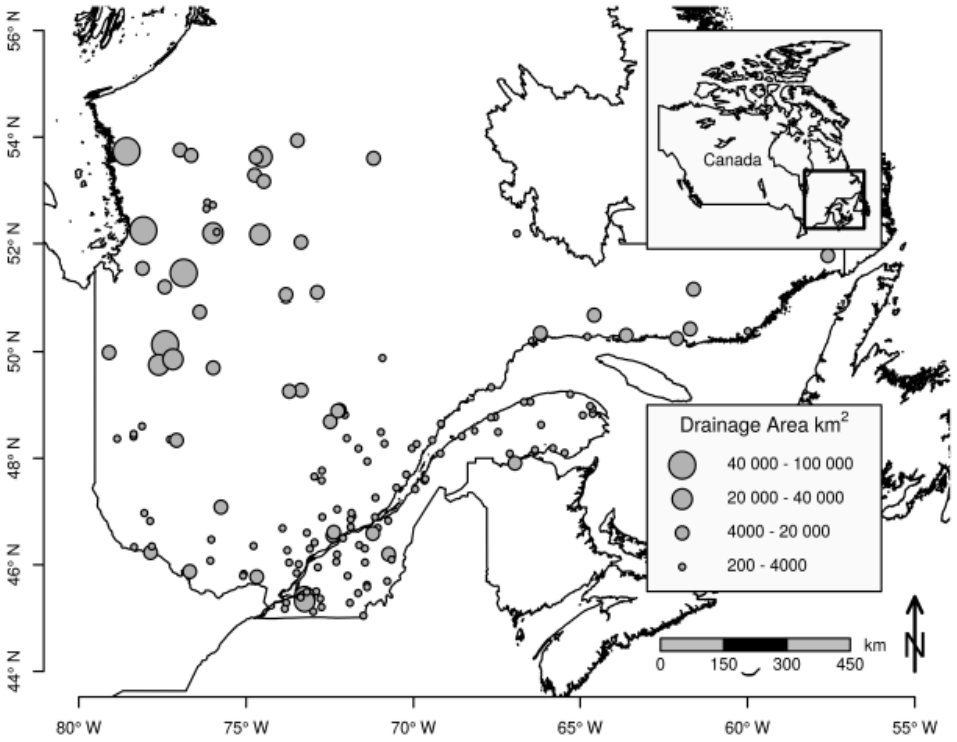
Proposed methodologies



Data

- Southern Quebec, Canada;
- 151 hydrometric stations;
- 5 physio-meteorological variables
- 3 hydrological variables: QS10, QS50 and QS100;
- Annual maximum discharge :

  1900 - 2002



Results

A. Nonlinear modeling of the hydrological processes

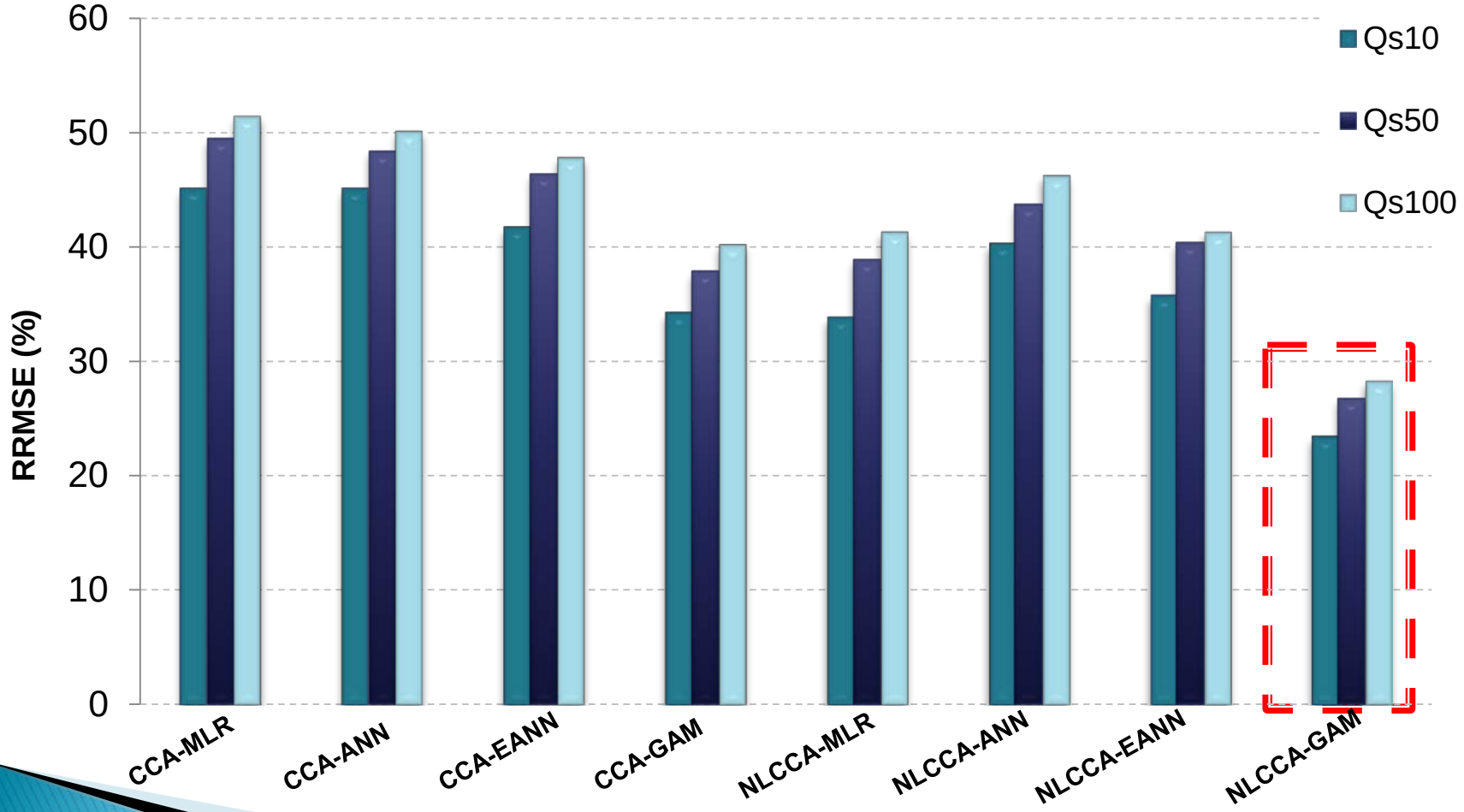
Leave-one-out cross-validation procedure

For k=1…N sites

- Remove temporary the k^{th} site from the dataset, **ungauged site**
- Train the RFA model using the remaining sites
- Compare regional and at-site estimated quantiles (RMSE, BIAS)

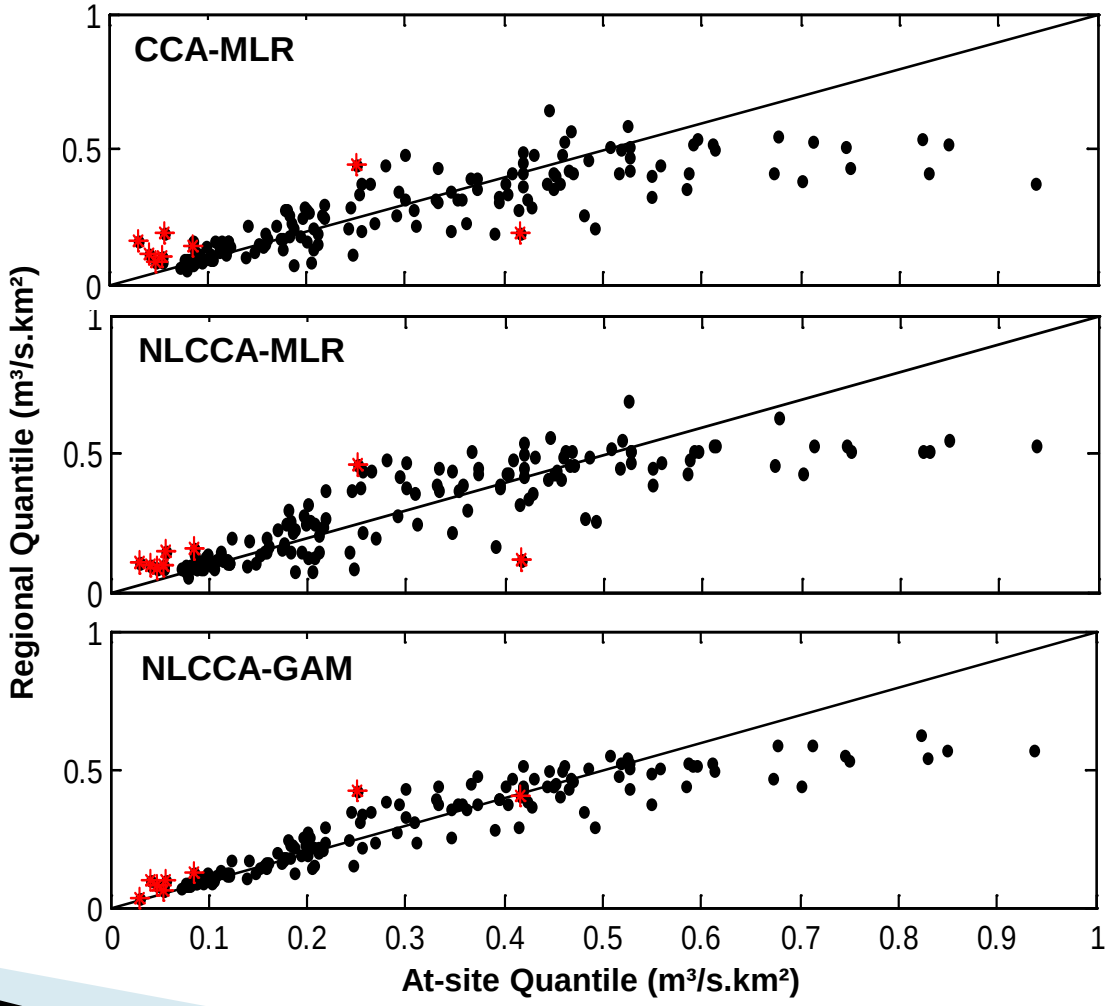
Results

A. Nonlinear modeling of the hydrological processes



Results

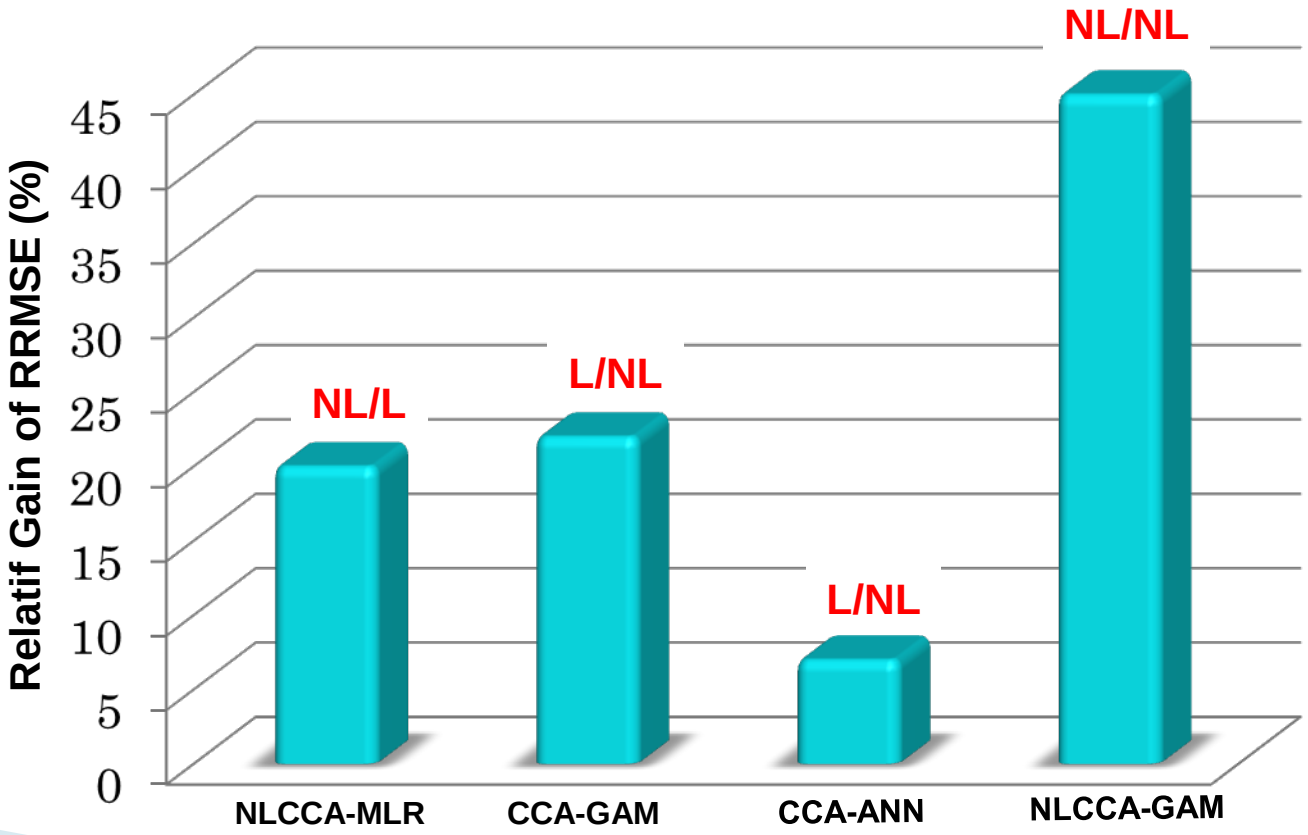
A. Nonlinear modeling of the hydrological processes



Results

A. Nonlinear modeling of the hydrological processes

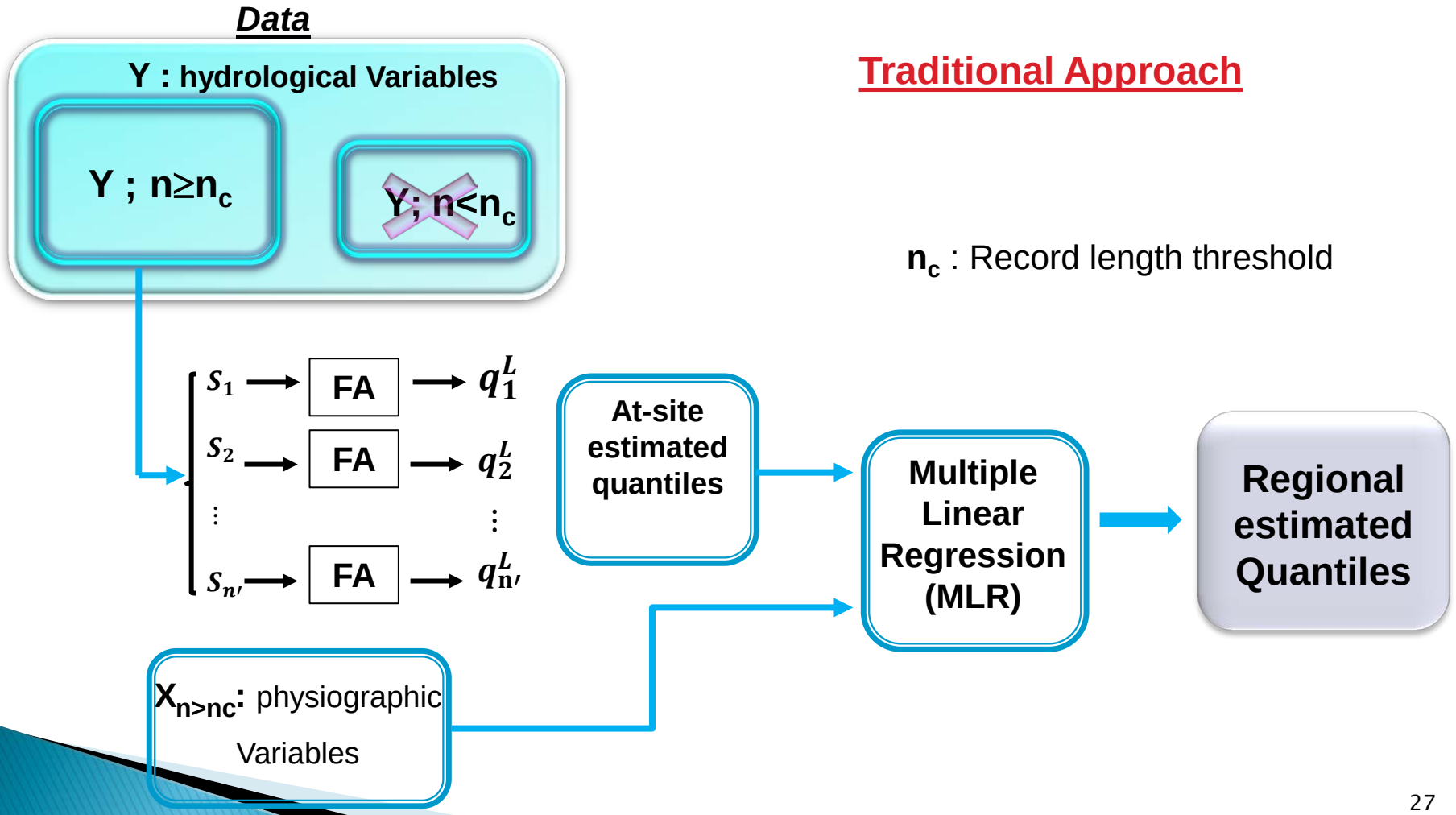
Relatif Gain of RRMSE compared to the CCA-MLR model, Qs_{100}



B. Exploitation of the hydrological information

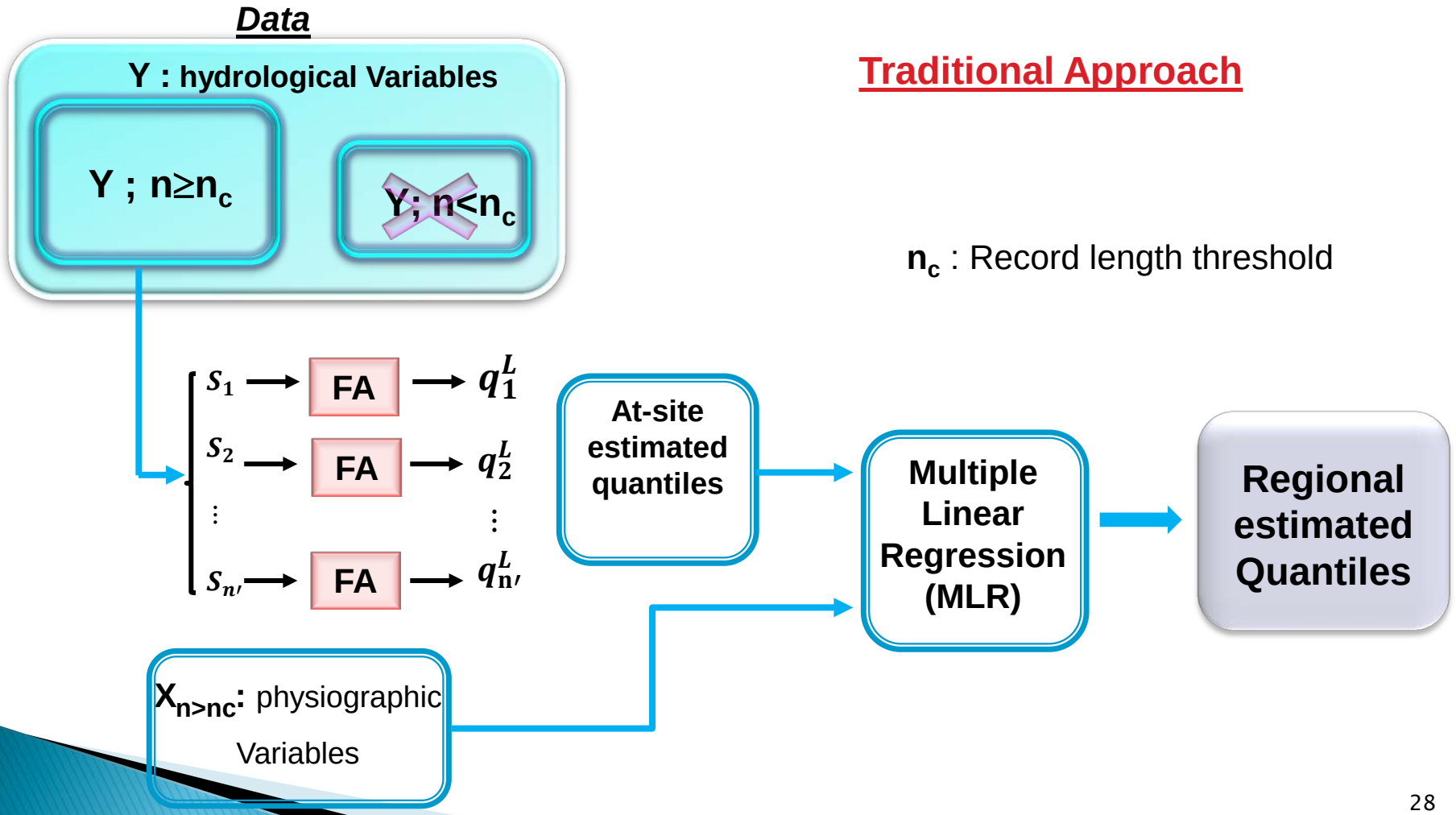
Proposed methodologies

B. Exploitation of the hydrological information



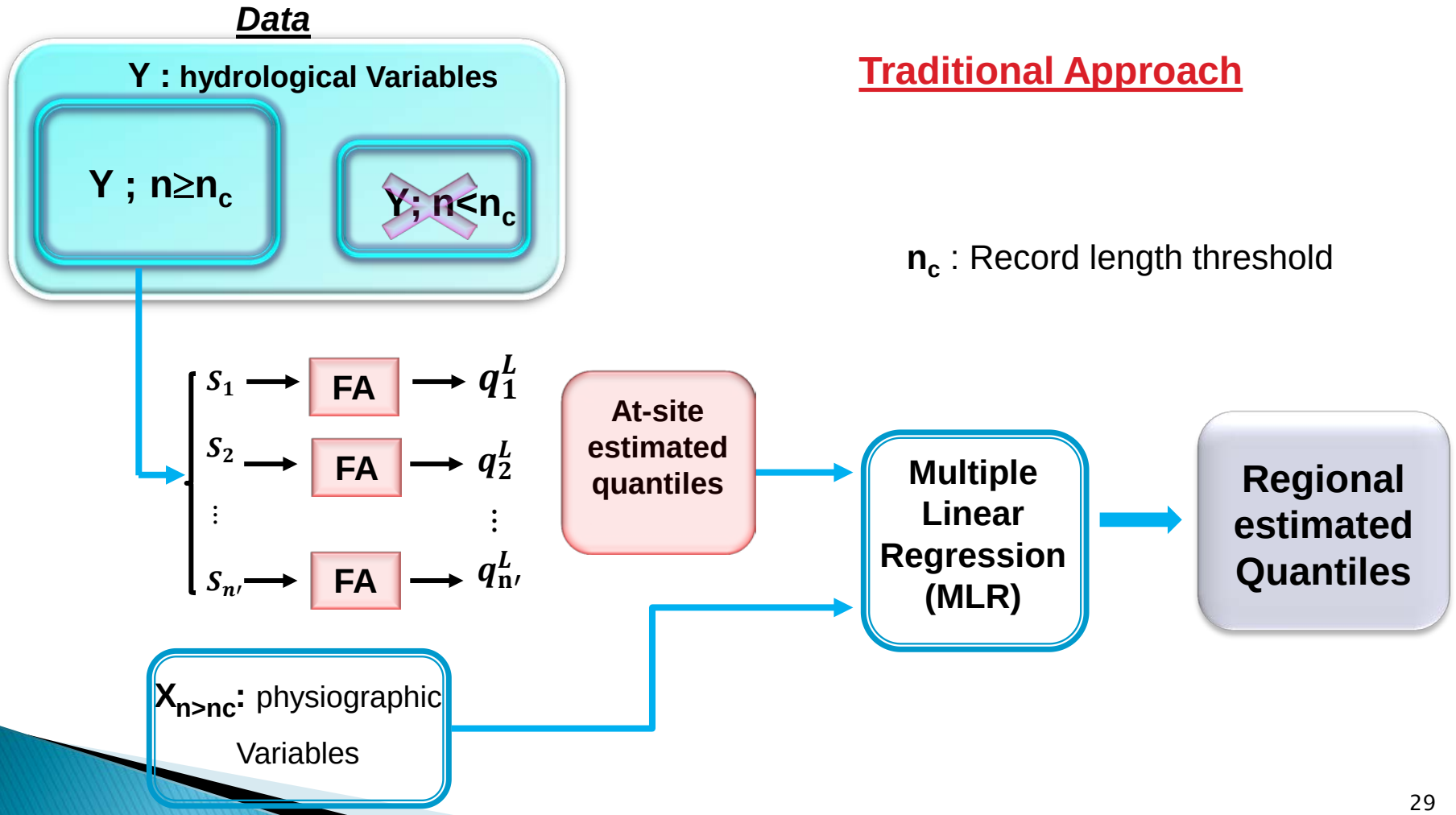
Proposed methodologies

B. Exploitation of the hydrological information



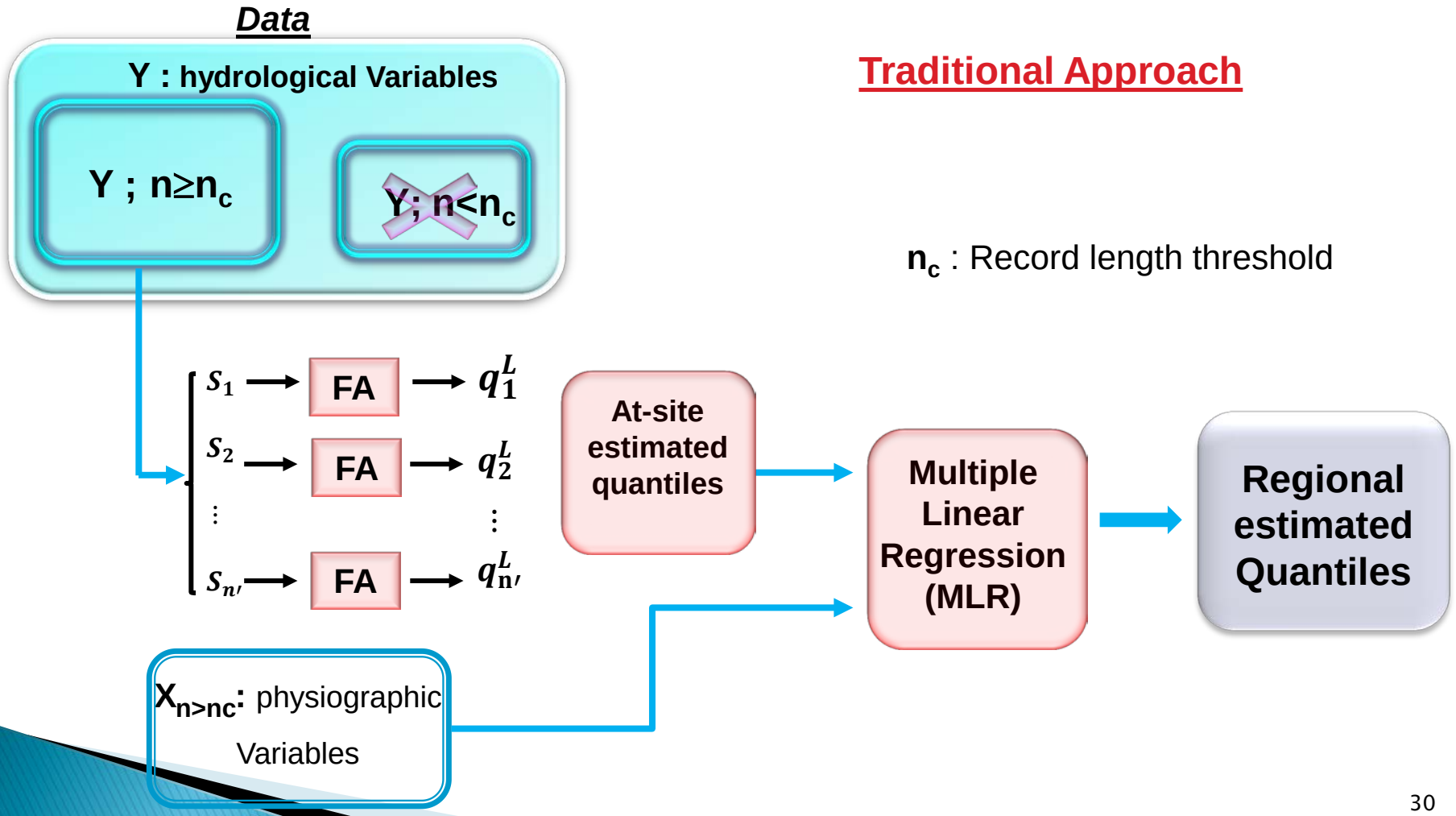
Proposed methodologies

B. Exploitation of the hydrological information



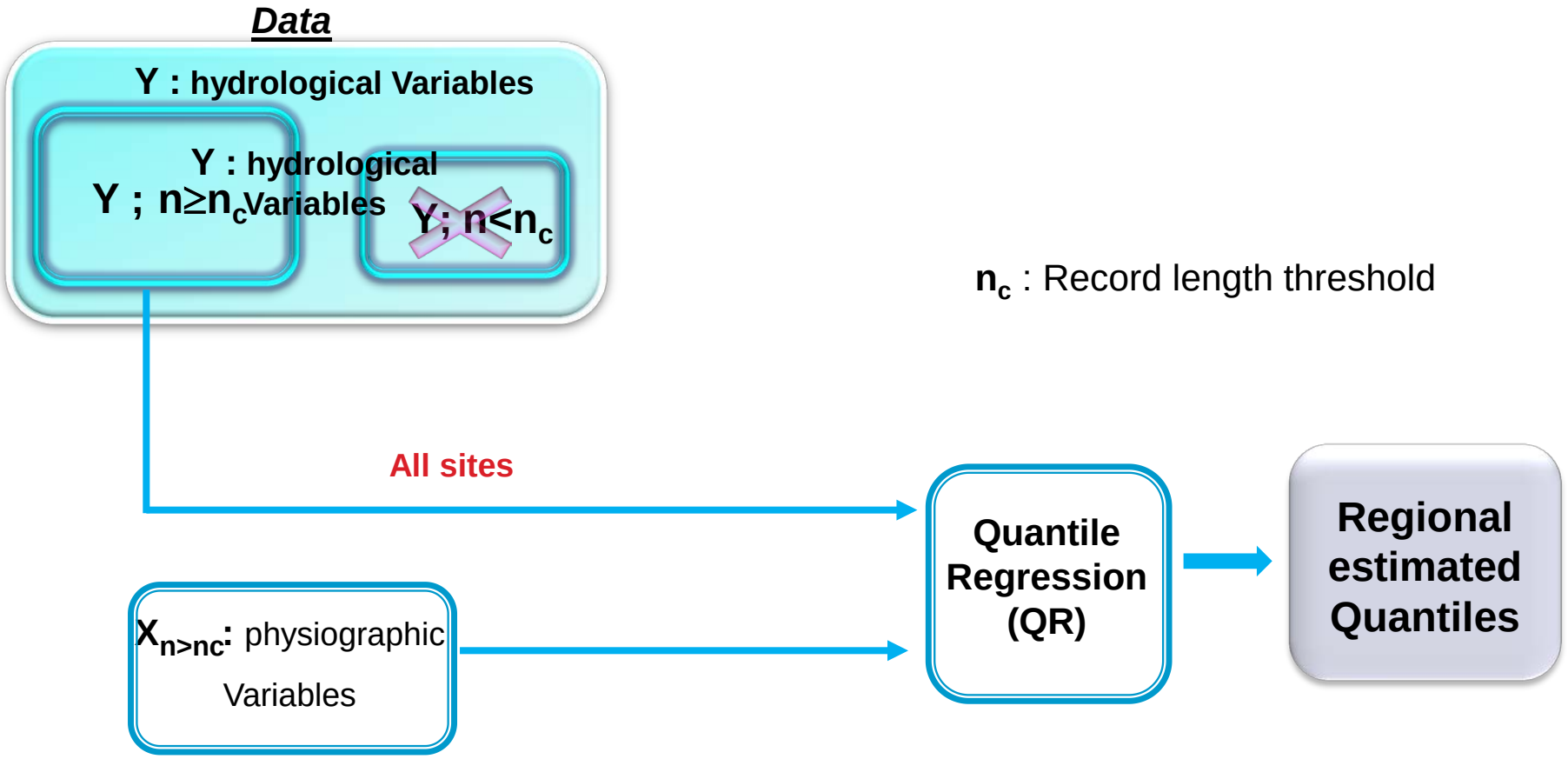
Proposed methodologies

B. Exploitation of the hydrological information



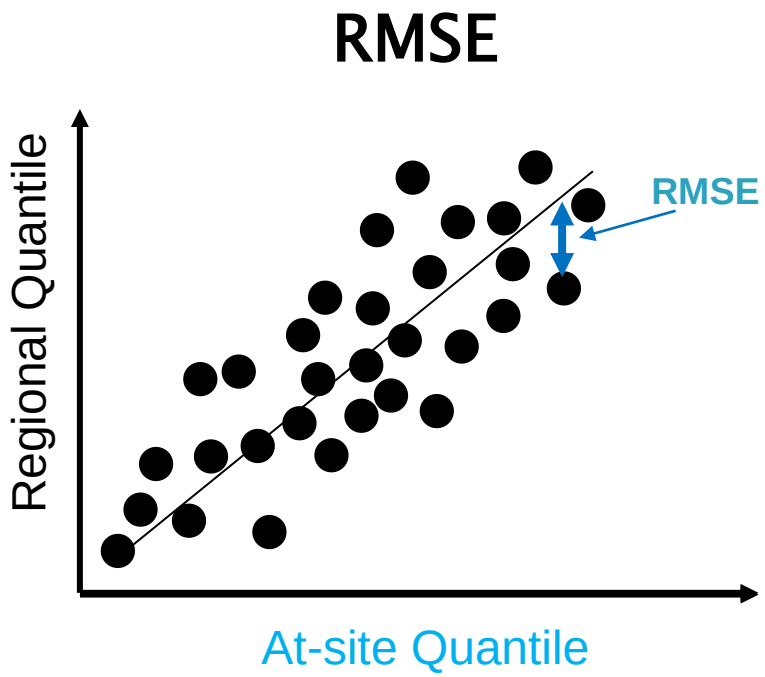
Proposed methodologies

B. Exploitation of the hydrological information

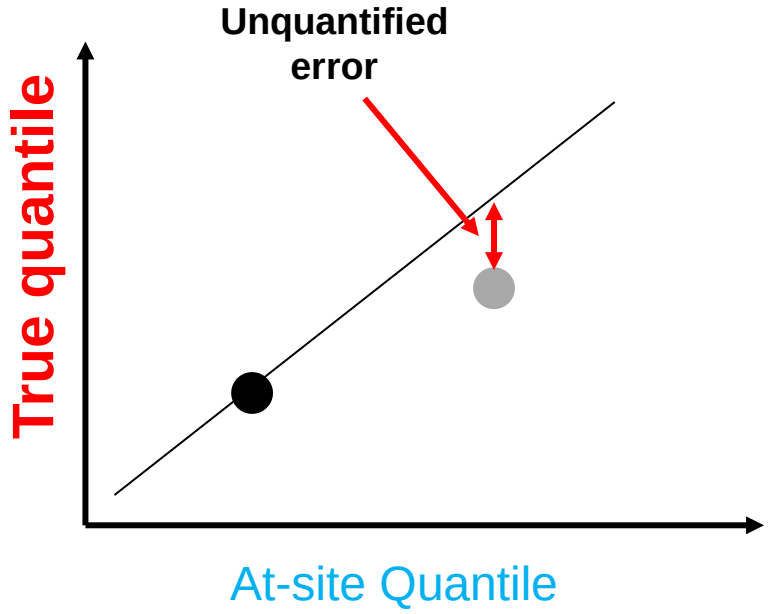


Proposed methodologies

- **Evaluation**



reference



Proposed methodologies

- **Evaluation**

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(q_L - \hat{q}_i^R\right)^2}$$

- **Proposed criterion: Mean Piecewise loss function**

$$MPLF\left(p\right) = \frac{10^3}{n}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\rho_p\left(y_{ij} - \hat{q}_{ip}^R\right) \quad ; \quad p \in (0,1)$$

$$\rho_p\left(u\right) = \begin{cases} u\left(p-1\right) & if\ u < 0 \\ up & if\ u \geq 0 \end{cases}$$

n : Total number of observations at all sites

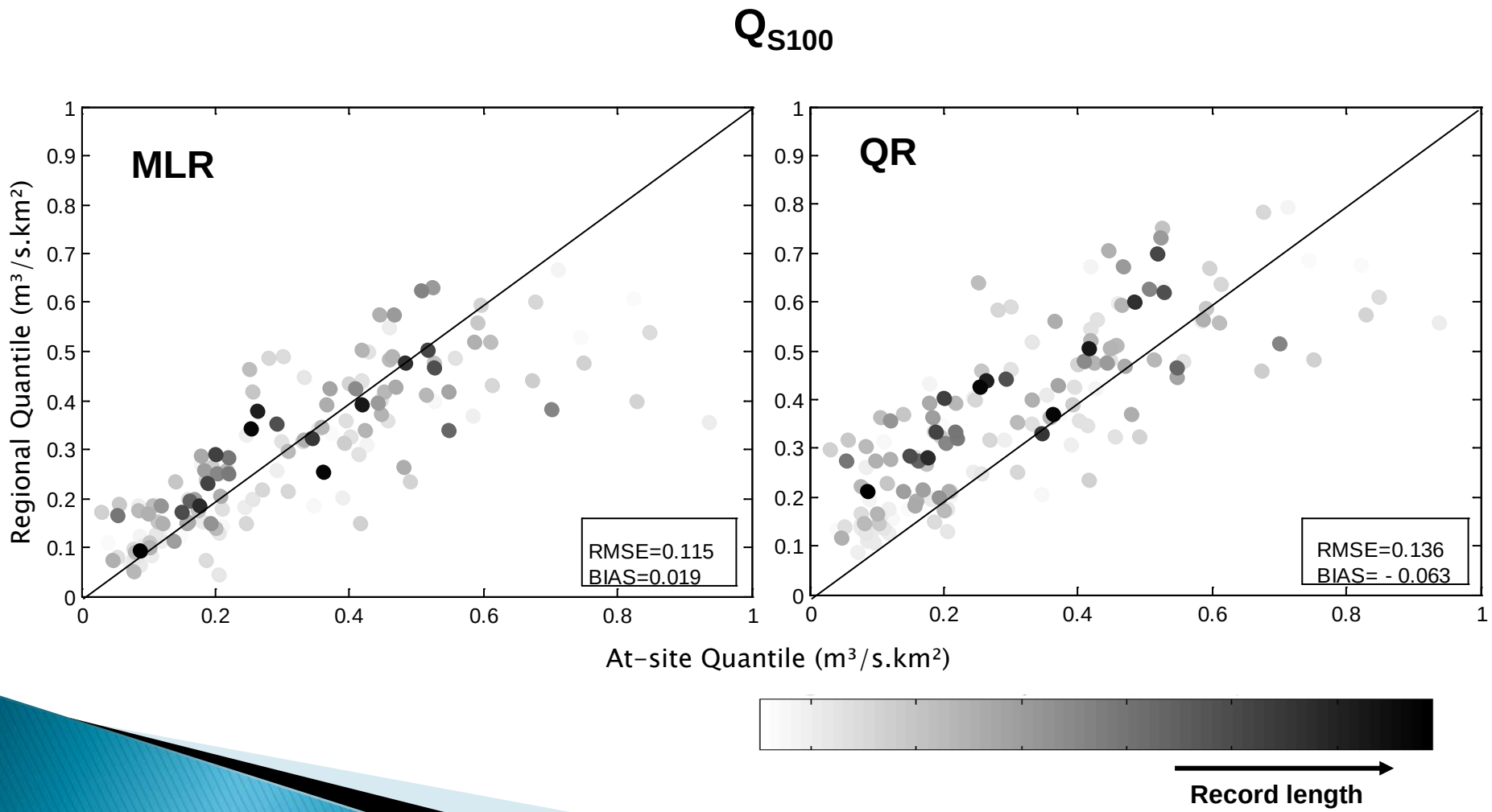
N : Number of sites

$\boldsymbol{n_i}$: year index at the site i

$\boldsymbol{\rho_p}$: Check function

Results

B. Exploitation of the hydrological information



Results

B. Exploitation of the hydrological information

	Q_{S10}		Q_{S50}		Q_{S100}	
	MLR	**QR**	**MLR**	**QR**	**MLR**	**QR**
MPLF ($m^3/s.km^2$)	16.07	**15.43**	6.62	**5.30**	4.65	**3.43**

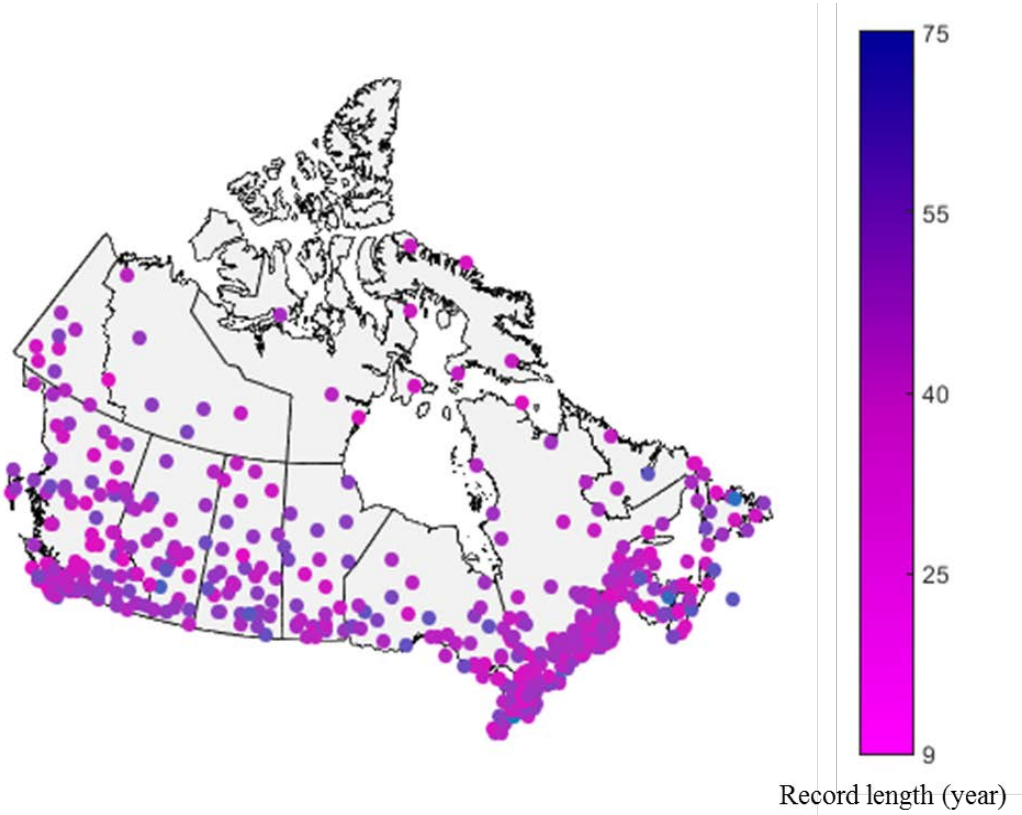
MPLF: Mean Piecewise loss function

Regional IDF curves

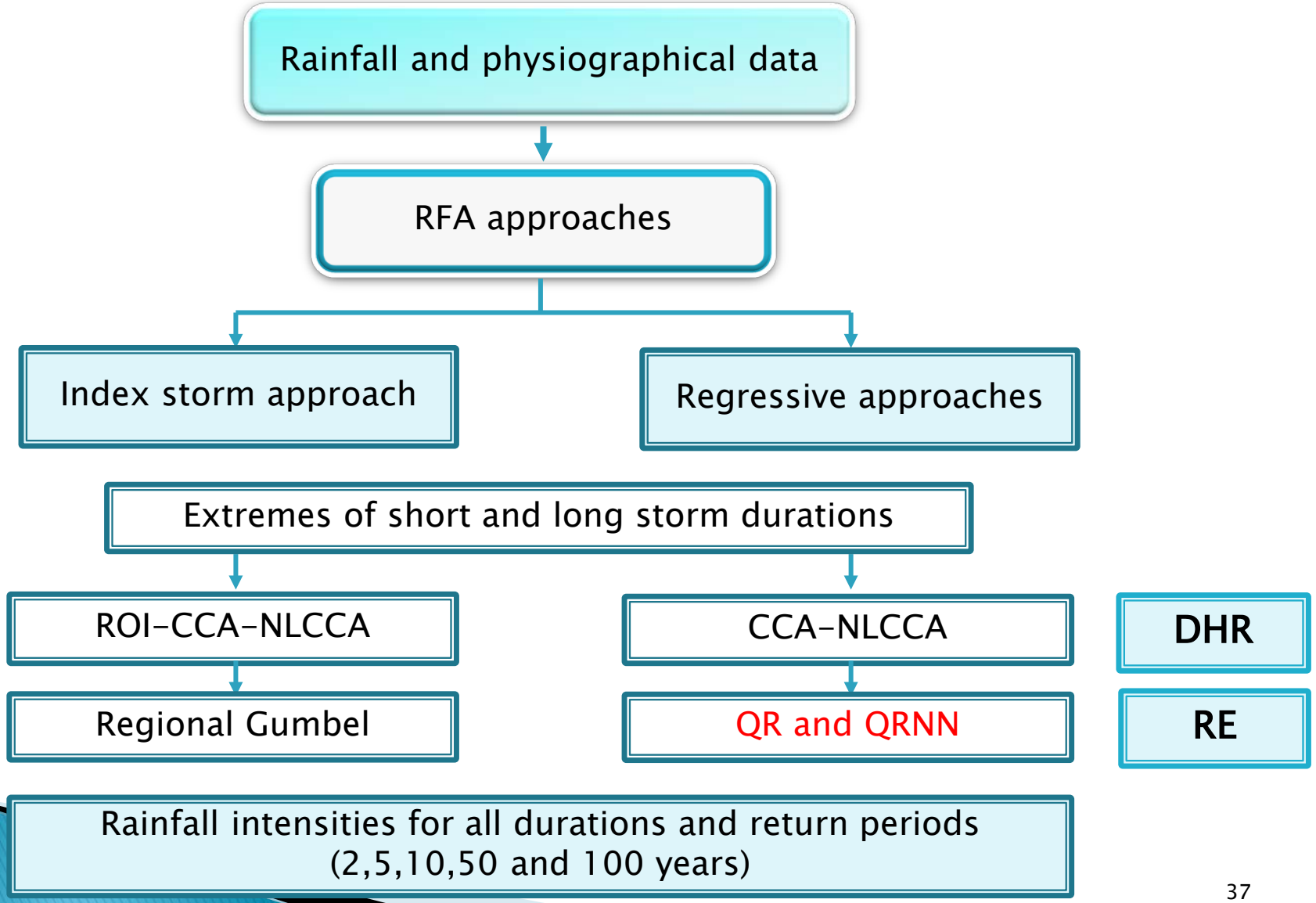
- **The occurrence frequency** of a rainfall event of given intensity and duration is important to establish **a measure of risk**.
- Estimating Intensity-Duration-Frequency Curves at ungauged sites

Data:

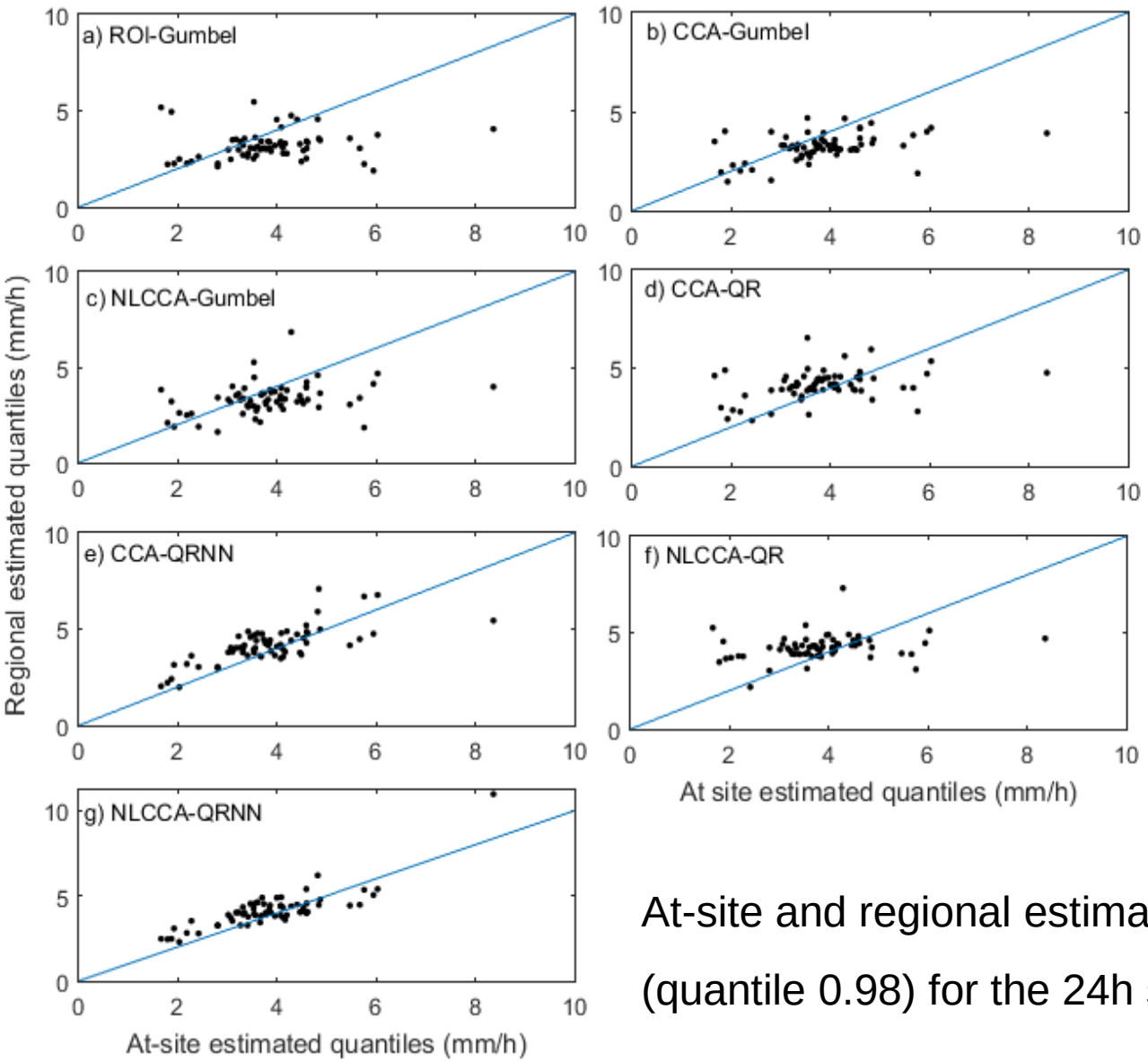
- Annual rainfall maxima (ECCC)
- 564 stations, Canada, 1905 - 2013
- Durations: 5, 10, 15, 30-min and 1, 2, 6, 12, 24-h
- Longitude, latitude, elevation, aspect, slope and surface roughness



Regional IDF curves

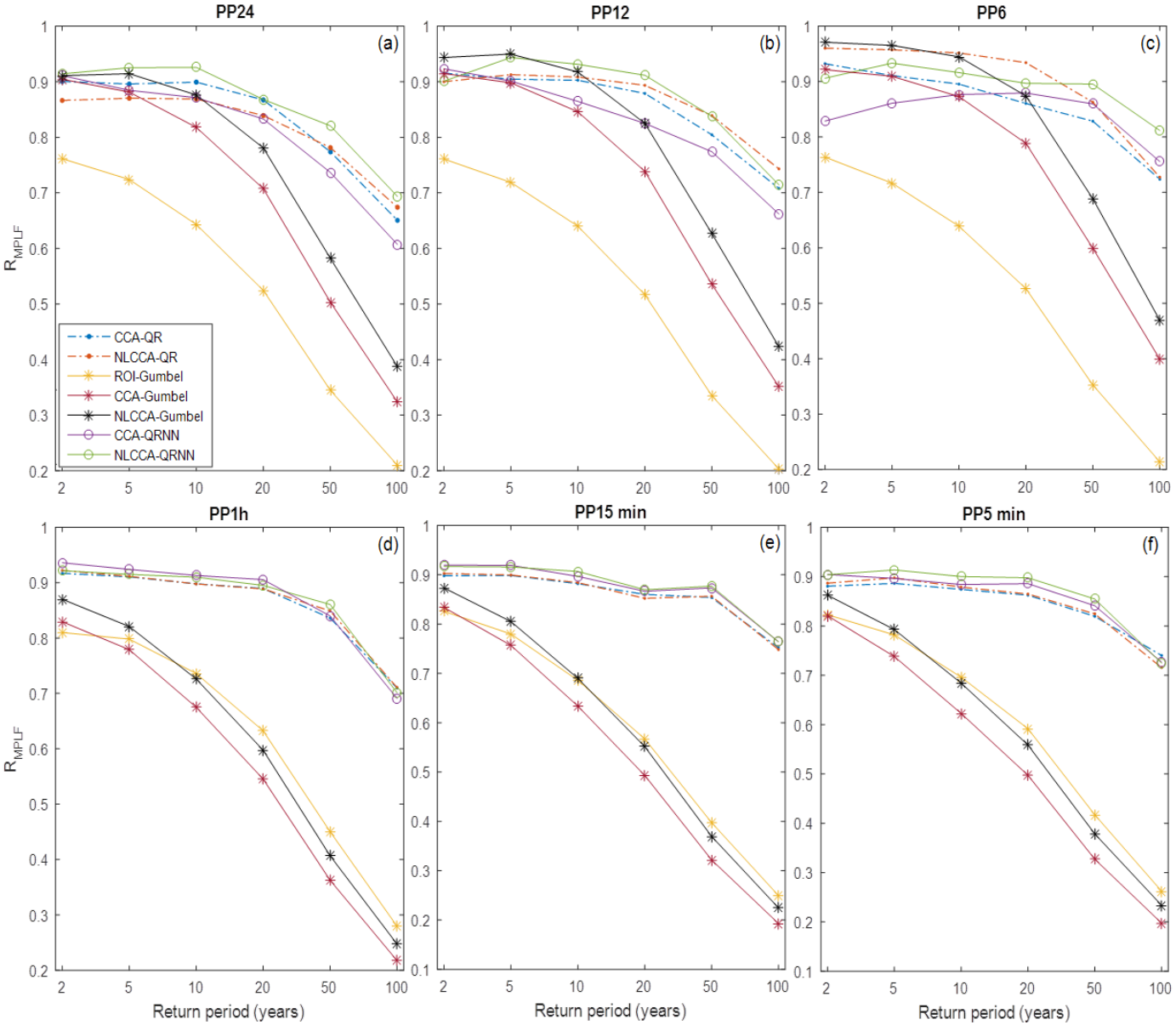


Results – Regional IDF curves



At-site and regional estimated 50-years return period (quantile 0.98) for the 24h storm duration

Results – Regional IDF curves



$$R_{MPLF} = \frac{MPLF_0}{MPLF_M}$$

Conclusions

- Considering Non-linearity in the two RFA steps leads to better results than linear cases;
- Considering a non-linear component at the first step (DHR) or at the second step (RE) of the RFA process leads to similar results;
- A direct RFA approach based on QR is a promising method for the estimation and evaluation of flood quantiles at-sites with short to medium record lengths;
- A QR-based approach (under linear and non linear versions) provides better estimates of IDF curves at ungauged sites as compared to the commonly used approach, the index storm.

Publications

1. D. Ouali, F. Chebana et T.B.M.J. Ouarda (2016a). "Non-linear canonical correlation analysis in regional frequency analysis". **Stoch Environ Res Risk Assess**. DOI 10.1007/s00477-015-1092-7.

2. D. Ouali, F. Chebana et T.B.M.J. Ouarda (2017). "Fully nonlinear regional hydrological frequency analysis". **Journal of Advances in Modeling Earth Systems**, 9(2), 1292-1306.

3. D. Ouali, F. Chebana et T.B.M.J. Ouarda (2016b). "Quantile regression in regional frequency analysis: a better exploitation of the available information". **Journal of Hydrometeorology**. DOI: 10.1175/JHM-D-15-0187.1

4. D. Ouali, A.J. Cannon. "Estimation of rainfall Intensity–Duration–Frequency curves at ungauged locations using quantile regression methods". Submitted.

Thank you …!