APCC Training Program on
"Generation of regional climate data derived from statistical downscaling techniques"

Hyung-Il Eum
APEC Climate Center (APCC)

APCC
APEC CLIMATE CENTER

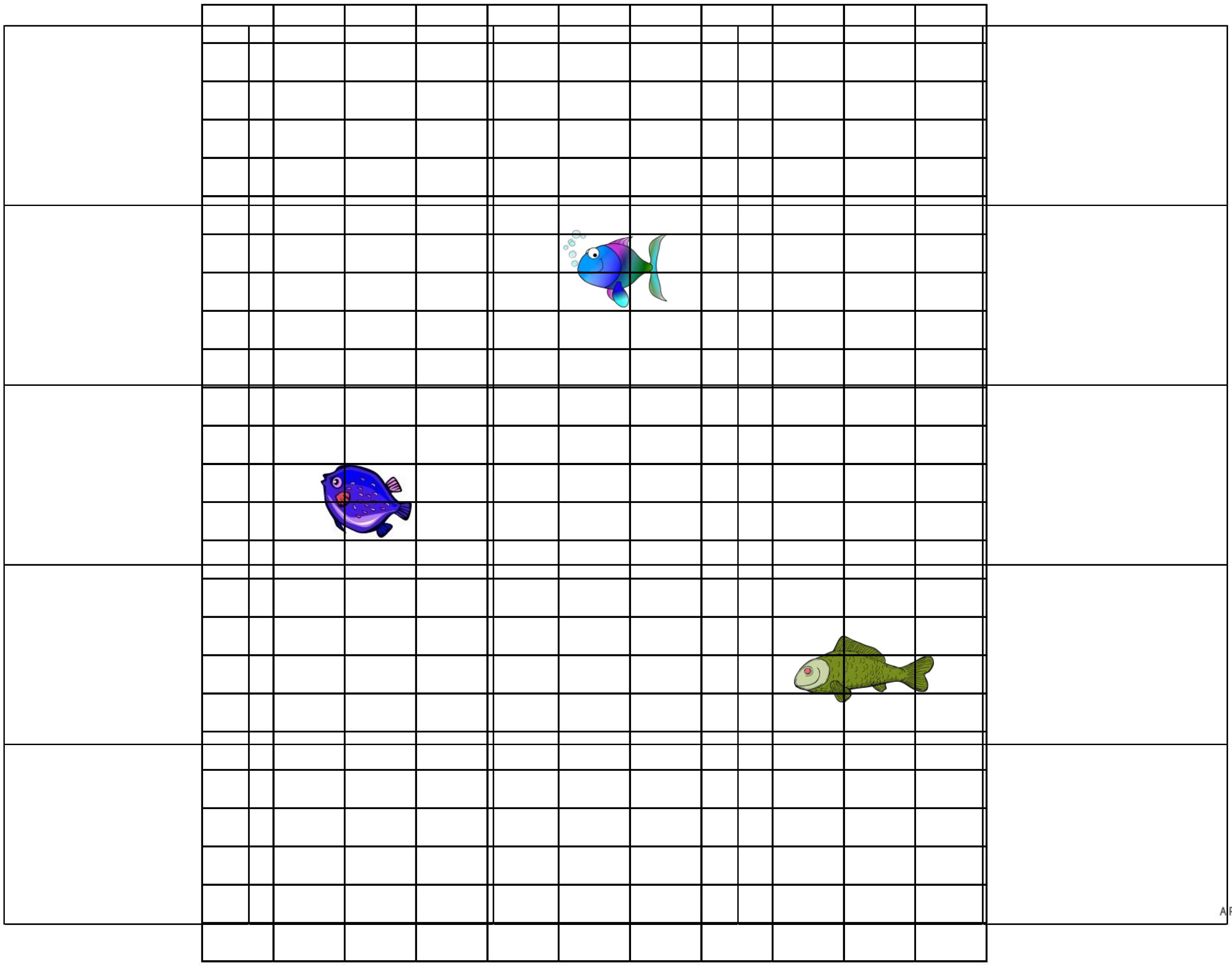
Statistical Downscaling Methods:

Hyung-Il Eum
APEC Climate Center (APCC)

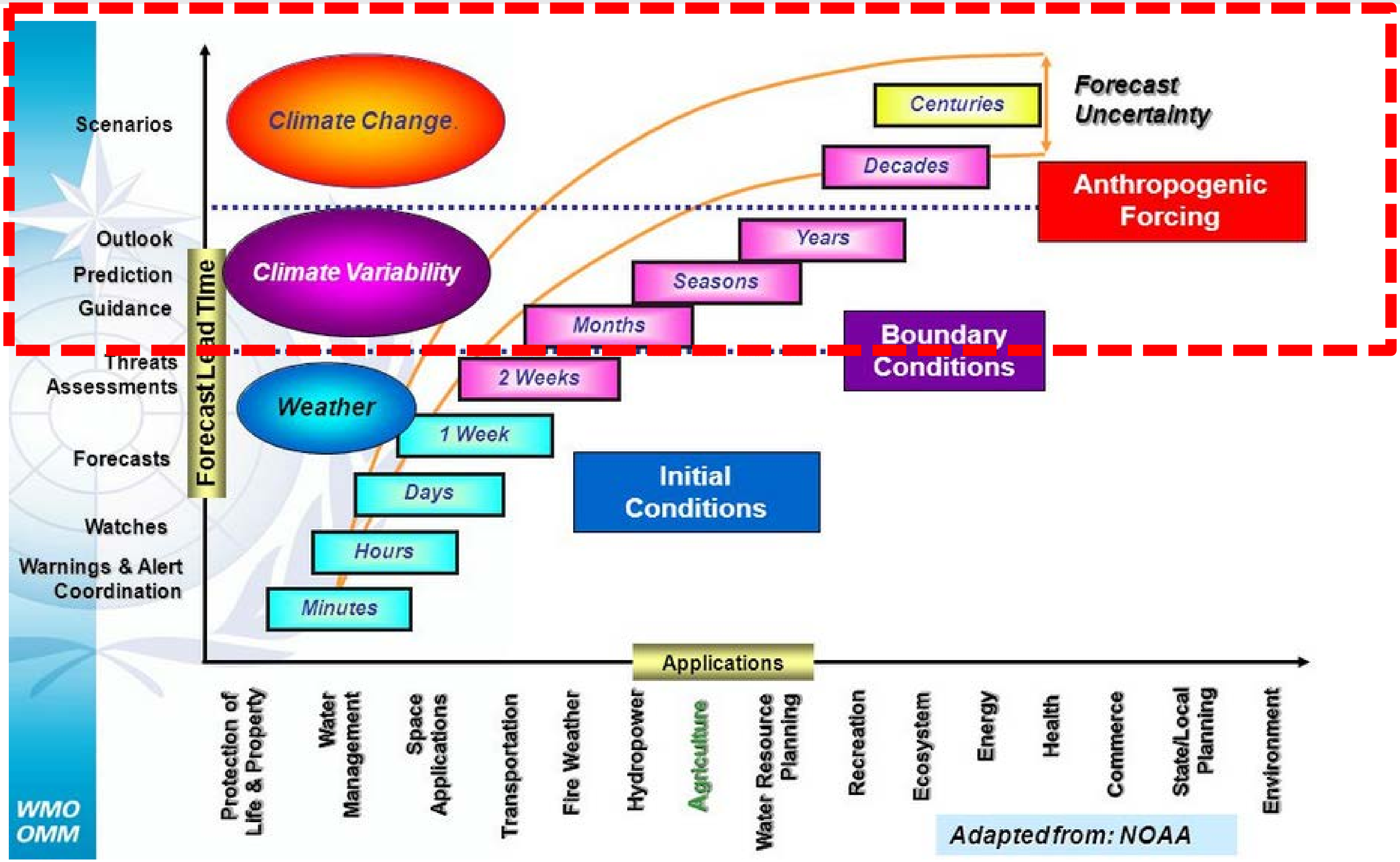
Contents

- **Why do we need downscaling?**
- **Dynamical VS Statistical downscaling**
  - Pros and cons
- **Statistical background**
  - Probability distribution
  - Quantile (Percentile)
- **Statistical downscaling methods**
  - Perfect Prognosis (PP)
  - Model Output Statistics (MOS)
  - Long-term trend preserving methods

Why?



Climate prediction framework

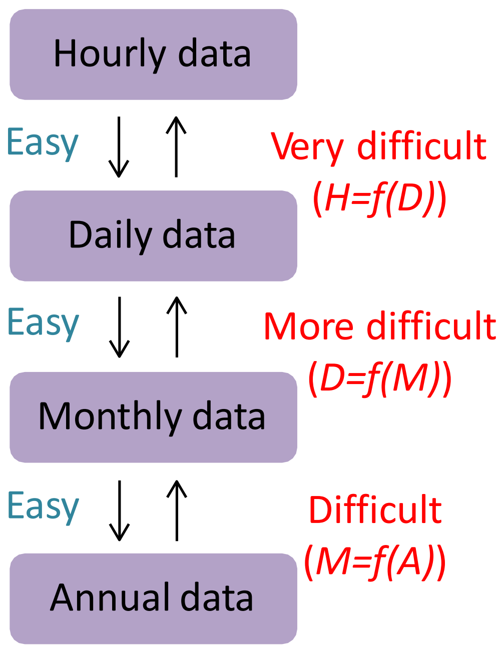


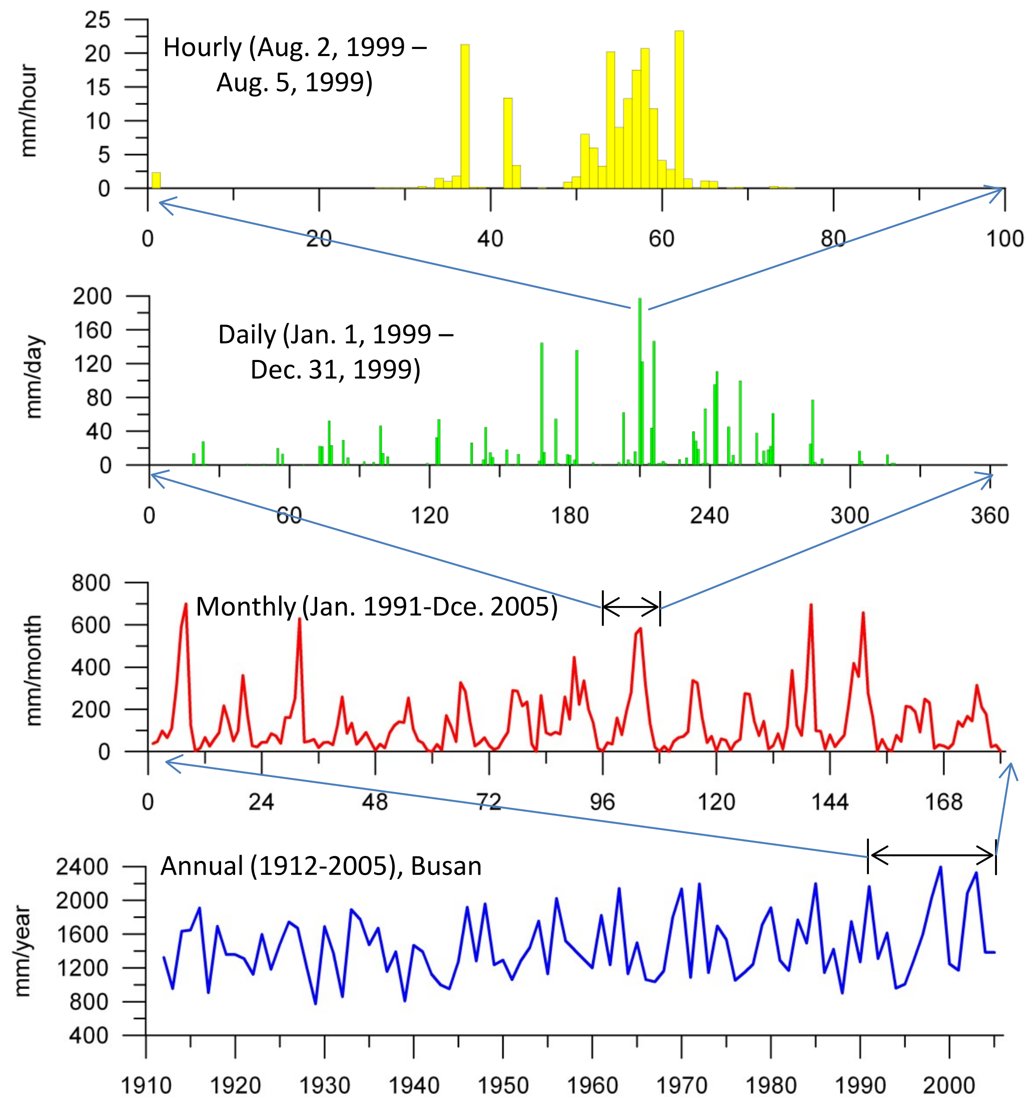
Temporal scale of GCMs

- **Temporal scales in climate forecasts**
  - Predictable signal of seasonal climate
    - Surface ocean and land condition on longer scales (monthly – seasonal)
  - Short time scales
    - Dominated by atmospheric “weather noise”
- **APCC Multiple Models Ensemble (MME)**
  - 6-month forecasts at monthly time step
- **Need of accurate sub-daily or daily climate information**
  - Agricultural or hydrologic applications

Temporal disaggregation

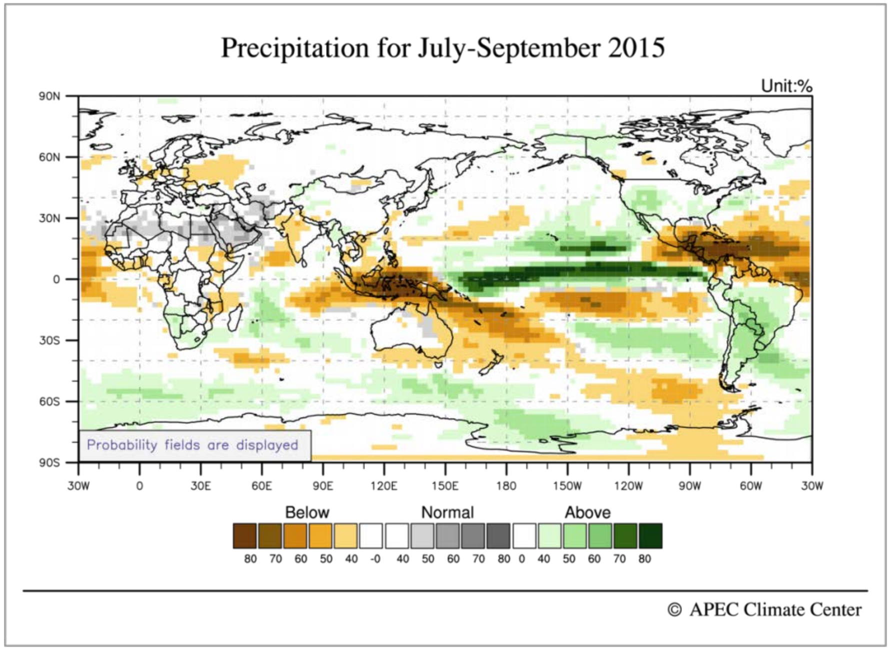
- **Temporal scale problem**

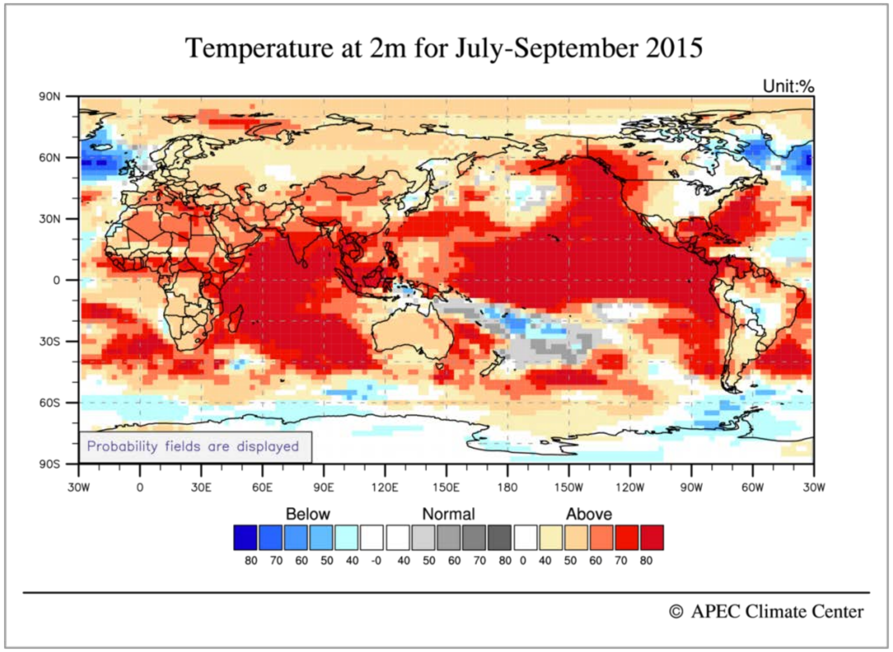




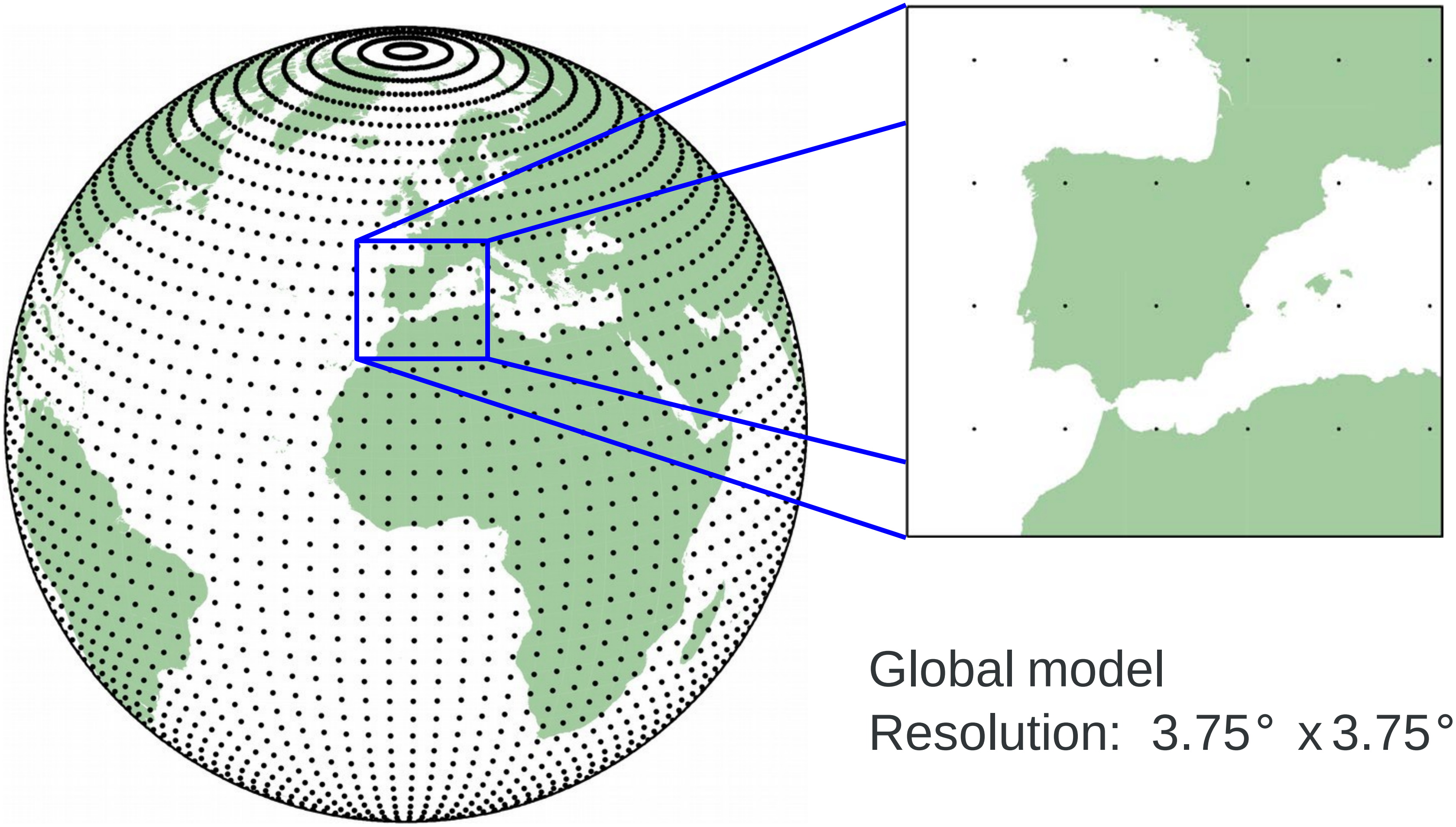
Spatial scale

- **Spatial scale in climate forecasts**
  - Horizontal resolution 100 to 300 km





A GCM resolution

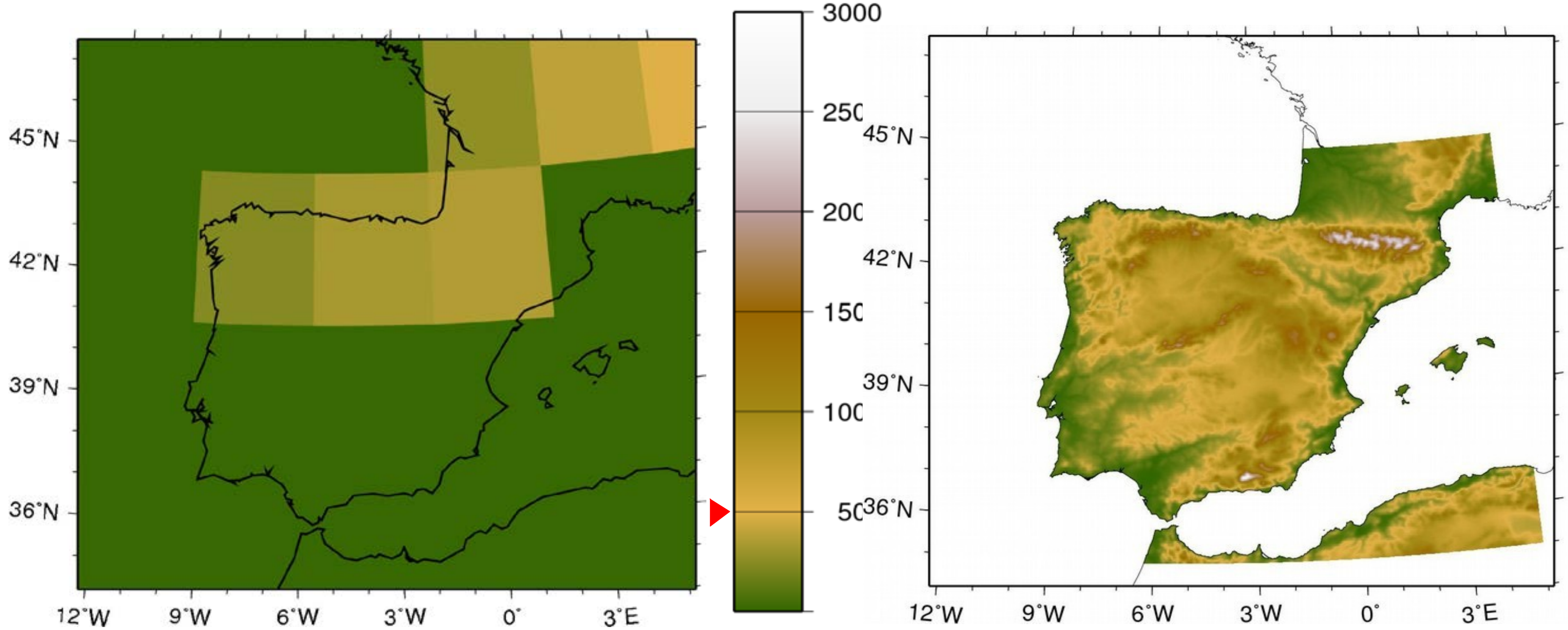


Global model
Resolution: 3.75° x 3.75°

Source: http://www.meteo.unican.es

APEC CLIMATE CENTER

Spatial resolution issue

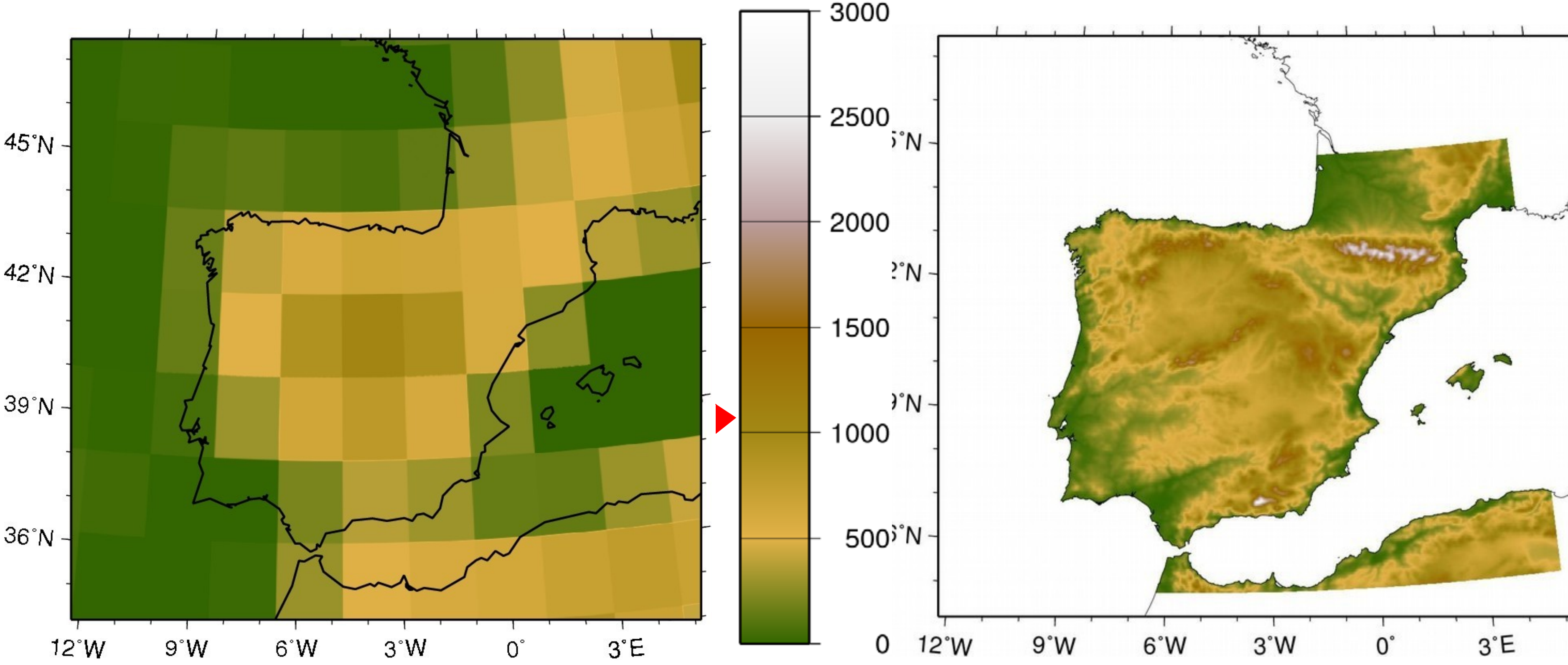


GCM at 3.75° | **Real world**

Source: http://www.meteo.unican.es

APEC CLIMATE CENTER

Spatial resolution issue

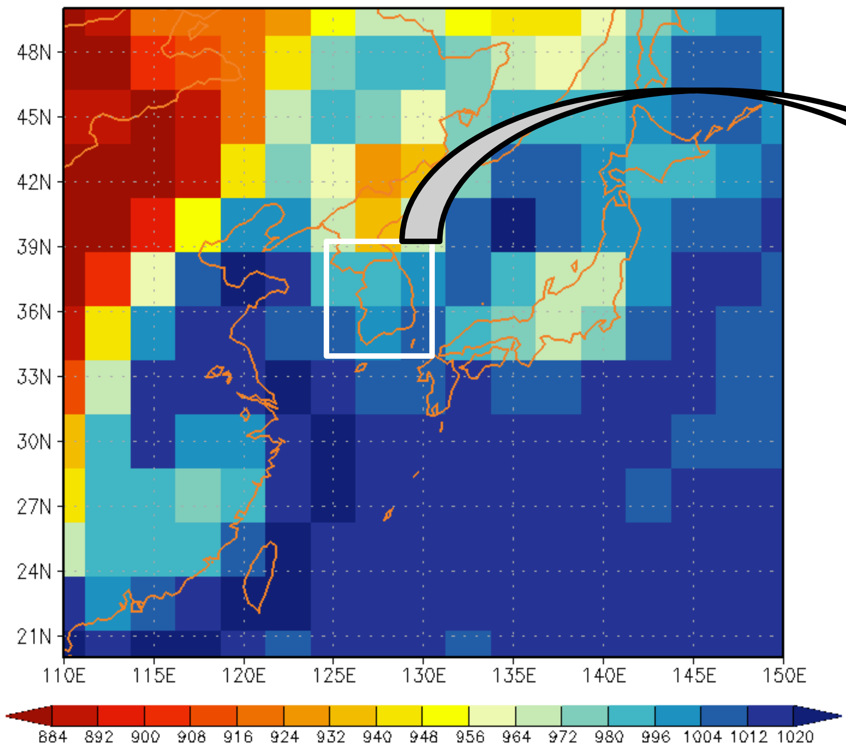


GCM at 2.5°

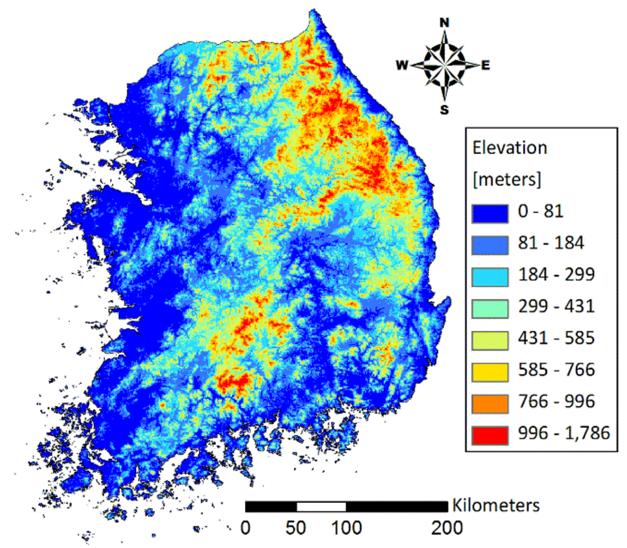
Real world

Source: http://www.meteo.unican.es

Spatial resolution issue

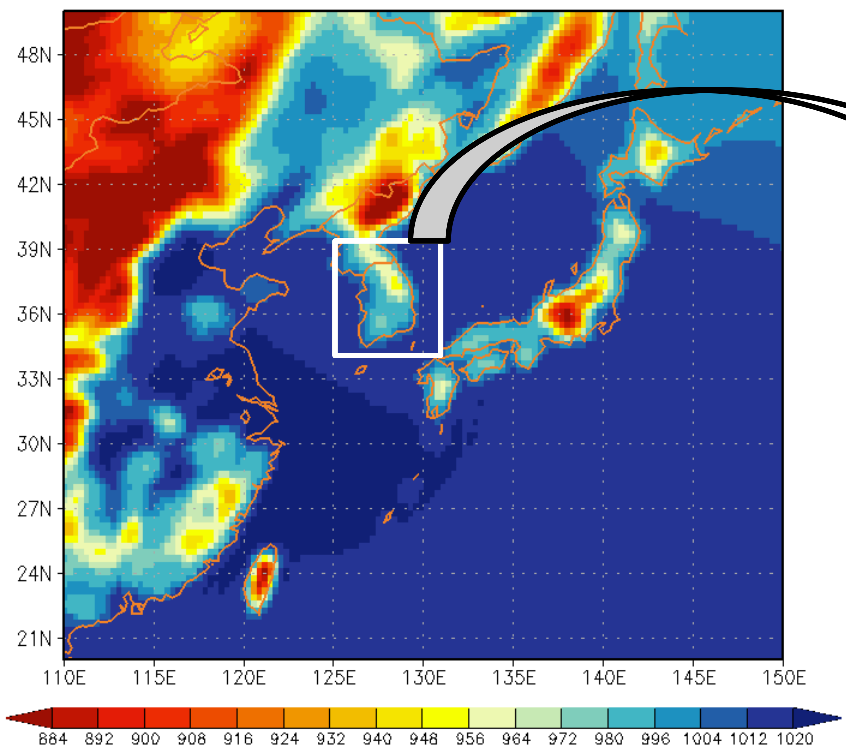


CFS at 2.5°

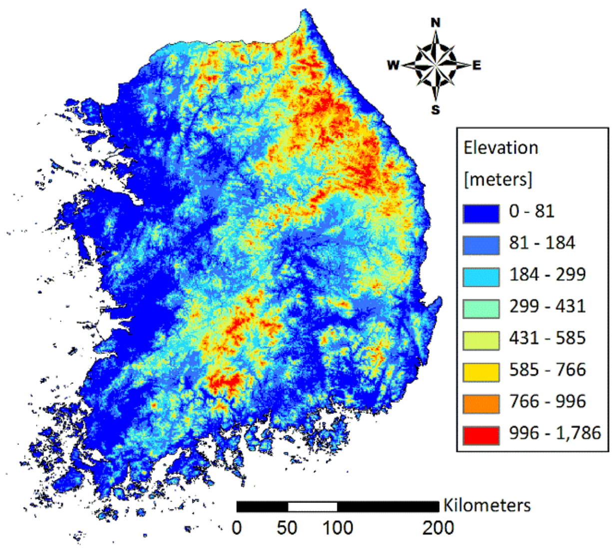


Real world

Spatial resolution issue



RCM (GRIMs) at 30 km



Real world

The gap between climate predictions and regional applications (1)

➤ Spatial scale problem

Observed precipitation (10 X 10km)

1	2	3	4	2	9	5	10
5	4	6	7	2	7	6	11
10	10	10	11	9	11	6	14
13	15	16	15	4	12	8	16
6	11	10	5	1	2	8	5
9	2	9	12	12	13	13	6
10	7	3	10	10	15	15	6
3	12	14	1	10	9	8	7

Perfect GCM (20 X 20km)

3	5	5	8
12	13	9	11
7	9	7	8
8	7	11	9

(1+2+5+4)/4 = 3

Perfect GCM (40 X 40km)

8.25	8.25
7.75	8.75

(1+2+5+4+3+4+6+7+10+10+10+11+13+15+16+15)/16 = 8.25

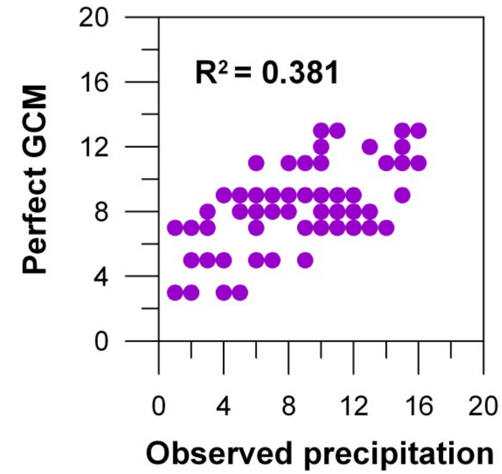
The gap between climate predictions and regional applications (2)

Perfect GCM (20 X 20km)

3	5	5	8
12	13	9	11
7	9	7	8
8	7	11	9

Spatial disaggregation

3	3	5	5	5	5	8	8
3	3	5	5	5	5	8	8
12	12	13	13	9	9	11	11
12	12	13	13	9	9	11	11
7	7	9	9	7	7	8	8
7	7	9	9	7	7	8	8
8	8	7	7	11	11	9	9
8	8	7	7	11	11	9	9

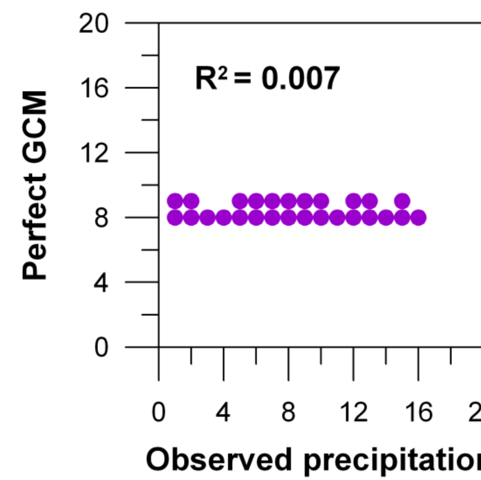


Perfect GCM (40 X 40km)

8	8
8	9

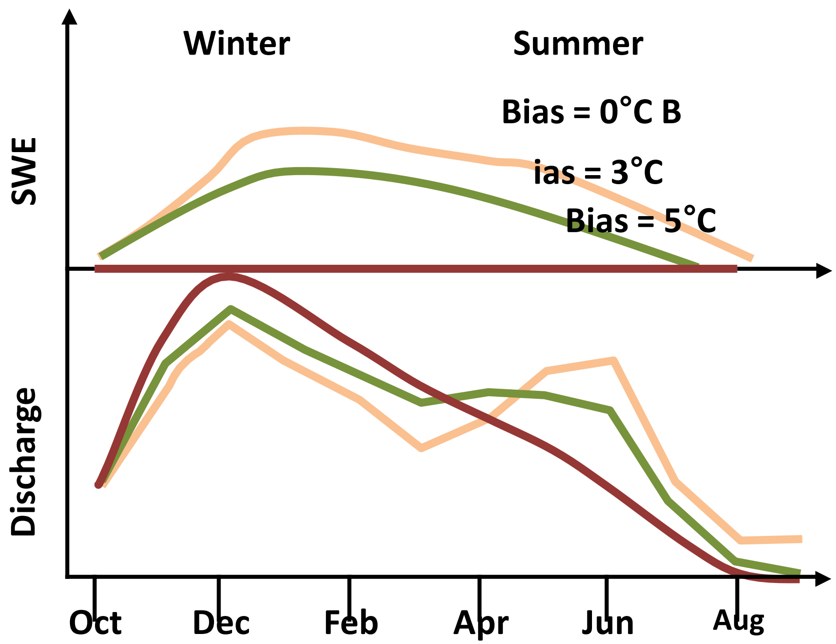
Spatial disaggregation

8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8
8	8	8	8	9	9	9	9
8	8	8	8	9	9	9	9
8	8	8	8	9	9	9	9
8	8	8	8	9	9	9	9



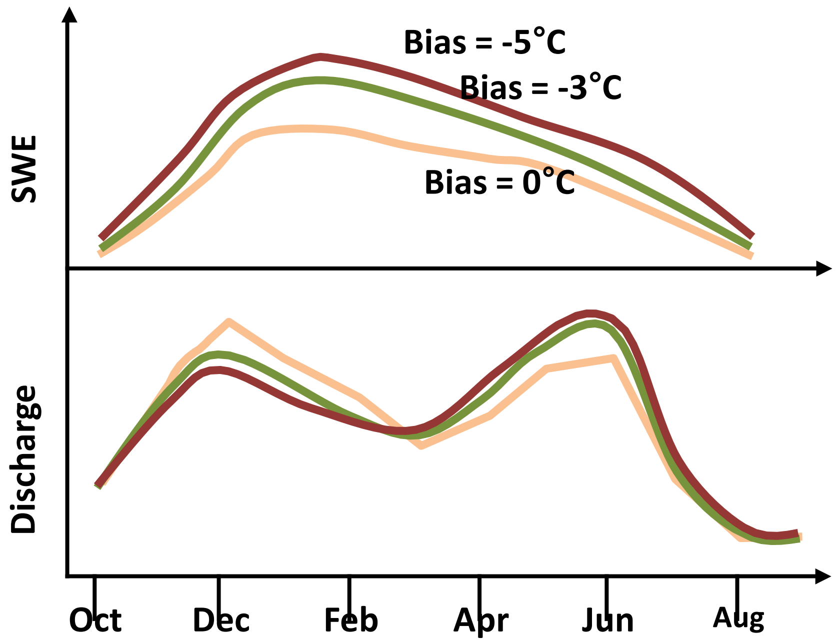
Potential impacts of systematic bias (1)

➢ Warm bias in winter temperature



Potential impacts of systematic bias (2)

➢ Cold bias in winter temperature



copyright@Jung (2013)

APEC CLIMATE CENTER

Numerical weather prediction models

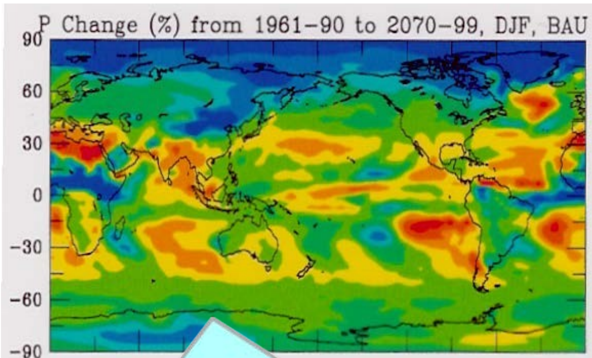
- **Physical parameterization**
  - Theory
  - Simplified assumptions for unknown variables

Difference in climate models

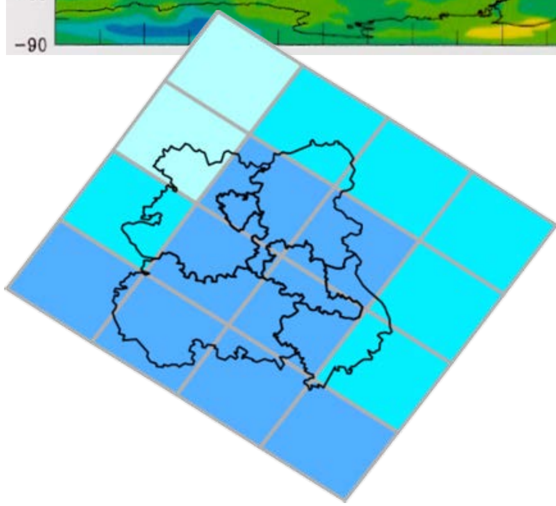
GCMs limitations

Climate models

300km

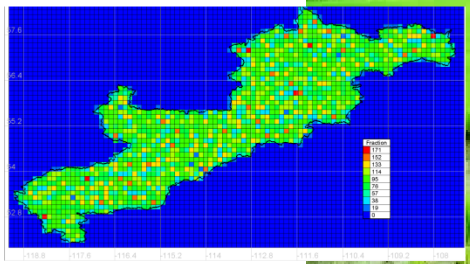


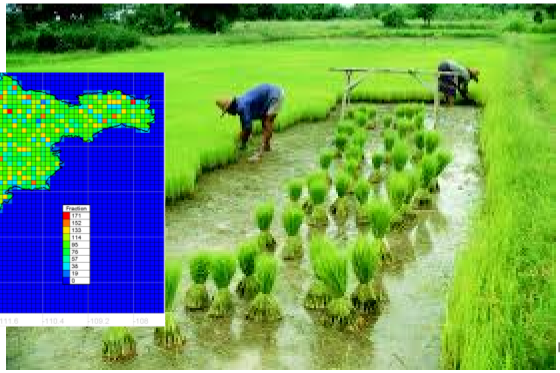
50km



Regional applications require...

~5 - 10 km

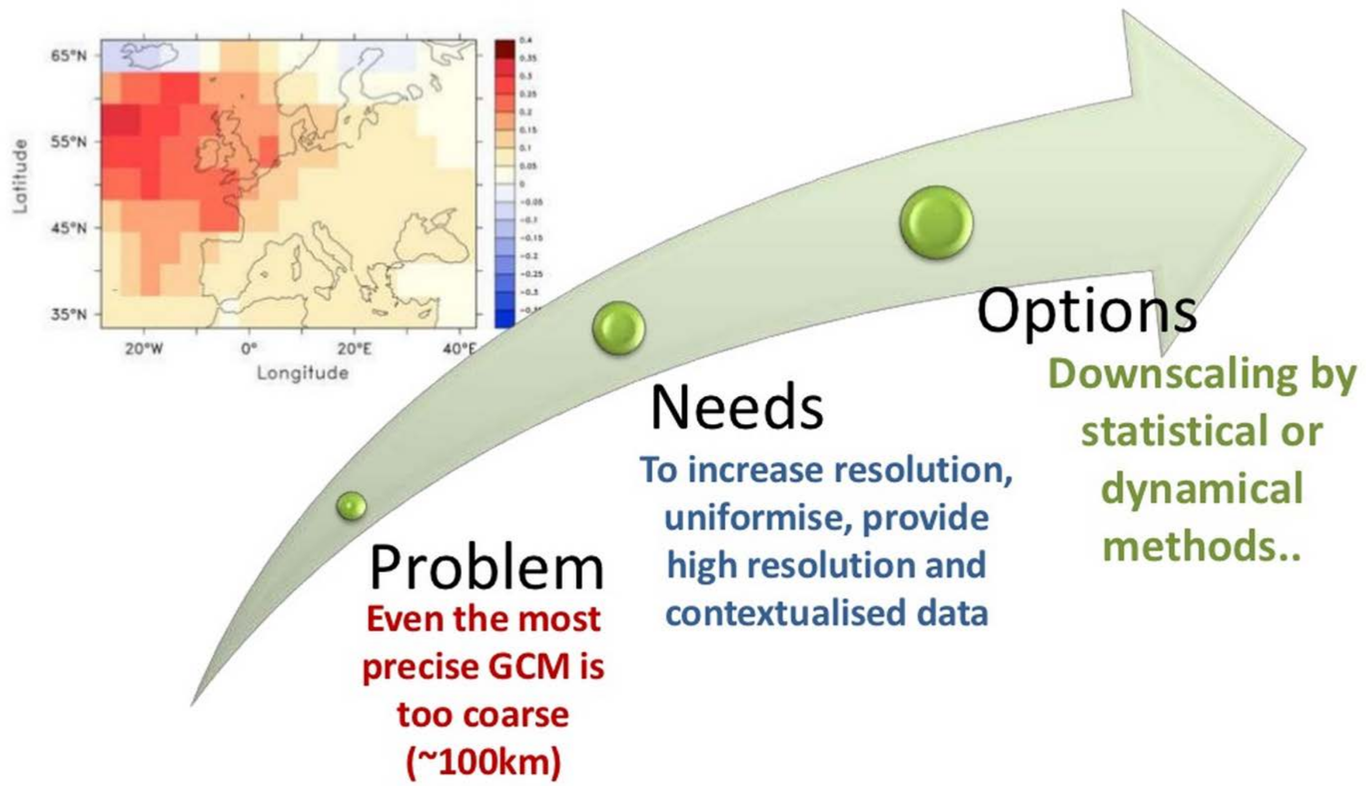




~ 1 km

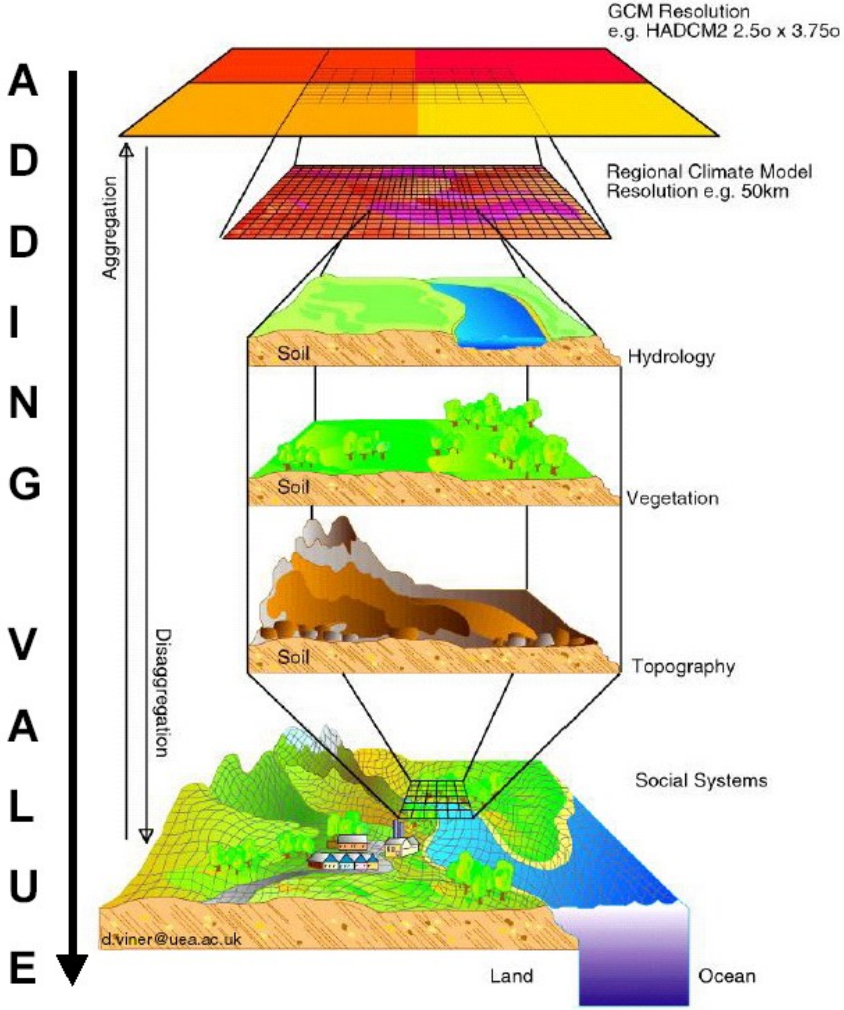
Needs & Options to use GCM's information

➢ **Downscaling**



Source: Navarro et al. (2012)

Downscaling



Downscaling techniques try to adapt the **coarse global model output to the local features** of a given region.

Source: ORASECOM 2010

Statistical Background

APCC
APEC CLIMATE CENTER

Sources of Uncertainty

- **Randomness (uncontrollable)**
  - Inherent unexplainable variability of nature
- **Lack of information/understanding**
  - Parameter uncertainty
  - Modeling uncertainty
  - Sampling uncertainty (≠ data uncertainty)
- **Error/inaccuracy**
  - Data uncertainty
  - Operational uncertainty

Statistics and Probability

- **Statistics VS Probability**
- **Statistics**
  - Methods for drawing inferences about the properties of a population based on the properties of a sample from that population
- **Probability**
  - Methods for calculating the likelihood of an event given known population characteristics

Random Variable and Random Events

- **Random variable**
  - A mathematical vehicle for representing an event in an analytical form
  - Equally likely
- **Random events**
  - Mapped into the real line through the random variable X

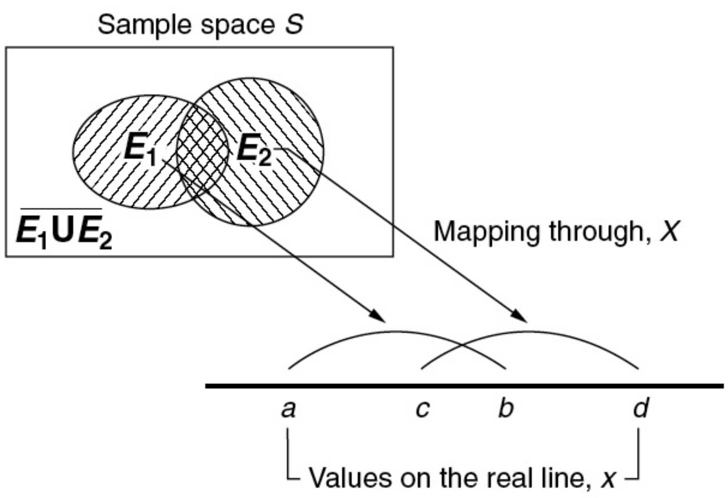
Random Variable and Random Events

$$E_1 = (a < X \leq b)$$

$$E_2 = (c < X \leq d)$$

$$E_1 E_2 = (c < X \leq b)$$

$$\overline{E_1 \cup E_2} = (X \leq a) + (X > d)$$



Population VS Sample

- **Population**
  - All characteristic values from all elements
- **Sample**
  - A part of population
  - A source of systematic error in statistics
- $\boldsymbol{\mu}$ **vs** $\bar{X}$
- $\boldsymbol{\sigma}$ **vs** $\bar{S}$

Probability

➢ Statistics VS Probability

- From population ?

➢ Probability VS Likelihood

- Sum -----> 1 ??

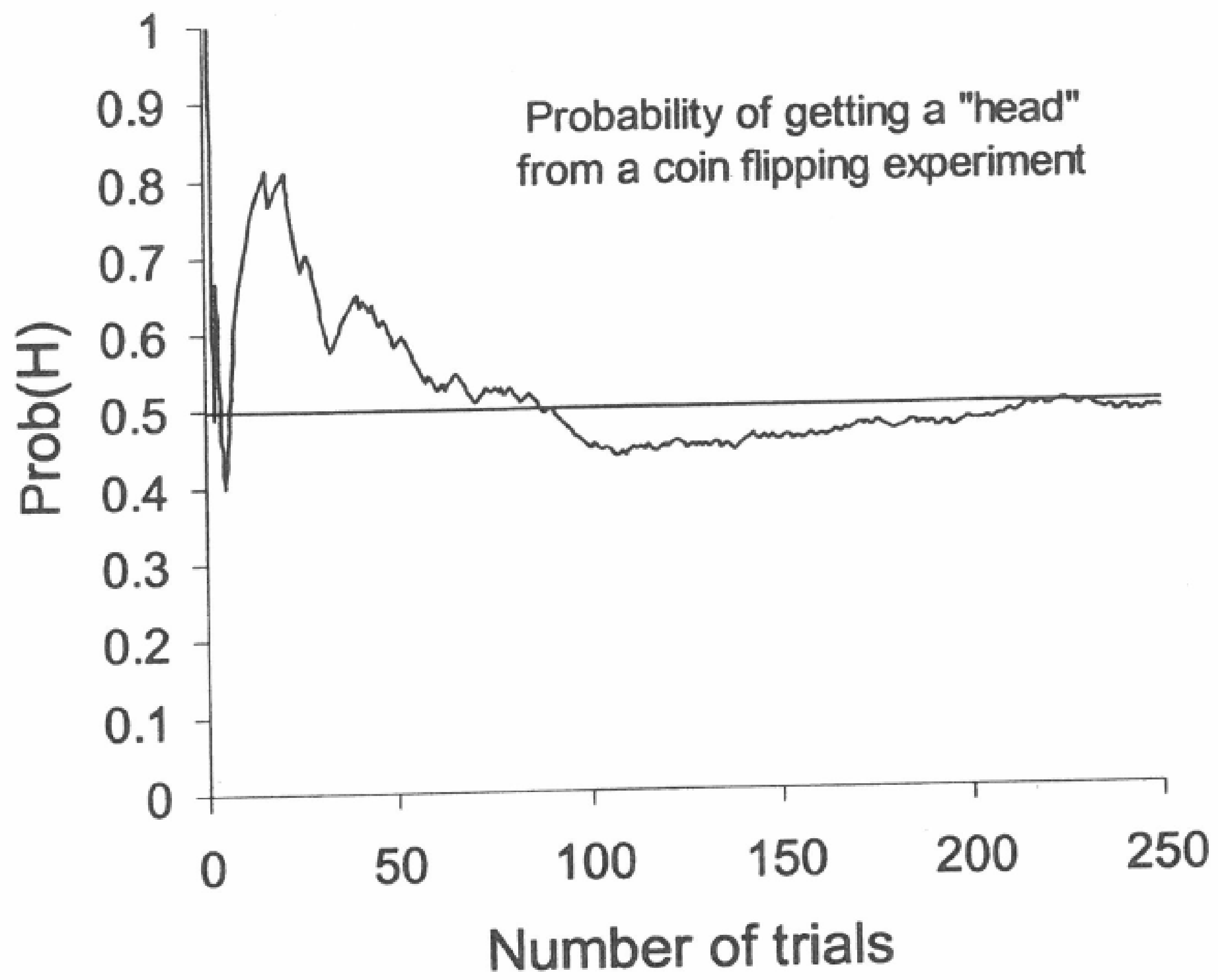
➢ Probability

- Prob(A)= n_a/n
- Prob(A)= $\lim_{n \to \infty} \frac{n_a}{n}$

Probability

- **A measure of how likely an event will occur**
- **A number expressing the ratio of favorable outcome to the all possible outcomes**
- **Probability is usually represented as P(.)**
  - P (getting a club from a deck of playing cards) = 13/52 = 0.25 = 25 %
  - P (getting a 3 after rolling a dice) = 1/6

Coin flip



Probability of getting a "head"
from a coin flipping experiment
Prob(H)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
50
100
150
200
250
Number of trials

Axioms of Probability

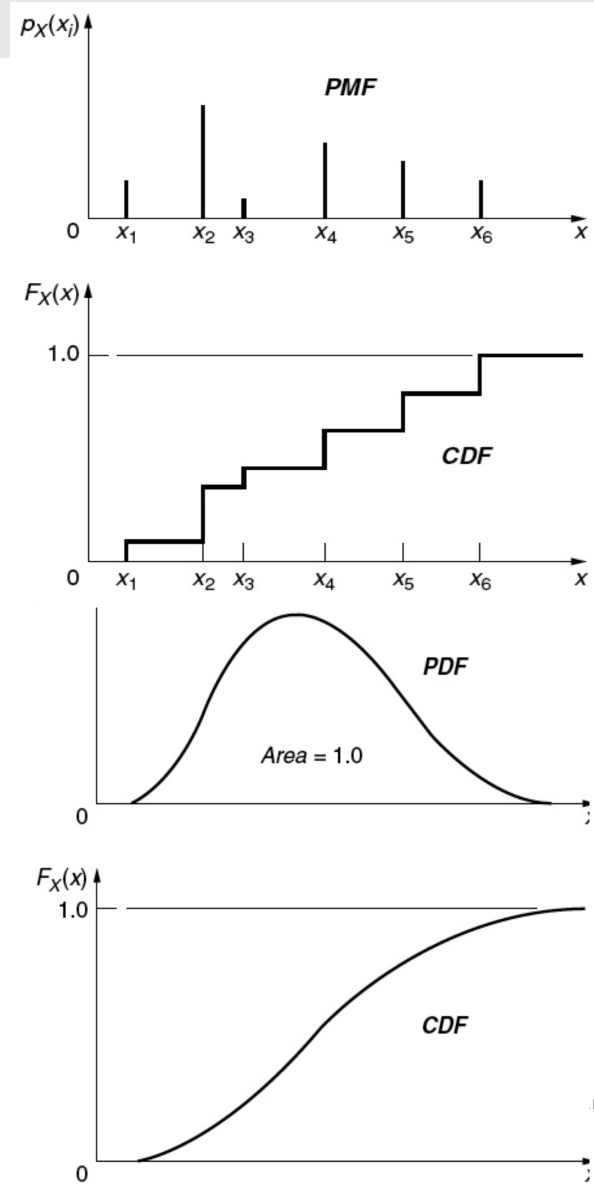
- **P(S) =1**
- **0 ≤ P(A) ≤ 1**
- **P(AUB) = P(A) + P(B)**
  - if A & B are mutually exclusive
  - Independent VS mutually exclusive
- **P(AUB) = P(A) + P(B)-P(A ∩B)**
- **P(AUBUC) = ?**

Probability Mass/Density Functions (PMF/PDF)

- Random variable X
  - $x_1, x_2, \ldots, x_n$

- $f_X(x_i) = F_X(x_i) - F_X(x_{i-1})$

- $dP_X(x) = p_X(x)dx$

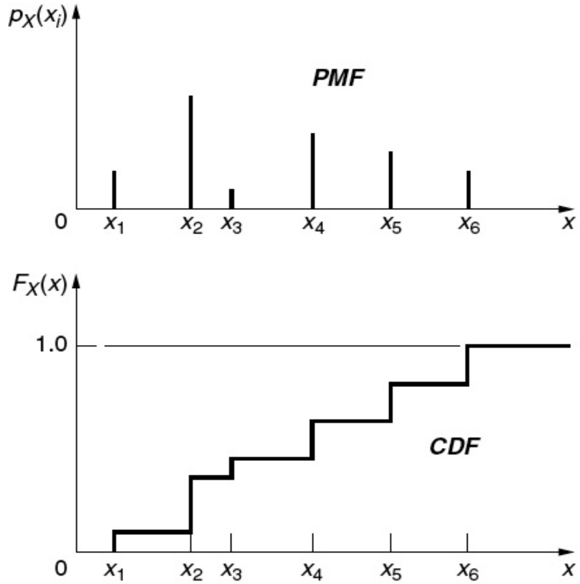


IMATE CENTER

Cumulative Mass/Density Functions (CMF/CDF)

- $F_X(x) = \sum_{x_i \le x} f_X(x_i)$

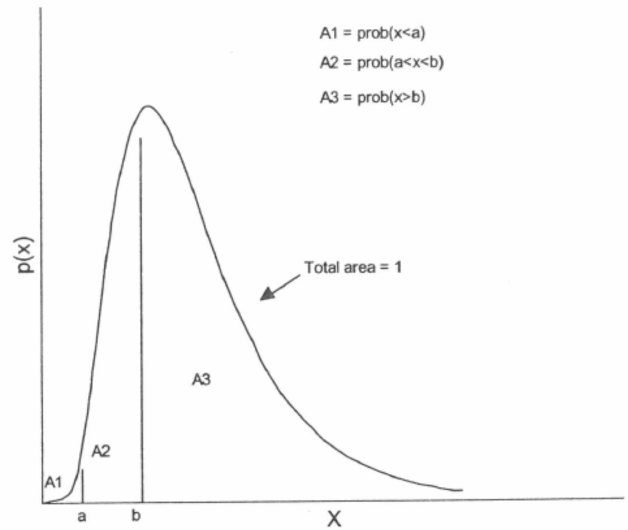
- $P_X(x) = Pr(X \le x) = \int_{-\infty}^{x} p_X(t)dt$

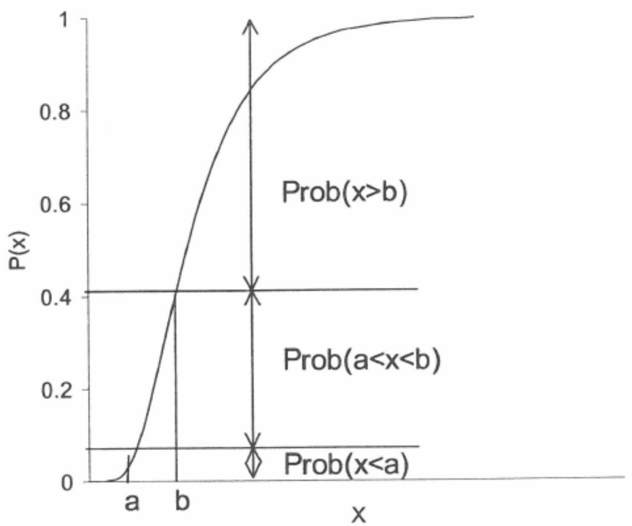


Probability Concept for Continuous R.V.

- $Pr(a \le x \le b) = \int_a^b p_X(t)dt = P_X(b) - P_X(a)$

- **Pr(x=d)=?**



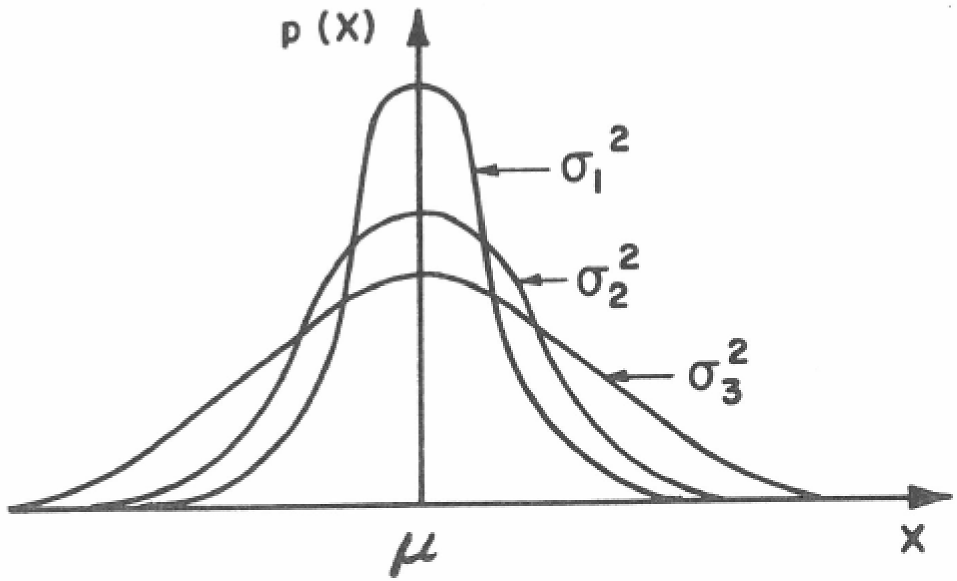


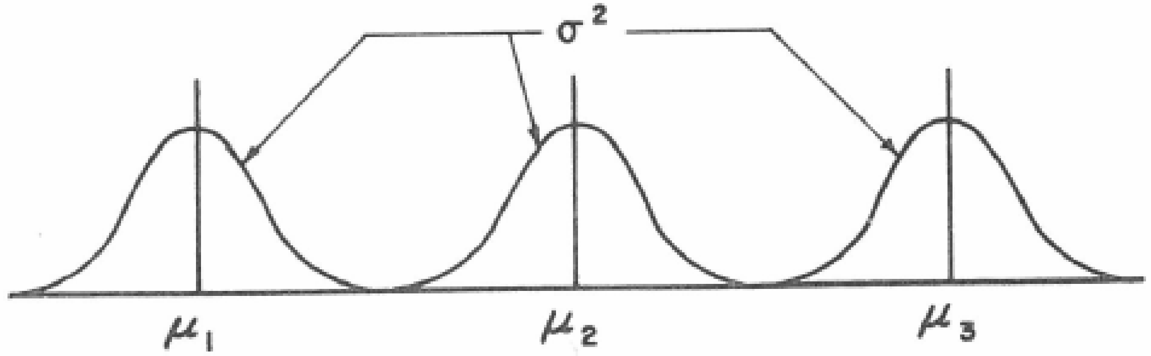
Gaussian or normal distribution

- **Most widely used**
- **2 parameter distribution**
- $p_X(x) = \frac{1}{\sqrt{2\pi\theta_2^2}} e^{-\frac{1}{2}\left(\frac{x-\theta_1}{\theta_2}\right)^2} \quad -\infty < X < \infty$
- $\theta_1 = \mu,\ {\theta_2}^2 = \sigma^2$

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad -\infty < x < \infty$$

- $N(\mu, \sigma^2)$


p (X)
σ1 2
σ2 2
σ3 2
X
μ


σ 2
μ 1
μ 2
μ 3

68-95-99.7 law

- **$N(\mu, \sigma^2)$**
  - Cover 68 % of all data between $\mu - \sigma$ and $\mu + \sigma$
  - Cover 95 % of all data between $\mu - 2\sigma$ and $\mu + 2\sigma$
  - Cover 99.7 % of all data between $\mu - 3\sigma$ and $\mu + 3\sigma$

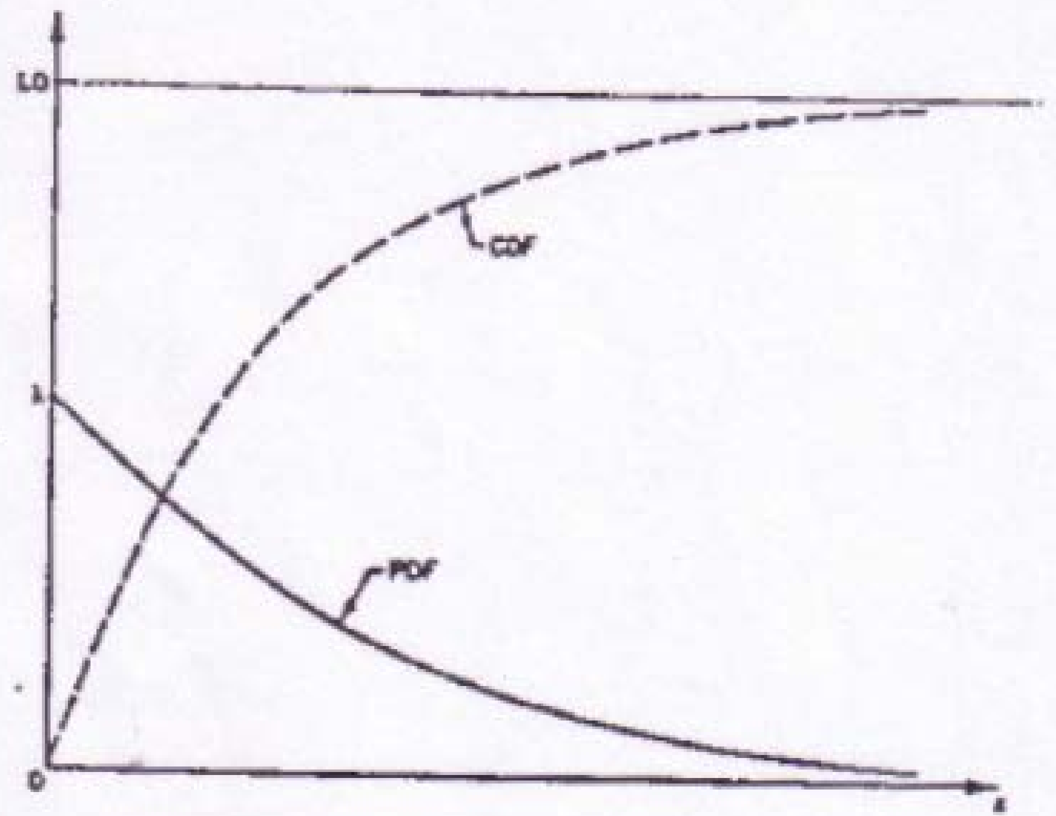
Exponential Distribution

- **Probability distribution function**
  - $p_X(x) = \lambda e^{-\lambda x} = f_T(t:\lambda)$ for x>0
- **Population parameters**
  - Mean=E[X]=1/ λ
  - Variance=Var[X]=1/ λ^2
  - Skewness=2
- **Parameter estimates**
  - $\lambda = \frac{1}{\bar{x}}$
- **Special case of the gamma distribution ($\eta = 1$)**

Example

- **Find the maximum likelihood estimator for the parameter λ of the distribution $p_x(x)= \lambda e^{-\lambda x}$ for X>0**
  - L($\boldsymbol{\lambda}$)=$\prod_{i=1}^{n} \boldsymbol{\lambda e^{-\lambda x}} = \lambda^n e^{-\lambda \sum x_i}$
  - l($\boldsymbol{\lambda}$)=lnL($\boldsymbol{\lambda}$)=nln($\boldsymbol{\lambda}$)- $\lambda \sum x_i$
  - $\frac{\partial l(\lambda)}{\partial \lambda} = \frac{n}{\lambda}$- $\sum x_i$=0
  - $\hat{\lambda} = \frac{n}{\sum x_i} = \frac{1}{\bar{x}}$

Exponential Distribution



PDF and CDF of the exponential distribution

Gamma distribution (1)

- **Probability distribution of the time to the nth occurrence, which is the sum of n independent r.v. $T_1+T_2+\ldots+T_n$ from the exponential distribution**
  - Waiting time of nth occurrence
- **Probability distribution function**
  - $p_x(x) = \dfrac{\lambda^\eta x^{\eta-1} e^{-\lambda x}}{\Gamma(\eta)} \quad X, \eta, \lambda > 0$

$\Gamma(\eta)$ is the gamma function having the properties

$$\left.\begin{array}{ll} \Gamma(\eta) = (\eta - 1)! & \text{For } \eta = 1, 2, 3, \ldots \\ \Gamma(\eta + 1) = \eta\Gamma(\eta) & \text{For } \eta > 0 \\ \Gamma(\eta) = \int_0^\infty t^{\eta-1} e^{-t}\, dt & \text{For } \eta > 0 \\ \Gamma(1) = \Gamma(2) = 1; \Gamma(1/2) = \sqrt{\pi} & \end{array}\right\}$$

Gamma distribution (2)

$$E(X) = \frac{\eta}{\lambda}$$

$$Var(X) = \frac{\eta}{\lambda^2}$$

$$\gamma_X = \frac{2}{\sqrt{\eta}}$$

- **E[x]=** $\int_0^\infty x \frac{\lambda^\eta}{\Gamma(\eta)} x^{\eta-1} e^{-\lambda x}\, \boldsymbol{dx} = \frac{\lambda^\eta}{\Gamma(\eta)} \int_0^\infty x^\eta e^{-\lambda x}\, dx$

$$= \frac{\lambda^\eta}{\Gamma(\eta)} \frac{\Gamma(\eta+1)}{\lambda^{\eta+1}} \int_0^\infty \frac{\lambda^{\eta+1}}{\Gamma(\eta+1)} x^\eta e^{-\lambda x}\, dx = \frac{\eta}{\lambda}$$

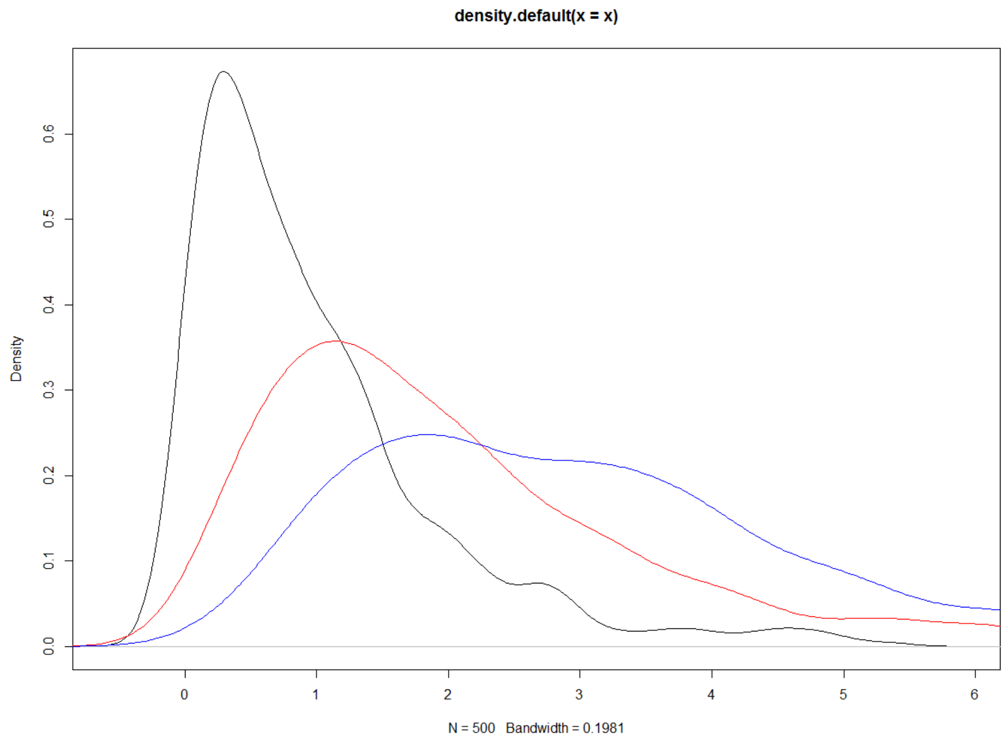
- **Parameter estimates**

$$\hat{\lambda} = \frac{\bar{x}}{s^2}$$

$$\hat{\eta} = \frac{\bar{x}^2}{s^2}$$

Gamma distribution (3)

```
> set.seed(600)
> x <- rgamma(500,shape=1,scale=1)
> plot(density(x))
> x <- rgamma(500,shape=2,scale=1)
> lines(density(x),col="red")
> x <- rgamma(500,shape=3,scale=1)
> lines(density(x),col="blue")
>
```



Extreme Value (EV) Family

- **The extreme value of a set of r.v. is also random**
- **The pdf of extreme value r.v. in general depends on the sample size and the parent distribution**
  - Y: largest from X
  - $P_Y(y) = \mathrm{Prob}(\mathrm{Y} \leq y) = \mathrm{Prob}(all\ X \leq y)$
  - If x ~ iid
    - Prob($X_1 \leq y$)* Prob($X_2 \leq y$)*…*Prob($X_n \leq y$)=$[P_X(y)]^n$

EV distributions

Type	Extreme Value	Parent Distribution
Type I	largest	normal, lognormal, exponential, gamma
	smallest	normal
Type II	largest or smallest	Cauchy
Type III	largest	beta
	smallest	beta, lognormal, exponential, gamma

Extreme Value Type I (Gumbel)

- **PDF**
  - $f_X(x) = \frac{1}{\alpha}\exp\left[-\frac{x-\xi}{\alpha} - \exp\left(-\frac{x-\xi}{\alpha}\right)\right]$
    - α: scale parameter
    - ξ: location parameter
- **CDF**
  - $F_X(x) = \exp\left[-\exp\left(-\frac{x-\xi}{\alpha}\right)\right]$

Extreme Value Type I (Gumbel)

➢ Parameter estimation

- Method of moments
  - $\hat{\alpha} = \frac{s_X}{1.283}$
  - $\hat{\xi} = \bar{X} \mp 0.45 s_X$
- MLE
  - $\hat{\alpha} = \frac{s_X \sqrt{6}}{\pi} = 0.7797 s_X$
  - $\hat{\xi} = \bar{X} - 0.5772 \hat{\alpha}$

Quantile or Percentile

- $P_X(x_p) = 1 - p = 1 - \frac{1}{T}$
  - p: Exceedance probability
  - p=$\frac{i-a}{n+b}$

TABLE 18.3.1 Alternative Plotting Positions and their Motivation*

Name	Formula	a	T_1	Motivation
Weibull	$\frac{i}{n+1}$	0	$n+1$	Unbiased exceedance probabilities for all distributions
Median†	$\frac{i-0.3175}{n+0.365}$	0.3175	$1.47n+0.5$	Median exceedance probabilities for all distributions
APL	$\frac{i-0.35}{n}$	~0.35	$1.54n$	Used with PWMs [Eq. (18.1.13)]
Blom	$\frac{i-3/8}{n+1/4}$	0.375	$1.60n+0.4$	Unbiased normal quantiles
Cunnane	$\frac{i-0.40}{n+0.2}$	0.40	$1.67n+0.3$	Approximately quantile-unbiased
Gringorten	$\frac{i-0.44}{n+0.12}$	0.44	$1.79n+0.2$	Optimized for Gumbel distribution
Hazen	$\frac{i-0.5}{n}$	0.50	$2n$	A traditional choice

* Here a is the plotting-position parameter in Eq. (18.3.6) and T_1 is the return period each plotting position assigns to the largest observation in a sample of size n.

† For $i = 1$ and n, the exact value is $q_1 = 1 - q_n = 1 - 0.5^{1/n}$.

EQTIL (IMSL) & quantile() in R

- $\boldsymbol{Q(p) = (1 - f)x_j + f x_{j+1}}$
  - $j = int[p(n+1)], f = p(n+1) - j$
- $p = \frac{\{Q-(1+j)x_j + j x_{j+1}\}}{(n+1)(x_{j+1} - x_j)}$

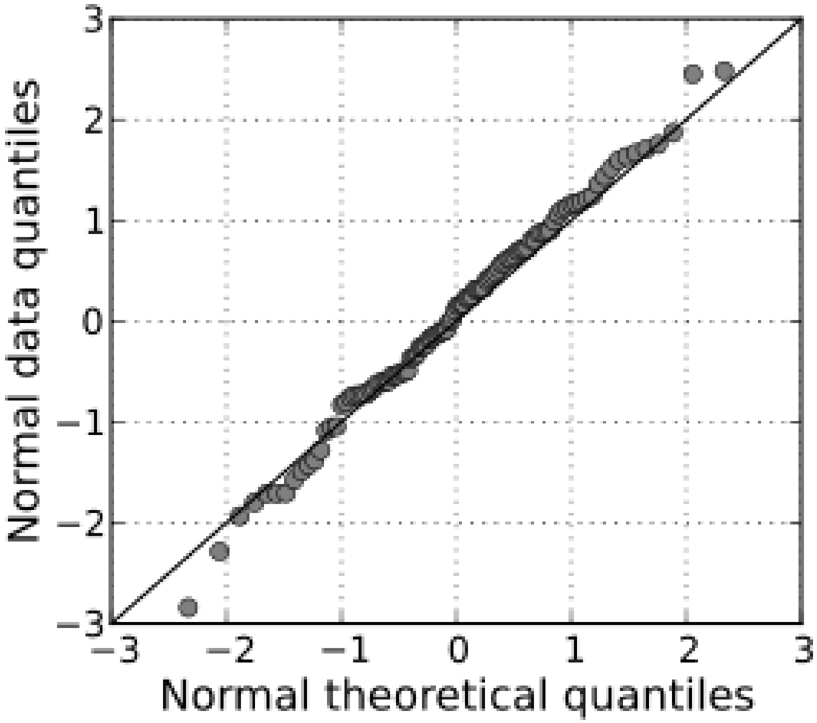
- **Examples…**
  - With Excel
  - PERCENTILE.EXC (MS office 2010)

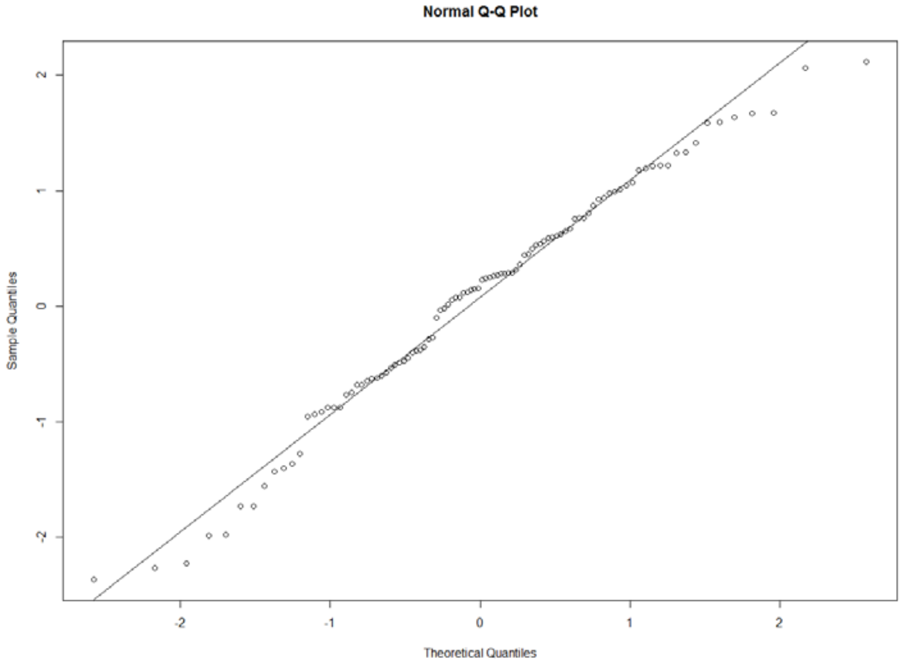
Quantile-Quantile (Q-Q) plot

➢ Graphical method

- Two probability distributions

```
> x<-rnorm(100, mean = 0, sd = 1)
> qqnorm(x)
> qqline(x)
```



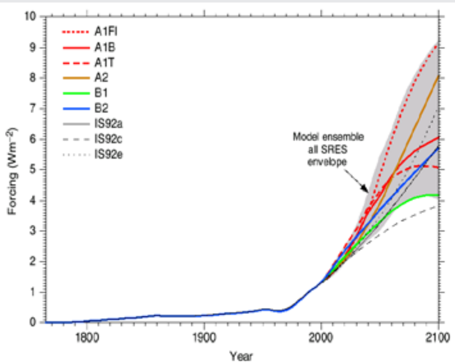


Statistical Downscaling Methods

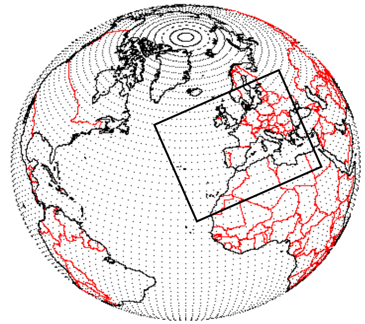
APCC
APEC CLIMATE CENTER

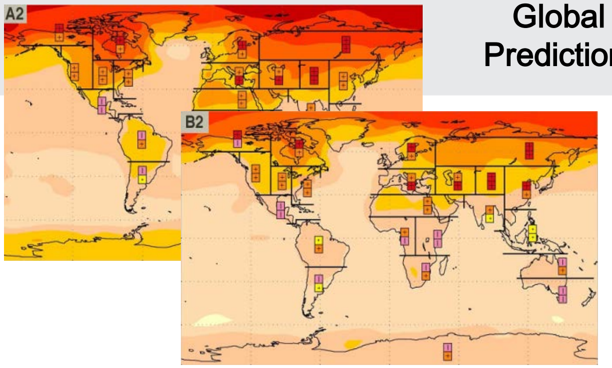
Downscaling techniques

Emission Scenarios



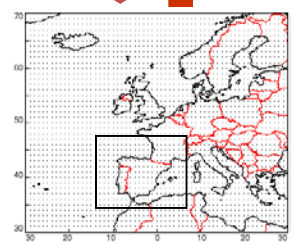
GCM



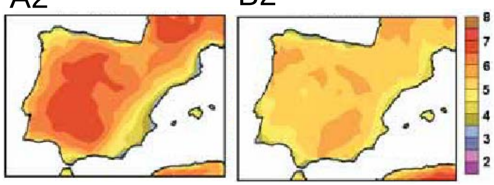


Global Predictions

RCM



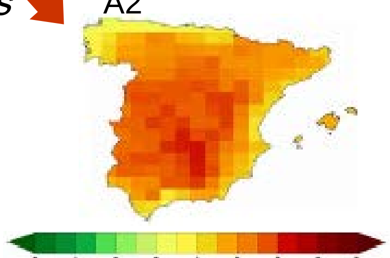
__Dynamical Downscaling__ runs regional climate models in reduced domains with boundary conditions given by the GCMs.



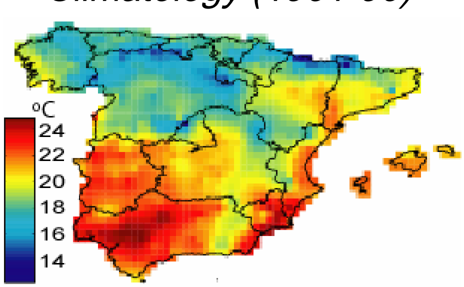
Historical Records

$$Y = f\,(X;\theta)$$

Different tecniques



Climatology (1961-90)

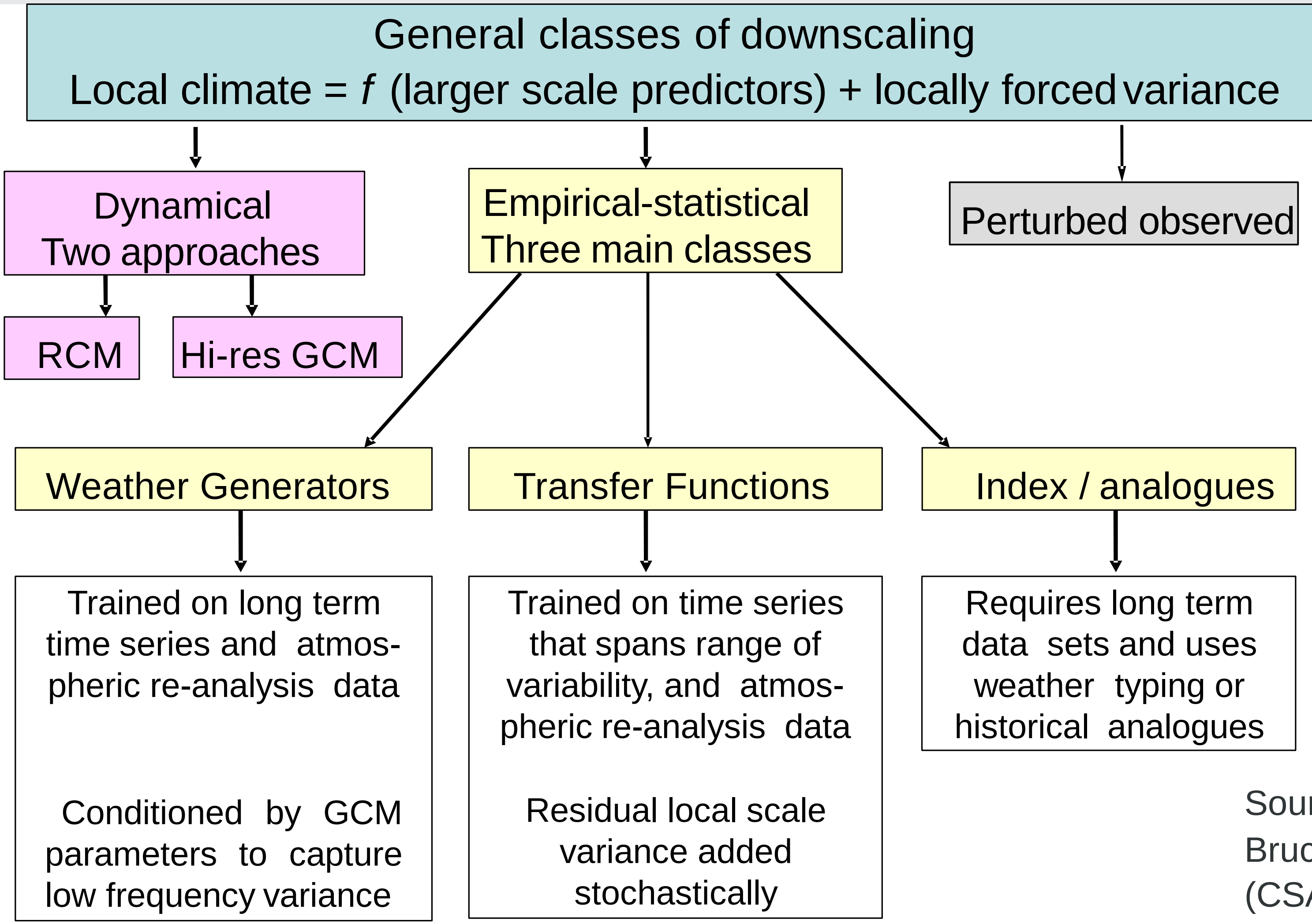


__Statistical Downscaling__ is based on empirical models fitted to data using historical records.

Source: http://www.mdm.unican.es

APEC CLIMATE CENTER

Downscaling techniques



Source:
Bruce Hewitson
(CSAG)

Pros and cons

➢ Dynamical and Statistical Downscaling

Method		
Statistical	• Easy and efficient to apply • Fine resolution • Able to directly incorporate observations into method • Apply to all GCMs	• Needs a reliable long-term observed data • Does not account for non-stationarity in the predictor-predictand and relationship • Affected by biases in underlying GCM • Dependent on choice of predictors
Dynamical	• Physical-based process • Produce numerous variables • Subdaily data available • Providing climate information at un-gauged points	• Computationally intensive • Limited resolution (25-50km) • Dependent on RCM parameterization • Considerable internal-variability (systematic bias)

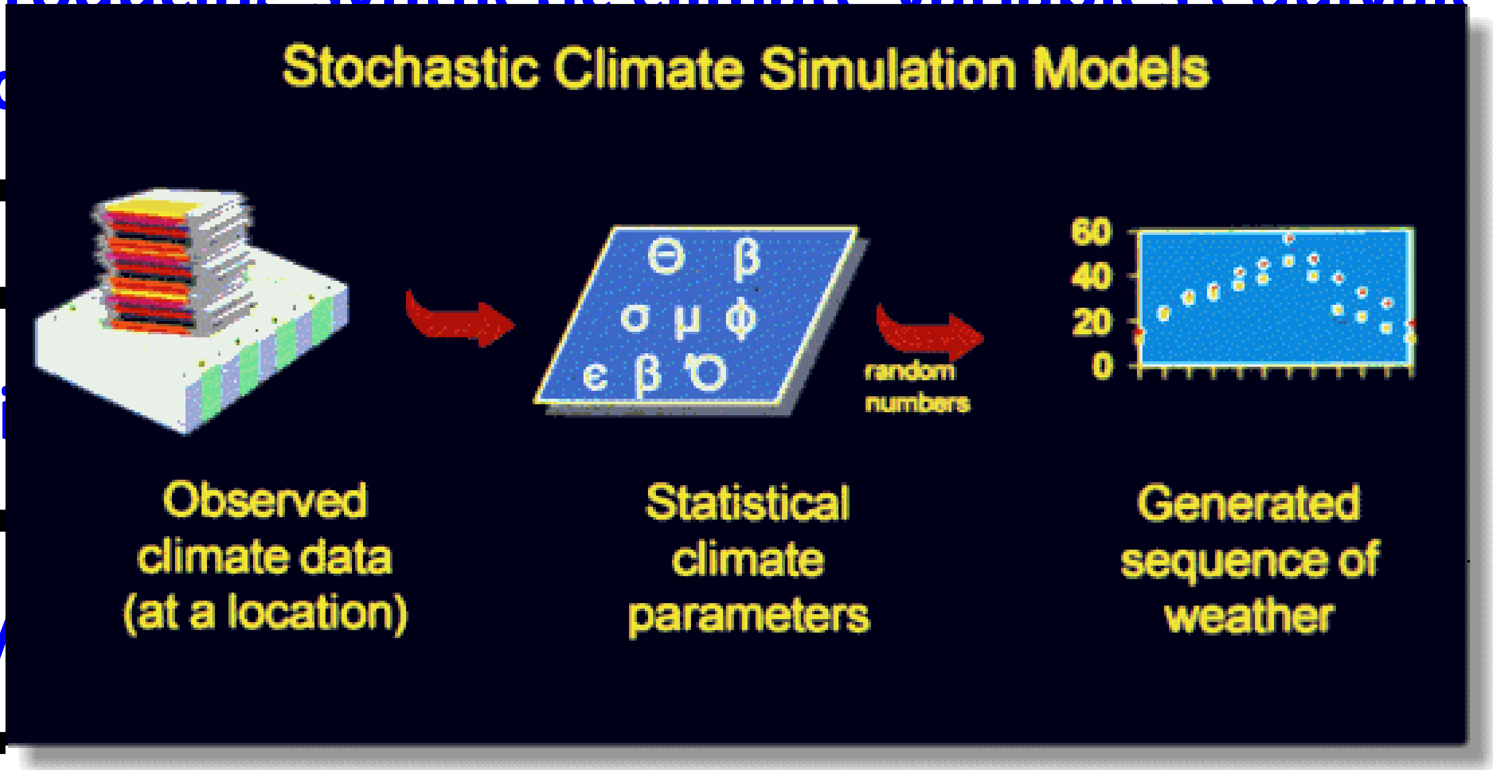
Reference: Cooney, 2012, Downscaling Climate Models: Sharpening the Focus on Local-Level Changes, *Environ Health Perspect* 120:a22-a28

Weather Generator (WG)

- Producing synthetic climate variables equivalent
  to

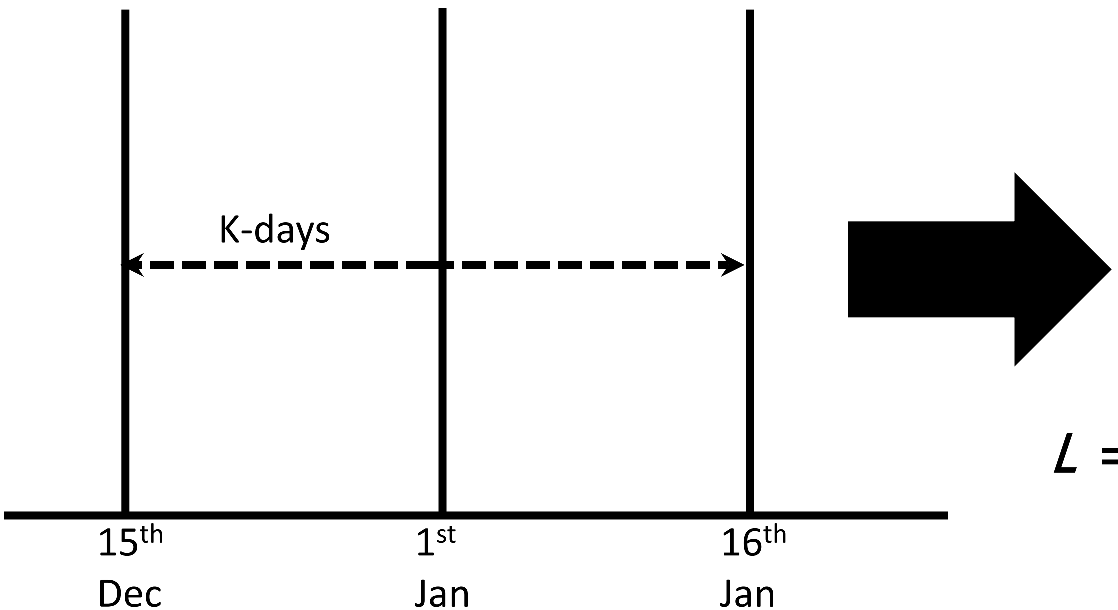
- Si

- M



Weather Generator (WG)

➢ K-nearest neighbor (K-NN) WG



Collecting historical data

$$\begin{cases} Dec15/1976 \\ Dec15/1977 \\ \vdots \\ Jan16/2015 \end{cases}$$

$$L = (k+1) \times N - 1 \text{ days}$$

K-NN WG

- **Mahalanobis distance for all candidates**

$$d_k = \sqrt{\left(\overline{X}_t - \overline{X}_k\right) C_t^{-1} \left(\overline{X}_t - \overline{X}_k\right)^{\mathrm{T}}}$$

- **Sorting the Mahalanobis distances d_k from the smallest to the largest**
- **Generate random number and determine the nearest neighbor using discrete probability distribution**

$$w_k = \frac{1/k}{\sum_{i=1}^{K} 1/i}$$

Yates et al. (2003)

$$p_j = \sum_{i=1}^{j} w_i$$

Sharif and Burn (2006)

K-NN WG

- **Estimate a conditional standard deviation σ for *K* nearest neighbors, and bandwidth λ. Then, use the perturbation process.**

$$y_{i,t}^{j} = x_{i,t}^{j} + \lambda \sigma_i^j z_t \qquad \lambda = 1.06\,\sigma\,K^{1/5} \qquad \lambda_a = x_{*,t}^{j} / 1.55 \sigma_*^j$$

- **Repeat previous steps to produce time series as long as users want**

K-nearest neighbor (K-NN) WG

Climate Change Impact Assessment Using K-Nearest Neighbor Weather Generator: Case Study of the Nakdong River Basin in Korea

Hyung-Il Eum[1]; Slobodan P. Simonovic, M.ASCE[2]; and Young-Oh Kim, M.ASCE[3]

Abstract: This study presents a methodology for assessing impacts of climate change in regional-scale hydrology using the K-nearest neighbor weather generator (WG) model in combination with the outputs of global circulation models (GCMs) and hydrologic rainfall-runoff model. The Nakdong River Basin in Korea is used as a case study. The study applies a systematic approach to select the variables for the WG model from five meteorological variables available in the basin. In addition, the GCMs' projections based on a B1 emission scenario are incorporated into the proposed WG model to reflect the impacts of climate change, consequently generating the meteorological series for 60 years. Meteorological data for historic and two climate scenarios (dry and wet) are generated. The generated time series of meteorological variables are used with the streamflow synthesis and reservoir regulation rainfall-runoff model to calculate the streamflow at 25 sub-basins. The results demonstrate that the WG models combined with the GCMs' outputs are able to provide future weather conditions for the assessment of climate change impacts. One of the major findings of the study is the potential severity of drought impacts that may increase the historical drought duration in the basin up to three times.

DOI: 10.1061/(ASCE)HE.1943-5584.0000251

CE Database subject headings: Climate change; Streamflow; Case studies; Korea, South.

Author keywords: Weather generator; Climate change; Streamflow.

Introduction

Climate change is generally defined as a long-term significant

spatial resolution problem. Dynamic downscaling incorporates the GCMs' results directly as a boundary condition using the complex physical-based algorithms to describe atmospheric pro-

[1]Associate Professor, Dept. of Civil and Environmental Engineering, Seoul National Univ., 599 Gwanak-ro, Gwanak-gu, Seoul 151-742, Korea.

Table 1. Variable Combinations for the WG Models

WG models	WG-5Var	WG-4Var	WG-3Var	WG-2Var
Variables	Precipitation, temperature, relative humidity, wind speed, sea level pressure	Precipitation, temperature, relative humidity, wind speed	Precipitation, temperature, relative humidity	Precipitation, temperature

Table 2. Median Absolute Bias of Four WG Models

	Youngju		Gumi		Jinju	
WG models	Precipitation (mm)	Temperature (°C)	Precipitation (mm)	Temperature (°C)	Precipitation (mm)	Temperature (°C)
WG-2Var	10.10	0.18	7.12	0.10	9.84	0.23
WG-3Var	8.77	0.13	6.41[a]	0.18	9.54	0.09
WG-4Var	14.19	0.19	11.42	0.17	16.14	0.17
WG-5Var	12.36	0.19	8.09	0.17	14.70	0.17

[a]Best performance (lower number represents the better performance).

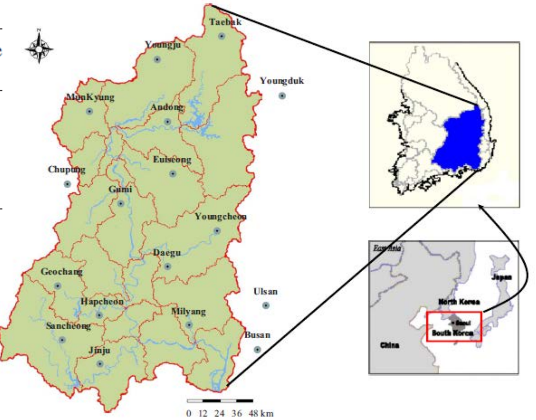
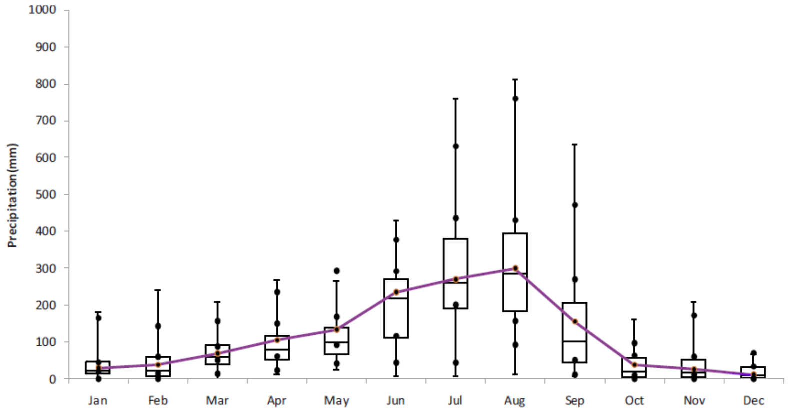
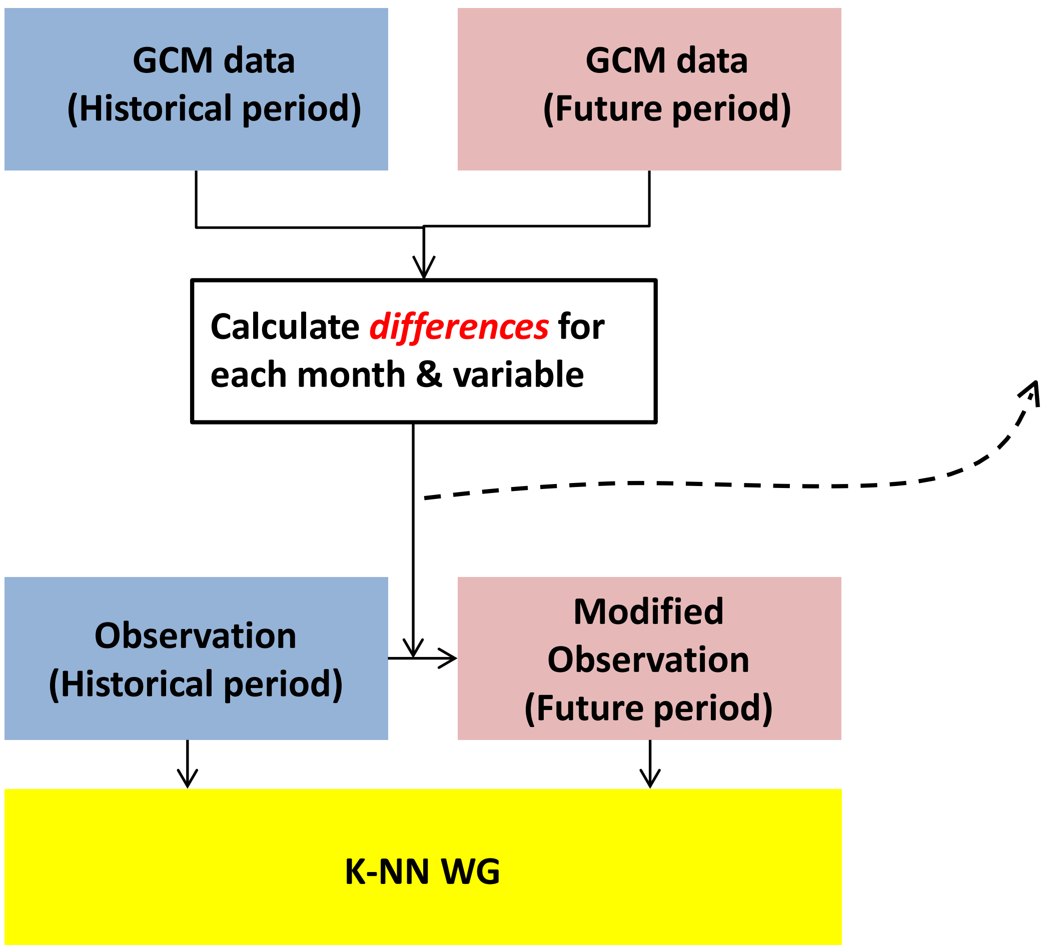


Fig. 1. Nakdong River Basin, Korea



(a) WG-3Var

K-NN WG under Climate Change



$$P_{fut} = P_{his} \times P_{delta}$$

$$T_{fut} = T_{his} + T_{delta}$$

K-NN WG under Climate Change

➢ **Modify observations with climate signals of GCMs**

Table 3. Difference in Precipitation between the Historic and the B1 Climate Change Scenario

Month	CSIRO mk 3.0	GFDL cm 2.1	MIROC3	ECHAM5	CCSM3
January	−0.147	0.213	0.113	−0.572	0.391
February	−0.004	0.089	0.051	0.584	0.195
March	0.042	−0.011	0.358	0.009	0.125
April	−0.031	−0.280	0.791	1.025	−0.121
May	−0.306	0.030	0.555	0.154	0.100
June	0.178	0.467	1.473	−0.743	0.287
July	0.016	−0.117	0.143	0.564	−0.019
August	−0.594	0.088	0.895	−0.387	0.569
September	−0.433	0.321	0.225	0.310	0.434
October	−0.562	−0.371	−0.185	−0.145	−0.320
November	−0.121	0.126	0.136	−0.352	0.105
December	−0.054	0.217	−0.140	0.358	0.488
Sum of differences	**−2.016**	0.772	**4.416**	0.803	2.233

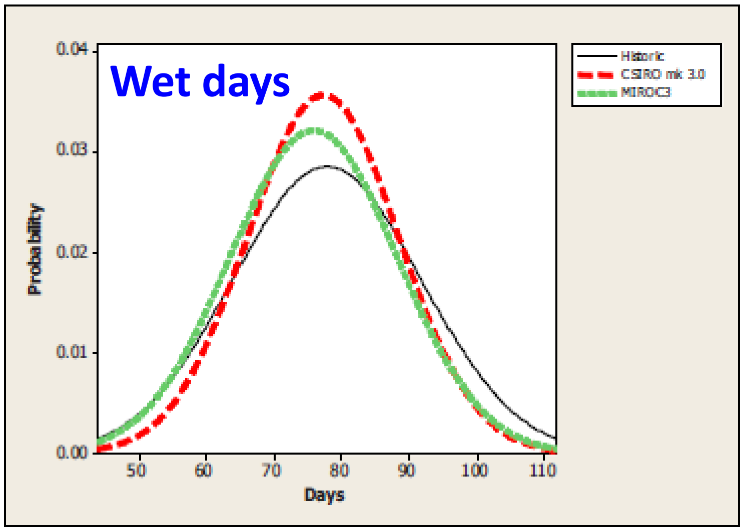
Note: Unit: mm/day.

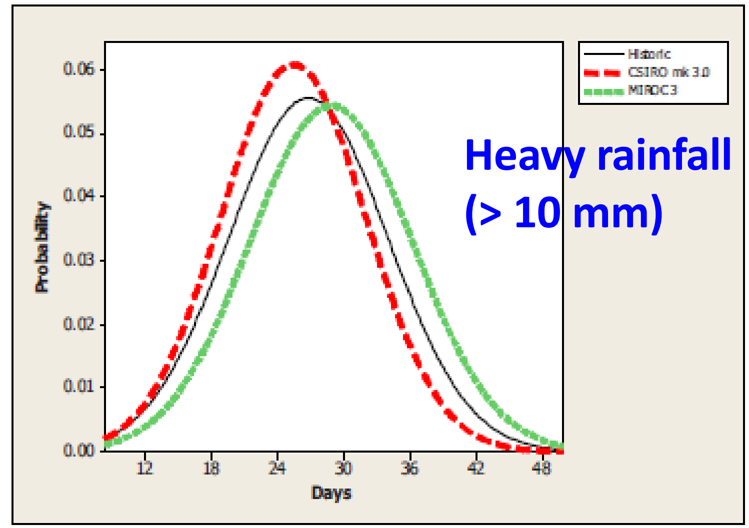
Table 4. Difference in Temperature between the Historic and the B1 Climate Change Scenario

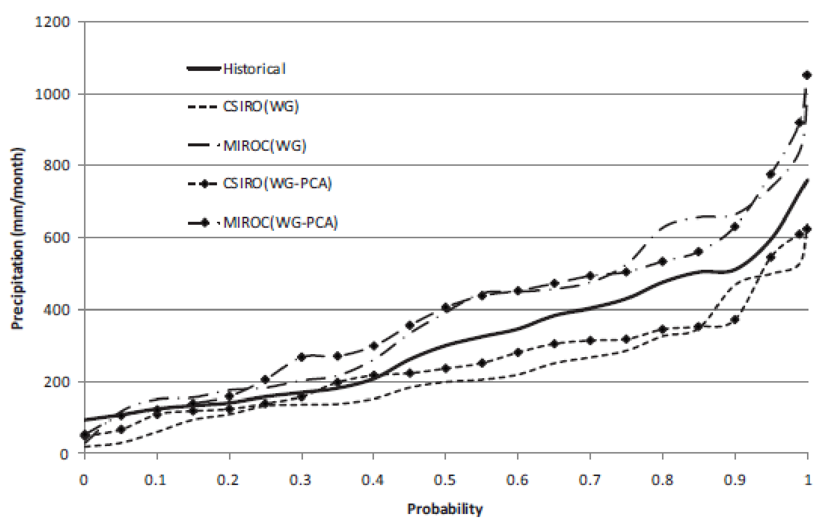
Month	CSIRO mk 3.0	GFDL cm 2.1	MIROC3	ECHAM5	CCSM3
January	0.42	1.86	2.12	0.72	0.57
February	0.54	0.80	1.87	1.38	1.16
March	0.54	1.19	1.66	0.65	0.92
April	0.40	1.03	1.53	0.29	0.90
May	0.43	1.15	1.41	0.37	0.65
June	0.54	0.86	1.71	0.68	1.27
July	0.56	0.96	1.33	0.60	1.40
August	0.18	1.01	1.58	0.80	0.86
September	0.16	0.95	1.69	0.66	1.29
October	0.09	1.03	1.78	0.24	1.34
November	0.06	0.66	1.53	0.12	1.19
December	0.72	0.84	1.74	1.14	0.87
Average	**0.39**	1.03	**1.66**	0.64	1.03

Note: Unit: °C

K-NN WG under Climate Change



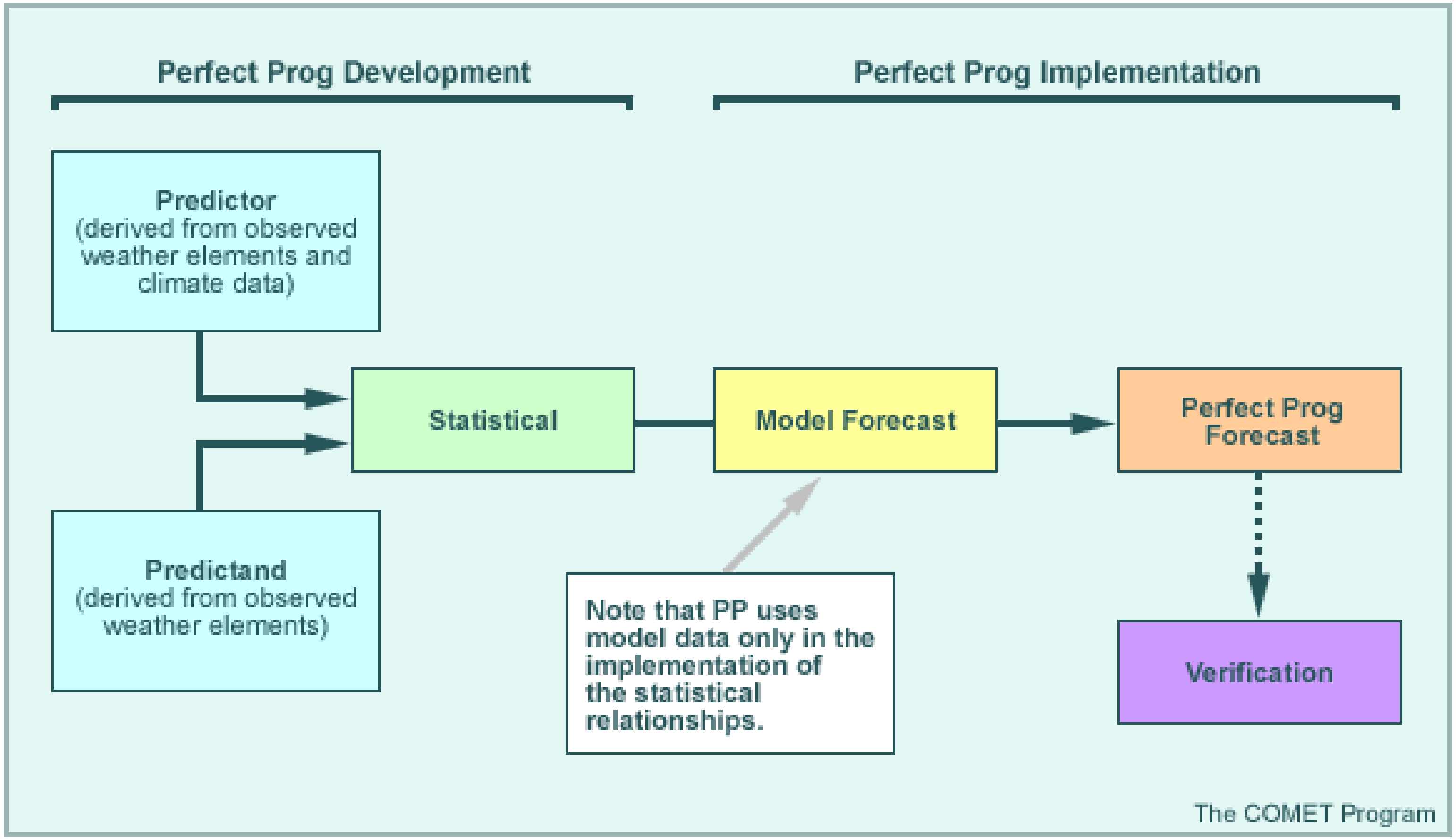




APEC CLIMATE CENTER

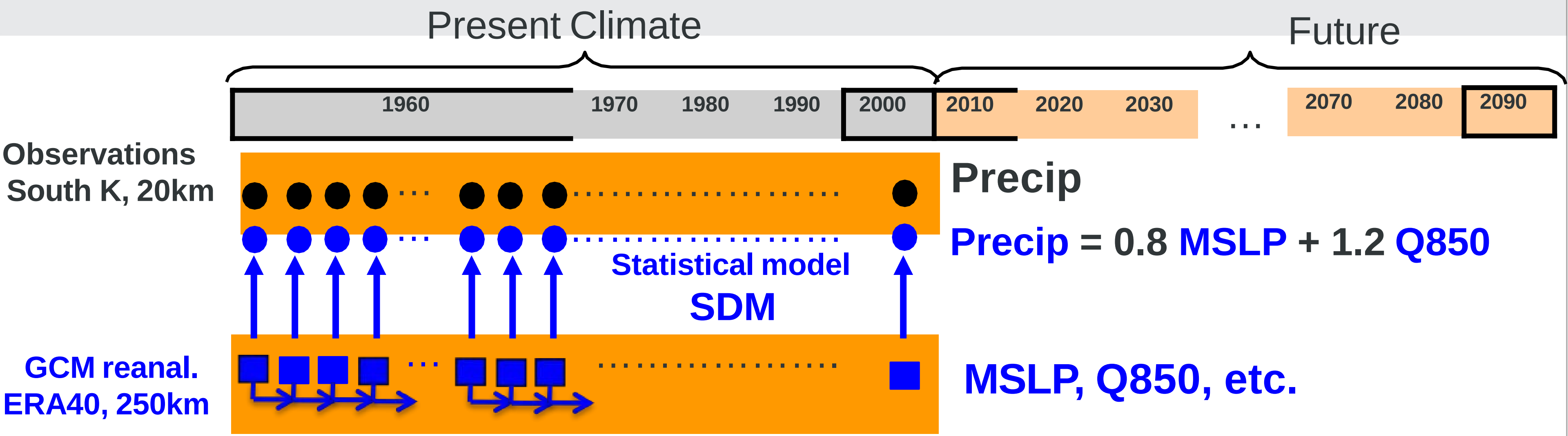
Statistical downscaling techniques

➢ Perfect Prognosis (PP)

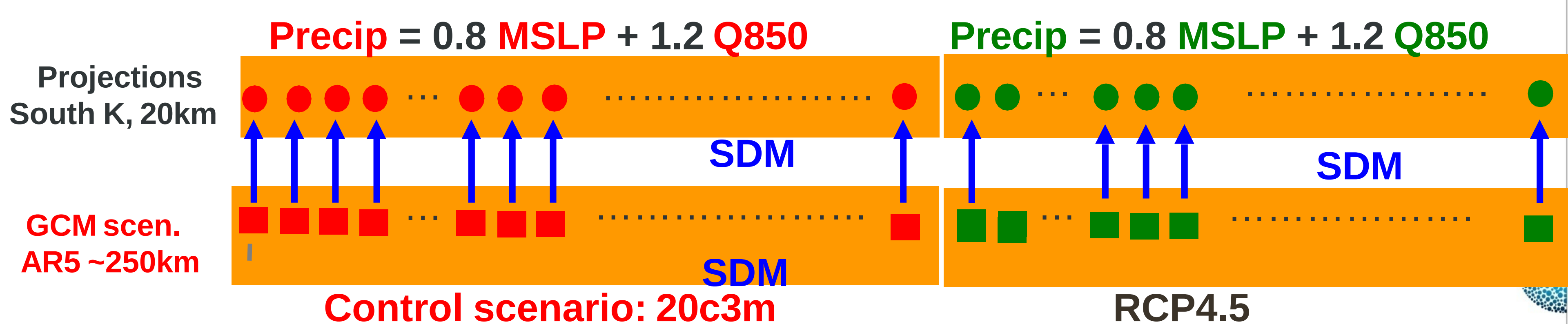


Source: http://www.met.tamu.edu/class/metr452/models/2001/output.html#Statistical Weather Forecasting

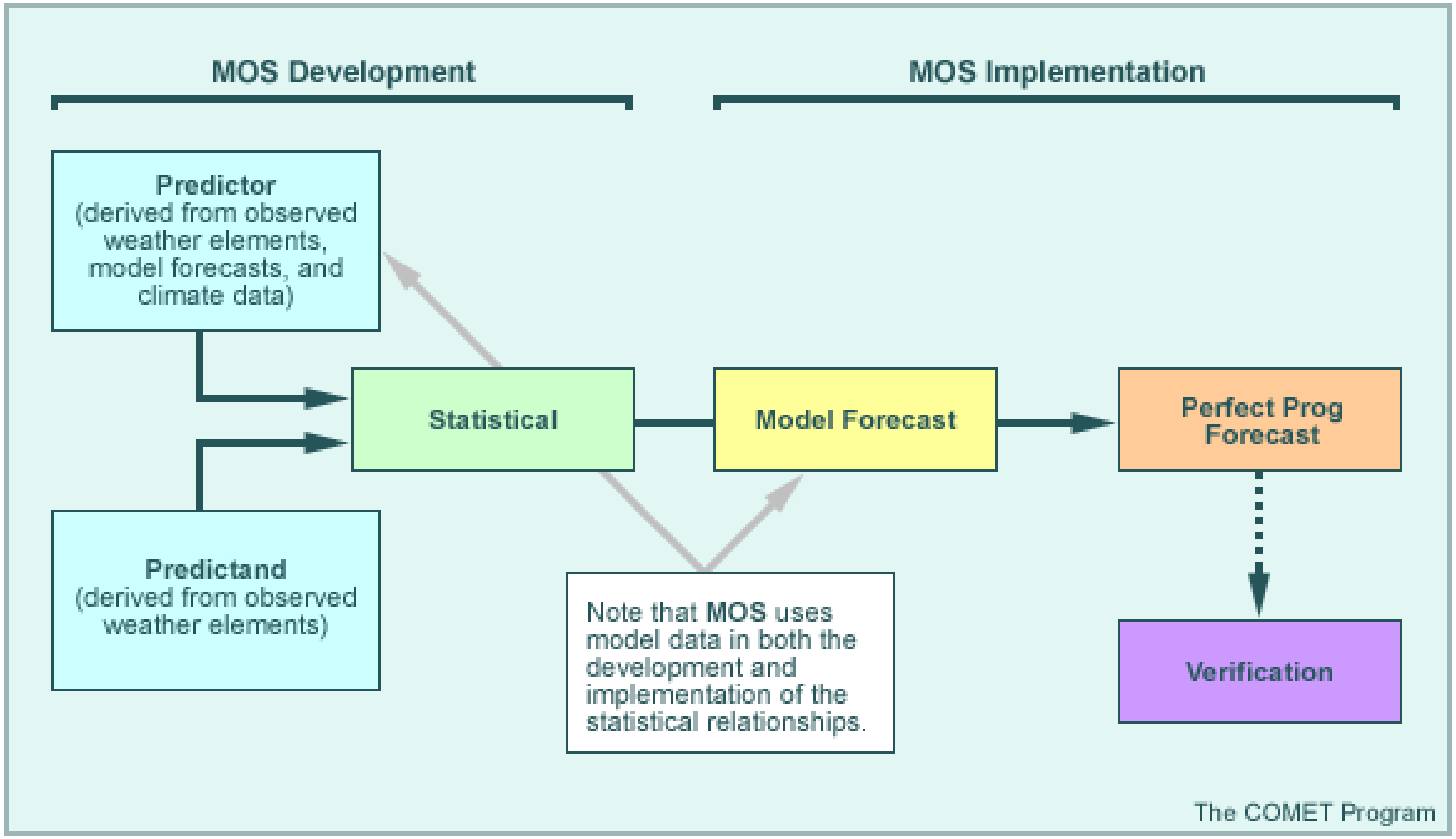
Perfect Prognosis (PP)



- **Assumption 1:** Reanalysis choice
- **Assumption 2:** Choosing consistent predictors:
- **Assumption 3:** Stationarity/robustness: SDM SDM

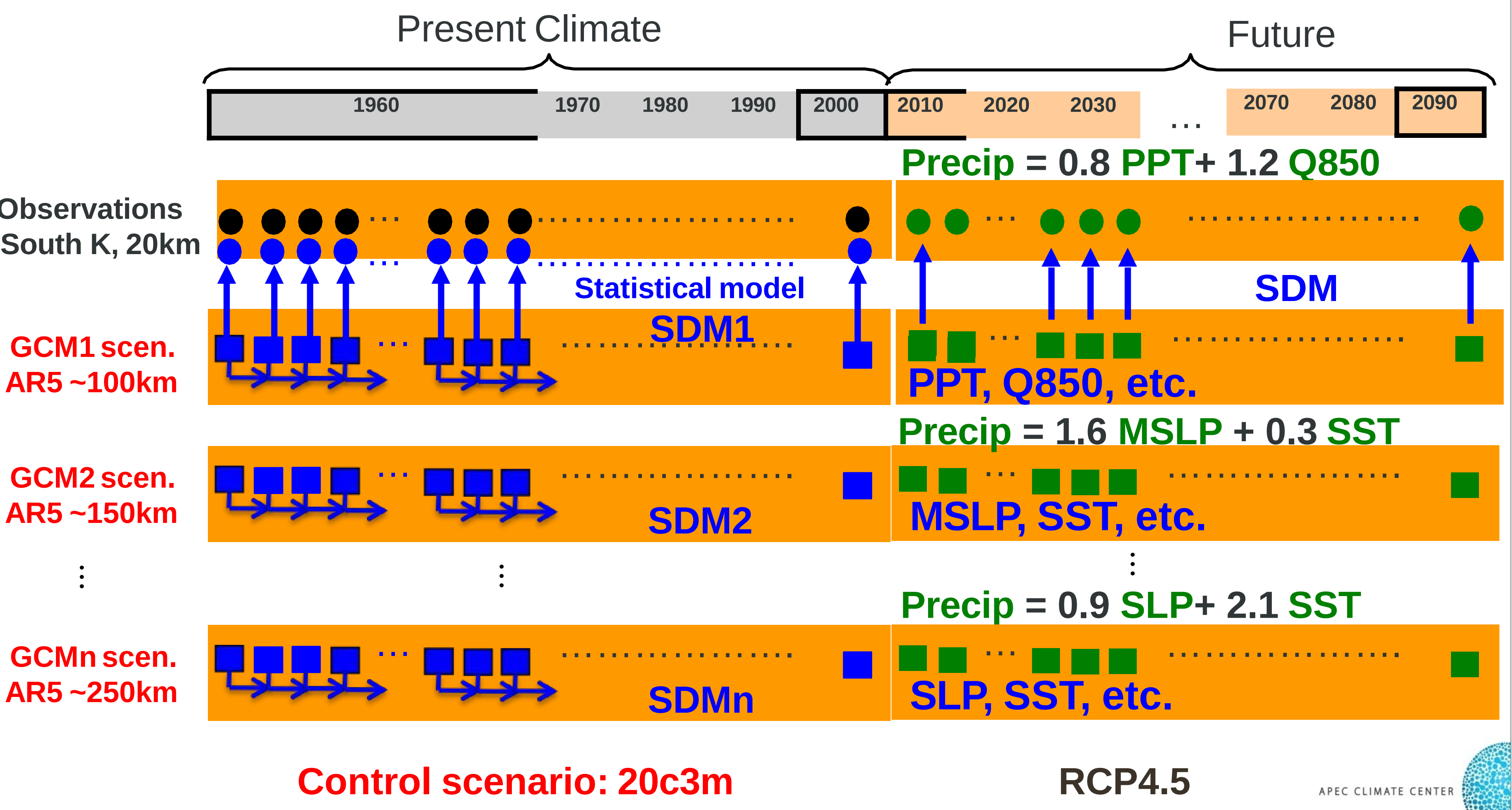


Model Output Statistics (MOS)



Source: http://www.met.tamu.edu/class/metr452/models/2001/output.html#Statistical Weather Forecasting

MOS



APEC CLIMATE CENTER

MOS

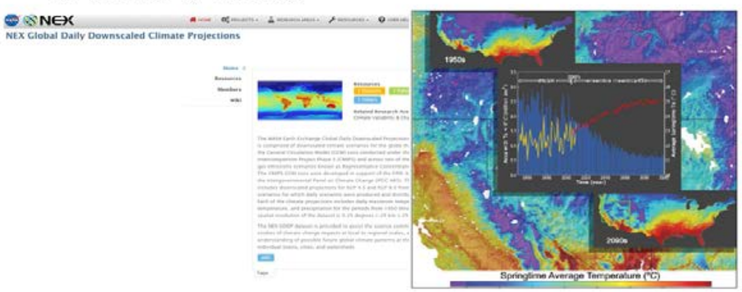
- **Predictors**
  - From the global (or regional) model for both training and downscaling phases
- **Need the model output**
  - Day-to-day correspondence with observations
- **These methods can work with the variable of interest as predictor**
  - Local precipitation can be derived from the direct model precipitation forecasts

MOS models

- **Most popular**
  - Bias Correction/Spatial Disaggregation (BCSD)
  - Bias Correction/Constructed Analogs (BCCA)

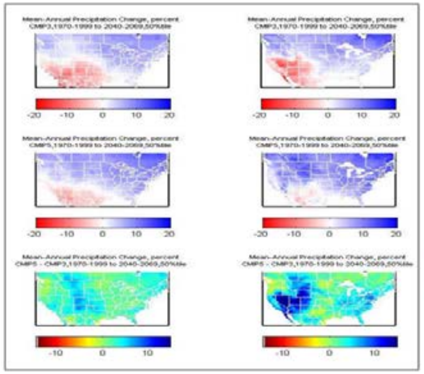
NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP)

- **BCSD**
  - 1950 through 2100 at 0.25 degrees (~25 km x 25 km)
  - 21 GCMs × 2 RCPs



Downscaled CMIP3 and CMIP5 climate and hydrology projections (DCHP)

- **BCSD & BCCA**
  - 1950-2099 at 1/8° (~12km)

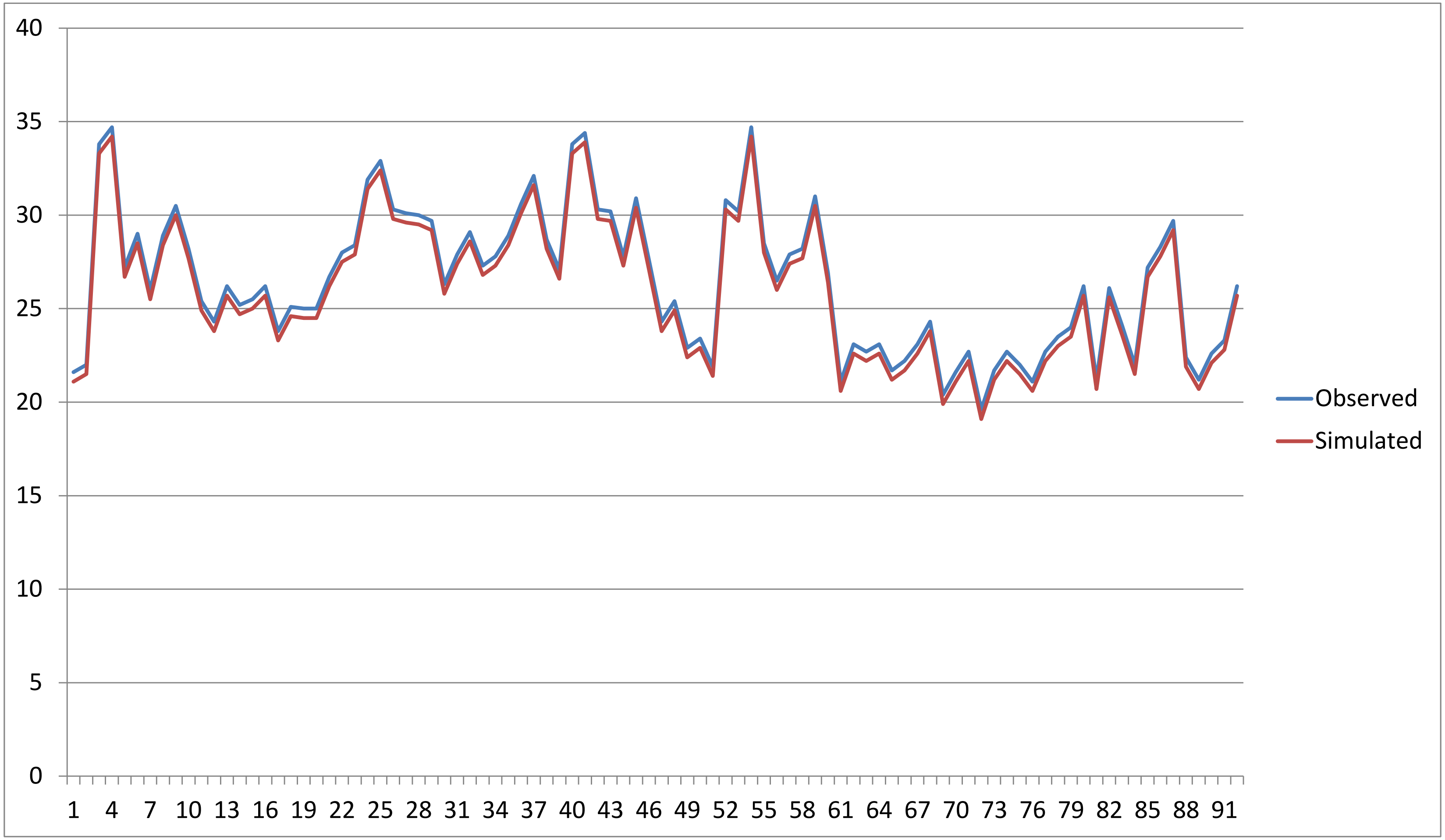


http://gdo-dcp.ucllnl.org/downscaled_cmip_projections/dcpInterface.html#Welcome

Model error

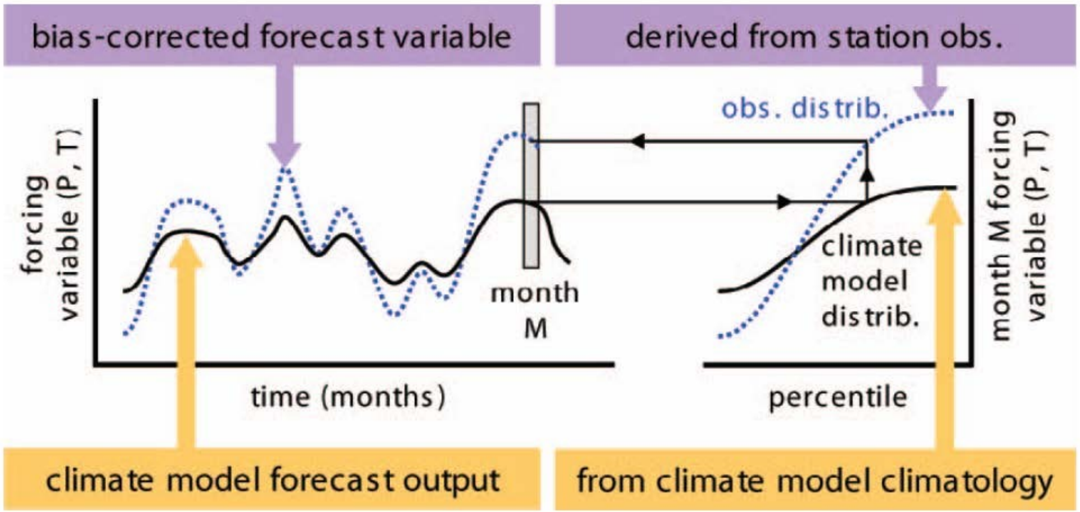
- **Error**
  - Systematic
    - Bias
  - Random
    - Ex: measurement errors
- **Bias correction**

Bias-correction



Bias Correction/Spatial Disaggregation (BCSD)

- **Statistical bias correction of GCM simulations**
  - Quantile mapping
- **Spatial downscaling to fine scale (i.e. stations)**
- **Temporal disaggregation from monthly to daily**



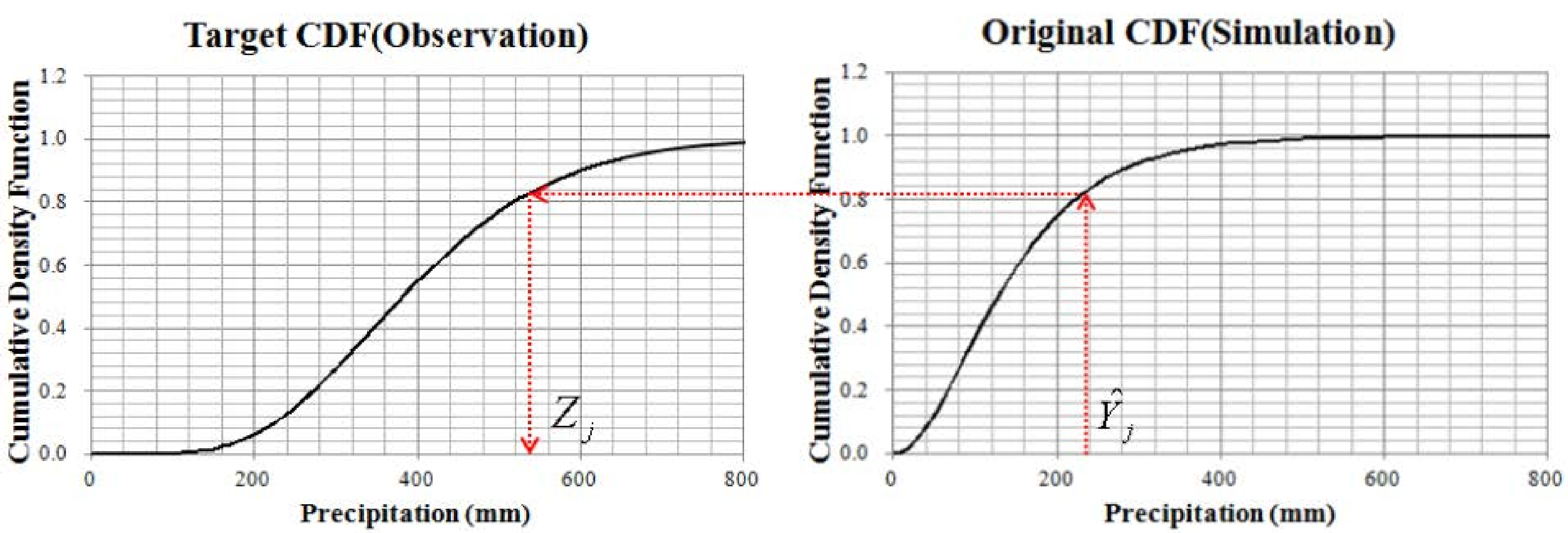
Source: Wood et al 2006, BAMS

Bias Correction/Spatial Disaggregation (BCSD)

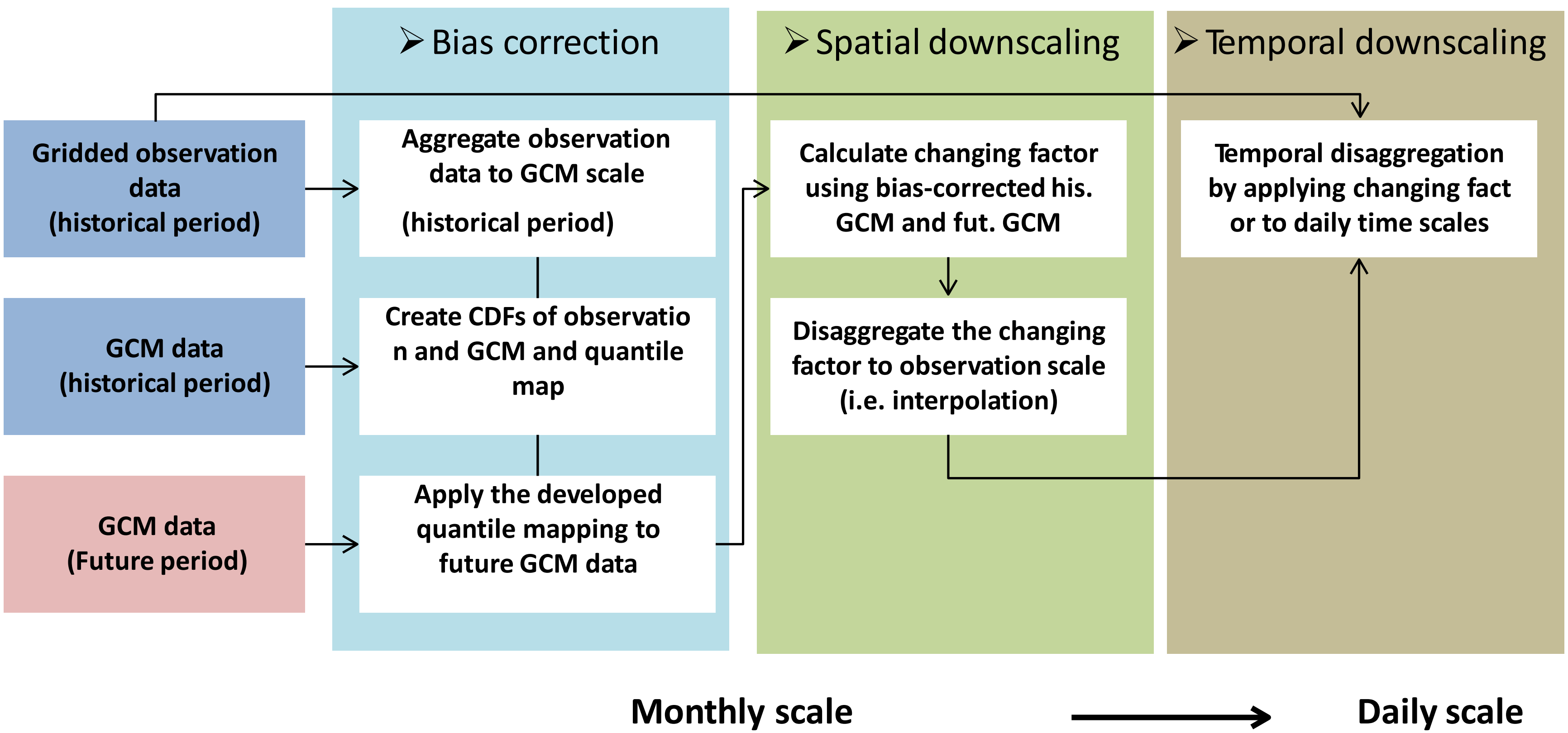
➢ Bias-correction: Quantile mapping

- Monthly data set at large grid points (e.g. GCM points)

$$Z_j(t) = F_{\text{obs}}^{-1}\left[F_m\{\hat{Y}_j \;\; (t)\}\right]$$



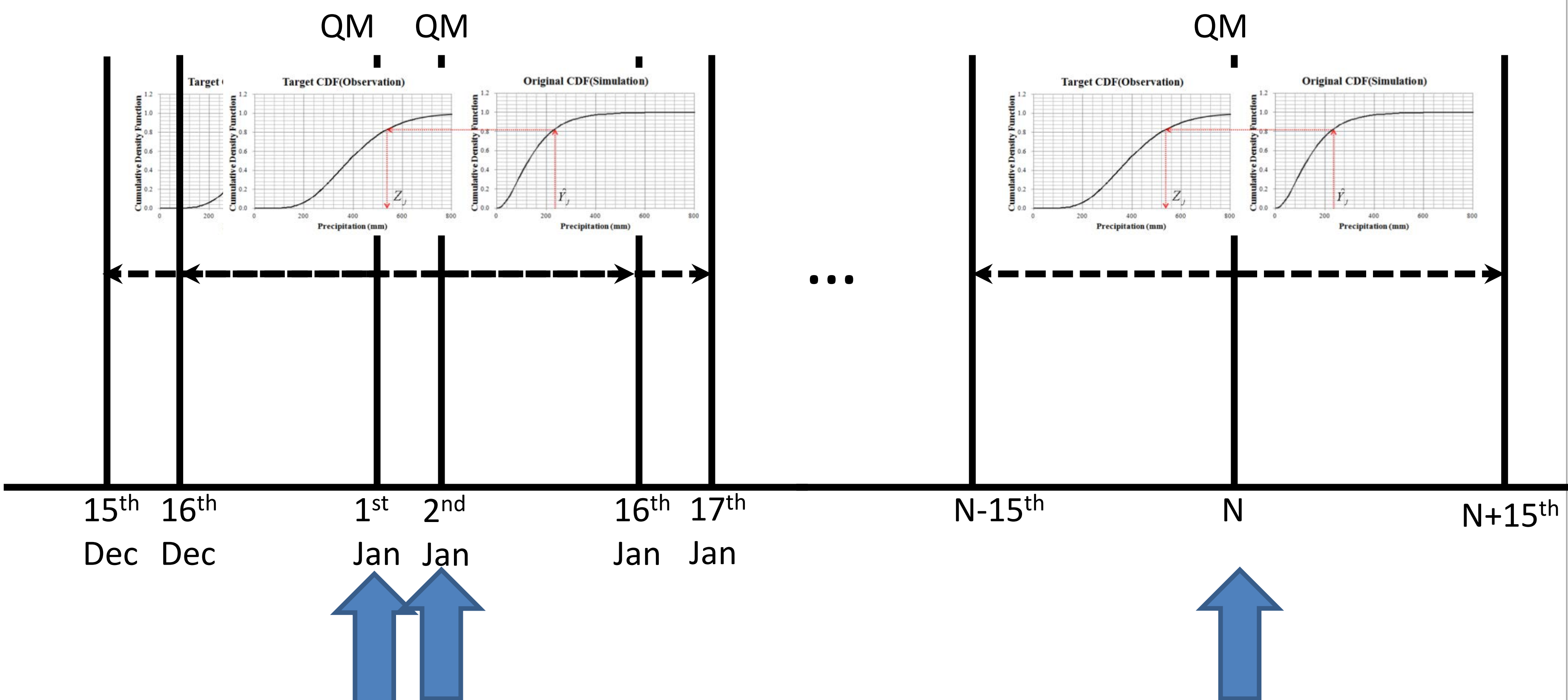
Bias Correction/Spatial Disaggregation (BCSD)



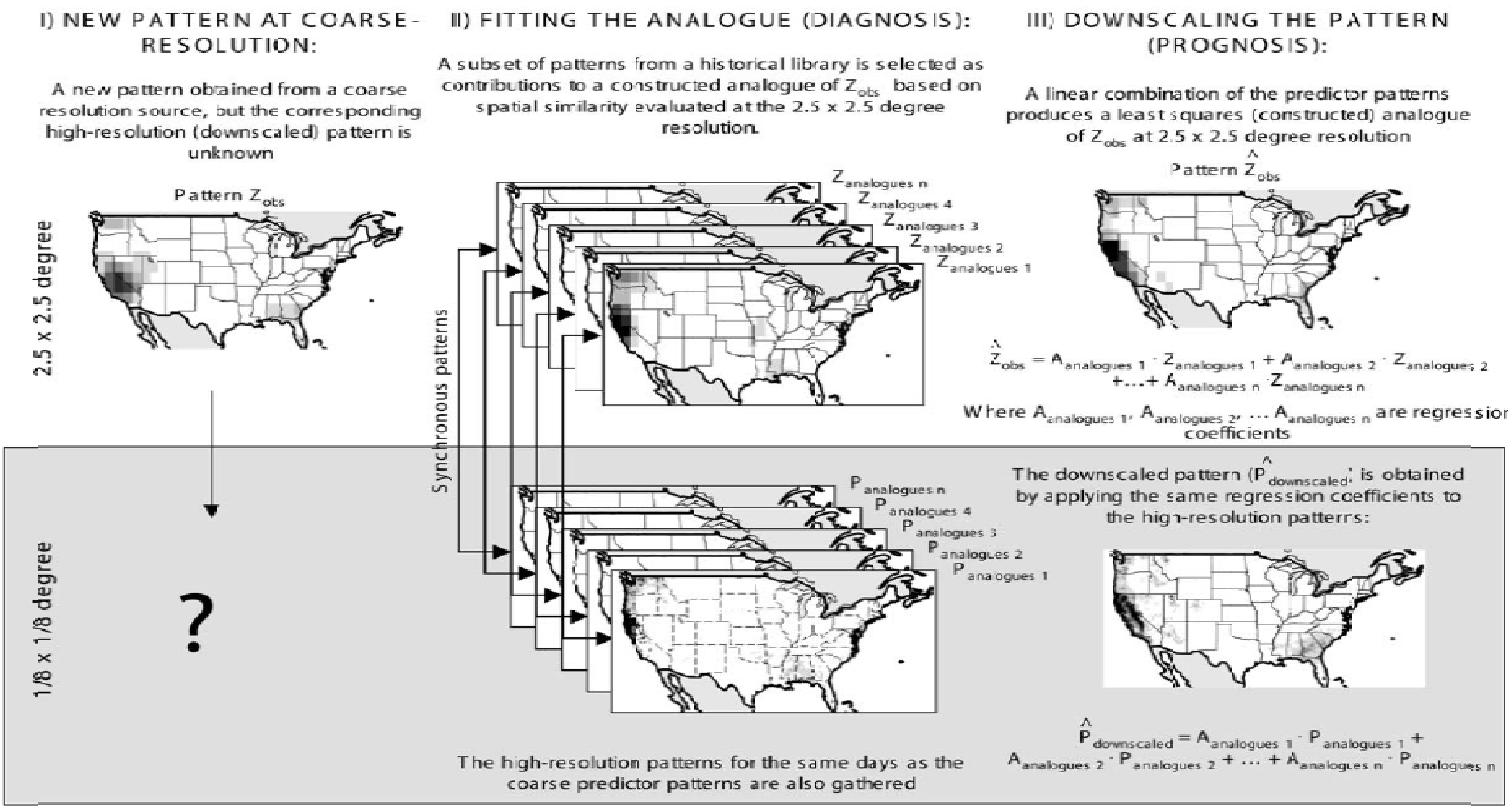
Daily BCSD (1)

- **Spatial disaggregation**
  - Interpolating daily GCM output to finer grid points
- **Bias correction**
  - Quantile mapping
    - Sample distribution of ±15-day moving window
- **No need of temporal disaggregation**
- **Values out of historical range**
  - Gumbel for precipitation
  - Normal distribution for temperature

Daily BCSD (2)

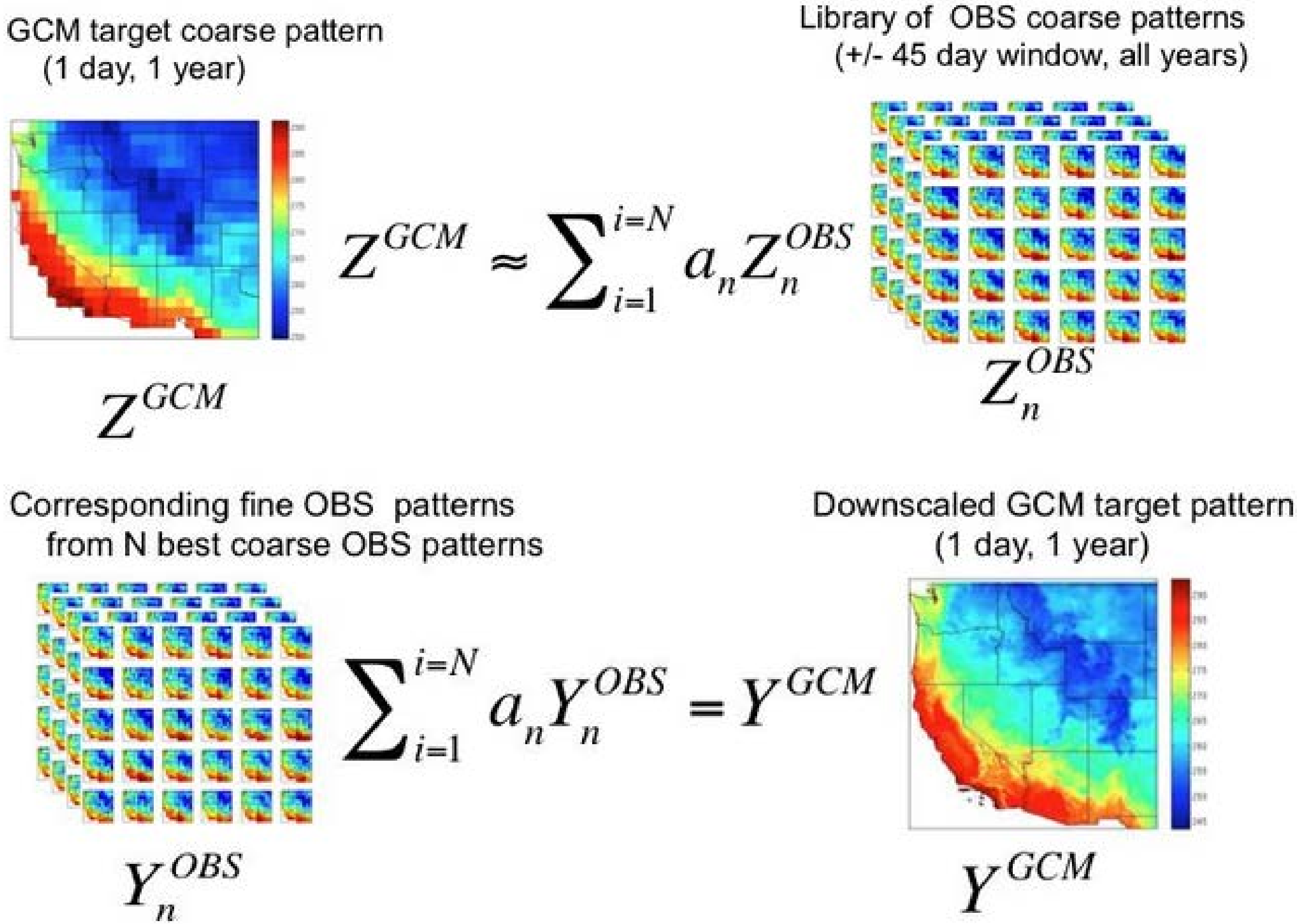

QM
QM
QM
Target CDF(Observation)
Original CDF(Simulation)
Cumulative Density Function
Precipitation (mm)
...
15th Dec
16th Dec
1st Jan
2nd Jan
16th Jan
17th Jan
N-15th
N
N+15th

BCCA (1)



Source: Hidalgo et al. (2008)

BCCA (2)



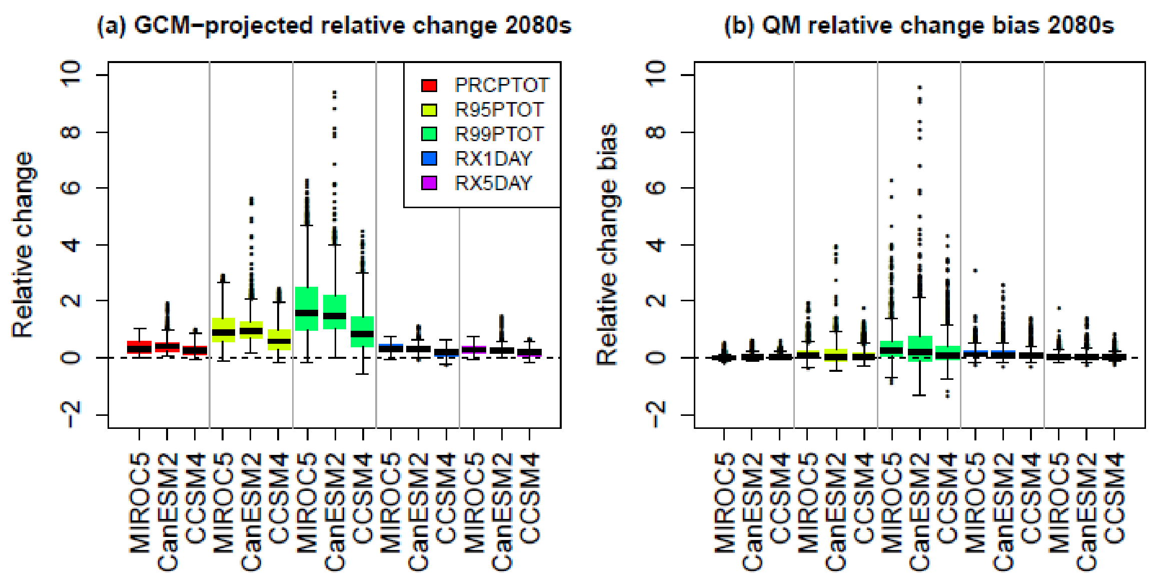
Source: http://climate.northwestknowledge.net/MACA/MACAmethod.php

What if does a long-term trend exist?

➢ QM Equation

- Extrapolation required
  - Gumbel distribution for PRCP
  - Normal for temperature (i.e. TMAX, TMIN)

Inflation by frequent extrapolation

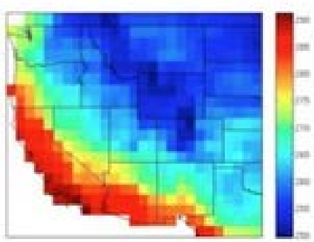


Source: Cannon et al. (2015)

What if does a long-term trend exist?

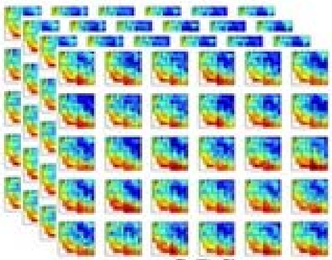
➢ BCCA

- Hard to find similar weather patterns from historical data
- Not reliable climate projections



Z^{GCM}

$$Z^{GCM} \approx \sum_{i=1}^{i=N} a_n Z_n^{OBS}$$



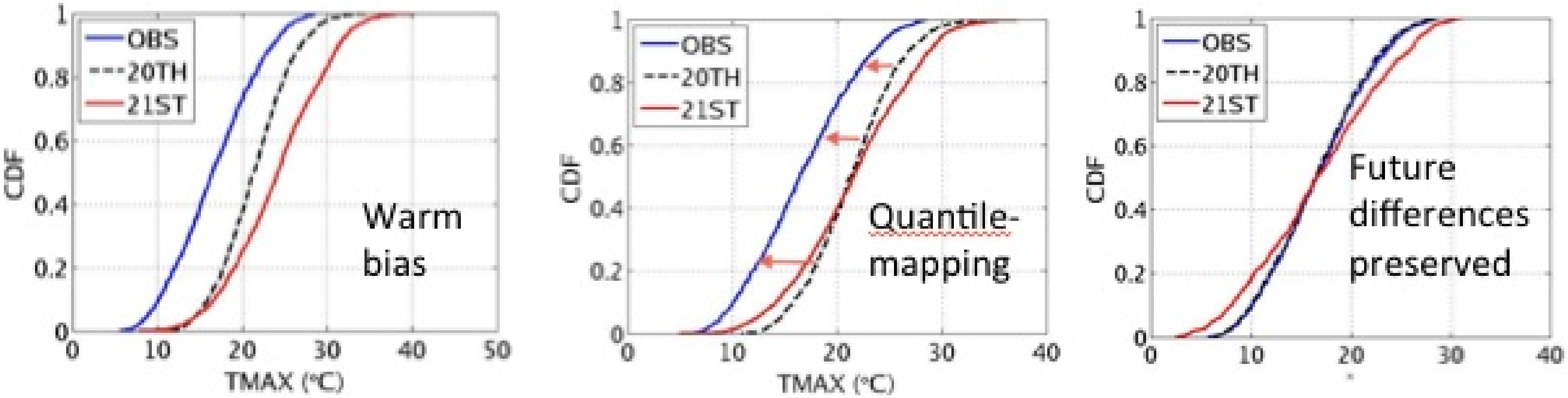
Z_n^{OBS}

Few analogs similar to future spatial structure

Detrended quantile mapping (DQM)

- **Account for changes in the projected values**
  - Removing the modelled trend in the long-term mean
  - Reimposing it after QM

$$\hat{x}_{m,p}(t) = F_{o,h}^{-1}\left(F_{m,h}\left(\frac{\bar{x}_{m,h}x_{m,p}(t)}{\bar{x}_{m,p}(t)}\right)\right)\frac{\bar{x}_{m,p}(t)}{\bar{x}_{m,h}}$$



Raw GCM Data | After Epoch Adjustment | After Bias Correction

Source: http://climate.northwestknowledge.net/MACA/MACAmethod.php

Quantile Delta Mapping (QDM)

- **Preserving model-projected relative changes in quantiles**
- **Correcting systematic biases in quantiles**
- **Step 1: Calculate percentile from modelled CDF**

$$\tau_{m,p}(t) = F_{m,p}\{x_{m,p}(t)\}$$

- **Step 2: Relative change in quantiles between historic and future**

$$\Delta_m = \frac{F_{m,p}^{-1}\{\tau_{m,p}(t)\}}{F_{m,h}^{-1}\{\tau_{m,p}(t)\}}$$

Quantile Delta Mapping (QDM)

- **Step 3: Bias-correction from observed values**

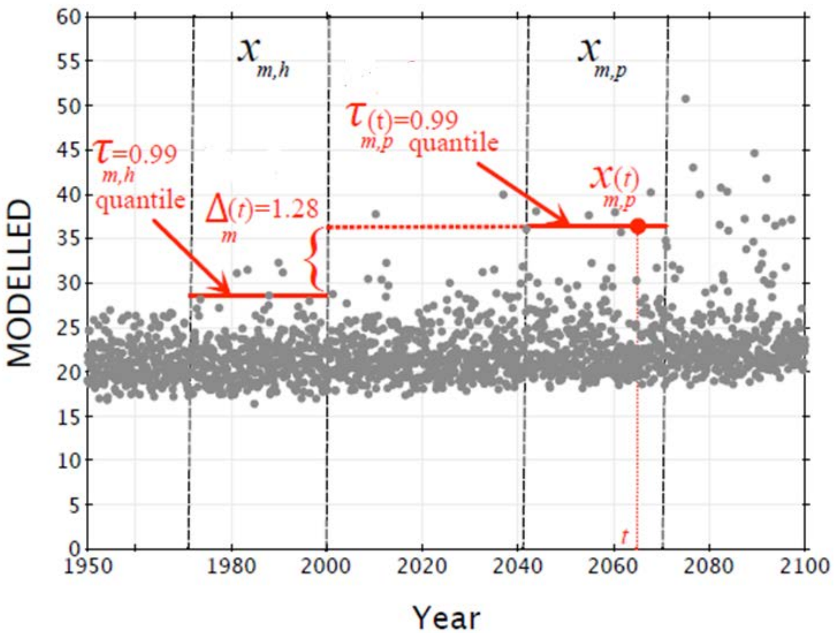
$$\hat{x}_{o:m,h:p}(t) = F_{o,h}^{-1}\{\tau_{m,p}(t)\}$$

- **Step 4: Applying he relative change (Step 2) to bias-corrected value in (Step 3)**

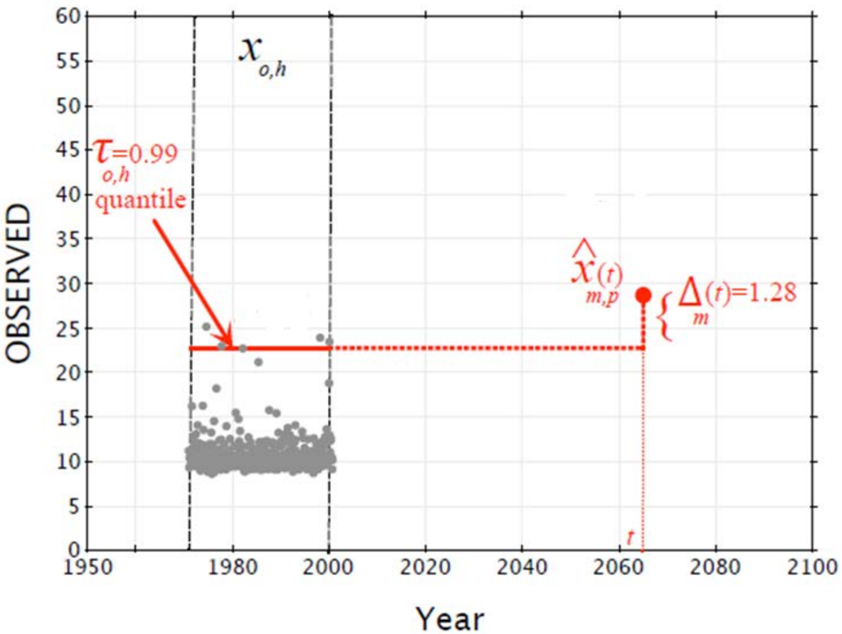
$$\hat{x}_{m,p}(t) = \hat{x}_{o:m,h:p}(t)\ \Delta_m$$

Quantile Delta Mapping (QDM)

Step 1-2



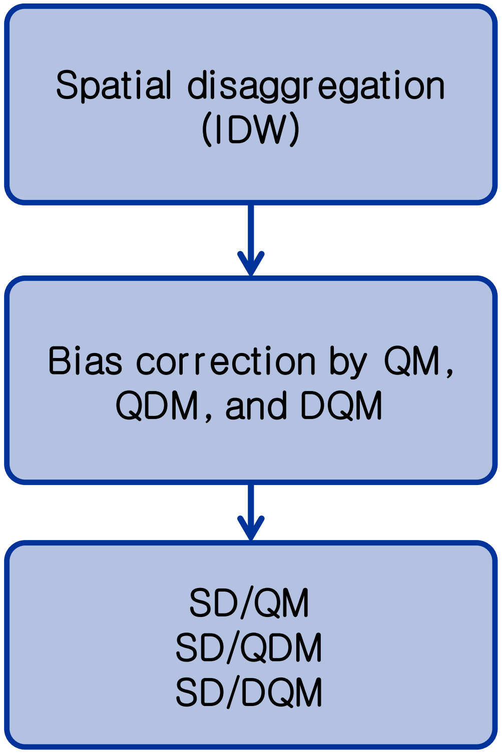
Step 3-4



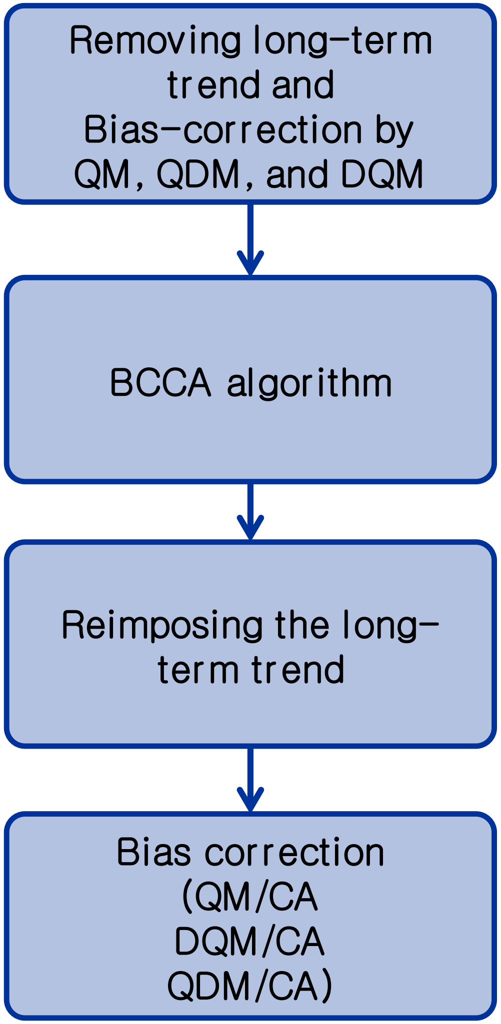
Source: Cannon et al. (2015)

SDMs

BCSD

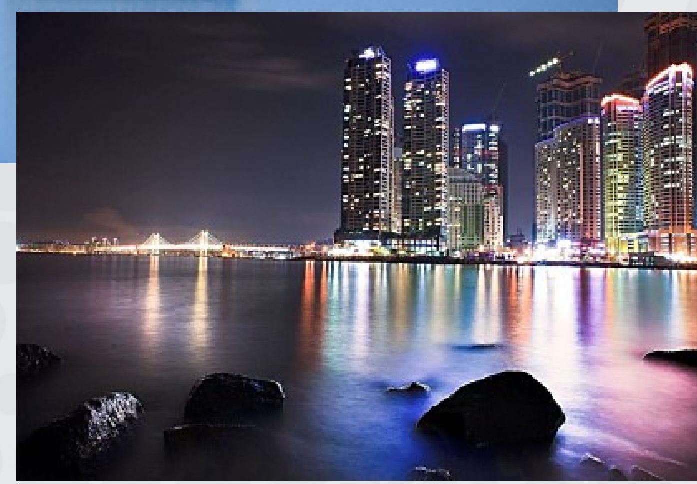


BCCA



Summary

- **Dynamical VS Statistical**
  - Simple and efficient
  - Stationarity
- **Basic statistics**
  - Normal, Gamma, and EV distributions
  - Quantile
- **PP VS MOS techniques**
  - Weather generator
  - MOS models
    - BCSD; BCCA
  - Long-term trend preserving methods



APCC
APEC CLIMATE CENTER