

Downscaling of weather within climate

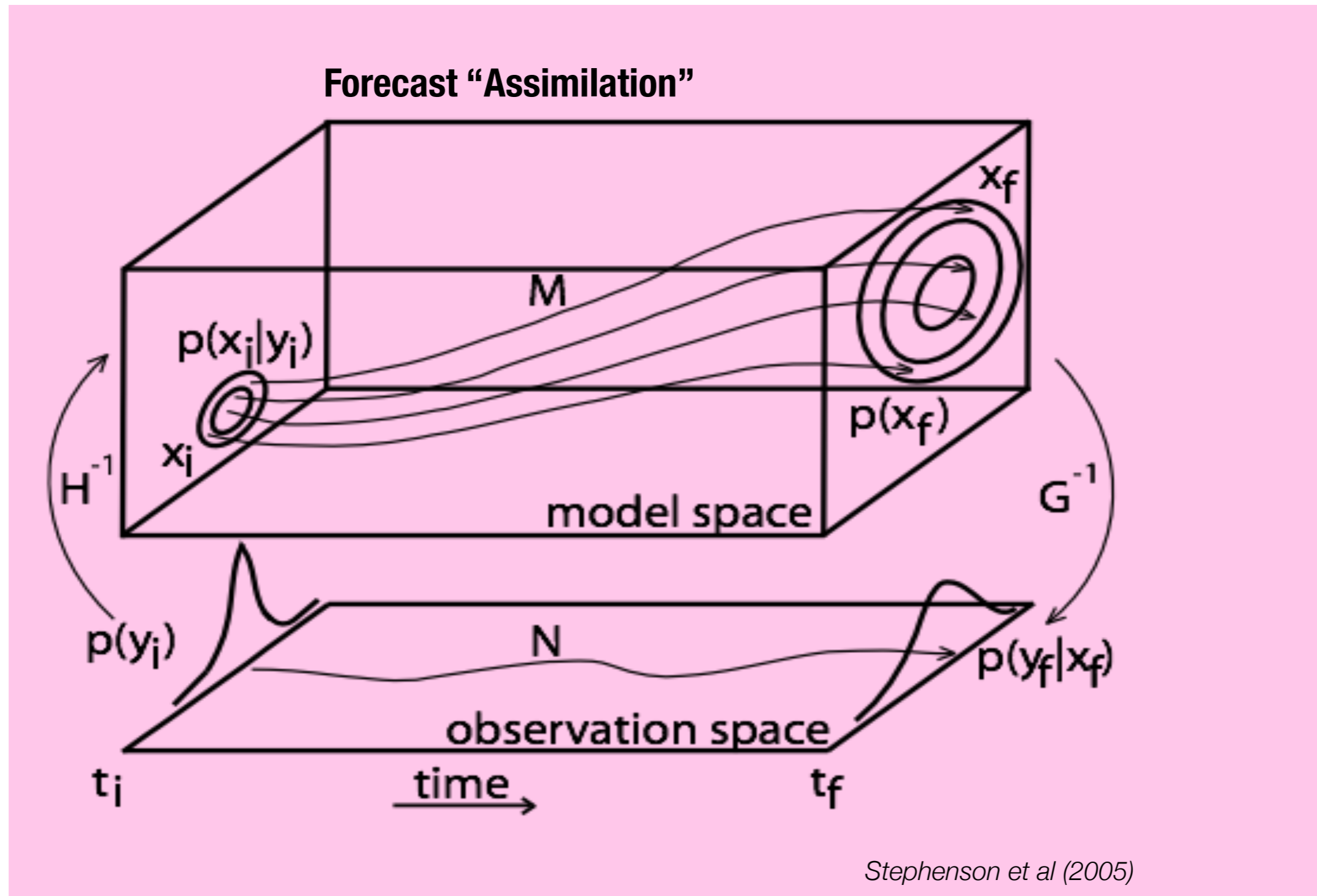
Andrew W. Robertson, IRI

International Research Institute for Climate and Society (IRI)
New York, USA



how to make tailored seasonal forecasts?

combining GCM forecasts and observational data

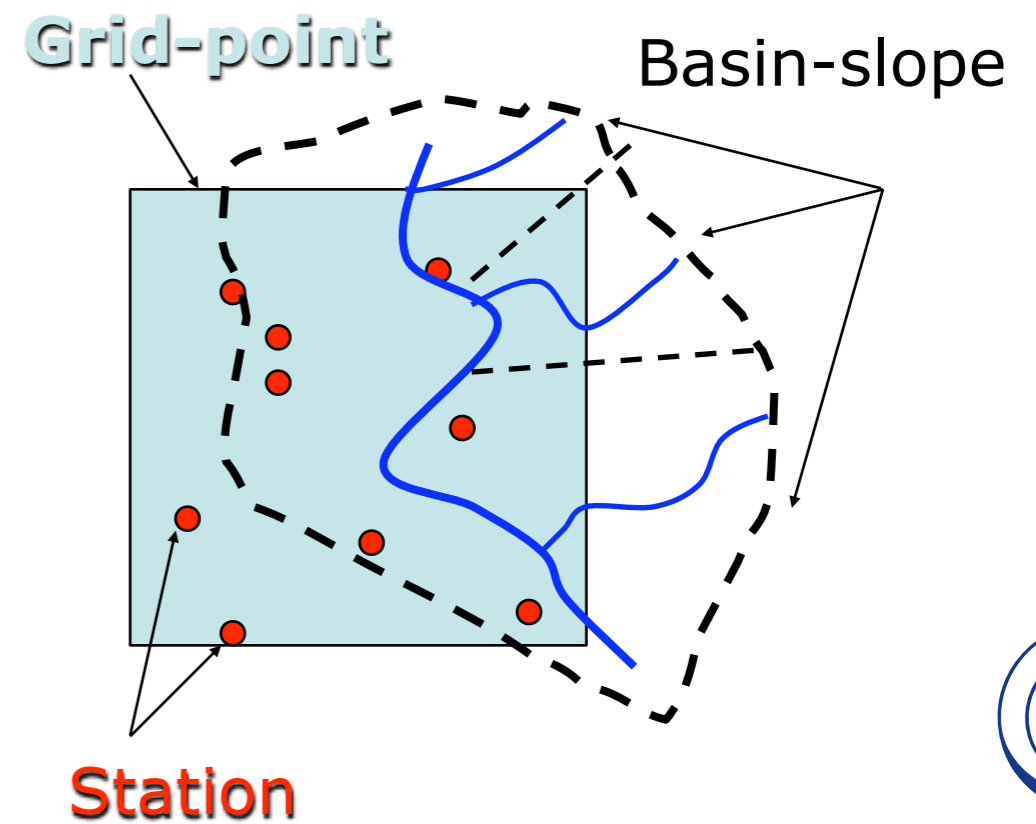


From weather forecasts to climate prediction...

- Weather forecast (up to few days)
- Medium-range forecast (1–2 weeks)
- Climate prediction (months to seasons)
 - in climate prediction, the weather needs a statistical treatment

... and from climate to weather : spatial and temporal disaggregation of the climate predictions

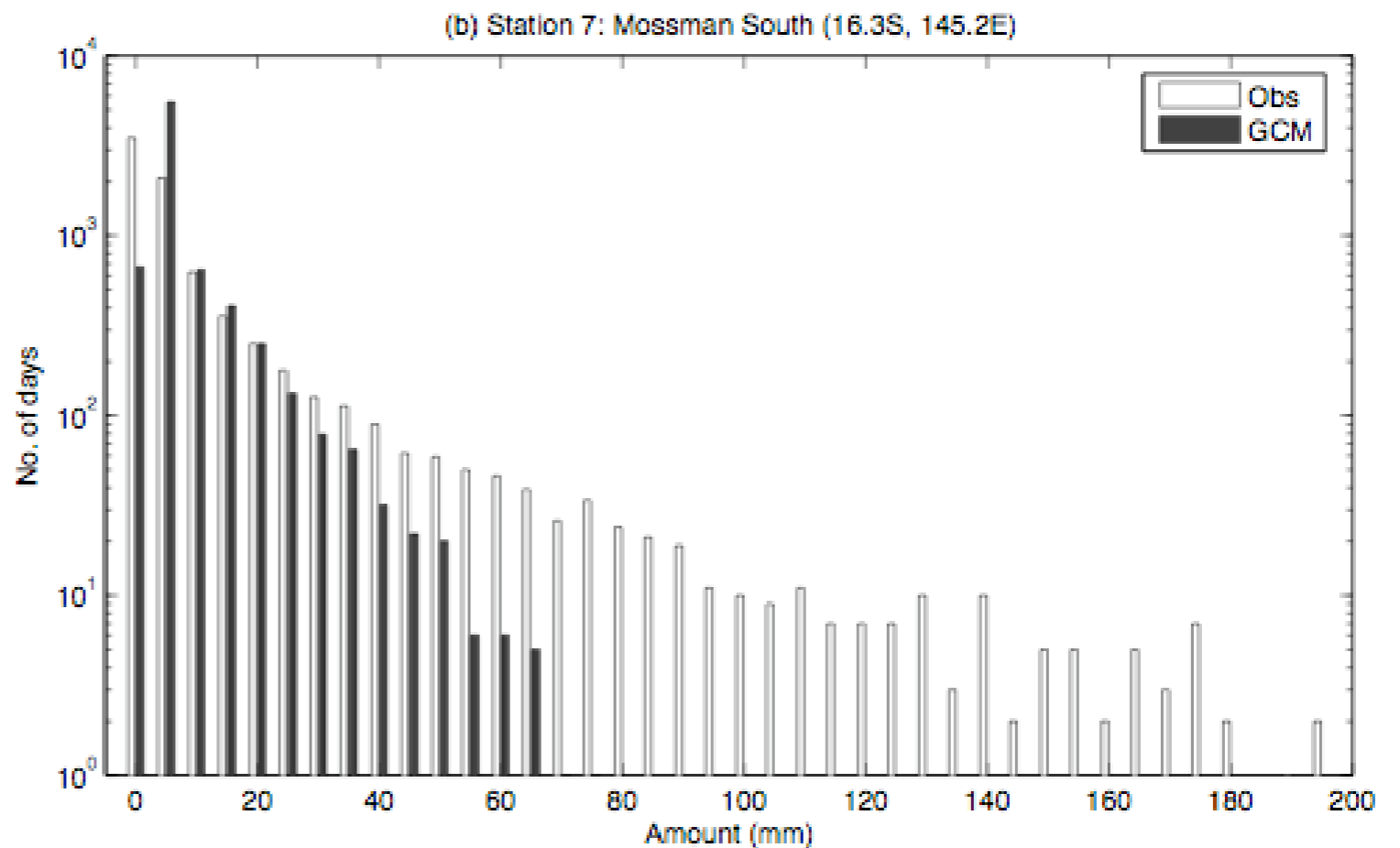
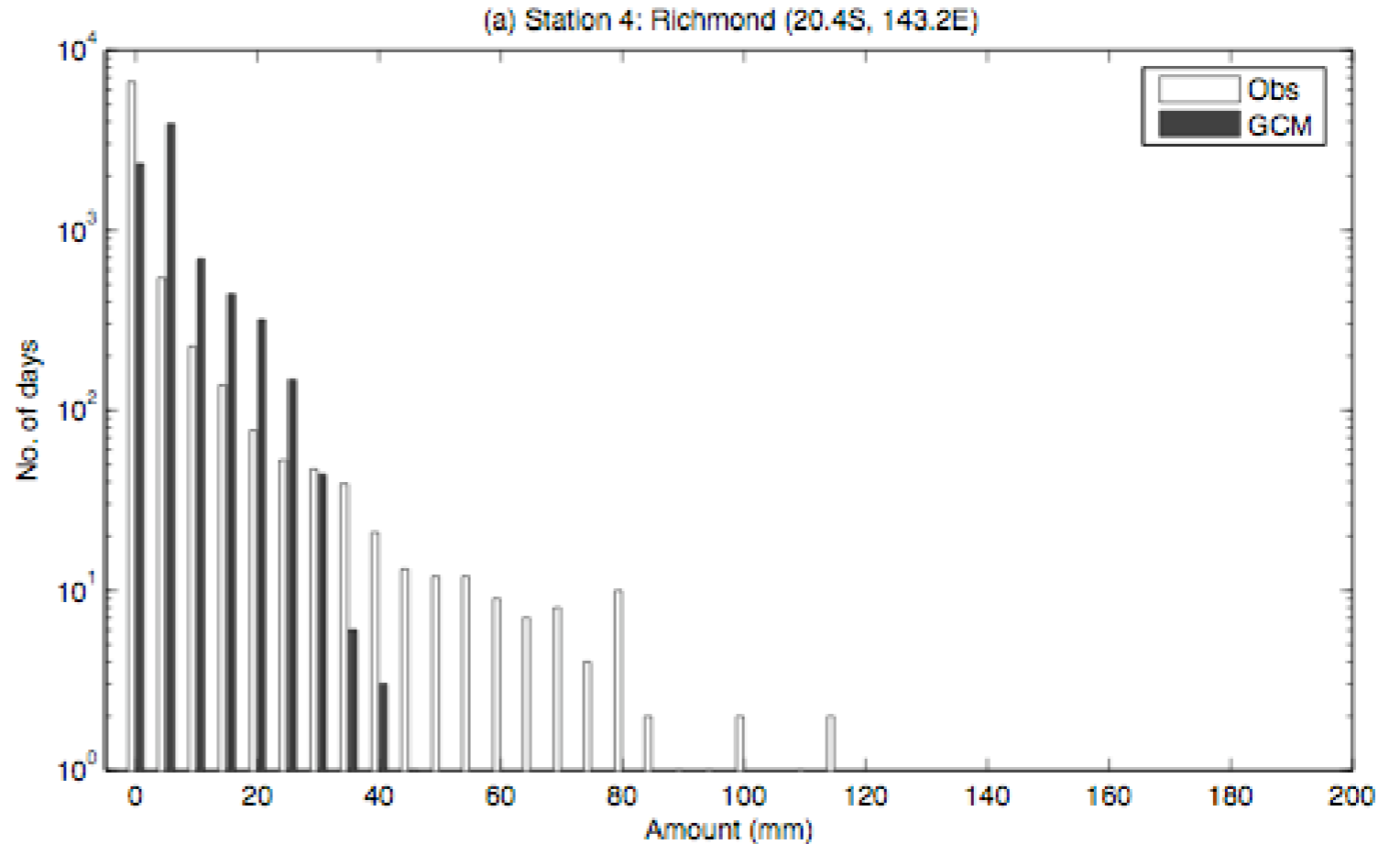
- goal is to extract useful information from seasonal forecasts, and it's often the weather (within the climate) that counts
 - agriculture: crops are very sensitive to dry spells (and potentially wet spells)
 - flood risks for reservoir management
 - malaria transmission may be more sensitive to distribution of rainfall within the season (e.g. puddling vs. downpours)
- beyond seasonal averages
- predictability of the “second kind”, because we are not trying to predict what will happen on a particular day (Lorenz)



Example of GCM vs. Station Daily Rainfall Distributions

(Queensland in Summer)

... need for calibration



two basic ways to make seasonal predictions of weather statistics

1. Regression (i.e. MOS) of seasonal means using the statistic of interest as the predictand:

$$\mathbf{y} = \mathbf{Ax} + \mathbf{e}$$

(get probability distribution from error properties of best estimate prediction)

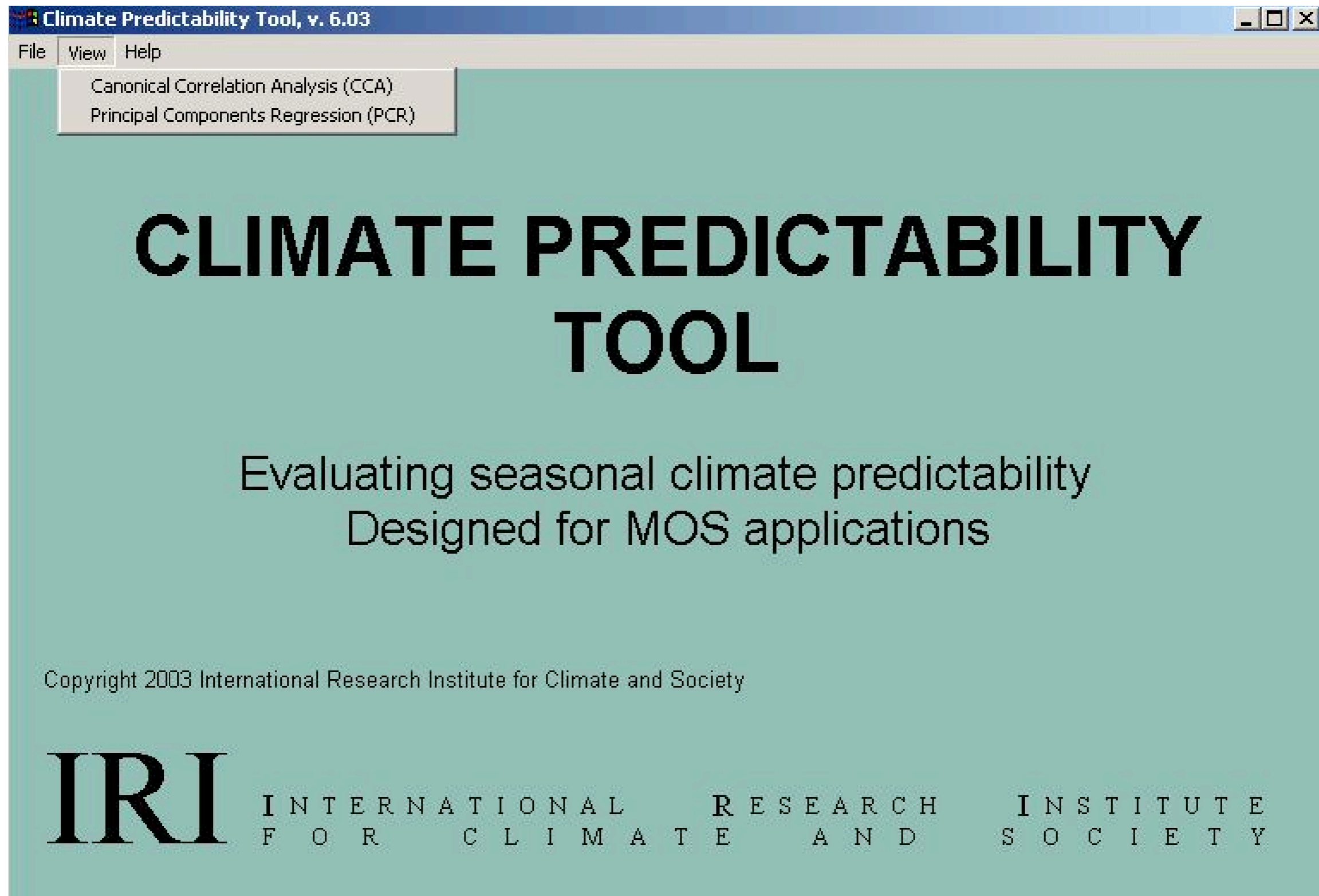
where \mathbf{y} is a seasonal-average quantity:

- i. number of dry days at a station
 - ii. number of dry spells of 10 days or longer
- train the regression model using observed estimates of \mathbf{y} , and GCM hindcasts \mathbf{x} . Then, given a GCM forecast \mathbf{x}_f , we can predict $\mathbf{y}_f = \mathbf{Ax}_f$ directly.

2. Use a method to obtain daily rainfall sequences from the GCM forecast. Then simply construct the needed statistic from the predicted daily sequences. This is sometimes called *temporal downscaling*.



seasonal averages: “CPT”



Principal Components Regression

PROJECT:

Explanatory (X) variables:

Response (Y) variables:

Training data file:

Training data file:

X input file:

HULM-OUG_OND5002prc1

browse

Y input file:

SST_OND1950-08.tsv

browse

Number

First year

First year

Minimum

Maximum

Climate Predictability Tool, v. 9.10 - Results Window

File Tools Customise Help

Project:

Progress: 100%

Actions:

Checking for missing values ...
Data read successfully

Beginning analysis ...
Calculating climatologies and thresholds ...
Optimizing cross-validation ...
Training period: 1950-1980

CUR

Number of No

1

2

3

4

Constructing model
Identifying categories
Done!

Screen plots

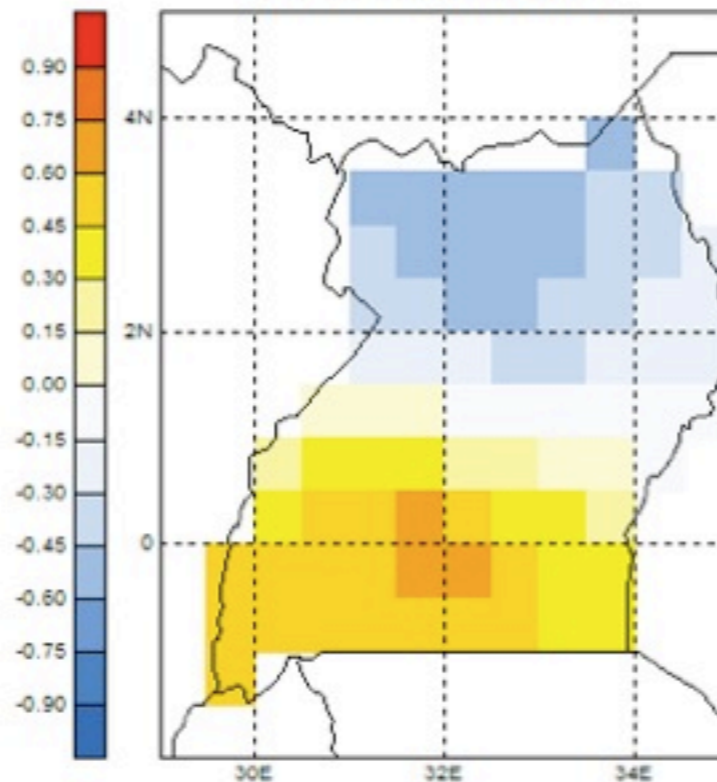
X EOFs Scree Plot



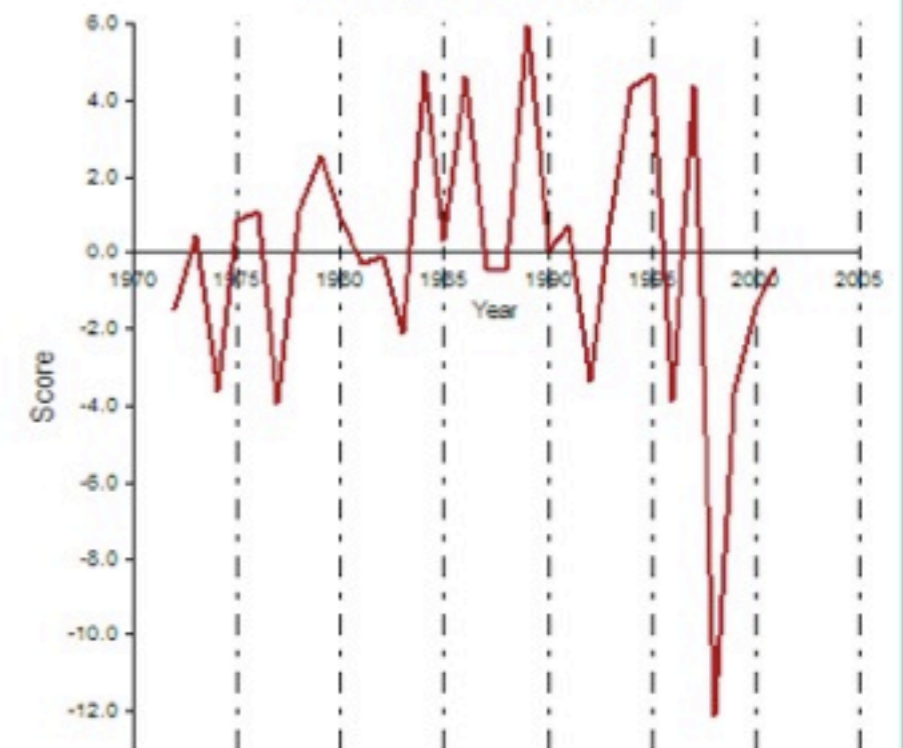
X EOFs

EOF: 29

X Spatial Loadings (EOF2)



X Temporal Scores (EOF2)



Statistical Downscaling Perfect Prognosis v MOS

Perfect Prognosis

Step 1: Establish relationship between *observed* large-scale field (e.g temperature) and *observed* high resolution target variable pattern

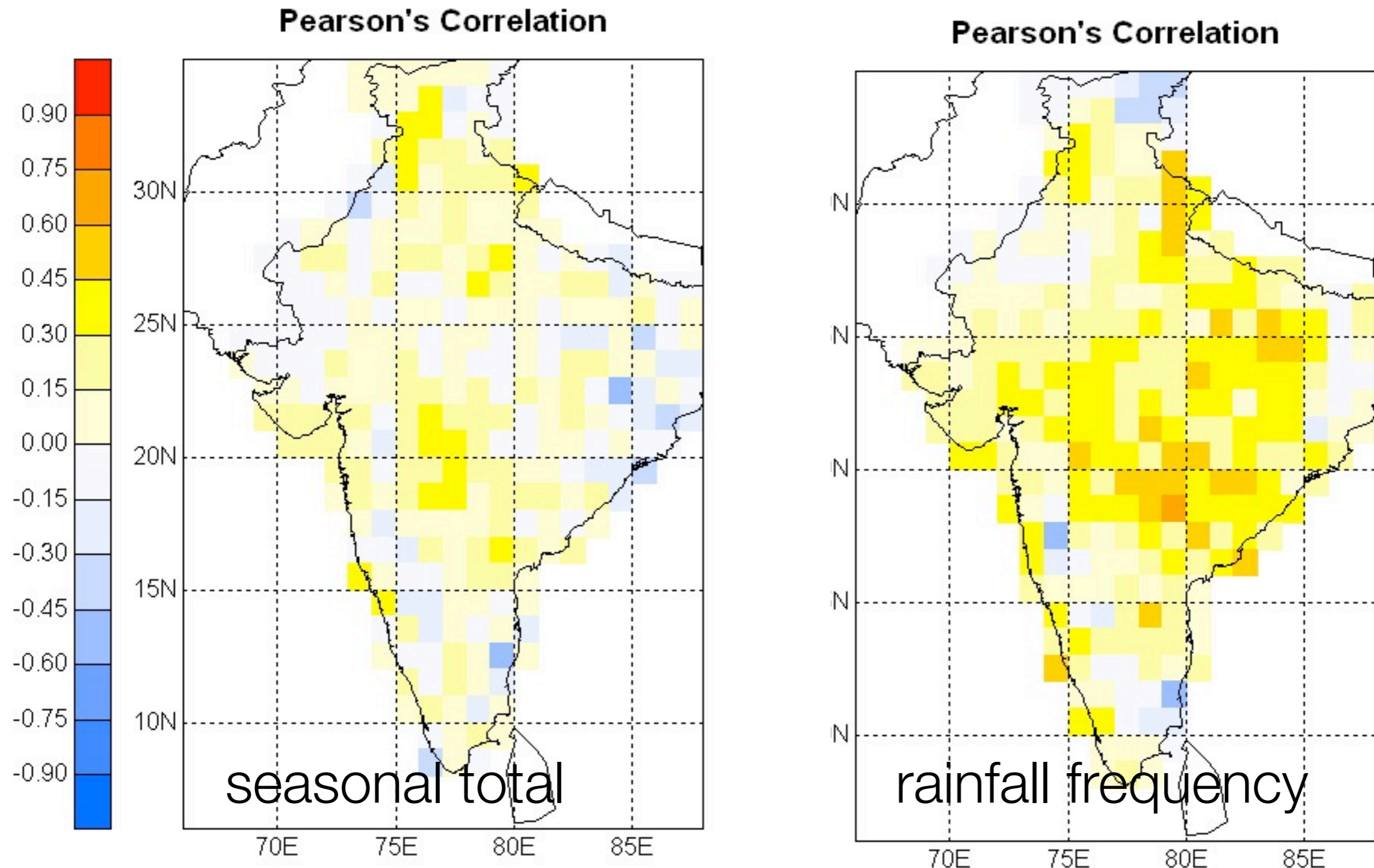
Step 2: Take output from GCM and apply the above statistical relationship to predict high resolution target variable

Model Output Statistics (MOS) Correction

Establish statistical relationship (say over 1951-2000) between *GCM predicted* large-scale field (e.g. low-level winds) and *observed* target variable at required scale

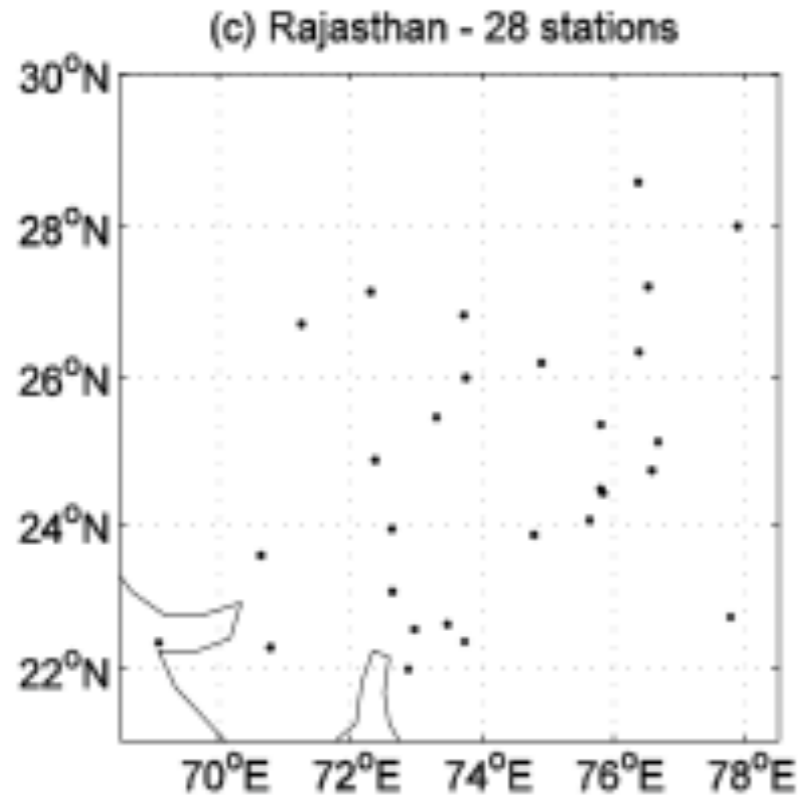
seasonal regression of rainfall statistics with CCA

Indian monsoon rainfall



JJAS rainfall correlation skill ECHAM4-CA: made from June 1

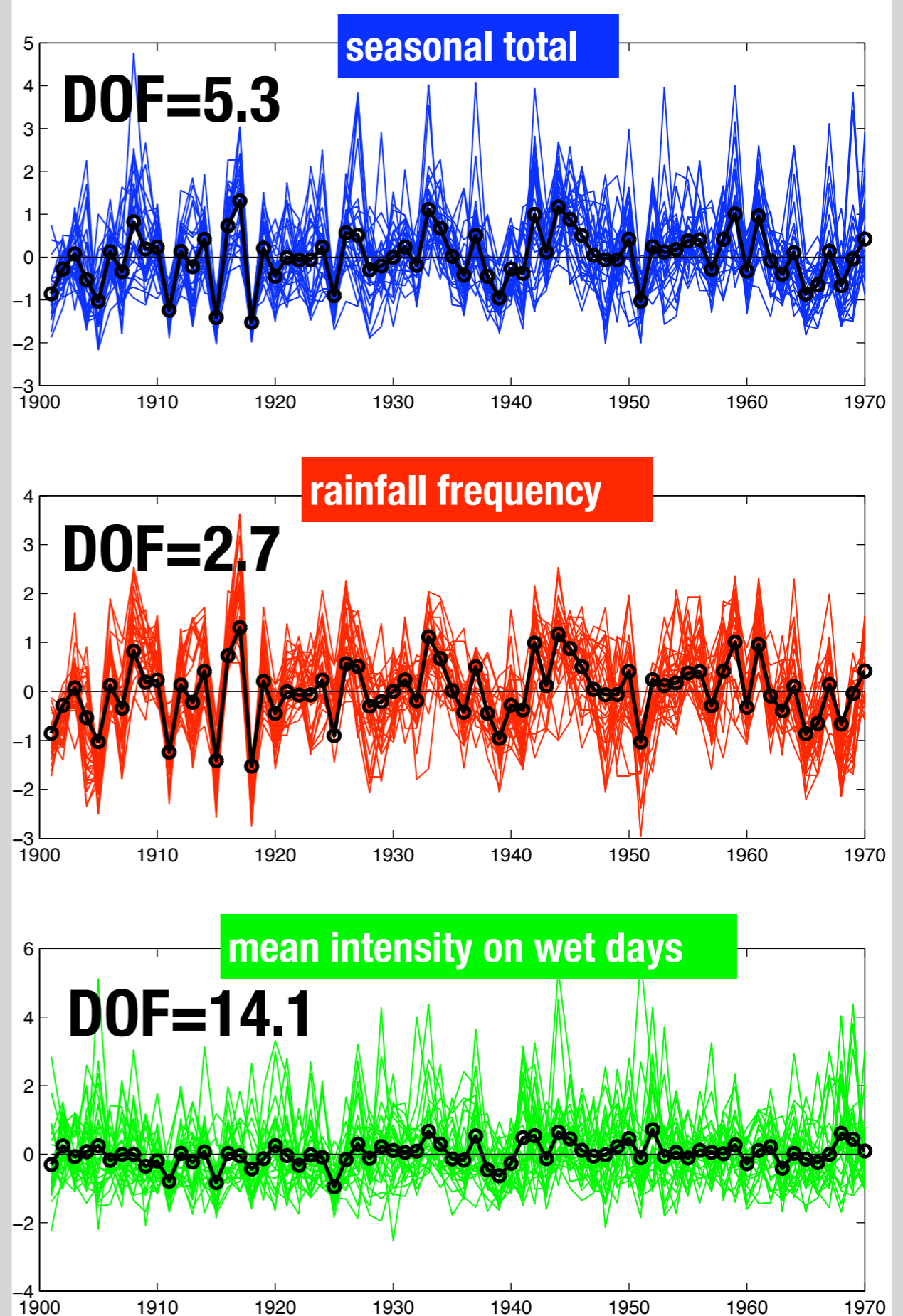
seasonal rainfall anomalies at individual stations



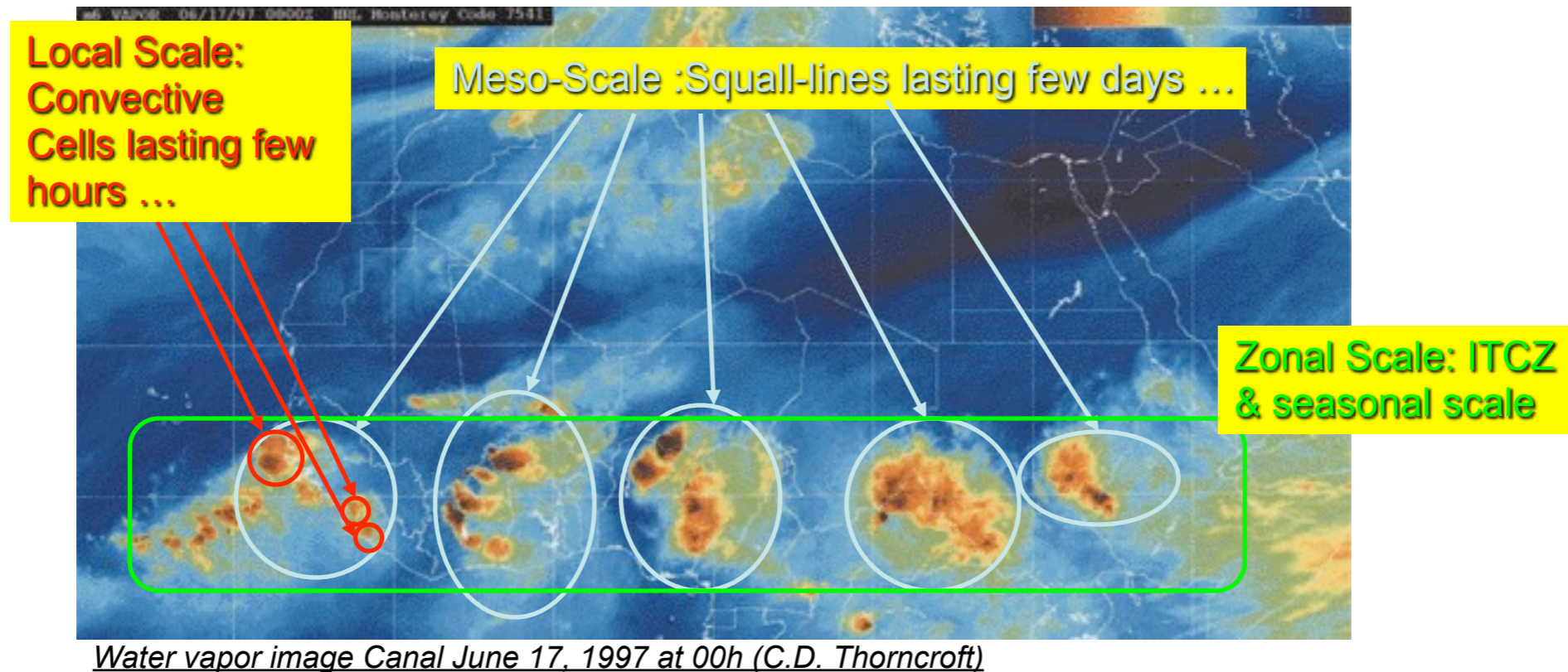
Jun-Sep
1900-70

seas. amount =
(no. of wet days) x (mean intensity on wet days)

Rainfall at individual stations and station-average



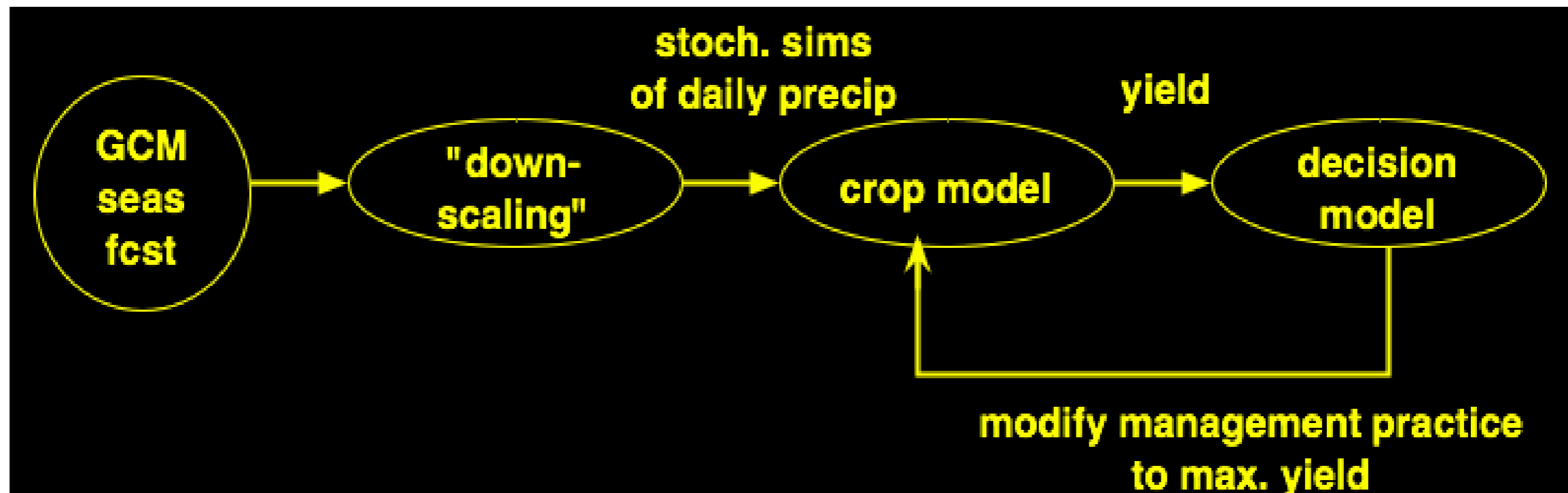
why are seasonal anomalies of rainfall frequency more coherent than intensity?



- intense convection is very fine scale: may hit or miss a raingauge
- organization of rainfall is larger scale
- climate forcings integrate across a season, preferentially acting on occurrence

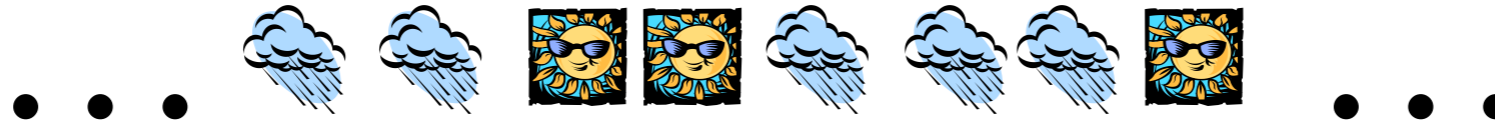
GCM-based stochastic daily weather sequences for climate risk management

.. crop models need daily time sequences



.. as do malaria models and hydrologic models

Approaches to temporal downscaling



1. Statistical transformation of daily GCM output
 - Local scaling
 - Quantile-quantile mapping
2. Historical analog techniques
 - Use various subsets of past data based on a seasonal-mean predictor(s), or even daily GCM output
3. Stochastic weather generators
 - Parameters estimated from seasonal (or monthly) GCM predictions
 - Hidden Markov model

Example of a Weather Generator

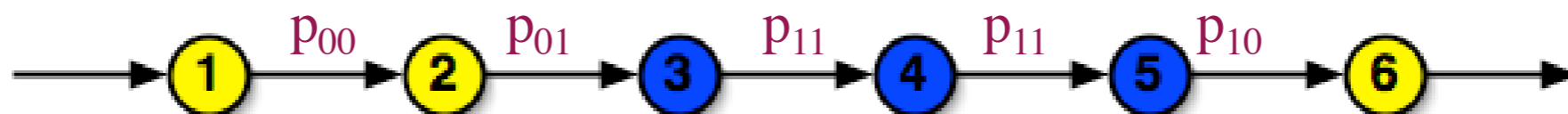
variable		model	parameters
PREC (precipitation)	occurrence	Markov chain (order=1)	p_{11}, p_{01} [monthly]
	amount	Gamma distribution	μ, σ [monthly]
SRAD (solar radiation) TMAX (max. temperature) TMIN (min. temperature)		AR model: $x^*(t+1) = Ax^*(t) + Be$	A, B: 3x3 matrices + 3 × (wet/dry) × (avg's/std's)

$$x^* = \frac{x - m(x)}{s(x)}$$

where: $m(x)$, $s(x)$ depend on day of the year and precipitation occurrence

Dubovsky et al. (2001)

- Markov process used to model persistence (“today depends on yesterday”)
- Solar radiation and Tmax are likely to be lower on wet days, so precipitation is usually chosen as the driving variable with others conditioned on it
- Daily-total precipitation is often exactly zero, so most weather generators treat occurrence and amount separately
- Simulation proceeds by drawing random numbers between 0 and 1



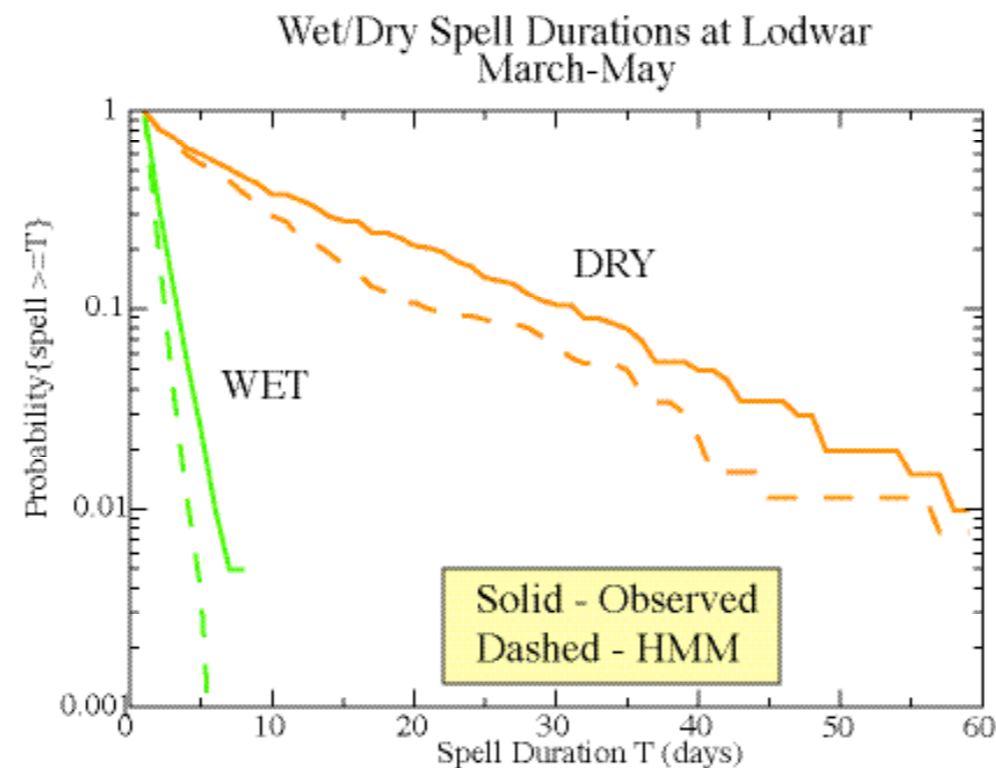
Markov chain:

- Wet- and dry-spell lengths (T days) follow a geometric distribution

$$\Pr\{\text{dryspell} = T\} = p_{01}(1 - p_{01})^{T-1}, \quad T = 1, 2, 3, \dots$$

$$\Pr\{\text{wetspell} = T\} = p_{10}(1 - p_{10})^{T-1}, \quad T = 1, 2, 3, \dots$$

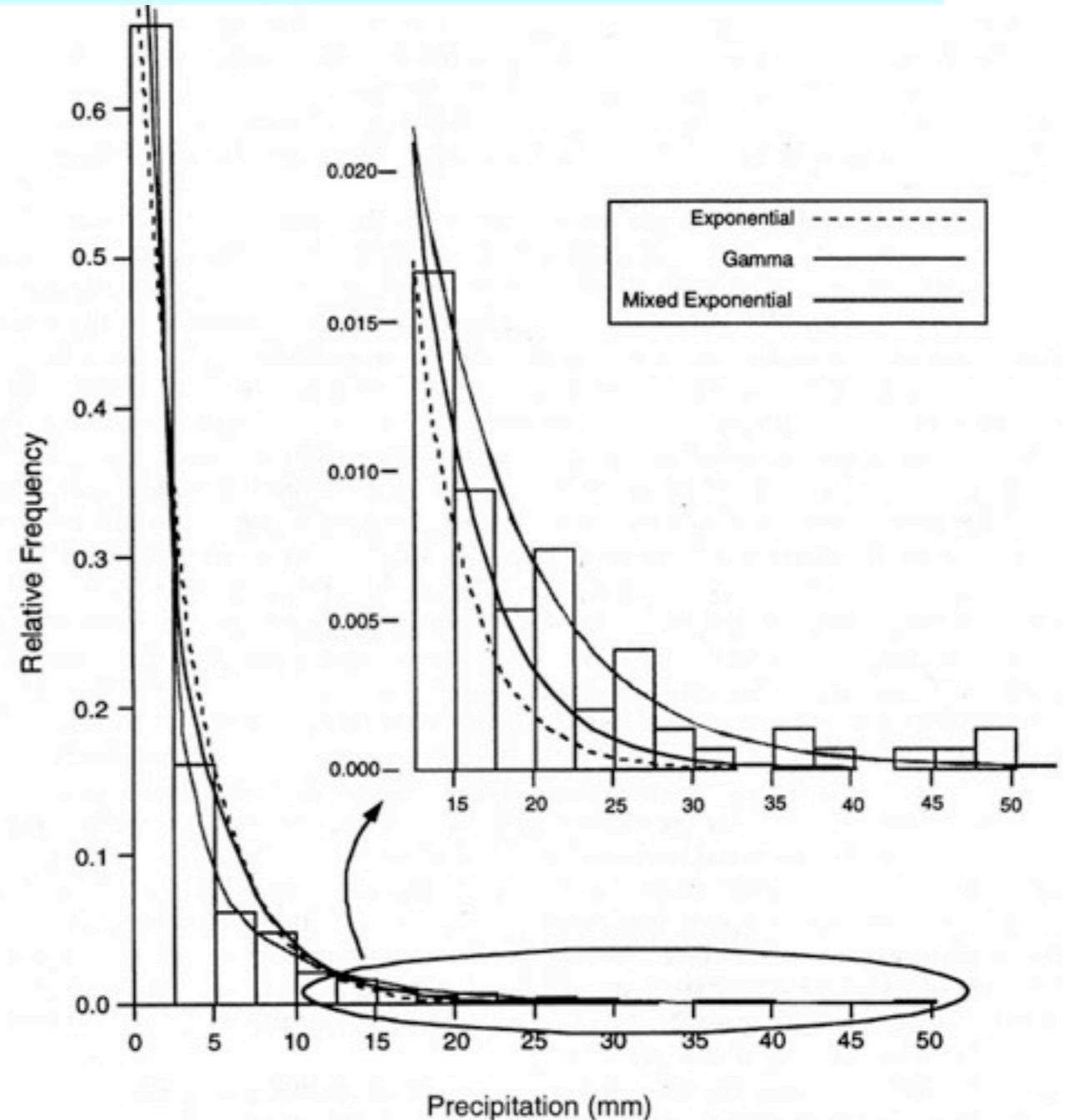
(just the multiplicative law of probabilities for independent events:
probability of “success” is multiplied by a sequence of $T-1$ consecutive failures to transition out of spell)



(Note $1-p_{01}=p_{00}$
and $1-p_{10}=p_{11}$)

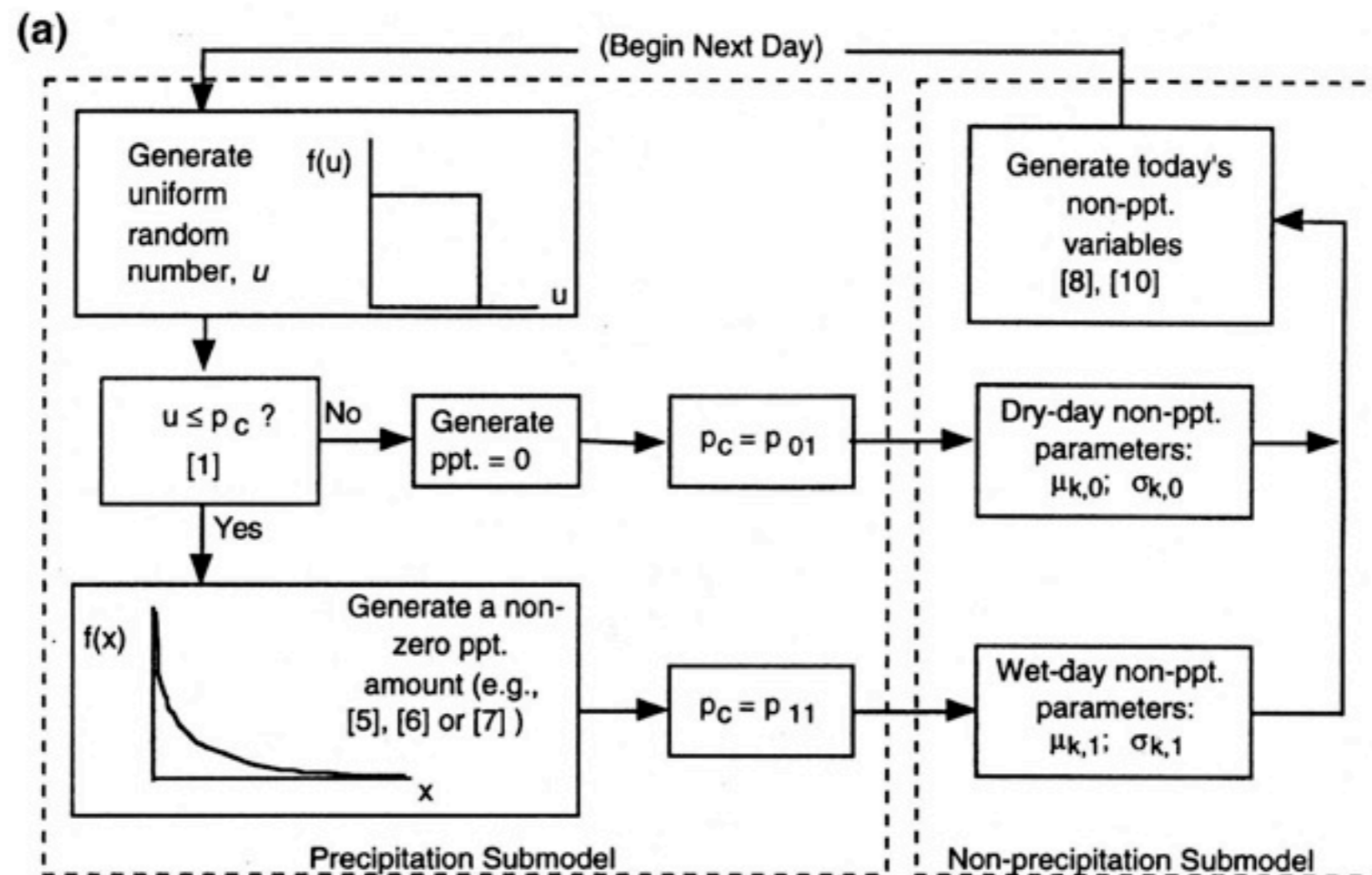
Precipitation Amount

- Distribution of daily amount is strongly skewed to the right
- Exponential is simplest reasonable model. It requires only one parameter μ , yet reproduces the strong positive skewness
- Two-parameter gamma distribution is most popular choice (shape α and scale β)
 - $\alpha=1$ yields exp distrib, while the extra flexibility improves the fit



Simulation

- Draw a random number $u \in [0,1]$.
- If day $t-1$ was dry, then day t is simulated to be wet if $u < p_{01}$..

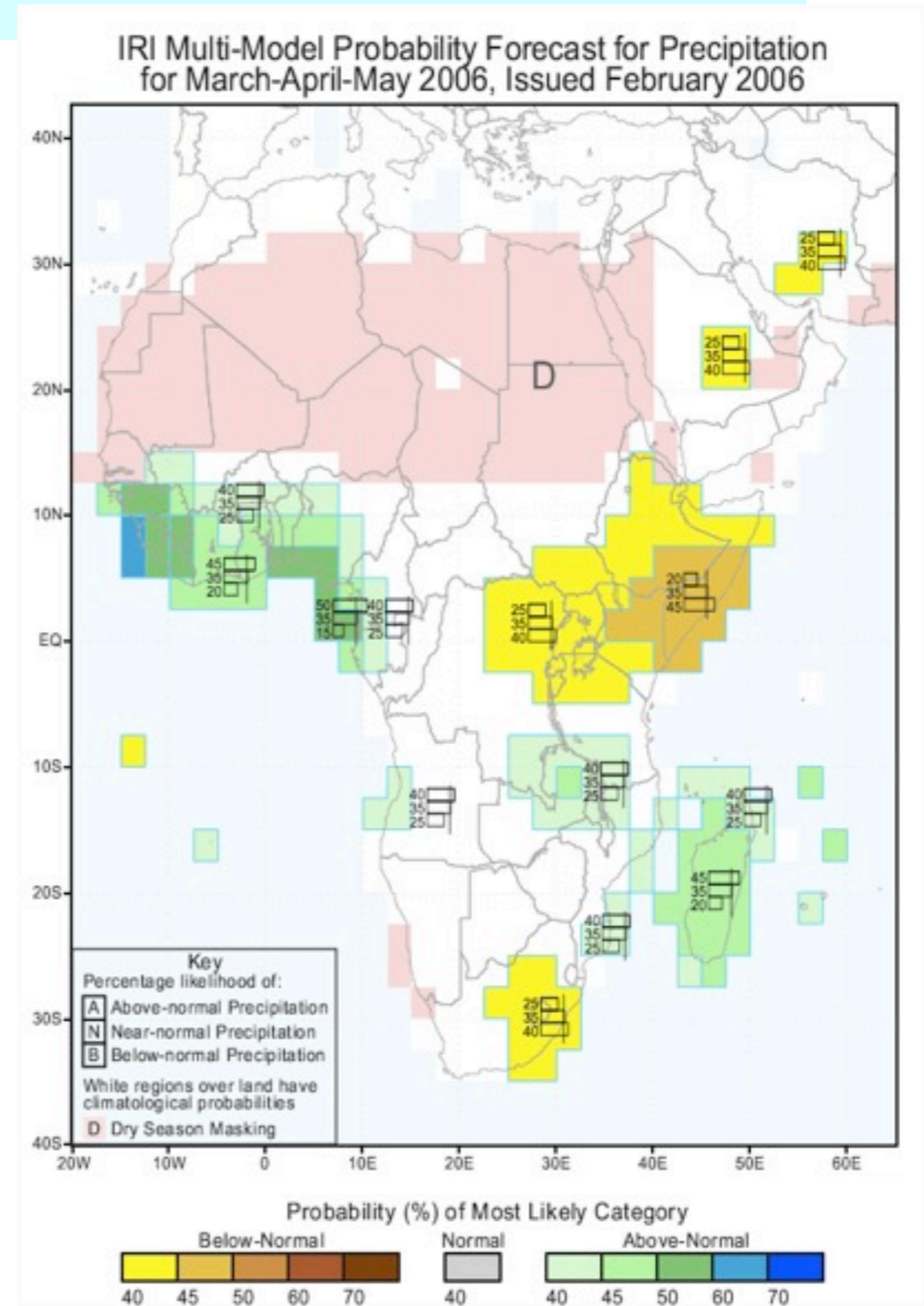


WGs for downscaling seasonal forecasts

- Use climatological parameters
- Then *rescale* the generated daily values such that their *monthly means* exactly match the monthly GCM prediction (Hansen)
 - Additive offsets for temperatures & insolation parameters
 - Multiplicative adjustment for precip., with repeated (stochastic) generation to match target

Realizations of daily weather in forecast seasonal climate

- Forecasts are often in terms of tercile-category probabilities $\{p_B, p_N, p_A\}$
- Could make WG parameters depend on these probabilities



- Forecast parameters can be estimated from the historical record (Briggs & Wilks 1996, Wilks 2002) by:
 1. computing their values averaged over *all* years in each tercile of the local precip record
$$\{\bar{\pi}^{(B)}, \bar{\pi}^{(N)}, \bar{\pi}^{(A)}\}$$
 2. and weighting according to the forecast $\{p_B, p_N, p_A\}$

$$\pi(p_B, p_N, p_A) = p_B \bar{\pi}^{(B)} + p_N \bar{\pi}^{(N)} + p_A \bar{\pi}^{(A)}$$

- Similar approach can be taken to condition WG parameters on ENSO phase (Woolhiser et al. 1993)

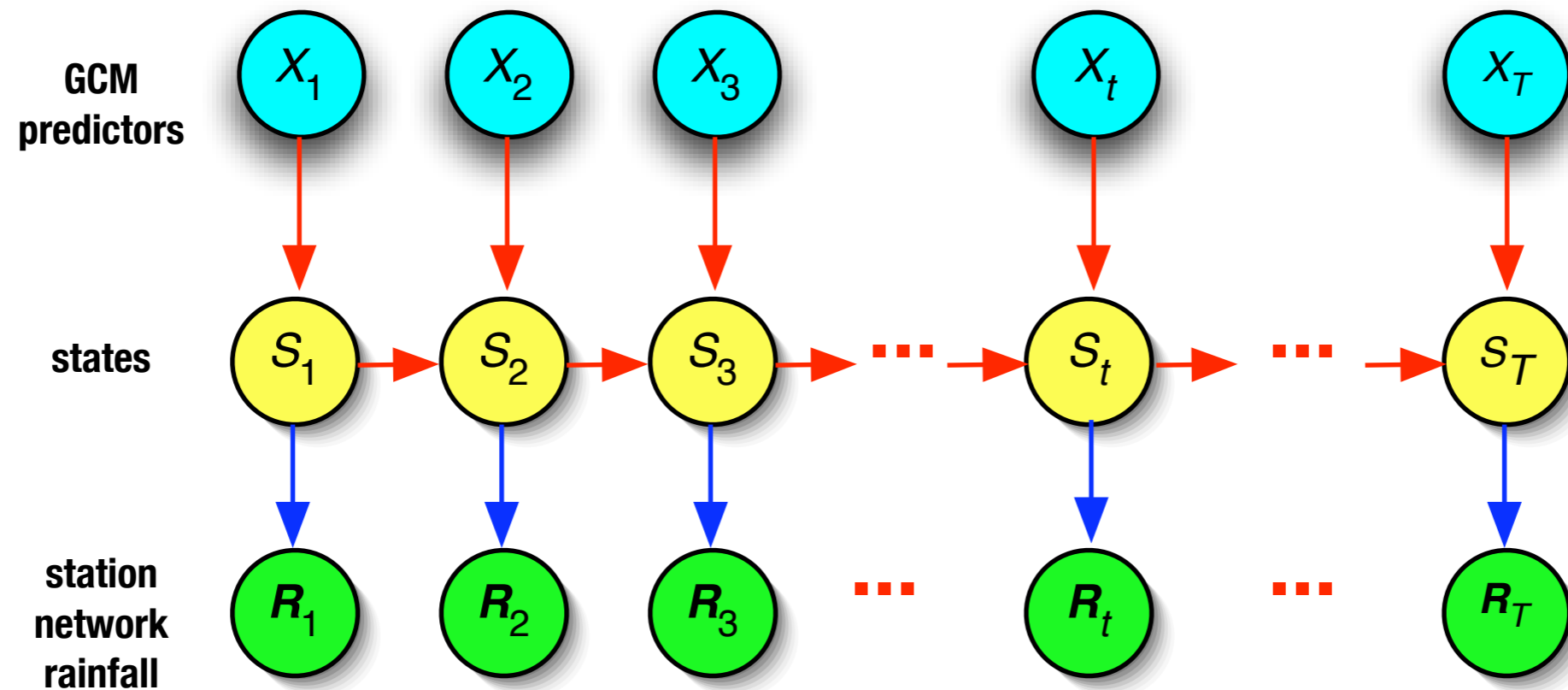
Multi-site extension

- Run a series of WG's in parallel
- Use spatially correlated random numbers (Wilks, 1998)
- Use a Hidden Markov Model

Downscaling with a NHMM

- A Markov chain of discrete atmospheric states (rather than precip occurrence as in the WG)
- Extra state “layer” is identified with physical atmospheric states
 - Assumes that rainfall depends only on the weather state
 - Dynamical controls influence these states, I.e. *non-stationary transition probabilities*
- We can either (a) take the predicted daily sequences of X from an ensemble of GCM predictions, or (b) seasonal-mean values (and repeat them daily)

downscaling daily weather sequences with a Non-homogeneous Hidden Markov Model



Transition probabilities modulated by X

Rainfall is conditionally dependent on the weather state

.. daily sequence of rainfall vectors

daily rainfall states

1

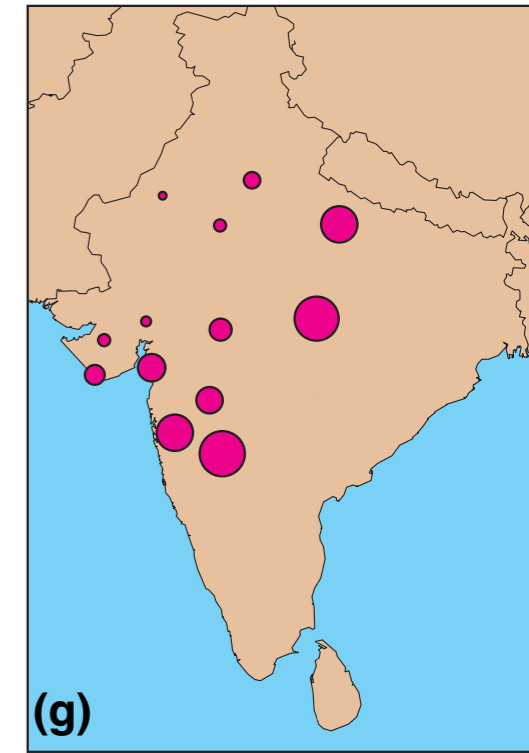
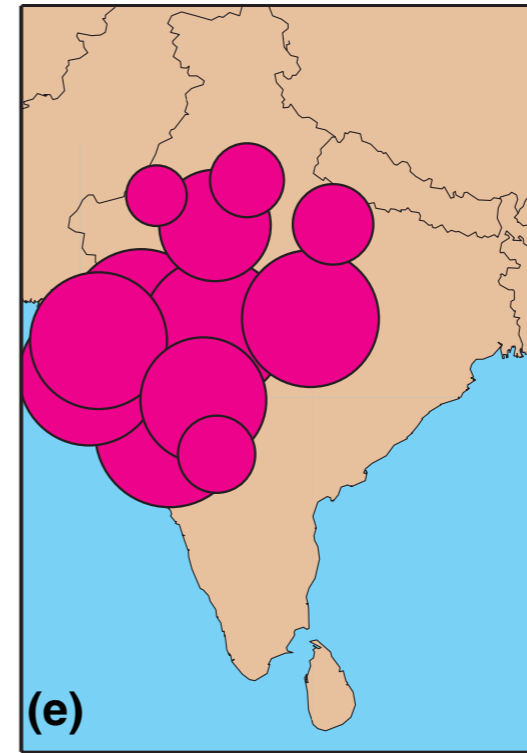
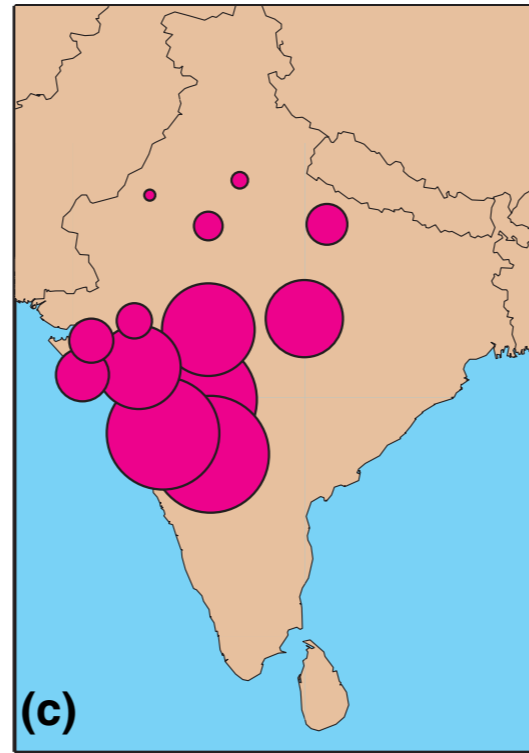
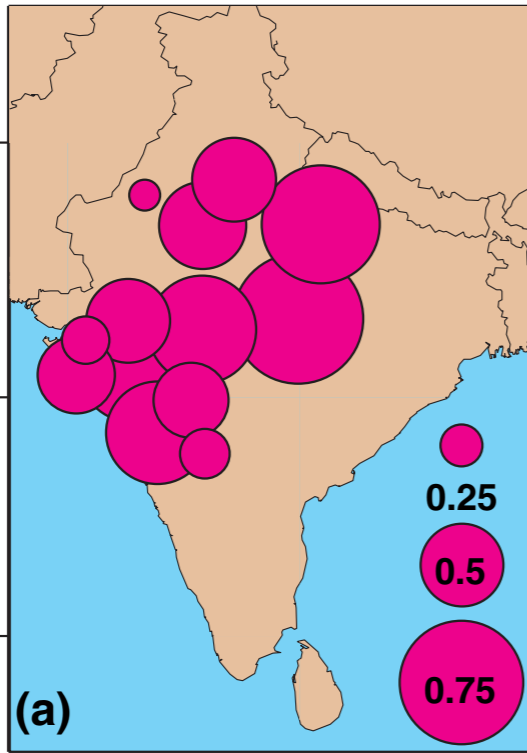
2

3

4

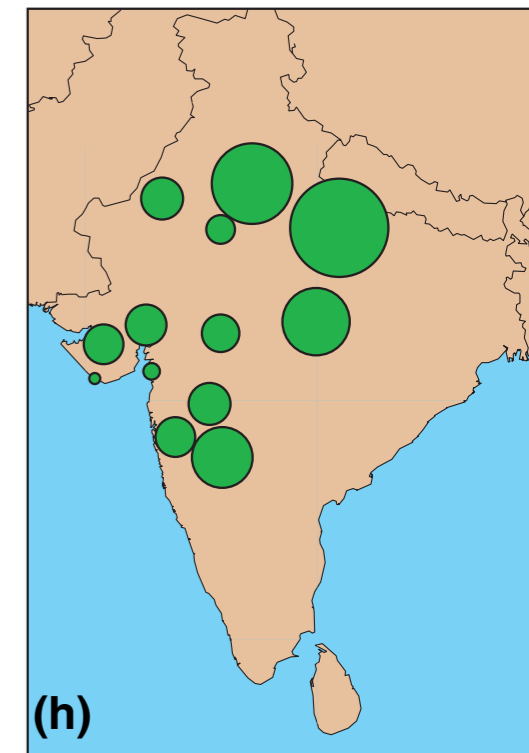
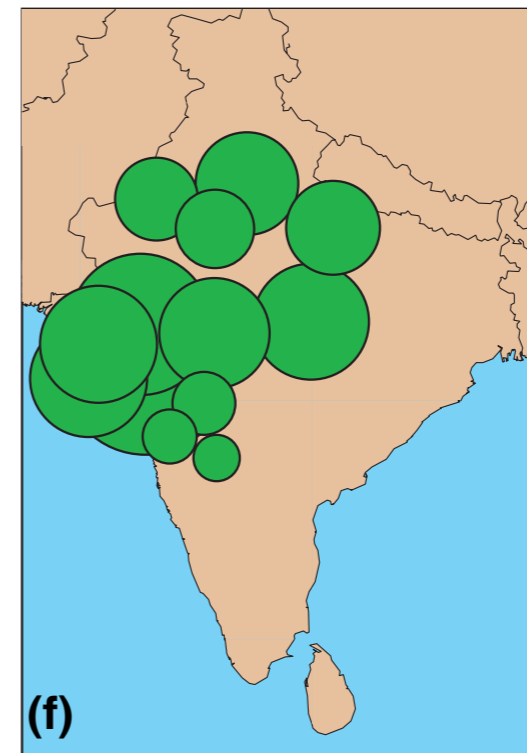
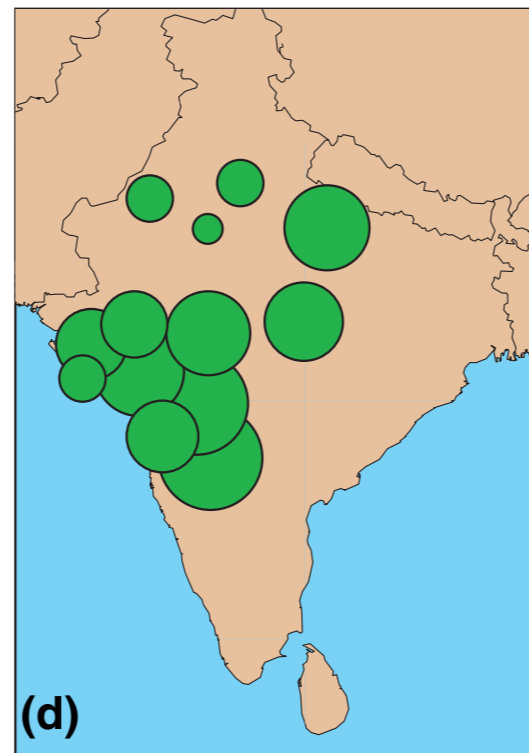
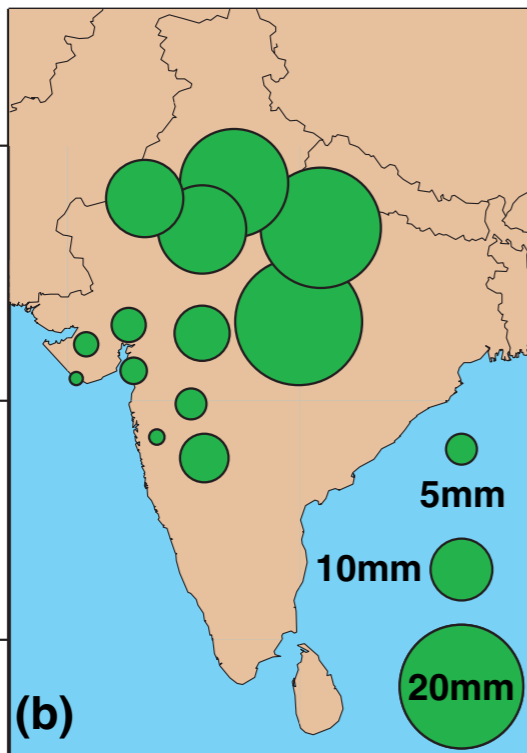
Daily
Rainfall
Probability

30°
20°
10°



Rainfall
Amount
on Wet
Days

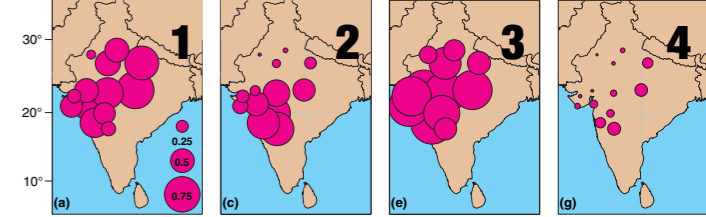
30°
20°
10°



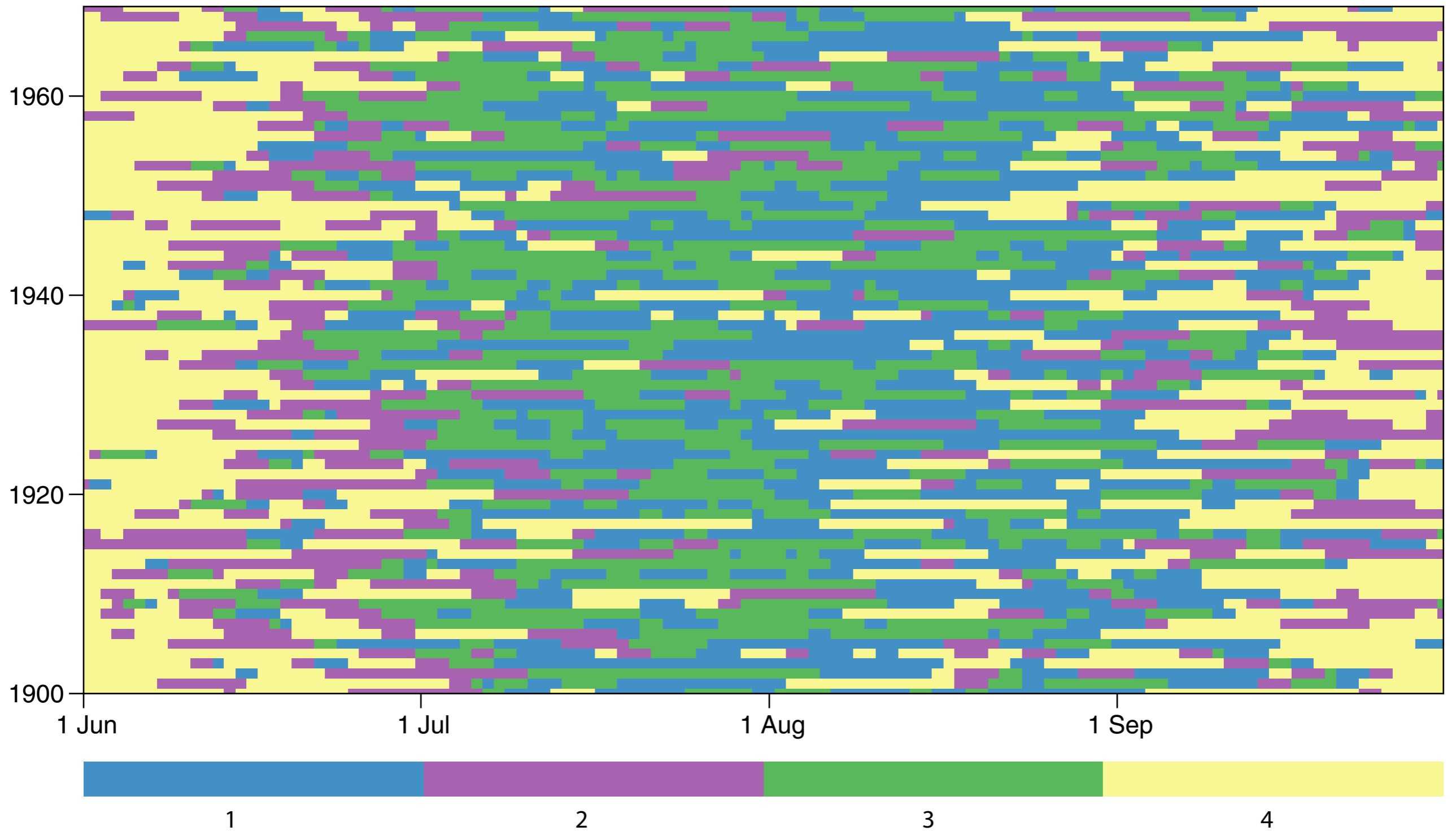
(d)

(f)

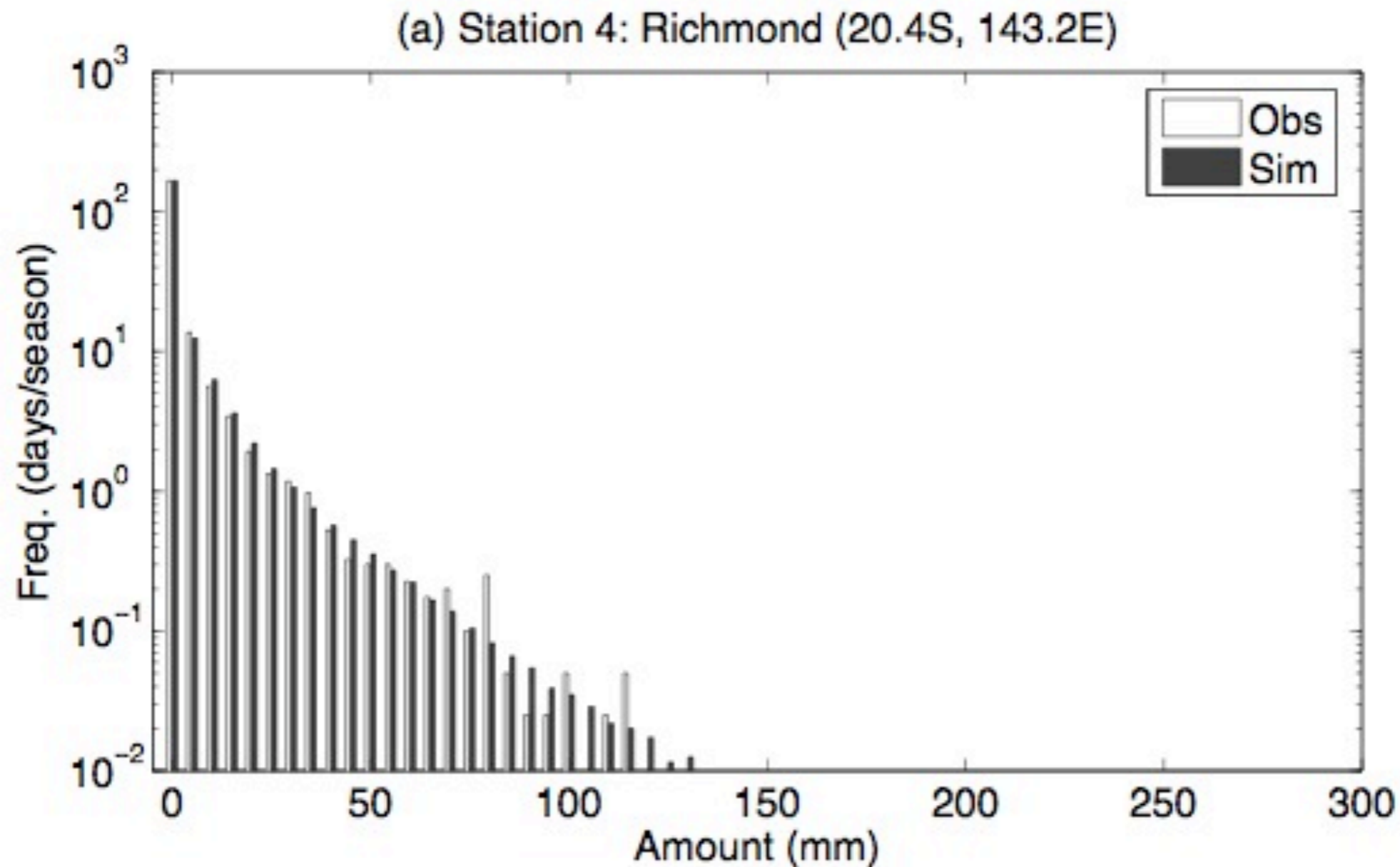
(h)



temporal sequence of states



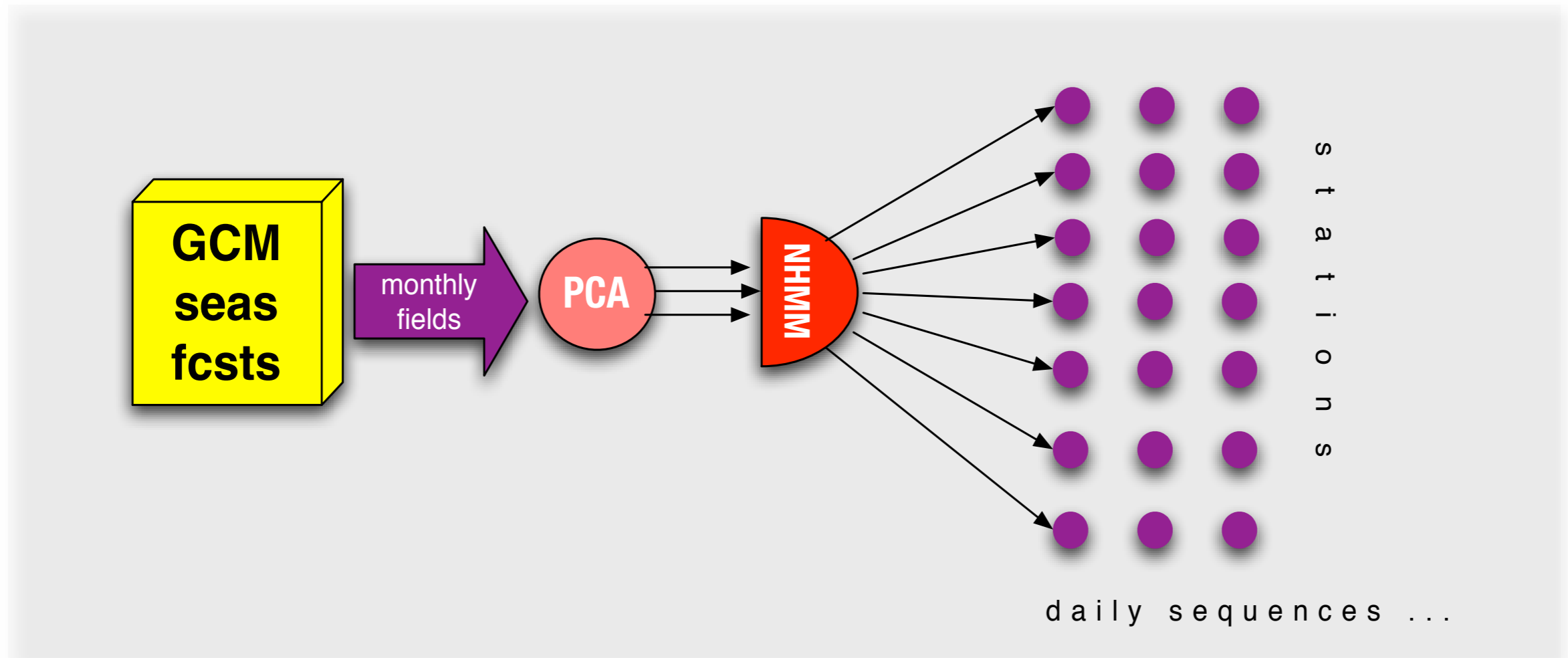
Rainfall amount distribution obtained from an HMM



From Queensland Australia (Oct–Apr)

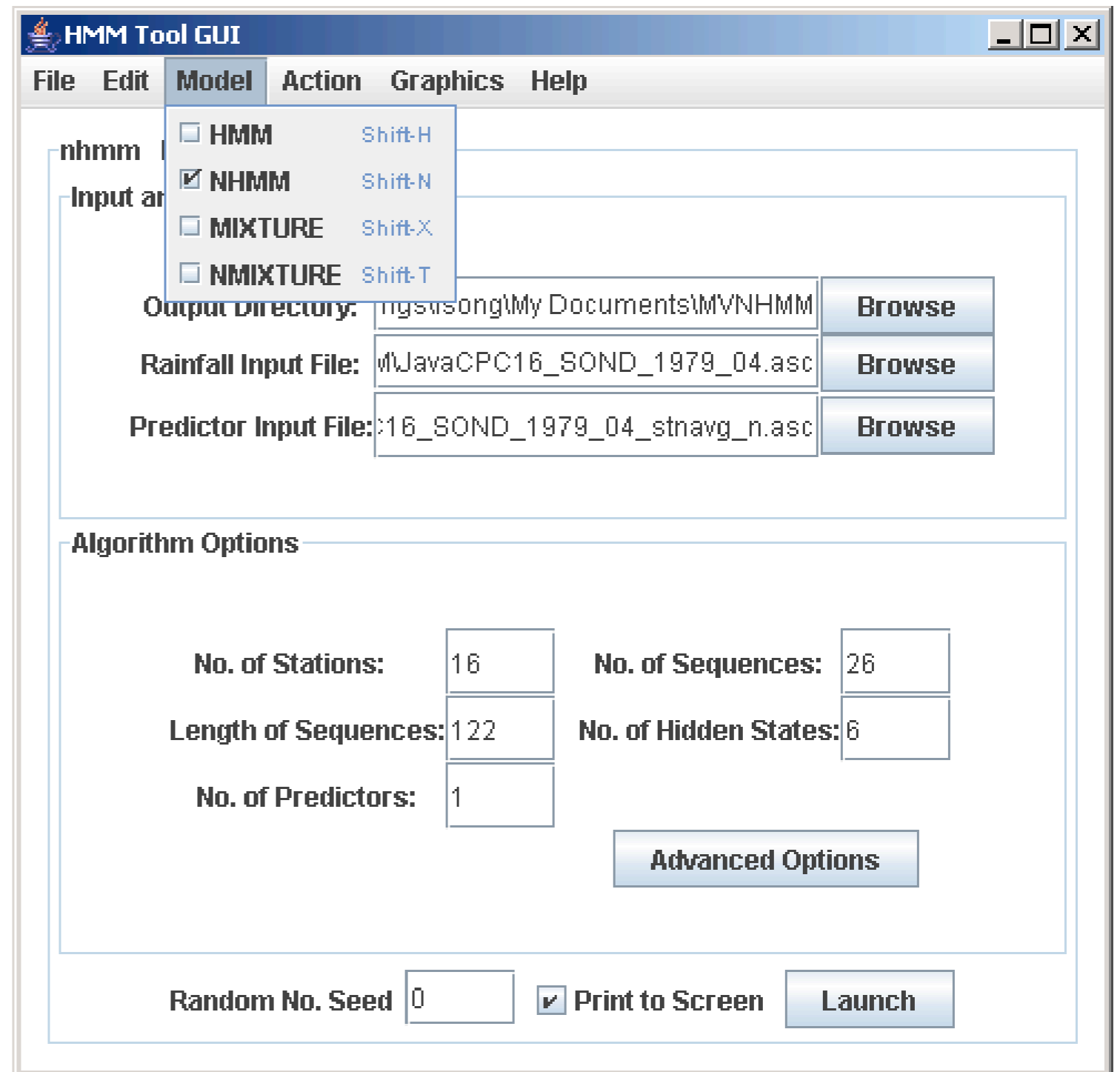
Non-zero amounts modeled by mixed exponential distribution

conditioning stochastic daily weather sequences on GCM seasonal forecasts

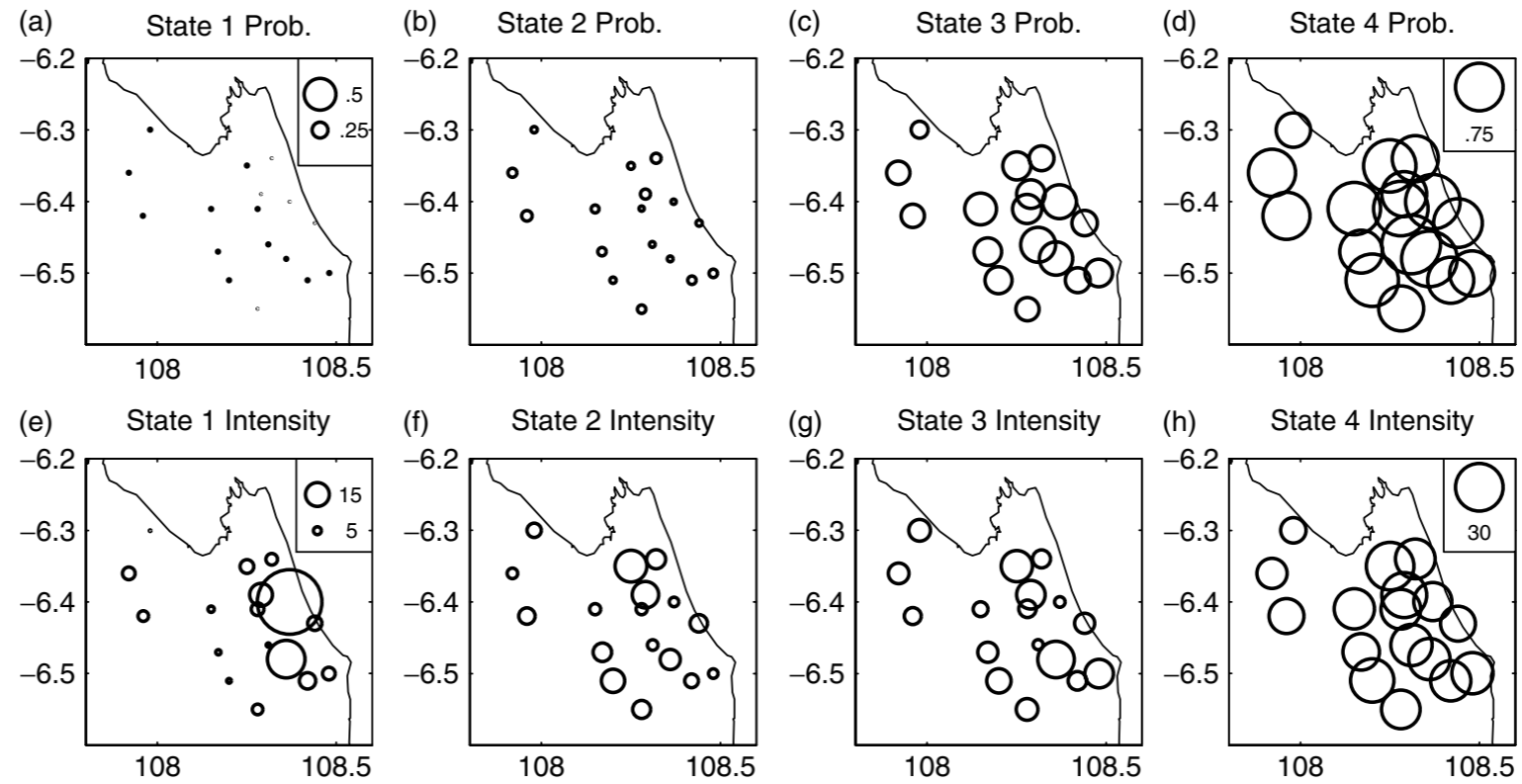


toolboxes for downscaling in time

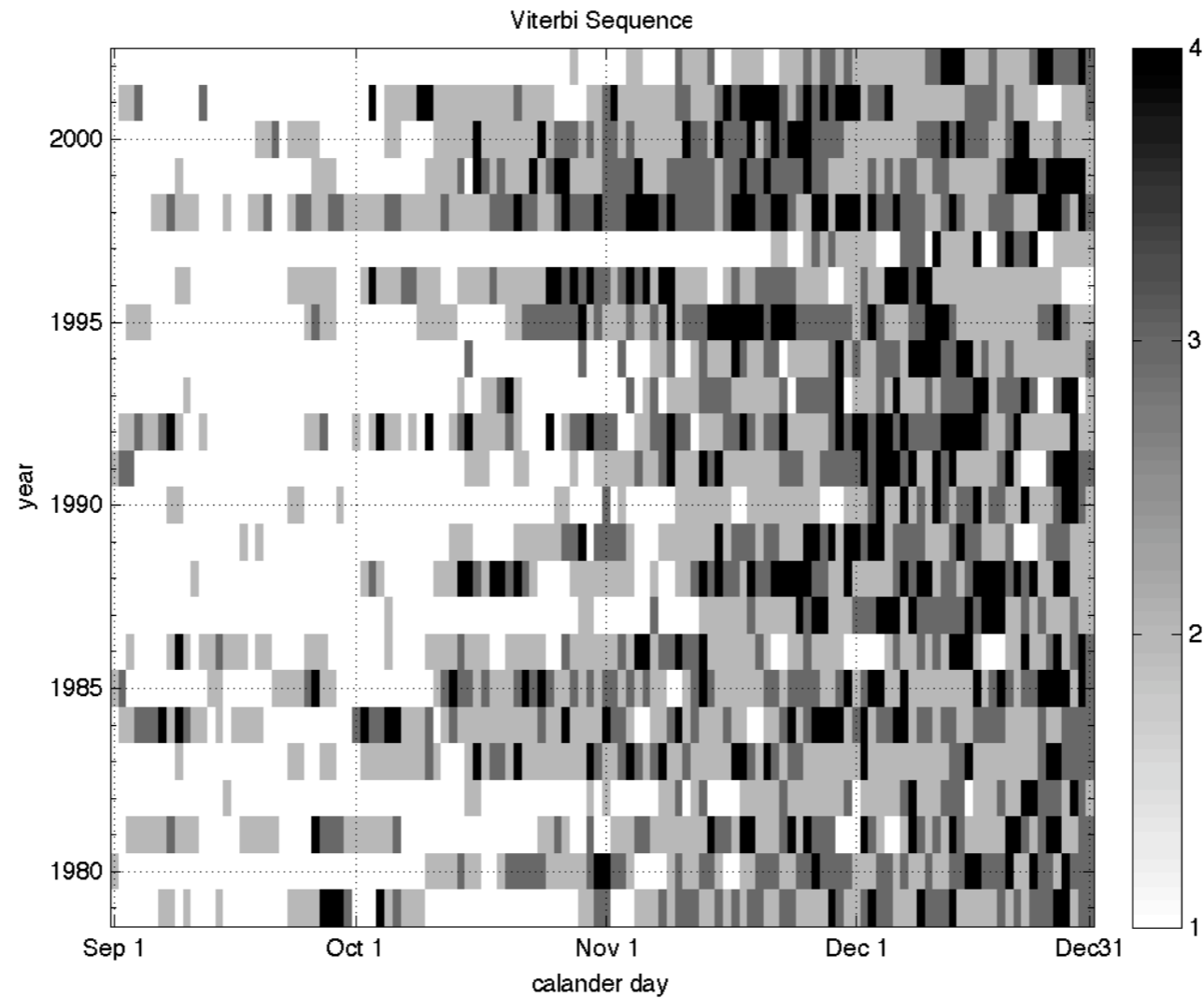
- toolboxes for constructing stochastic daily weather sequences conditioned on GCM outputs
 - ▶ HMM
 - ▶ KNN/weather typing



HMM of daily rainfall over Indramayu district, Java

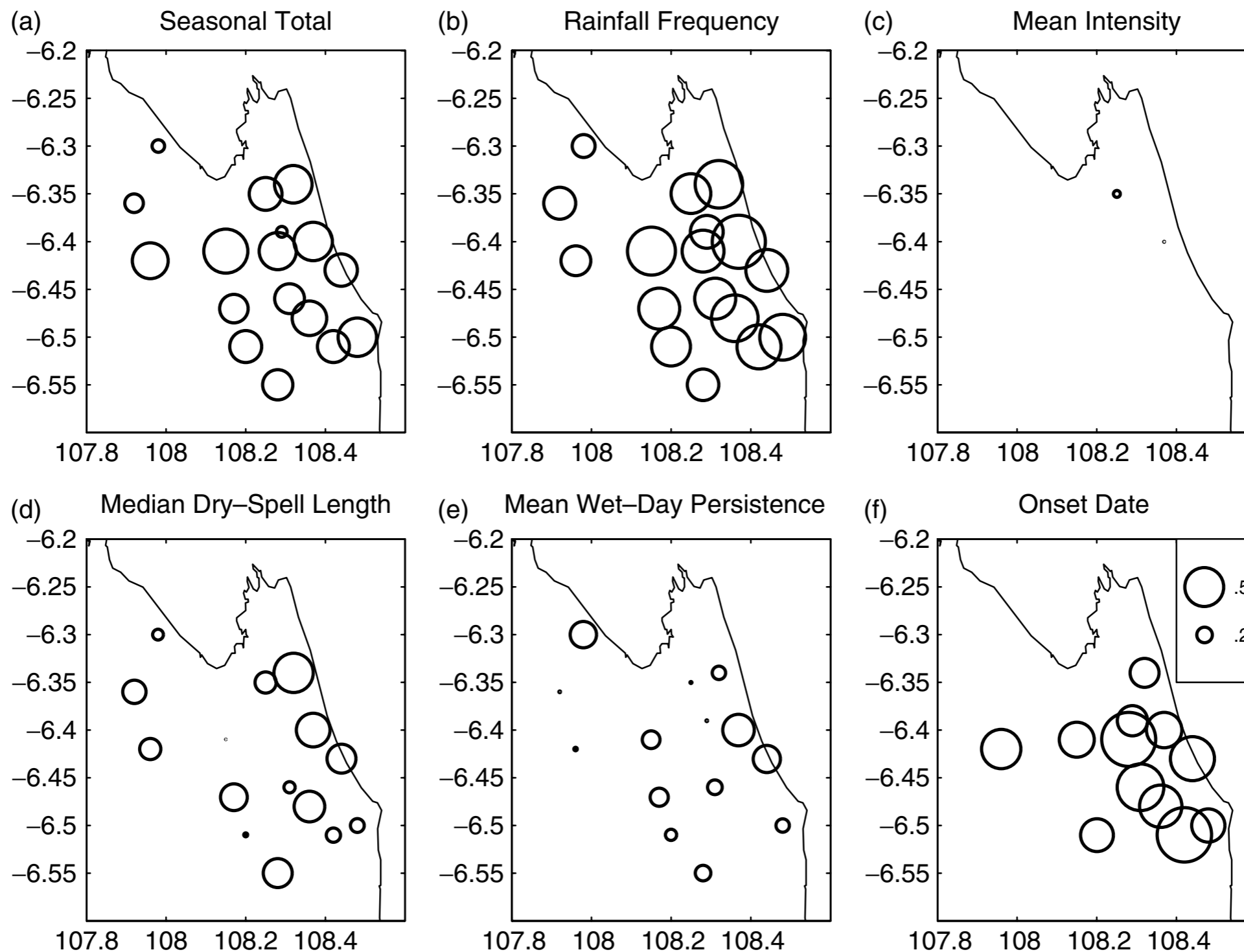


HMM
states



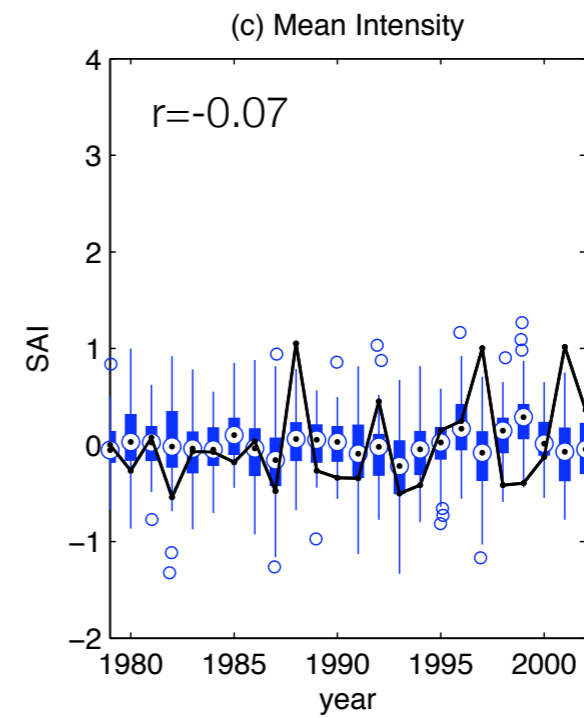
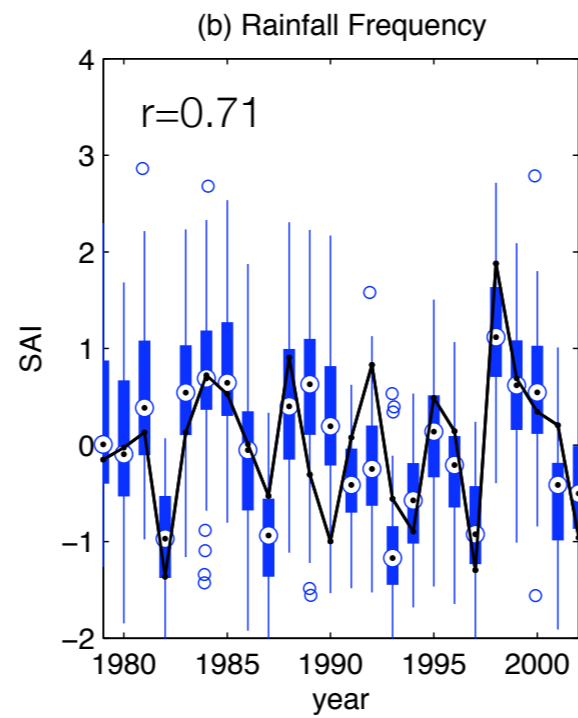
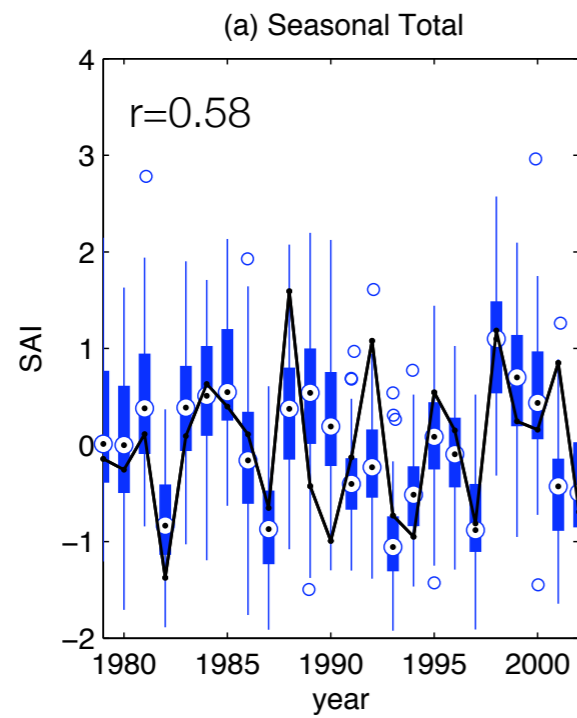
HMM
state
sequence

anomaly correlation skill of HMM downscaling, Indramayu district, Java

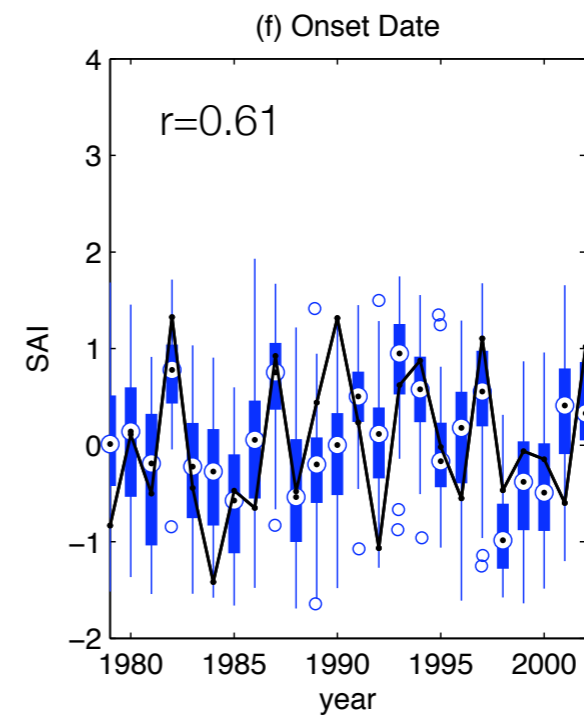
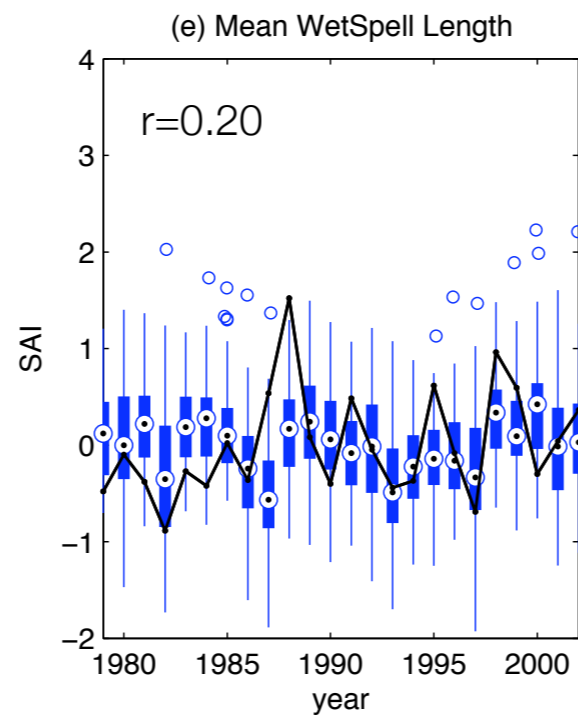
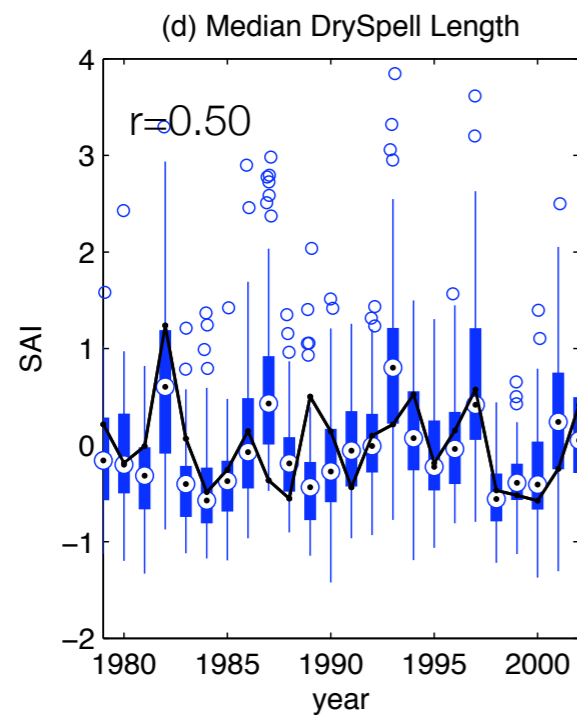


Sep-Dec season hindcasts made from Aug 1 (ECHAM-CA; 1979-02; station rainfall; NHMM)

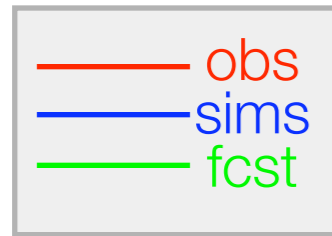
the forecasts: standardized station average



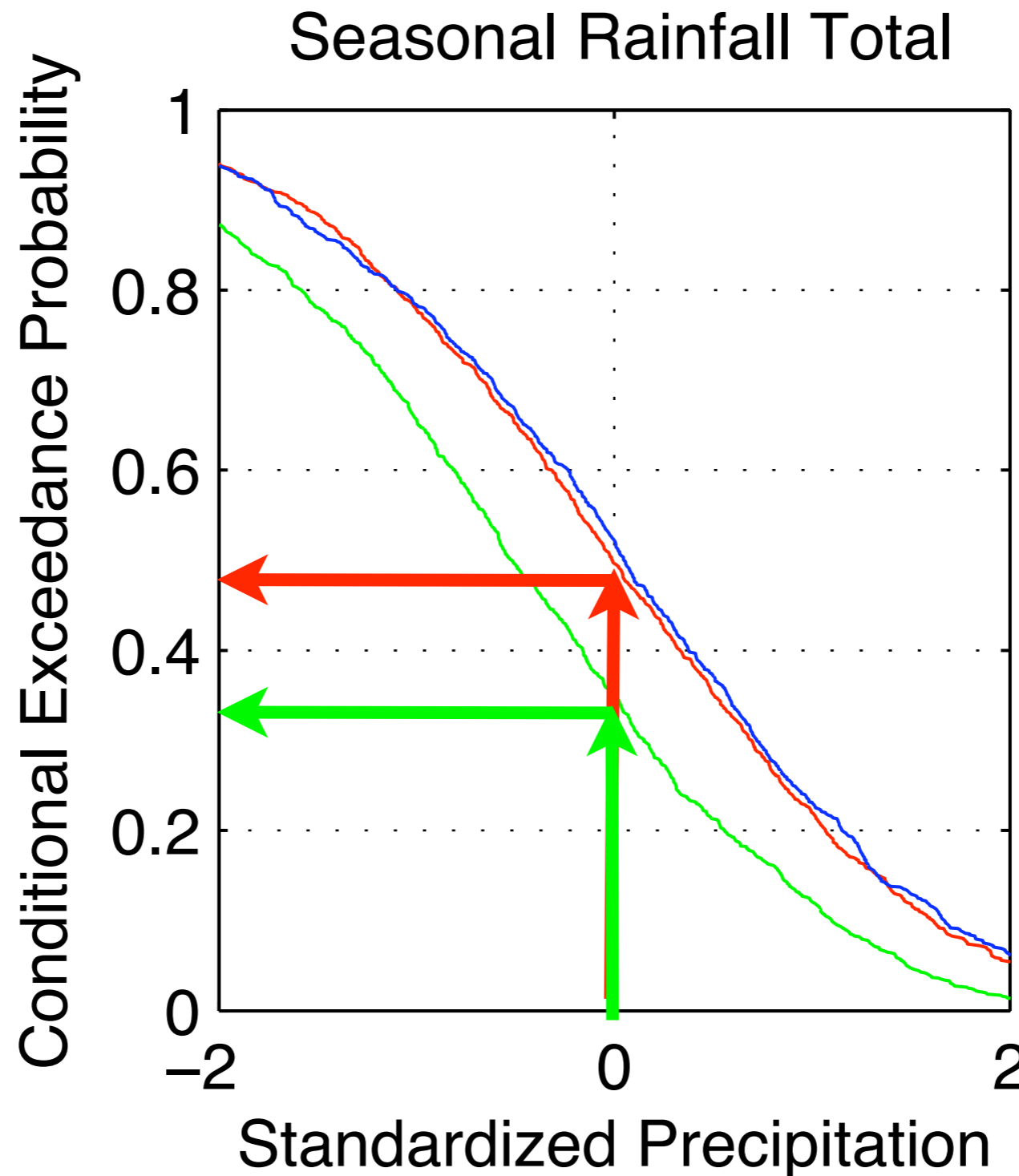
obs



tailored forecast for 2007 for Indramayu district, Indonesia



probability of exceeding
a given threshold



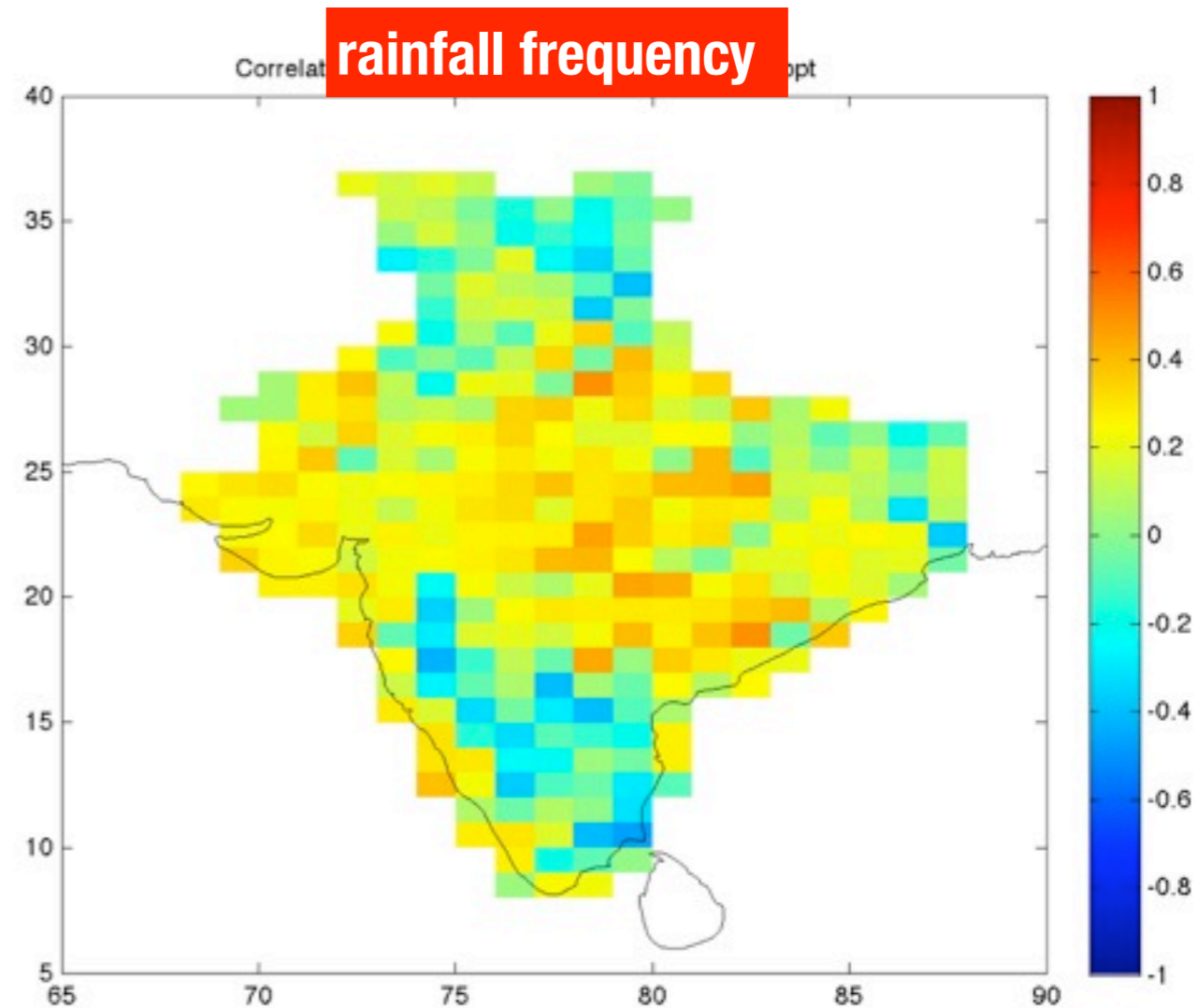
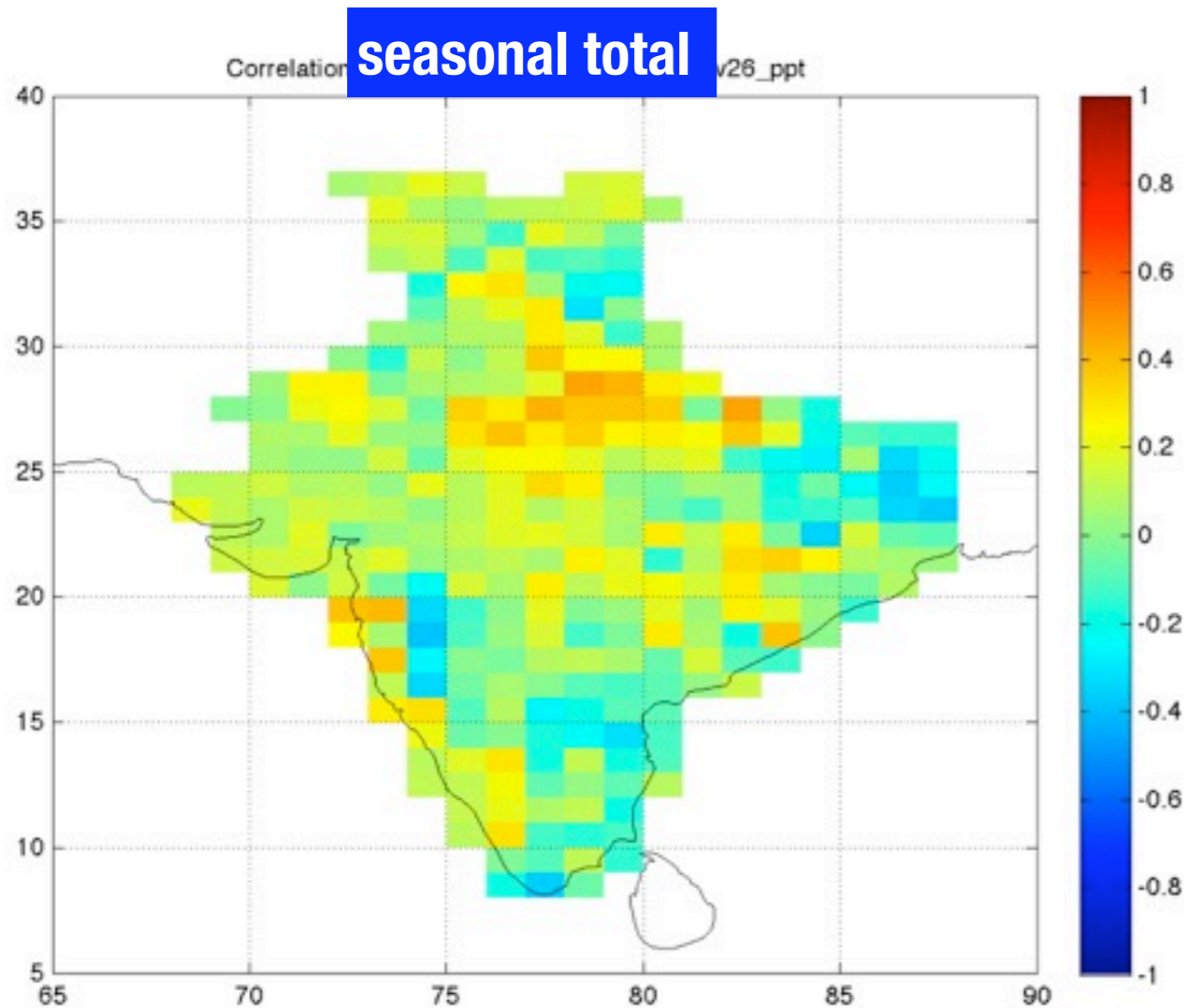
summary

- two basic ways to make seasonal predictions of weather statistics
 - regression (“MOS”) of seasonal statistics of daily rainfall
 - generate stochastic daily weather sequences
- stochastic daily sequences can be obtained by
 - quantile-quantile mapping of GCM’s daily rainfall
 - resampling (analog) methods
 - stochastic weather generators (SWG)
- hidden Markov model provides a multi-station SWG
- weather statistics such as rainfall frequency and monsoon onset date often exhibit surprisingly high seasonal predictability

additional slides



Jun–Sep anomaly correlation skill: NHMM[ECHAM4 precip (65E-200E, 5N-35N) , IMD]

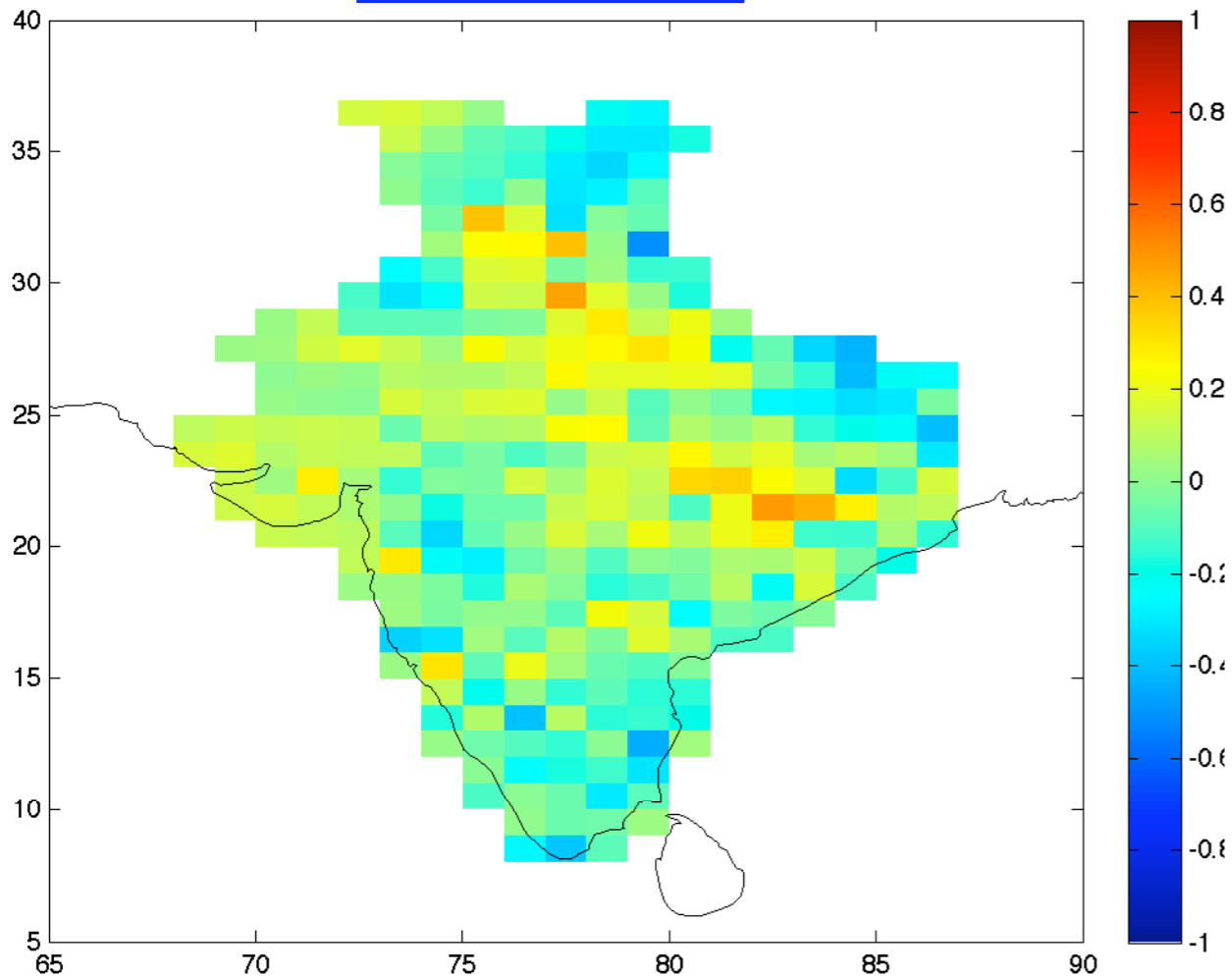


Non-homogeneous hidden Markov model (NHMM): builds a (cross-validated) logistic regression between predictor EOF time series and rainfall state probabilities.

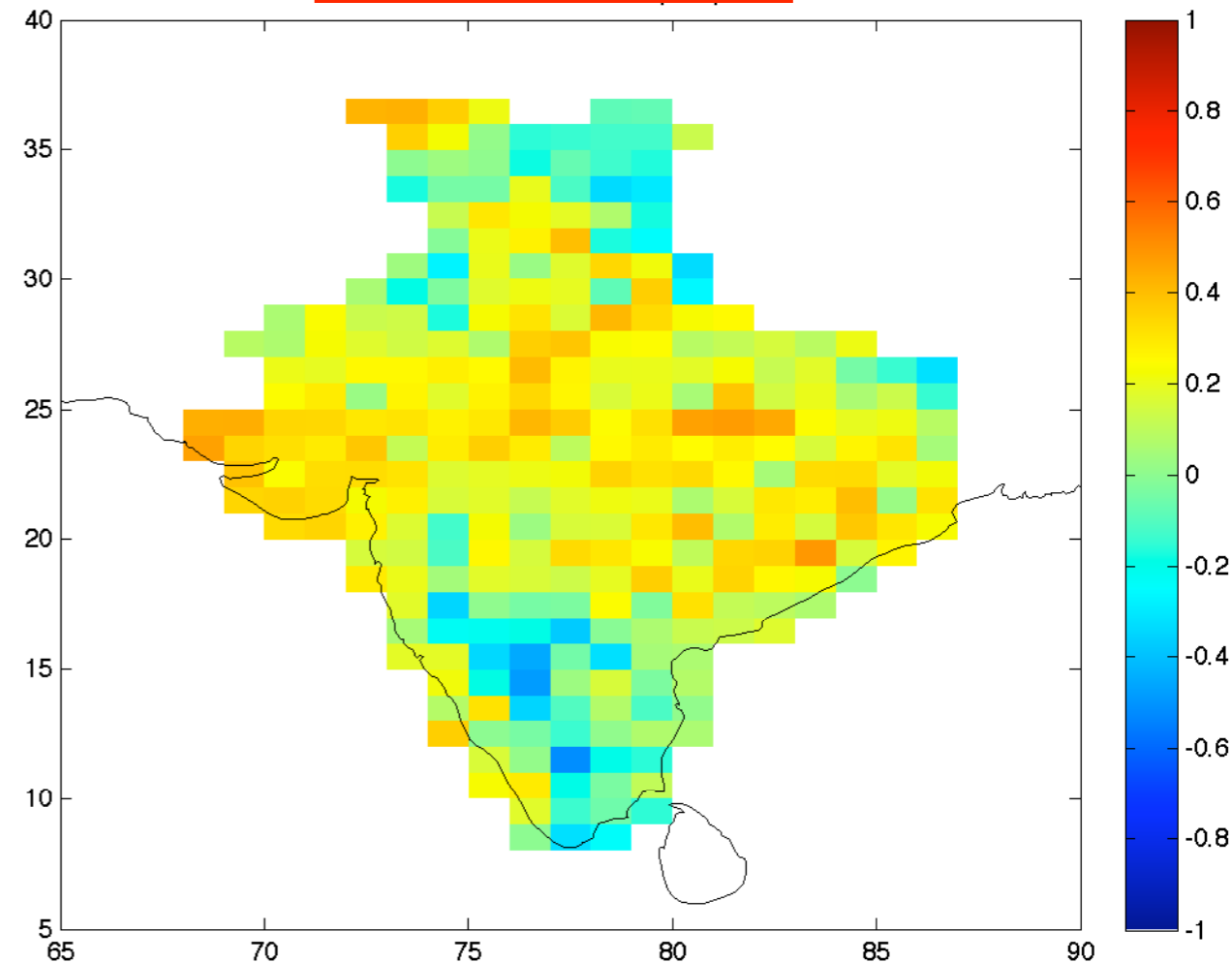
Jun–Sep anomaly correlation skill:

CCA[ECHAM4 precip (60E-90E, 5N-30N) , IMD]

seasonal total



rainfall frequency



Canonical correlation analysis (CCA): identifies linear combinations of the observation and predictor EOFs that are maximally correlated.