Perfect prognosis statistical downscaling

Joanna Wibig
Department of Meteorology and Climatology
University of Lodz

Validating Regional Climate Projections,
4th VALUE Training School – Triest, 26-30 Oct 2015

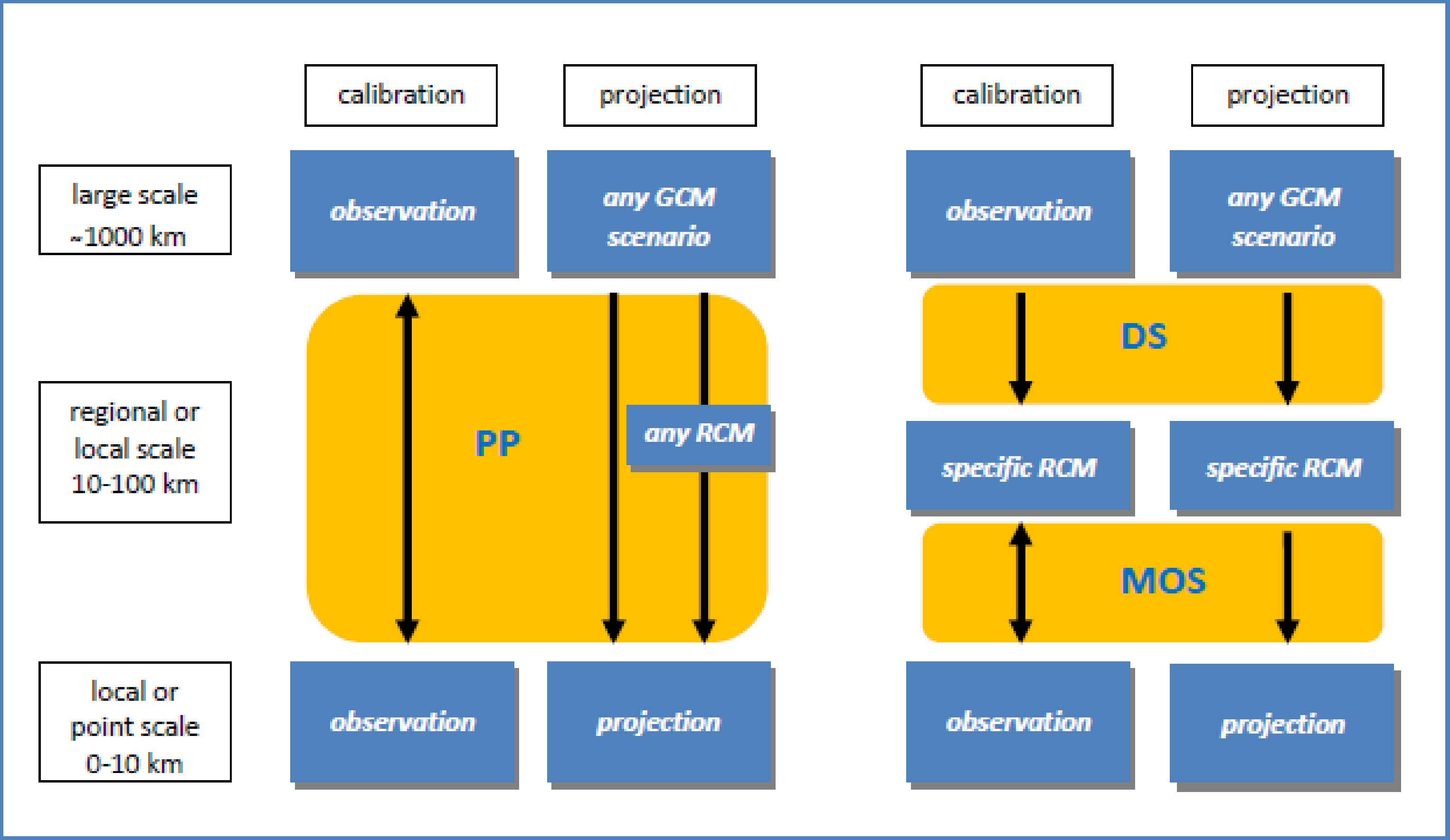
OUTLINE

- *Definition of downscaling and perfect prognosis methods.*
- *Predictors and preprocessing*
- *Reduction of dimensionality*
- *Regression methods*
- *Analogs methods*
- *Weather classification methods*
- *Weather generators*

What is downscaling ?

It is understood as a process linking large-scale variables with small-scale variables.

Types of downscaling methodologies



The PP approaches establish the statistical relationship between large-scale predictors and regional or local-scale predictands

$$y = f(X, g) + \varepsilon$$

predictand
i.e. the output data of a statistical model

y – local or regional predictand

X – large scale predictors

g – local features

ε – small scale noise

predictor
i.e. the input data used in a statistical model

PREDICTORS & PREPROCESSING

The requirements that the predictors should fulfil:

- ***a strong statistical relationship between predictors and predictand should exist,***
- ***the relationship between predictors and predictand should be stable in time,***
- ***suitable predictors should be reasonably well simulated by GCMs,***
- ***predictors should capture the global warming signal***

REDUCTION of DIMENSIONALITY

A reduction of dimensionality:

- ***The predictor, being a large-scale variable, is defined at a huge number of grid points.***
- ***Decomposing the field variable into a small number of modes of variability***

EOF's (Empirical Orthogonal Functions)

The large-scale variability can be described in terms of orthogonal empirical functions (EOFs) (Lorenz 1956; North et al. 1982; Benestad 2001).

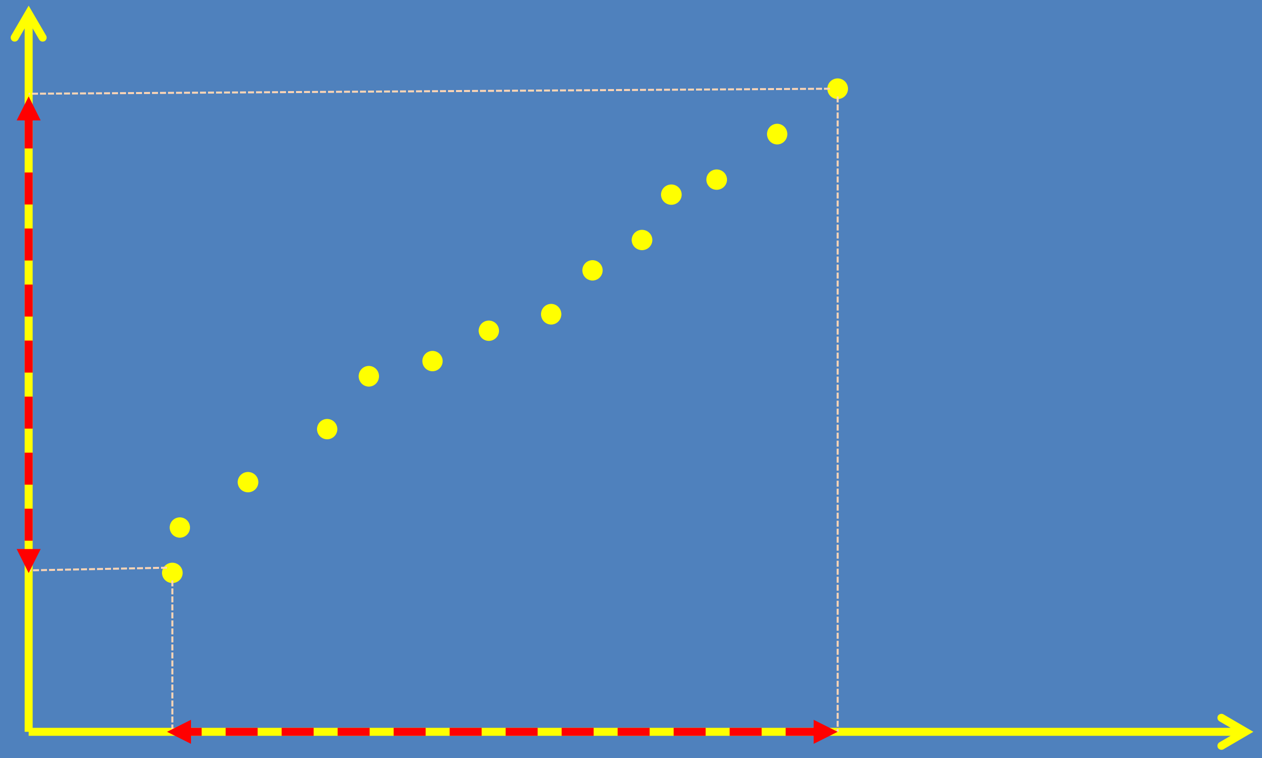
The spatial structures of EOFs describe a set of spatially coherent 'modes' that describe the variations in the gridded data.

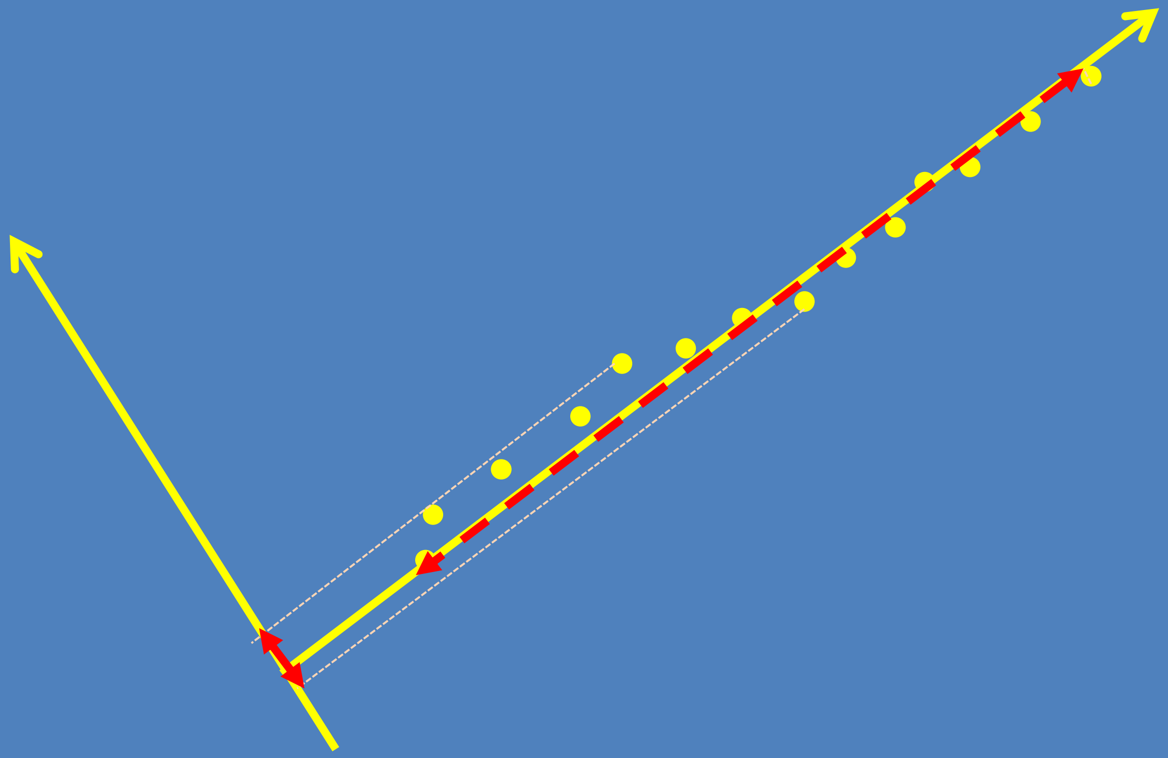
The leading modes describe the structures that are most pronounced and have the greatest spatial scales.

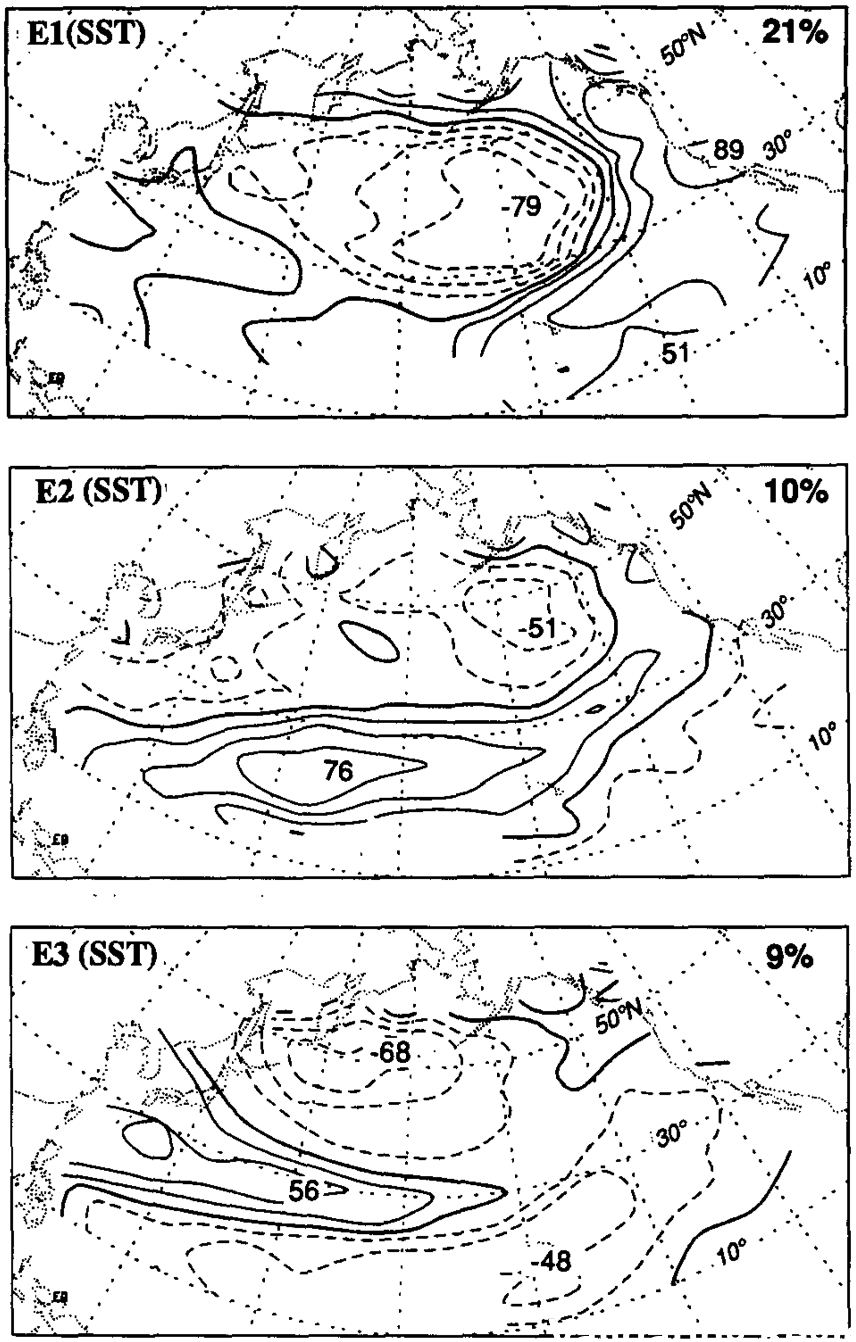
Only a small number of leading EOFs describe the major part of field variability (Wilks 1995).

Each spatial EOF pattern is associated with a vector of weights, often referred to as a 'principal component' (PC), describing how strongly this pattern is present at any time of the record.

The PCs are the basis for the downscaling model calibration, for instance a multiple regression against the predictand.







EOFs

The first three EOFs of normalized Pacific SST based on data from 39 winters (1946-1985). The values are correlation coefficients between PCs and SST anomalies at each grid point. The fraction of variances explained by each PC is given in upper right-hand corner.

Wallace et al., 1992, JClim

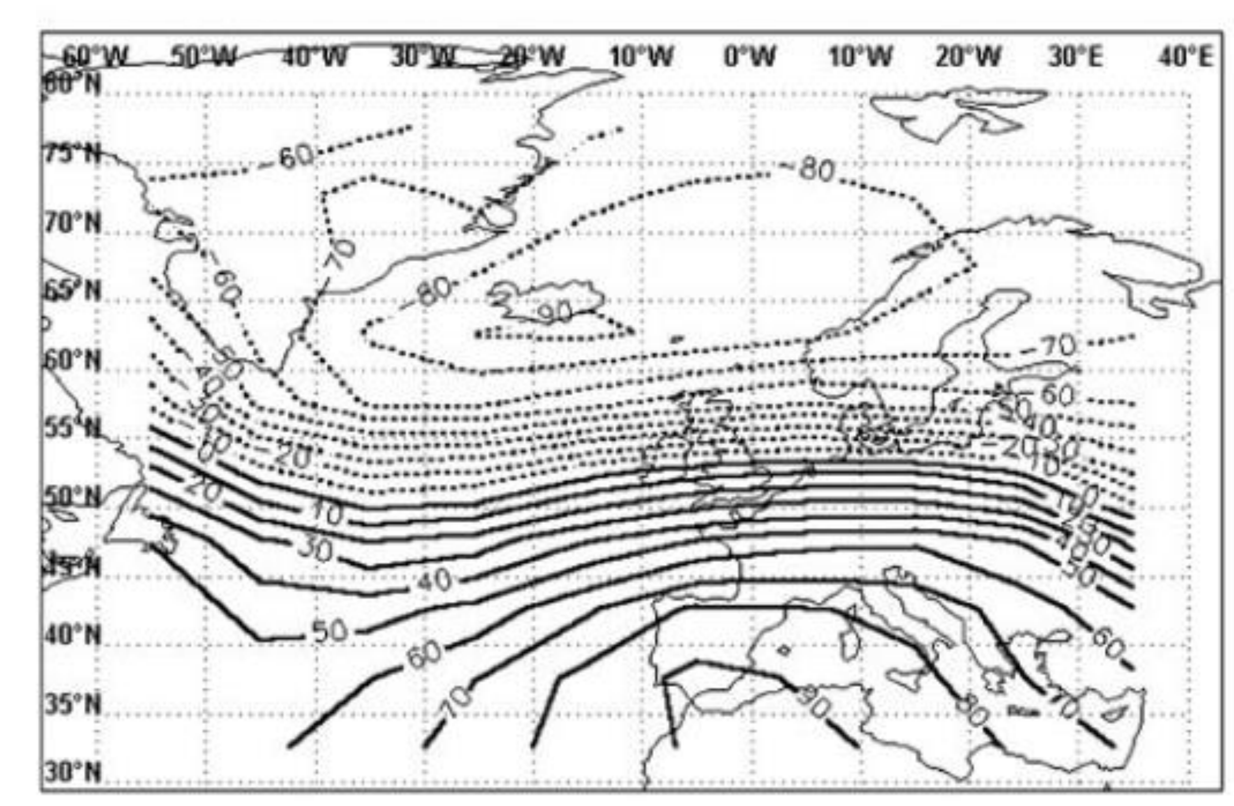
INDICES:

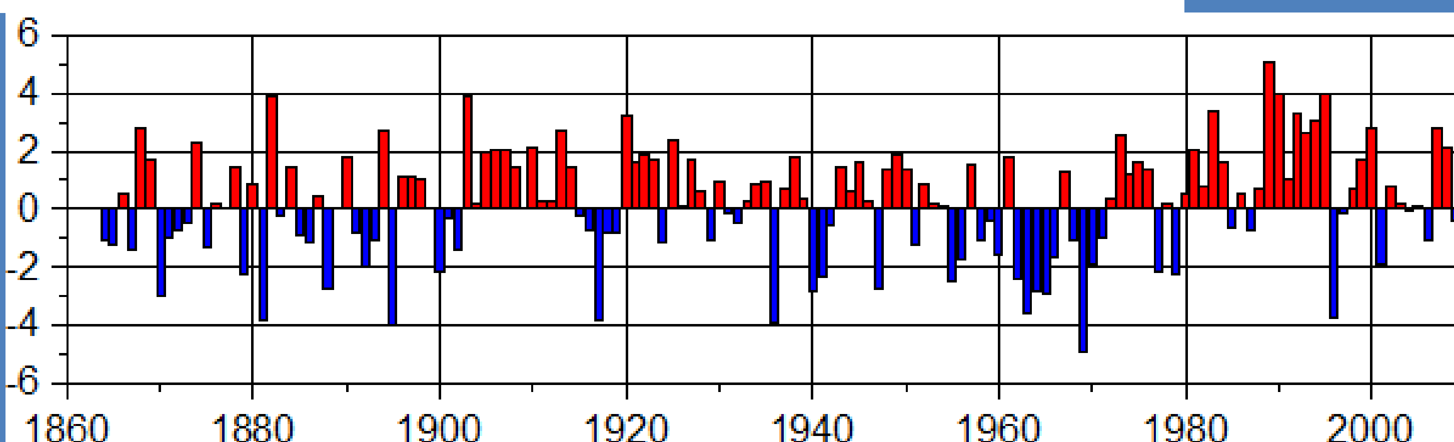
In the case of SLP fields, these can be the indices of zonal and meridional flow, vorticity, or other indices, such as the North Atlantic Oscillation index (Conway and Jones 1998; Wilby and Wigley 2000).

WEATHER TYPES:

The large-scale field, usually SLP or geopotential height, is mapped into a set of categories—weather types—by a clustering algorithm

Index NAO:





Castro-Diez, Y., et al., 2002, GRL, 29: NO. 0, 10.1029/2001GL014042

REGRESSION MODELS

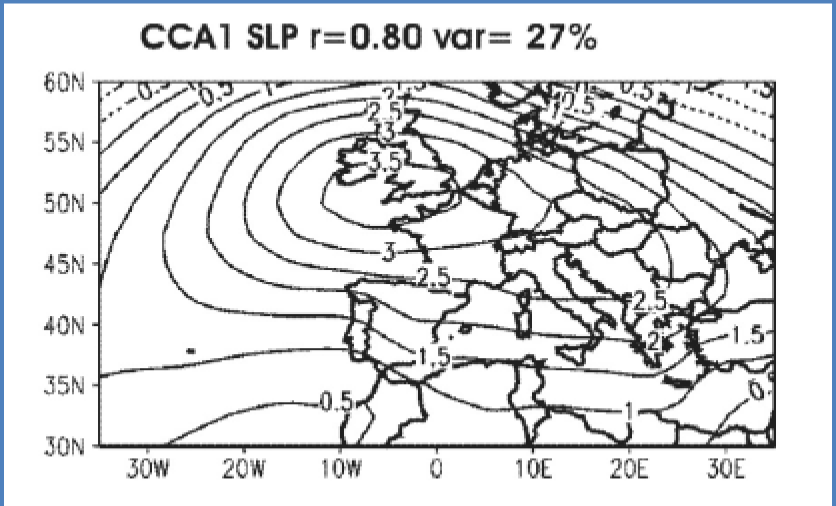
They include linear and nonlinear relationships between predictors and the predictand:

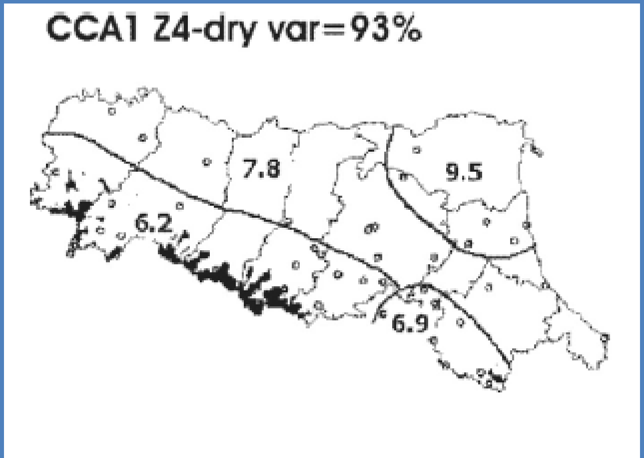
- ❑ *multiple regression (Murphy 1999) - minimises the root-mean-square errors (distance between predictions and observations)*
- ❑ *CCA method (Busuioc et al. 1999) - maximises the correlation between predictand and predictor fields*
- ❑ *SVD method (Bretherton et al. 1992) - maximises the covariance between predictand and predictor fields*

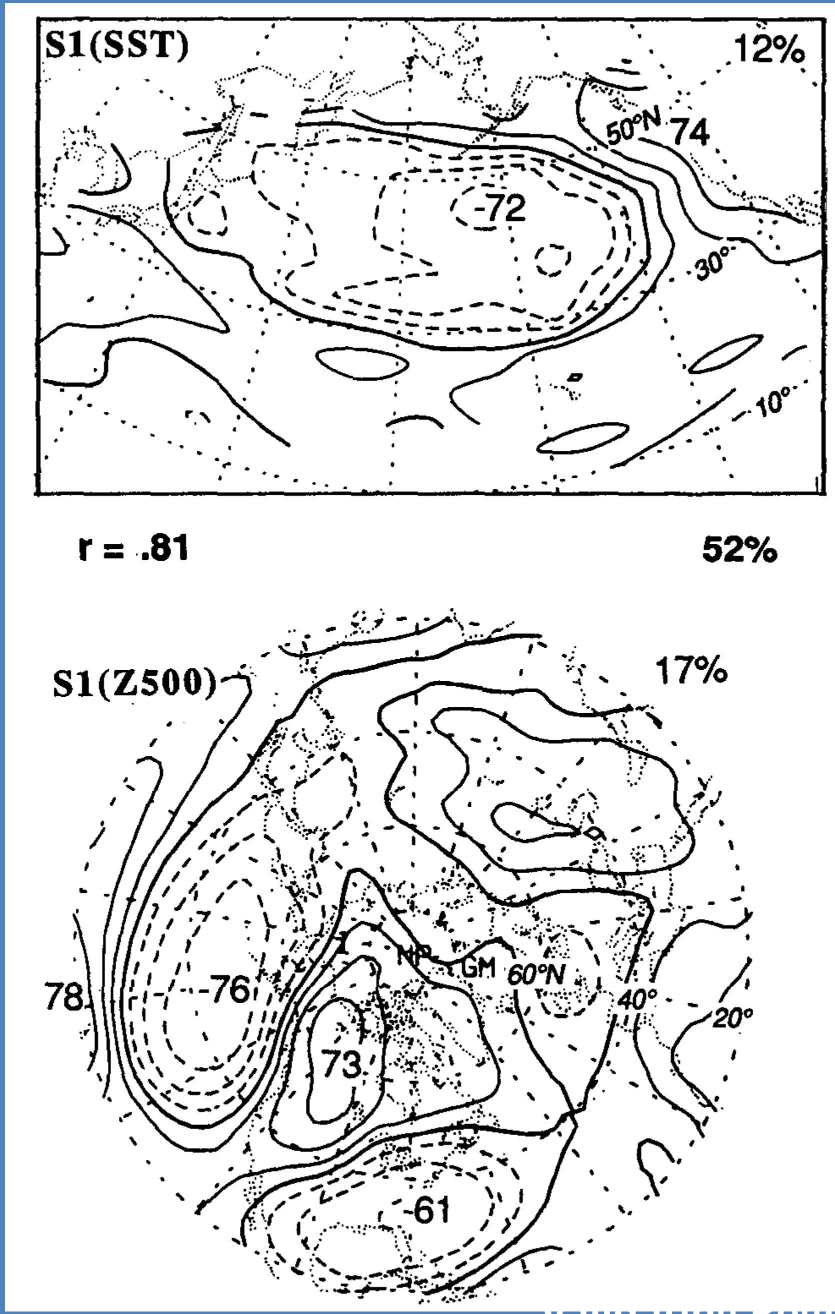
Artificial neural networks also represent nonlinear regression models (Crane and Hewitson 1998).

Busuioc, et al., 2008, IJoC

- ***The transformation procedure should generate a predictor that has high predictive power, that is, explains a high percentage of the variability of the predictand.***
- ***Some methods, such as canonical correlation analysis (CCA) or the singular value decomposition (SVD) method, directly seek the modes having the highest correlation or covariance with the predictand field, while others do not.***





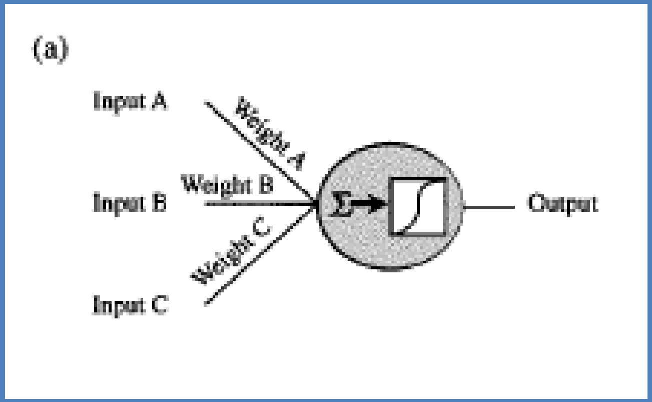


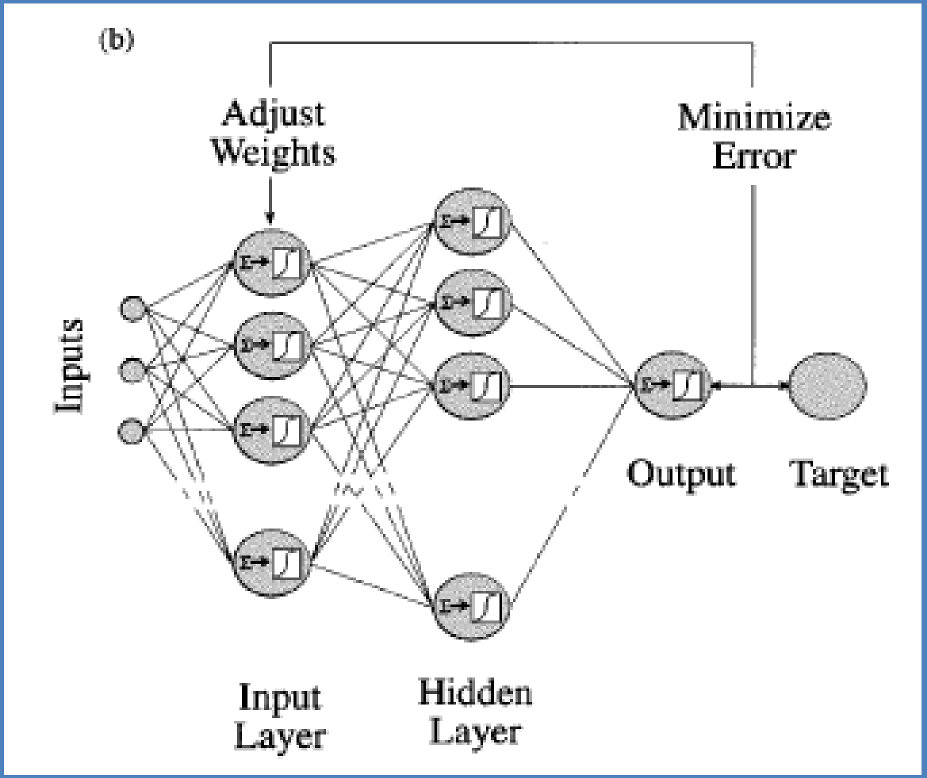
SVD

Heterogeneous correlation pattern for the first mode in the direct SVD expansion. The temporal correlation coefficients between between the corresponding expansion coefficients is printed on the left and SCF (squared covariance fraction) expressed as a percent - on the right.

Wallace et al., 1992, JClim

AN ARTIFICIAL NEURAL NETWORK





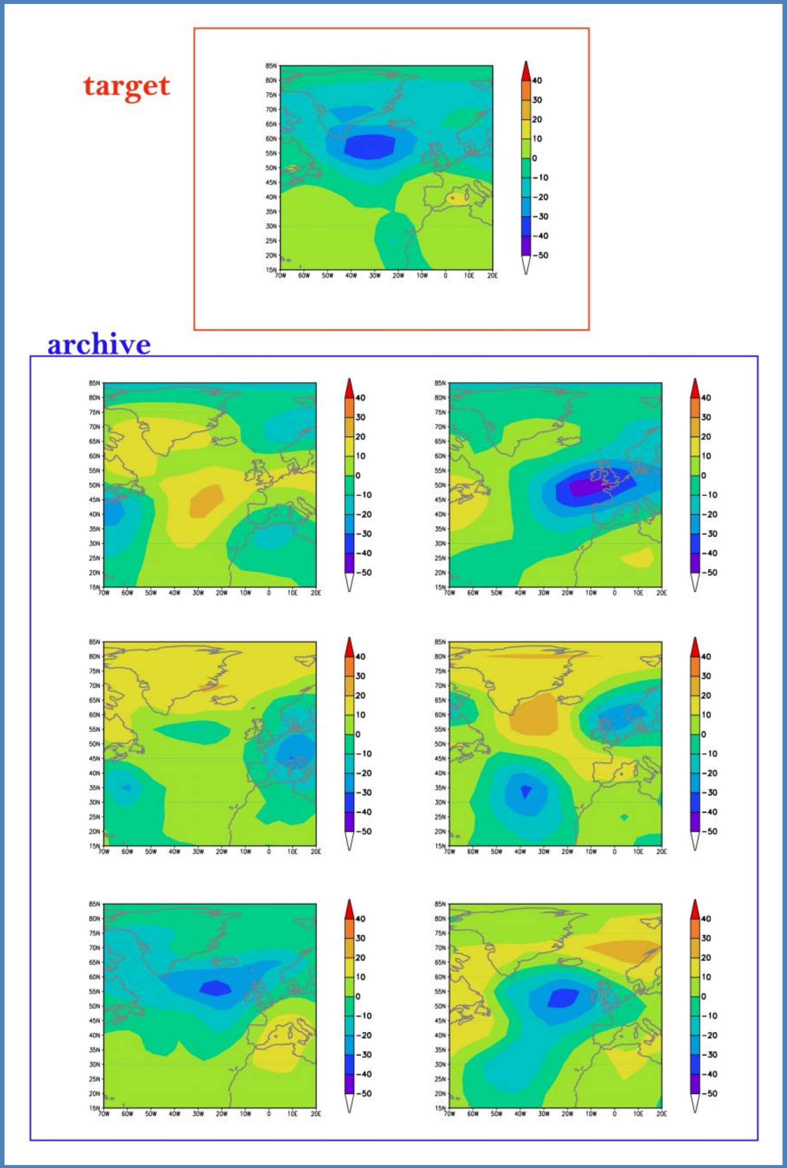
The structure of an artificial neural network set a) An individaul processing node; b) The arrangement of nodes in the input, hidden and output layers

Crane & Hewitson, 1998, Int. J. Clim.

LINEAR & NONLINEAR TECHNIQUES

The ANALOG method:

Picking the date from the archive most closely resembling the day for which prediction is made.



QUESTION:

How to determine the past events that are the most similar to the one in question?

Variants of ANALOG method:

- ***Definition of similarity***
- ***Search for the most similar pattern or a few the most similar***
- ***Search for similar patterns or similar evolution***
- ***Use of methods reducing the dimensionality of a large-scale field***

Similarity measure examples:

- ***Euclidean***

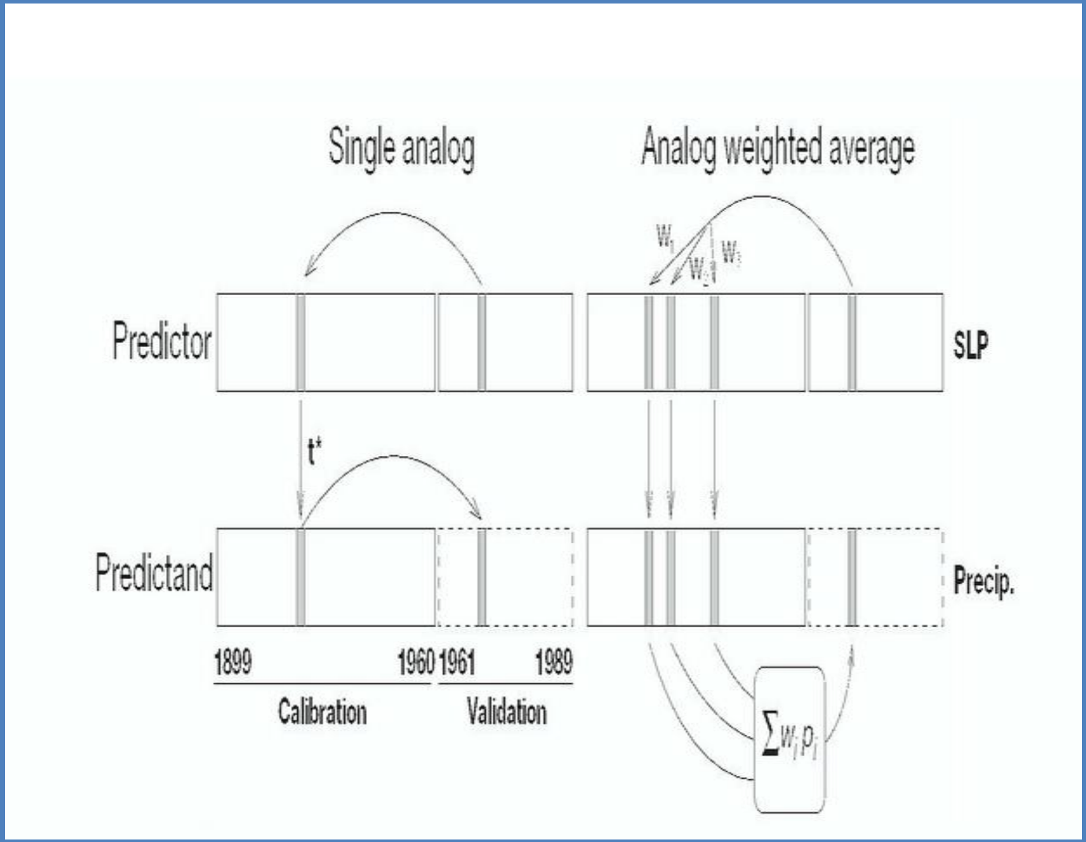
$$S = \sum_{i=1}^{n} (x_i - y_i)^2$$

- ***Weighted euclidean***

$$S = \sum_{i=1}^{n} w_i \cdot (x_i - y_i)^2$$

- ***Absolute deviations***

$$S = \sum_{i=1}^{n} \left| x_i - y_i \right|$$



Fernandez and Sáenz, Climate Research 2003

- Search for several the most similar and average the local variable
- Local noise is averaged out increasingly
- The PDF is increasingly more normal

A search for similar evolution:

For instance taking the similarity measure:

$$S = \sum_{i=1}^{N} \sum_{\tau=0}^{T} \left[x(i,t) - y(i,t-\tau) \right]^2$$

A reduction of dimensionality:

- *Each N-gridpoint field is reduced to a few leading indices that characterize the large-scale important features: positions of high/low pressure centres, intensity of zonal winds, synoptic situation*
- *Such reduction can be performed by PCA (but non necessarily)*

A drawbacks associated with analog models:

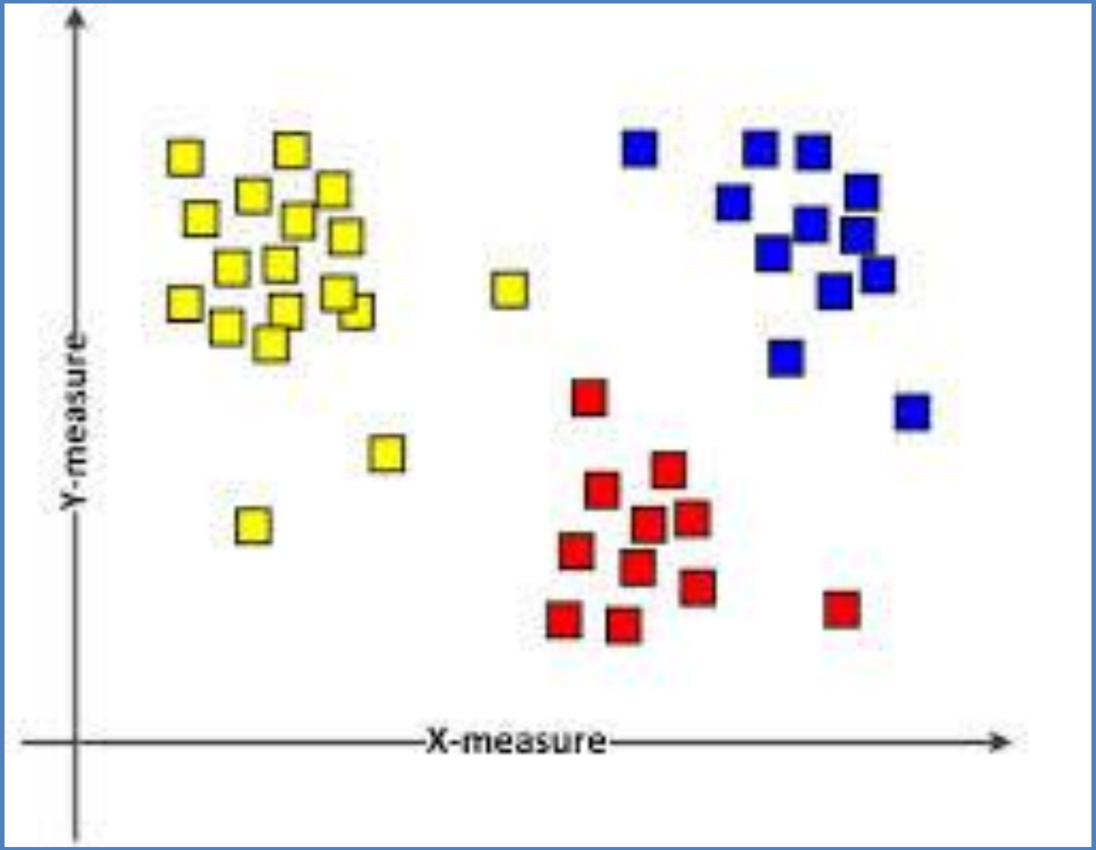
- ❑ *The values outside the range of the calibration set cannot be extrapolated*
- ❑ *Nonstationary relationships between the large-scale and local climate cannot be accounted for*
- ❑ *Large training samples are necessary*
- ❑ *A consistency in the order of consecutive days cannot be ensured*

The CLASSIFICATION and/or CLUSTER method:

The idea is that one type of weather tends to bring one type of precipitation (temperature) type.

Cluster analysis bases the predictions on a number of closest states (Wilks 1995), either by taking the mean of the days with close matches or by using the observed values for all days that match the predicted state, and constructing a statistical distribution

Clusters tend to consists of points which a close to the each other compared to the members of other clusters.



Measuring of the cluster to cluster distance:

- ***Single linkage***
- ***Complete linkage***
- ***Average linkage***
- ***Centroid clustering***
- ***Ward's minimum***

$$d_{j,i} = \left\| \bar{x}_i - \bar{x}_j \right\| = \left[\sum_{k=1}^{K} \left(x_{i,k} - x_{j,k} \right)^2 \right]^{0.5}$$

Decide the cluster and then choose a random observation from a set of days associated with this cluster

Weather generators:

Models used for Monte-Carlo simulation are regarded as random number generators with outputs statistically resembling daily weather data at a given location.

Although stochastic model outputs behave statistically like weather data, it is not expected that any particular simulated weather sequence will be duplicated in weather observations at a given time in either the past or future.

Wilks & Wilby, 1999, Progress in Physical Geography

In the majority of weather generators all elements are conditioned on precipitation presence or absence.

In particular there is two step way of daily precipitation total generation.

- ❑*Checking if the day is dry (no precipitation) or wet (precipitation >0 mm)*
- ❑*On a wet day the precipitation amount is generated.*

The precipitation occurence can be decided in so called Markov process:

First order Markov process:

- ***There are four possibilities:***

W-W, W-D, D-W, D-D

- ***If the previous day was wet, there is a probability p_{11} that the next day will be wet and $p_{10} = 1 - p_{11}$ that the next day will be dry***

- ***If the previous day was dry, there is a probability p_{01} that the next day will be wet and $p_{00} = 1 - p_{01}$ that the next day will be dry***

The precipitation occurence can be decided in so called Markov process:

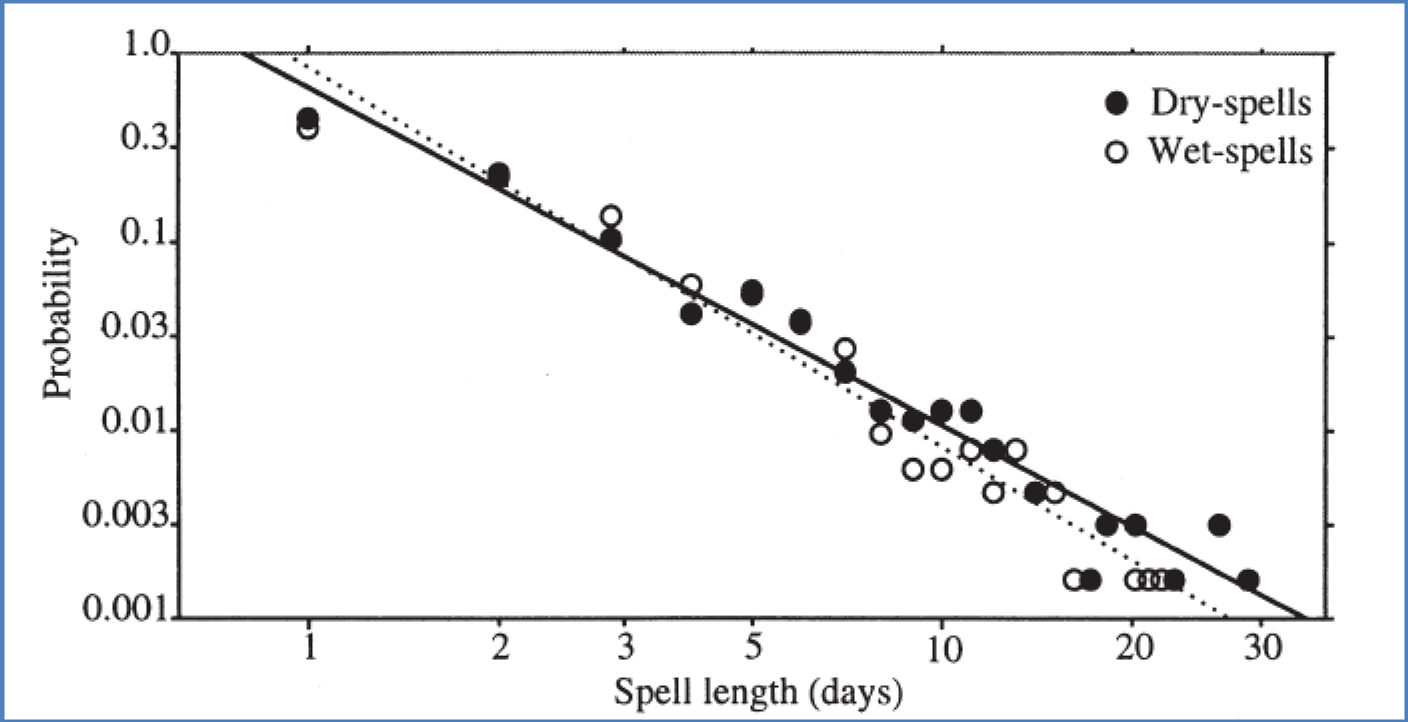
Second order Markov process:

- ***There are eight possibilities:***

WW-W, WW-D,WD-W, WD-D, DW-W, DW-D, DD-W, DD-D

- ***If the previous day was wet, and the day before the previous was also wet there is a probability p_{111} that the next day will be wet and $p_{110} = 1-p_{111}$ that the next day will be dry ....***

An alternative way is to simulate the precipitation occurrence from spell-length model:



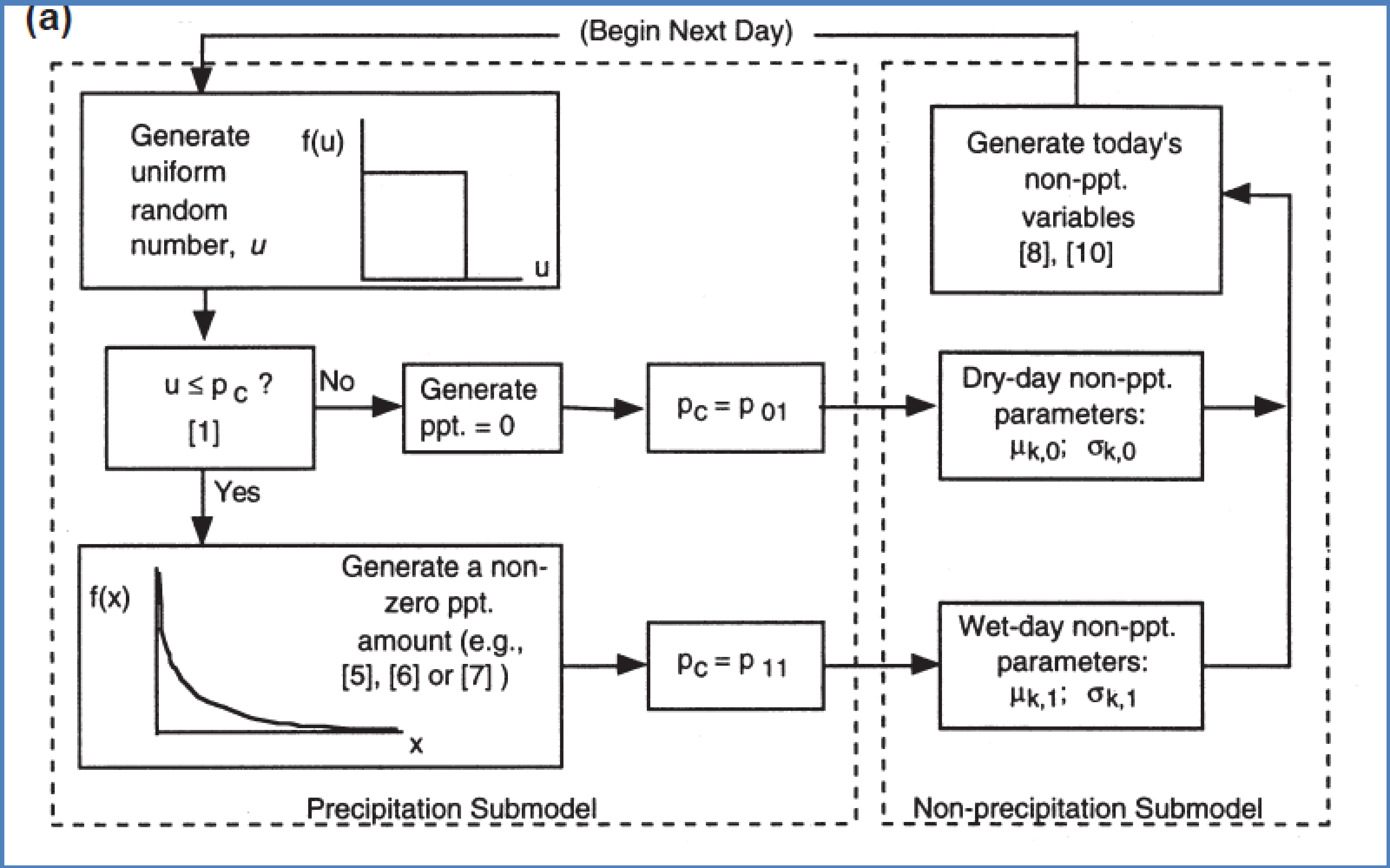
Wilks & Wilby, 1999, Progress in Physical Geography

precipitation occurrence:

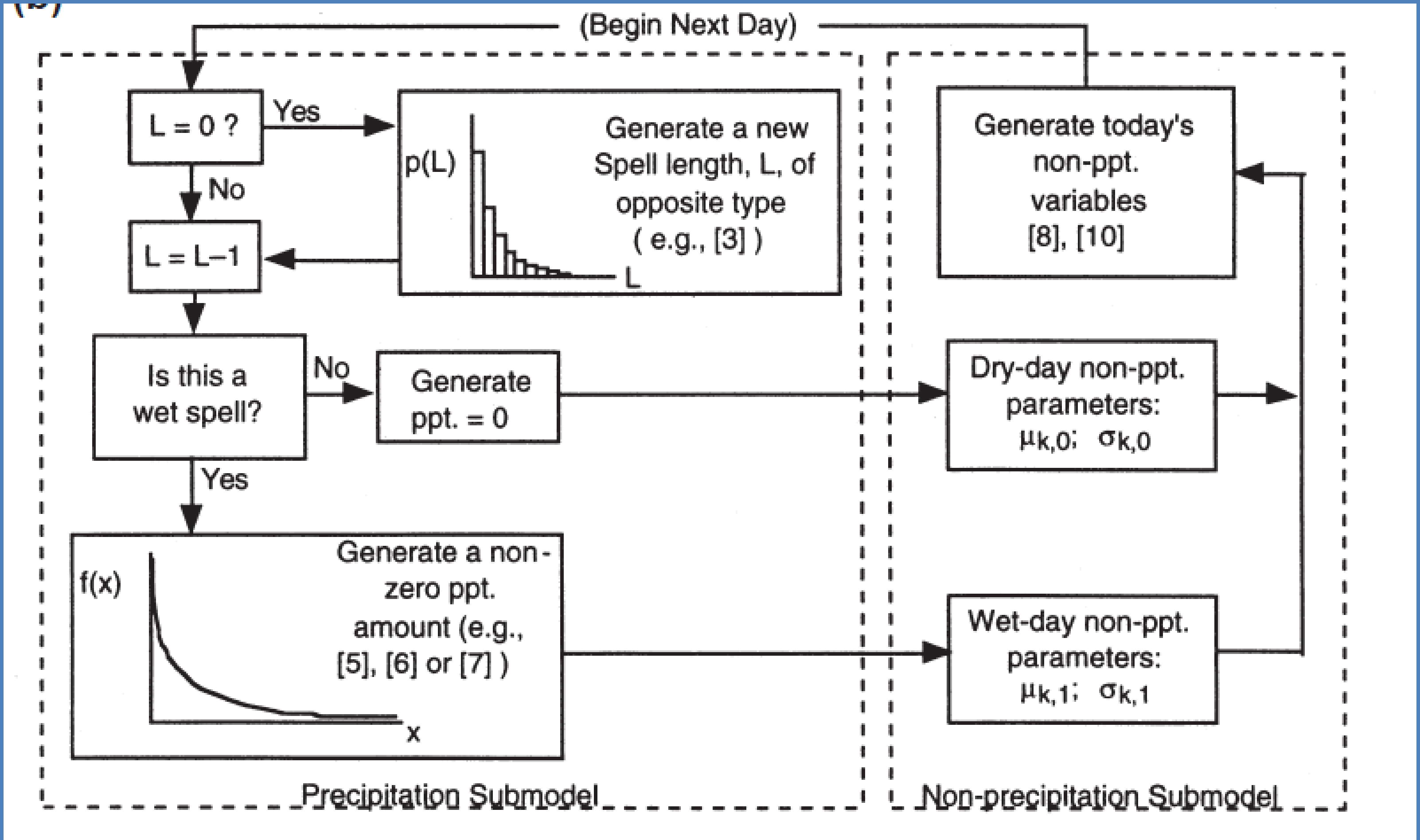
- *first or second order Markov chain*
- *spell-length model - alternating wet and dry series approximated by semiempirical probability or geometric or truncative negative binomial or negative binomial or mixed geometric distributions,*

precipitation amount:

- *exponential distribution*
- *mixed exponential distribution*
- *two parameter gamma or Weibull distribution*
- *kappa distribution*
- *log-normal distribution*
- *cubic root normal distribution*
- *semi-empirical distributions for each month/season*
- *multistate Markov approach*



Wilks & Wilby, 1999, Progress in Physical Geography



Wilks & Wilby, 1999, Progress in Physical Geography