

Downscaling Seasonal Precipitation in Perú using clustering techniques (self-organizing maps SOM)

<http://grupos.unican.es/ai/meteo>



José M. Gutiérrez
Antonio S. Cofiño
Carmen Sordo
CSIC / Univ. de Cantabria, SPAIN



Rafael Cano
Cristina Primo
Ins. Nac. de Meteorología, SPAIN

Thanks to (collaborators):

PERÚ

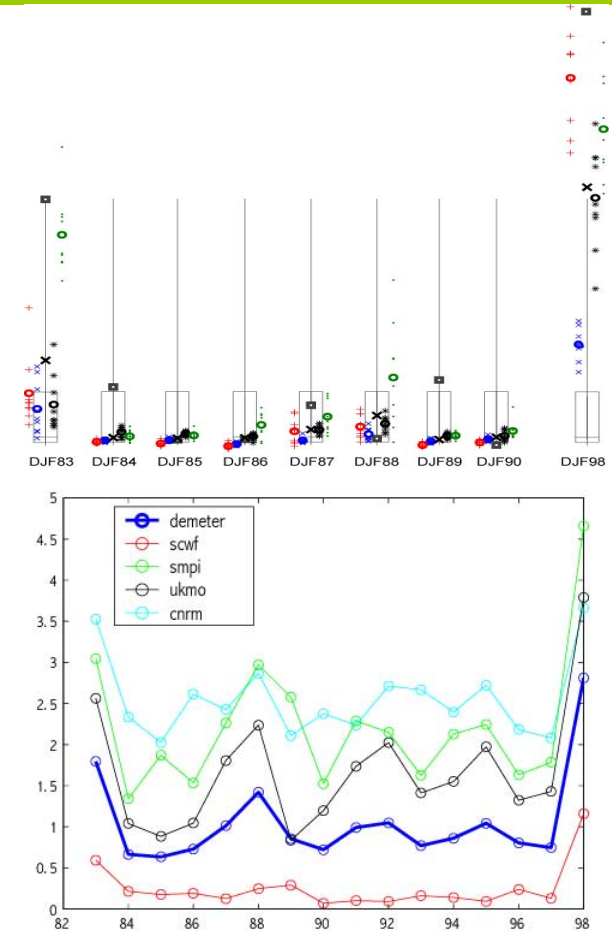


AI

**met
group**

**National Service of
Meteorology and Hidrology**

**National Geophysical
Institute**



<http://grupos.unican.es/ai/meteo>

DEMETER (Kiel 7-8 July, 2003)

1

REMINDER: The SOM Clustering Model

SOM
Analog
technique

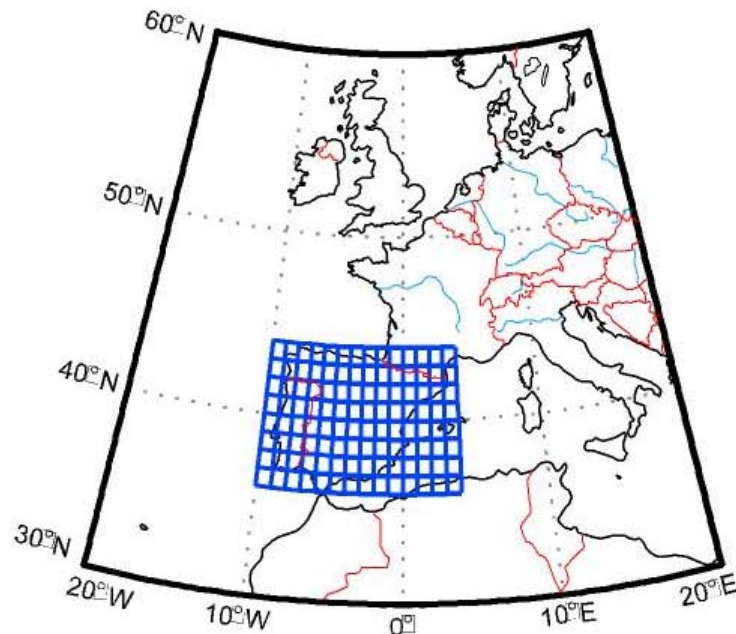
Tropical
latitudes

Middle
Latitudes

Future
work

AI
met
group

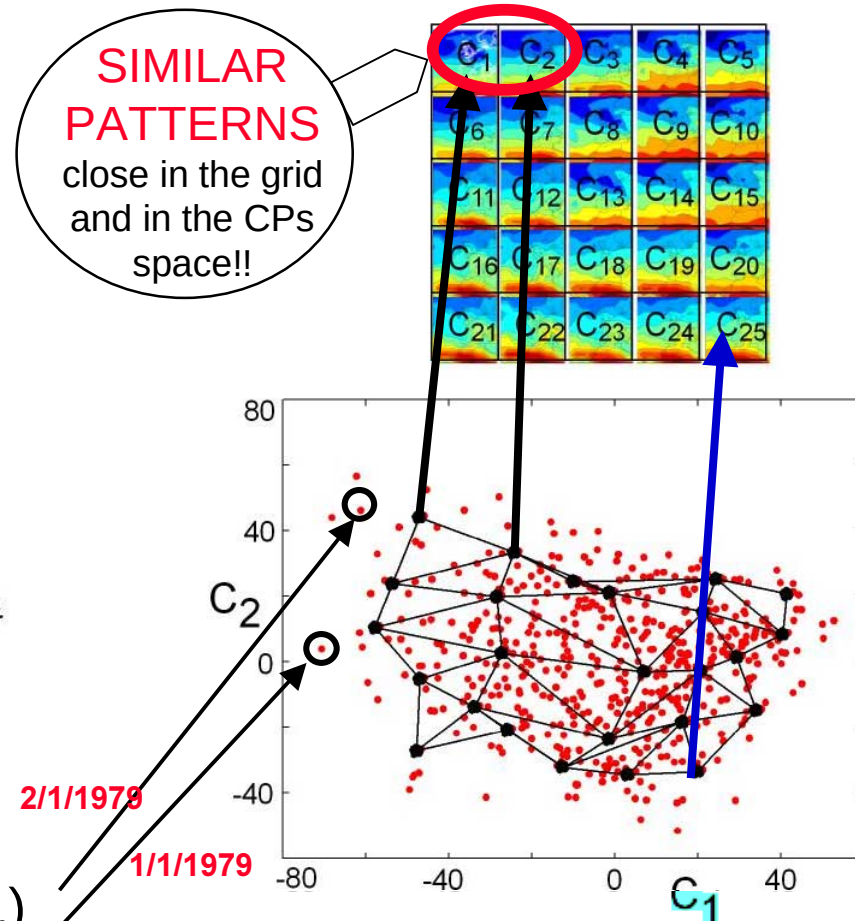
- Perú (Piura, Mantaro)
- Iberian Peninsula



Atmospheric daily circulation:

$\vec{ac} = (T(1000 \text{ mb}), T(850 \text{ mb}), \dots, Z, H, \dots)$

Dimension reduction: EOFs
100 dimensions are kept.



5 x 5 Self-org.
map (SOM)

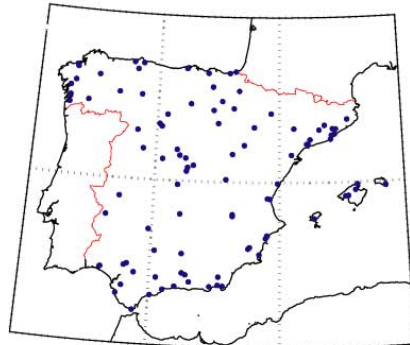
Projection of ERA-15
using the first 2 PCs

5500 ERA-15 daily patterns are clustered using a self-organizing map.

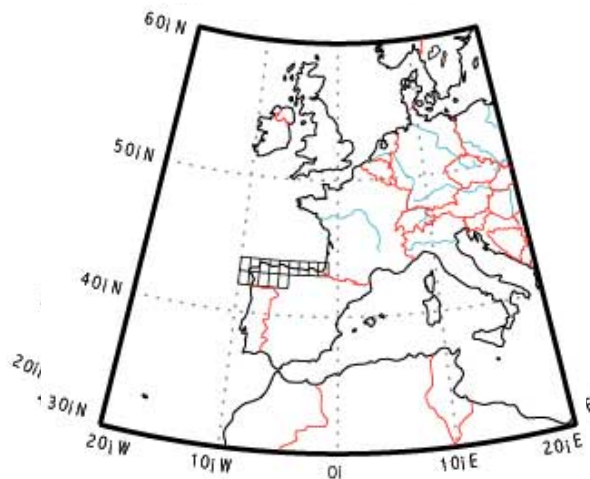
Clustering Downscaling Model vs Standard Analog

Short-range daily forecast :

- ✓ lead time D+1
- ✓ 100 stations in Spain
- ✓ 730 days (1997-1998)

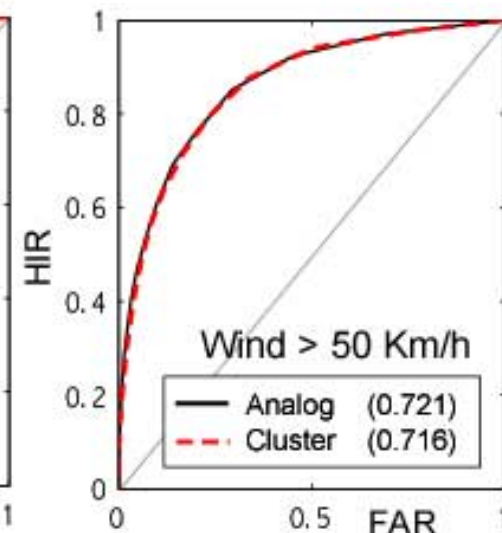
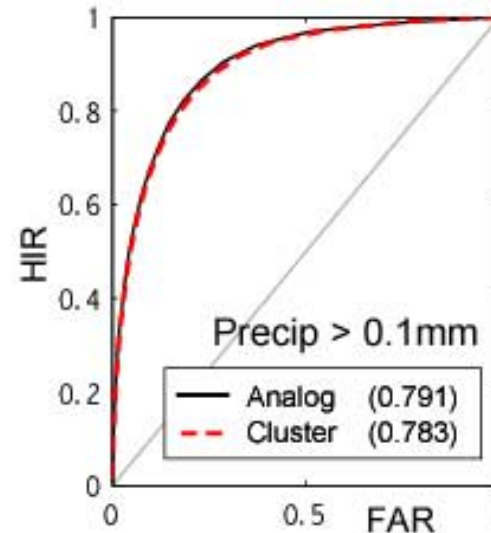


48 Days in advance (04, 12, 24Z)

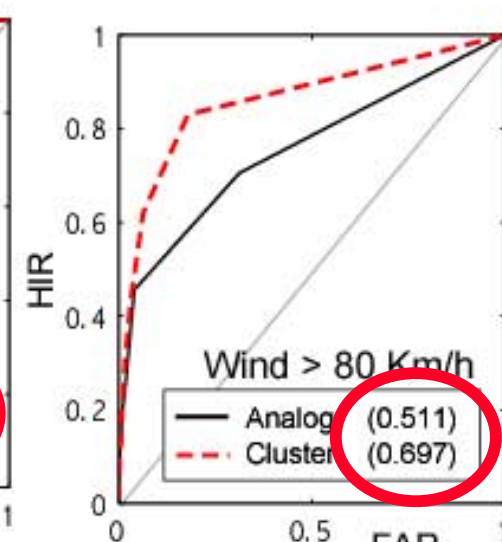
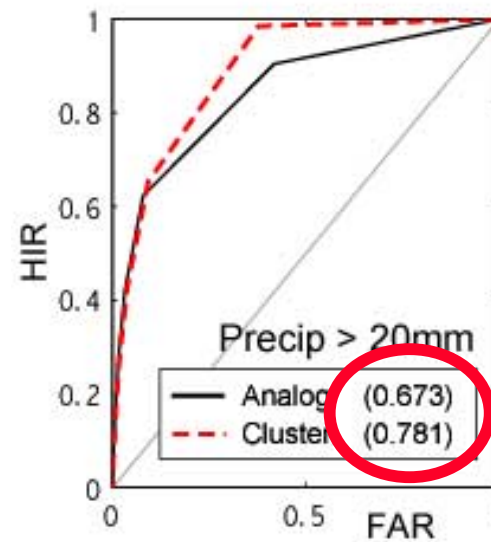


Rainfall

Max. wind



Normal event



Tail event

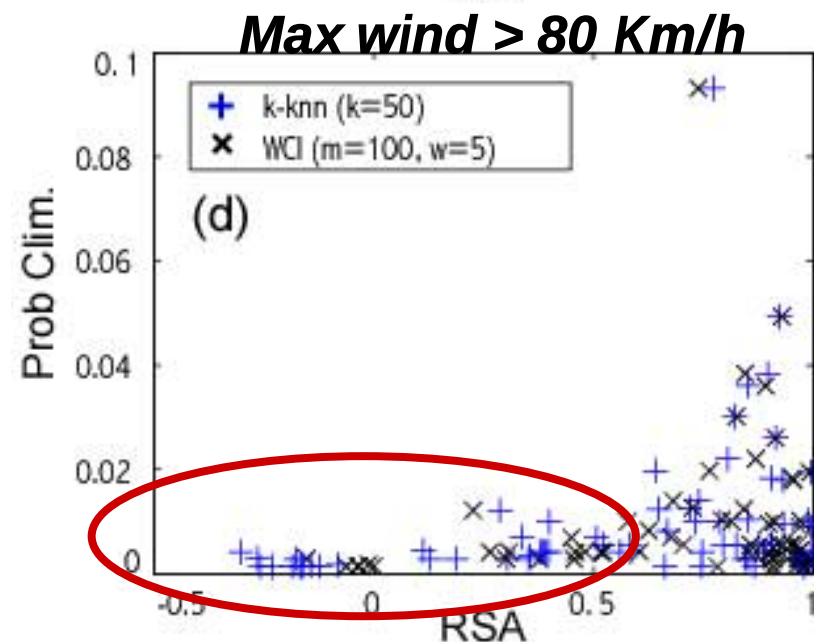
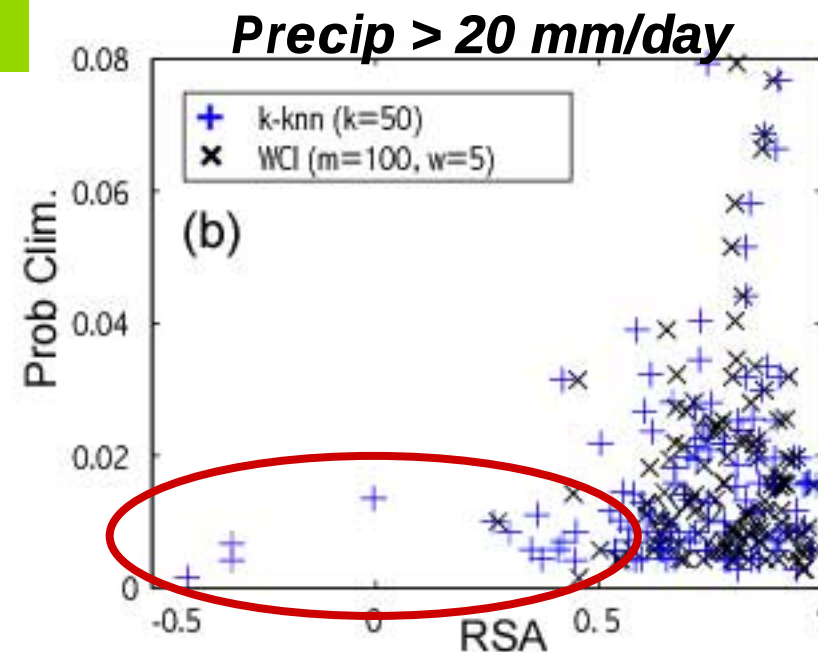
AI
met
group

Individual Skill of Stations

The clustering method clearly outperforms the standard analog method in those stations with lower climatic frequency for the event.

J.M. Gutiérrez, A.S. Cofiño, R. Cano, and M.A. Rodríguez, "Clustering Methods for Statistical Downscaling in Short-Range Weather Forecast", **Monthly Weather Review**, to appear (2004).

AI
met
group

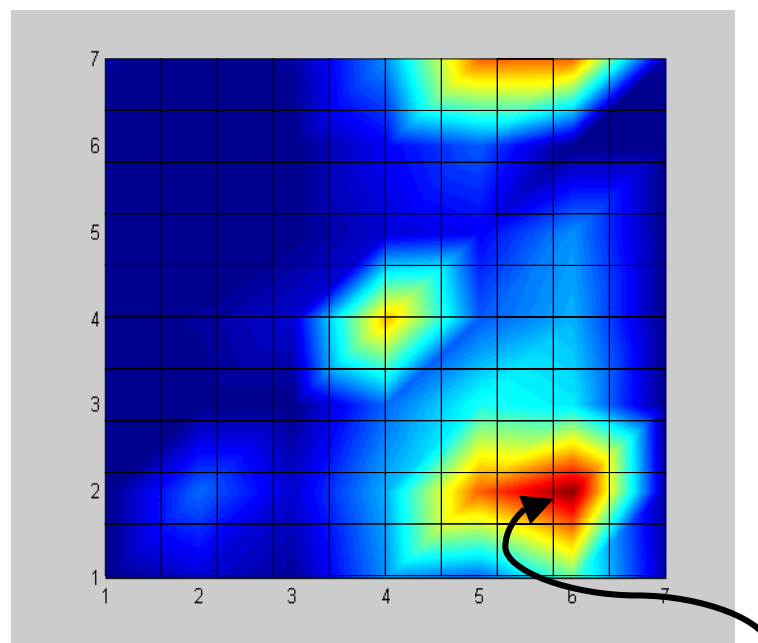


Clustering Downscaling. Seasonal Forecast.

For a given month, the daily ERA-15 patterns define a "regular" density on the SOM grid (the frequency of points falling in each of the clusters).

For instance for January (1979-1993), we get a climatological SOM density:

12 / 1 / 1979
12 / 2 / 1979
...
12 / 31 / 1993



Most frequent cluster for January

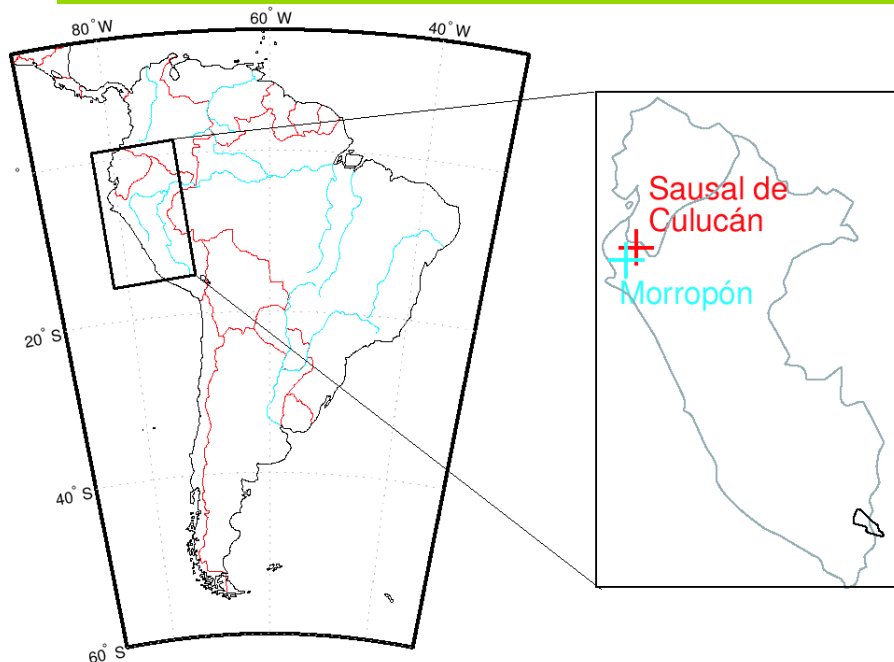
From each of the DEMETER models:

- ECMWF
- UKMO
- MPI
- CNRM

we obtain 90 daily forecasts (x 9 ensemble members) for a given season (e.g., DJF).

These patterns also define a density on the SOM. Differences and similarities with the climatological one indicate anomalies.

Seasonal Forecast. Downscaling Precipitation in Perú



We considered:

a 2D "surface oceanic" pattern:

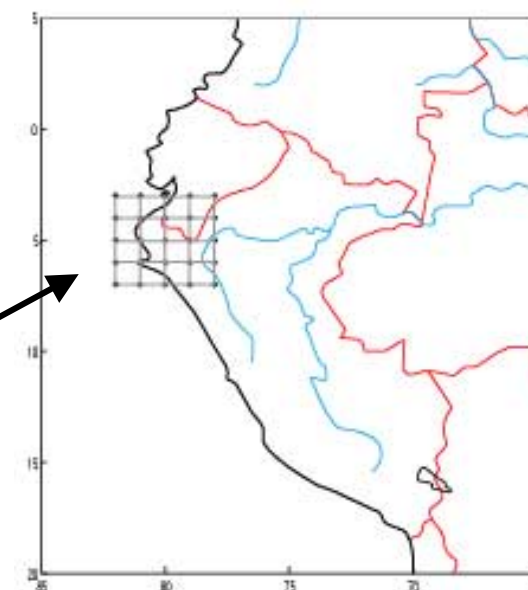
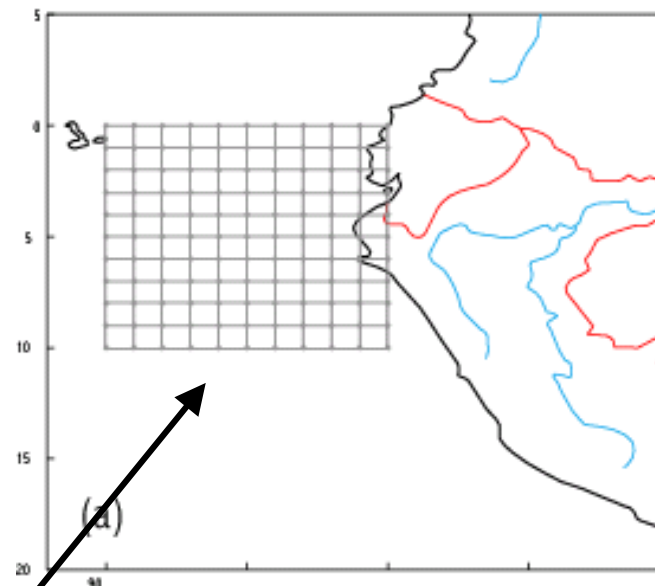
✓ **area "El Niño 1-2".**

✓ **Variable: MSLP**

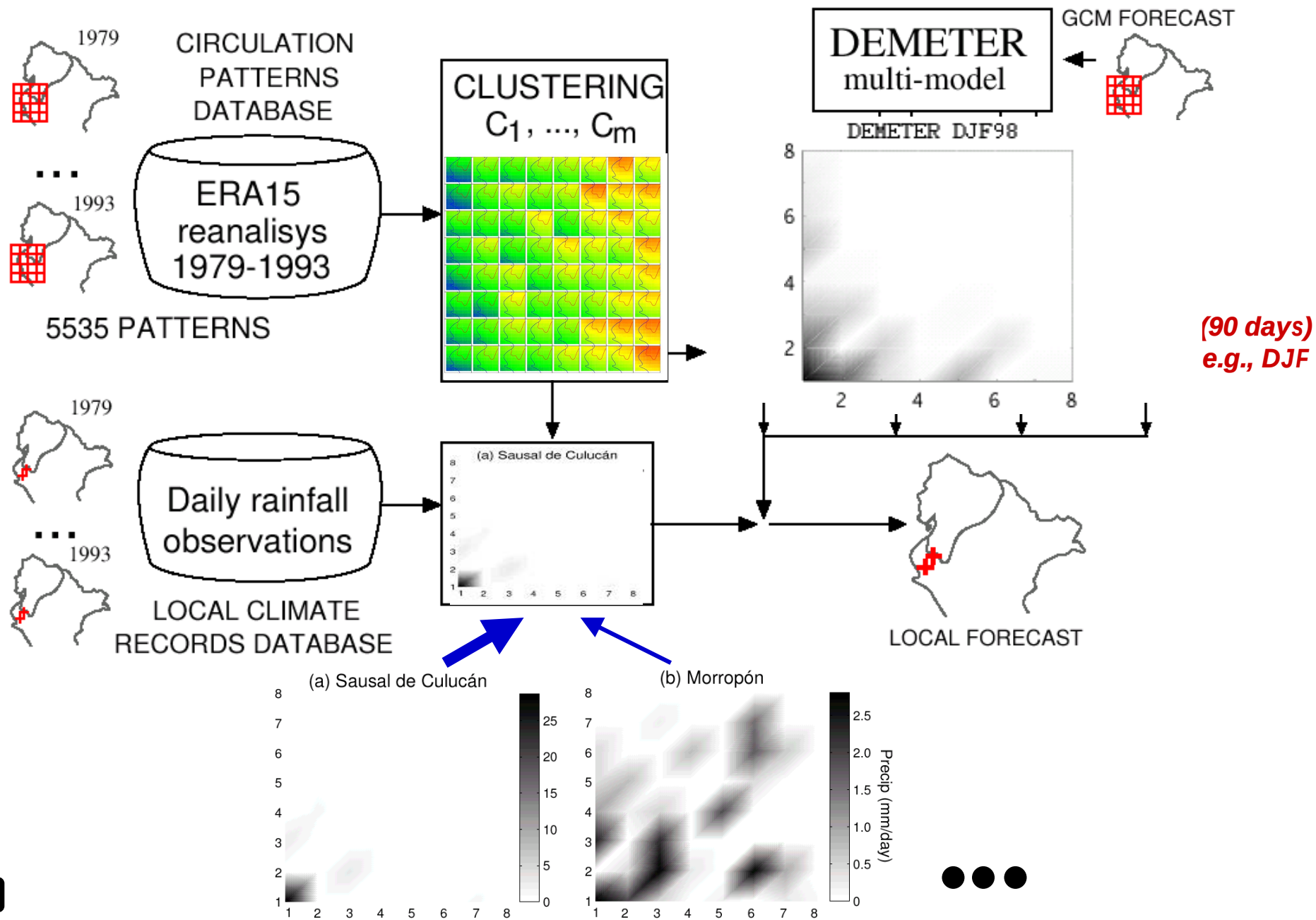
a 3D atmospheric pattern:

✓ **regional scale.**

✓ **Z, H, and T at 500mb and 850mb**

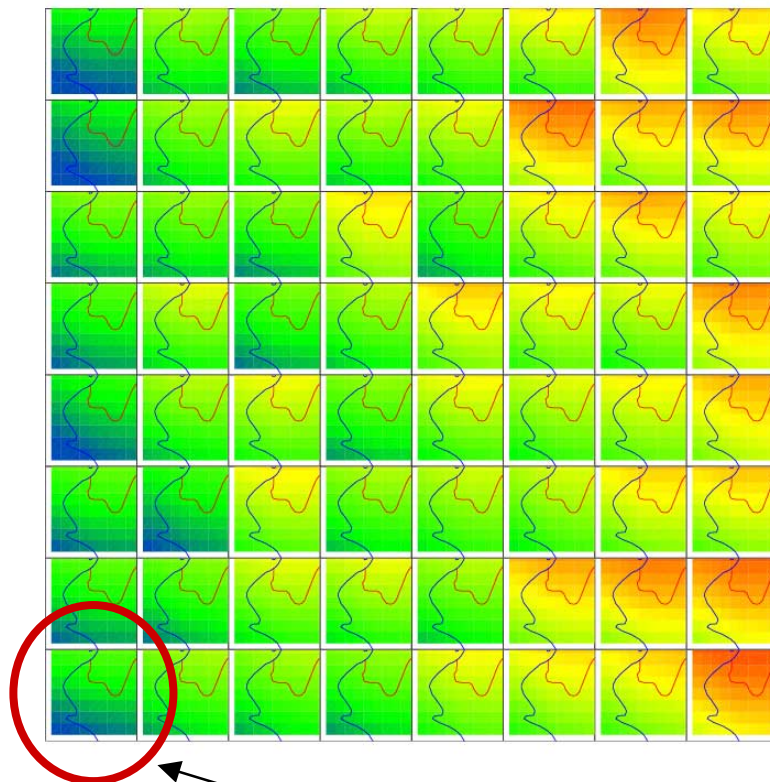


**AI
met
group**

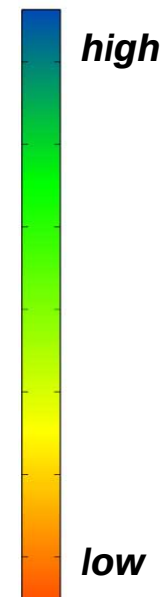
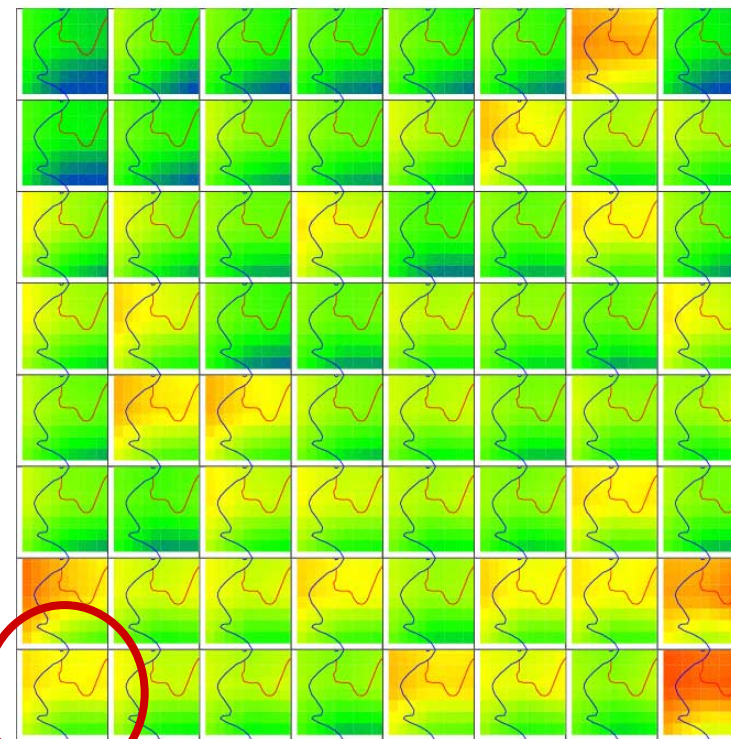


SOM Atmospheric Patterns (prototypes)

Z500 time: 00Z



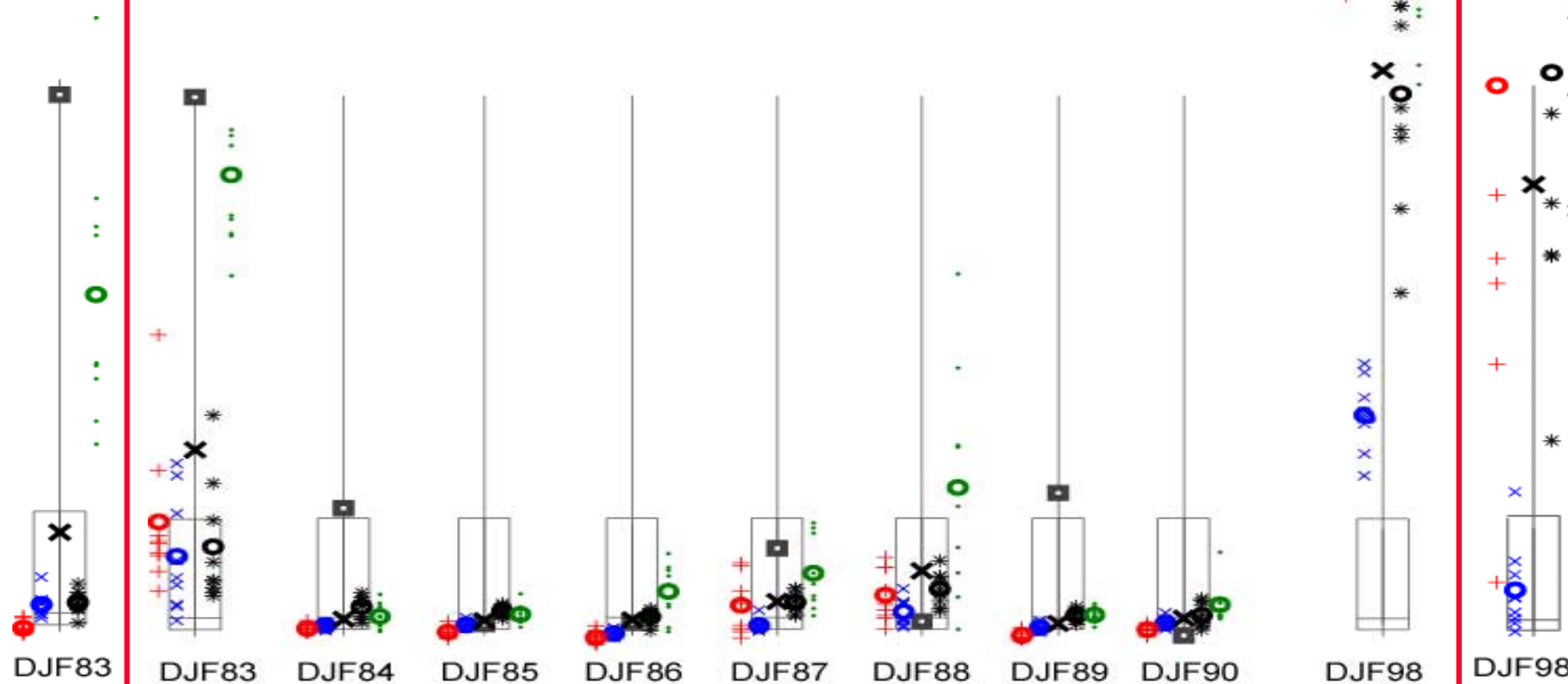
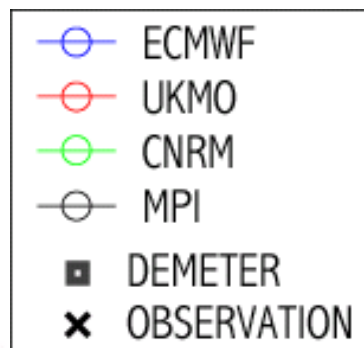
Z850 time: 00Z



Pattern associated with "El Niño" rainfall in **Sausal de Culucán**

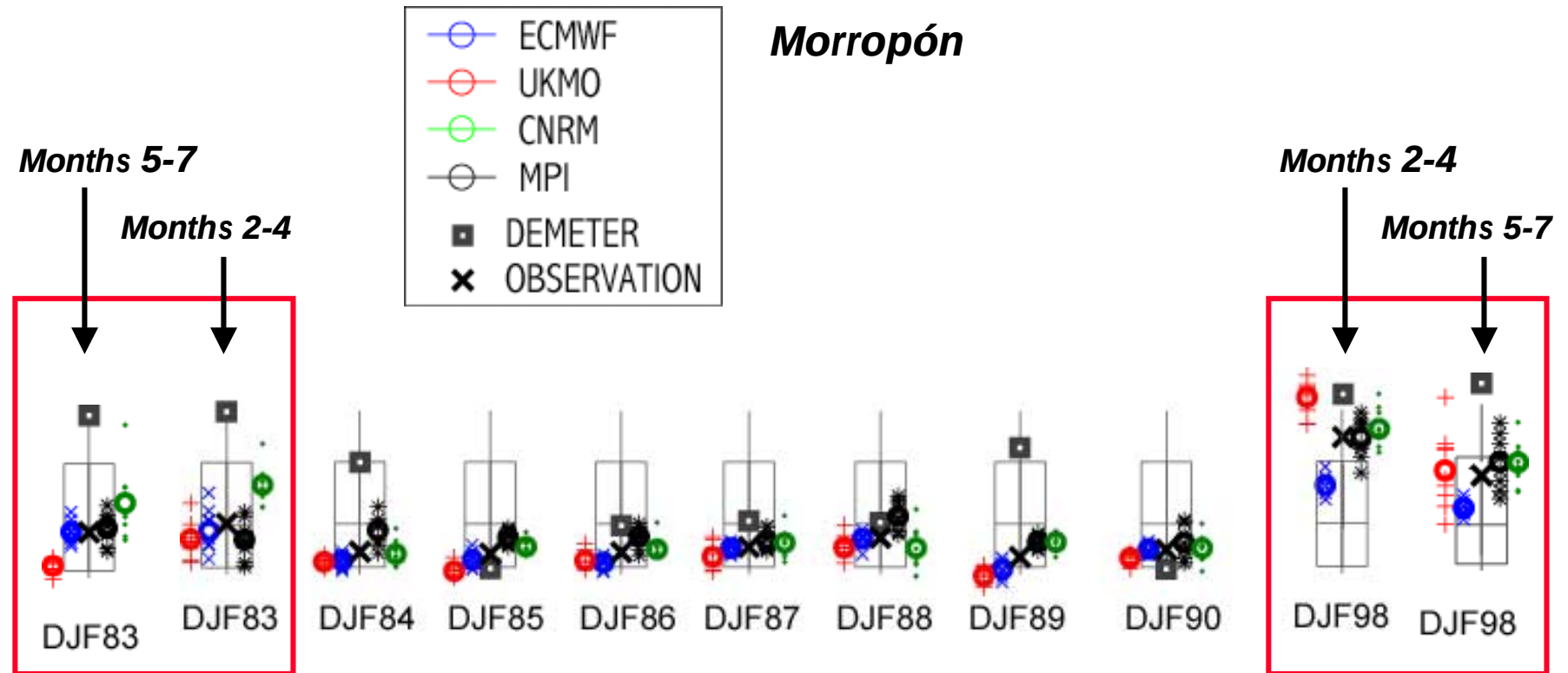
Results for months 2-4 (5-7) DJF 1983-1998

Sausal de Culucán



AI
met
group

Morropón



How **reliable** and how **anomalous** is a multimodel ensemble forecast?

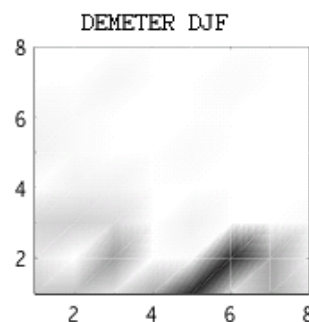
Reliability has to do with dispersion of ensemble members.

Anomaly has to do with distance to climatology.

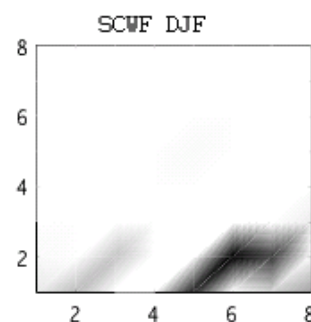
SOM Density Plots (months 2-4 Winter DJF)

**Climatology
1979-1994**

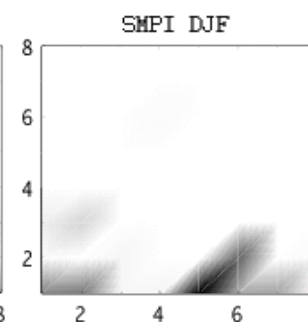
Multi Model



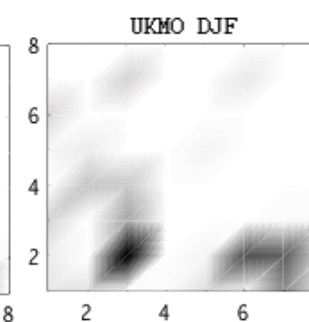
ECMWF



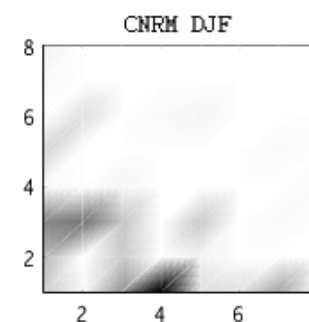
MPI



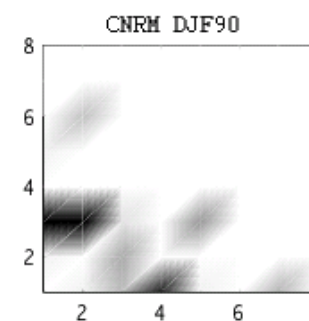
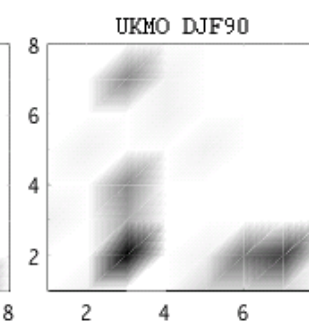
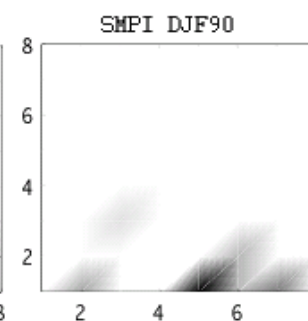
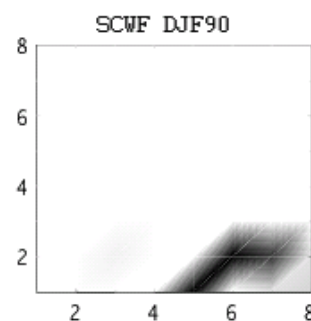
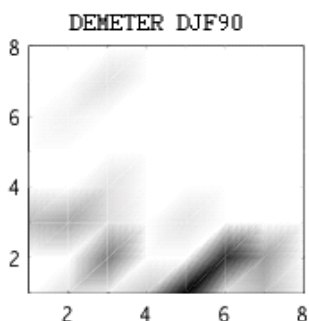
UKMO



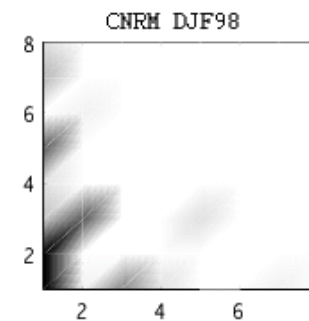
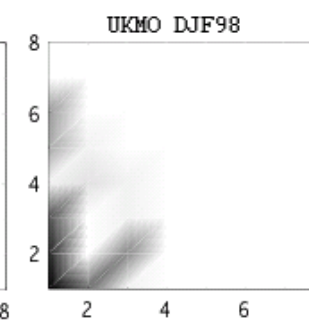
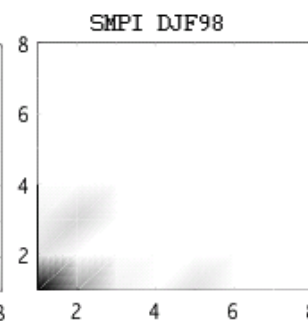
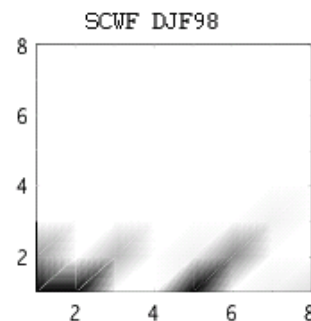
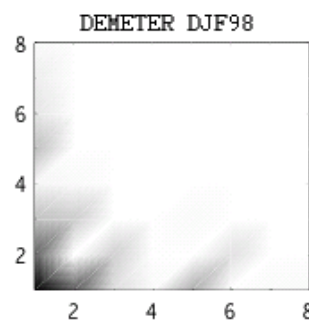
CNRM



1990



"El Niño" 1998



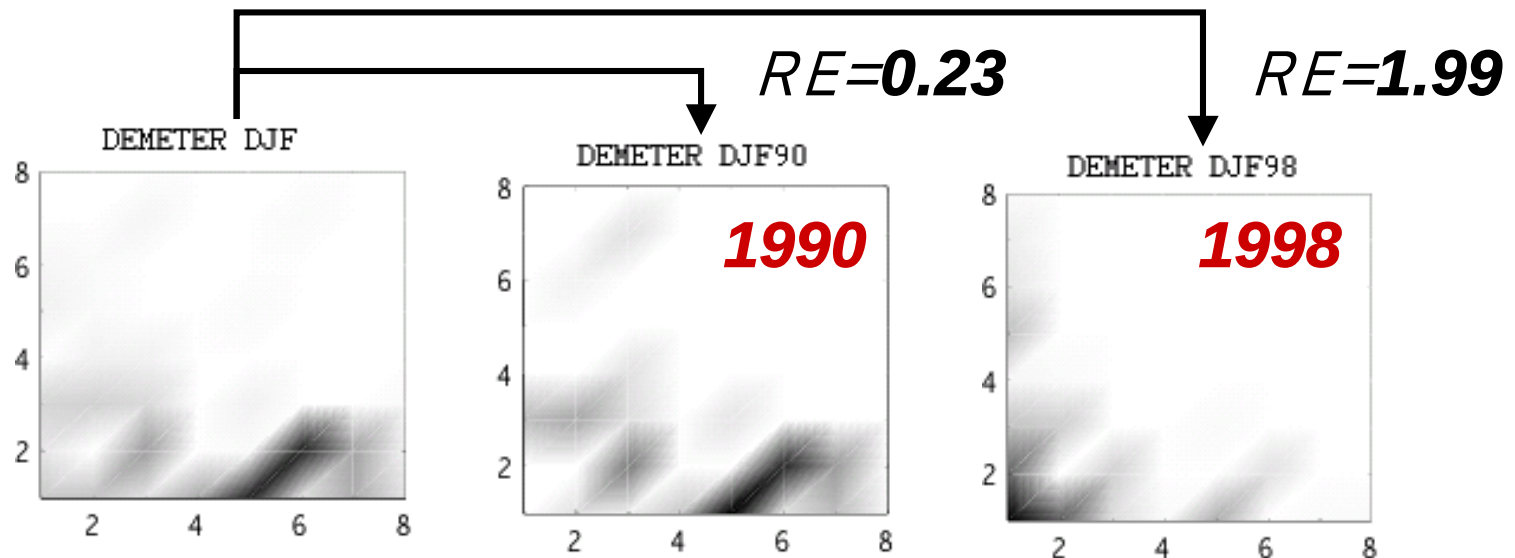
**AI
met
group**

Quantifying Anomalies & Predictability. Relative Entropy

The relative entropy characterizes the similarity between two distributions:

- ✓ *P_i (DEMETER climatology) and*
- ✓ *Q_i (distribution for a particular forecast year, e.g., "El Niño" 1998)*

$$RE = -\sum Q_i \log Q_i / P_i \quad \geq 0 \text{ and grows as the distributions become "different" (in terms of entropy, or disorder).}$$



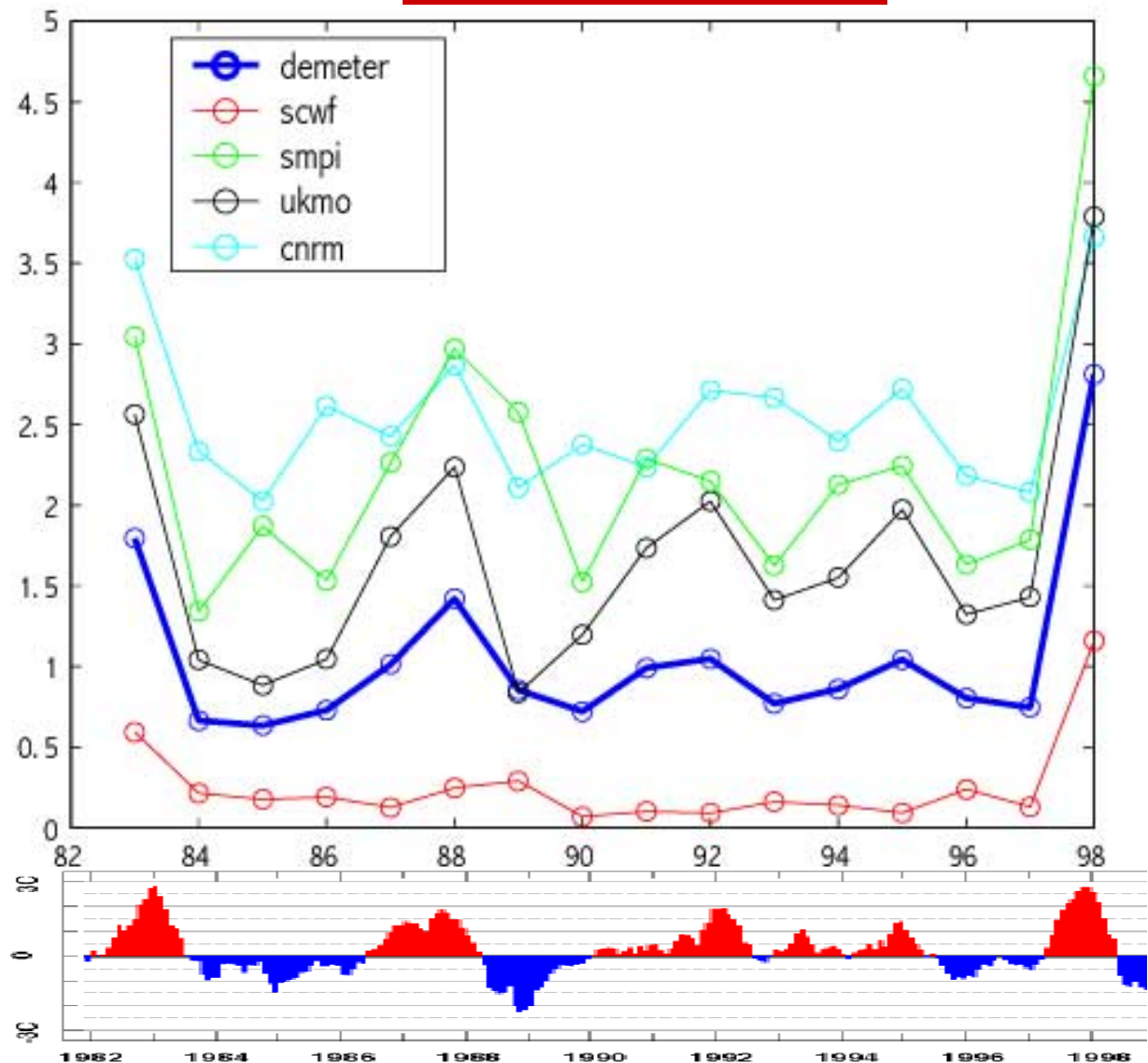
Relative Entropies for 1982-1998

High entropy values indicate an anomaly in the seasonal patterns.

The higher the entropy the larger the deviation from the climatology.

DEMETER output for "El Niño" years are associated maximum values of relative entropy.

Months 2-4 (DJF)

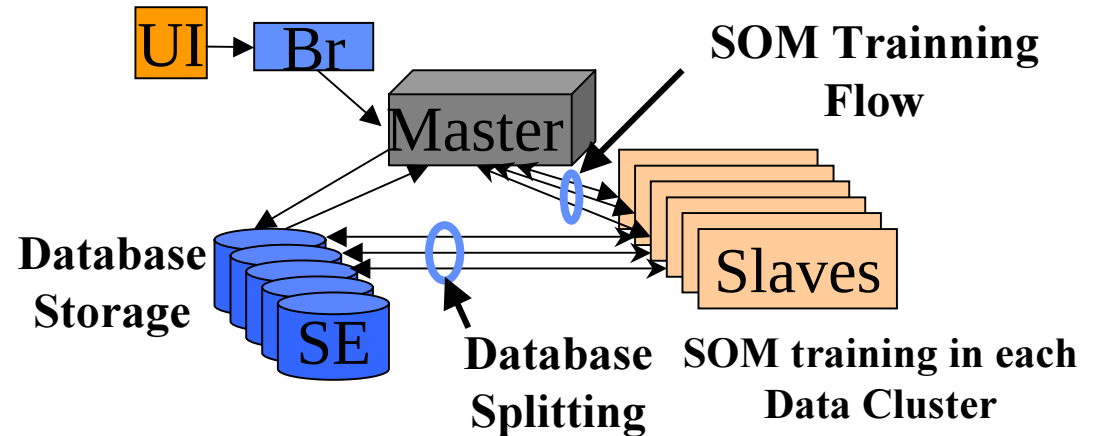


Web Portal for data-mining applications

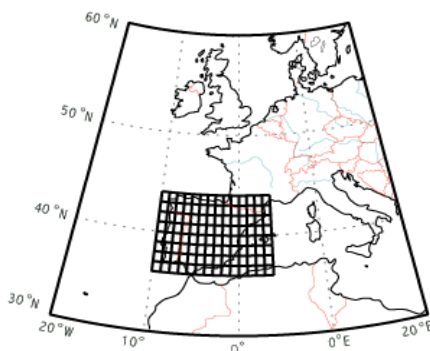
5th EU Framework

CROSSGRID

<http://www.eu-crossgrid.org>

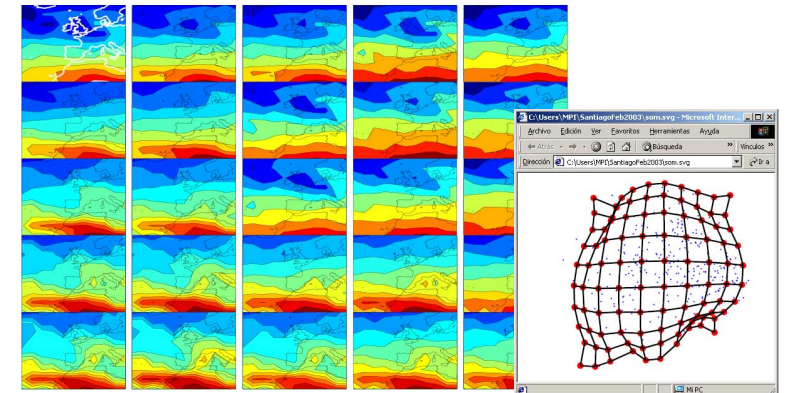


Developing clustering tools in an Interactive & user friendly framework:



Interactive selection of parameters
✓ grid area,
✓ variables, etc.
through a Web portal.

Output. XML-based interactive graphic format, binary files, etc.



**AI
met
group**