Metrics for CORDEX

William Gutowski[1] & Richard Jones[2]

[1]Iowa State University, Ames, Iowa, USA

[2]Met Office, Hadley Centre, Exeter, UK

With input from S. Solman, C. Lennard & Many Modelers

General Aims and Plans for CORDEX

- Provide a set of *regional climate scenarios* (including uncertainties) covering the period 1950-2100, for the majority of the populated land-regions of the globe.
- Make these *data sets readily available and useable* to the impact and adaptation communities.
- Provide a *generalized framework for testing and applying* regional climate models and downscaling techniques for both the recent past and future scenarios.
- Foster coordination between regional downscaling efforts around the world and *encourage participation* in the downscaling process by local scientists/organizations

Metrics for CORDEX

Goals

1. Model performance versus variety of observations
2. Succinct
3. Side-by-side comparison of models on same graph

Questions to keep in mind

A. What is missing or not needed?
B. Should the goals be modified?

Metrics for CORDEX

Two categories

A.Basic assessment common to all regions to provide a baseline

B. Region-targeted assessment to be determined by regional activities

VIA input

CORDEX: Basic Metrics

Fields: Core Data Variables

- 2-m Temperature
- 2-m Daily Tmax
- 2-m Daily Tmin
- 2-m Specific Humidity
- 10-m Wind Speed
- Precipitation
- Mean Sea-Level Pressure

CORDEX: Basic Metrics

Maps of Biases

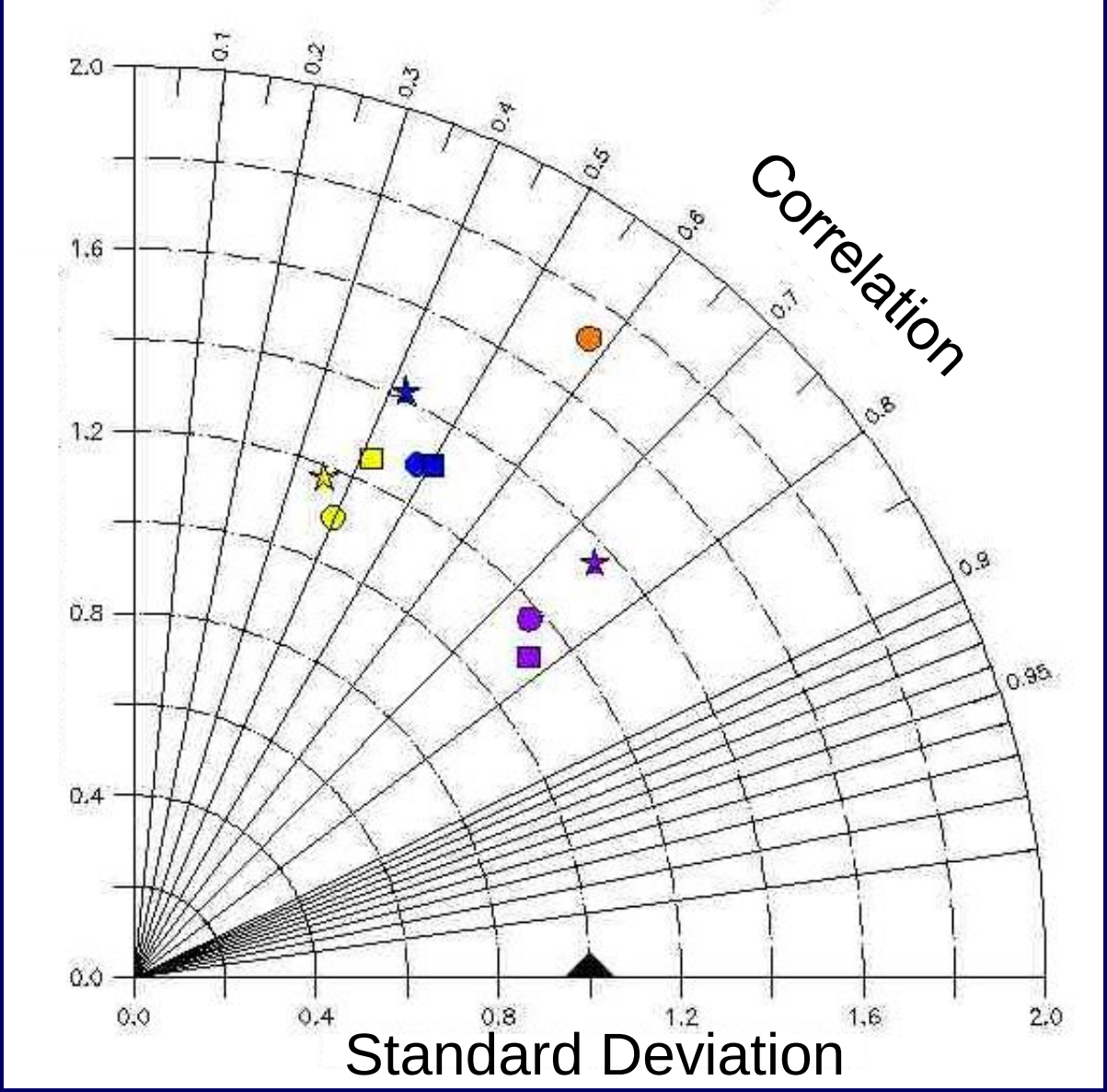
Å2-m Temperature
ÅPrecipitation
ÅMean sea-level pressure

Taylor Diagrams

ÅSeasonal means
ÅInterannual variability ï spatial correlation
ÅInterannual variability ï temporal correlation

All fields should be interpolated to the common grid for model output when constructing model-observation differences.

What is a Taylor Diagram?

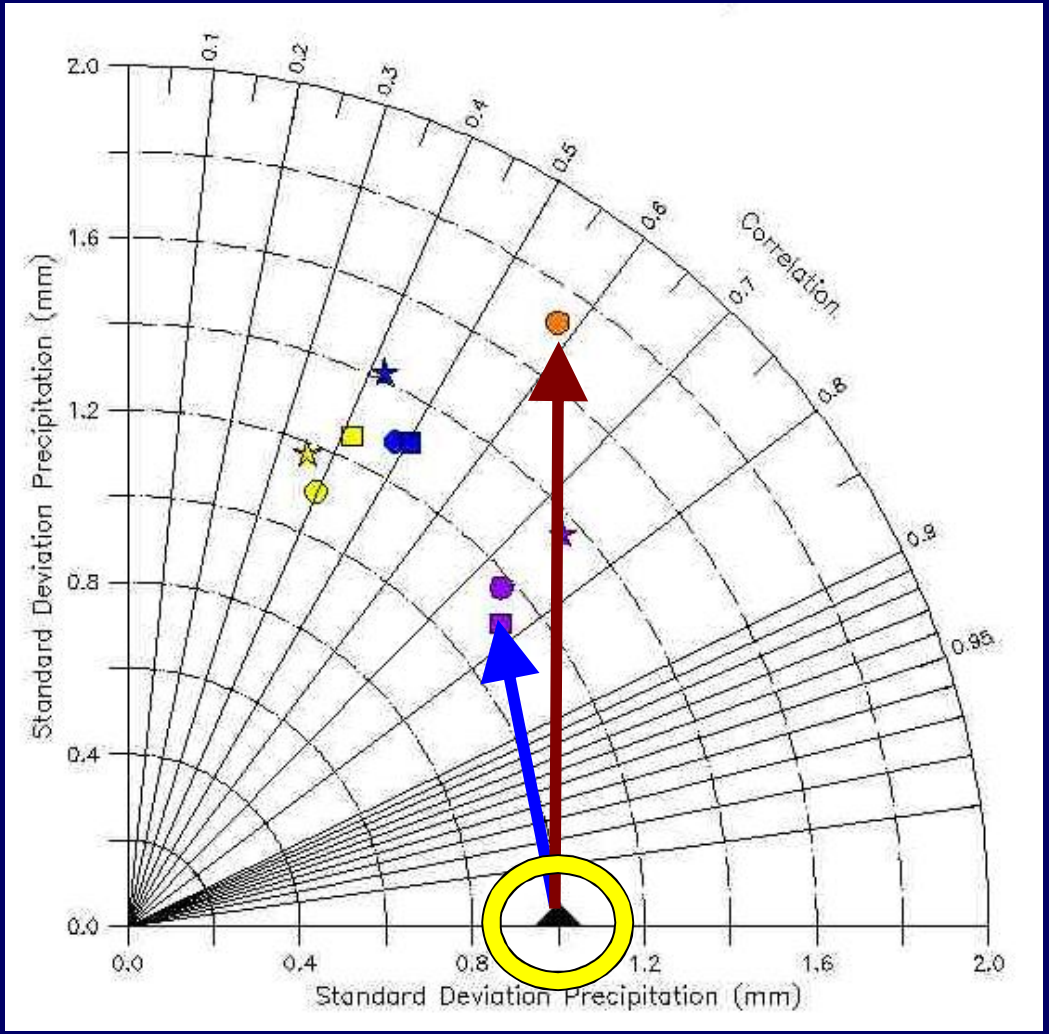


What is a Taylor Diagram?

Correlation = correlation with reference field

Standard Deviation = Amplitude of variability about the mean (normalized by reference variability)

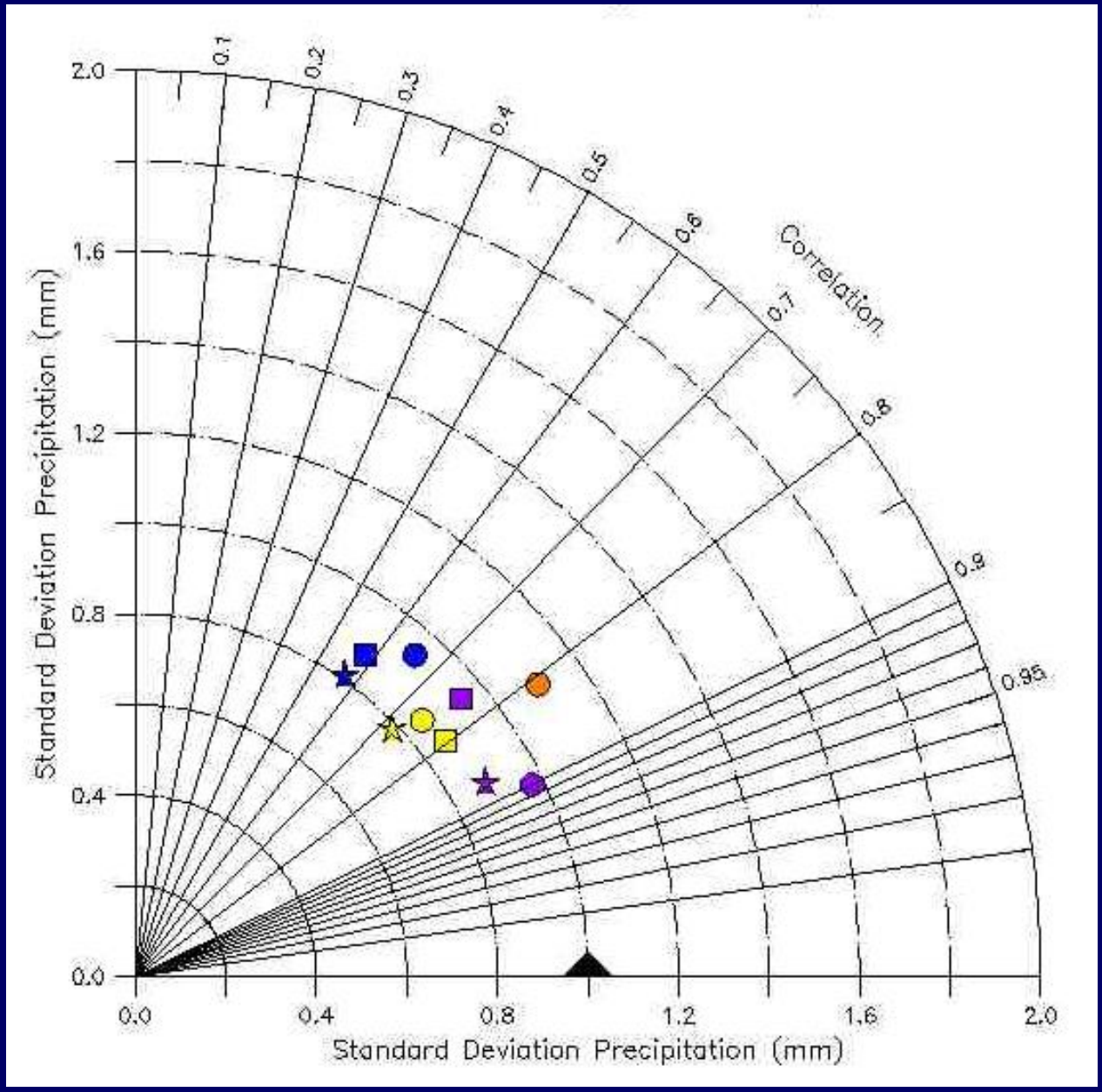
Distance to reference point = Bias-corrected RMS difference (i.e., a measure of error)



January Precipitation

Different colors = Different models

Different shapes = Different domain sizes

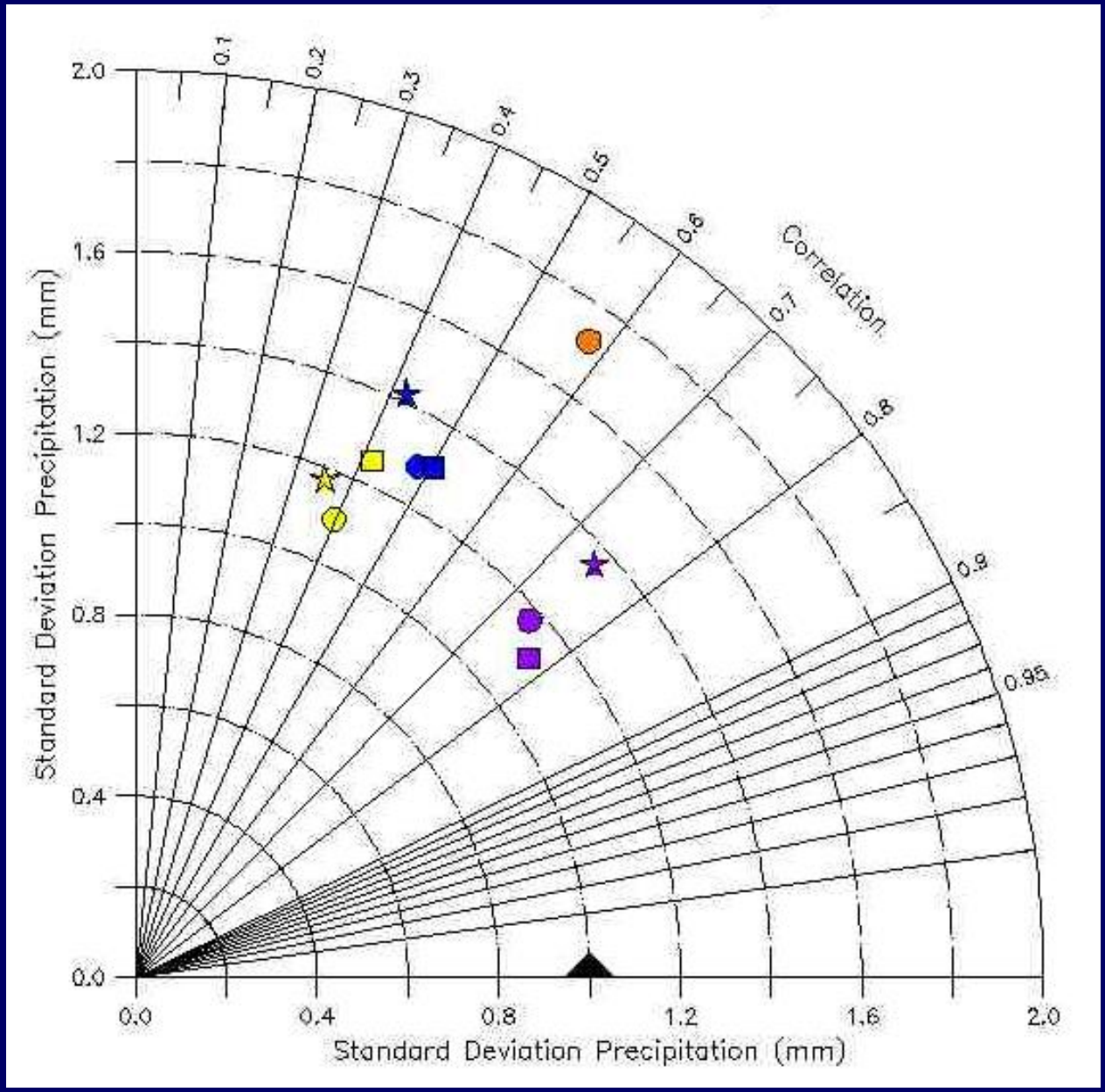


(C. Anderson, 2005)

June Precipitation

Different colors = Different models

Different shapes = Different domain sizes



(C. Anderson, 2005)

CORDEX: Basic Metrics

Reference Fields

Å**CRU** (monthly mean, 0.5º, 1901-2006) - 2m mean/max/min temperature, precipitation
Å**Willmott** (monthly mean, 0.5º, 1901- 2008) - 2m mean temperature, precipitation
Å**TRMM 3B42** (3-hour, 0.25º, 1998-2009) - precipitation
Å**ERA-INTERIM** (3-6 hour, daily, 0.75º, 1989-2009) - 2m mean/max/min temperature, precipitation, zonal and meridional wind

All fields should be interpolated to the common grid for model output when constructing model-observation differences.

CORDEX: Taylor Diagrams

Seasonal Means (DJF, MAM, JJA, SON)

Measures: Error in spatial patterns of mean fields

Interannual Variability: Spatial Correlation (seasonal maps)

Measures: Error in spatial patterns of interannual variability

Interannual Variability: Temporal Correlation (maps of annual time-series correlation)

Measures: Error in times series of interannual variability

CORDEX Region-Targeted Metrics

A. Metrics defining model quality relevant to region-specific processes, e.g. monsoonal circulation

B. Metrics relevant to users in the region, e.g. onset of dry spells, start of growing season, etc.

What should region groups chose for metrics?

(i) Sub-regions?
(ii) Time periods?
(iii) Diurnal or other periodic behavior?
(iv) Targeted processes?

NOTE: Need reliable observations for any metric.

CORDEX Region-Targeted Metrics

B. Metrics relevant to users in the region, e.g. onset of dry spells, start of growing season, etc.

Challenge to VIA Applications and Training:

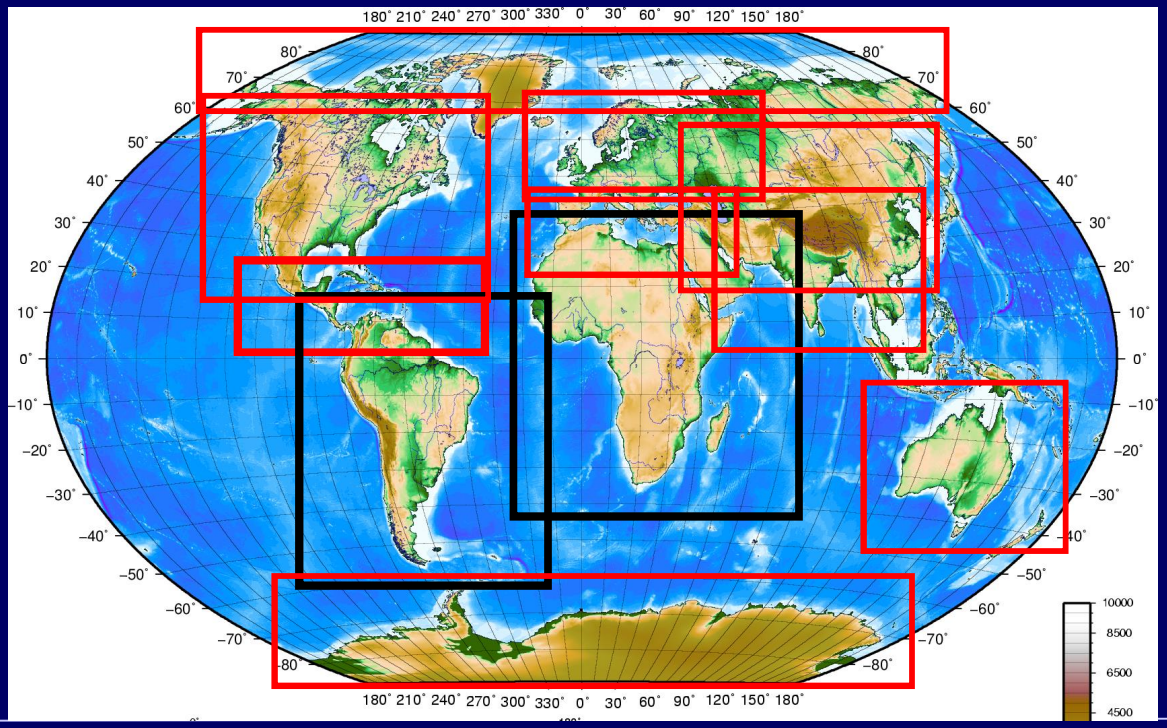
- What additional metrics are relevant and useful?
- How would you want to assess models providing output?

NOTE: Need reliable observations for any metric.

Recall This General Aim for CORDEX

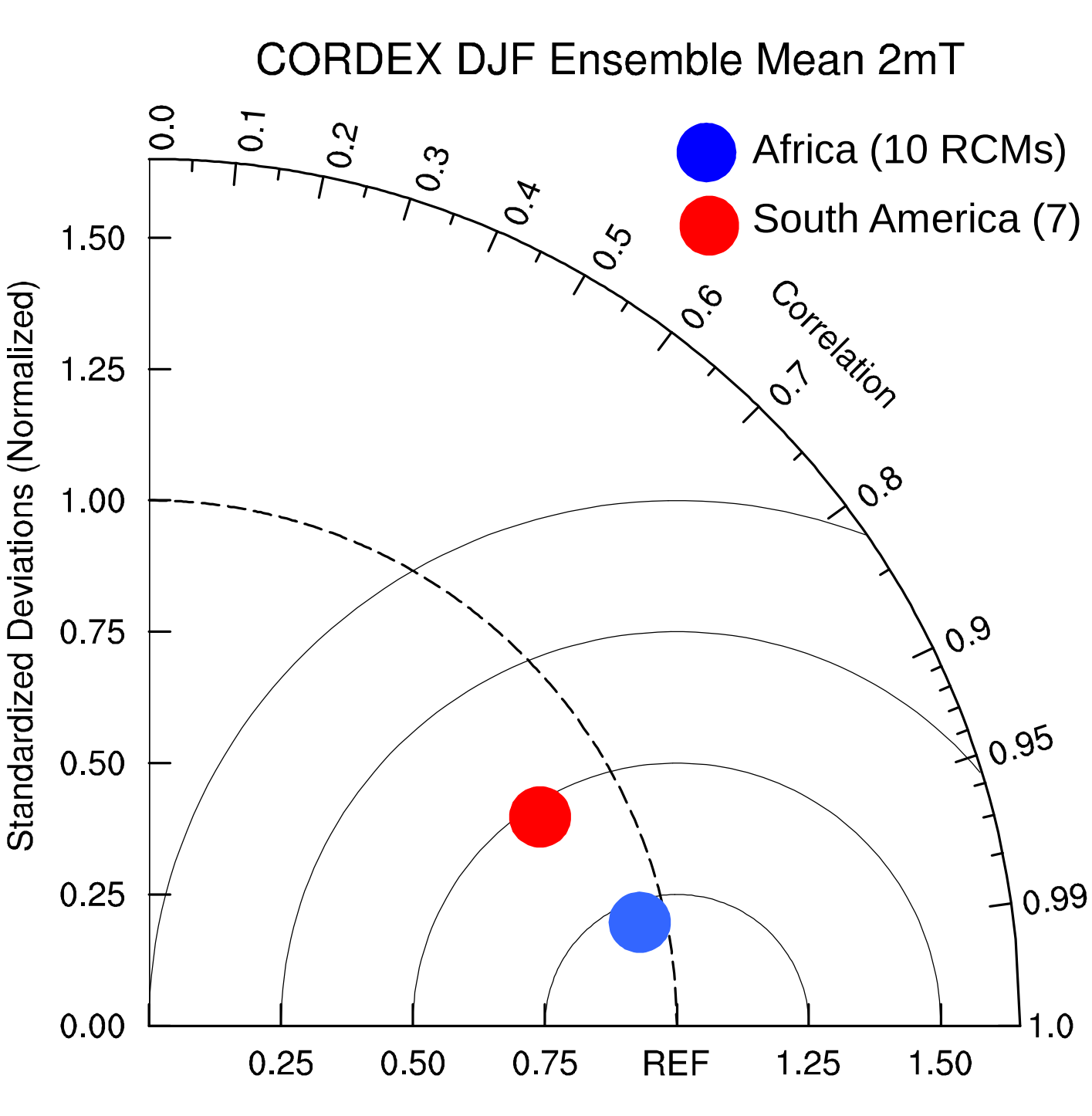
- Provide a *generalized framework for testing and applying* regional climate models and downscaling techniques for both the recent past and future scenarios.

How does downscaling compare across CORDEX regions?
(e.g., ñTransferabilityòï Takle et al., 2007, *Bulletin AMS)*



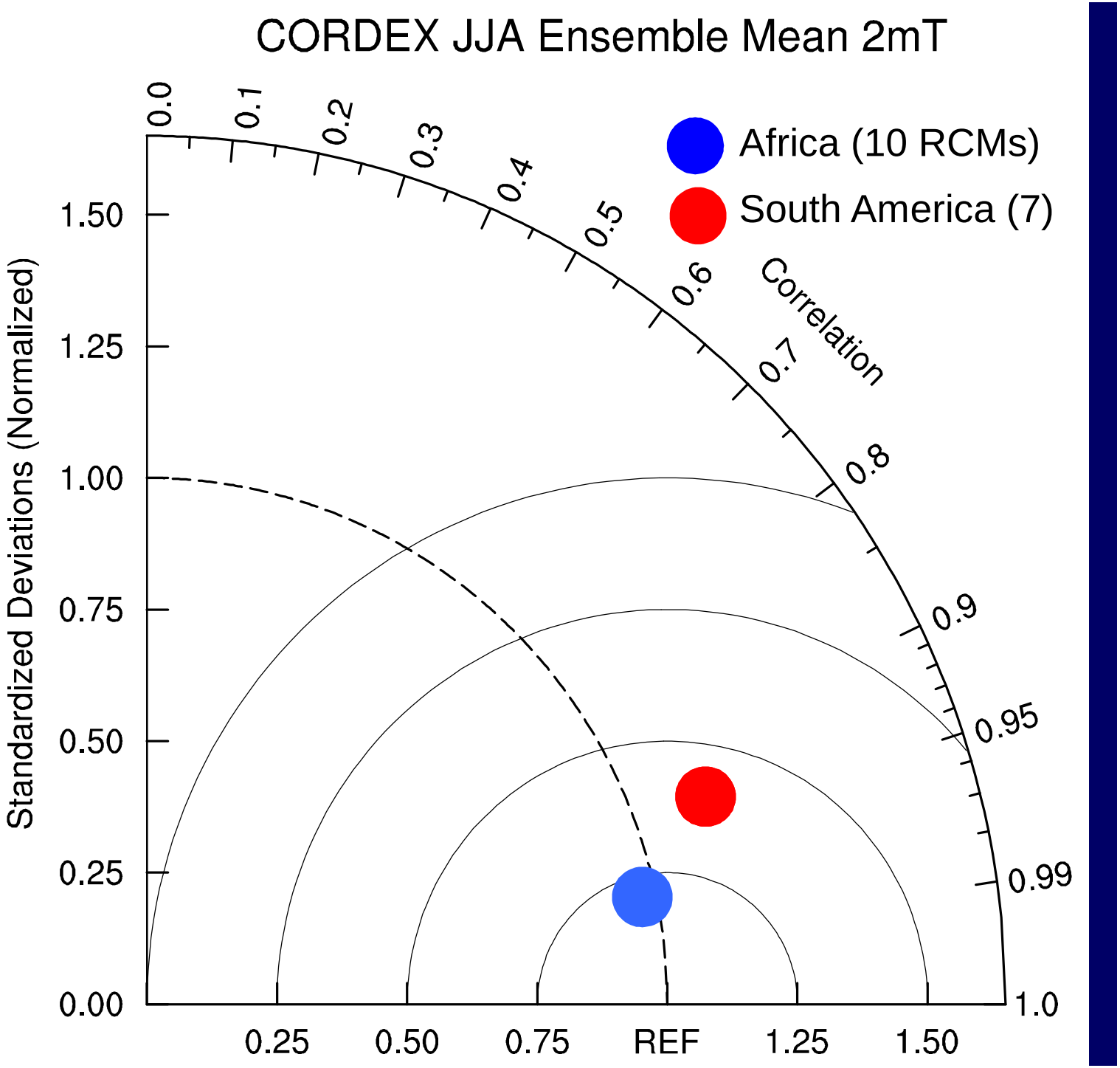
RCMs vs. ERA-Interim

Developing a Pan-CORDEX View



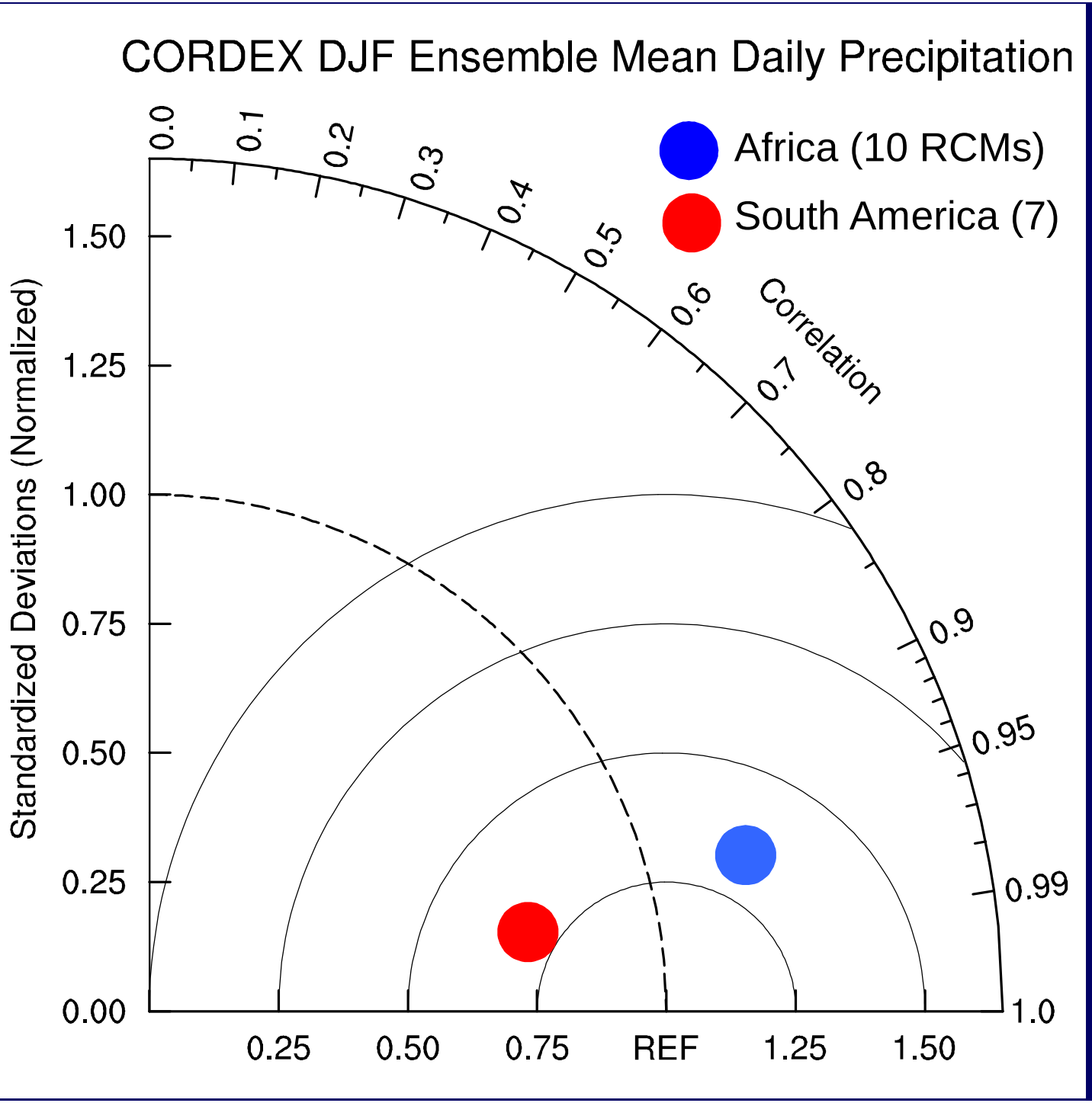
RCMs vs. ERA-Interim

Developing a Pan-CORDEX View



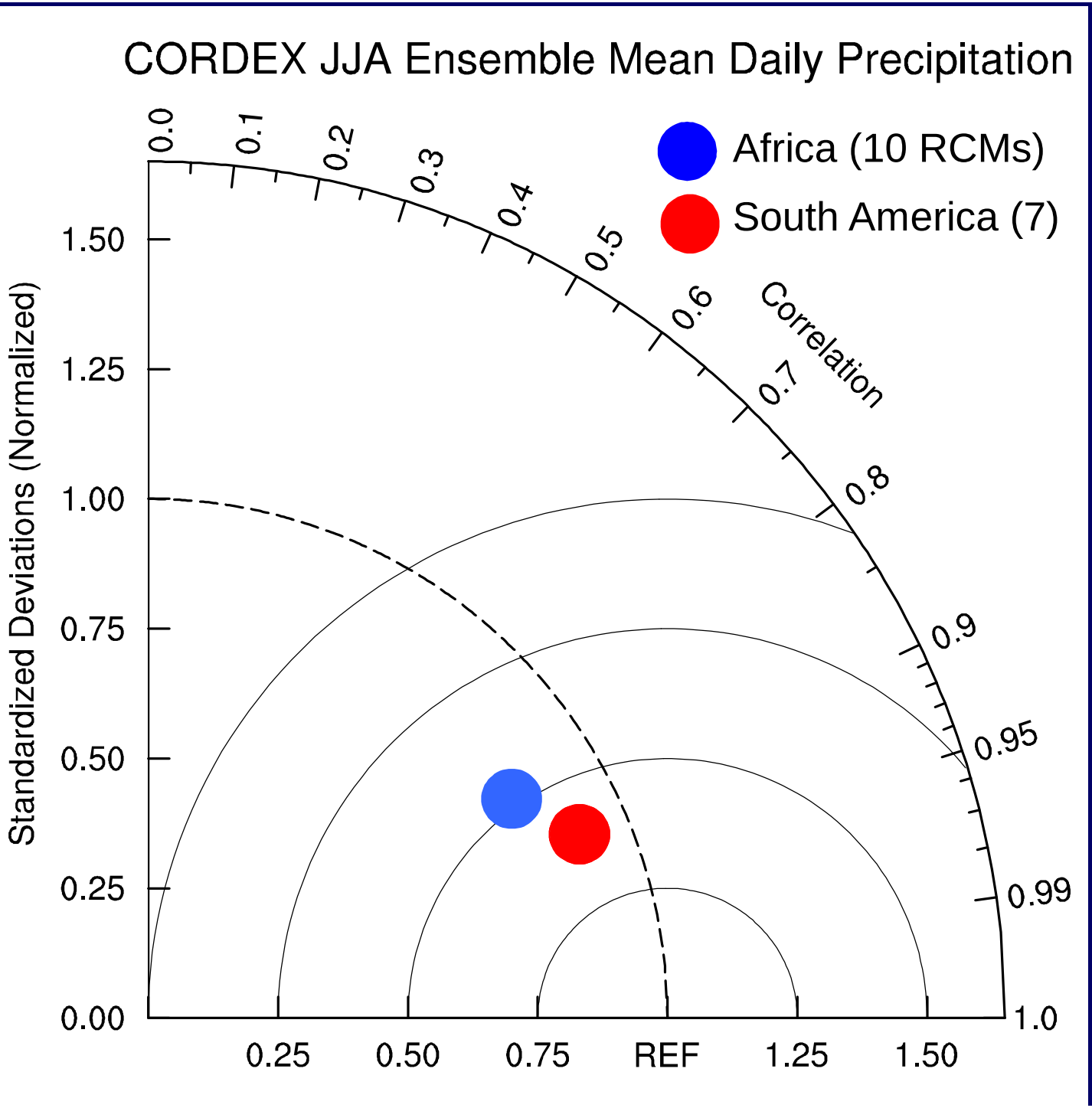
RCMs vs. ERA-Interim

Developing a Pan-CORDEX View



RCMs vs. ERA-Interim

Developing a Pan-CORDEX View

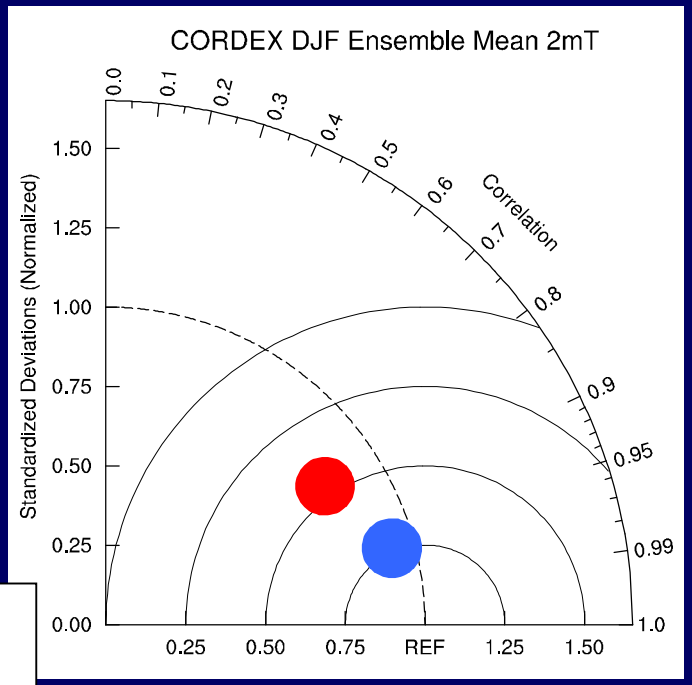


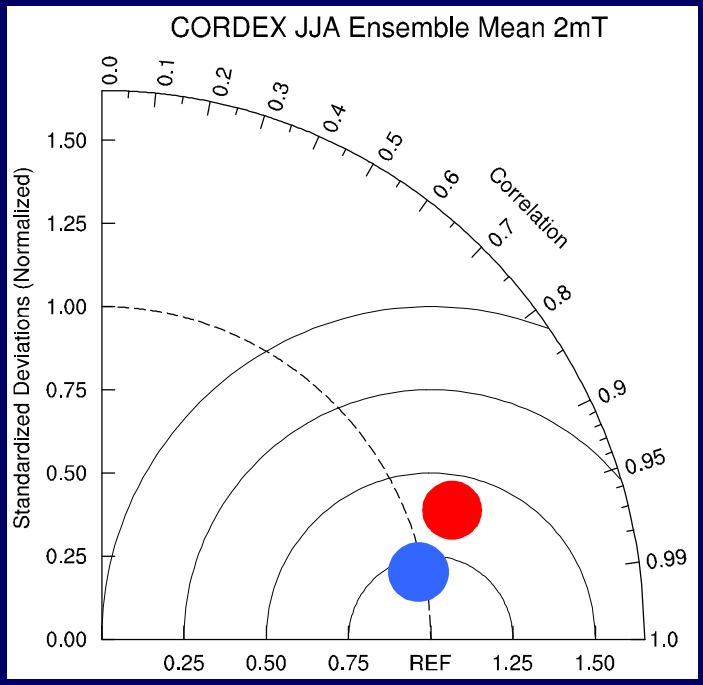
RCMs vs. ERA-Interim

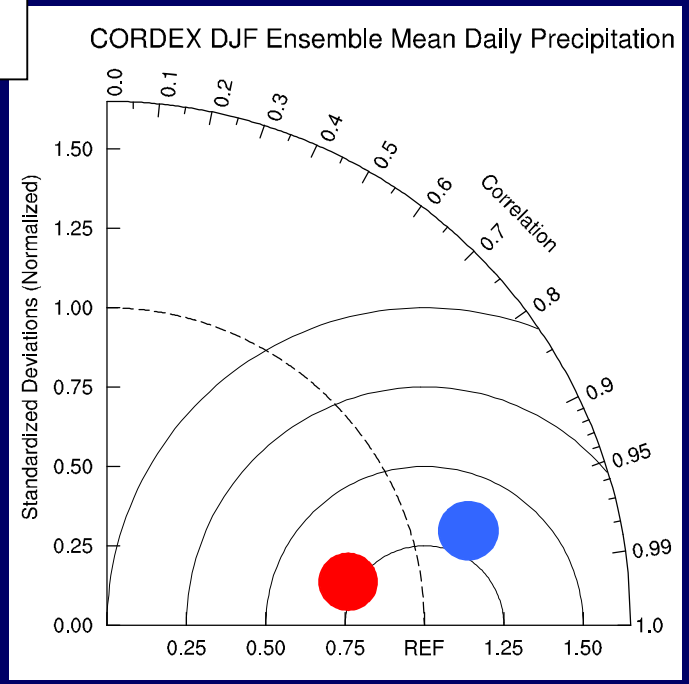
Africa (10 RCMs)

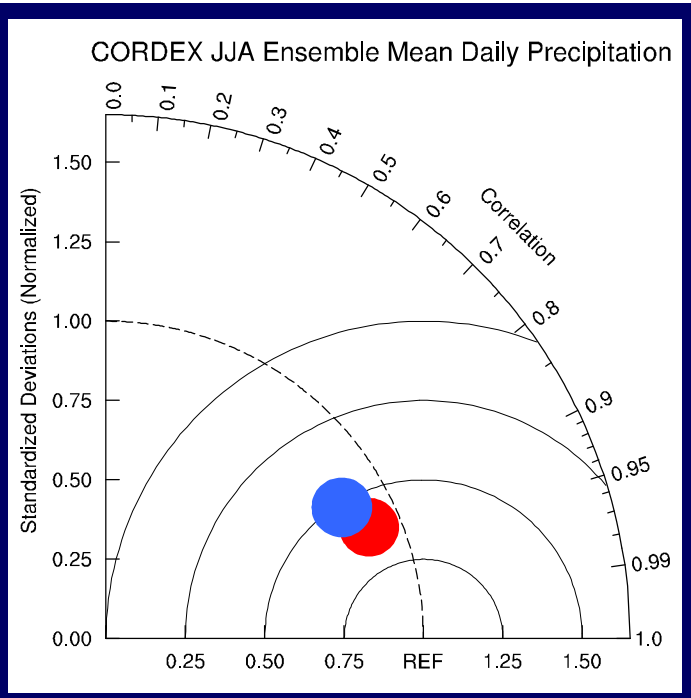
South America (7)

Developing a Pan-CORDEX View

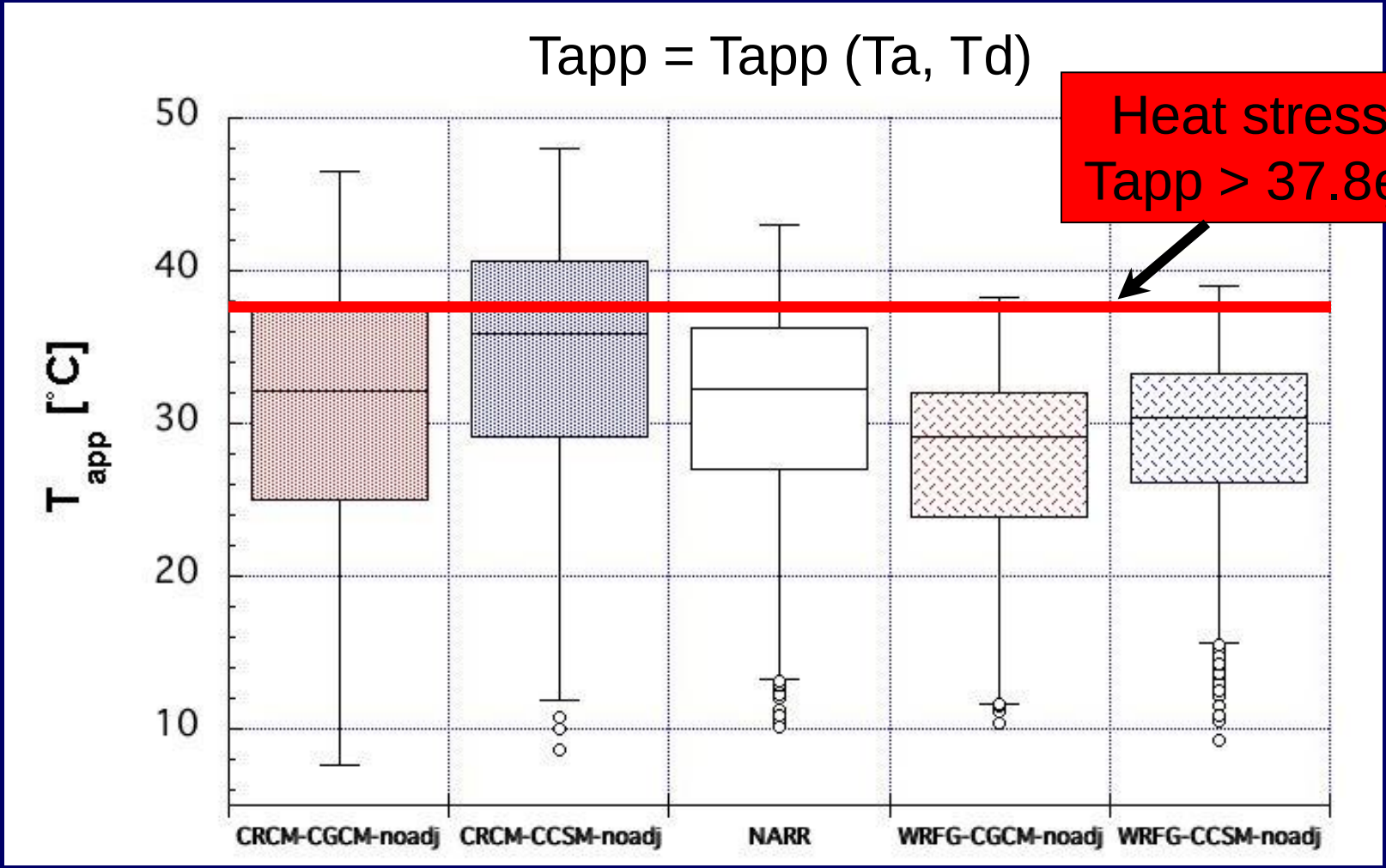




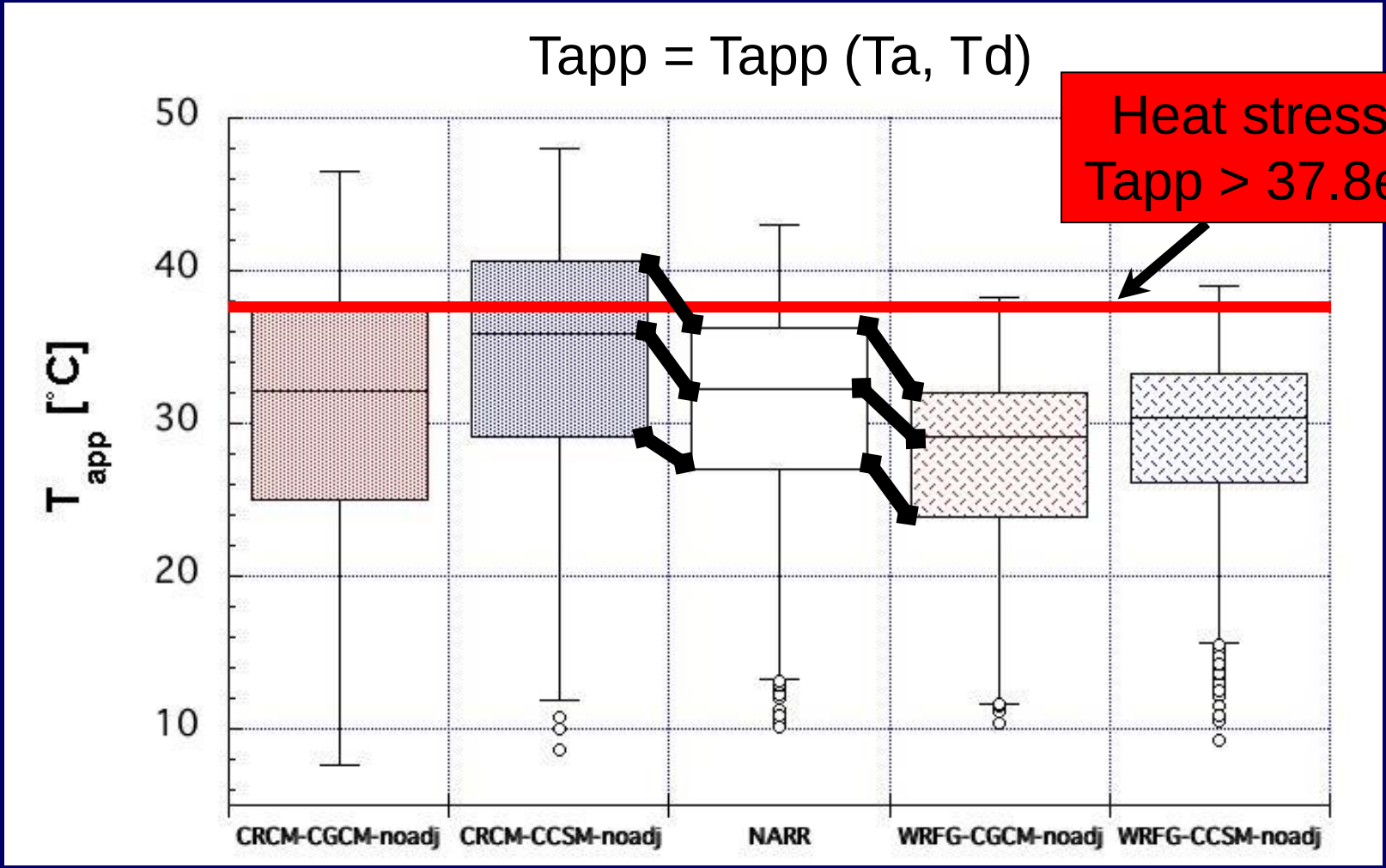




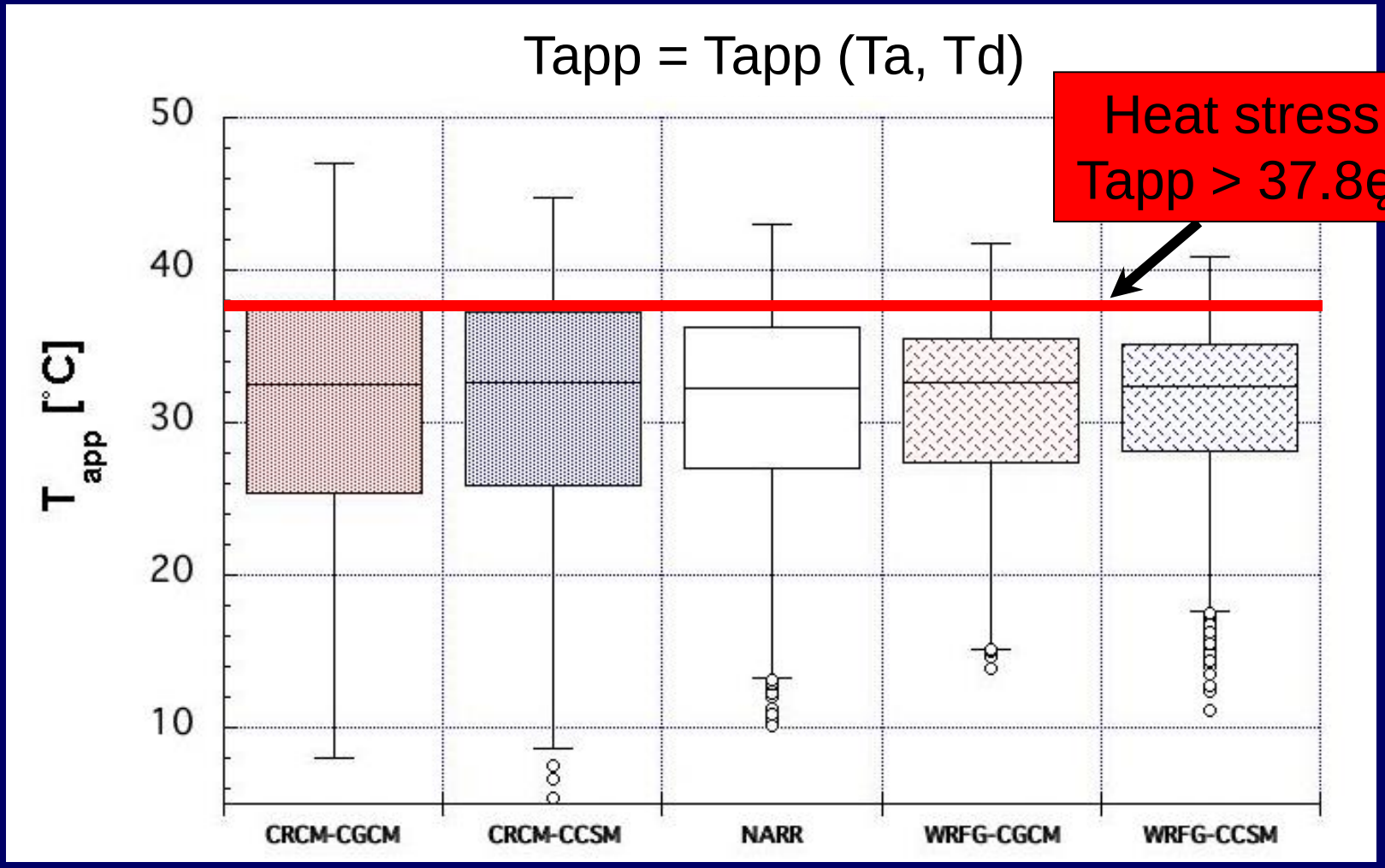
Alternative Metrics?



Alternative Metrics => Bias Correction?



Alternative Metrics => Bias Correction?



Questions to keep in mind

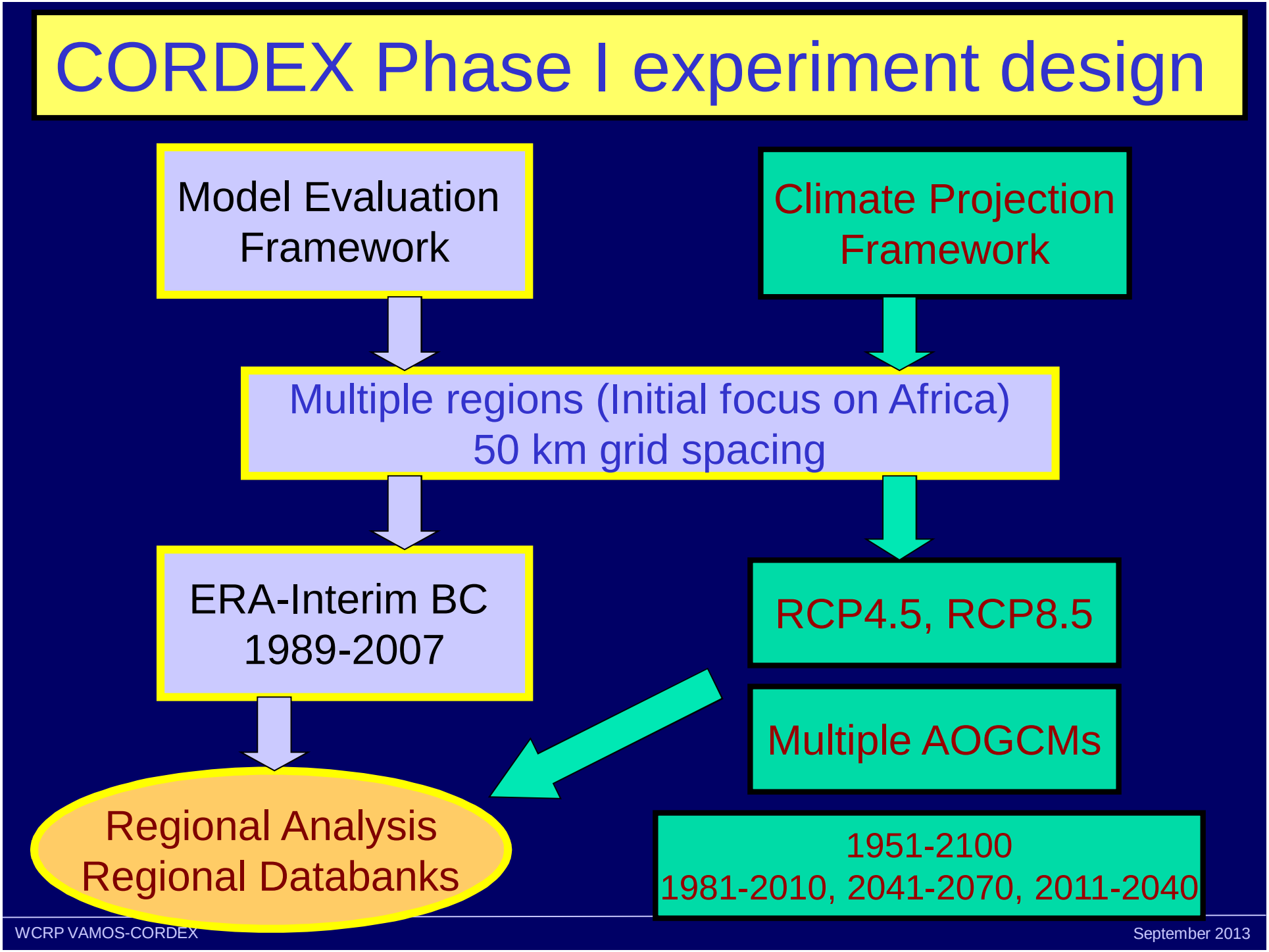
A. What is missing or not needed?

B. Should the goals be modified?

¡Gracias!

Additional Slides

CORDEX Phase I experiment design



CORDEX Statistical Downscaling: Baseline Requirements

- users choose their predictors,
  (based on standard GCM atmospheric variables)
- priority predictands: precipitation, Tmax and Tmin
- time resolution: daily
- spatial resolution: Ò0.5ę(RCM baseline)

- Standard downscaling: to a grid
- Special downscaling: to stations to maximize value to end users
- Present-day downscaling uses ERA-interim reanalysis as predictors
- Projected changes follow RCM requirements, as feasible
  (e.g., scenarios, time periods, etc)

CORDEX metrics

Document currently being finalised. Guiding principles:

1. Should be a summary statement about model performance compared to a range of available observations
2. Should be succinct
3. Should allow side-by-side comparison of models on the same graph

Proposal is Taylor diagrams of e.g. correlation, rms of seasonal means, spatial variability, interannual variability etc and maps of surface temp, precip and pmsl biases

Two categories:

A. Basic assessment common to all regions to provide a baseline

B. Region-targeted assessment to be determined by regional activities defining:

(i) Metrics defining model quality relevant to region-specific processes, e.g. North America monsoon;

(ii) Metrics relevant to users in the region, e.g. onset of, duration and dry-spell length within rainy season;

Observations and reanalysis

- ✓ **CRU** (monthly mean, 0.5º, 1901-2006)
  2m mean/max/min temperature, precipitation
- ✓ **Willmott** (monthly mean, 0.5º, 1901- 2008)
  2m mean temperature, precipitation
- ✓ **TRMM 3B42** (3-hour, 0.25º, 1998-2009)
  precipitation
- ✓ **ERA-INTERIM** (3-6 hour, daily, 0.75º, 1989-2009)
  2m mean/max/min temperature, precipitation,
  zonal and meridional wind

all data sets are interpolated onto the 0.44º rotated grid by CDO, bilinear interpolation (remapbil)

Diagnostics

Å two periods:
1998-2008 (TRMM)
1992-2006 (3 year spin up, CRU ends in 2006)

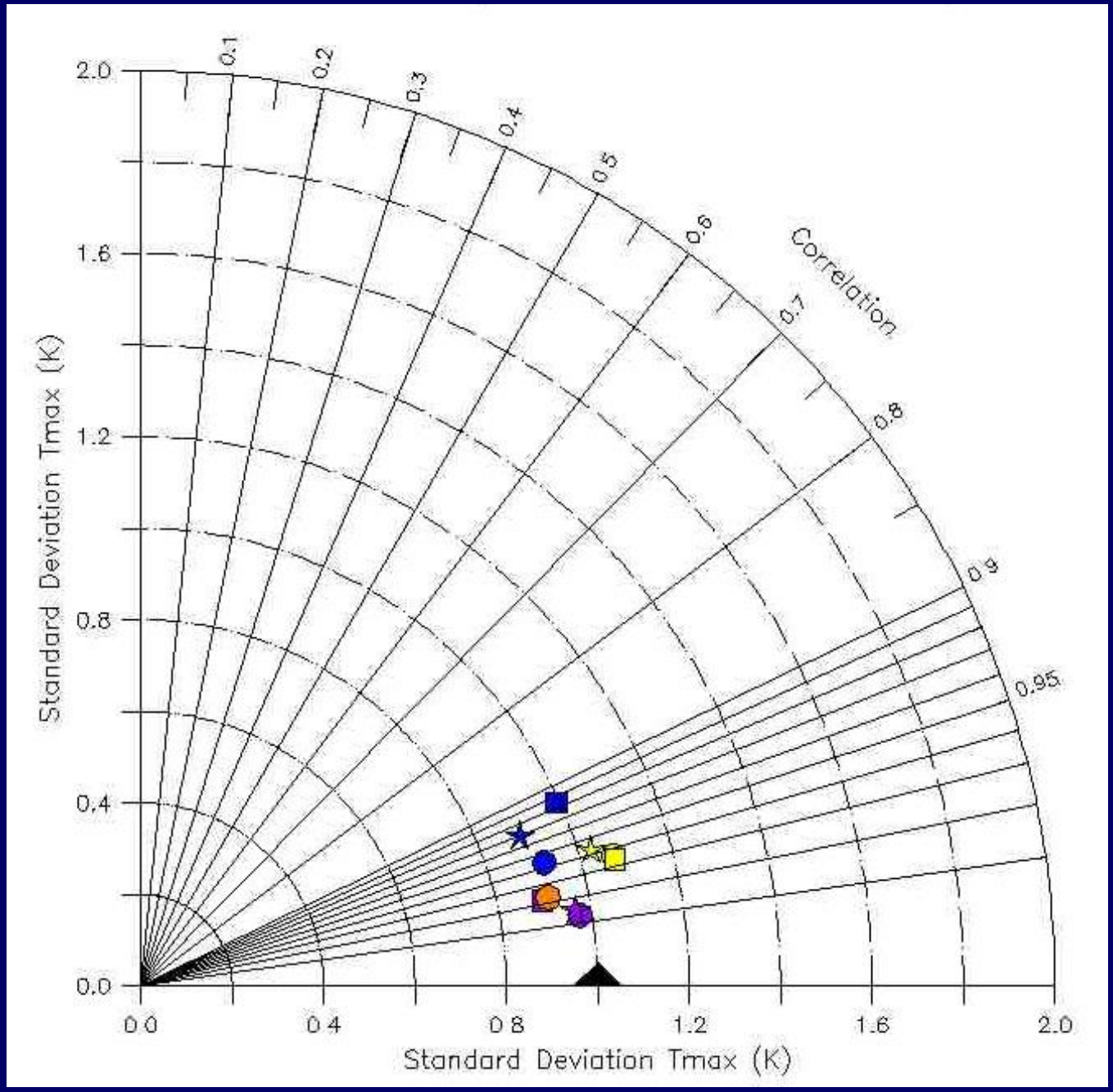
Å four seasons:
JFM, AMJ, JAS, OND

seasonal mean
interannual variability (standard deviation)
annual cycle (daily data, 50-day low pass filter)
diurnal cycle (3-hour data)

January Tmax

Different colors = Different models

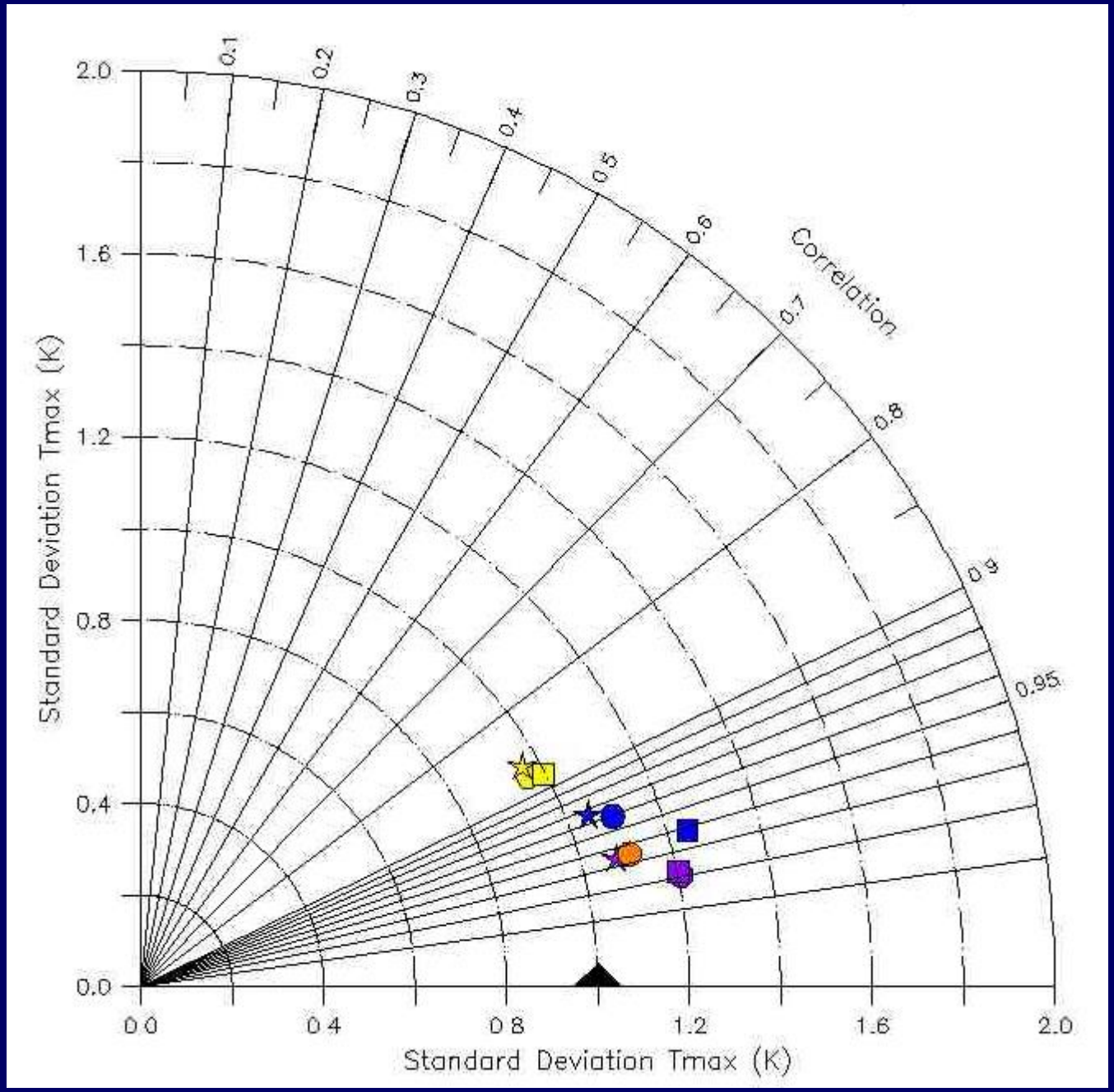
Different shapes = Different expts.



July Tmax

Different colors = Different models

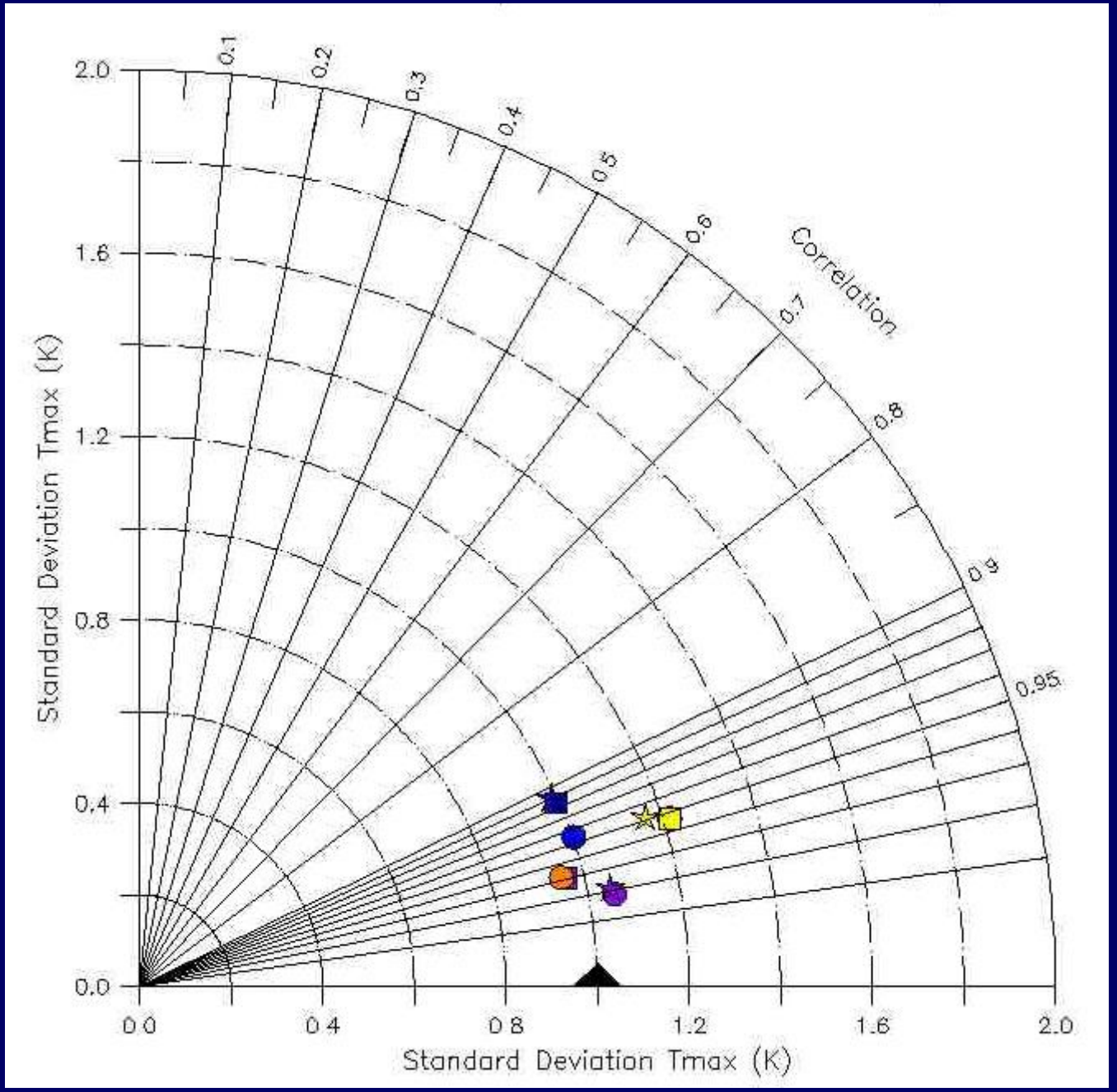
Different shapes = Different domain sizes



January Tmin

Different colors = Different models

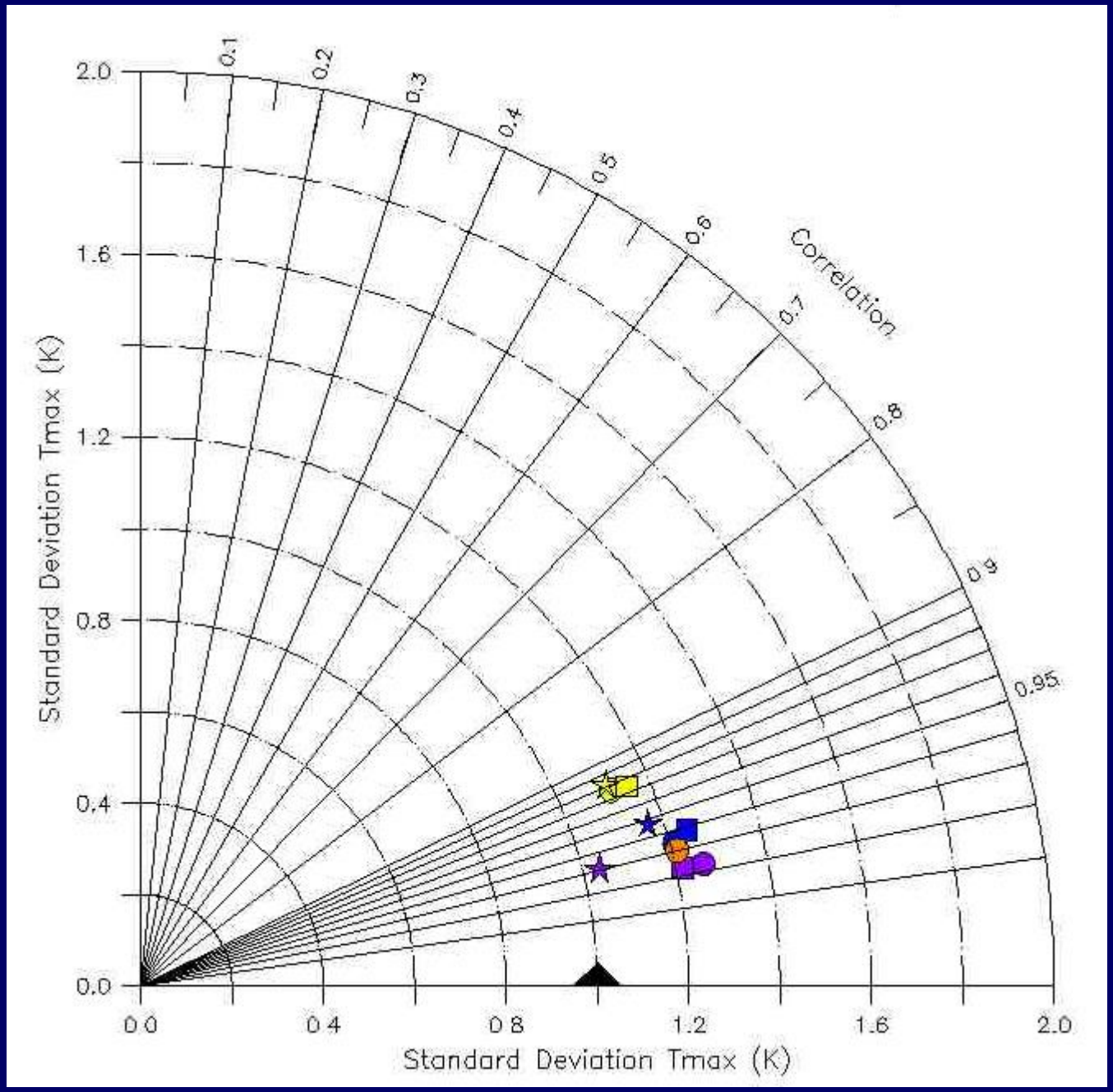
Different shapes = Different domain sizes



July Tmin

Different colors = Different models

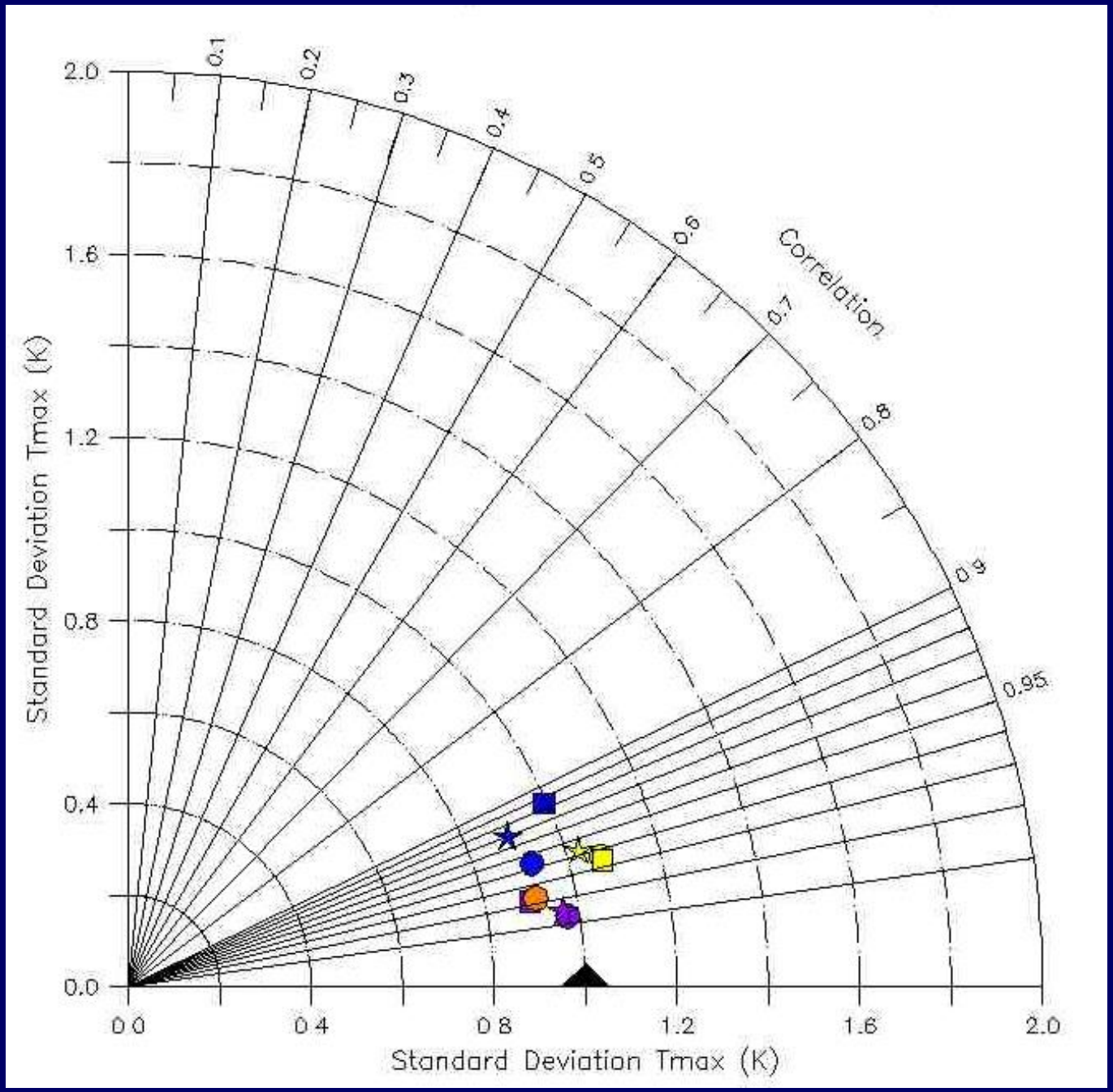
Different shapes = Different expts.



January 500 hPa Heights

Different colors = Different models

Different shapes = Different expts.



June 500 hPa Heights

Different colors = Different models

Different shapes = Different expts.