A comparison of statistical downscaling techniques for daily precipitation: Results from the CORDEX Flagship Pilot Study in South America

Bettolli ML, Gutiérrez JM, Iturbide M, Baño-Medina J, Huth R, Solman S, Fernández J, da Rocha RP, Llopart M, Lavín-Gullón A, Coppola E, Chou S, Doyle M, Olmo M, Feijoo M.

CORDEX-ICRC, Beijing 14-18 October 2019

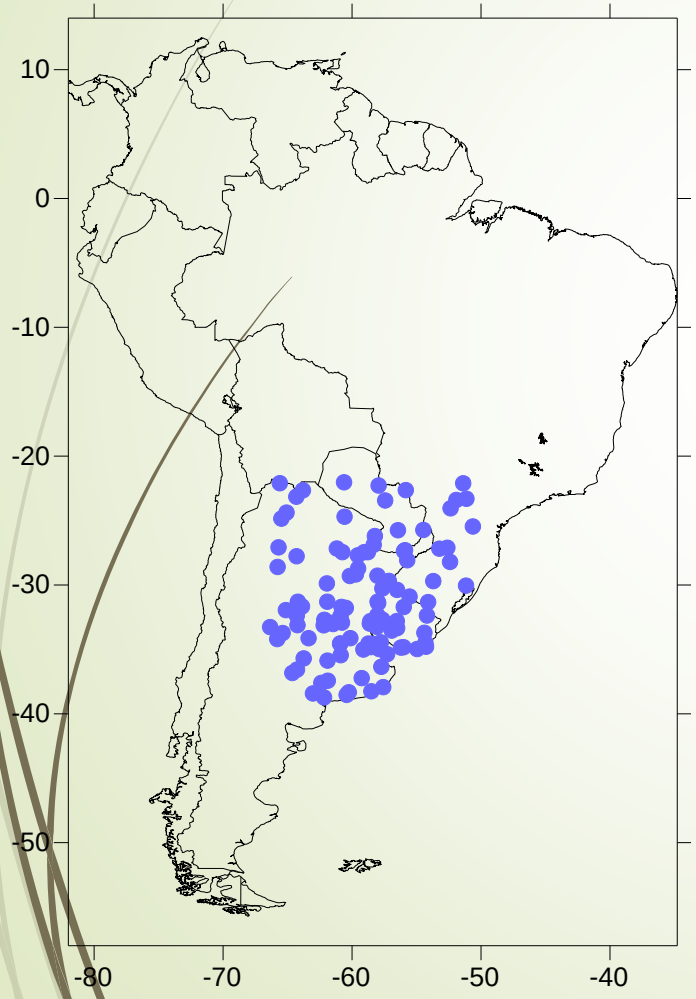
Objective

- to intercompare different statistical downscaling techniques in simulating daily precipitation in SESA with special focus on extremes.

Objective

- to intercompare different statistical downscaling techniques in simulating daily precipitation in SESA with special focus on extremes.
  - To evaluate the sensitivity to the reanalysis choice
  - To evaluate the sensitivity to predictor variables

Strategy and experiment design



ESD Simulations

- **Approach:**

  Perfect Prognosis

- **Predictors:**

  ERA-Interim reanalysis
  JRA reanalysis

- **Predictands**:
  Station Data (100): daily Pr, Tx and Tn
  MSWEP: daily Pr

- **Season:**

  October to March

- **Training and Test:**
  Cross validation k-folding strategy:
  6 folds containing 5 consecutive years in the period 1979-2009
  Independent Test period: 2009-2010

Strategy and experiment design

Generalized linear model (GLM)

Analog Method (AN)

Method	Configuration	Predictor Variables
GLM_pc	PCs (95% variance)	Z500, V850, Z1000, Q700, Q850, T700, T850
GLM_pc.C	PCs Circulation Variables (95% variance)	Z500, V850, Z1000
GLM_l4	Local predictor values in the four nearest grid boxes.	Z500, V850, Z1000, Q700, Q850, T700, T850
GLM_ls	Combination of local and spatial predictors (PCs 90%Variance)	Local: Q850 Spatial: V850, Z500,Z1000
AN_pc	Nearest neighbor, PCs (95% variance)	Z500, V850, Z1000, Q700, Q850, T700, T850
AN_pc_C	Nearest neighbor, PCs Circulation Variables (95% variance)	Z500, V850, Z1000
AN_l16	Nearest neighbor, Local predictor values in the four nearest grid boxes.	Z500, V850, Z1000, Q700, Q850, T700, T850

Strategy and experiment design

Generalized linear model (GLM)

Analog Method (AN)

Method	Configuration	Predictor Variables
GLM_pc	PCs (95% variance)	Z500, V850, Z1000, Q700, Q850, T700, T850
GLM_pc.C	PCs Circulation Variables (95% variance)	Z500, V850, Z1000
GLM_l4	Local predictor values in the four nearest grid boxes.	Z500, V850, Z1000, Q700, Q850, T700, T850
GLM_ls	Combination of local and spatial predictors (PCs 90%Variance)	Local: Q850 Spatial: V850, Z500,Z1000
AN_pc	Nearest neighbor, PCs (95% variance)	Z500, V850, Z1000, Q700, Q850, T700, T850
AN_pc_C	Nearest neighbor, PCs Circulation Variables (95% variance)	Z500, V850, Z1000
AN_l16	Nearest neighbor, Local predictor values in the four nearest grid boxes.	Z500, V850, Z1000, Q700, Q850, T700, T850

Strategy and experiment design

Generalized linear model (GLM)

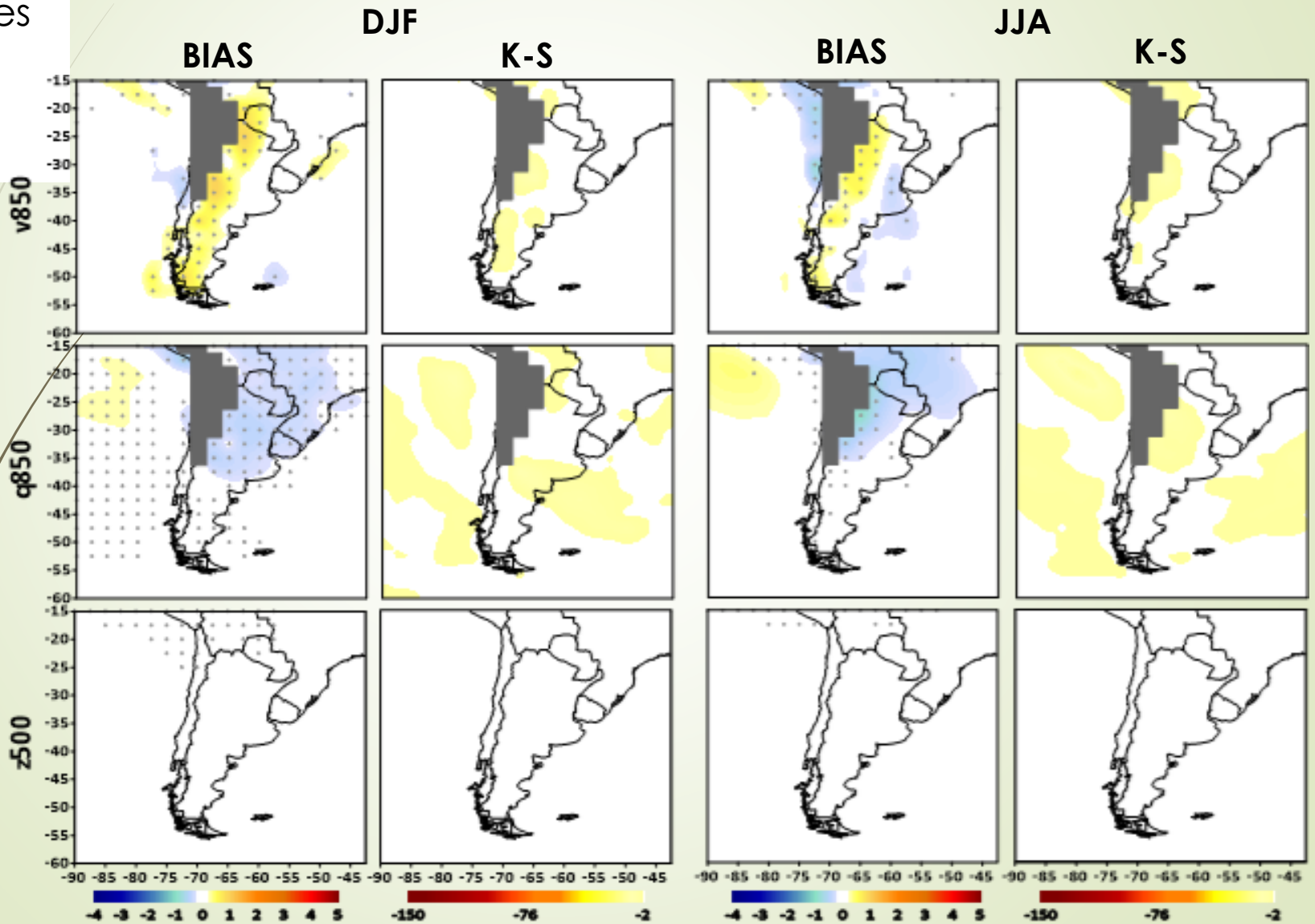
Analog Method (AN)

Method	Configuration	Predictor Variables
GLM_pc	PCs (95% variance)	Z500, V850, Z1000, Q700, Q850, T700, T850
GLM_pc.C	PCs Circulation Variables (95% variance)	Z500, V850, Z1000
GLM_l4	Local predictor values in the four nearest grid boxes.	Z500, V850, Z1000, Q700, Q850, T700, T850
GLM_ls	Combination of local and spatial predictors (PCs 90%Variance)	Local: Q850 Spatial: V850, Z500,Z1000
AN_pc	Nearest neighbor, PCs (95% variance)	Z500, V850, Z1000, Q700, Q850, T700, T850
AN_pc_C	Nea Vari	
AN_l16	Nea valu	

The simulations were performed in collaboration between the University of Buenos Aires and the University of Cantabria (**Climate4R**)

Results

Differences Between JRA and ERA-I



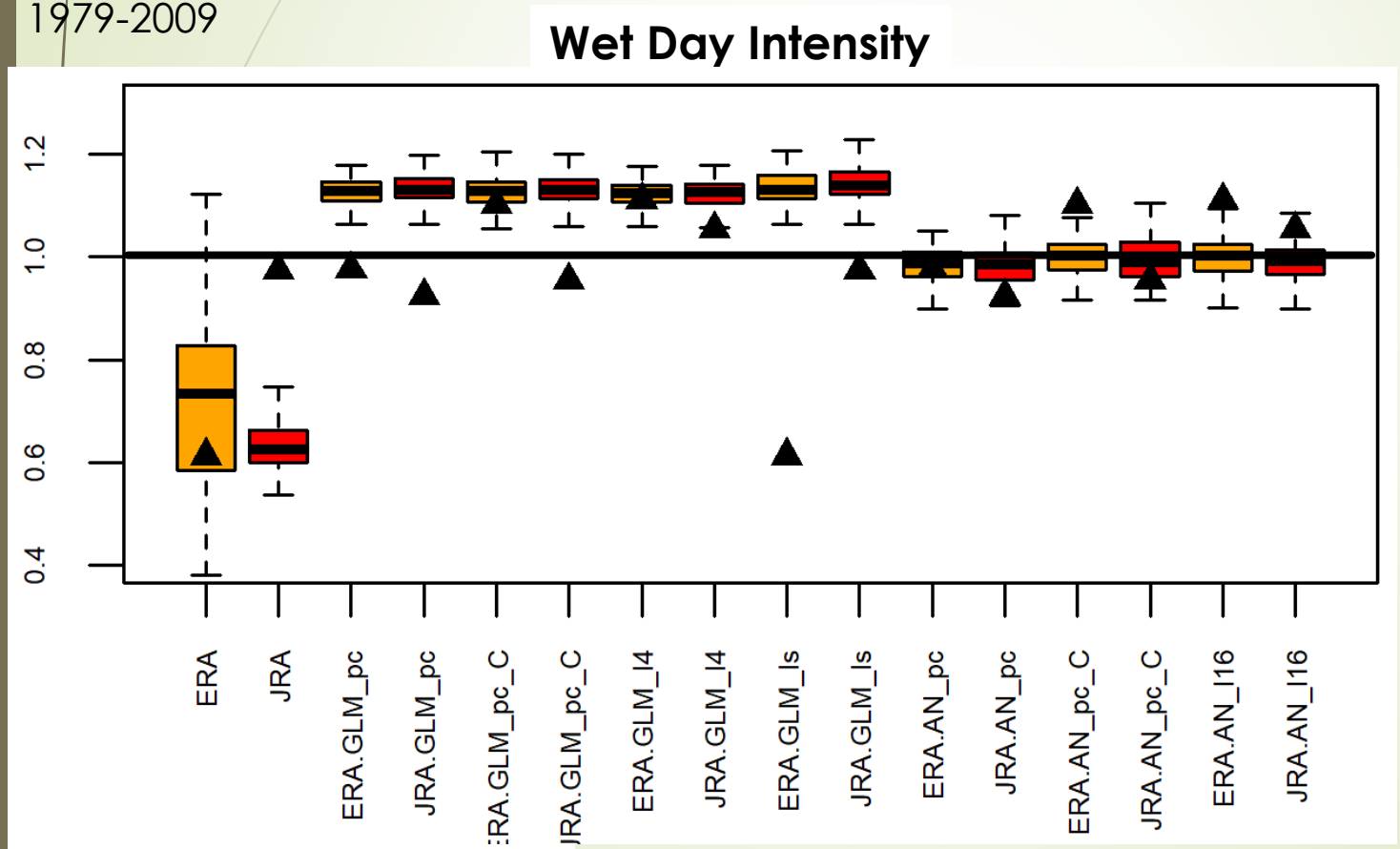
Results

ERA-I

JRA

1979-2009

Warm Season 2009/10



Results

ERA-I

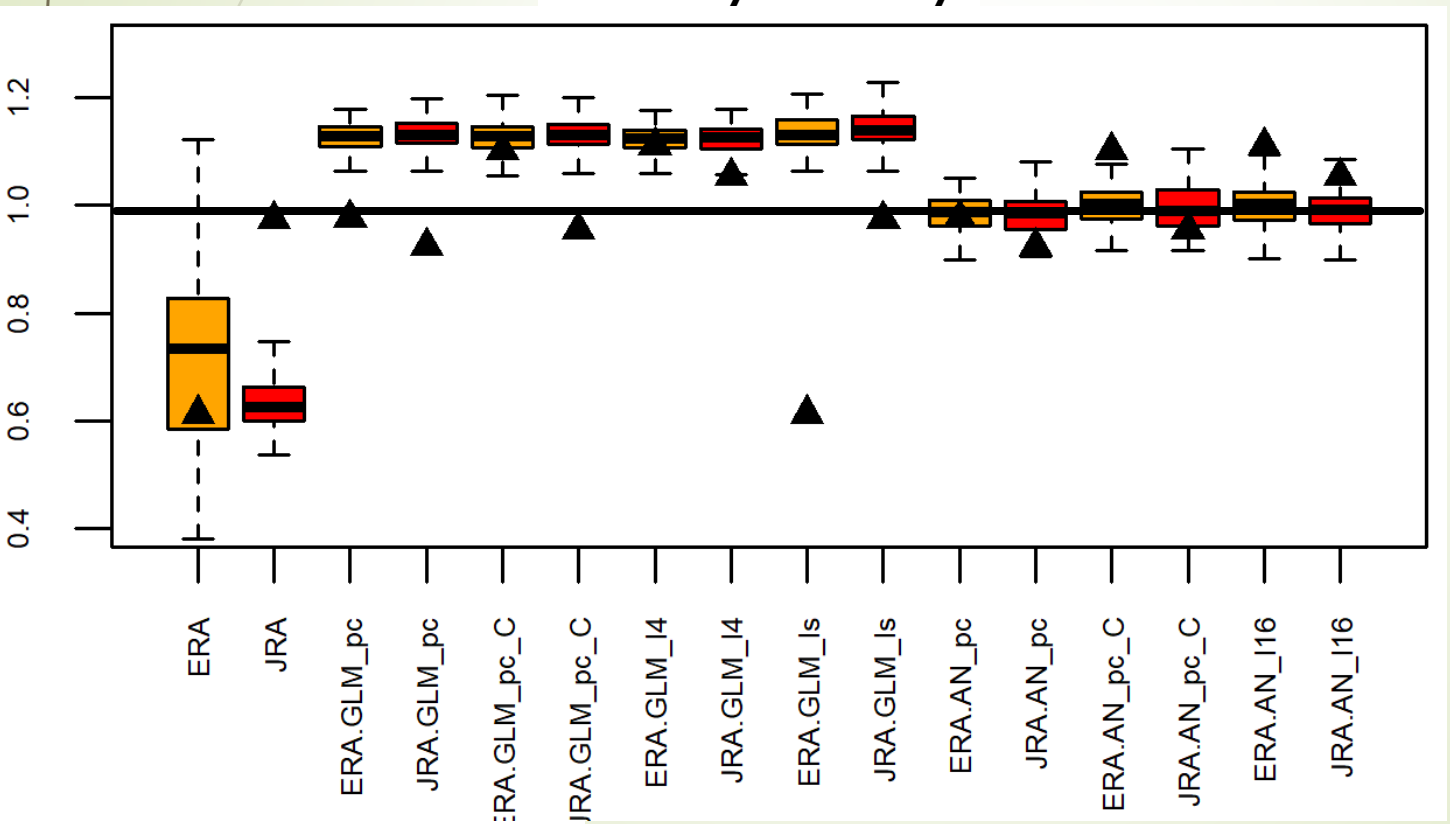
JRA

1979-2009

Warm Season 2009/10

Raw data: Underestimate
GLM: overestimate
AN: OK
2009/10: considerable spread

Wet Day Intensity



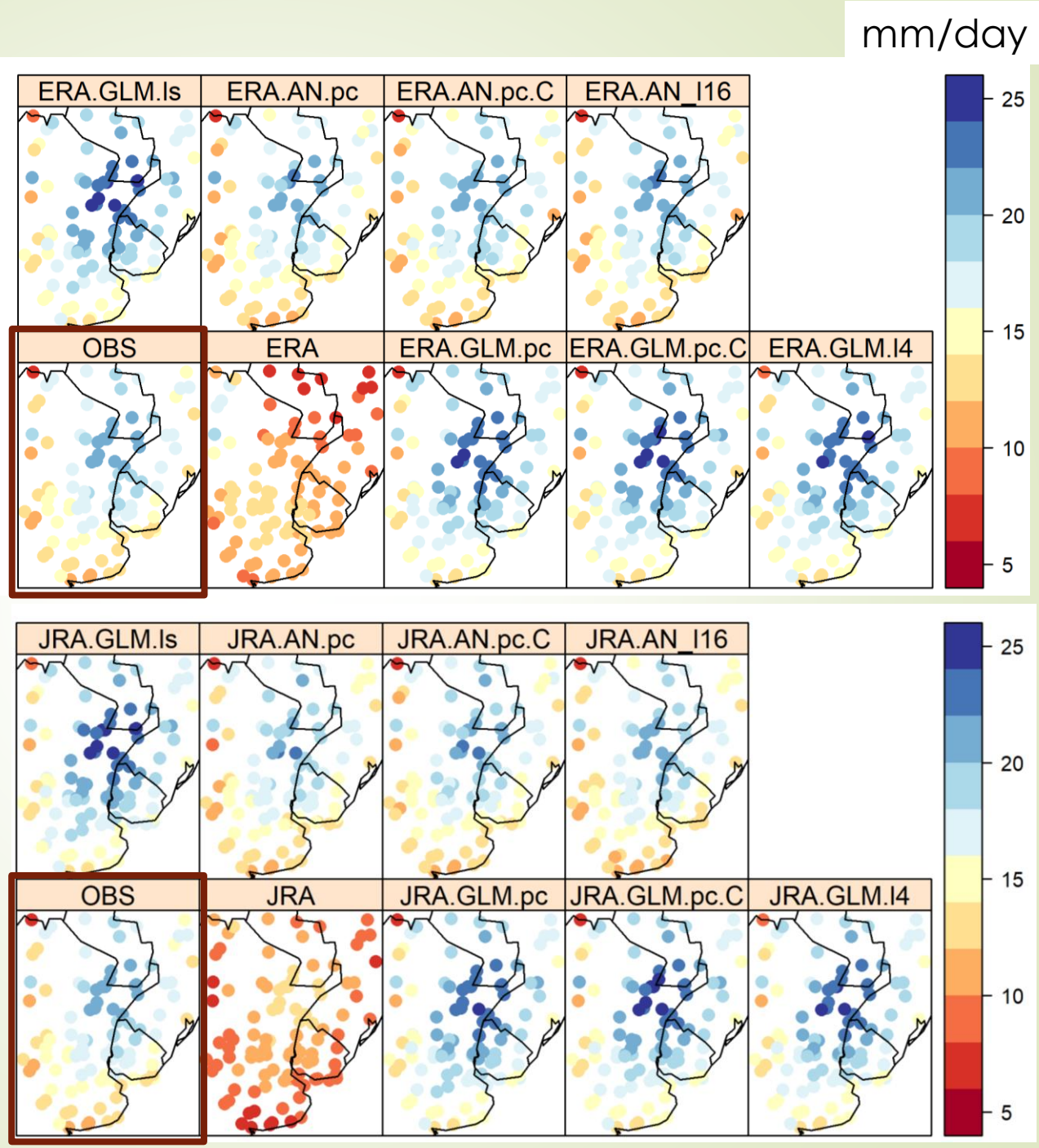
Ratio downscaled/OBS

Results

Wet Day Intensity

1979-2009

Even tough the **GLM** tended to overestimated the values, they **are able to reproduce the spatial** behavior of the wet day intensity.



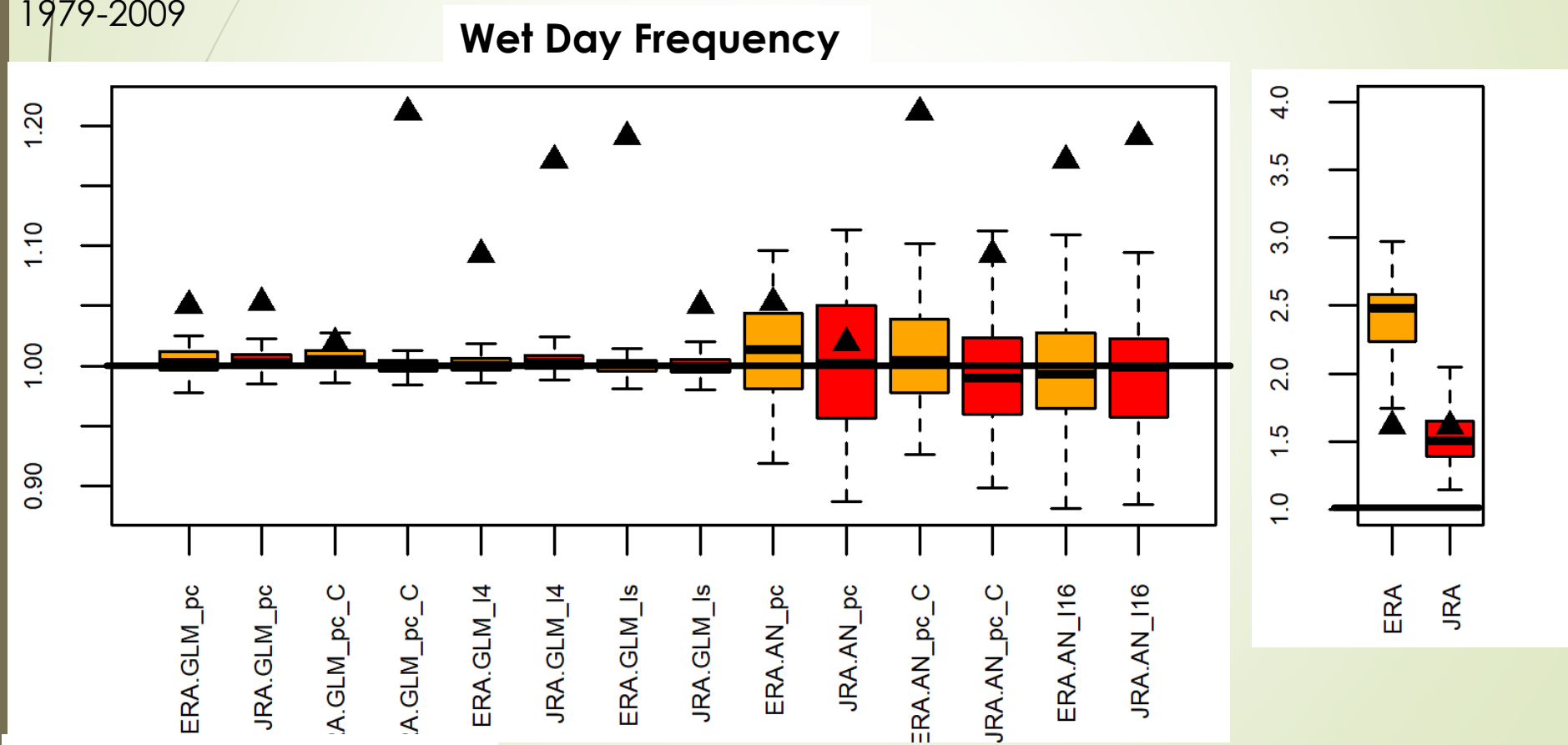
Results

ERA-I

JRA

1979-2009

Warm Season 2009/10



Results

ERA-I

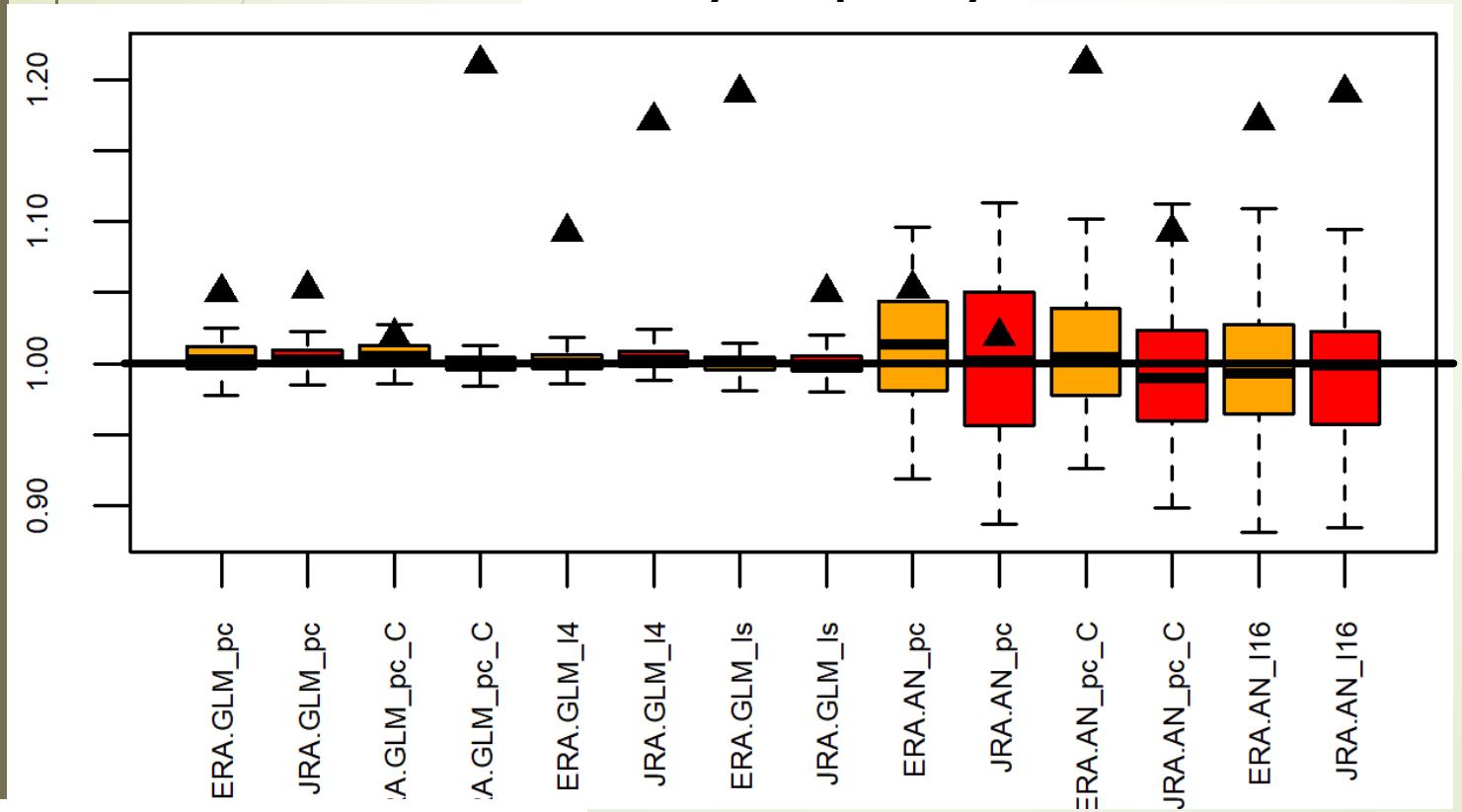
JRA

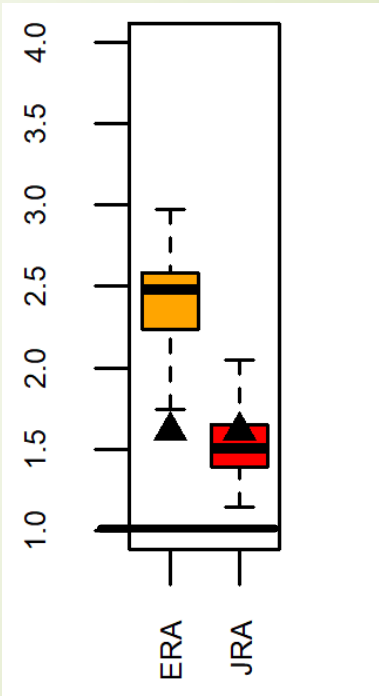
1979-2009

Warm Season 2009/10

Raw data: Overestimation
GLM: OK
AN: Spatial spread in performances
2009/10: considerable spread

Wet Day Frequency





Ratio downscaled/OBS

Results

ERA-I

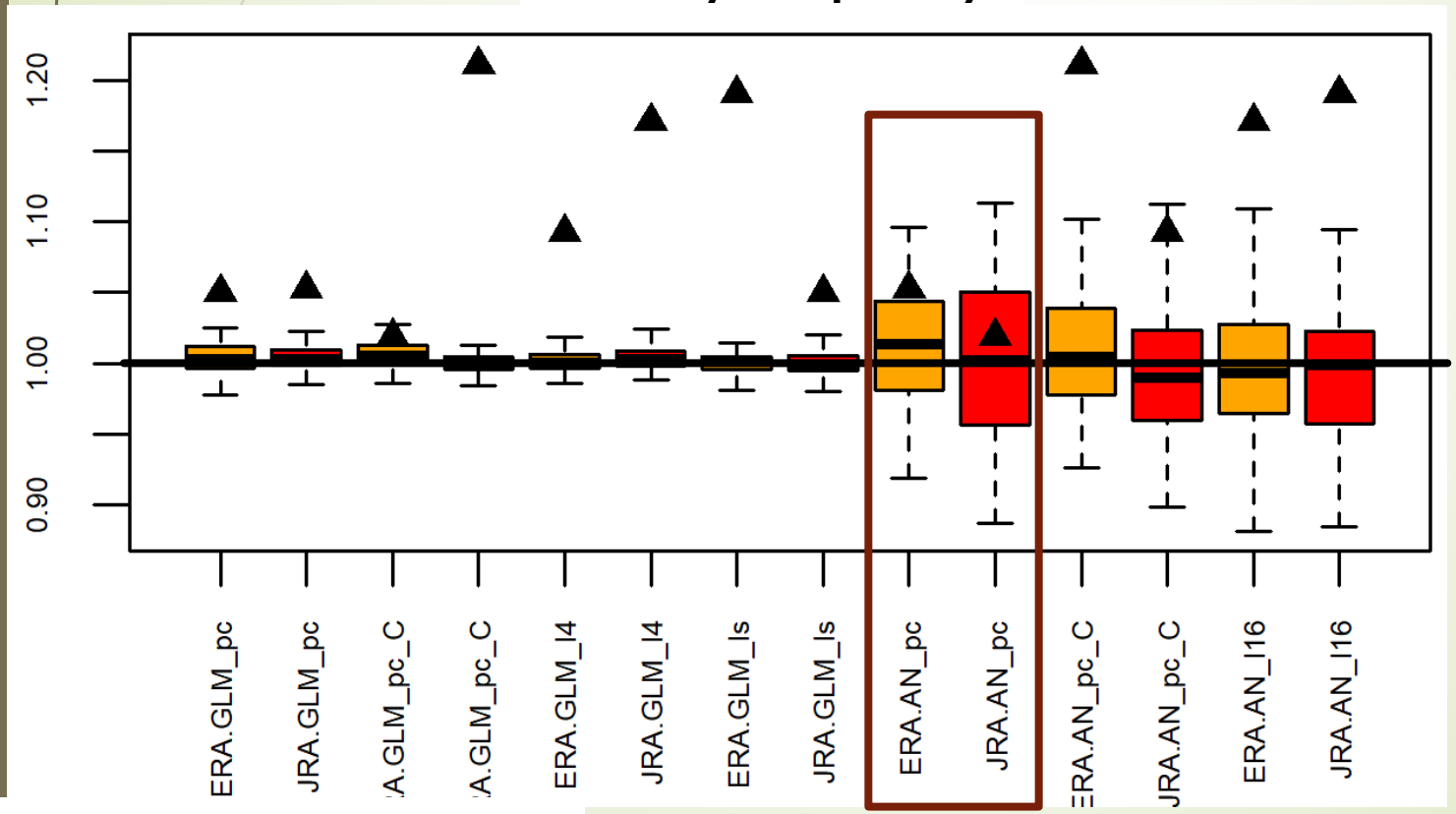
JRA

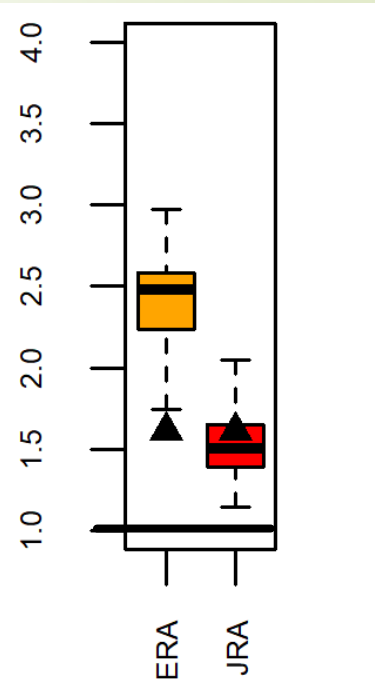
1979-2009

Warm Season 2009/10

Raw data: Overestimation
GLM: OK
AN: Spatial spread in performances
2009/10: considerable spread
Except for the **AN** that considers the full set of predictor variables

Wet Day Frequency





Ratio downscaled/OBS

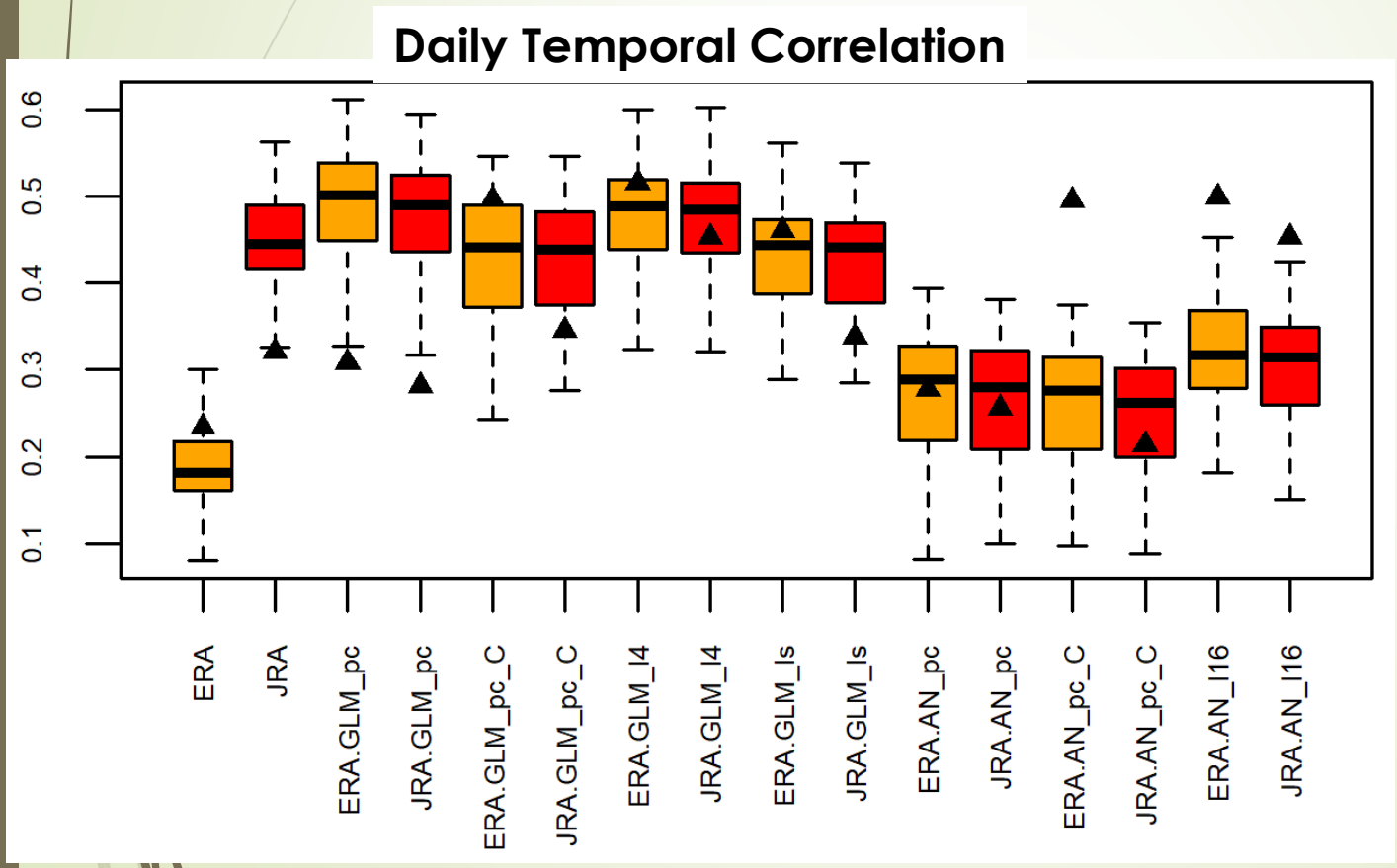
GLM: performs best

Results

ERA-I

JRA

Warm Season 2009/10



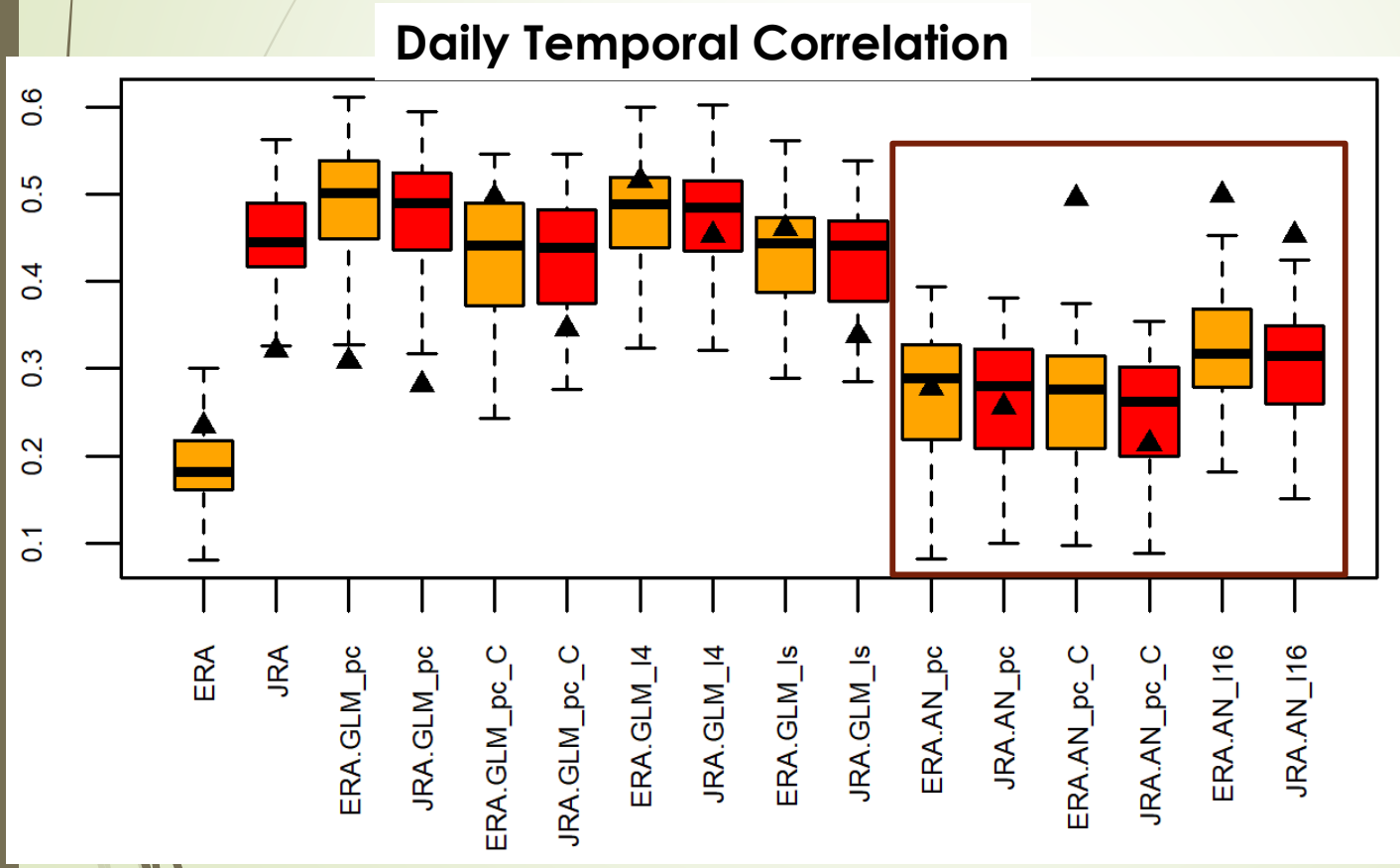
Results

ERA-I

JRA

Warm Season 2009/10

GLM: performs best
2009/10: some differences depending on the reanalysis choice and the predictor set are evident .



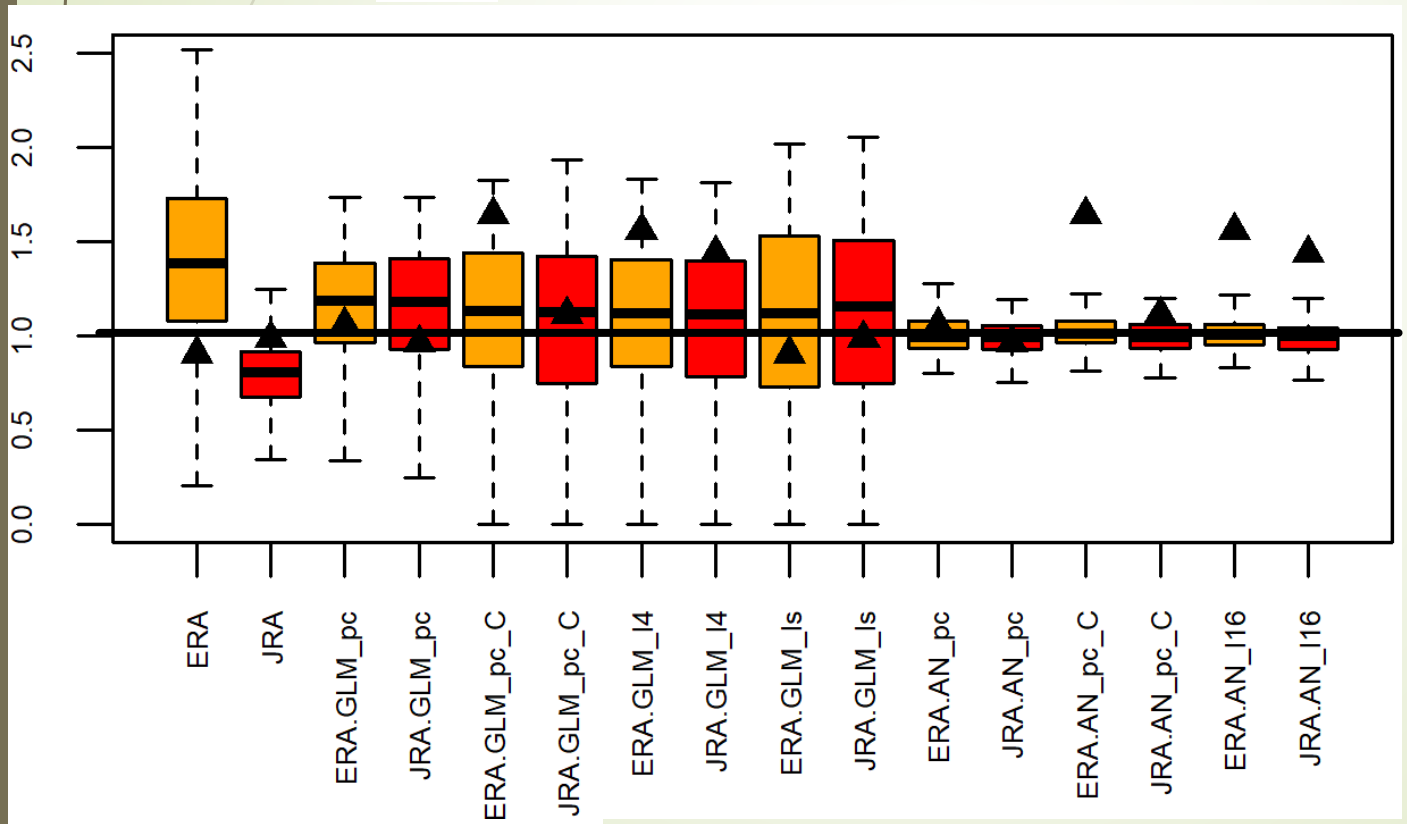
Results

All methods show similar performances but
GLM: present more spread

ERA-I
JRA
1979-2009

Warm Season 2009/10

R20



Ratio downscaled/OBS

Results

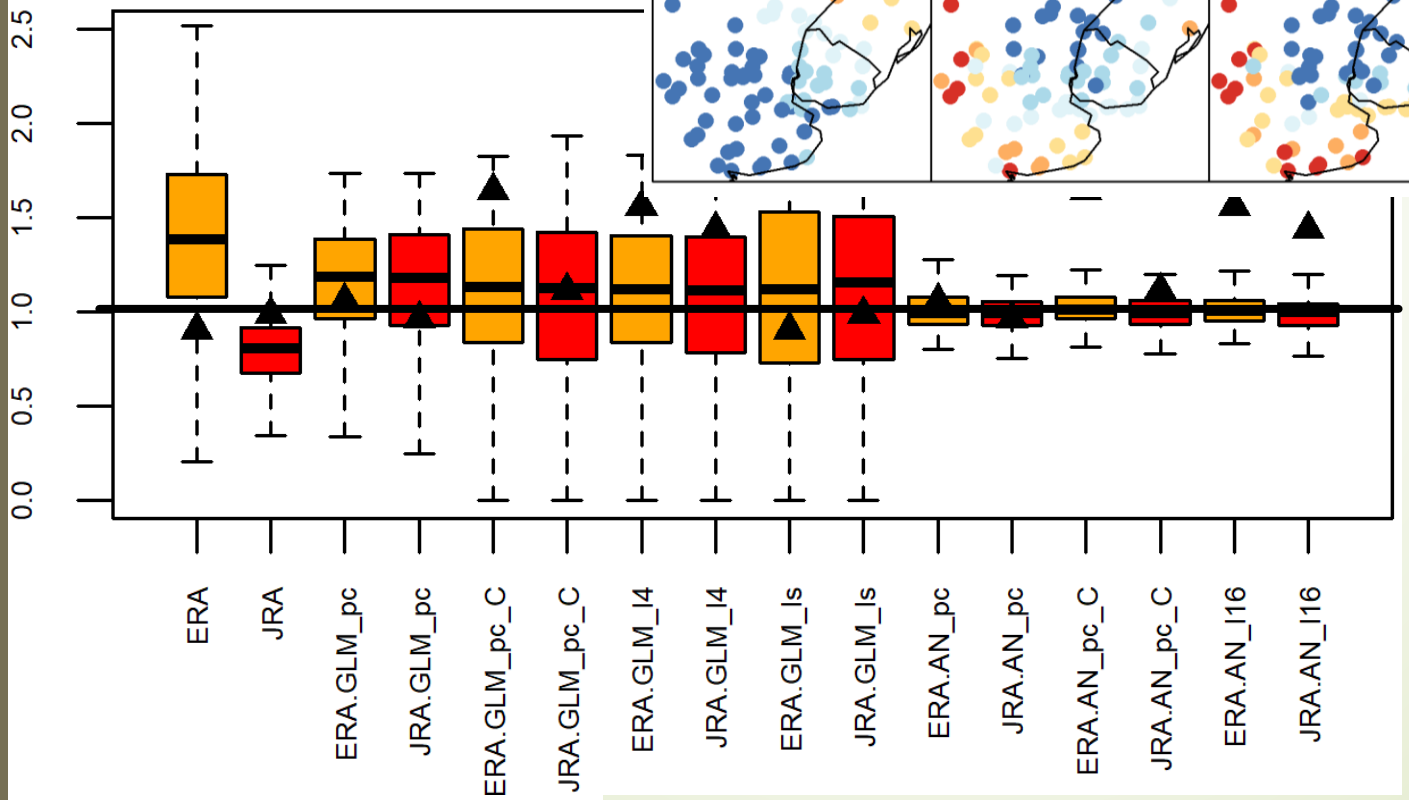
ERA-I

JRA

1979-2009

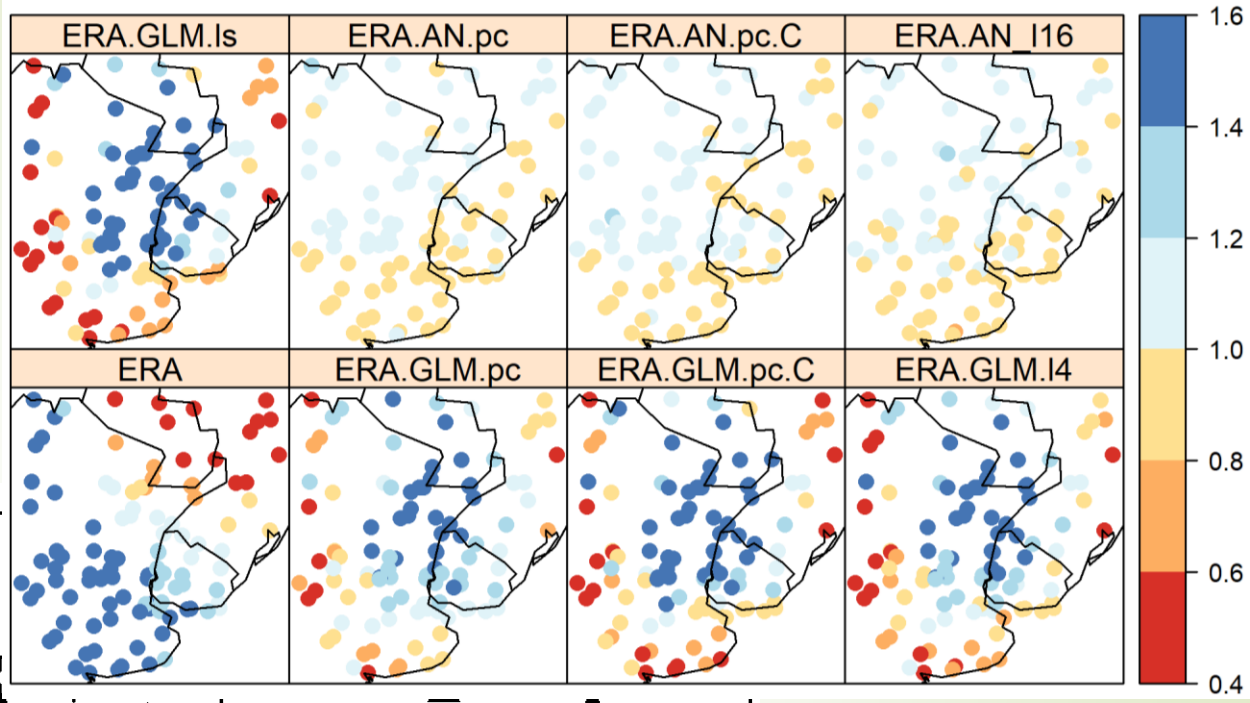
Warm Season 2009/10

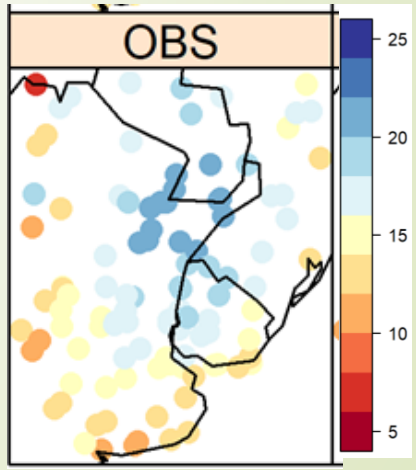
R20



Ratio downscaled/OBS

R20 (Ratio)





Wet Day Intensity

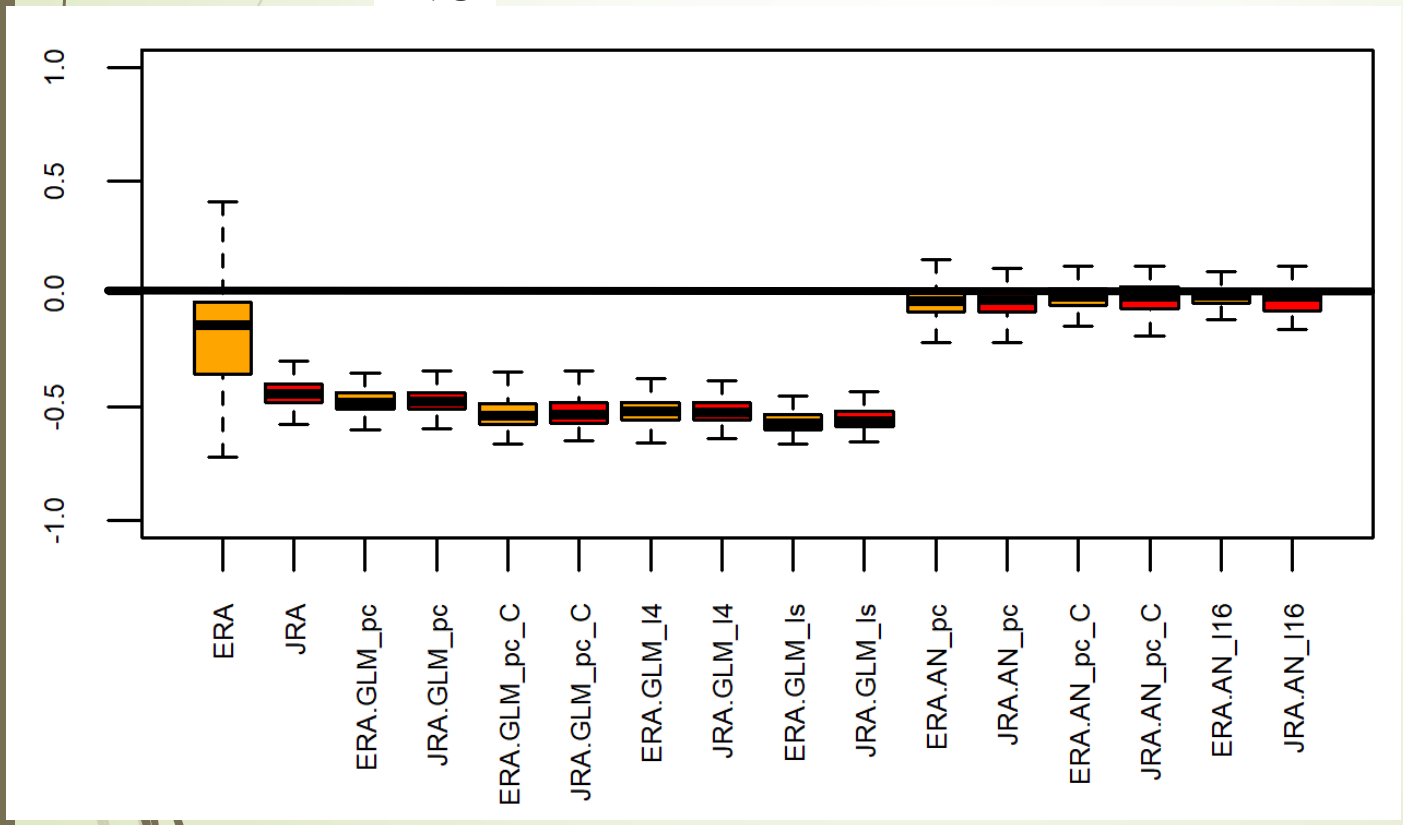
Results

Raw data and GLM: underestimate the P98
AN: perform best

ERA-I
JRA
1979-2009

Warm Season 2009/10

P98



Relative Bias

Concluding remarks

- The results show that the methods are generally more skillful when combined predictors including temperature and humidity at low levels of the atmosphere are considered.
- The performance of the models is also sensitive to reanalysis choice.
- The methods show overall good performance in simulating daily precipitation characteristics over the region, but no single model performs best over all validation metrics and aspects evaluated.

Thanks!