How reliable are decadal climate predictions?

D. Verfaillie[1,*]**, F. J. Doblas-Reyes**[1,2]**, M. Donat**[1]**, N. Pérez-Zanón**[1]**,**
B. Solaraju-Murali[1]**, V. Torralba**[1] **and S. Wild**[1]
[1] Earth Sciences Dept, Barcelona Supercomputing Center (BSC), Barcelona, Spain
[2] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain
* Now at Georges Lemaître Centre for Earth and Climate Research, Earth and Life Institute, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

EGU 2020, Session CL3.3
5 May 2020

European Climate Prediction system

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

Initialised decadal predictions (INIT) vs. other sources of future information available to stakeholders
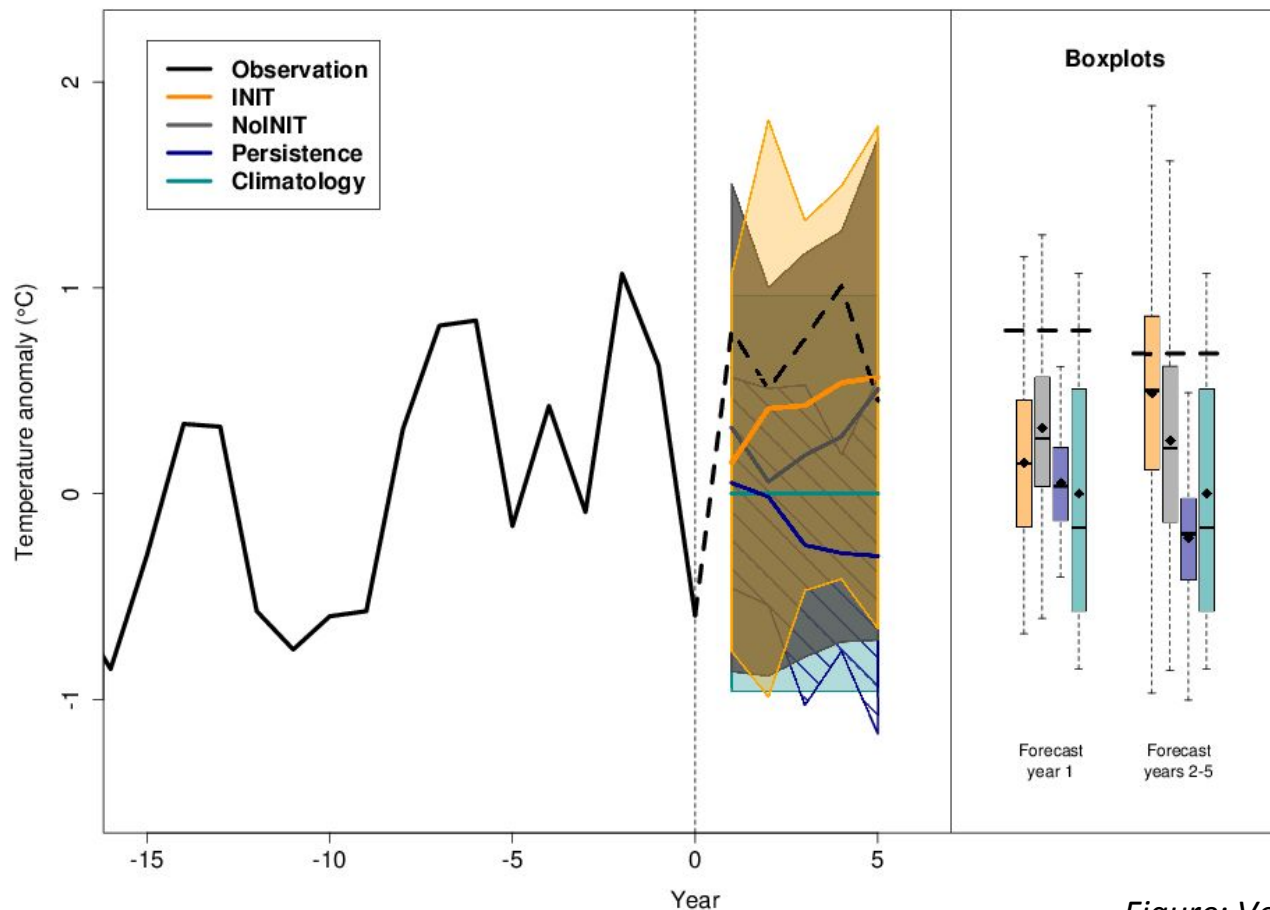→ Different **time series**

Figure: Verfaillie et al.

Initialised decadal predictions (INIT) vs. other sources of future information available to stakeholders
→ Different **probability distributions**
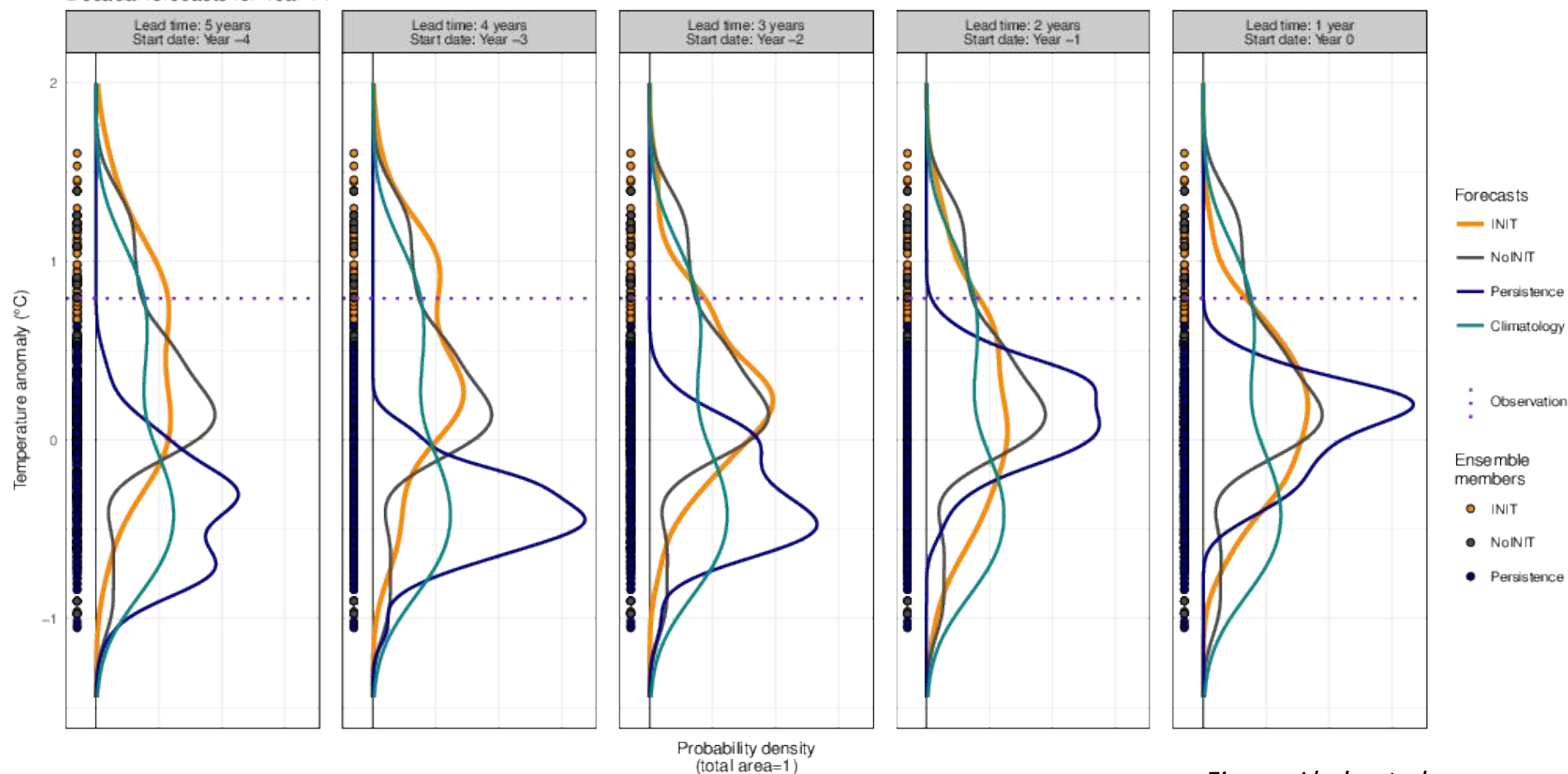

Decadal forecasts for Year +1
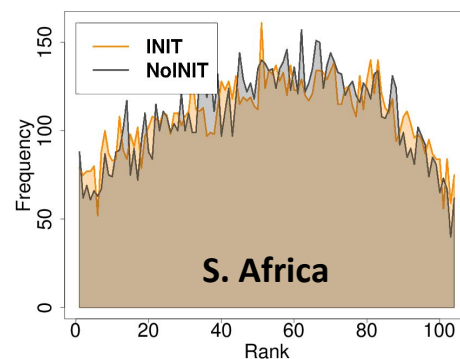
Figure: Lledo et al.
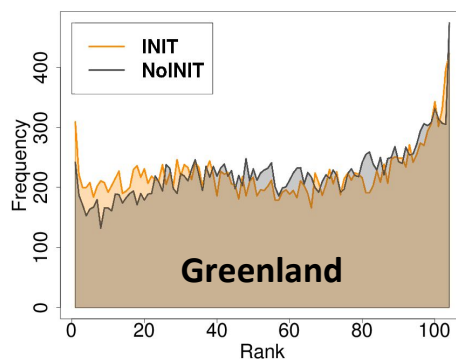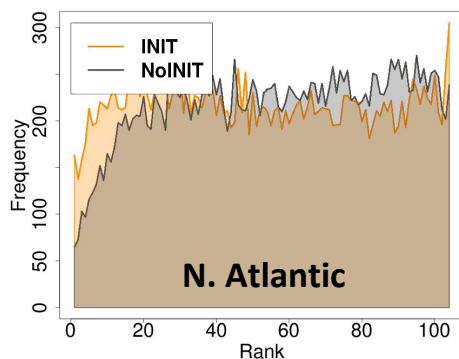
Comparing INIT and NoINIT

Generally done in terms of forecast accuracy (skill scores)

Here: impact of initialisation in terms of **reliability**
= *agreement between the predicted probabilities and observed relative frequencies of a given event*

Different tools:
- **rank histograms** → illustrate if ensemble members and verifying observation come from the same probability distribution, in which case the rank histograms are flat

Verfaillie et al., under rev.
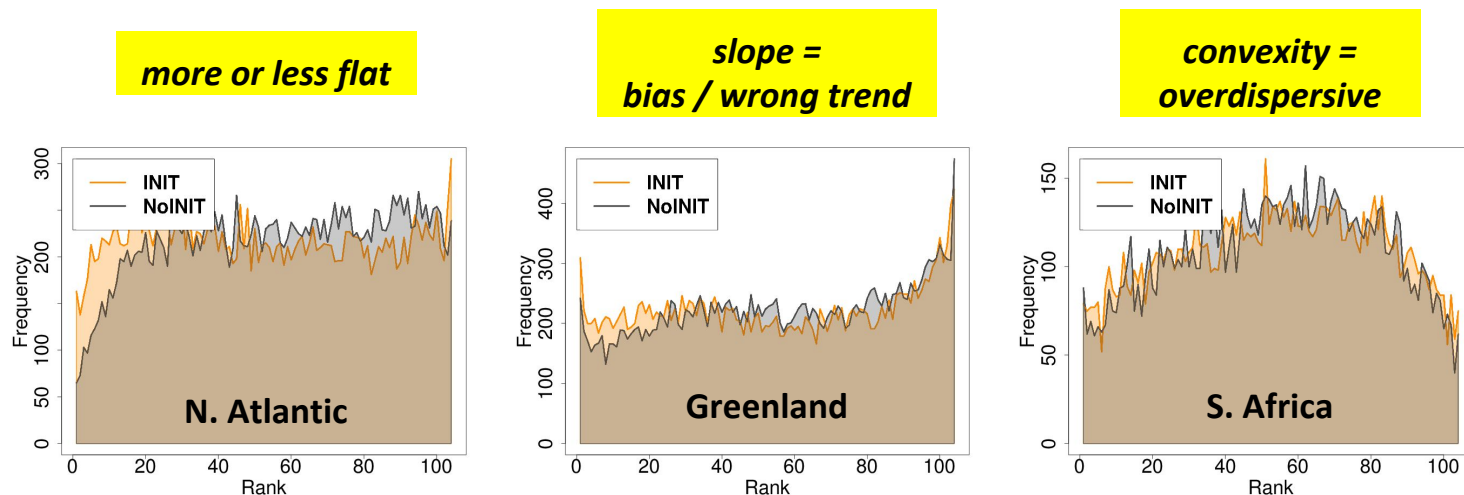
Comparing INIT and NoINIT

Generally done in terms of forecast accuracy (skill scores)

Here: impact of initialisation in terms of **reliability**
= *agreement between the predicted probabilities and observed relative frequencies of a given event*

Different tools:
- **rank histograms**
- **Joliffe and Primo (2008) test statistics: Pearson X^2**
  + decomposition into "**slope**" & "**convexity**" (& residual)

more or less flat

*slope =
bias / wrong trend*

*convexity =
overdispersive*

N. Atlantic

Greenland

S. Africa

5

Verfaillie et al., under rev.

Comparing INIT and NoINIT

Near-surface air Tº, **multi-model** set of 12 different models (CMIP5 & SPECS, both INIT & NoINIT), **30 different regions**
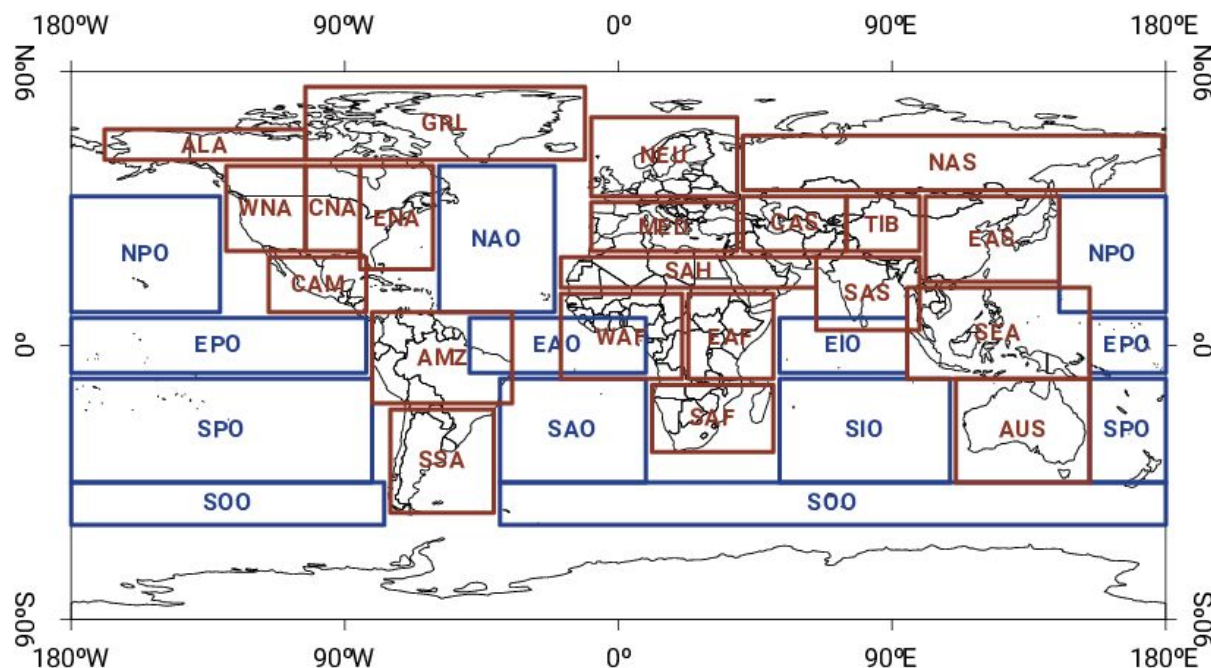

SOO: Southern Ocean
SAO: South Atlantic Ocean
EAO: Equatorial Atlantic Ocean
NAO: North Atlantic Ocean
SIO: South Indian Ocean
EIO: Equatorial Indian Ocean
SPO: South Pacific Ocean
EPO: Equatorial Pacific Ocean
NPO: North Pacific Ocean
AUS: Australia

AMZ: Amazon Basin
SSA: Southern South America
CAM: Central America
WNA: Western North America
CNA: Central North America
ENA: Eastern North America
ALA: Alaska
GRL: Greenland
MED: Mediterranean Basin
NEU: Northern Europe

WAF: Western Africa
EAF: Eastern Africa
SAF: Southern Africa
SAH: Sahara
SEA: Southeast Asia
EAS: East Asia
SAS: South Asia
CAS: Central Asia
TIB: Tibet
NAS: North Asia

Verfaillie et al., under rev.
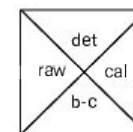
MM, forecast year 1, uncorrected data
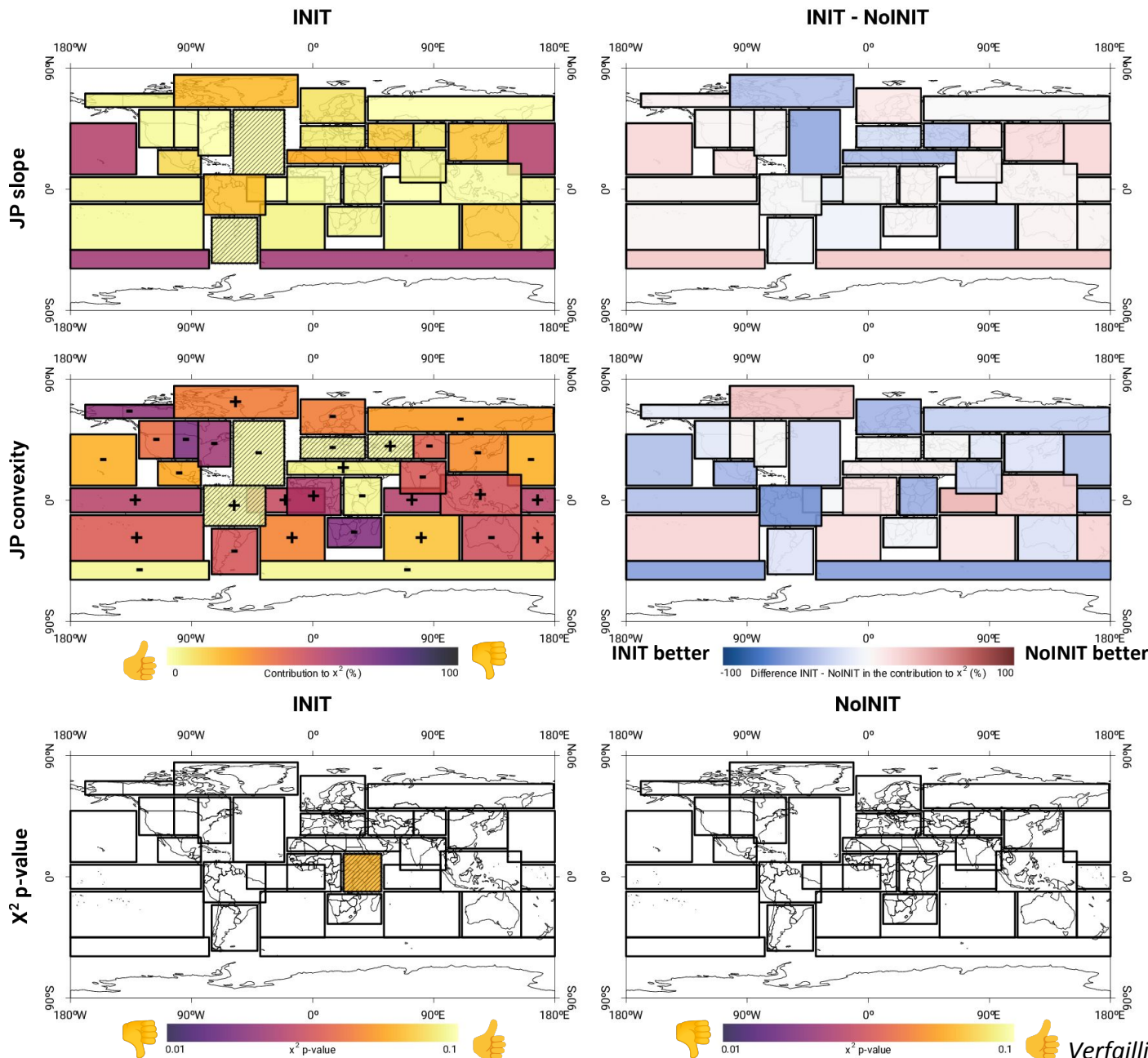
Comparing INIT and NoINIT

- Both uncorrected INIT & NoINIT generally **unreliable**
- **Biased**, incorrect **trend** and/or **dispersion** errors
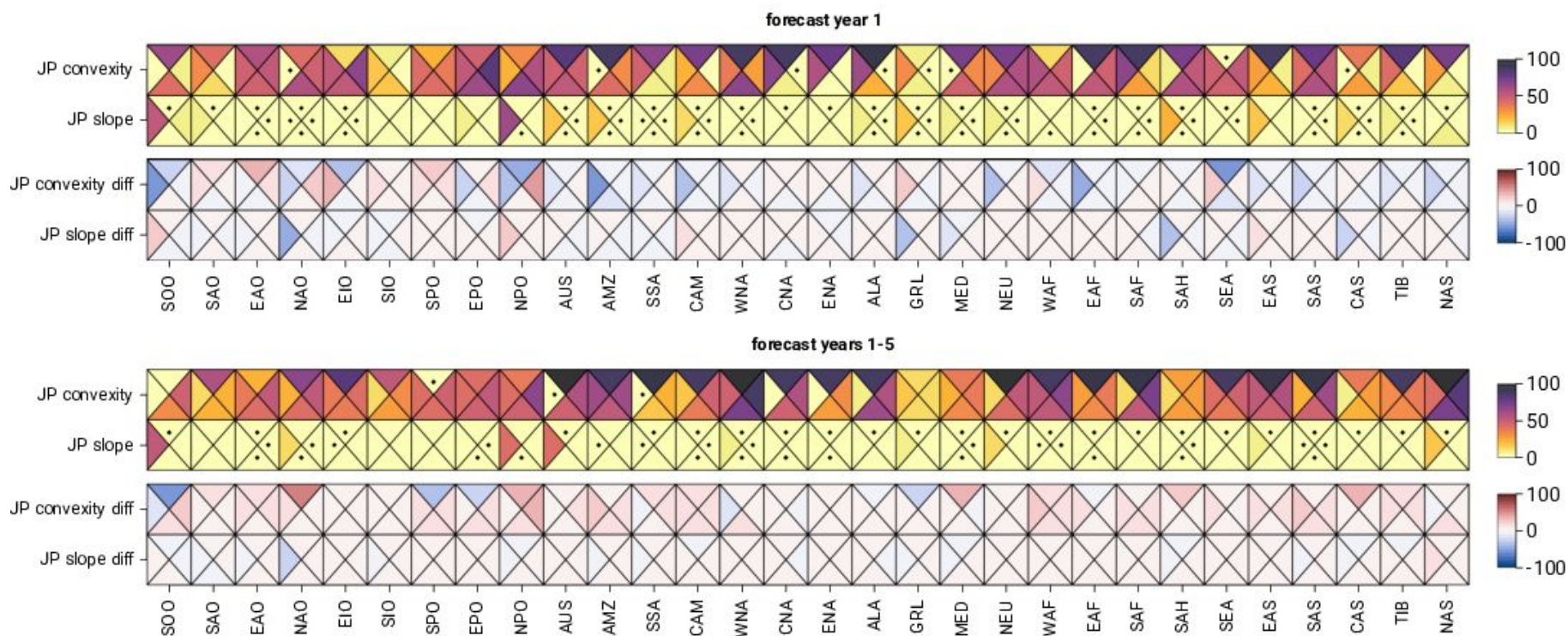- Some regions (e.g., NAO) without bias or dispersion error, yet unreliable (**residual** parameter)

MM
Impact of forecast time
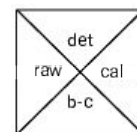# Impact of different post-processing methods

raw = uncorrected
det = detrended
b-c = bias-corrected
cal = calibrated

- no slope/convexity error (p ≥ 0.05)
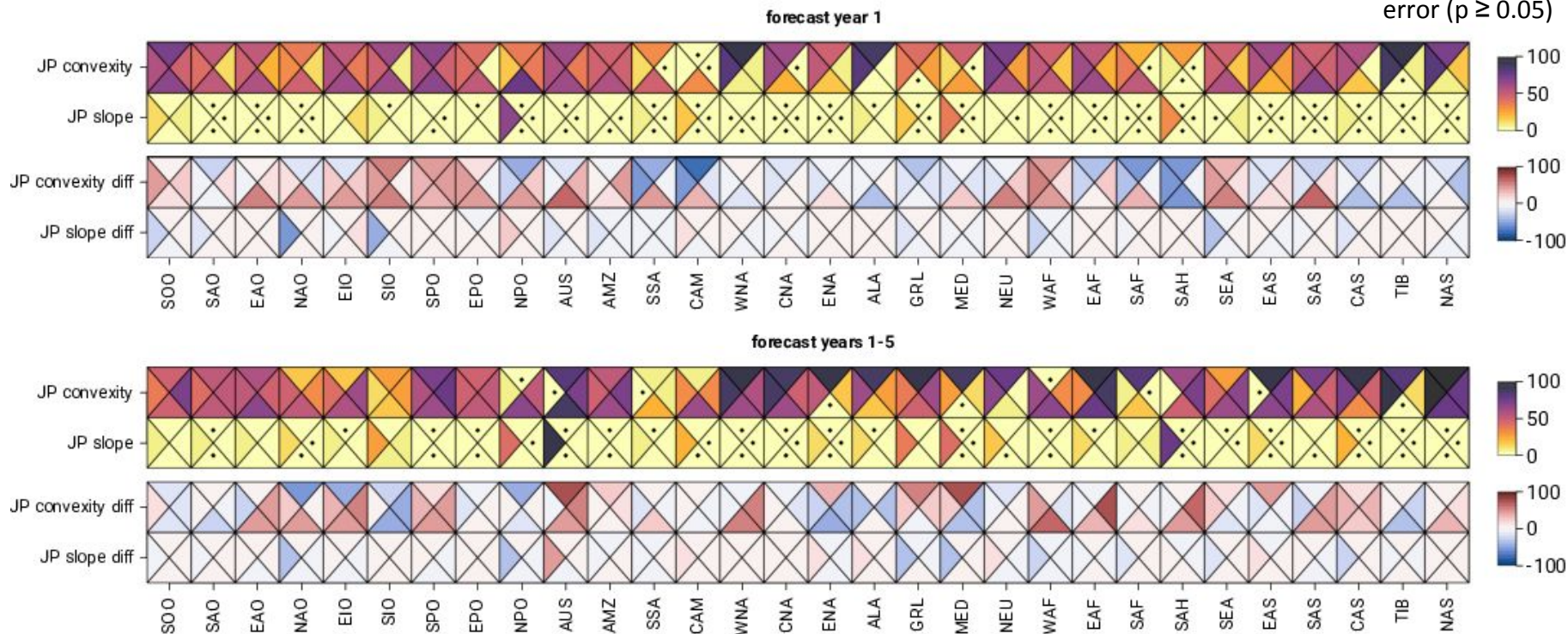

forecast year 1

forecast years 1-5

- Effect of **initialisation** on reliability small, mostly limited to forecast year 1
- **Bias correction** & **calibration** necessary for reliable forecasts

9

Verfaillie et al., under rev.

Impact of different model ensembles: NCAR

raw = uncorrected
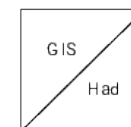det = detrended
b-c = bias-corrected
cal = calibrated

• no slope/convexity error ($p \geq 0.05$)

- **Combination of models** more important than **ensemble size** of each individual forecast system
- Due to large range of **model physics & initialisation approaches** → error compensation
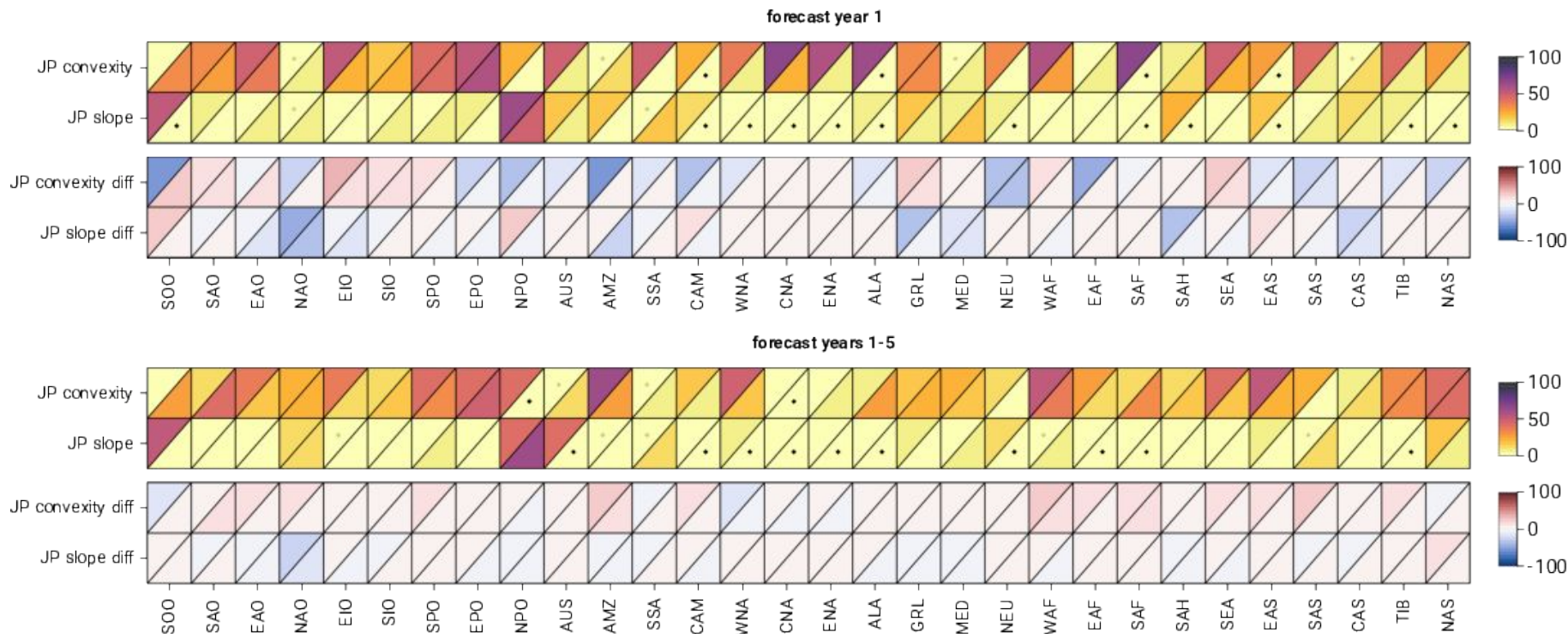
10

Verfaillie et al., under rev.

MM, uncorrected data
Impact of the reference obs dataset

GIS = GISTEMP
Had = HadCRUT4

• no slope/convexity error (p ≥ 0.05)

- Impact of different reference datasets
- Measure of **observational uncertainty**

Verfaillie et al., under rev.

Main conclusions

- Both uncorrected INIT & NoINIT generally **unreliable**
- **Biased**, incorrect **trend** and/or **dispersion** errors
- Some regions (e.g., NAO) without bias or dispersion error, yet unreliable (**residual** parameter)
- Effect of **initialisation** on reliability small, mostly limited to forecast year 1
- **Bias correction** & **calibration** necessary for reliable forecasts
- **Combination of models** more important than **ensemble size** of each individual forecast system
- Due to large range of **model physics & initialisation approaches** → error compensation
- Impact of different reference datasets
- Measure of **observational uncertainty**

Thanks!

deborah.verfaillie@bsc.es

EUCP
European Climate Prediction system

BSC
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA