Environment and Climate Change Canada | Environnement et Changement climatique Canada

JWGFVR

Joint Working Group on Forecast Verification Research:

Updates and future plans, including process based verification

Barbara Casati (ECCC), Caio Coelho *(CPTEC/INPE), Thomas Haiden (ECMWF)*

WGNE meeting, 3rd Nov 2021

Outline:

Slides 1-3: the JWGFVR past 20 years
Slides 4-6: some 2020-2021 progresses
The rest of the slides: process-based verification
Last slide: let's discuss the future plans!

Canada

Joint Working Group in Forecast Verification Research

https://community.wmo.int/activity-areas/wwrp/wwrp-working-groups/wwrp-forecast-verification-research

WORLD METEOROLOGICAL ORGANIZATION
Weather · Climate · Water

WWRP WORLD WEATHER RESEARCH PROGRAMME

JWGFVR

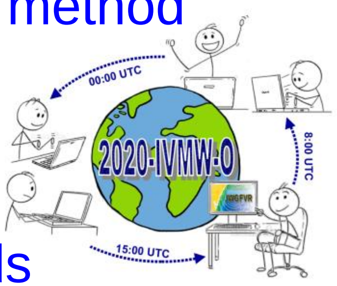
Mission: JWGFVR plans and facilitates the development and application of improved diagnostic verification methods to assess and enable improvement of the quality of weather forecasts, including forecasts from numerical weather and climate models. It also collaborates on forecast verification with WGNE and WCRP, and engages in the plans and implementation of the verification component of WWRP projects from their outset.

Promote good verification practices :

- Verification tutorials
- Verification web-page
- WMO recommendation reports and standards for operational centers
- Verification software

Advance verification research:

- Spatial verification method intercomparisons
- International workshops on verification methods
- Verification challenges
- Special issues & publications



Support verification activities in the other **WWRP/WMO projects and Working Groups**

HIWeather

Subseasonal-to-Seasonal S2S Prediction Project

YOPP YEAR OF POLAR PREDICTION

Verification Workshops and Tutorials

- 30 July – 1 Aug **2002, Boulder**: “Making Verification More Meaningful” (Barb Brown).
- 15-17 Sept **2004, Montreal**: 2nd International Verification Workshop (Laurie Wilson)
- 31 Jan – 2 Feb **2007, Reading**: 3rd International Verification Workshop & Tutorials (Anna Ghelli) https://www.ecmwf.int/en/learning/workshops-and-seminars/past-workshops/2007-international-verification-methods
  Ebert and Ghelli (2008) ed. Met Apps Special Issue; Casati et al (2008) review article
- 4 -10 June **2009, Helsinki**: 4th International Verification Workshop & Tutorials (Pertti Nurmi) https://space.fmi.fi/Verification2009/
- 1-7 Dec **2011, Melbourne**: 5th International Verification Workshop & Tutorials (Beth Ebert)
  Ebert et al (2013) review article
- 13-19 March **2014, New Delhi**: 6th International Verification Workshop & Tutorials (Raghu Ashrit)
- 3-11 May **2017, Berlin**: 7th International Verification Workshop & Tutorials (Martin Goeber)
  https://www.7thverificationworkshop.de
  Dorninger et al (2018) ed. Met Zet special issue; Dorninger et al (2020) ed. Met Apps special issue.
- 9-20 November **2020, online,** 8th International Verification Method Workshop (Barbara Casati & Manfred Dorninger) https://jwgfvr.univie.ac.at Casati et al (2021) BAMS workshop summary
- 21-25 June **2021, online:** MPE-CDT + JWGFVR verification summer school
  https://mpecdt.ac.uk/mpe-cdt-jwgfvr-forecast-verification-summer-school

Spatial verification methods

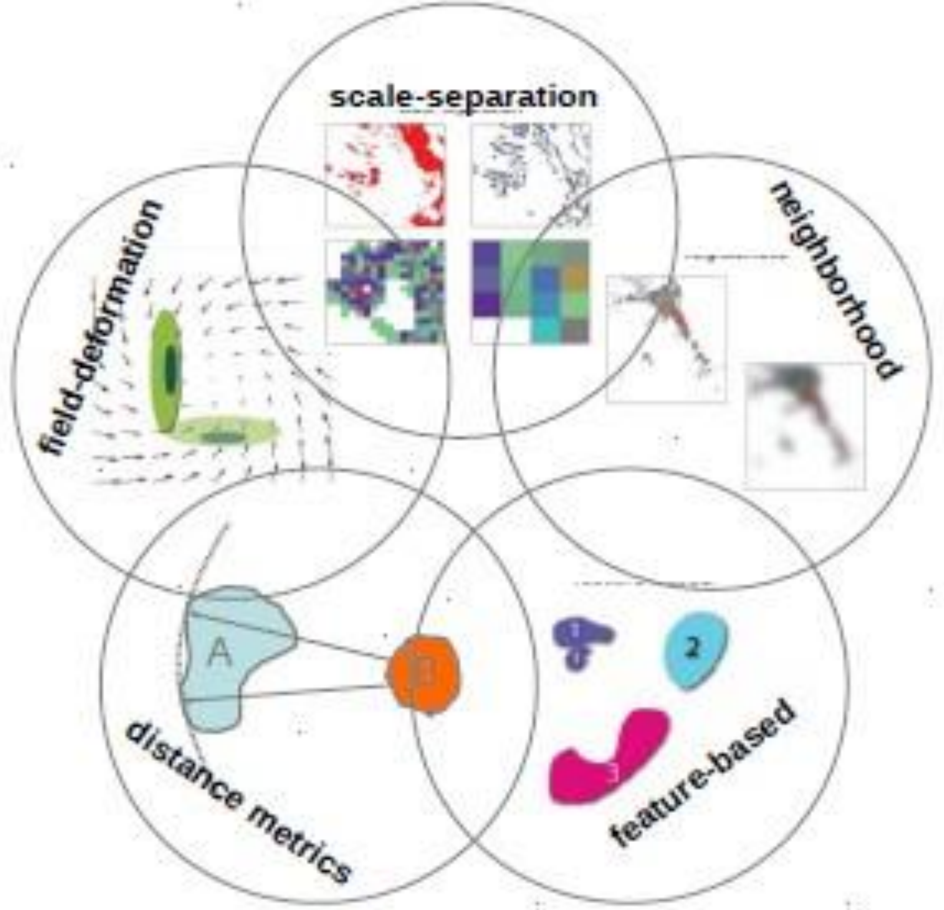
- Account for **coherent spatial structure** and the presence of **features**
- Aim to provide information on **error in physical terms (meaningful verification): e.g.** assess **scale structure** and **displacement error** (separately from **intensity error**)
- Account for **small time-space uncertainties** (avoid **double-penalty** issue)

Spatial method inter-comparisons:

- Spatial Verification Inter-Comparison Project (ICP): Gilleland et al (2010), BAMS
- Mesoscale Verification Intercomparison in complex Terrain (MesoVICT): Dorninger et al (2018), BAMS

http://www.ral.ucar.edu/projects/icp includes an list of more than 200 peer-review articles

Open source community verification tools: R spatialVx package, MET and METplus



2020-2021: JWGFVR achievements (1/3)

June 2021 tutorials: MPE-CDT and JWGFVR verification summer school
https://mpecdt.ac.uk/mpe-cdt-jwgfvr-forecast-verification-summer-school

Strengthening connections with various WMO groups (follow up from autumn 2020 SSC and WGNE recommendations):

- **Tropical Cyclones project:** R. Ashrit on committee
- **Paris 2024 RDP** (Valery Masson, telecon 25th Feb 2021): B. Casati on committee
- **Aviation2RDP** (Peter Li, telecon 15th March 2021): S. Landman on committee
- **Process diagnostics with WGNE** (C. deMott, J. Baker, telecon 9th June 2021):
  - ✓ a small subgroup of the JWGFVR (T. Haiden, M. Mittermaier, B. Casati, C. Coelho) is following this up through teleconferences, aiming to learn, test, and further develop existing process-based verification methods, leveraging possibly on the existence of currently available datasets (e.g. YOPPsiteMIP).
- **Continue contributing to S2S, PPP and HIW verification activities**:
  - 2nd WMO verification challenge: Develop and Demonstrate the Best New Verification Metric Using Non-Traditional Observations

2020-2021: redefining some of the priority verification research areas (2/3)

Following up from the November 2020 around the clock International Verification Method Workshop online (2020-IVMWO, https://jwgfvr.univie.ac.at), a **BAMS meeting summary article** was produced (Casati et al., 2021, soon to be published) **highlighting the priority verification research areas** in need of further developments, which includes the following:

- **Observational uncertainty and representativeness:** the WG is planning to follow up on this topic with DAOS and the data assimilation community
- **Process-diagnostics by conditional verification and verification of the relationships between variables:** the WG has started following-up on this topic with WGNE
- **Addressing the complexity of Earth System Modeling:** the WG is starting to broaden the verification research view considering **coupling** (e.g., ocean and sea-ice, land-atm interactions), and looking into ways to enhance synergies with different research communities (e.g., ocean, sea-ice, surface)

2020-2021: conducted verification research (3/3)

(see Sept 2021 SSC report/presentation or Extra Material for details)

- **NWP verification against own analysis by exploiting data assimilation confidence mask.**
- **Process-based 2-meter temperature forecast verification over China at 3-km resolution: conditional assessment under overcast, cloudy, partly cloudy, and clear sky conditions.**
- **Inter-comparison performance assessment of sub-seasonal to seasonal (S2S) prediction project models.**
- Evaluating sub-seasonal forecasts for water management in Brazil.
- **Seamless rainfall prediction skill comparison between global and reginal ensemble sub-seasonal prediction systems over India.**
- Bias correction and verification of lagged S2S daily precipitation ensemble predictions for driving downstream applications over India.
- **End-to-end verification of the ensemble precipitation-to-river-flow forecasting chain.**
- Feature-based evaluation of Chlorophyll-a blooms.
- Model diagnostic evaluation and warnings.
- **Evaluation of daily precipitation features in climate simulations including persistence, intermittency, spatial structure (size and orientation).**

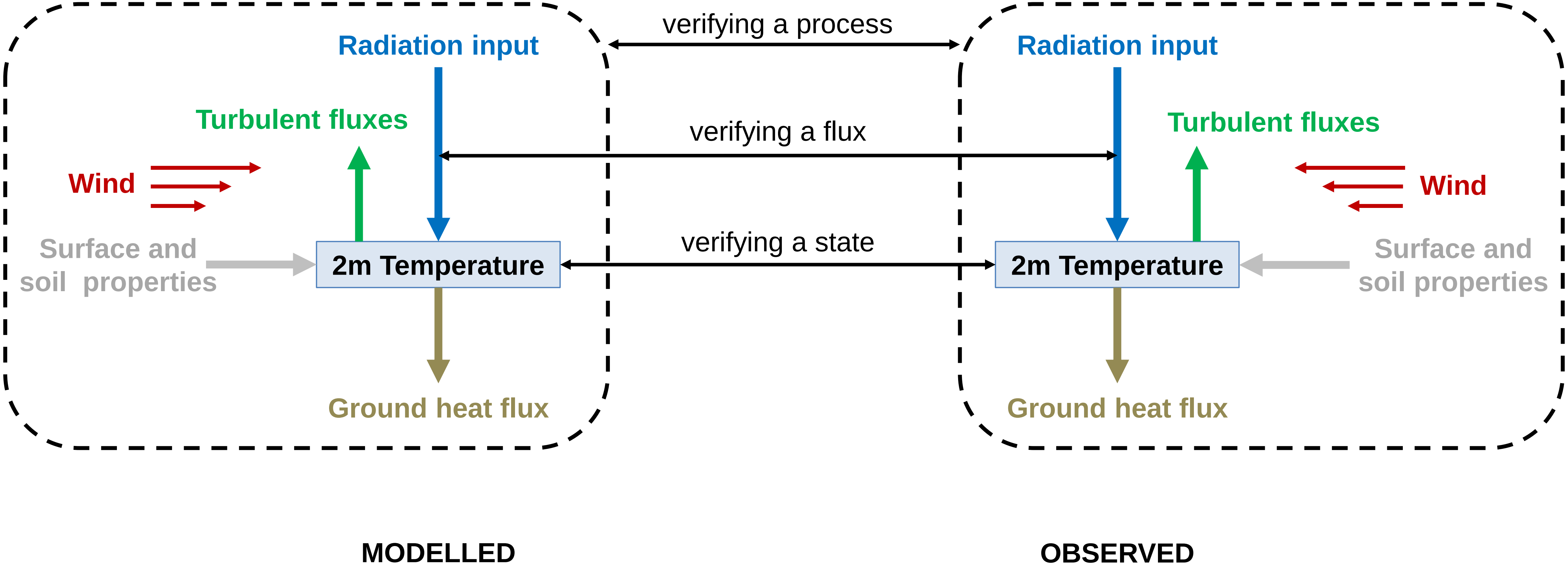
Outcomes from the 2020IVMWO (https://jwgfvr.univie.ac.at)
Error tracking and **process diagnostics**: most attended sessions!

- **Error (back) tracking techniques**, ensemble sensitivity analysis, and consequent relaxation experiments: a dynamical approach to error characterization, analyze the model error propagation in association with large-scale circulation (e.g. Magnusson 2017; Jung et al 2014; Lawrence 2019).
  - Relates to conditioning on weather types / composites / PCA / teleconnection studies;
  - Machine Learning could be exploited in these techniques (but not their interpretation).

- **Processes diagnostics** focus on verifying the **relationships between multiple variables**, which mirrors the physical process(es) interrelating such physical variables: e.g. Baker et al. 2021, Day et al 2020 / Miller et al 2018, look at the correlation and/or indices describing the relationship between different variables, and verify these correlations/indices.
- **Multivariate and multidimensional**: ought to include several **process-oriented variables** (e.g. fluxes) and explore relationships not only spatially (in 2D), but also **vertical profiles and time-series**; Need beyond traditional **co-located multivariate observations** (e.g. supersites).
- Process diagnostics can be improved by **conditional verification**, which can help stratify in a physically meaningful way (e.g. cloud versus clear-sky). **Multivariate statistics** also plays a role.

- Need to coordinate development of diagnostics with WGNE modelling community

Process verification

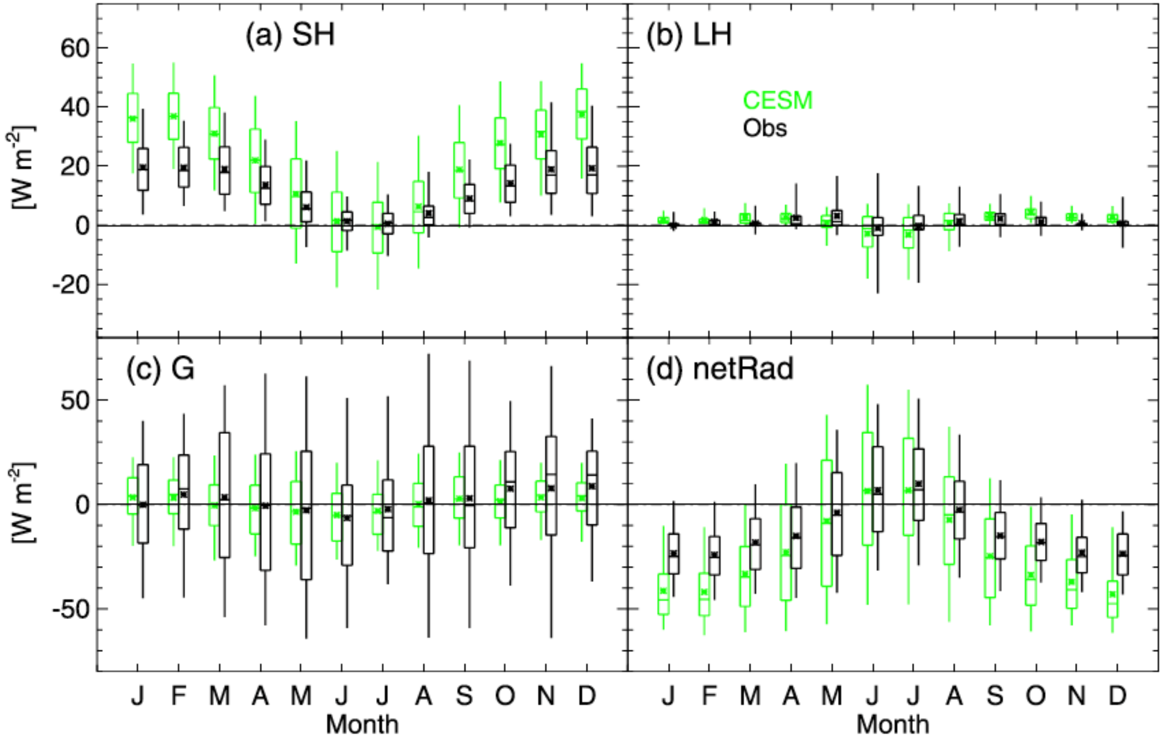


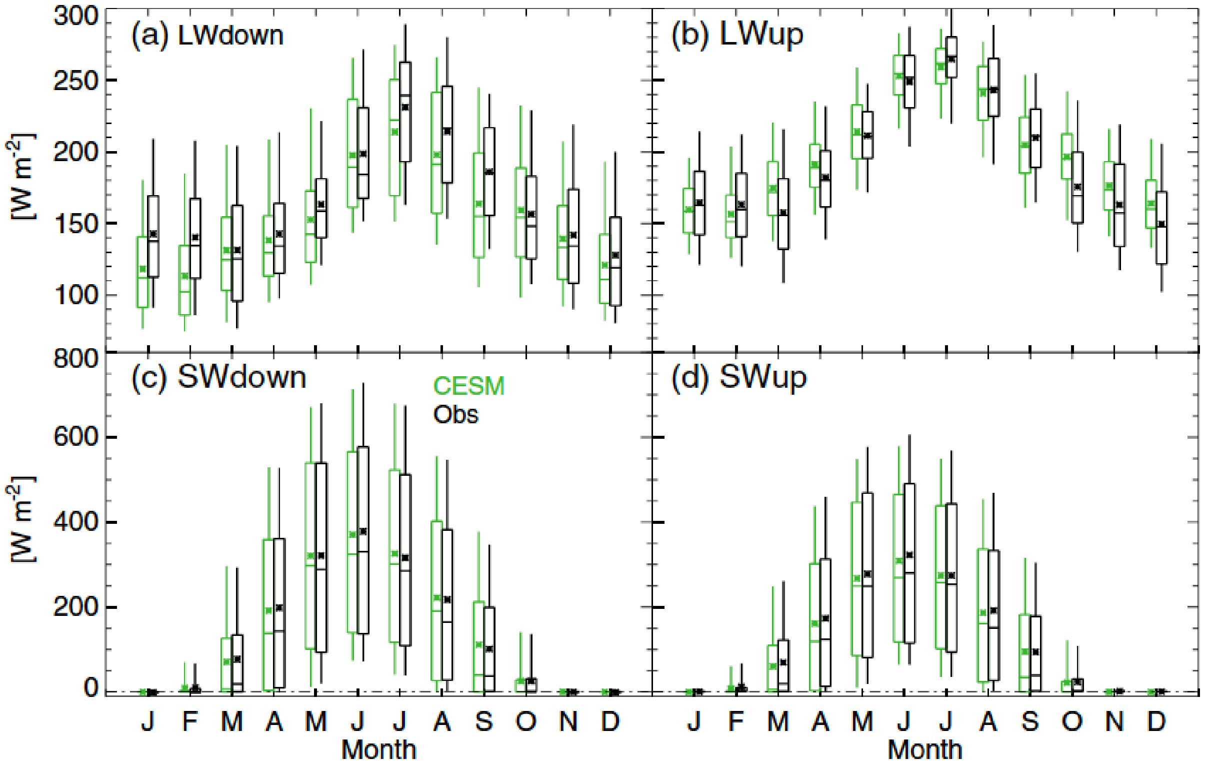
Slide from Thomas Haiden (ECMWF)

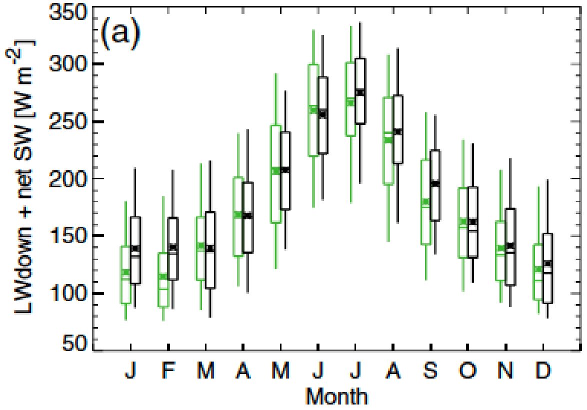
Examples from Miller et al (2018) and Day et al (2020)

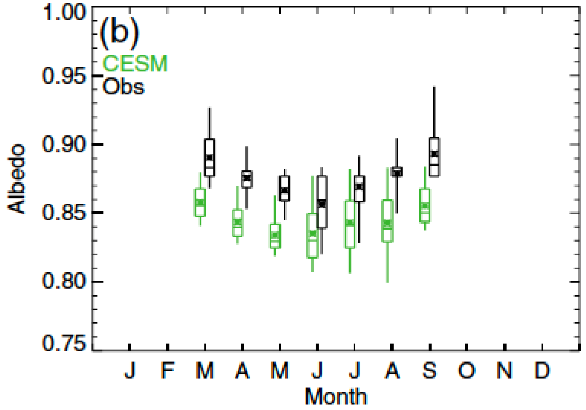
Analyze surface energy budget components to disentangle sources of T_{surf} biases and identify canceling errors:

LW↓ – LW↑ + SW↓ – SW↑ = – SH – LH – G



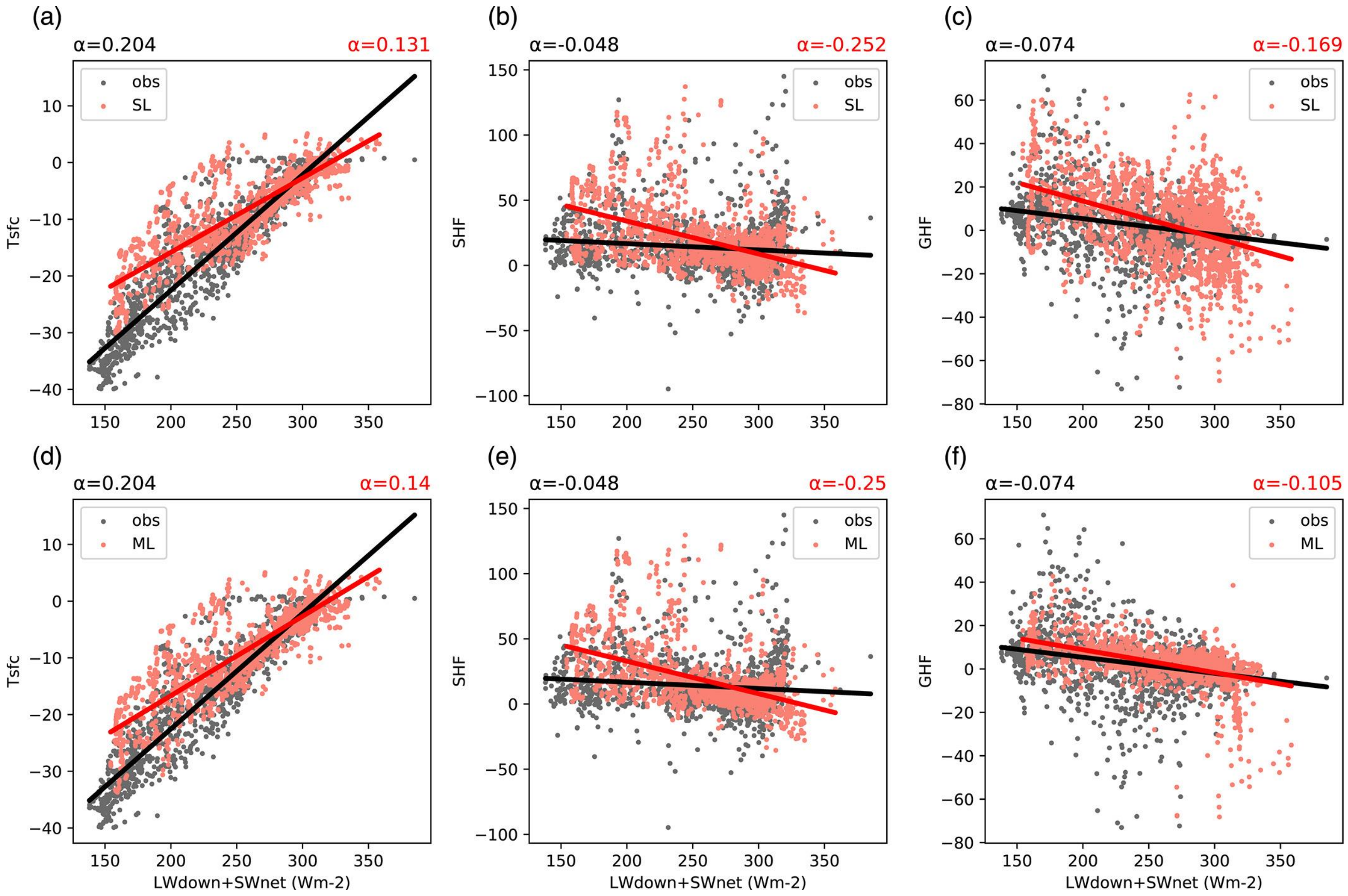


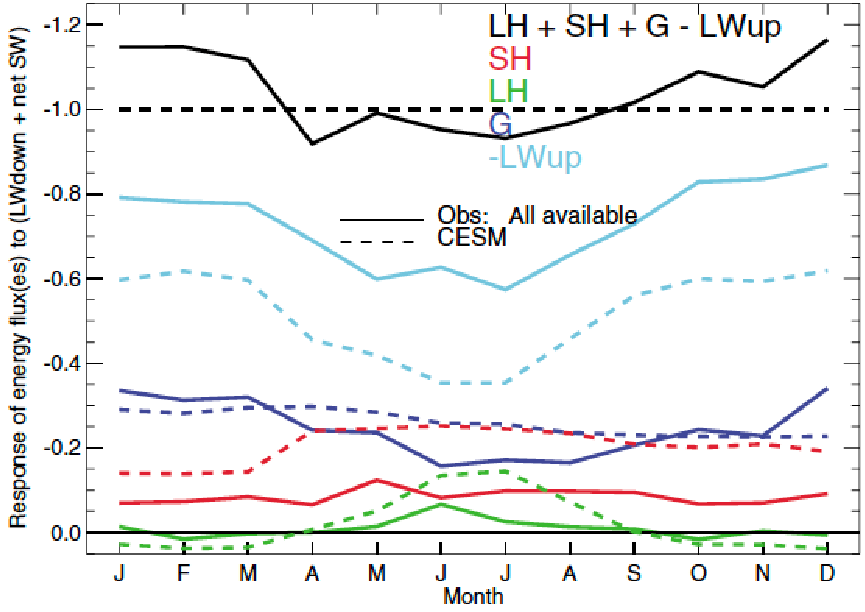




Examples from Day et al (2020) and Miller et al (2018)

Analyze the response of each Surface Energy Budget component to the radiative forcing = LW↓ + SW↓ – SW↑ = LW↑ – SH – LH – G



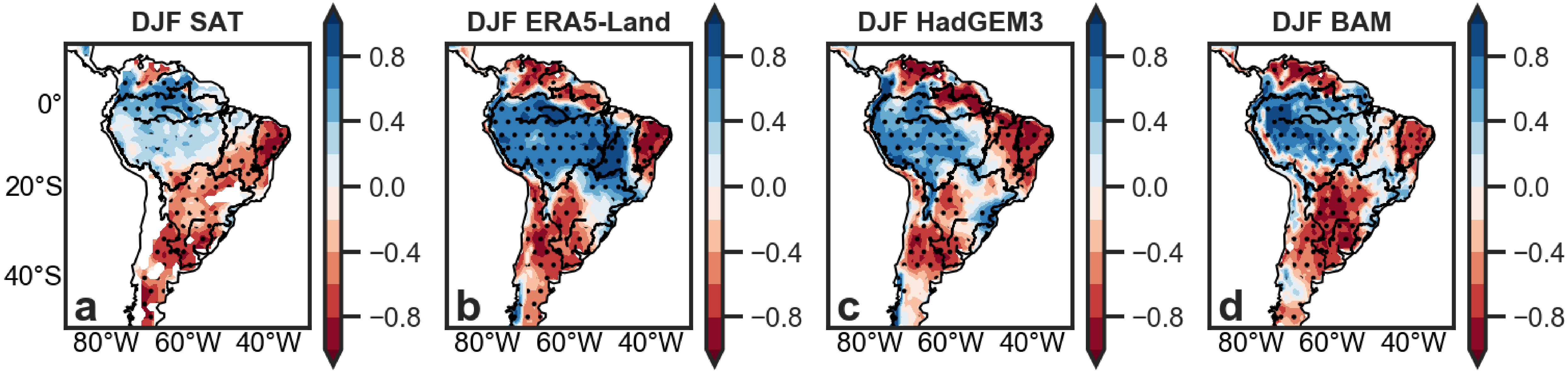


The relationship between the SEB variables here is approximated by the slope of the regression line

Examples from Baker et al (2021)

Analyze land-atmosphere interactions in climate model simulations by calculating some diagnostics describing the processes relating soil moisture, soil temperature, evapotranspiration and precipitation (based on previous literature), and then verifying the representation of these relationships spatially.

Temperature-Evapotranspiration Index (Seneviratne et al 2010)
in DJF estimated from satellite products, reanalysis, HadGEM3 and BAM-1.2



RED = land-surface influence on atmosphere
BLUE = atmospheric influence on land surface

- ERA5 and models capture dipole in evaporative regime across continent
- ERA5 and BAM-1.2 underestimate L-A coupling over Cerrado
- All tend to overestimate land-surface influence along Atlantic coast

Example from Baker et al (2021)

Two-legged metric (Dirmeyer et al. , 2011)

in DJF estimated from satellite products, reanalysis, HadGEM3 and BAM-1.2

Surface leg

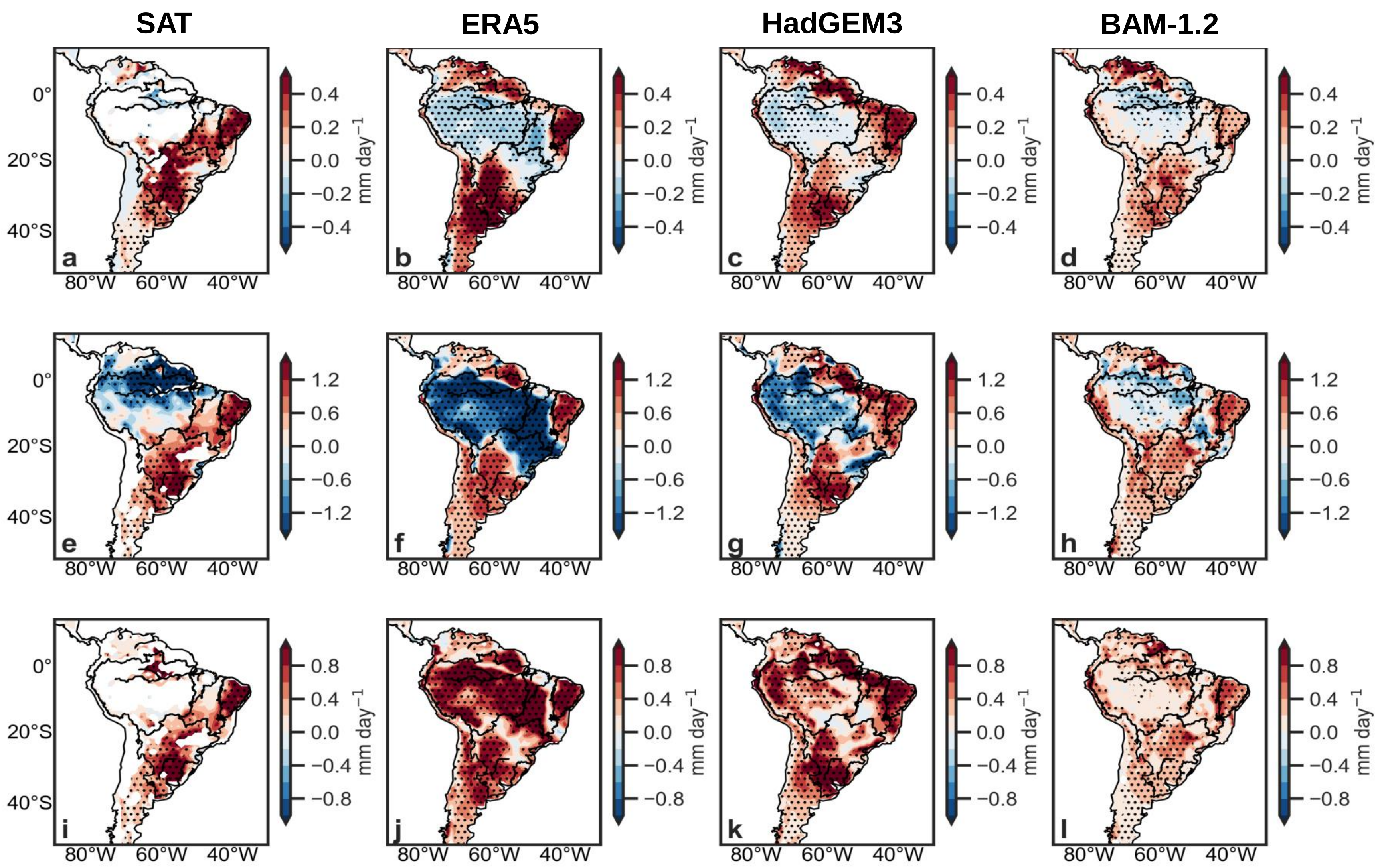
$$\frac{dET}{dSM} \cdot \sigma SM$$

Atmospheric leg

$$\frac{dP}{dET} \cdot \sigma ET$$

Surface-to-atmosphere moisture transfer pathway

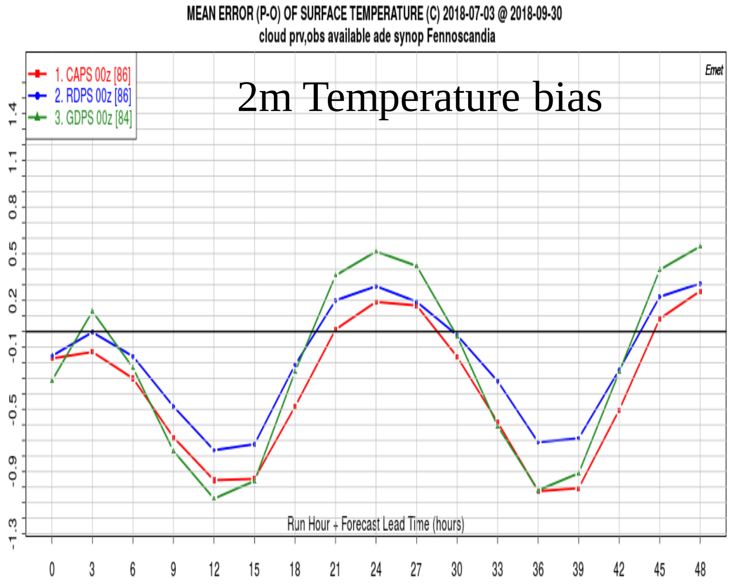
$$\frac{dP}{dSM} \cdot \sigma SM$$

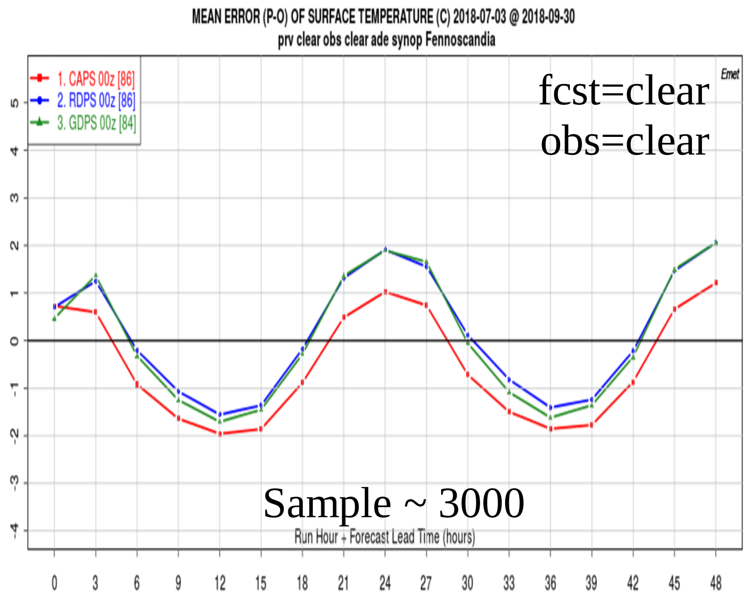


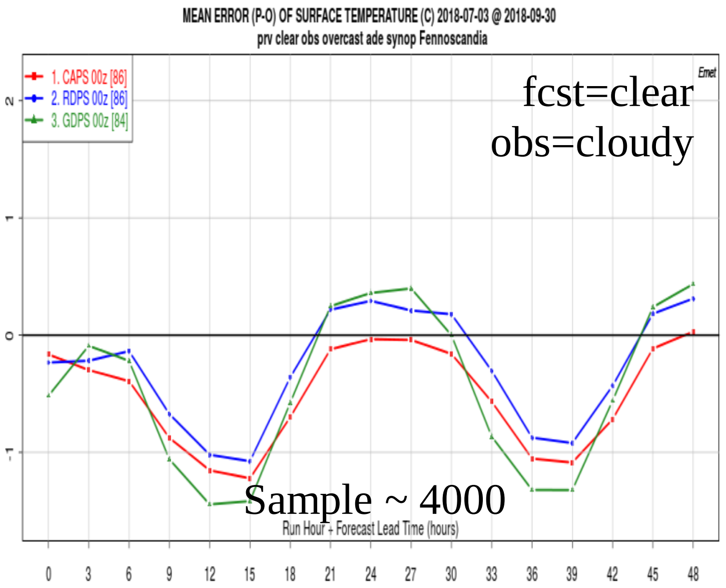
Partition feedback pathways:

1. Regression between surface state to surface flux
2. Regression between surface flux to atmospheric state
3. Compute the product and multiply by variability of surface state
4. End result represents the sensitivity of the total pathway

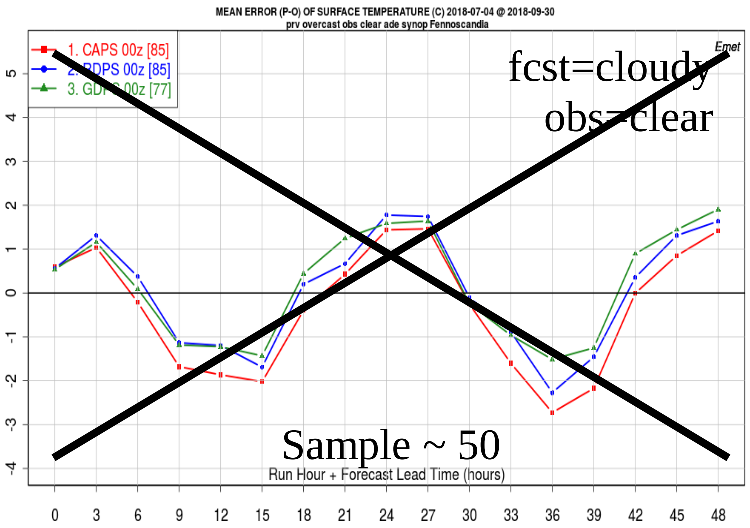
Conditional verification: physically meaningful stratification

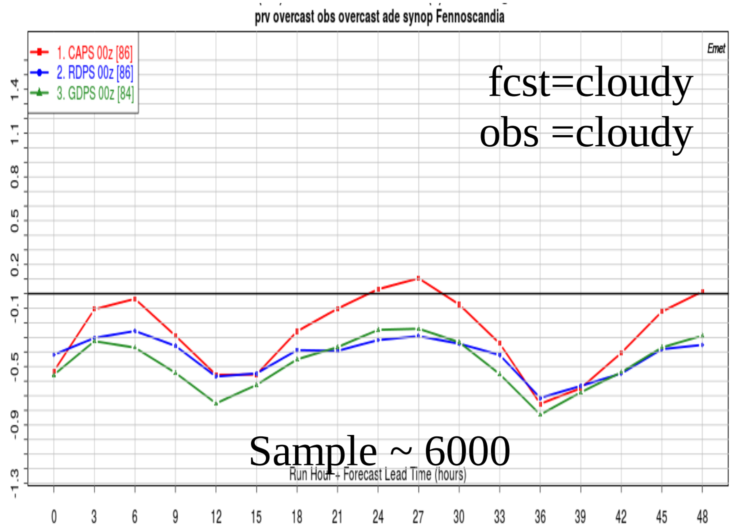






Example: 2m temperature bias as function of leadtime (0-48h), conditioned on cloud cover; Canadian Deterministic Prediction Systems: CAPS (3km) RDPS (10km) GDPS (25km), Fennoscandia, summer 2018





Challenges

1. Develop process-oriented verification research as **complementary** strategies to current model developers´ practices; in collaboration with the modeling community, we need to understand and seize the **added value** (and define the border line) for assessing how far process-oriented verification research should be pursued, to avoid duplicating model assessment efforts already performed by model developers
   - Example: corroborate/demonstrate that the behaviour of **the case study is representative of the whole period**, that **the super-site** processes are **representative of the whole domain** … with added statistical inference.
   - **Operationalization**: need to start simple (e.g. well-known processes related to well-observed variables …).
2. Analysing interactions of **different components of the earth system:** atmosphere, ocean and sea-ice, surface, ... (implies also linking with the different obs and modelling communities).

3. **Observations, observations, observations … and Quality Controlled observations!**
   - It might be less a matter of new scores, and more the level of completeness of observations …
   - Super-sites provide a suitable playground, but eventually need broader spatial coverage (e.g. satellites …)
   - Representativeness issue will be part of the challenges …

Future Plans (to the end of 2023, and hopefully beyond …)

Maintaining JWGFVR Legacy:

- Organize the 9th **International Verification Methods Workshops** (IVMW)
- Deliver **verification tutorials**
- Keep advancing and operationalize **spatial verification methods** (http://projects.ral.ucar.edu/icp/)
- Unify all **web resources** developed by the group in the past 20 years, as reference and legacy
- Keep supporting verification research activities in **WMO projects and WG** (PPP, HIW, S2S, Paris2024RDP, AvRDP2, Tropical Cyclones, …)

Re-newed Research Foci:

- **Processes diagnostics and ESM verification** (including the interaction of different variables and model components) in collaboration with modellers / WGNE and other WG (e.g., YOPPsiteMIP in PPP; Paris2024 for urban BL)
- **Verification for targeted downstream communities** (aviation, hydrology, urban)
- **Exploitation of data assimilation knowledge in forecast verification:** representativeness and observations uncertainty
- **Join efforts on model evaluation with the climate community** (both for upstream -modeling- and downstream -e.g. post-processing- use)

Environment and Climate Change Canada / Environnement et Changement climatique Canada

JWGFVR

Thank you for your attention!

JWGFVR Members: Barbara Casati (ECCC, co-chair), Caio Coelho (CPTEC, co-chair), Raghu Ashrit (NCMRWF), Marion Mittermaier (UK Met Office), Jing Chen (CMA), Manfred Dorninger (U. Vienna), Eric Gilleland (NCAR), Thomas Haiden (ECMWF), Stephanie Landman (SAWS), Chiara Marsigli (DWD)

Canada

Extras

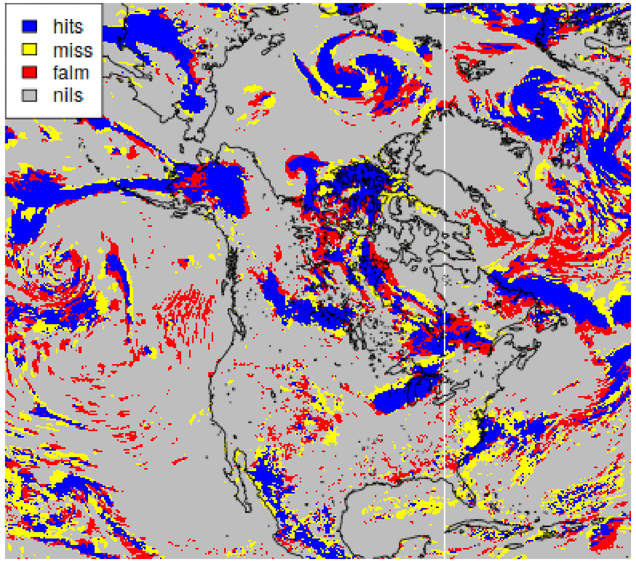
2020-2021 conducted verification research

- **NWP verification against own analysis by exploiting data assimilation confidence mask.**
- **Process-based 2-meter temperature forecast verification over China at 3-km resolution: conditional assessment under overcast, cloudy, partly cloudy, and clear sky conditions.**
- **Inter-comparison performance assessment of sub-seasonal to seasonal (S2S) prediction project models.**
- Evaluating sub-seasonal forecasts for water management in Brazil.
- **Seamless rainfall prediction skill comparison between global and reginal ensemble sub-seasonal prediction systems over India.**
- Bias correction and verification of lagged S2S daily precipitation ensemble predictions for driving downstream applications over India.
- **End-to-end verification of the ensemble precipitation-to-river-flow forecasting chain.**
- Feature-based evaluation of Chlorophyll-a blooms.
- Model diagnostic evaluation and warnings.
- **Evaluation of daily precipitation features in climate simulations including persistence, intermittency, spatial structure (size and orientation).**

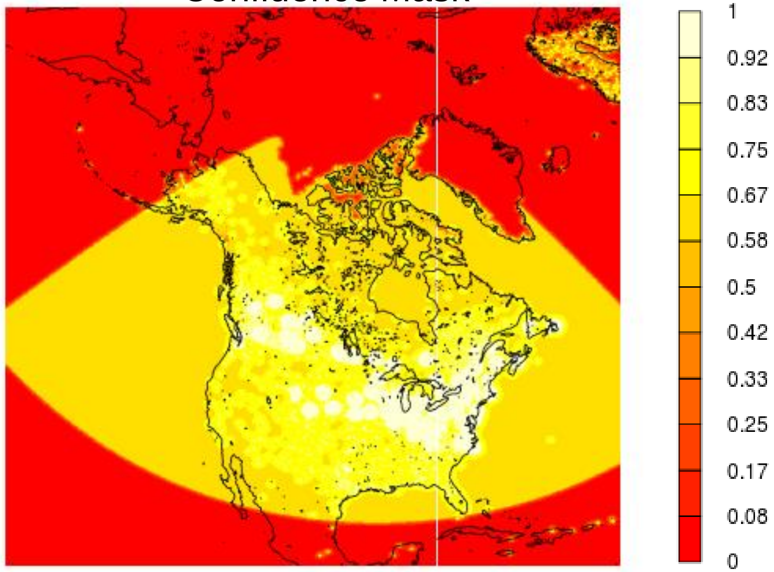
B.Casati, V.Fortin, F.Lespinas, D.Khedhaouiria: "**NWP verification against own analysis by exploiting Data Assimilation confidence mask** " Wea&For, in preparation

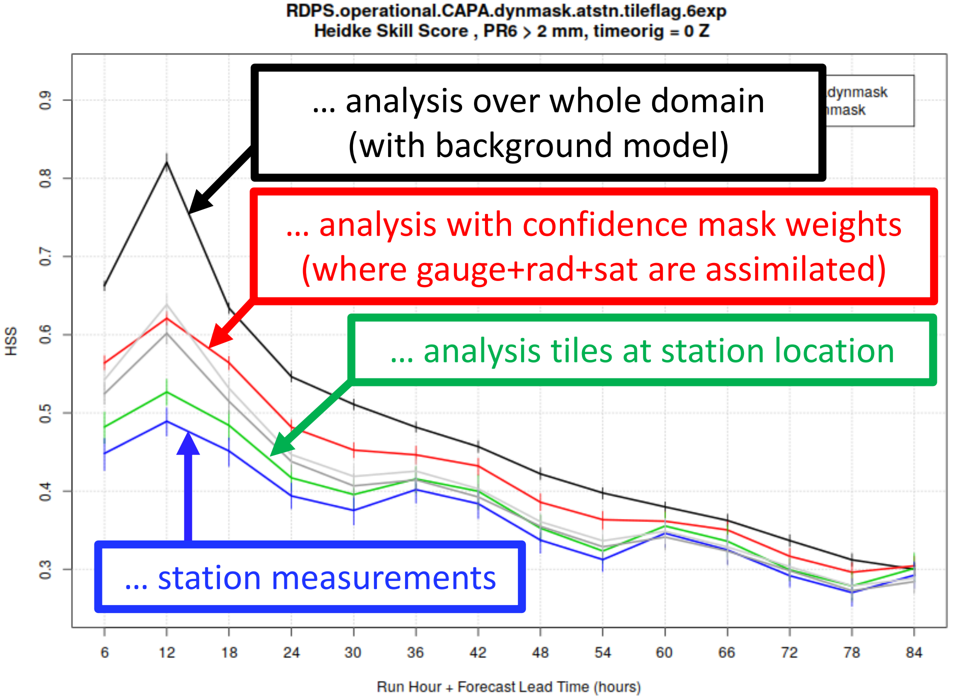
Environment and Climate Change Canada

Contingency Table Counts



Canadian Precipitation Analysis Confidence Mask





Aim: propose a verification approach against model-based analysis which **exploit Data Assimilation knowledge**, **by weighting verification scores with a DA confidence mask** which:

1. Reduces the background model influence (assigns zero weight if analysis = background)
2. Gives larger weights where/when more observations are assimilated
3. Assigns larger/smaller weights based on the confidence/uncertainty associated to the verifying analysis / assimilated observations

Process-based T2m forecast error of GRAPES-MESO 3km model

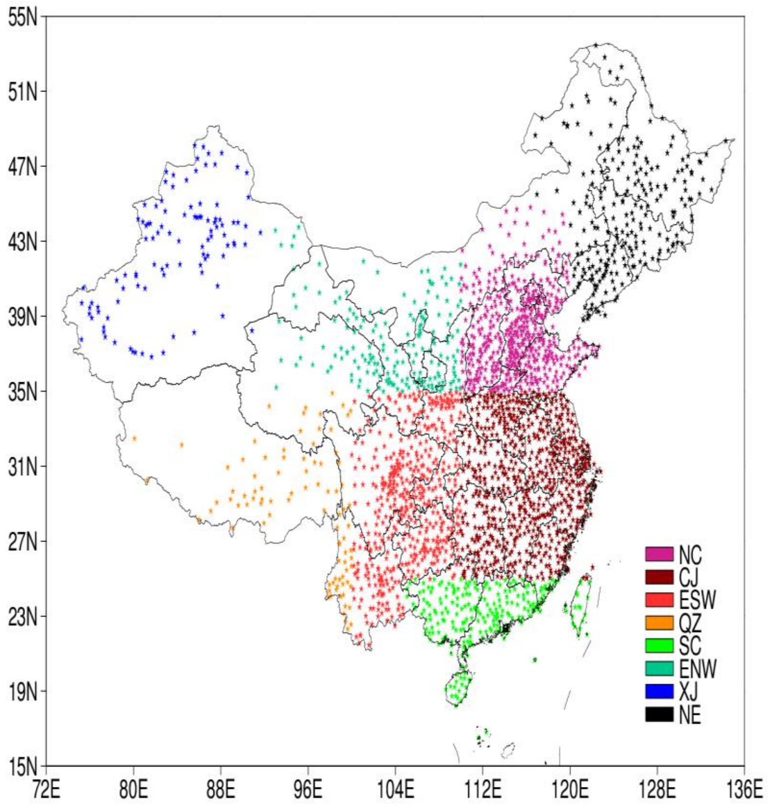


Fig 1:Distribution of 2420 stations

J. Chen (CMA)

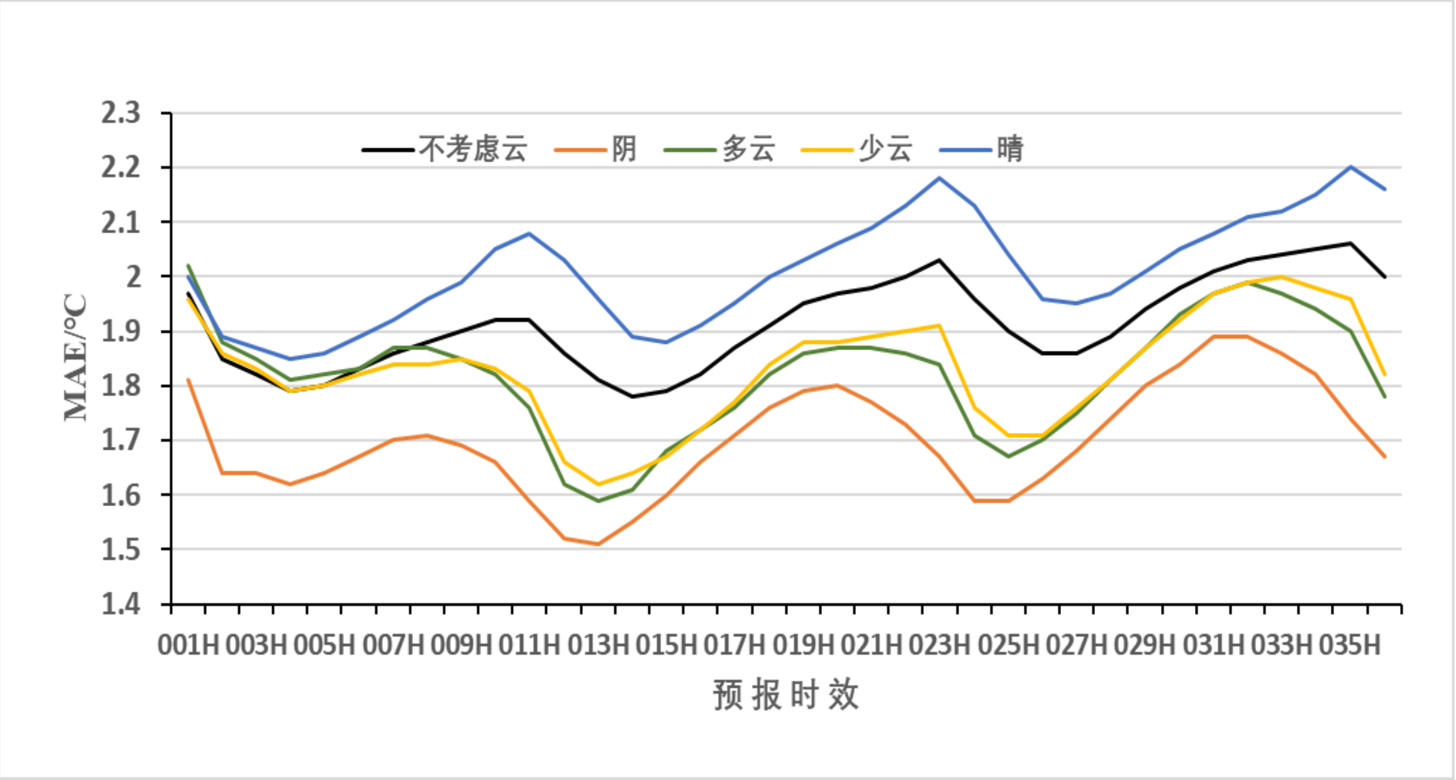
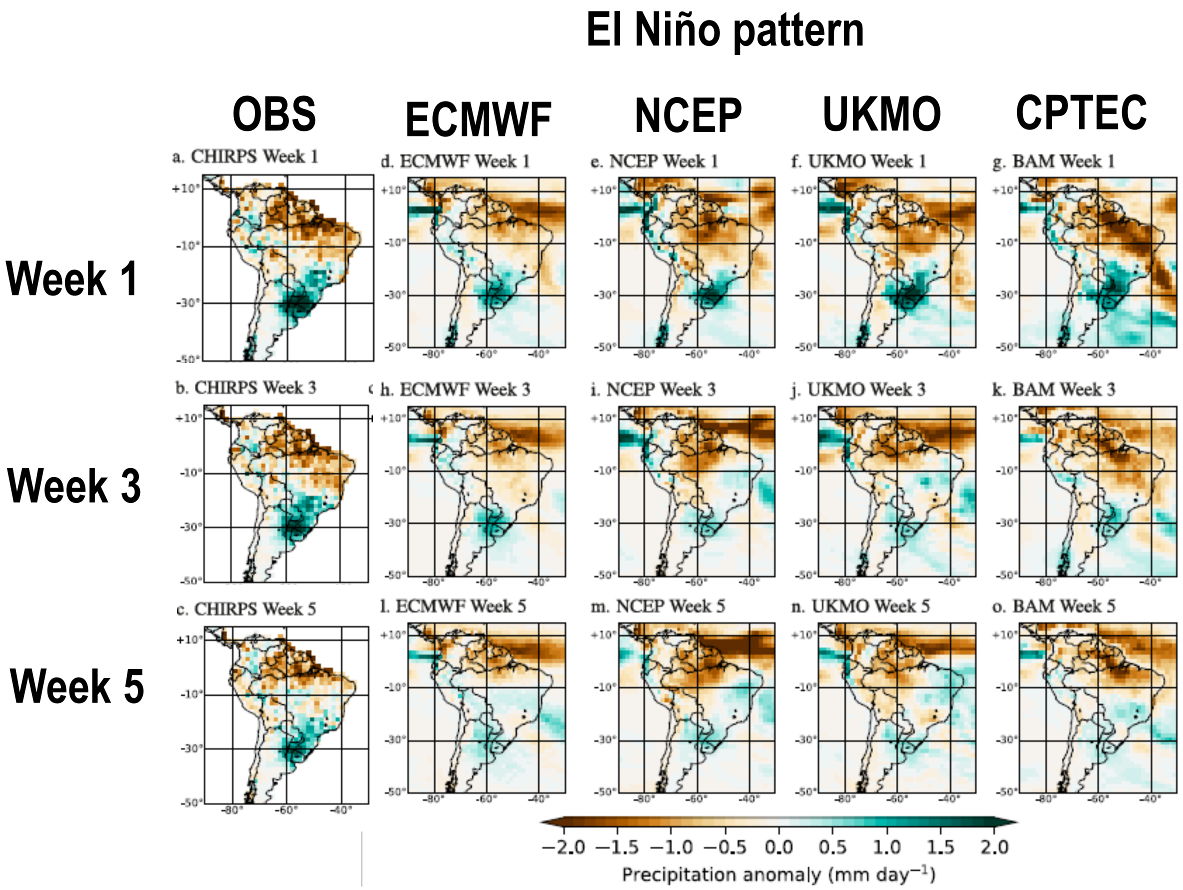
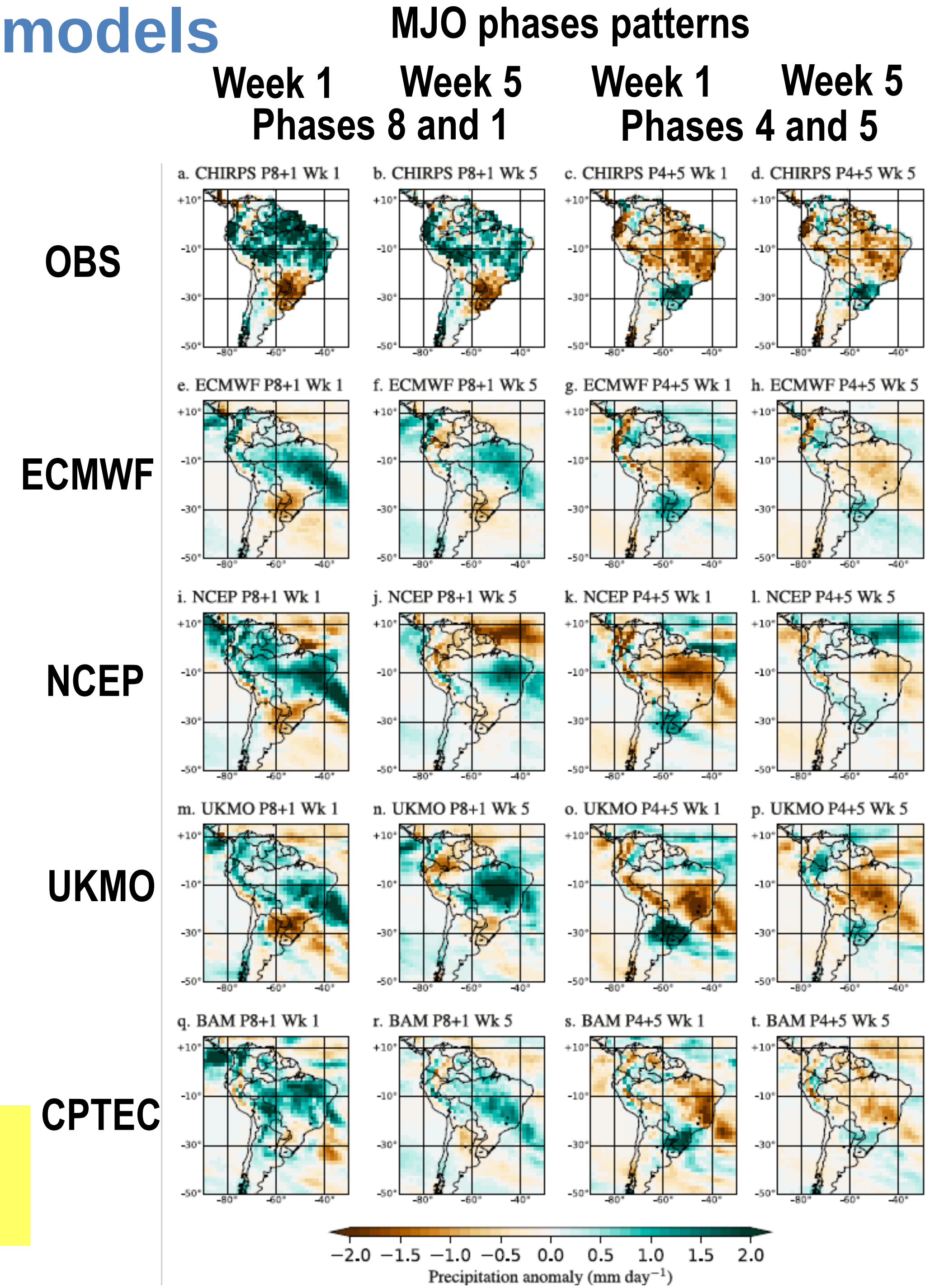


Figure 2. Process-based mean absolute error for 0-36h leading time forecast of temperature at 2m height by GRAPES-MESO 3km model from initial time 00UTC and 12UTC, 1 Jul. to 31 Dec.,2020

— all, — overcast, — cloudy, — partly cloudy, — clear sky,

How well CPTEC and S2S prediction project models represent El Niño and MJO phase precipitation patterns?





Kingaman, N. P., Young, M., Chevuturi, A., Guimarães, B., Guo, L., Wooolnough, S. J., Coelho, C. A. S., Kubota, P. Y., Holloway, C. E., 2021: Subseasonal prediction performance for austral summer South American rainfall. WEATHER AND FORECASTING, 36 (1), 147-169.

Met Office

Ministry of Earth Sciences

Seamless rainfall prediction skill comparison between GloSea5 and NEPS-G ensemble prediction systems

Newton Fund

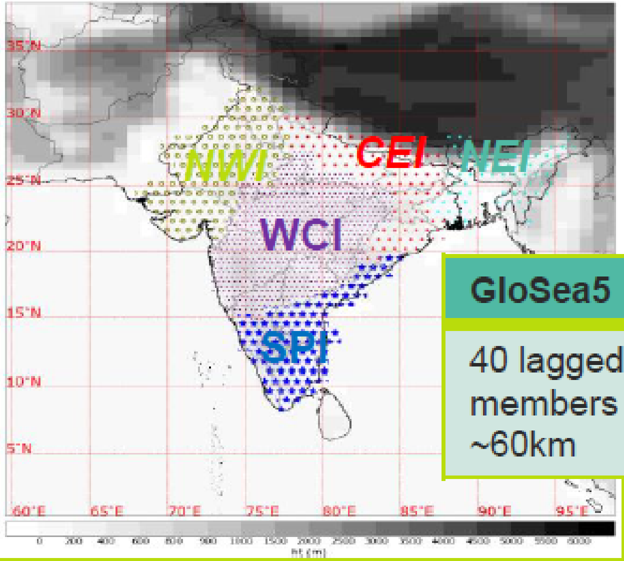
Motivation

Compare a range of time windows between S2S and NWP ensemble type systems

Observations

Daily GPM regridded to the GloSea5 grid

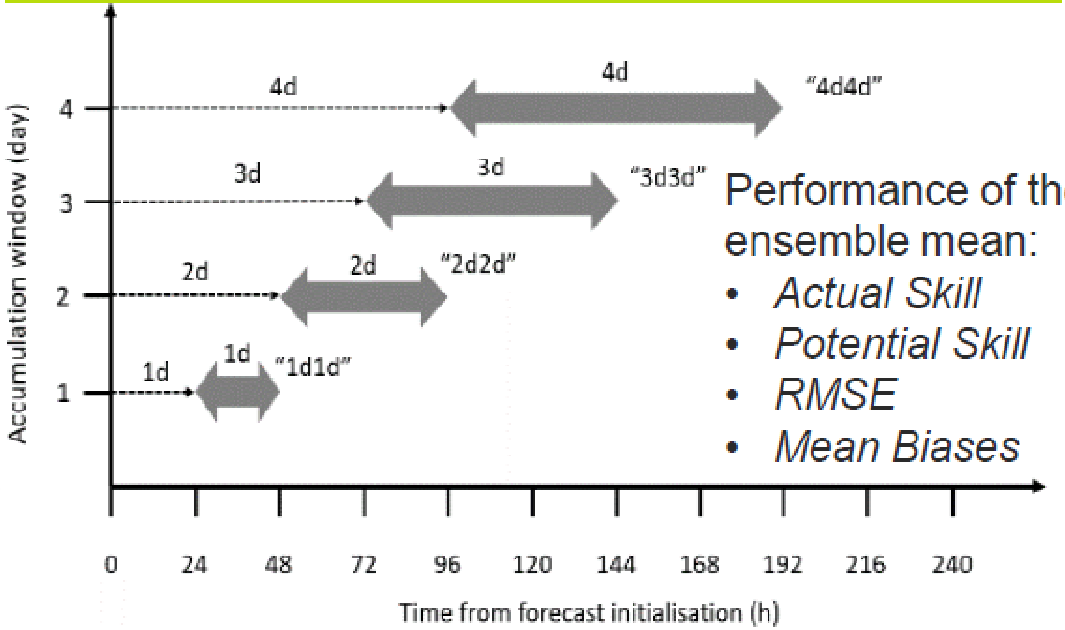
Models & domain



GloSea5	NEPS-G
40 lagged members @ ~60km	23 members @ ~12km

Raw model, i.e. no bias adjustments!

Method (adapted from Wheeler et al. 2016)

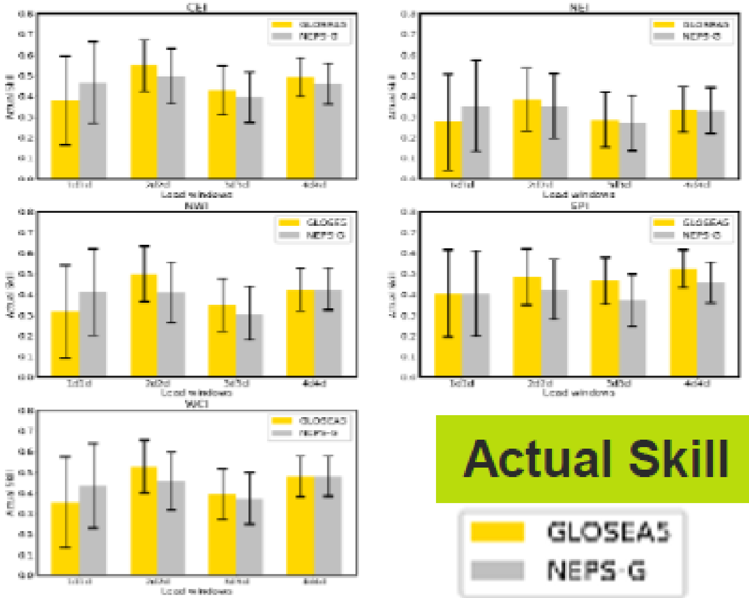


Performance of the ensemble mean:

- *Actual Skill*
- *Potential Skill*
- *RMSE*
- *Mean Biases*

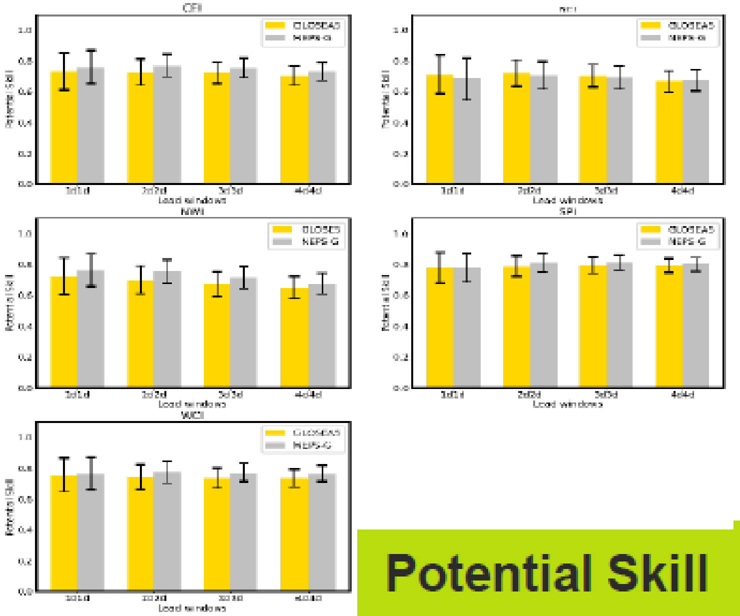
Focus on early lead times where modelling systems overlap.

- **GloSea5 has greater actual skill** over all Indian homogeneous regions except for the 1d1d lead window
- **NEPS-G regional potential skill is almost unanimously higher.**



Actual Skill

GLOSEA5
NEPS-G



Potential Skill

Kolusu et al. (2021), EMS

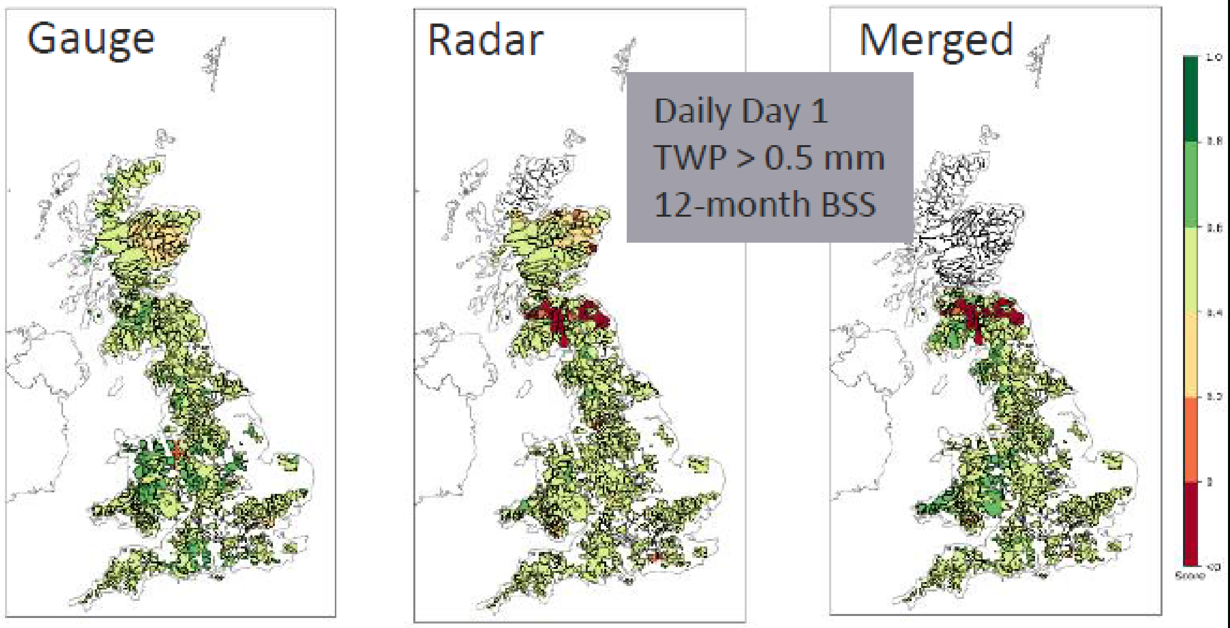
Met Office

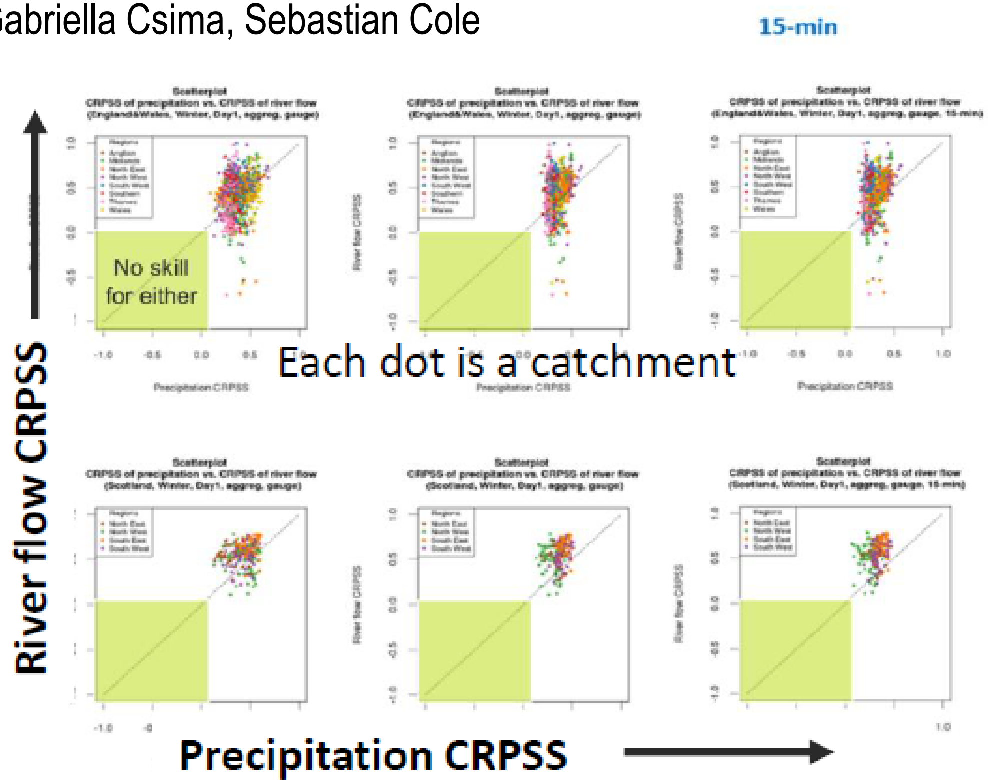
UK Centre for Ecology & Hydrology

End-to-end verification of the Ensemble precipitation-to-river-flow forecasting chain: How to maximize the skill for the user and does the uncertainty propagate?

Marion Mittermaier, Seonaid Anderson, Ric Crocker, Robert Moore, Steve Cole, Gabriella Csima, Sebastian Cole

Different observation types can lead to larger differences in scores than regional or temporal variations.



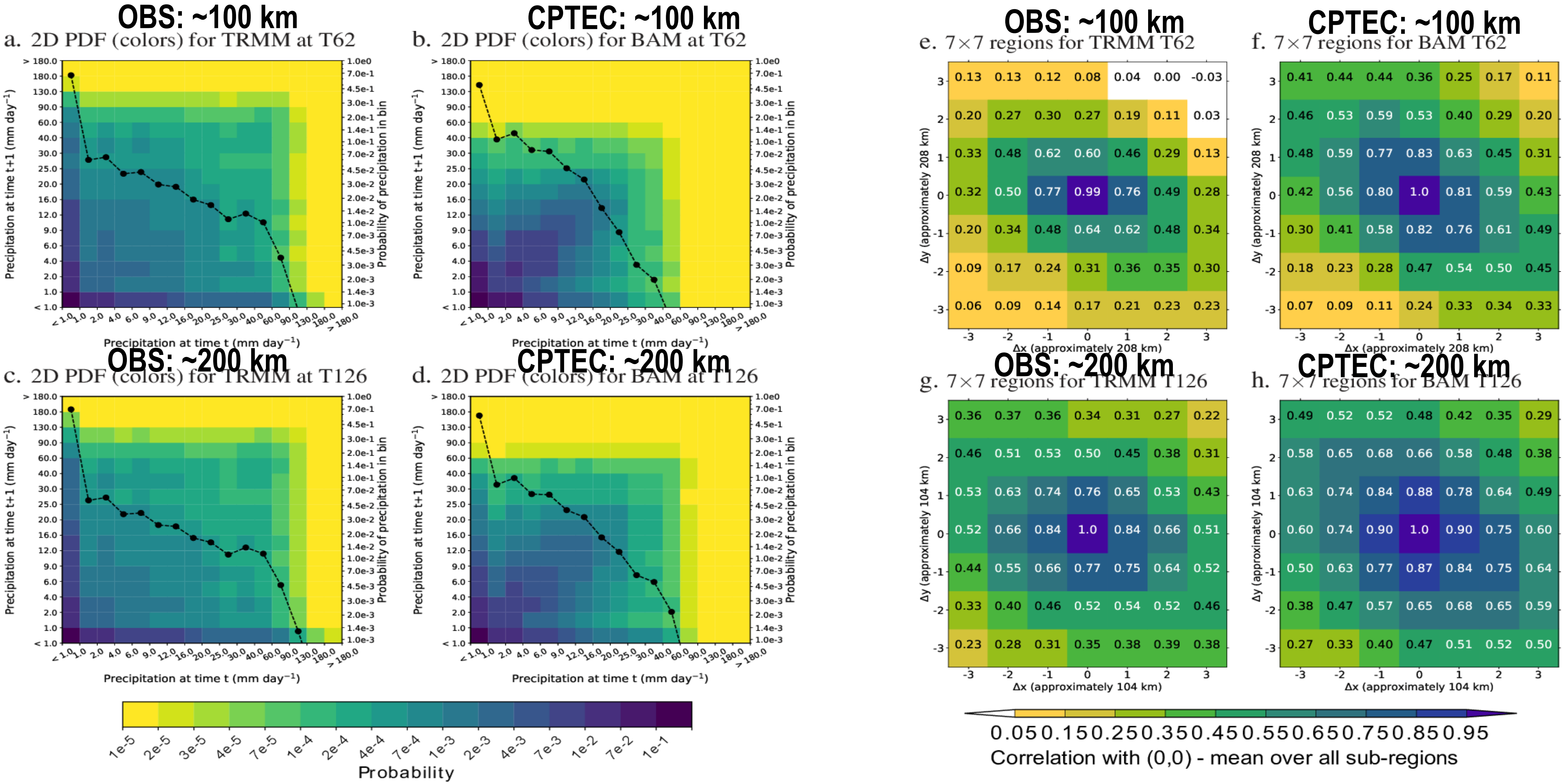


CRPS for catchment precipitation ensemble forecasts driving the G2G ensemble for FFC; how do the errors/skill in the precipitation translate to the river flow?
Anderson et al. (2019) and another paper to follow.

- **Comparing 15-mim precip driving hydrological model to produce 15-mim river flow**
- **Range of river flow CRPSS much wider, some catchments showing distinct insensitity to precipitation forecast skill**
- **Precipitation forecast skill and uncertainty is only one of the factors which will influence river flow forecast skill**

Report for FFC, EA, SEPA and NRW available here

Evaluation of daily precipitation features in climate simulations: persistence, intermitency, spatial structure (size and orientation)



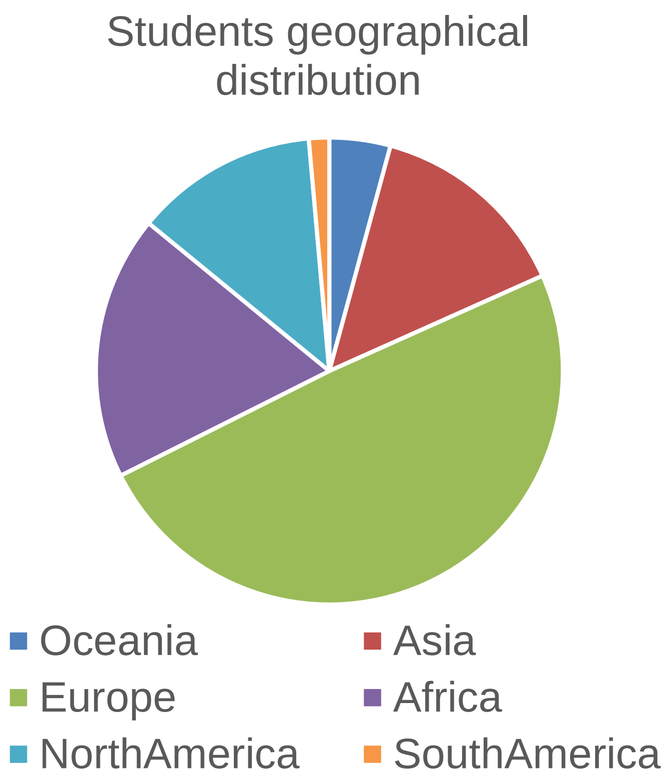
Caio A. S. Coelho, Dayana C. de Souza, Paulo Y. Kubota, Simone M. S. Costa, Layrson Menezes, Bruno S. Guimarães, Silvio N. Figueroa, José P. Bonatti, Iracema F. A. Cavalcanti, Gilvan Sampaio, Nicholas P. Klingaman, Jessica C. A. Baker, 2021: Evaluation of climate simulations produced with the Brazilian global

Delivery of online verification summer school

Jointly organized by the Mathematic of Planet Earth – Centre for Doctoral Training (MPE-CDT) and JWGFVR 21-25 June 2021

https://mpecdt.ac.uk/mpe-cdt-jwgfvr-forecast-verification-summer-school/

- 71 students from several counties around the world
- 10 theoretical lectures (50% external WG), complemented by some practical assignments + several Q&A drop-in sessions.
- Four complementary group projects, expanding on the lecture assignments. At the end of the tutorial week they presented, and the best two team projects were prized.
- Covered topics: verification basic concepts, traditional continuous and categorical scores, probabilistic and ensemble, inference, experimental design, sub-seasonal to seasonal, climate indices, spatial verification methods, verification of high impact weather



Support of host institution was fundamental for i) communication + organizational tasks; ii) IT arrangements (telecon + video-recordings); iii) host + open-access verification data; iv) providing basic codes for reading data.