Joint Working Group on Forecast Verification Research report

Caio Coelho
CPTEC/INPE

Marion Mittermaier
Met Office

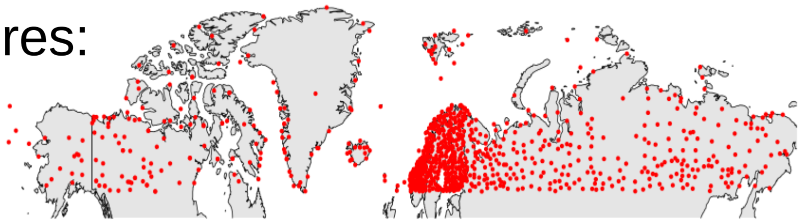
Outline

- Overview of working group involvement with WMO projects/activities
- Documents completed or in preparation:
  - Mesoscale Verification Inter-Comparison over Complex Terrain (MesoVICT) BAMS paper (final workshop)
  - Process-oriented verification
  - Novel observations
  - Precipitation review

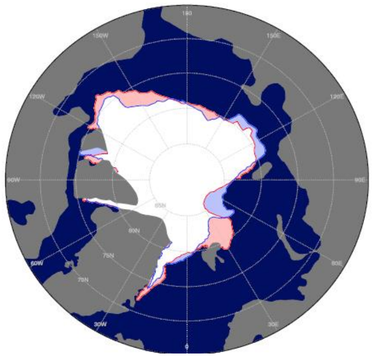
Update on global surface verification activities and enabling inter-comparison:

- Stratification of 2m temperature by dominant land-surface type or location
- Land-sea split of daily precipitation performance

YOPP core phase verification activities



1. **Operational summary verification** scores: YOPP is providing the framework for analyzing current verification practices in the Polar Regions: reveal issues, investigate solutions, **propose novel approaches**



2. Verification of **sea-ice** prediction during YOPP: user-informative **distance metrics** alongside traditional scores (e.g. Baddeley + IIEE + categorical scores)

3. NWP **process evaluation** against high frequency multivariate observations **at the YOPP super-sites.**
- A unique dataset of paired NWP model output and multivariate high-frequency obs which enables detailed process-based diagnostics.
- Target processes: clouds micro- and macro-physics; aerosols and hydro-meteors micro-physics; radiation, turbulence and energy budgets; energy and momentum fluxes.

YOPP YEAR OF POLAR PREDICTION

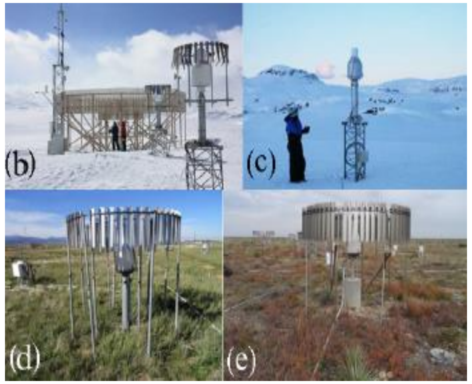
1. YOPP Operational summary verification scores

Tom Robinson, Barbara Casati (ECCC); Thomas Haiden, Martin Janousek (ECMWF); Morten Køltzow, Teresa Valkonen (Met Norway); Eric Bazile (MetFrance).

The activities consisted in comparing operational verification practices in the Polar Regions: exchange of objective verification scores during the YOPP Special Observing Periods (SOPs)

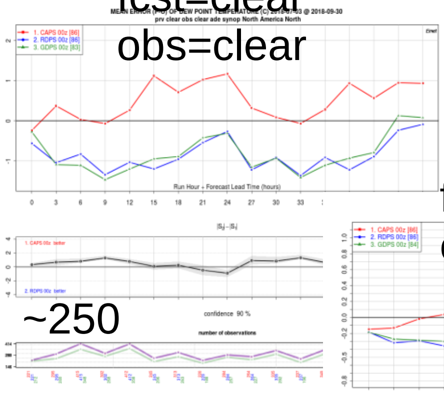
Key findings include:

2. Address Solid precipitation under-catch by using WMO-SPICE adjustment function



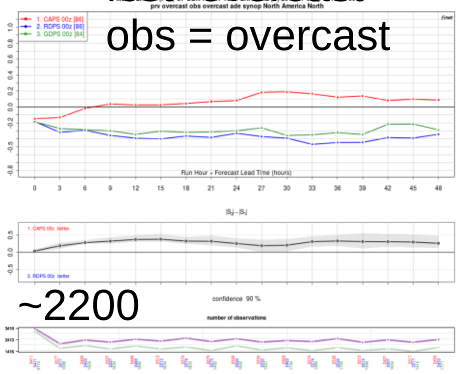
1. Apply (process driven) conditional verification

fcst=clear
obs=clear



~250

fcst=overcast
obs = overcast



~2200

3. Mitigate effects of network inhomogeneity by thinning or weighting ∝ station density

Images courtesy Barbara Casati

YOPP
YEAR OF POLAR PREDICTION

2. Verification of sea-ice prediction

Barbara Casati, JF Lemieaux, Ji Lei, Greg Smith (ECCC); Pam Posey, Julie Crout, Rick Allard (NRL); Bob Grumbine (NOAA); Malte Müller, Arne Melsom (MET Norway); Helge Goessling, Lorenzo Zampieri (AWI); Bill Merryfield et al (ECCC); Steffen Tietsche, Sarah Keeley, Jonny Day (ECMWF).

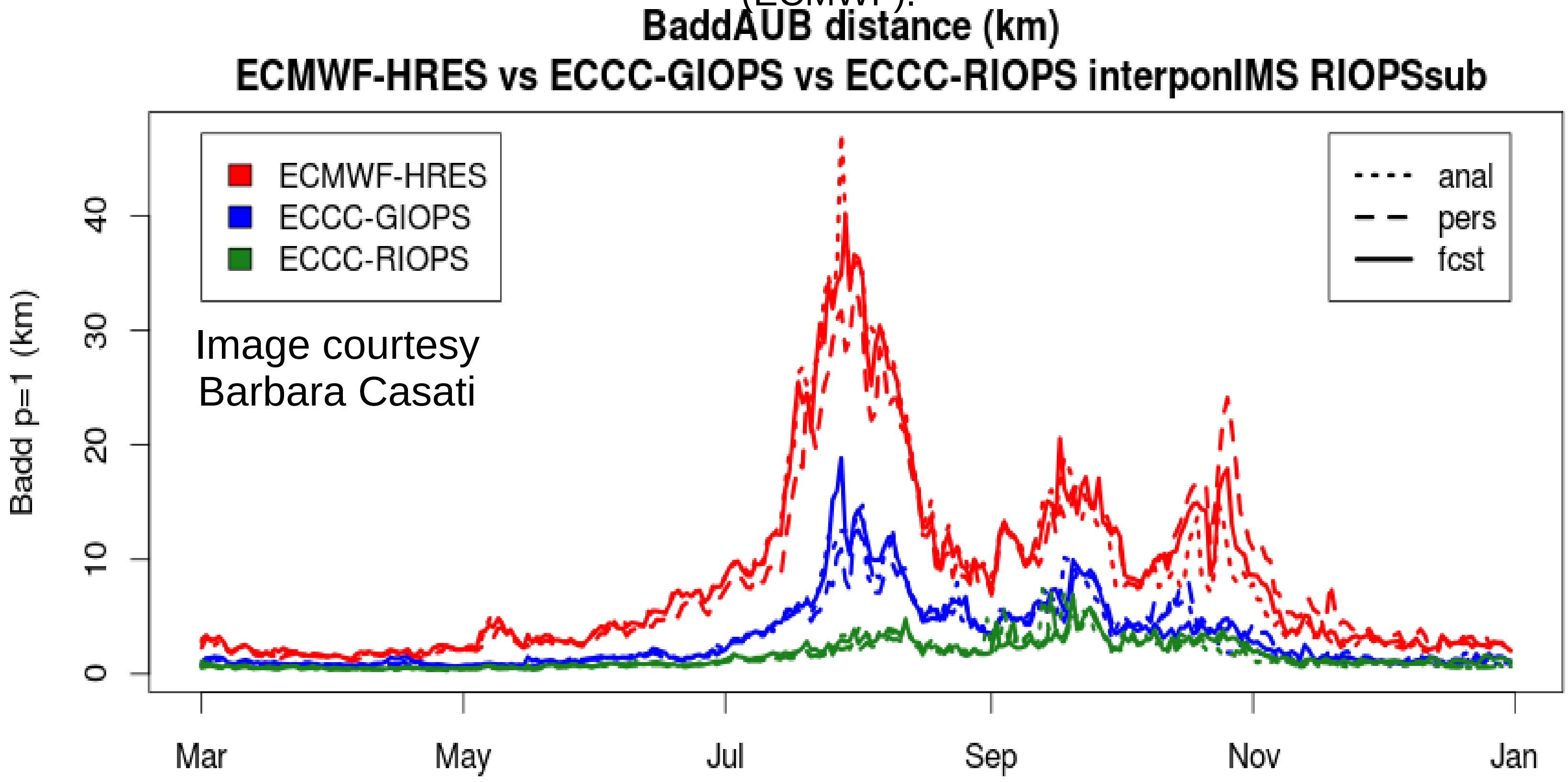


Image courtesy Barbara Casati

The sea-ice community has fully adopted user-informative distance metrics alongside traditional categorical scores.

- Currently: focus on **ice concentration, ice edge**
- Desiderata: **ice thickness, ice drift, ice pressure, MIZ**
- Challenge: exploit / improve **satellite products**

YOPP YEAR OF POLAR PREDICTION

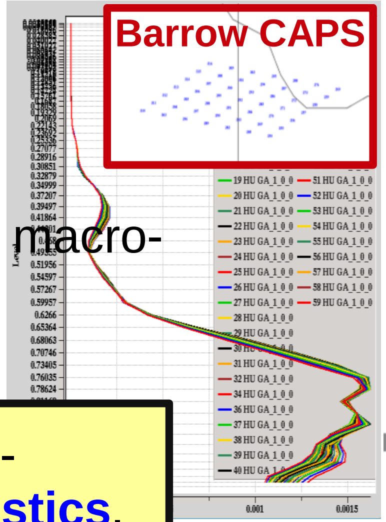
3. NWP process-based evaluation against high frequency multivariate obs at the YOPP super-sites

Gunilla Svensson (U. Stockholm); Taneil Uttal (IASOA,NOAA); Barbara Casati, Zen Mariani (ECCC); Jonny Day (ECMWF); Morten Køltzow (MetNo); Matthew Shupe (NOAA, Mosaic); Siri-Jodha Khalsa (NSIDC).

- **Arctic and Antarctic observatories**, furnished by suites of instruments that provide detailed measurements characterizing the vertical column of the atmosphere as well as the surface conditions and energy fluxes.
  - ➔ IASOA merged observatory data files

- Modelling centres (ECMWF, ECCC, Meteo France, …) are providing **NWP model output** at high frequency (on the order of model time-step) on model levels to enable comparison with the measurements available at the YOPP super-sites.

- **Target processes** include the representation of cloud micro- and macro-physics; aerosols and hydro-meteors micro-physics; radiation, turbulence and energy budgets; energy and momentum fluxes.



- This unique dataset of paired model output and multi-variate high-frequency observations enables detailed **process-based diagnostics**.
  - ➔ Open access via the YOPP data portal: https://yopp.met.no/

YOPP
YEAR OF POLAR PREDICTION

Forecast Verification activities

Caio Coelho

• Published forecast verification chapter on S2S book - Sub-Seasonal to seasonal prediction: The Gap Between Weather and Climate Forecasting (October 2018).

• Attended the Second International Conference on Subseasonal to Seasonal Prediction (S2S), at NCAR in Boulder (September 2018) and presented a proposed verification framework for sub-seasonal precipitation predictions.

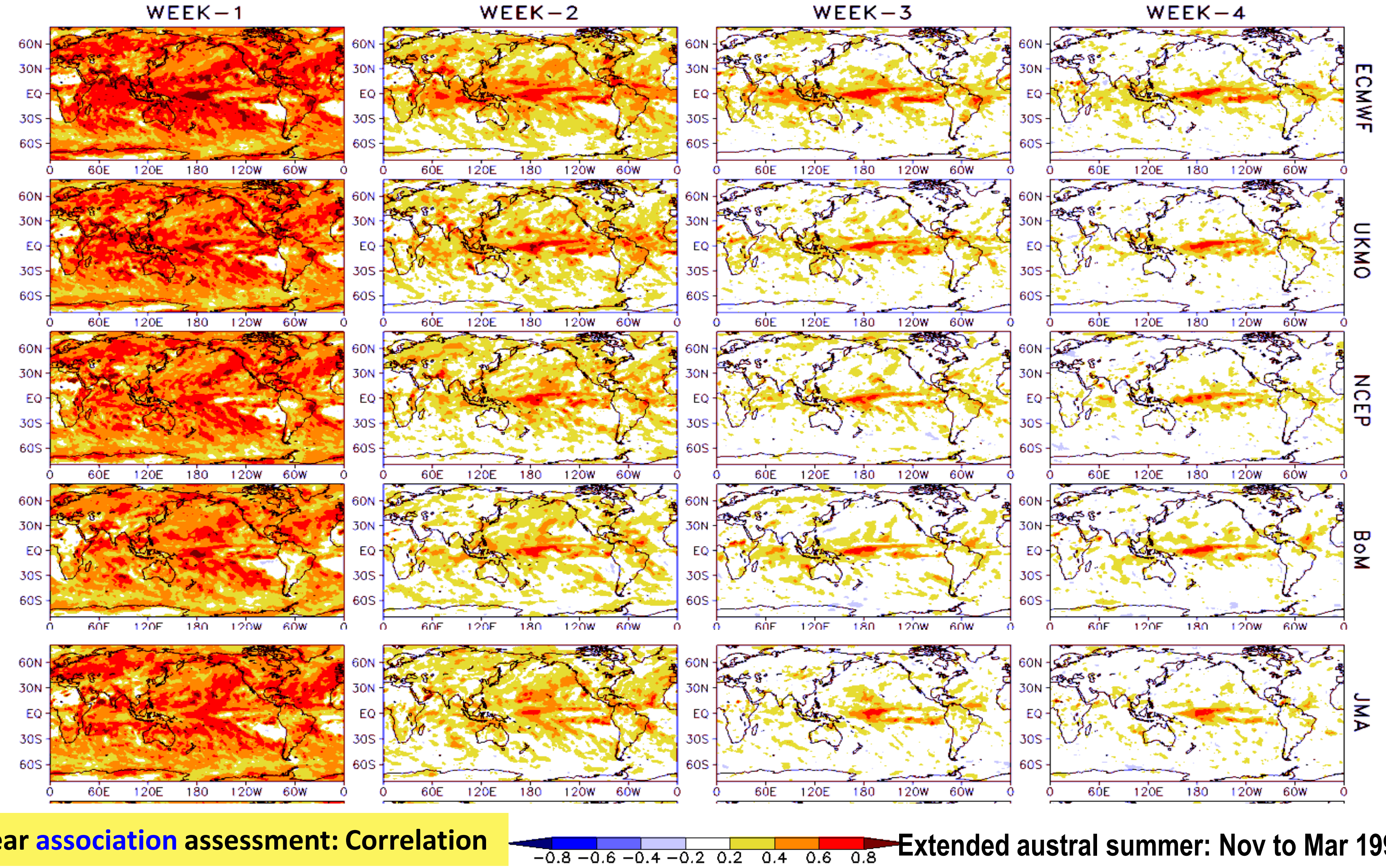
• Published paper proposing a verification framework for South American sub-seasonal precipitation predictions (December 2018).

• Published paper on the global precipitation hindcast quality assessment of all S2S project models (May 2019).

• Attended the Workshop on "Predictability, dynamics and applications research using the TIGGE and S2S ensembles" at ECMWF with PDEF and chaired working group discussion session on verification/calibration (April 2019).

• Provided recommendations on verification metrics for use in the 2nd Phase of the WGNE Aerosol project (Evaluating the impact of aerosols on Numerical Weather and Subseasonal Prediction), a joint collaboration involving WGNE, S2S and GAW.

How well in phase are sub-seasonal precip. predicted anomalies with the corresponding observations?



Linear association assessment: Correlation

Extended austral summer: Nov to Mar 1999-2009

Felipe M. de Andrade, Caio A. S. Coelho, Iracema F. A. Cavalcanti, 2019: Global precipitation hindcast quality assessment of the Subseasonal to Seasonal (S2S) prediction project models. Climate Dynamics

Plans for Research to operations (R2O) and S2S forecast and verification products development

Caio Coelho

- Recommend in collaboration with the Inter-Programme Expert Team on Operational Prediction on Sub-seasonal to Longer time scales (IPET-OPSLS) the verification scores to be computed by centres running operational sub-seasonal prediction models

- Disseminate via a wiki page the work performed by the S2S research community on calibration, multi-model combination, verification and forecast products generation, including software tools, web portals and publications

DOCUMENT / PUBLICATION

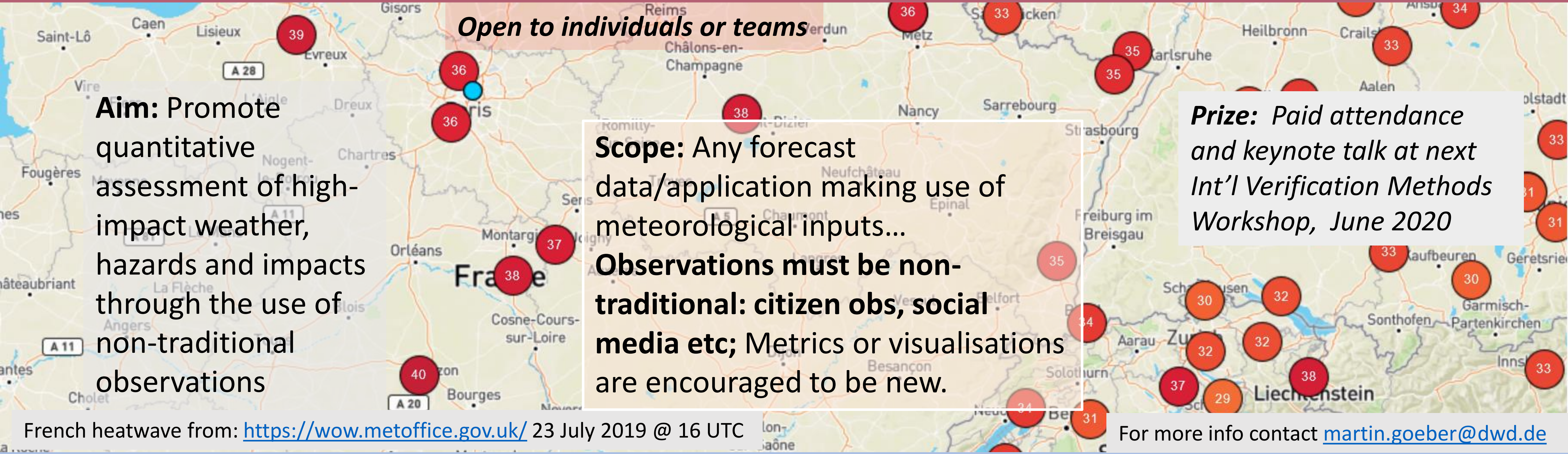
A value chain approach to optimising early warning systems

SOURCE(S): UNITED NATIONS OFFICE FOR DISASTER RISK REDUCTION (UNDRR)

HIWeather

Brian Golding, Beth Ebert, Marion Mittermaier, Anna Scolobig, Shannon Panchuk, Claire Ross and David Johnston, 2019: A value chain approach to optimising early warning systems. Contributing Paper to GAR 2019.

2nd Challenge to develop and demonstrate the best new forecast verification metric *using non-traditional observations*



Open to individuals or teams

Aim: Promote quantitative assessment of high-impact weather, hazards and impacts through the use of non-traditional observations

Scope: Any forecast data/application making use of meteorological inputs… **Observations must be non-traditional: citizen obs, social media etc;** Metrics or visualisations are encouraged to be new.

Prize: *Paid attendance and keynote talk at next Int'l Verification Methods Workshop, June 2020*

French heatwave from: https://wow.metoffice.gov.uk/ 23 July 2019 @ 16 UTC

For more info contact martin.goeber@dwd.de

Timeline :

- Launch, EMS, Copenhagen, September 2019
- Deadline for entries : 15 February 2020
- Announcement of winner : end March 2020

JWGFVR HIWeather

Run by WMO Joint Working Group on Forecast Verification Research in support of WWRP HiWeather Project

WG involvement in WMO projects/activities

- SWFDPs – Provide verification training in Nairobi for HIGHWAY (January) and at Africa SWIFT summer school in Ghana (July)

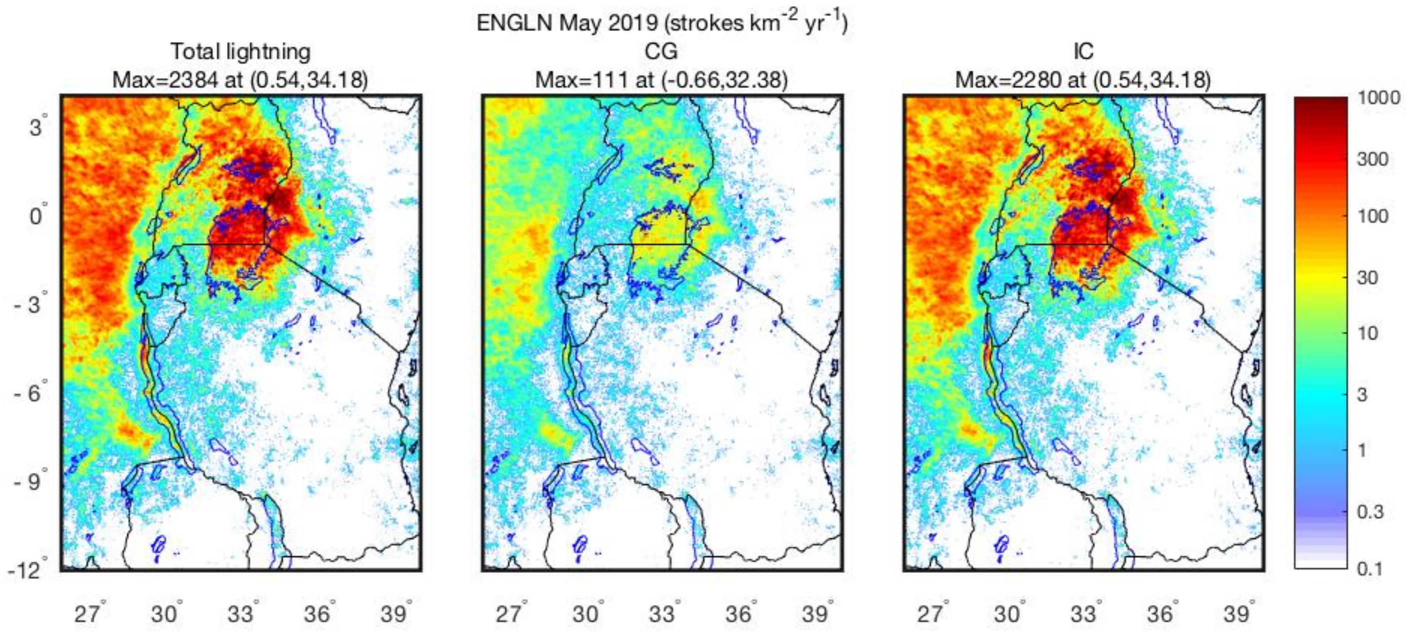
- AvRDP – Continue to contribute to project's verification needs, attending meetings and contribute to reports. Presented at recent meeting (remotely) in August (Pretoria, South Africa).

JWGFVR

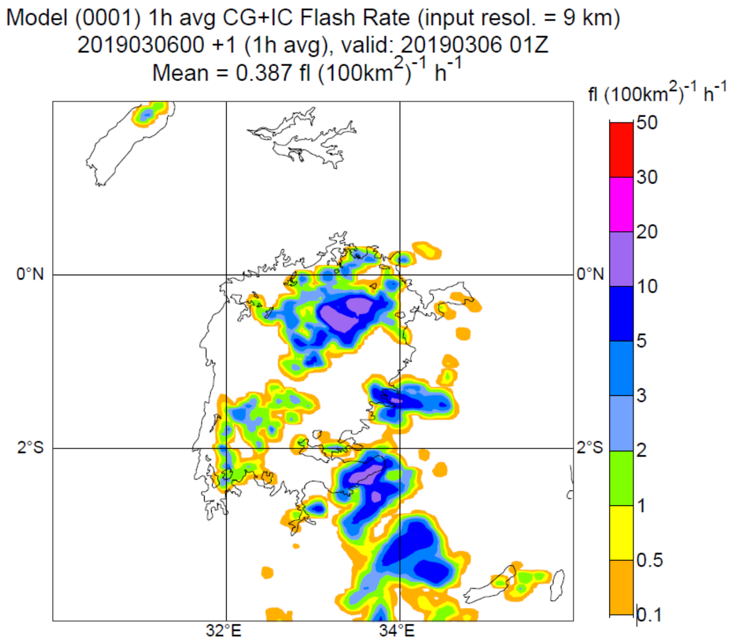
WG involvement in WMO projects/activities

- HIGHWAY/ L. Victoria – Funded by UK DfID. Focus is now on the evaluation of the lightning diagnostic and any warnings verified against lightning and other non-standard observations. Project concludes in March 2020.

Monthly lightning stats



EC forecast



Courtesy Steve Goodman, NASA

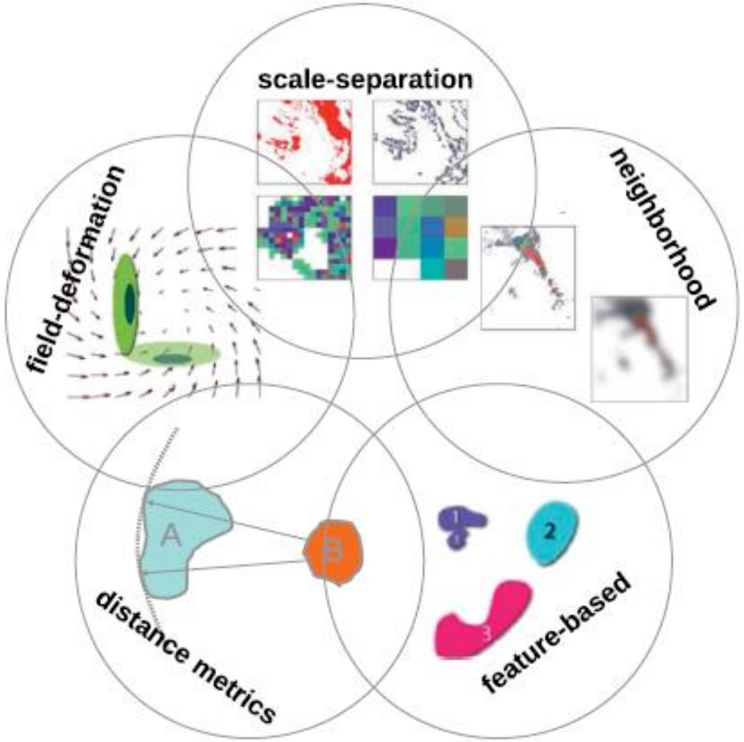
THE SETUP OF THE MesoVICT PROJECT

https://doi.org/10.1175/BAMS-D-17-0164.1
Final Form: 28 March 2018
Published online: 9 October 2018

7 publications so far.
Anticipate ~10 publications in total.

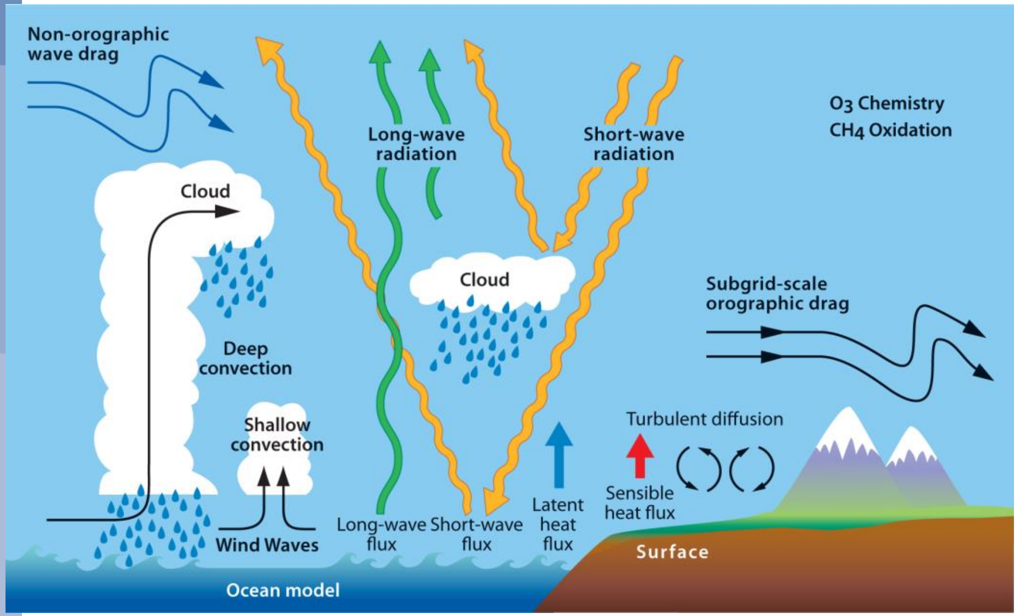
Manfred Dorninger, Eric Gilleland, Barbara Casati, Marion P. Mittermaier, Elizabeth E. Ebert, Barbara G. Brown, and Laurence J. Wilson

13th EMS/11th ECAM launch	Reading, UK	9-13 Sep 2013
1st MesoVICT workshop (kick-off)	Vienna, Austria	2-3 Oct 2014
15th EMS/12th ECAM	Sofia, Bulgaria	7-11 Sep 2015
16th EMS/11th ECAC	Trieste, Italy	12-16 Sep 2016
2nd MesoVICT workshop	Bologna, Italy	21-23 Sep 2016
7th International Verification Methods Workshop	Berlin, Germany	3-11 May 2017
EMS annual meeting	Budapest, Hungary	3-7 Sep 2018
MesoVICT final workshop	Vienna, Austria	8-9 July 2019



Final workshop held at University of Vienna, Austria, https://mesovict.univie.ac.at

Process-oriented verification document delivered



Process-oriented verification

Thomas Haiden, Barbara Casati, Caio Coelho, Eric Gilleland, Raghavendra Ashrit, and Manfred Dorninger

Joint Working Group for Forecast Verification Research, 28 Feb 2019

1. What is process-oriented verification?

Verification of forecasts from numerical weather prediction (NWP) models serves a wide range of purposes, including modelling-oriented forecast evaluation, performance monitoring for documentation and management, and user-oriented evaluation of final products. The main goal of modelling-oriented verification is to better understand, and ultimately reduce, model and data assimilation issues and errors in order to improve forecast quality. To achieve this goal, the verification methodology needs to be designed such that it allows identifying the role of specific model processes in the occurrence of forecast errors. Because this approach can include a range of methodologies we define process-oriented verification here not in terms of specific techniques (although examples are given below) but rather by its overall objective of improving process understanding. While this kind of verification has always been an integral part of NWP model development, some of it could be adopted to become part of operational NWP verification suites. Apart from generating verification results that can be acted upon more directly by model developers, it would provide additional insights for forecast users. Ideally, it would contribute to a more efficient research-to-operations and operations-to-research cycle in NWP, both within and between NWP centres. This report discusses methodologies that are already being used, but could perhaps be used more widely and systematically, in process-oriented verification of outputs from various NWP centres.

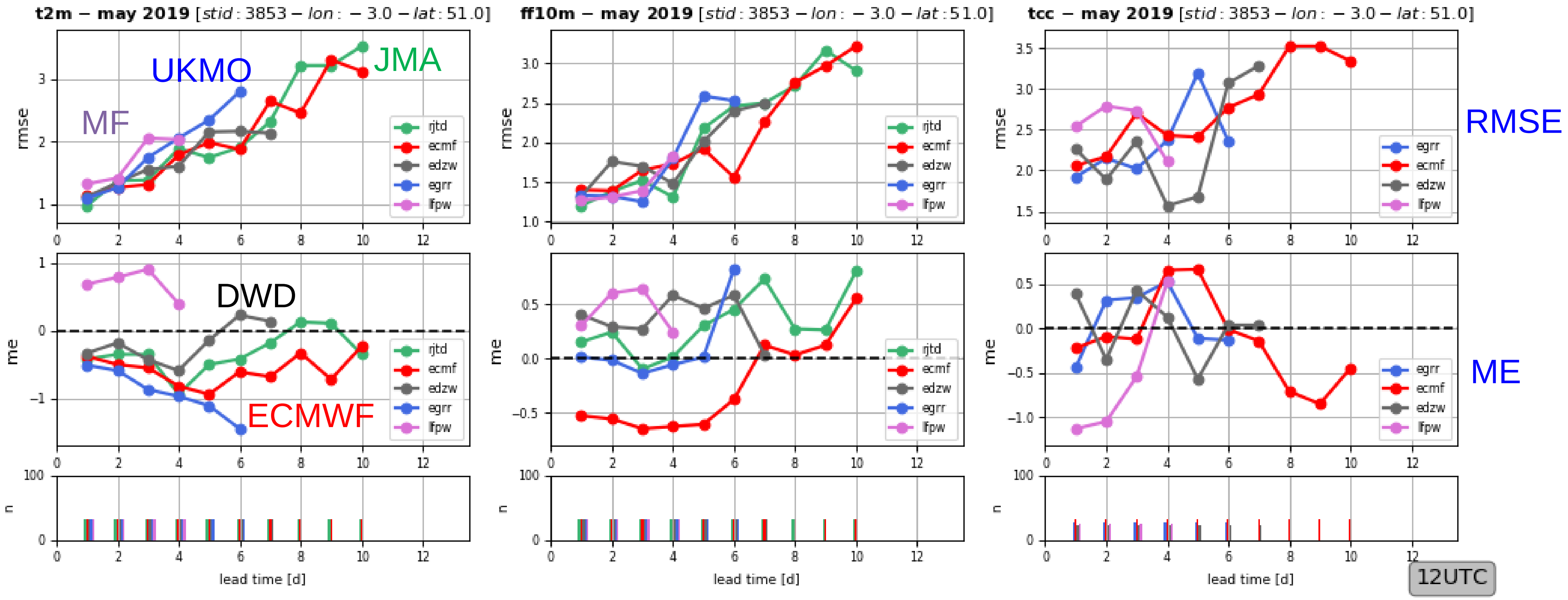
Two documents in preparation…

1. Led by Marion Mittermaier: “*How to “do” precipitation verification across space and time scales: A review of common challenges and potential solutions*” – provisional title
2. Led by Chiara Marsigli: “*Observations for high-impact weather and their use in verification*” - provisional title *(focus on thunderstorms and fog)*

WMO CBS exchange of surface scores

Participating centres increasing, with site-by-site variations!

2m temperature | 10m wind speed | Total cloud cover

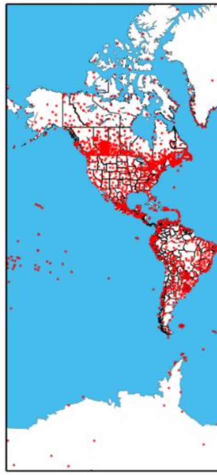


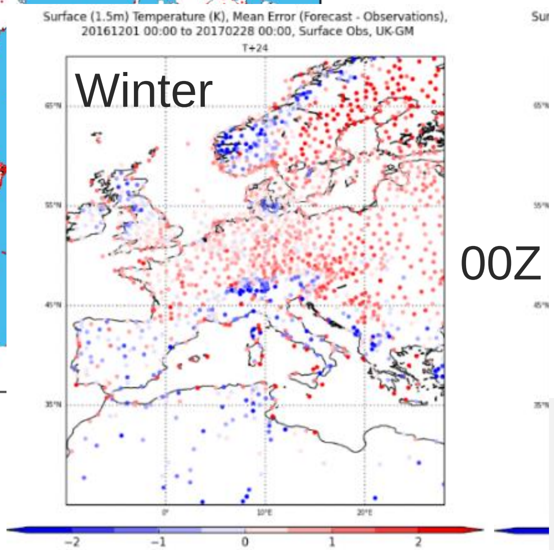
Scores for individual SYNOP stations from different models

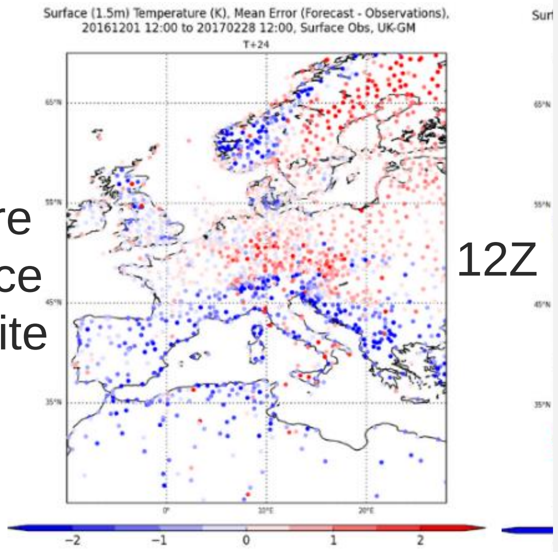
Aggregating temperature scores

Csima and Mittermaier, 2019

All the stations

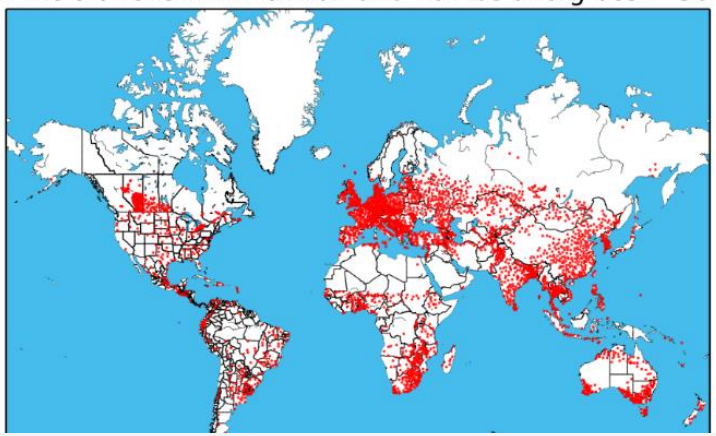






Local temperature performance may be quite different.

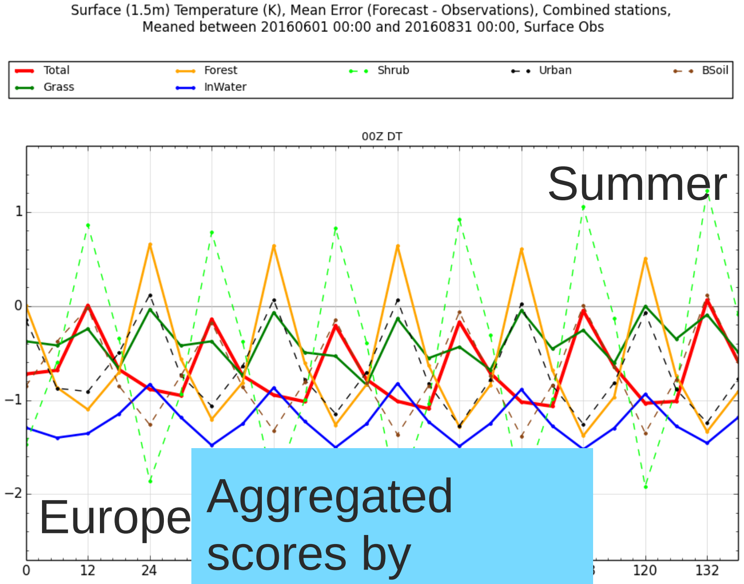
All the stations with not-nan and not-ice and grass > 50%



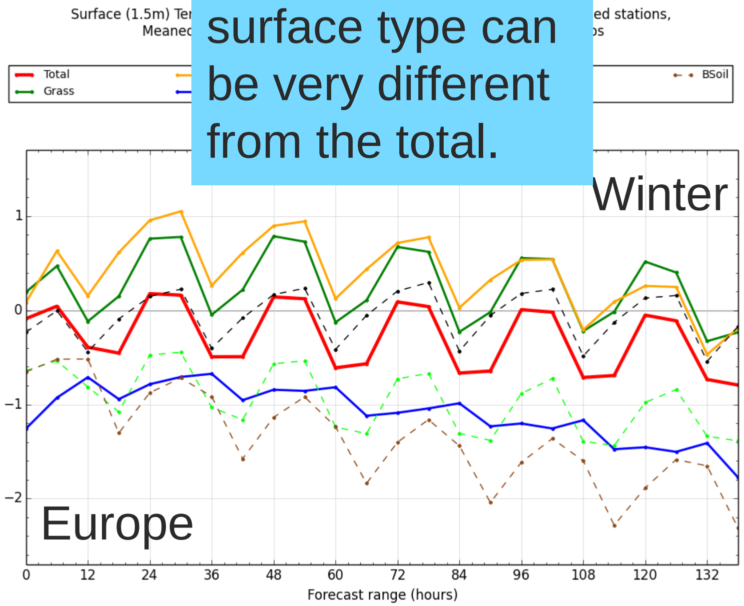
Observing sites are expected to be grass enclosures, unless this isn't possible (rock, sand, snow, ice).
This may also not be the case during the cold season in many mid-latitude locations (snow, ice).

Diagnosed forecast 1.5m T is a weighted average of temperatures for different land surface types.

We have 9 (sub-)tiles

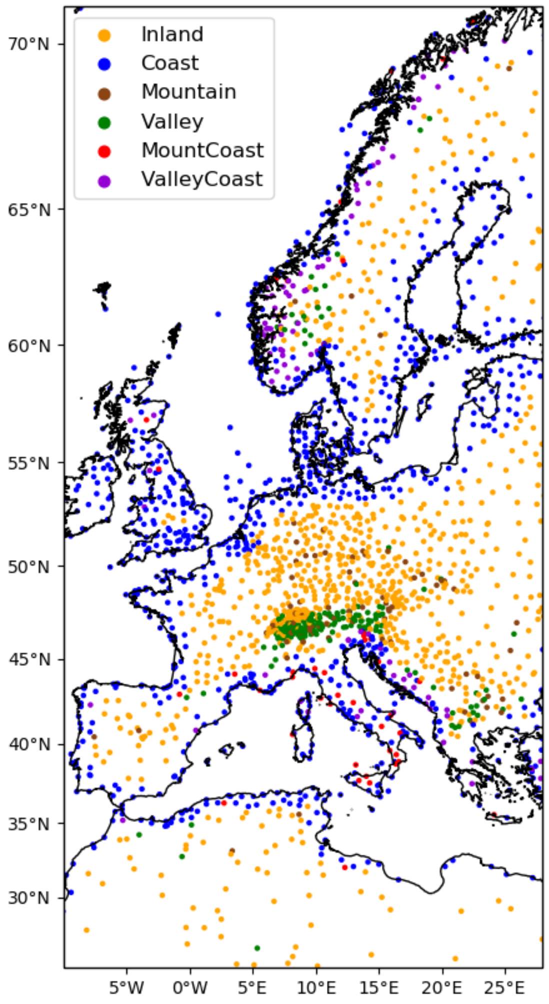


Aggregated scores by dominant land surface type can be very different from the total.

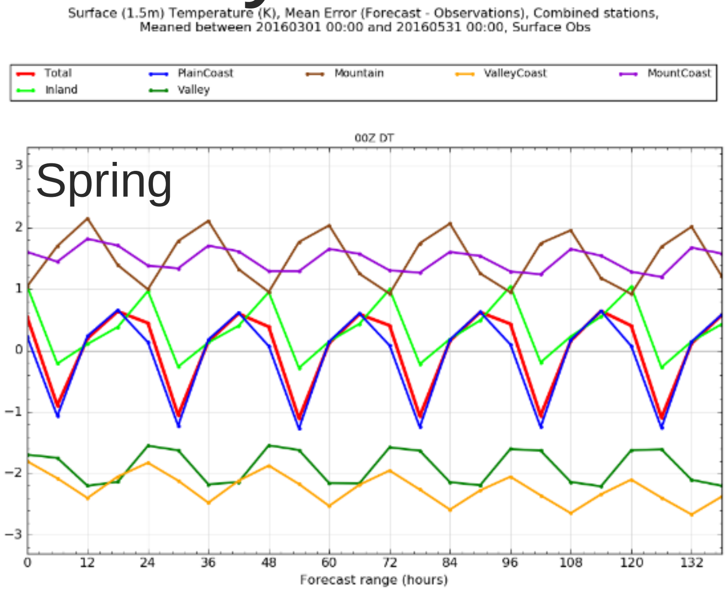


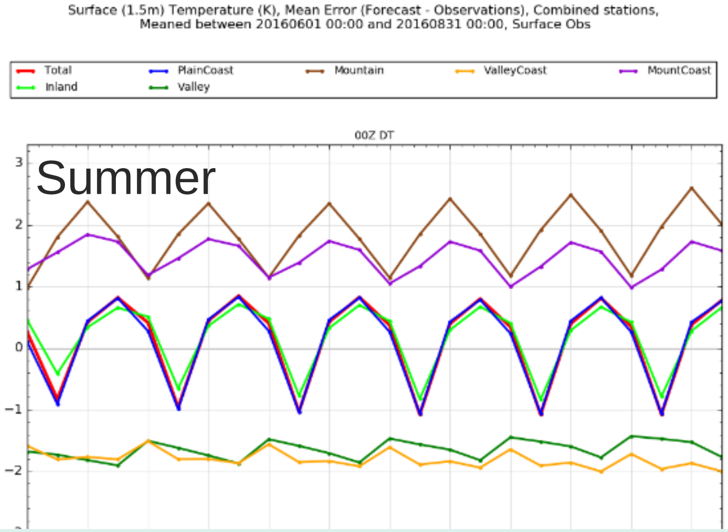
Stratification by location

Csima and Mittermaier, 2019

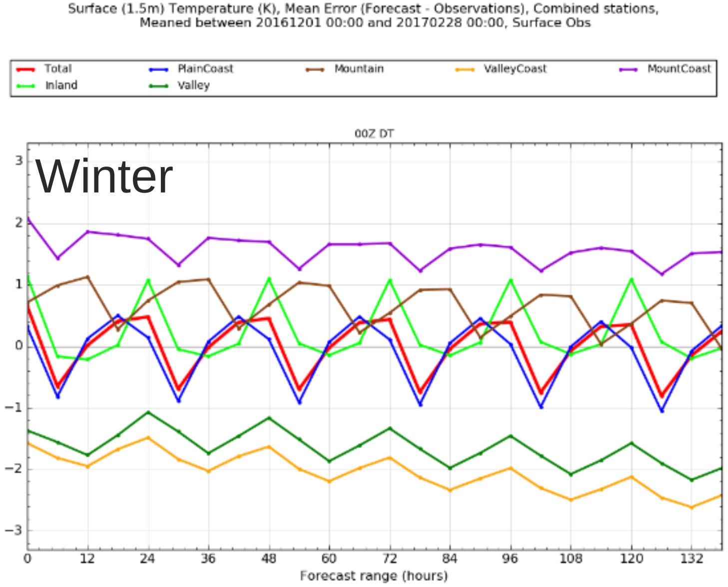


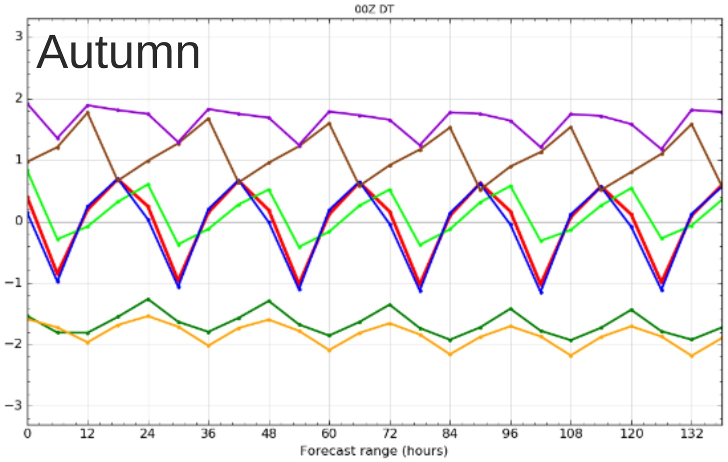
Inland	Coast	Mountain	Valley	Mount Coast	Valley Coast
4282	3424	156	528	98	239





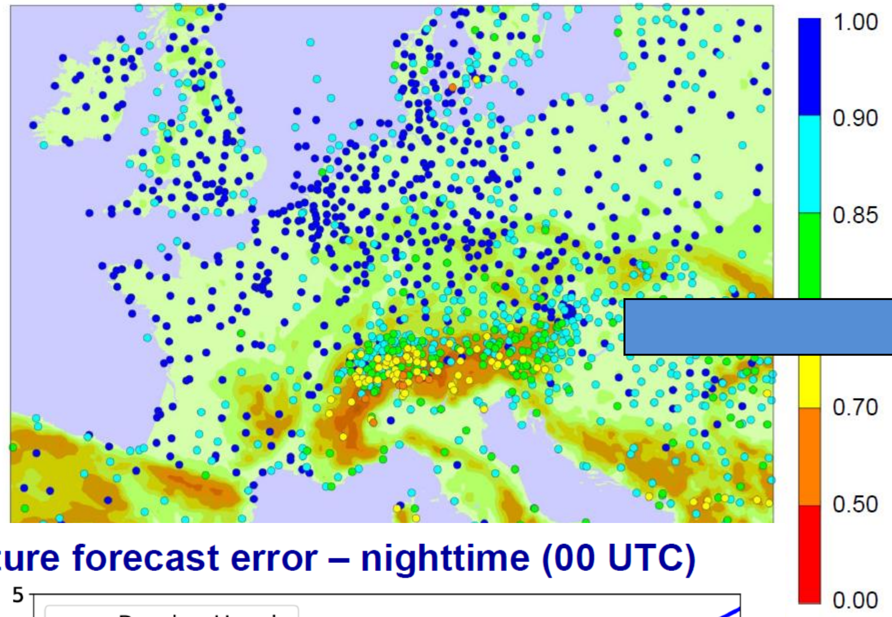
Flat coastal and inland sites dominate the total bias (they are the most numerous).



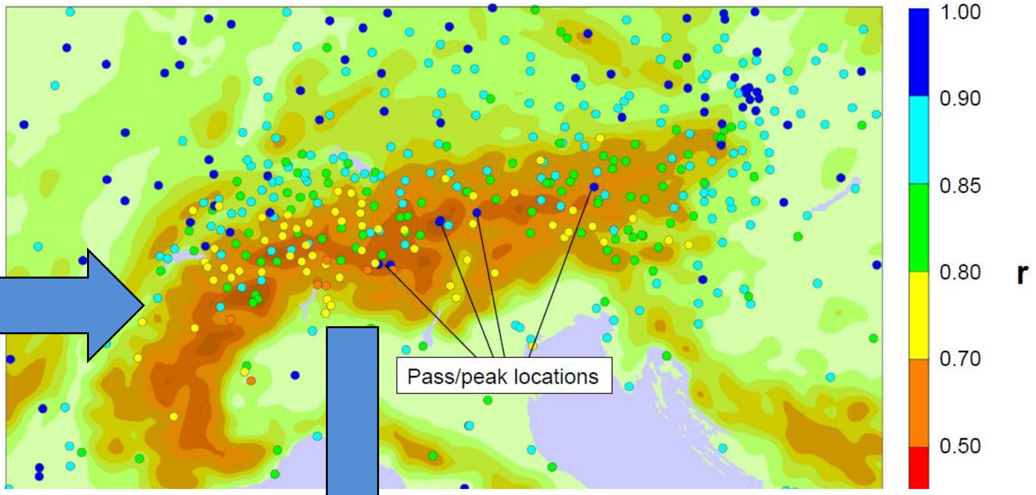


Forecast error growth in complex terrain (Haiden and Wedi, ICAM 2019)

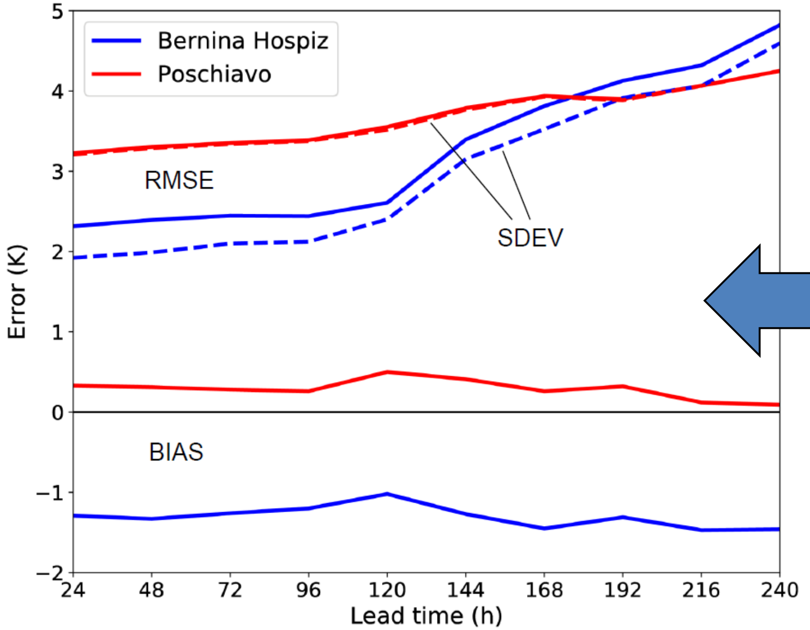
2m temperature forecast skill at day 3 – nighttime (00 UTC)



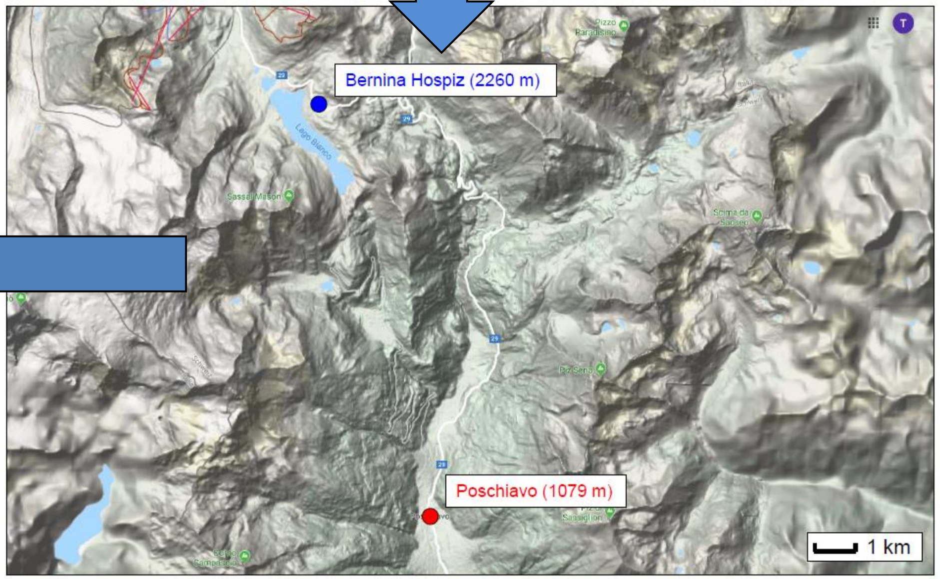
m temperature forecast skill at day 3 – nighttime (00 UTC)



2m temperature forecast error – nighttime (00 UTC)



Mountain and valley ation: example



Model topography

Bernina Hospiz: 2699 m

Poschiavo: 2344 m

Verifying precipitation globally

Up to now we tend to verify:

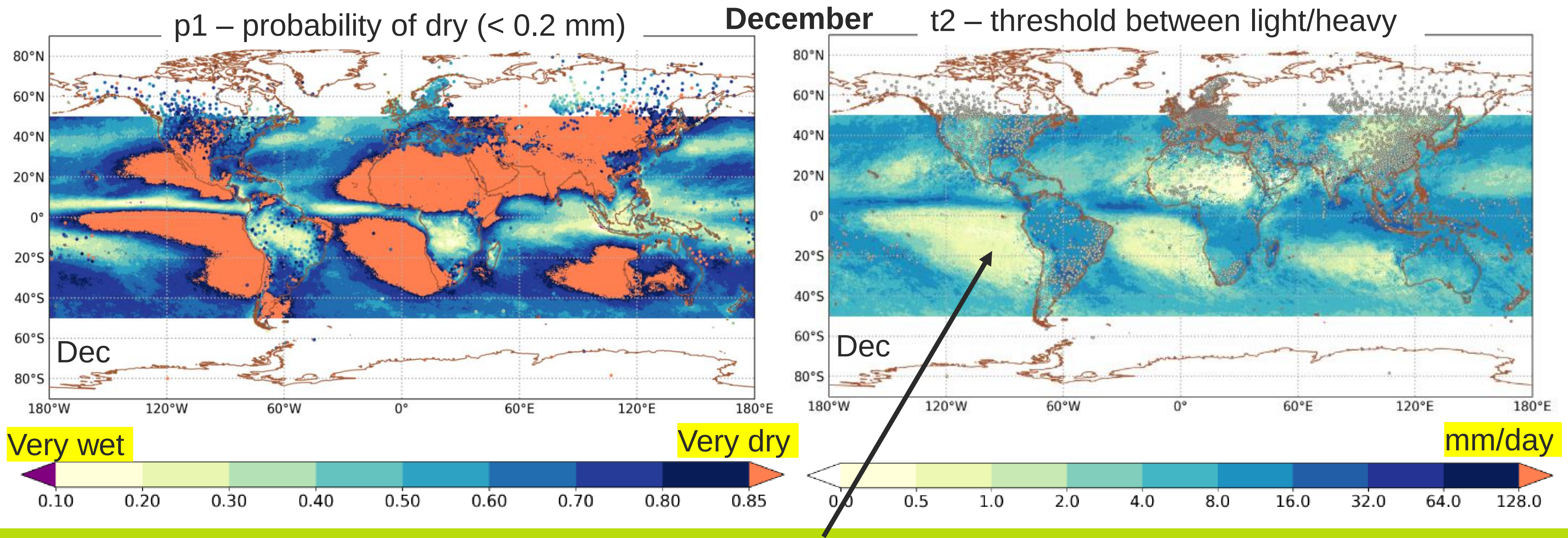
- Over land only
- Gauge based
- Aggregation – huge variations in climate, danger of false skill.

- SEEPS introduced the idea of using a climatology assessing the *local* performance (nearest model grid point) of the forecast providing a safe way of aggregating many (climatologically different) locations BUT variations in gauge observations between centres prevented us from recommending SEEPS for CBS.
- Here,
  - we create a satellite-based (TRMM) climatology 1998-2015 to assess “all” model grid points
  - Use SEEPS in conjunction with this climatology and GPM IMERG to calculate scores
- The aim is to:
  - Check how well the gauges are sampling the performance
  - Use a land-sea split to explore variations in performance (not been possible before)
  - Compare the gauge and satellite-based scores

North, Mittermaier and Milton, 2019: Using SEEPS with a TRMM-derived climatology to assess global NWP precipitation forecast skill. *In preparation/internal review.*

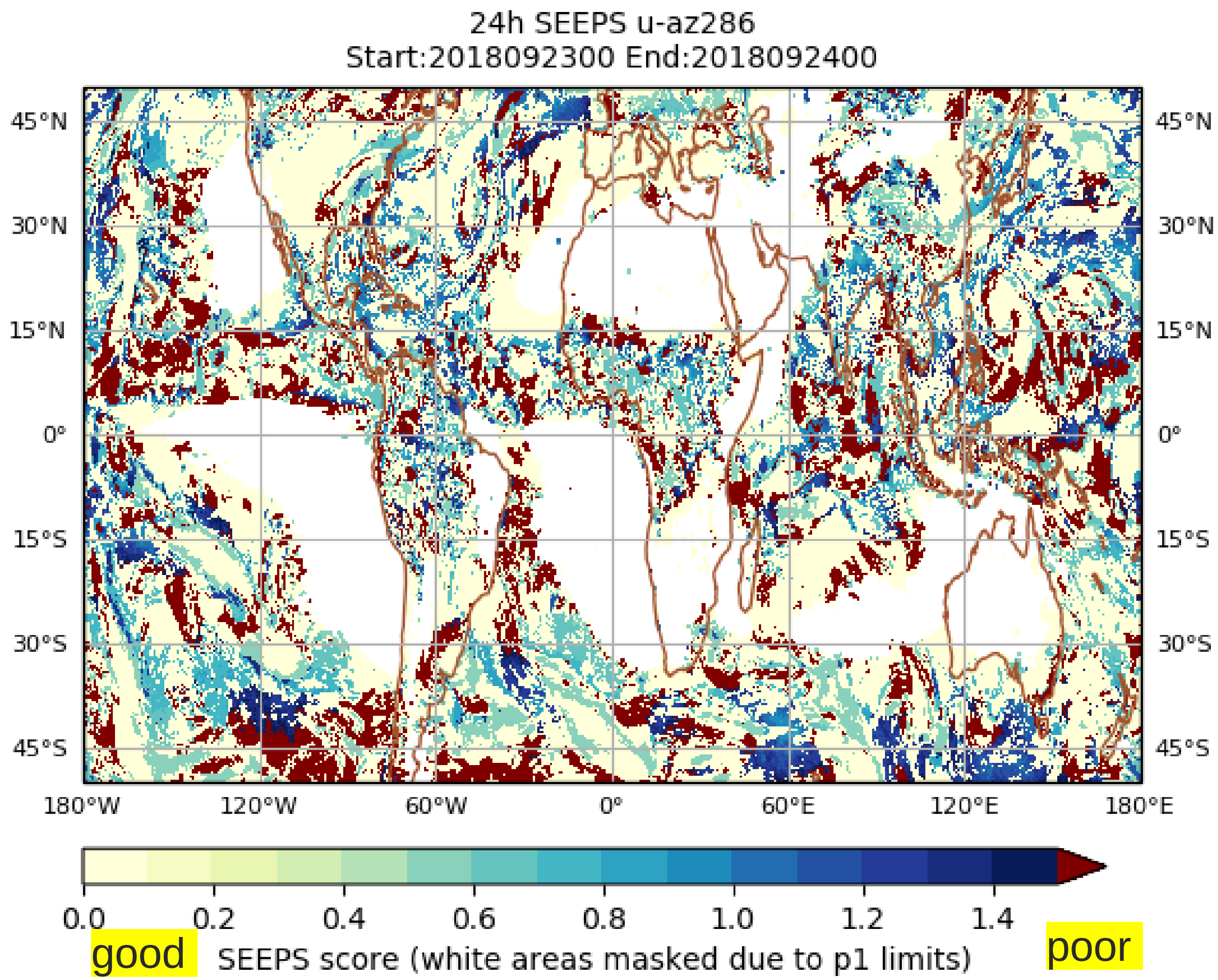
The concept behind SEEPS

- Stable Equitable Error in Probability Space (SEEPS)
- A verification metric that was designed for monitoring model precipitation skill using a climatology derived from rain gauges to provide a climatologically "aware" assessment
- See Rodwell et al. (2010), Haiden et al. (2012) for details.



North, Mittermaier and Milton, 2019

Daily SEEPS



Example day 3 SEEPS

= 0 is perfect
> 1 considered poor

Errors follow synoptic features/systems

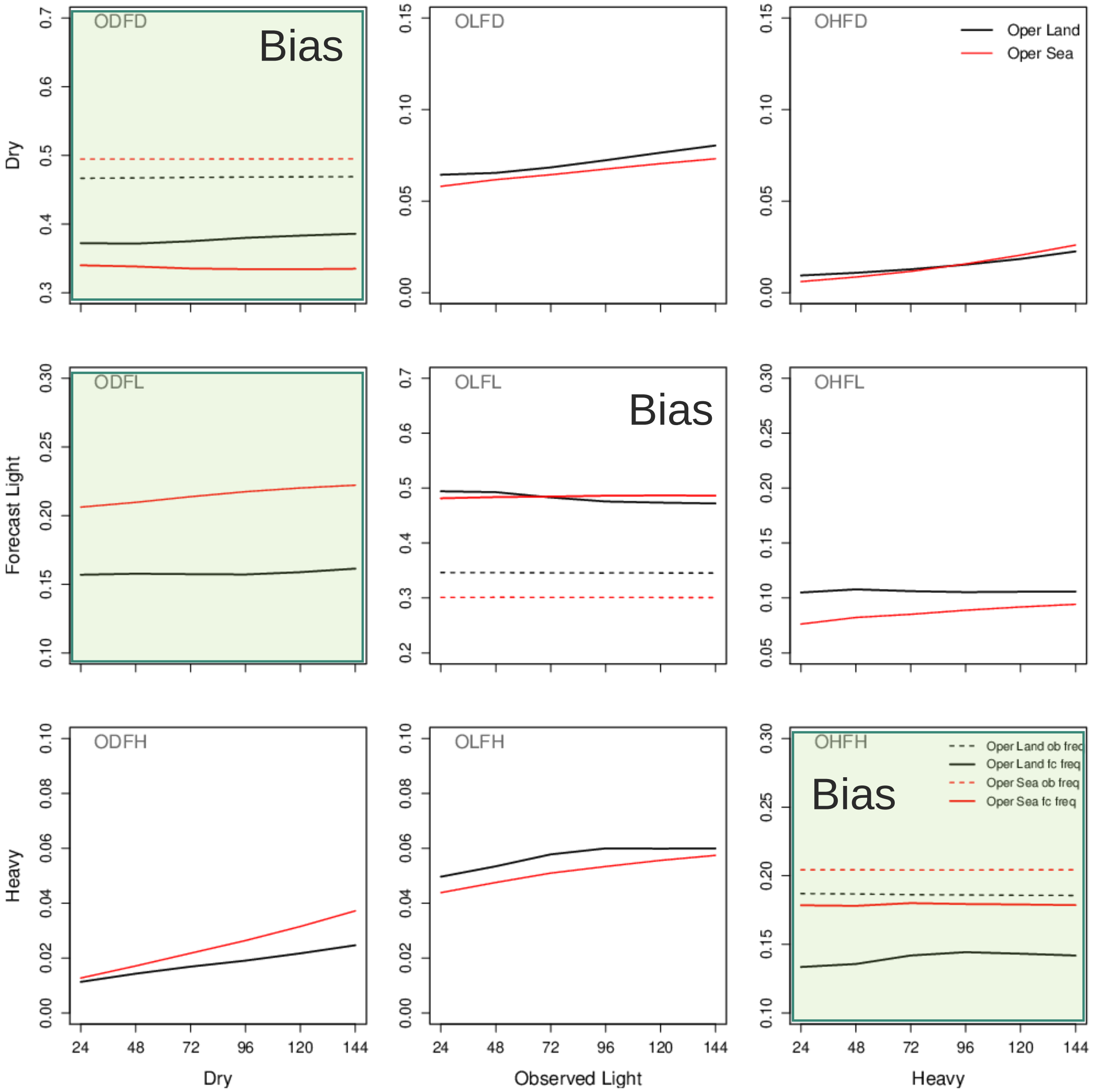
Regions of “gross errors”

Met Office Land-sea split

- Consider impact of model changes on water cycle/budget
- Slightly different biases: sea has a bigger bias for dry and heavy categories. Reverse for light.
- Main error contribution differences arise from “observed dry forecast dry” (ODFD) (land more), “observed dry forecast light” (ODFL) and “observed heavy forecast heavy” (OHFH) where the sea areas contribute far more
- This will be valuable for comparing the operational model with future upgrades

North, Mittermaier and Milton, 2019

Decomposition of global SEEPS aggregate



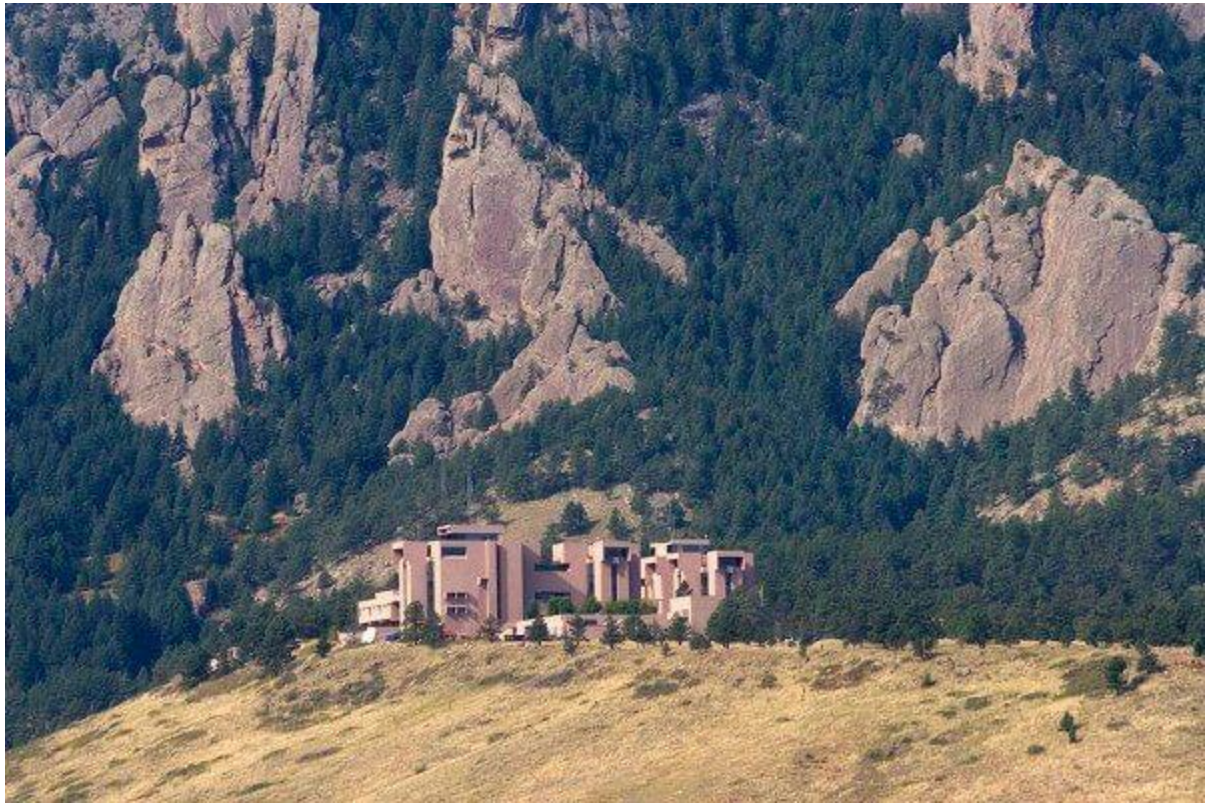
8th International verification methods workshop and tutorial?

June 2020

First announcement pending….. We hope.

If not….

joint annual meeting with WGNE in Boulder?



Thank you for your attention!

WG membership

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

Two vacancies – DA and climate evaluation (targeting earlier career scientists)