Joint Working Group on Forecast Verification Research report

Caio Coelho
CPTEC/INPE

Marion Mittermaier
Met Office

Environment and Climate Change Canada
Environnement et Changement climatique Canada

WORLD METEOROLOGICAL ORGANIZATION

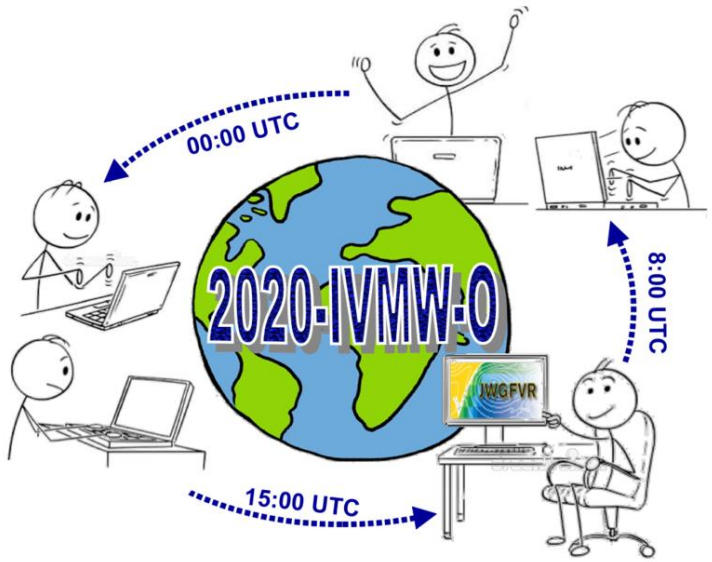
November 2020 around-the-clock International Verification Methods Workshop Online (2020-IVMW-O): Weeks of 9-13 and 16-20 Nov 2020

Chaired by Barbara Casati (ECCC)

https://jwgfvr.univie.ac.at

- 20 sessions
- 325 registrations
- 16 Keynote talks (30+10 min)
- 59 Oral presentations (15+5 min)
- 18 Poster presentations (5+5 min)

Total: 93 presenters !



The **November 2020 around-the-clock International Verification Methods Workshop Online (2020-IVMW-O)** is an outreach activity of the WWRP/WGNE Joint Working Group on Forecast Verification Research (JWGFVR). The workshop spans two weeks, from the 9th to the 13th and from the 16th to the 20th of November 2020, and consists in 2-hour online daily sessions, with live-stream presentations and discussion. This virtual event intends to fill the gap between the 7th and 8th International Verification Methods Workshop, since the latter has been postponed due to the COVID-19 pandemic.

The goal of the workshop is to **discuss recent aspects of verification research and keep the research community up-dated on new verification practices**, as applied to different types of weather forecasts and environmental predictions, on all spatial and temporal scales, from weather and sub-seasonal to seasonal and decadal, as well as for long climate projections. Participants are welcome from operational, research and forecast-user communities.

Workshop focus is on verification methodologies

Contributions from the World Meteorological Organization (WMO) endorsed High Impact Weather (HIW), Polar Prediction Project (PPP) and Sub-seasonal to Seasonal prediction project (S2S) are particularly encouraged.

Research themes of IVMW-O 2020 sessions:

- Error Characterization (tracking)
- Data Assimilation Techniques in Verification
- Representativeness and observation uncertainty (to what extent measurements at a single location are representative of averages over a larger area)
- Metaverification (properties of verification methods)
- Spatial Methods (incl. neighbourhood)
- Operational practices (incl. software/tools and visualization)
- Physical Process Diagnostics
- Ocean
- Sea-ice (Polar prediction)
- Sub-seasonal to Seasonal and Decadal Predictions + climate projections
- High Impact Weather (incl. user value chain)

2-hour sessions
Including discussions

JWGFVR

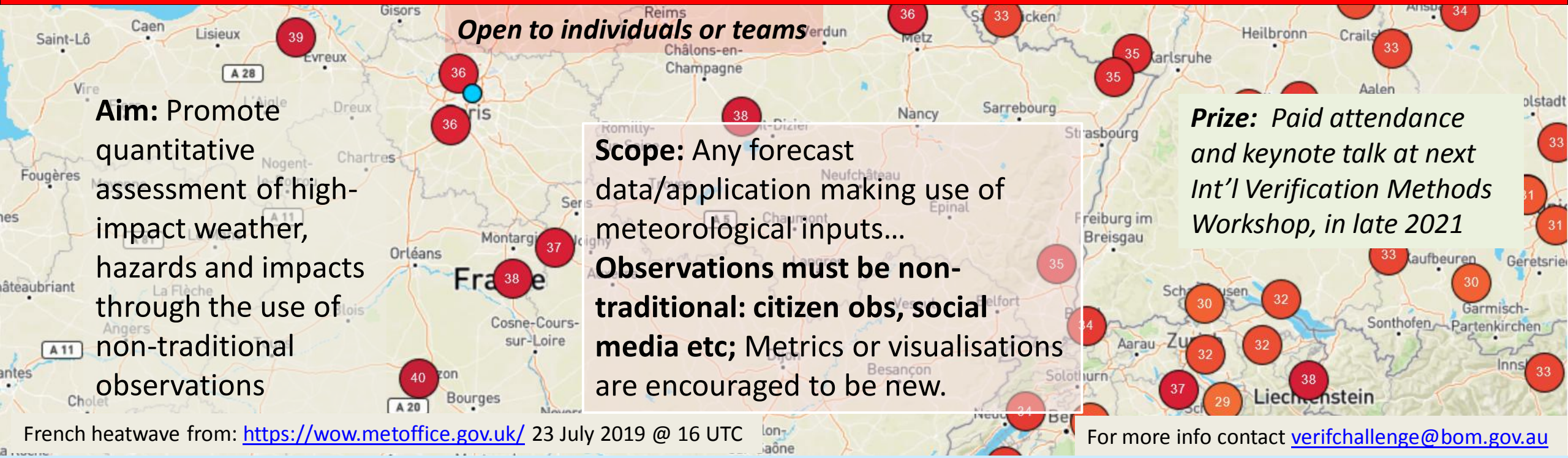
2nd Challenge to develop and demonstrate the best new forecast verification metric *using non-traditional observations*

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, in late 2021*



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

For more info contact verifchallenge@bom.gov.au

Timeline :

- Deadline for entries : 31 April 2021
- Announcement of winner : 30 June 2021
- Entry form available at WMO website: **https://community.wmo.int/news/2nd-international-verification-challenge**

JWGFVR HIWeather

Run by WMO Joint Working Group on Forecast Verification Research in support of WWRP HiWeather, S2S and PPP Projects

Document in preparation…

Review Led by Chiara Marsigli: "*Observations for high-impact weather and their use in verification*" *(focus on thunderstorms and fog)*

To be promoted as keynote talk at 2020 IVMW-O (Nov 2020)

For verifying forecasts of high-impact weather "traditional" observations often do not permit characterization of the phenomenon of interest

New observations or more generic quantities can be considered reference data (or proxies) for the verification: remote sensing datasets, datasets derived from telecommunication systems including cell phones, data collected from citizens, reports of impacts and claim/damage reports from insurance companies

Describes what is needed to transform these different sources of information about high-impact weather phenomena into objective data to be used for performing a statistical verification

- Contributes to the three WWRP societal challenges: high-impact weather, evolving technologies and urbanization
- Extends JWGFVR observation-focussed activities by stimulating the research community to explore the use of new sensor networks and other evolving technologies in urban settings: helps further address urb. and evol. tech. chall.

JWGFVR

Research to Operations and S2S Forecast and Verification products developments activities

Caio Coelho

• Created a wiki page for promoting and disseminating the work performed by S2S research community on calibration, multi-model combination, verification and forecast products generation, software tools, web portals and publications

• Developed in collab. with IPET-OPSLS (now ET-OCPS) proposed text with:

- designation criteria for Global Producing Centres for Subseasonal Forecasts (GPCs-SSF) and Lead Center(s) for Sub-seasonal Forecast Multi-Model Ensemble (LCs-SSFMME) within the WMO operational infrastructure
- recommended verification scores for sub-seasonal forecasts to be computed and disseminated by GPCs-SSF and LCs-SSFMME

Status: Tabled at EC-72 this year, which recommended to be tabled at EC-73 after a review by the first meeting of the new Infrastructure Commission (INFCOM-1)

• Provided recommendations to US Weather Research Science Working Group (WRSWG-S2S) to help develop research agenda based on requirements/gaps: verification and products

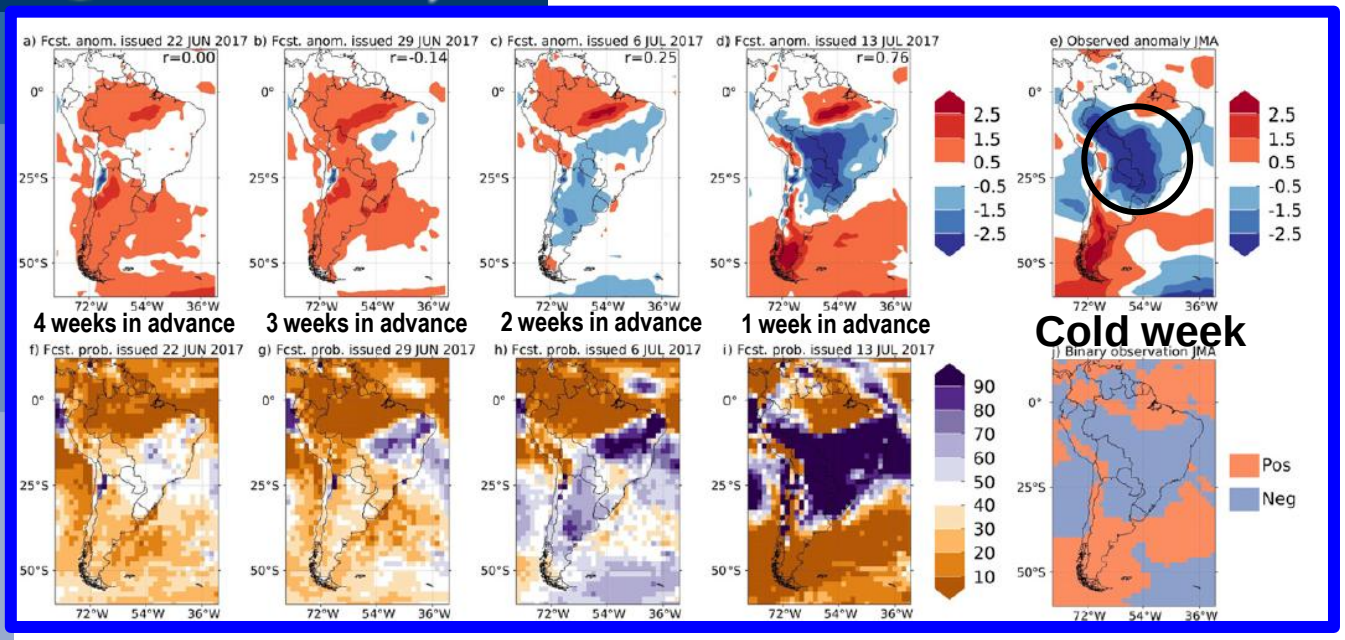
• Produced papers assessing quality of sub-seasonal predictions produced by S2S project models and new Brazilian model

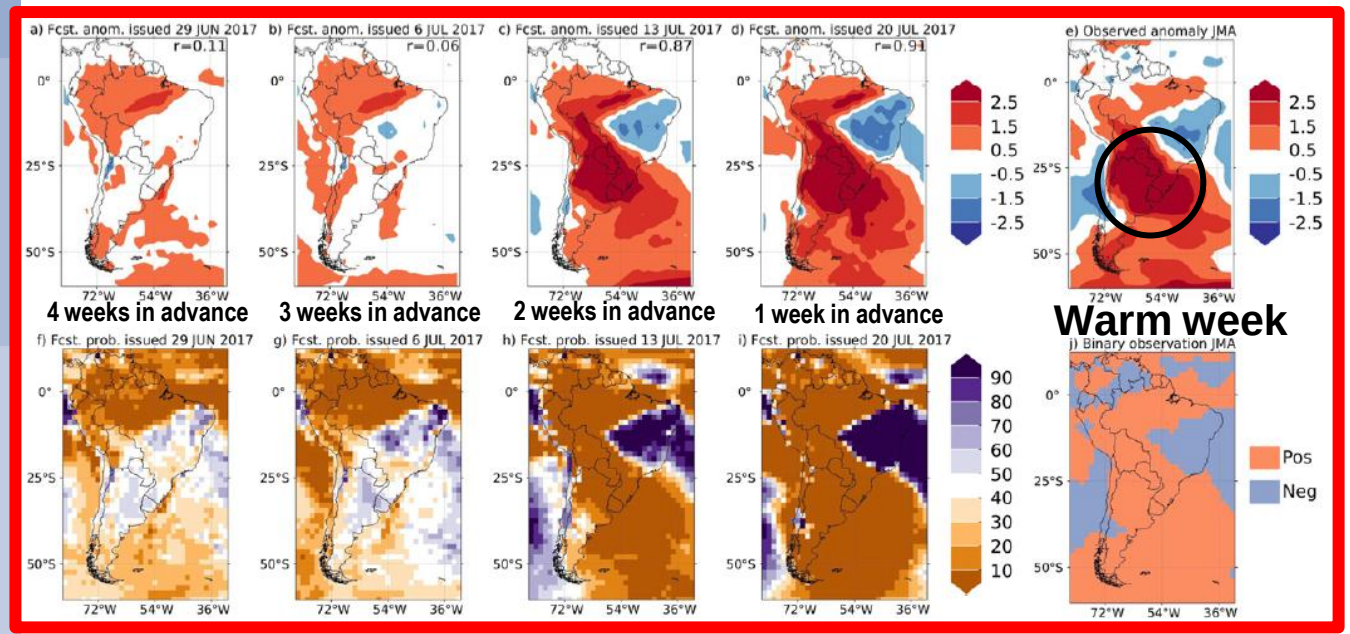
• Contributed article to S2S project newsletter

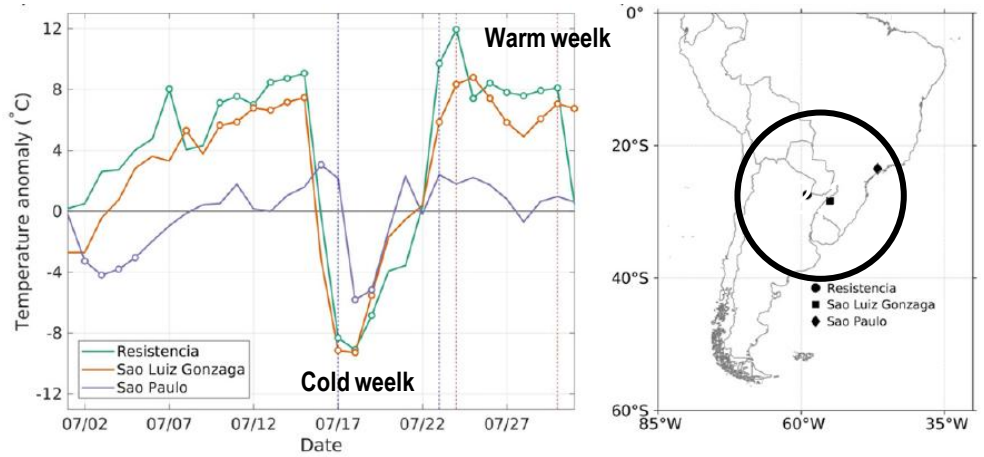
Research to Operations and S2S Forecast and Verification products developments activities

Mariano S. Alvarez, Caio A. S. Coelho, Marisol Osman, Mári A. F. Firpo, Carolina S. Vera (2020)

Assessment of ECMWF subseasonal temperature predictions for an anomalously cold week followed by an anomalously warm week in central and southeastern South America during July 2017. Weather and Forecasting. 25. 1871-1889.







- Applied verification framework of Coelho et al (2018) using 20 years of hindcasts and 3 years of real time predictions
- Cold week well predicted one week in advance
- Warm week well predicted two weeks in advance
- Consistent with regions where model showed good past performance

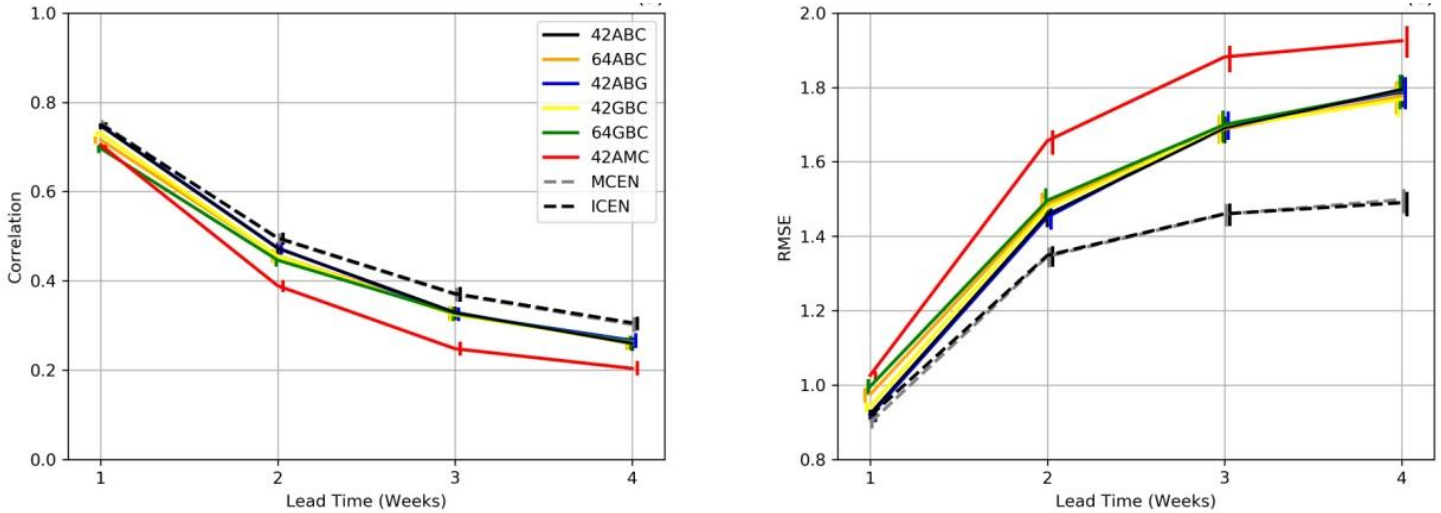
Coelho, C. A. S.; M. A. F Firpo. F. M. de Andrade (2018) A verification framework for South American sub-seasonal precipitation predictions. METEOROLOGISCHE ZEITSCHRIFT, v. 27, n. 6, p. 503-520, 2018.

Contributes to high-impact weather and urbanization challenges

Research to Operations and S2S Forecast and Verification products developments activities

Bruno S. Guimarães, Caio A. S. Coelho, Steve J. Woolnough, Paulo Y. Kubota, Carlos F. Bastarz, Silvio N. Figueroa, José P. Bonatti and Dayana C. de Souza (2020)

Configuration and hindcast quality assessment of a Brazilian global sub-seasonal prediction system. QJRMS, 146, Issue728, Part A, 1067-1084.



- Performance of weekly temperature anomalies averaged between 60ºN e 60ºS of several configurations (solid lines) of the Brazilian Global Atmospheric Model (BAM-1.2)
- Dashed black line represents the model configuration that presented best performance run in ensemble mode
- Vertical bars: 95% confidence intervals of computed metrics
- Promoting computation of confidence intervals in verification metrics to address sampling uncertainty

Research to Operations and S2S Forecast and Verification products developments activities

Caio Coelho

Wiki contains: Proposed questions to be addressed

http://s2sprediction.net/xwiki/bin/view/dtbs/R2O

The World Weather Research Programme (WWRP) has flagged improving forecasts of precipitation over land as an important area for S2S to focus research and services development efforts. In order to help advance scientific knowledge and the development of forecast and verification products in this priority area this sub-project invites the S2S research and operational communities to address the following questions:

- What is the current performance level of sub-seasonal precipitation forecasts over land? Over which continental regions can these forecasts be best trusted? How performance levels vary through the seasons of the year?
- What is the current capability of S2S models in anticipating the occurrence of extreme precipitation events over land (periods of deficit or excess precipitation)?

High-impact (extreme) events

- How well the main patterns of precipitation variability on the sub-seasonal time scale over various continental regions are represented in S2S prediction models?
- How best to combine and calibrate sub-seasonal precipitation forecasts over land in order to produce improved, combined and well-calibrated products and services?
- Are there identifiable opportunities for producing sub-seasonal precipitation forecasts over land with improved quality? For example, are forecasts produced during Madden and Julian Oscillation (MJO) and/or El Niño Southern Oscillation (ENSO) events more skilful than when neutral conditions are present? Are forecasts for active and break rainfall phases and dry/wet spells (or other quantities of interest) of adequate quality for developing forecast products for use in application sectors?

Promote addressing flow prediction quality: Application in hydrological sector

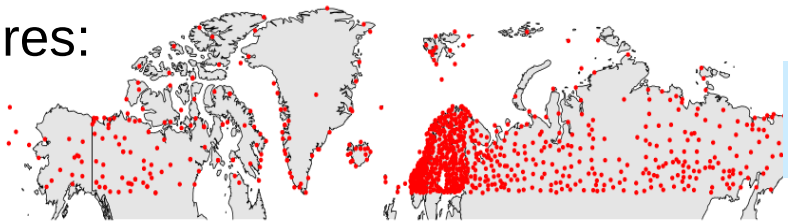
In order to address these questions the research and operational communities are encouraged to explore existing and develop novel methodologies for forecast calibration, combination and verification. Following the S2S verification chapter produced by the JWGFVR for the recent S2S book, it is particularly encouraged the identification of the most relevant forecast quality attributes for the target audiences (e.g. model and forecast developers, and various application sectors) in order to choose appropriate scores and metrics to be able to adequately address clearly and previously defined verification questions of interest. This practice helps performing a thorough assessment of sub-seasonal forecasts from both the probabilistic and deterministic points of view.

Contributes to high-impact weather (extremes), water and urbanization WWRP societal challenges

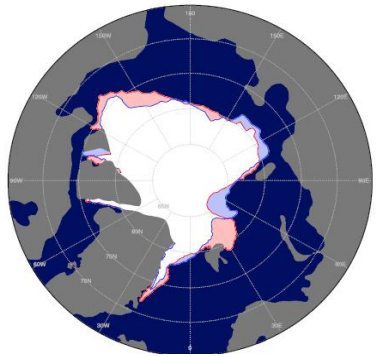
YOPP core phase verification activities

Barbara Casati

1. **Operational summary verification** scores: YOPP provided the framework for analyzing current verification practices in the Polar Regions: revealed issues, investigated solutions, **proposed novel approaches**



Improving operational verification practices



Promoting novel verification techniques: Areal and distance metrics

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.

Better understanding of surface fluxes exchanges and representation of radiation and energy budget

YOPP YEAR OF POLAR PREDICTION

YOPP Operational summary verification scores

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

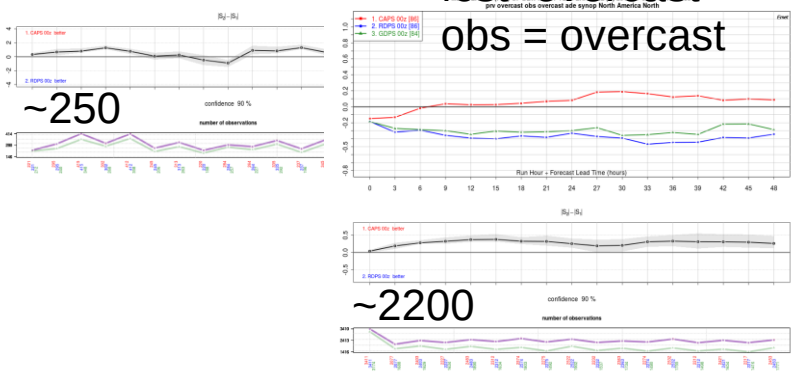
The activities consisted in comparing operational verification practices in the Polar Regions: exchange of objective verification scores during the YOPP Special Observing Periods (SOPs)

Key outcomes include:

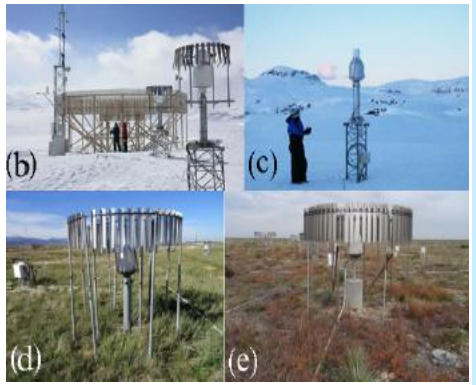
1. Apply (process driven) conditional verification

fcst=clear
obs=clear

fcst=overcast
obs = overcast



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



Promoting use of adjustment by operational communities (now widely used)

Impact on verification results and model ranking

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

Images courtesy Barbara Casati

YOPP YEAR OF POLAR PREDICTION

Accounting for representativeness error

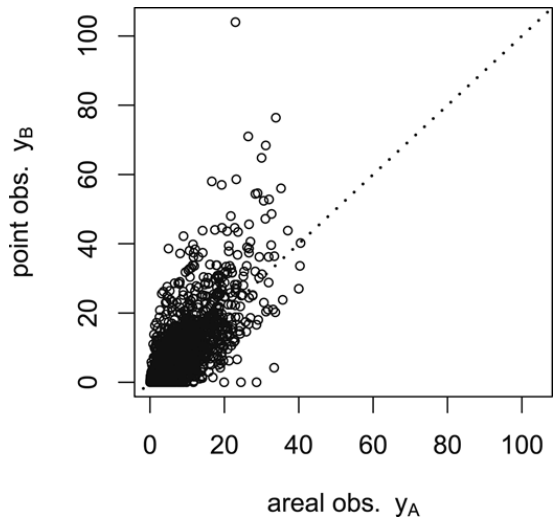
Thomas Haiden

Objective: account for scale mismatch between model output and point observation in ensemble verification. First target variable: daily precipitation.

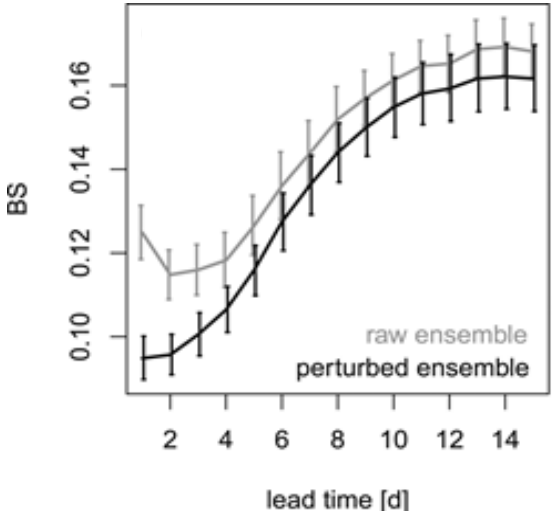
Method: application of a perturbed ensemble approach with perturbations generated from a parametric model. This model is based on an appropriate probability distribution for precipitation fitted with high-density observations.

Impact: large impact on probabilistic scores, large impact on forecast reliability

Scale mismatch



Impact on scores



Ben Bouallegue, Z., T. Haiden, N. J. Weber, T. M. Hamill, and D. S. Richardson, 2020: Accounting for Representativeness in the Verification of Ensemble Precipitation Forecasts. Mon. Wea. Rev., https://doi.org/10.1175/MWR-D-19-0323.1

Verification of Tropical Cyclone (TC) Strike Probability

Raghu Ashrit

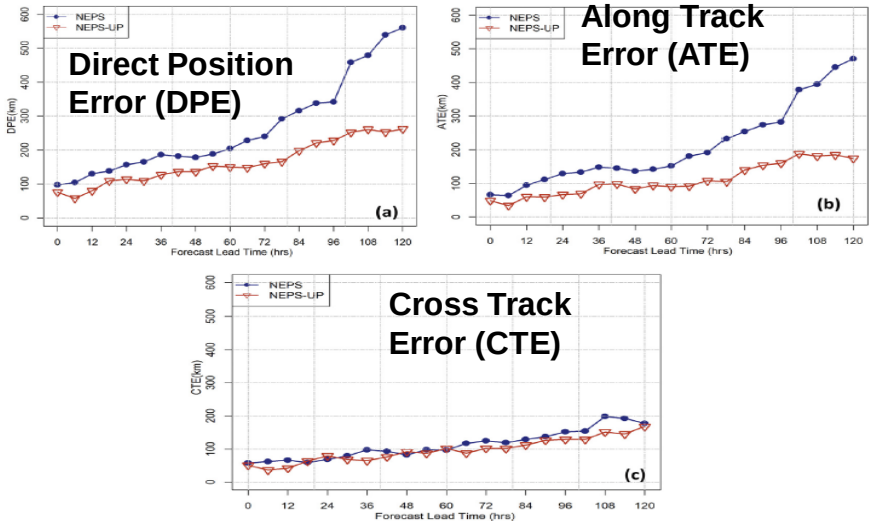
- **Objective** To asses the impact of increased grid resolution (from 35km to 11km) and decreased of ensemble size (from 44 to 23) on the North Indian Ocean Tropical Cyclone forecasts.

- **NCMRWF Ensemble Model upgrade in 2018 :**

Main differences between NEPS and the upgraded version NEPS-UP.

	NEPS	NEPS-UP
Time Step	12 min	5 min
Horizontal Resolution	33 km	12 km
Grid Size	800 × 600	2048 × 1536
Ensemble Size	44 member & 1 control	11 member & 1 control
Perturbation Generation	ETKF	ETKF also additional perturbations for SST and Deep Soil Temp
Forecast Duration	00 UTC cycle	00 and 12 UTC cycles
Operational Duration	December 2016–31 May 2018	1st June 2018 - till Date

Reduced Forecast Track Errors



Slightly improved Reliability and Discrimination of strike probability forecasts

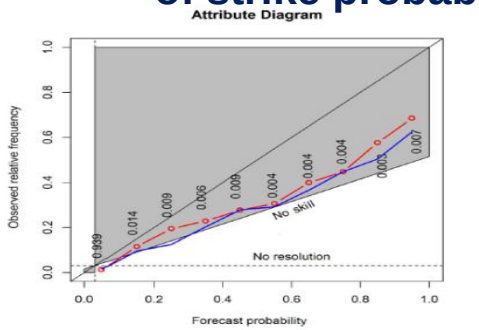


Fig. 3. Comparison of the Reliability Diagram for the strike probability forecasts from NEPS and NEPS-UP.

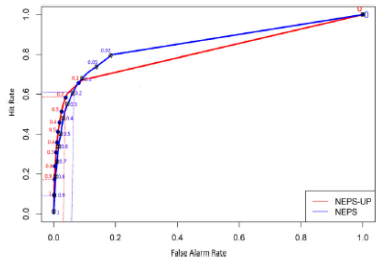


Fig. 4. A comparison of the Relative Operating Characteristics (ROC) curves between NEPS and NEPS-UP. The forecast probabilities are indicated by the dots and numbers on the curves. The horizontal lines indicate the hit rate and the vertical lines indicate the false alarm rate for each forecast probability.

The study documents the improved skill in NCMRWF Ensemble Model in TC forecasts with the upgrade in 2018 from NEPS (17km grid 44member) to NEPS-UP(12km grid 23member). The verification TC strike probability is the new addition to the traditional approaches.

Contributes to high-impact weather (extremes), water and urbanization WWRP societal challenges

Anumeha Dube, Raghavendra Ashrit, Sushant Kumar, and AshuMamgain, 2020: Improvement in Tropical Cyclone Forecasting through Ensemble Prediction System at NCMRWF in India. Tropical Cyclone Research and Review, 9(2), 106-116 DOI: 10.1016/j.tcrr.2020.04.003

Future plans and opportunities

- The Paris 2024 RDP: urban verification (e.g. air quality forecasts) in collaboration with GURME and WGNE (links with urbanization challenge)
- Connection with DAOS: Areas of overlap proposed by the SSC include
  - Assimilation of spatial patterns (as opposed to point observations)
  - Representativeness and observation uncertainty estimation using well-tuned DA and use in verification
  - Handling of extremes
  - Local versus global assimilation/verification, particularly at short time scales
- Promote use of new ensemble-based reanalyses in verification, including estimation of errors in reanalyses and incorporating their uncertainties in verification (link with the climate community)
- Promote research on use of novel observations for verification (links with evolving technologies challenge)
- Promote research on error (or performance) tracking in order to understand the source of forecast error (and sources of predictability)
- Promote process-diagnostics research incl. fluxes and exploiting satellite measurements: WGNE, YOPPsiteMIP (links with high-impact weather and water chall.)
- New area of research: Verification of impact(-based) forecasts and warnings of high impact weather. Promote activities to stimulate development of impact forecast verification (e.g. in the HIWeather project). (links with high-impact weather challenge)
- Connection with TMR: tropical cyclone verif. (links with high-impact weather, urbanization and water chall.)

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