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
- Summary of contributions to WWRP/WCRP projects
- Quantifying observation uncertainty on verification measures - A MesoVICT example
- Progress with process-oriented and surface verification

WG involvement in WMO projects/activities

- SWFDPs – Workshop in March 2018 (Viet Nam) including verification training

- Winter Olympics 2018 (PyeongChang) – Drafted verification plan. Actual verification to be performed after the Olympics.

- HIGHWAY/ L. Victoria – Funded by UK DfID. Research into the surface observation availability on GTS. Case studies into lightning diagnostics. Website for real-time forecasts and monthly verification statistics. To be concluded in March 2020.

- AvRDP – Working on verification guidelines for convection with focus on users' needs.

-TLFDP – WG member in SC contributing to verification planning. Currently in 3rd phase. End in Dec 2018.

- CBS – Flash flood verification. Developed guidance document.

WG involvement in WWRP/WCRP

- PPP – Verification activities focussed on Year of Polar Prediction (YOPP), internationally coordinated period of intensive observations, modelling, prediction, verification, user-engag. and education activities. YOPP is a key component of the Polar Prediction Project

- HIW – Contributing to verification activities through the evaluation task team

- S2S (joint WWRP/WCRP)– Produced chapter and scientific papers on S2S verification
- - Book: Sub-seasonal to seasonal prediction. The gap between weather and climate forecasting (1st Edition, Oct 2018). Chapter 16: Forecast verification for S2S time scales
  - A verification framework for South American sub-seasonal precipitation predictions
- - Global precipitation hindcast quality assessment of the Subseasonal to Seasonal (S2S) prediction project models

PPP activities:
Lead by Barbara Casati

Summary: ongoing YOPP verification activities

1. NWP **process-based 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.

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

3. Verification of **sea-ice** prediction during YOPP
   - User-informative distance metrics alongside traditional scores

Thank you!

barbara.casati@canada.ca

ELSEVIER

Sub-seasonal to Seasonal Prediction

1st Edition

The Gap Between Weather and Climate Forecasting

☆☆☆☆☆ Write a review

Editors: Andrew Robertson, Frederic Vitart

Paperback ISBN: 9780128117149

Imprint: Elsevier
Published Date: 26th October 2018

Page Count: 570

Chapter 16: Forecast verification for S2S time scales
Caio A. S. Coelho, Barbara Brown,
Laurie Wilson, Marion Mittermaier,
Barbara Casati

Overview of S2S verification methods and practices

Sampling strategies and information levels for sub-seasonal verification

Proposed framework for quantitative sub-seasonal precip. forecast quality assessment

- Level 1: Target week hindcast verification (11 ens. members)
  Similar to traditional seasonal forecast verification (~30 samples)
  Uses ECMWF S2S hindcasts intialized on Thu 14 April, 7 April, 31 March and 24 March of the 2016 calendar for the period 1996-2015 (20 samples)

- Large degree of differences in some characteristics of sub-seasonal hindcasts and real time forecasts, directly impacting the verification sample size.

- Level 2: All season hindcast verification (11 ens. members) Increased robustness
  In addition to the hindcasts produced for the four Thu initialization dates previously selected, aggregates hindcasts produced for nine additional initialization dates during the weeks of the previous and following month in order to incorporate in the sample all hindcasts initialized on Thu of March, April and May of the 2016 calendar (260 samples: 13 initialization dates times 20 years)
  MAM: Austral summer season, similar atmospheric features in S. American regions

- Level 3: All season near real time forecast verification (51 ens. members)
  Aggregate the real time forecasts produced on Thu during the 13 weeks of March, April and May of each of the past three years (2015, 2016 and 2017). This aggregation leads to a verification sample of 39 pairs of near real time forecasts and observations (39 samples: 13 initialization dates times 3 years)

Coelho, Caio A.S.; Firpo, Mári A.F.; de Andrade, Felipe M., 2018:A verification framework for South American sub-seasonal precipitation predictions. Meteorologische Zeitschrift.

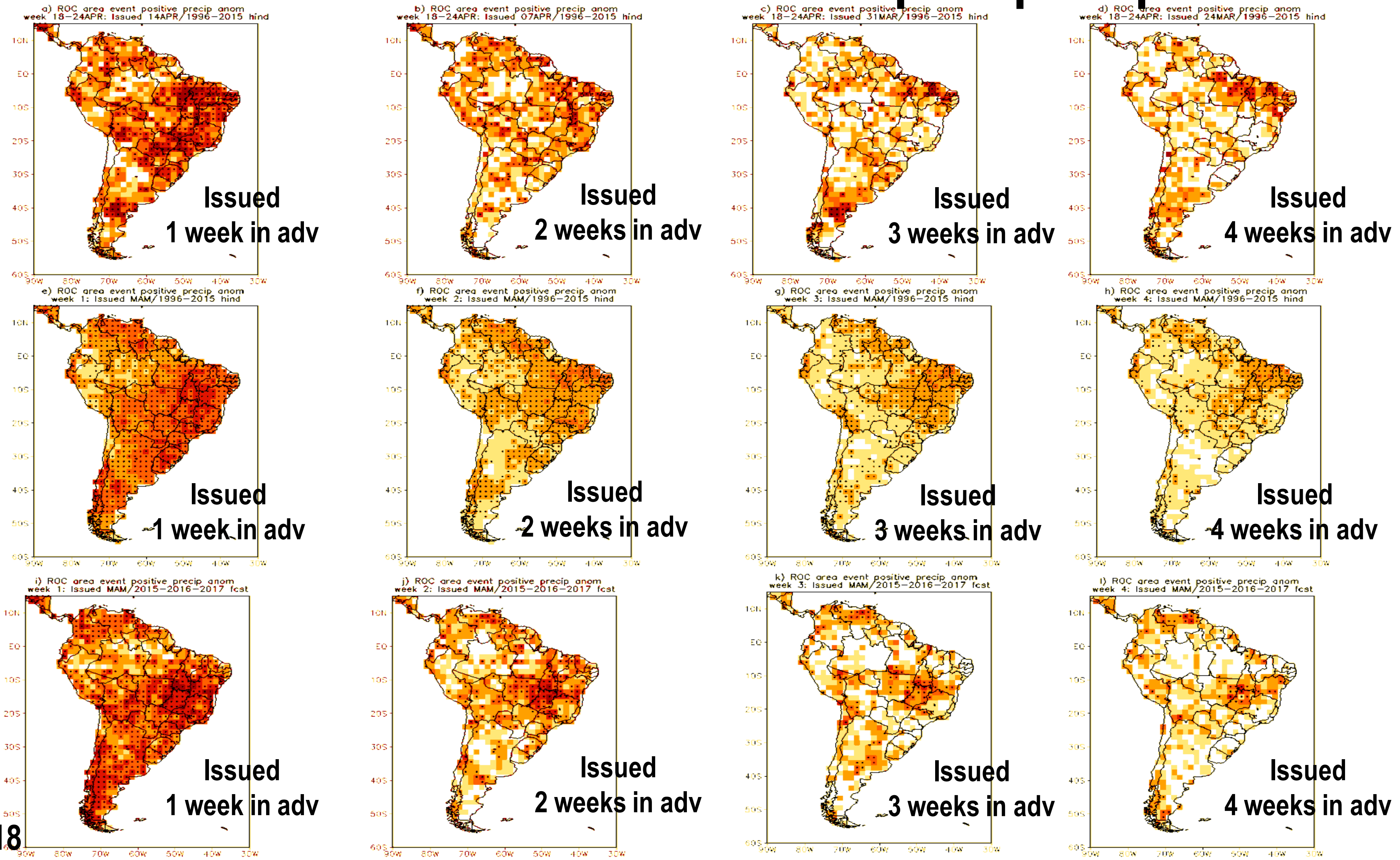
Discrimination comparative assessment: Area under the ROC curve for event pos. precip. anom.

ECMWF S2S database

Level 1:
Target week
hindcast verification

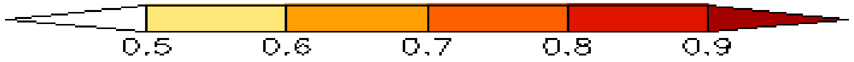
Level 2:
All season
hindcast verification

Level 3:
All season
near real time
forecast verification

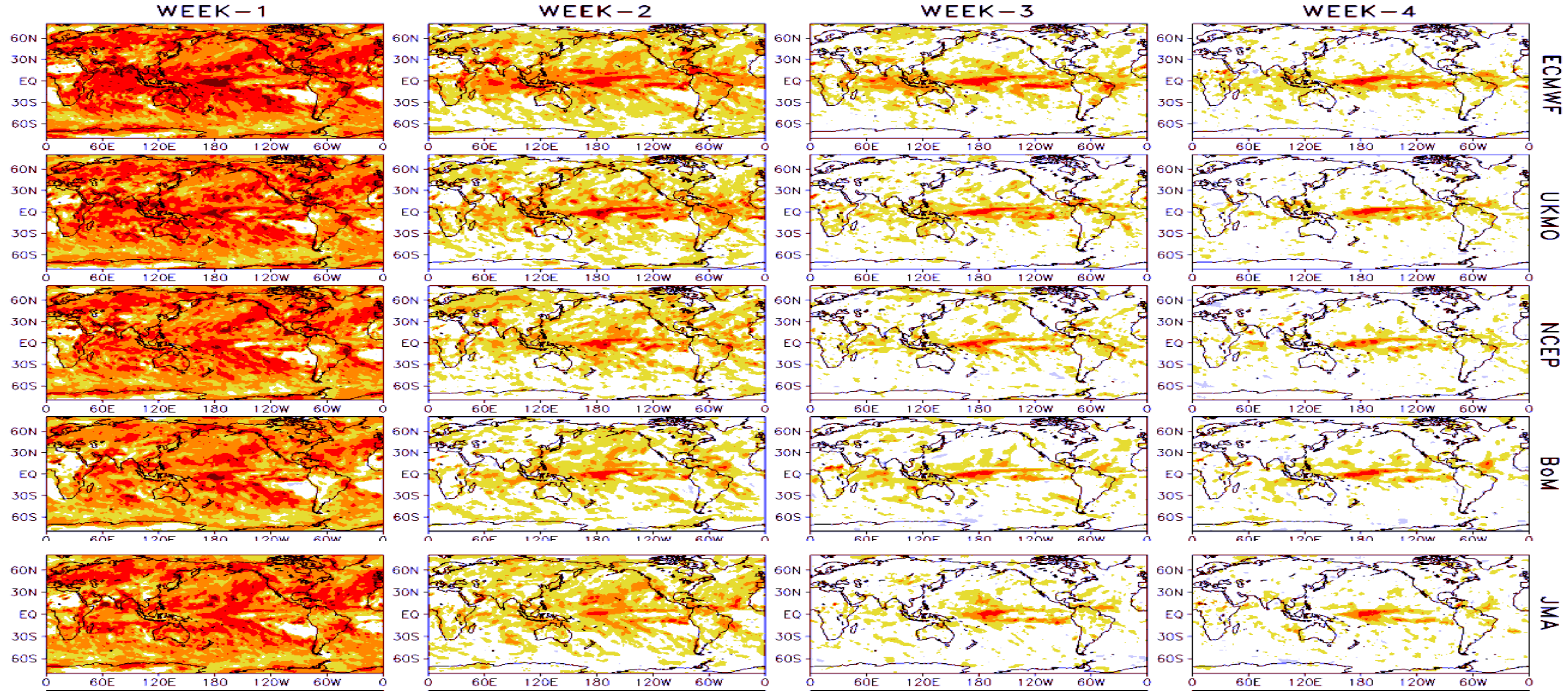


Coelho, Firpo, de Andrade, 2018

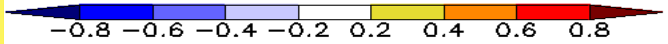
Discrimination: ability to distinguish events from non-events



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



Linear association assessment: Correlation



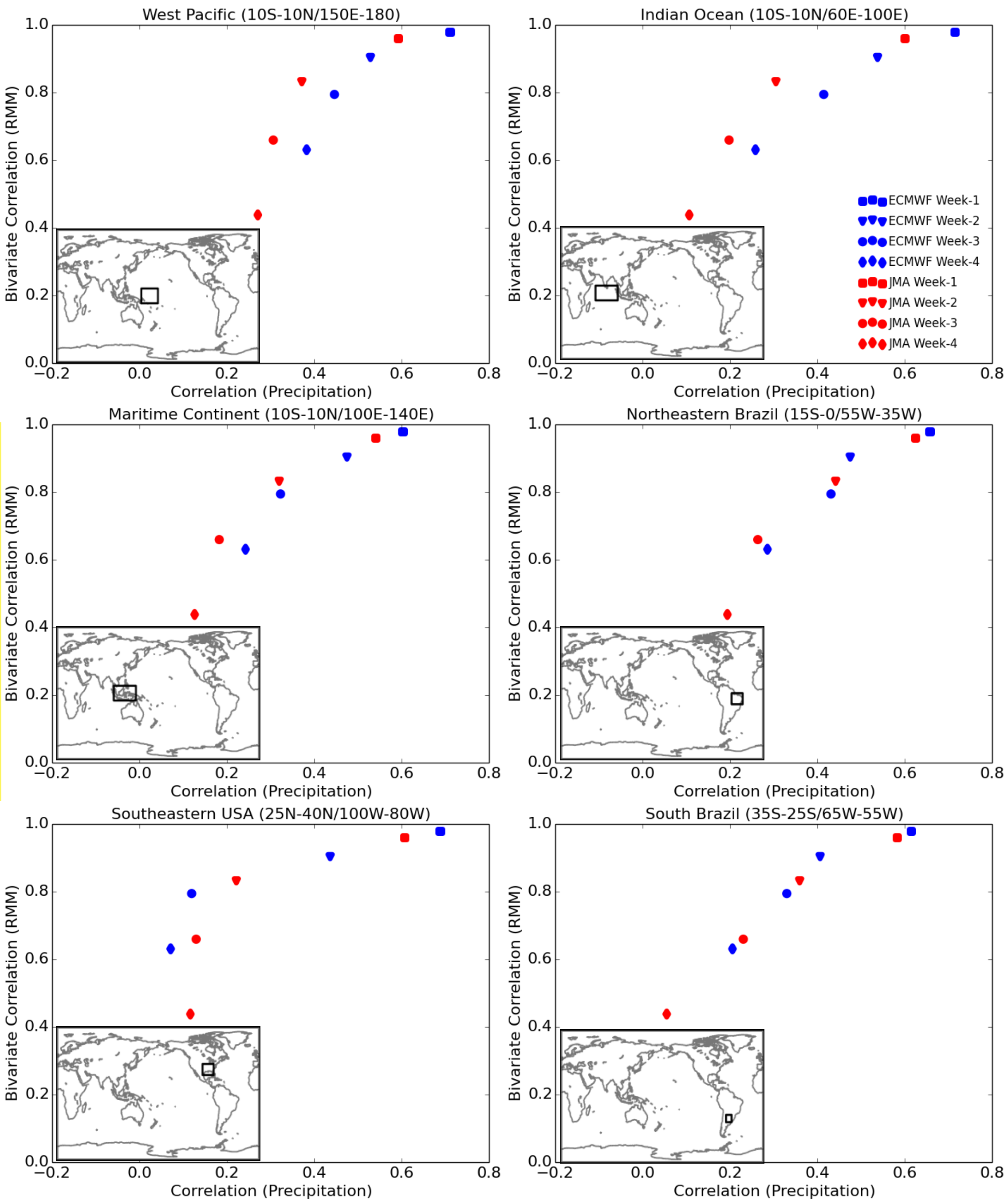
Extended austral summer: Nov to Mar 1999-2009

Felipe M. de Andrade, Caio A. S. Coelho, Iracema F. A. Cavalcanti, 2018: Global precipitation hindcast quality assessment of the Subseasonal to Seasonal (S2S) prediction project models. Climate Dynamics

What is the relationship btw MJO and precip. prediction ability in a coupled and an uncoupled model?

Linear association btw MJO and precip. prediction ability

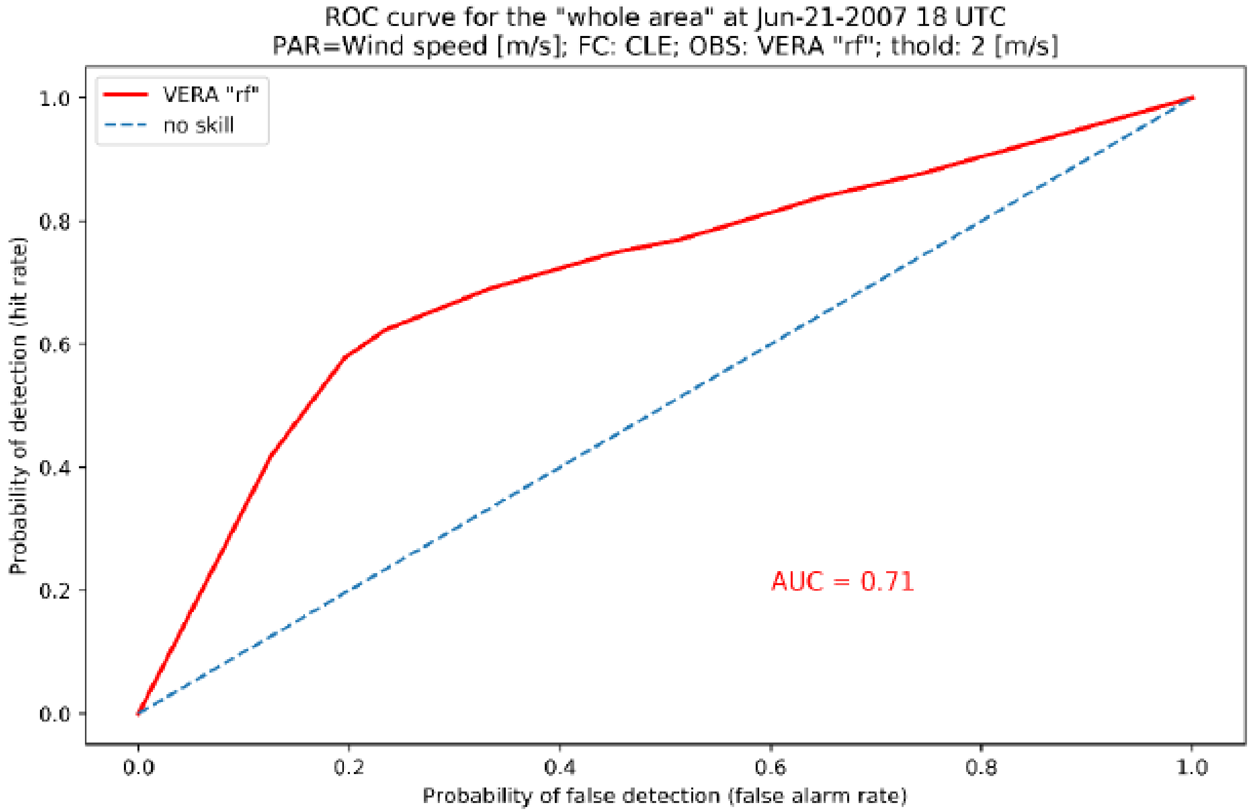
Ocean-atmosphere coupling likely has an important contribution for providing better subseasonal MJO and precipitation prediction ability, particularly on the tropical region.

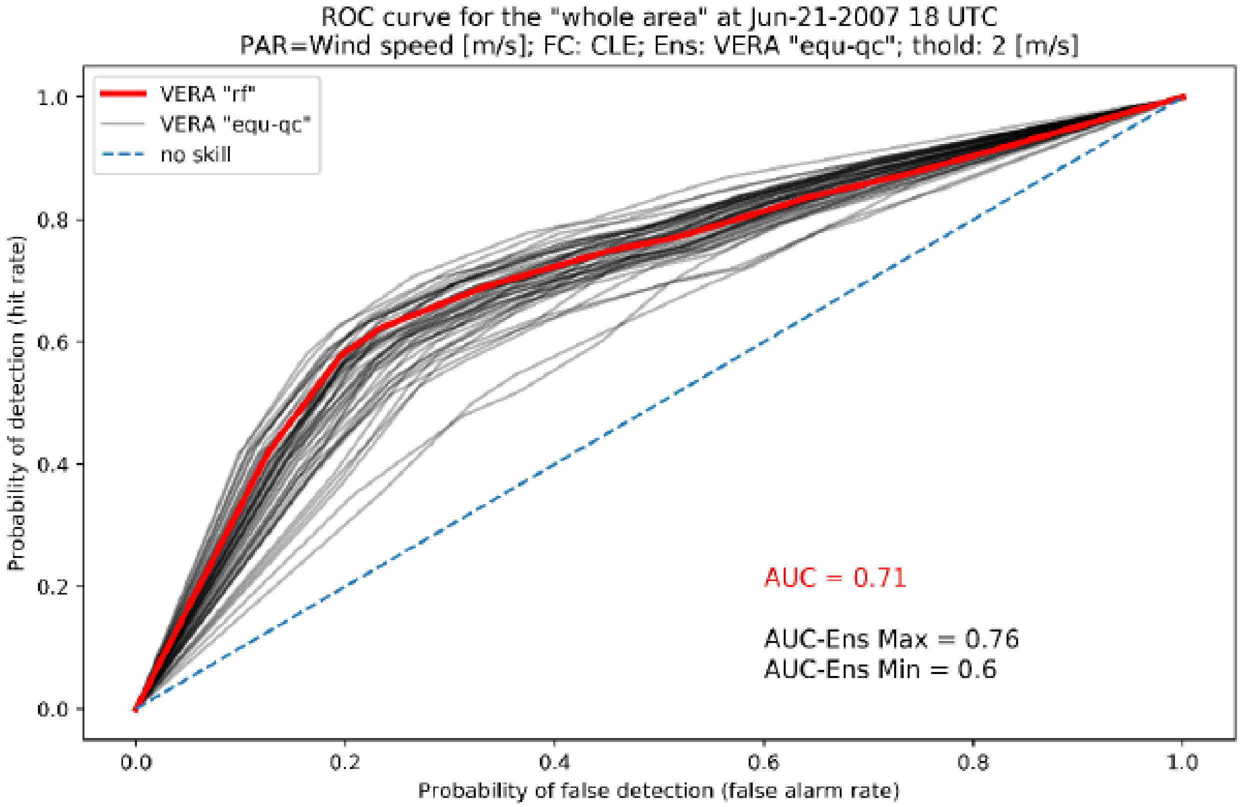


de Andrade, Coelho, Cavalcanti (2018)

Quantifying observation uncertainty on verification measures - A MesoVICT example

Manfred Dorninger and Simon Kloiber
University of Vienna, Department of Meteorology and Geophysics; Vienna, Austria
email: Manfred.Dorninger@univie.ac.at

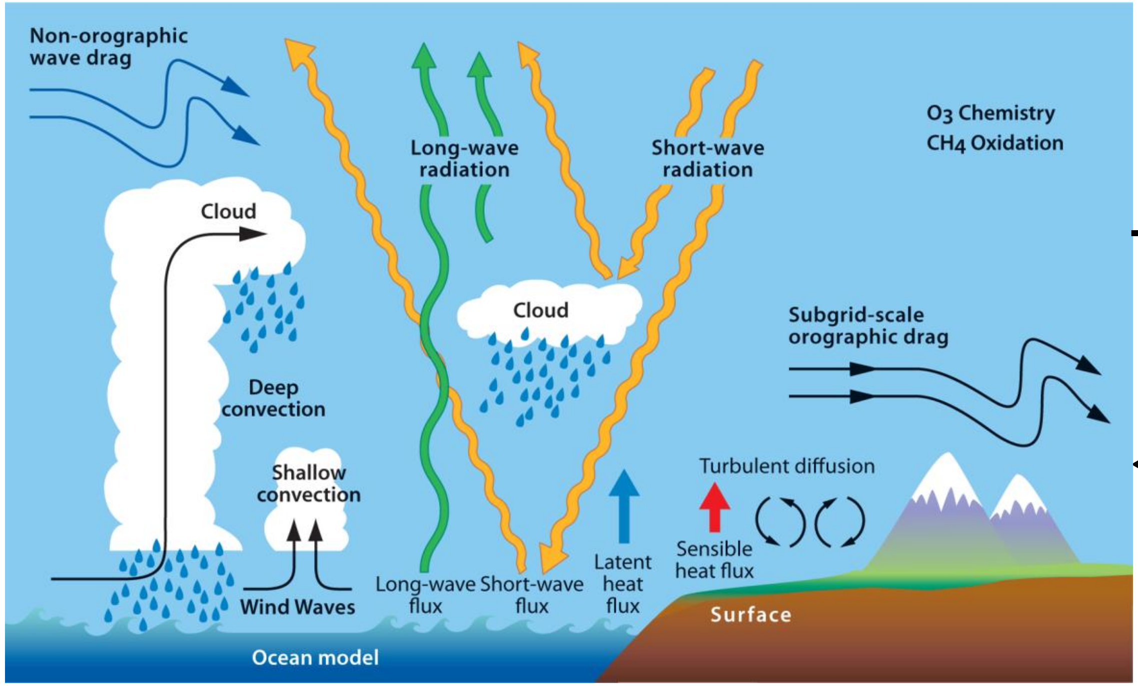




Use of analysis ensemble allows quantification of uncertainty in verification scores

Process-oriented verification

Thomas Haiden, ECMWF and JWGFVR



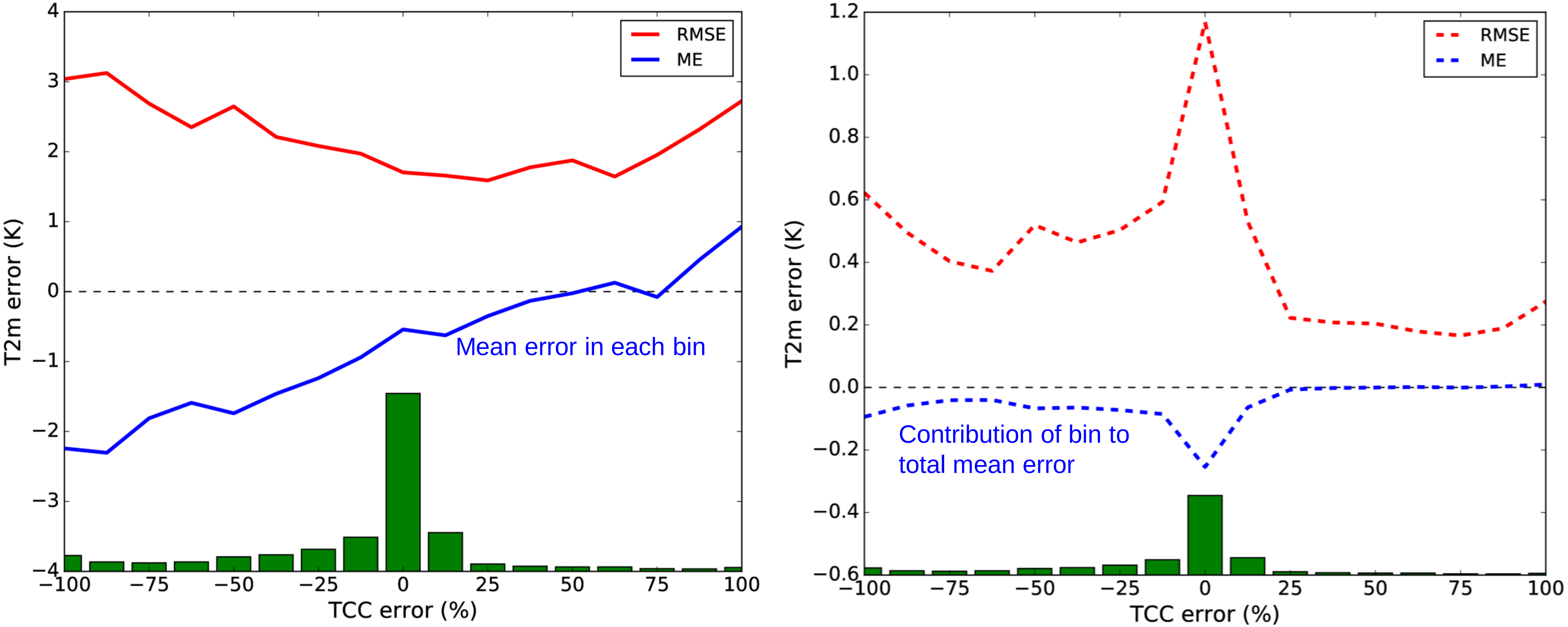
Why is process-oriented verification of interest?

- **Potential to improve the research to operations process in NWP**
- Verification that helps to identify causes of issues allows more efficient feedback to model developers
- WGNE-32: JWGFVR to provide a document on process-oriented verification (draft version to be circulated within JWGFVR by the end of 2018)
- What methods does it involve?
    - Conditional verification
    - Use of supersite observations
    - Model intercomparison
    - Combination of independent datasets
    - Other ..

Process-oriented verification methodology examples

Conditional verification: T2m error stratified by total cloud cover error

Winter in Europe at night



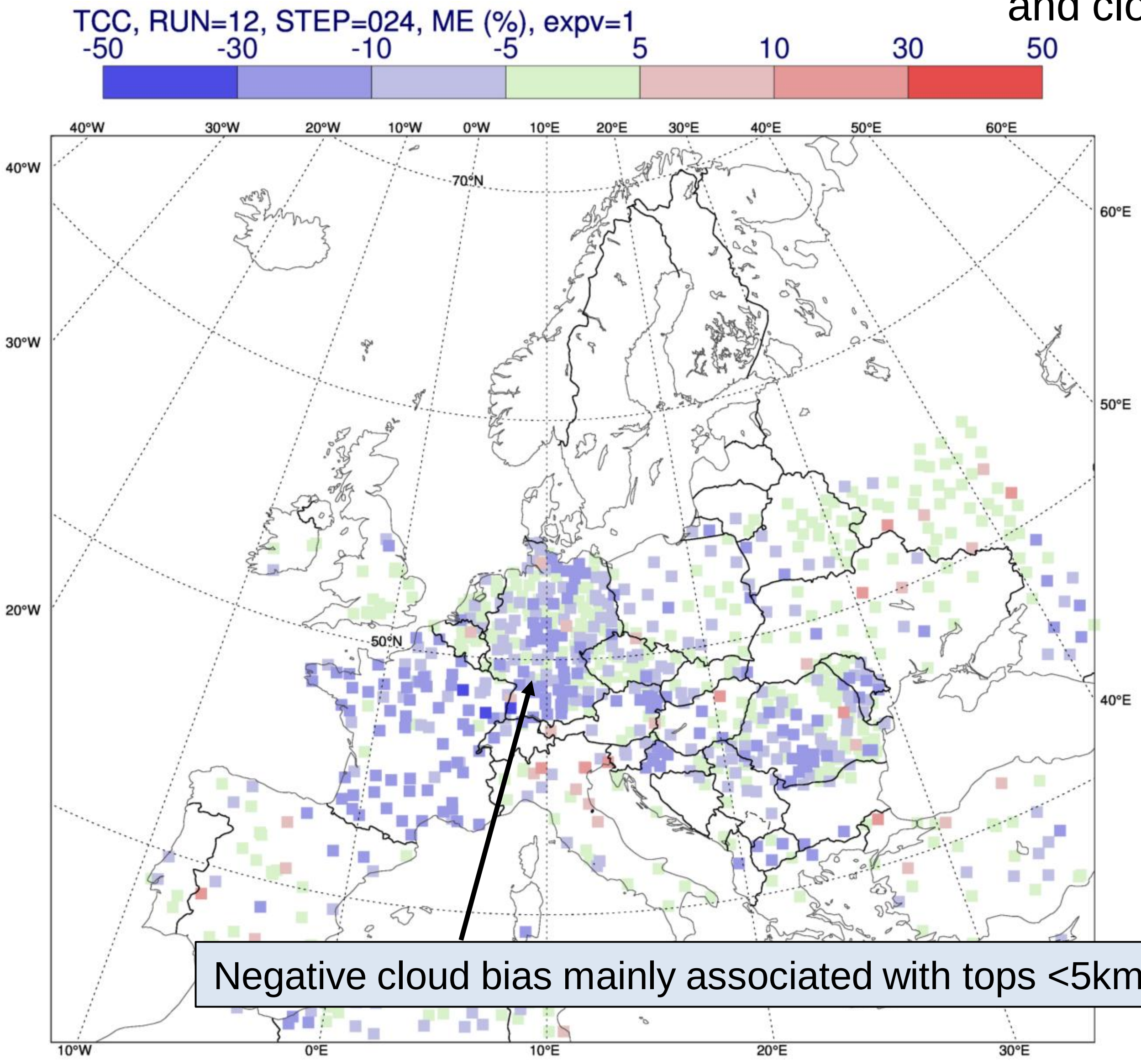
Overcast observed, clear sky predicted: T2m too low (neg. bias)

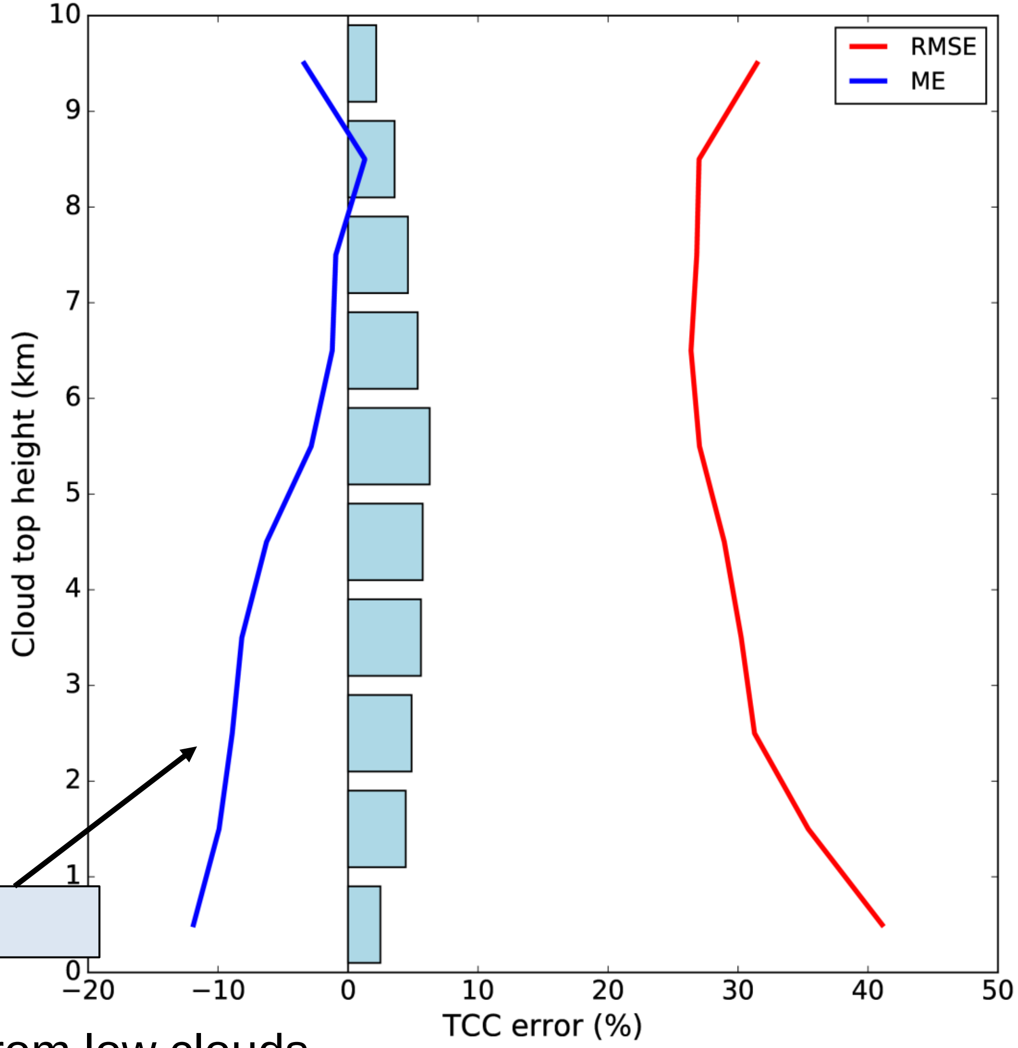
Total RMSE = 1.99 K
Total ME = -0.84 K

Reveals the importance for overall T2m bias of cases with/without TCC error

Conditional verification: TCC error stratification by cloud top height

Example of combined use of sfc. obs. (TCC from SYNOP), and cloud top height (from satellite) to gain insight





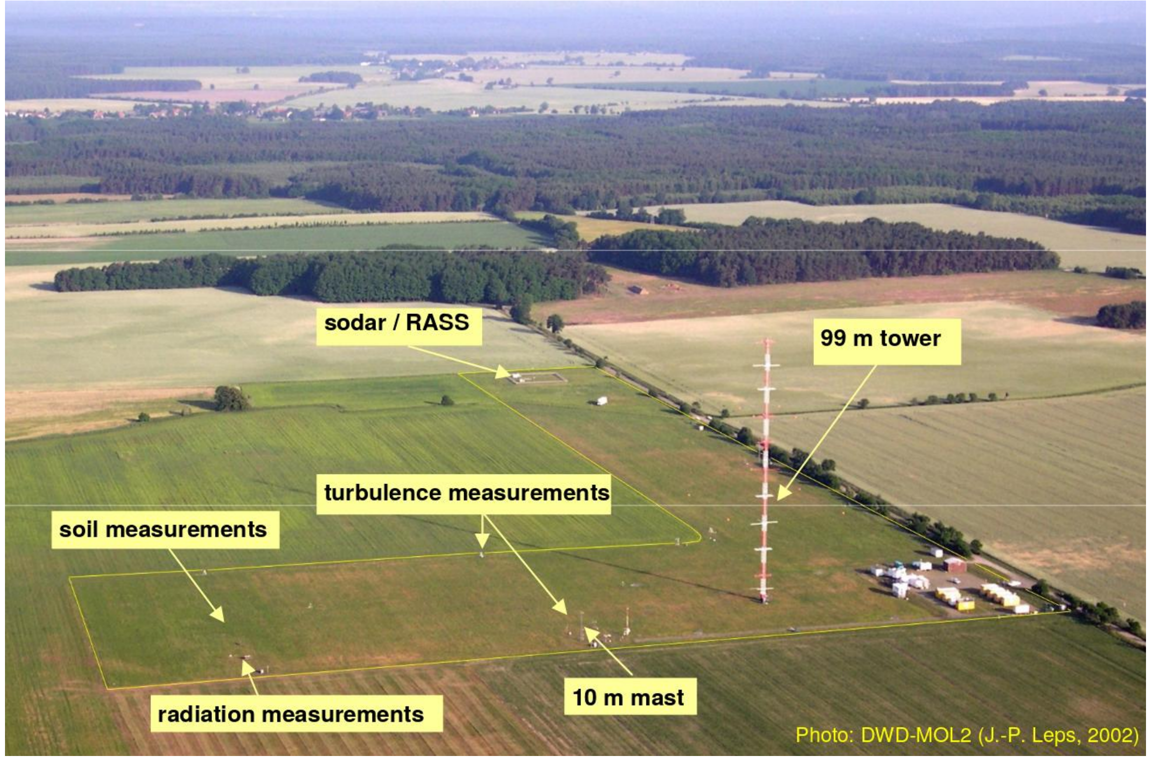
Negative cloud bias mainly associated with tops <5km

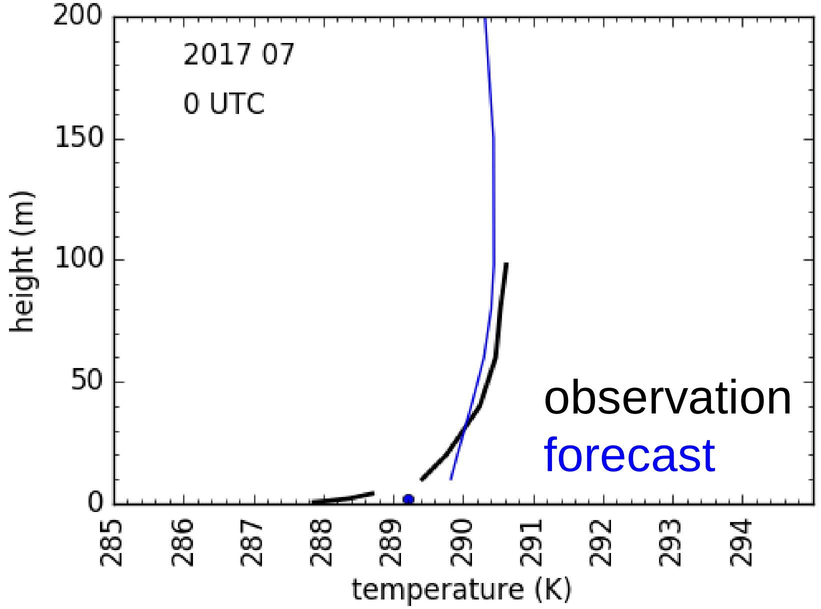
Large contribution to negative TCC bias comes from low clouds

Verification against supersite observations

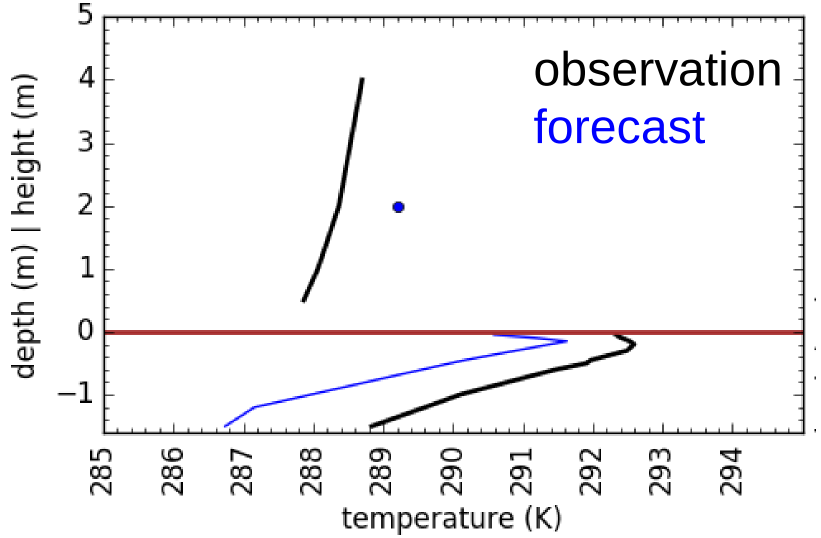
Verification against profile observations in air and soil

Lindenberg, Germany





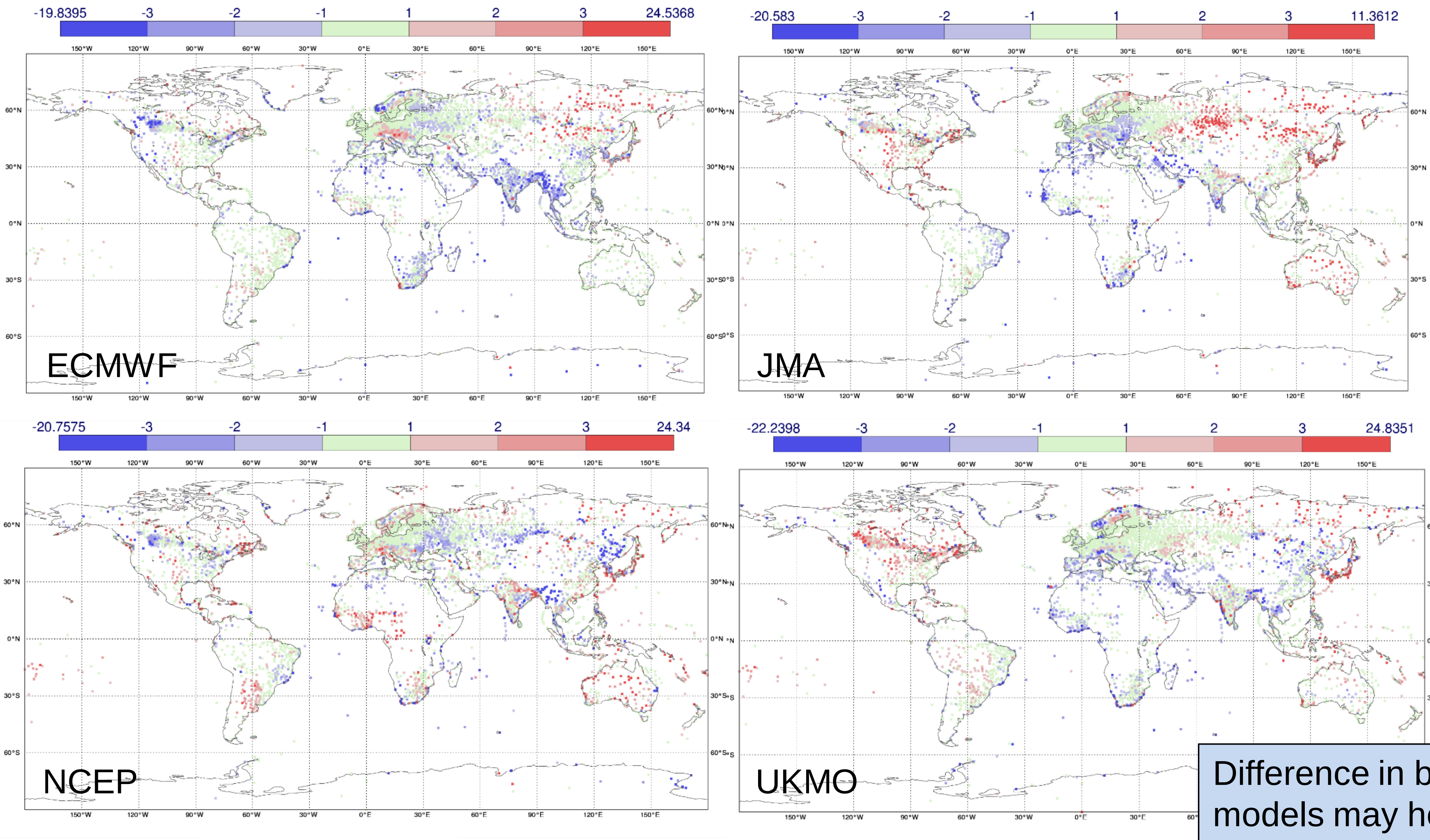
2m temperature too high
Soil temperature too low



A comprehensive set of observations allows to better constrain parameterizations

Model intercomparison

T2m bias of different global models (TIGGE)



Day 5
12 UTC
DJF 2016-17

Difference in biases between models may help to better understand their causes

New WMO guidelines for exchange of surface scores (T2m, 10m wind, total cloud cover, precip)

ECMWF Espaços Pesquisar Autenticação

WMO Lead Centre for Deterministic NWP Verification (LC-DNV)

Páginas

Notícias

ÁRVORE DE PÁGINAS

- Standard verification procedures
- **Exchange of WMO surface verifica**
  - Answered Questions on the surfac
  - SVS: ECMWF Implementation Not
- Lead Centre guidelines
- Reports

Páginas / WMO Lead Centre for Deterministic NWP Verification (LC-DNV)

Exchange of WMO surface verification scores

Criado por Martin Janousek, última alteração em fev 27, 2018

The standard procedures for verification of surface variables are defined in the Manual on the Global Data-processing and Forecasting System: Annex IV to the WMO Technical Regulations (available from WMO library at https://library.wmo.int/opac/index.php?lvl=notice_display&id=12793#.Wo8mHa0pFrk), appendix 2.2.34.

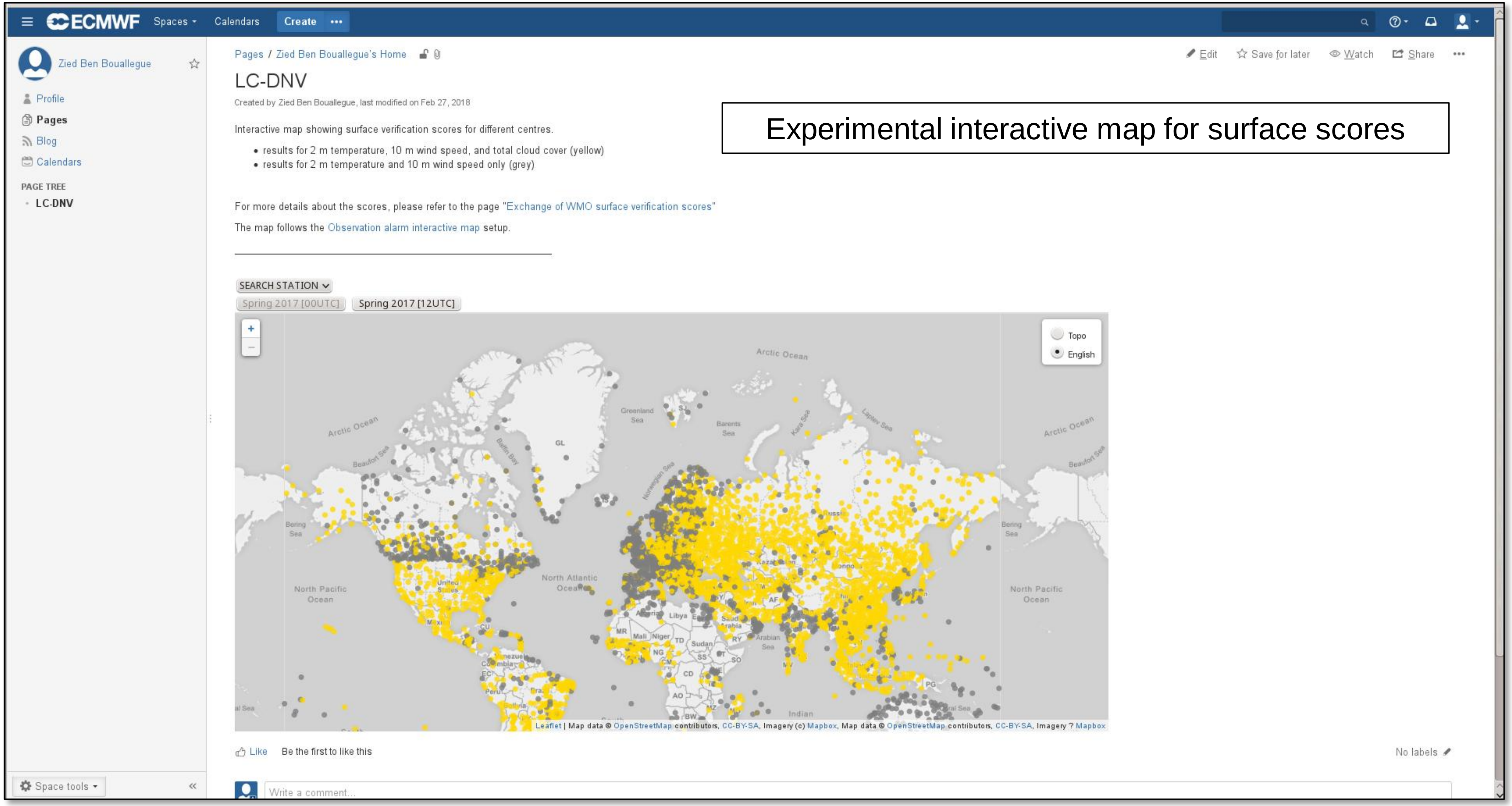
ECMWF in the capacity of LC DNV collects and archive the scores. Data are exchanged in an ASCII-based format which is simple but flexible and is similar to the format used WMO exchange of domain-averaged upper-air scores.

Resources

- The format of the reports of the station-based verification scores.
- The procedures for the exchange of (surface) verification reports.
- The remarks on the implementation of the surface verification reports at ECMWF.
- Answered Questions on the surface verification exchange

https://confluence.ecmwf.int/display/WLD/Exchange+of+WMO+surface+verification+scores

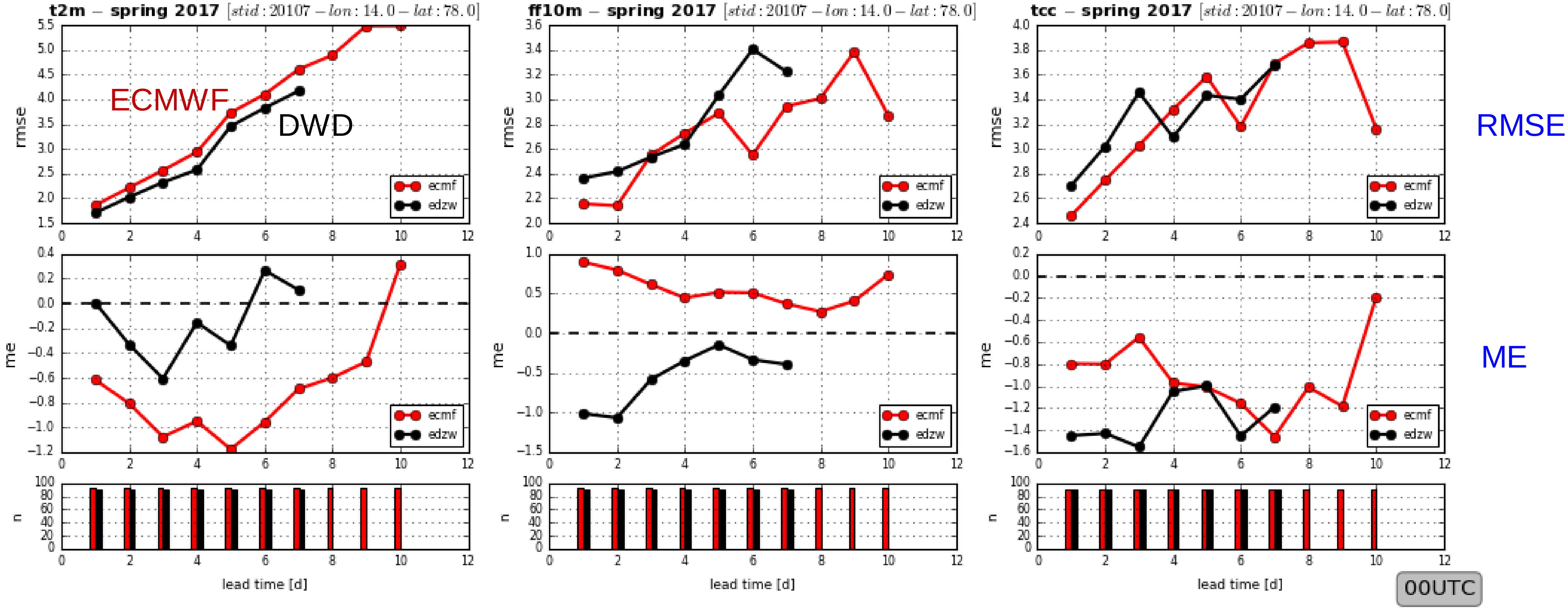
WMO CBS exchange of surface scores (please participate!)



https://software.ecmwf.int/wiki/display/~mozb/LC-DNV

WMO CBS exchange of surface scores (DWD and ECMWF so far)

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



Scores for individual SYNOP stations from different models

ECMWF

Some ingredients for process-based verification

- Decompose scores and metrics
- Stratify errors and perform conditional verification
- Constrain observation error using multiple datasets (e.g. in-situ, satellite)
- Use supersite/tower data
- Use Earth’s diverse geography to focus on specific regimes/processes
- Evaluate sensitivity to parameterization changes (also in single-column mode)
- Do the above for a range of models

These are already common in research, can we adopt in operational verification as well?

WG membership

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

Two vacancies.

Thank you for your attention!