

Uncertainty estimation in modern reanalysis systems



Climate Change

András Horányi, Per Dahlgren, Gionata Biavati, Hans Hersbach, Joaquin Muñoz-Sabater, Raluca Radu

European Centre for Medium-Range Weather Forecasts
(ECMWF)

ICR5, Rome, 13-17 November, 2017





Climate
Change

Introduction / rationale

- Uncertainty estimation is an **added value** to any meteorological analysis/forecasts → it is not different in reanalysis
- In many of the available reanalysis datasets uncertainty estimates are not available so far
- At ECMWF **CERA-20C** and **ERA5** include uncertainty estimation based on a 10-member EDA (Ensemble of Data Assimilations) system
- The main objective of the present talk is to give an overview about the **reliability diagnostics** for CERA-20C and ERA5 and give some hints about the use of EDA for reanalysis uncertainty estimation



Climate
Change

Spread-skill relationship (ensemble reliability)

- In a well-calibrated ensemble the ensemble spread should match the skill of the ensemble mean (spread-skill relationship, optimally diagonal line in a scatterplot)
- Some questions about the ensemble reliability:
 - Can the raw ensemble spread from EDA be directly used for the uncertainty estimation?
 - How to quantify the skill, i.e. what is the verification truth?
 - Should the errors in the analysis/observation be considered?
 - What guidelines can be given to the reanalysis users for the use of EDA uncertainty information?



Climate
Change

Some practical considerations

- CERA-20C and ERA5 include a 10-member EDA ensemble for computing the background errors (see relevant talks earlier this conference)
- Verification truth: high resolution IFS analysis (OPER), ERA-Interim, observations taking into account their respective errors
- In ERA5 the skill of the high-res product (rather than that of the ensemble mean) is of interest
- In ERA5 there are various additional EDA spread characteristics available (e.g. scaled ensemble spread, forecast spread)



Climate Change

Choice of verification truth + use of analysis/observation errors

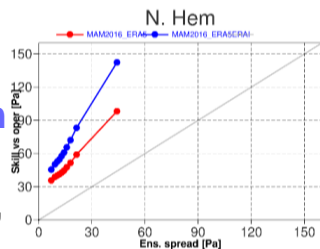
ERA5 raw MSLP spread-skill for spring, 2016

Skill=ERA5 vs OPER
Skill=ERA5 vs ERA-Interim

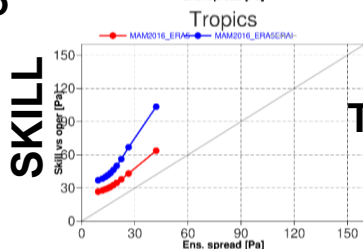
Errors in reference (ERA-I, OPER) **NOT** taken into account



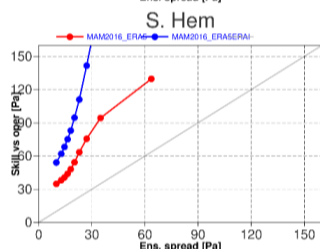
Looks better when verifying against operations, but it also depends on the period chosen (i.e. for older periods ERA-Interim is a better truth)



NH



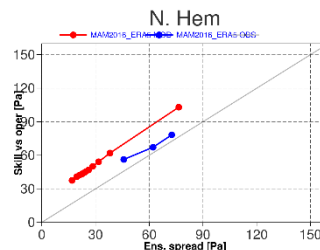
TROPICS



SH

SPREAD

MODEL SPACE



Skill=ERA5 vs OPER
Skill=ERA5 vs Observations

Errors in reference (OPER, obs) **ARE** taken into account



Very good reliability (slightly better against observations)

Yamaguchi et al., 2016 (QJ)
Saetra et al., 2004 (MWR)

MODEL VS. OBS. SPACE

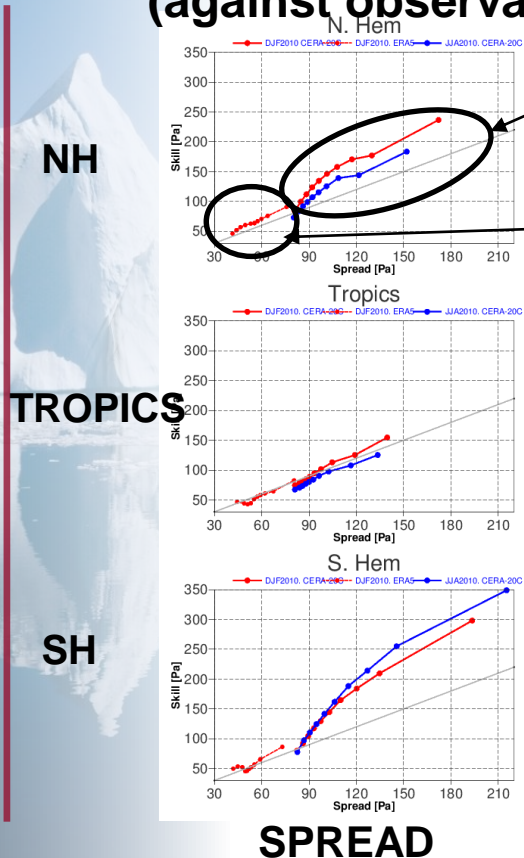




Climate Change

ERA5 vs. CERA-20C

Surface pressure spread-skill for 2010 (against observations taking into account the observation errors)



CERA-20C: summer and winter, 2010

ERA5 (winter, 2010)

SKILL

CERA-20C has a reasonably good reliability, though ERA5 is much better (better resolution and more observations)

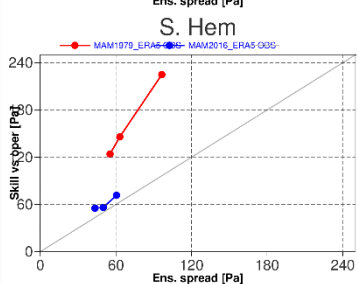
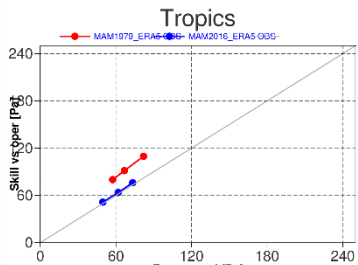
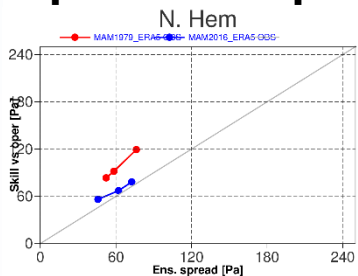




Climate Change

ERA5: different periods

Surface pressure spread-skill for 1979 and 2016 spring against observations



SKILL

1979 is under-dispersive especially at the Southern Hemisphere (it is anyway expected that the reliability depends on the time period too)

SPREAD



Climate
Change

Summary, conclusions, discussions

- The spread-skill relationship of both CERA-20C and ERA5 reanalysis systems seems satisfactory (especially ERA5)
- It is absolutely essential to take into account analysis/observation errors in the production of proper spread-skill relationship
- Though guidance to the ERA5 users might not be straightforward
 - We will provide our diagnostic results
 - And provide recommendations what to use and in what aspects the users should be careful at
 - Feedbacks are welcome
- Next steps: extend the results for more variables and more periods