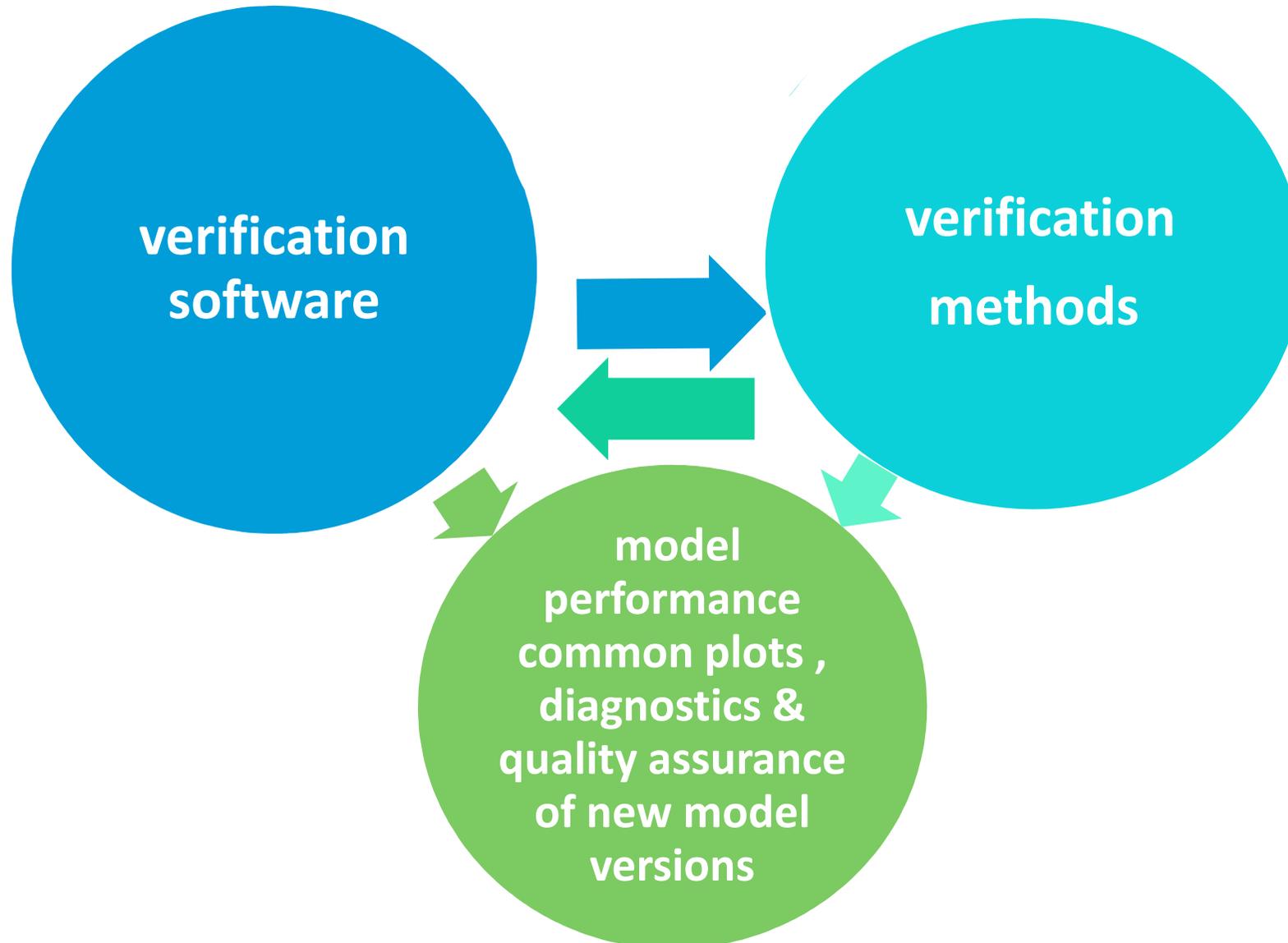


Verification Activities Overview

Flora Gofa

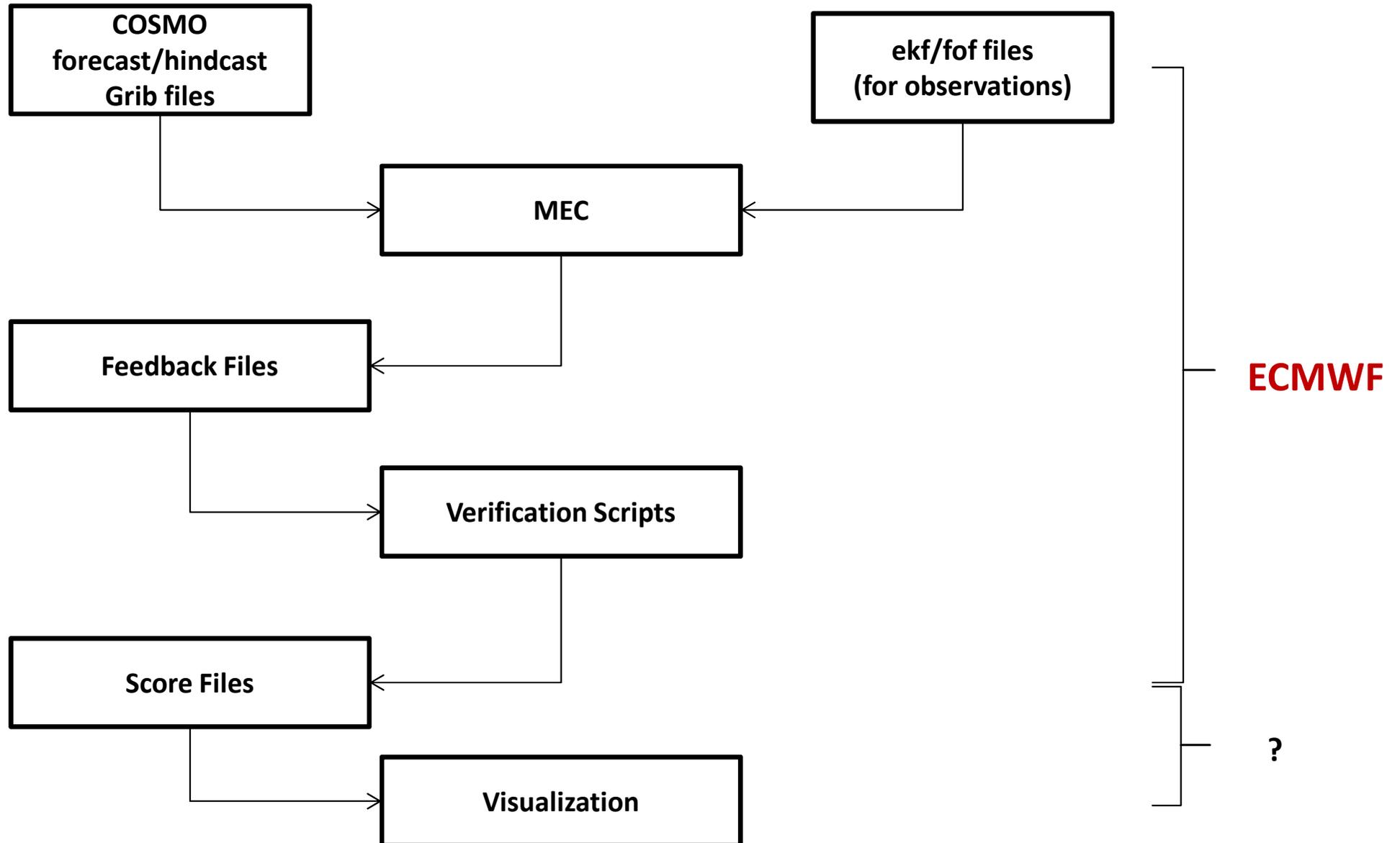
COSMO Working Group on Verification and Case studies

main activities



Feedback File Based Verification at DWD - **Rfdbk**

Workflow



Model equivalent Calculator MEC

Installation

- Sources: Fortran 2003/2008 and some C sources from DWD
- Makefile for gfortran is provided
- NetCDF, CGRIBEX (MPI Hamburg), GRIP-API (ECMWF), (MPI recommended)
- Fortran compiler, C compiler
- Sufficient memory to hold one model state (1 ensemble state)

Required model input

- **Grib or Grib2 files**
- **COSMO, ICON (EU Nest), IFS, HRM, ECHAM (not fully tested)**
- **PS, T, U, V, P, Q (mandatory, all model levels)**
- **T2M, TD2M, CLC, CLCT, CLCL, CLCM, CLCH, CLC, H_SNOW (optional)**
- **TOT_PREC, VMAX_10, TMIN_2M, TMAX_2M (optional, next release)**

Required observation input

- **fof/mon/cof/ekf/ver –files (existing fdbk files from nudging, LETKF or MEC)**
- **CDFIN (BUFR converted by bufrx2netcdf to NetCDF, BUFR in WMP-templates as used by DWD)**

Output

- ver-files, NetCDF feedback files including past forecasts



- Using feedback files for the verification means a huge reduction in workload as much of the tedious data preparation tasks are done within DA
- Rfdbk is a R interface for COSMO feedback files
- Main purpose of Rfdbk is to load feedback file content with R
- Additional functionalities useful for verification is implemented as well



Models

- 3 ICON global deterministic routines
- 3 ICON EU Nest deterministic routines
- 2 ICON global EPS
- 2 ICON EU Nest EPS
- 3 COSMO-DE deterministic routines
- 3 COSMO-DE-EPS ensemble routines
- IFS deterministic
- IFS EPS
- + Experiments

Observation systems

- SYNOP
- TEMP (radiosondes)
- SATOB (AMV)
- GPSRO (radio occultations)
- SCATT (scatterometer)
- AIREP (aircraft)
- PILOT (wind profiler)

Methods

- Deterministic: continuous and categorical
- EPS: ensemble and probabilistic

Visualization

- Lead-time
- Time series
- Station based

Aggregation

- Sub-domains
- Height bins or levels
- Lead-time to time of day conversion („hindcast mode“)



Choose score file(s):
CONT_BYSTATieu_icon-ieup1_iconp1-ieup_iconp-lme_icon-icon-iconp1_iconp1-iconp_iconp_VAR1-29-39-111-11

Load data Make plot

Initial hour
00

Valid hour
00

Variable
2 metre temperature (K)

Score
RMSE

Model name (adds column)
ieu_icon lme_icon

score difference

Lead-time (adds row)
024 048 072

Raster plot

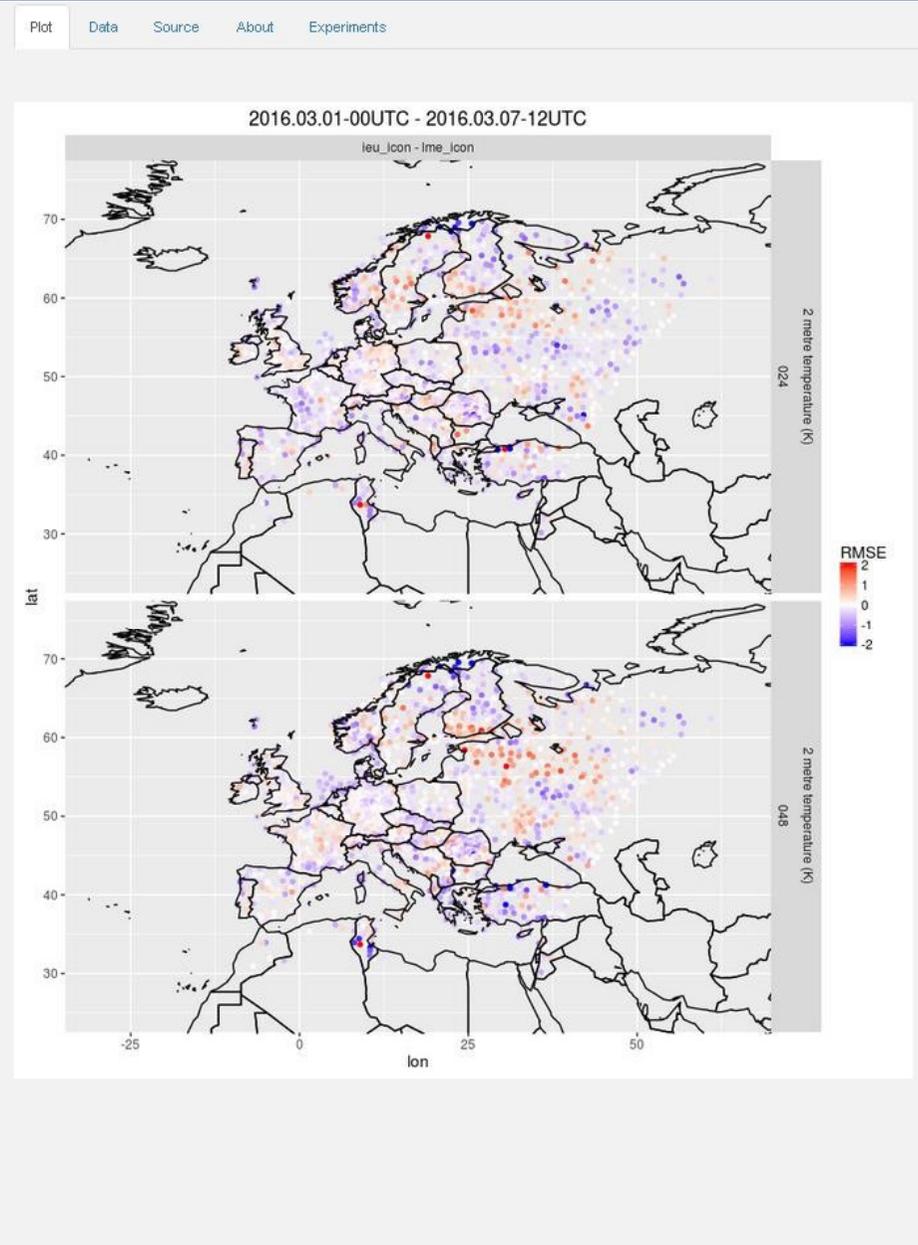
Longitude
-180 -160 -140 -120 -100 -80 -60 -40 -20 0 20 40 60 80 100 120 140 160 180

Latitude
-90 -80 -70 -60 -50 -40 -30 -20 -10 0 10 20 30 40 50 60 70 80 90

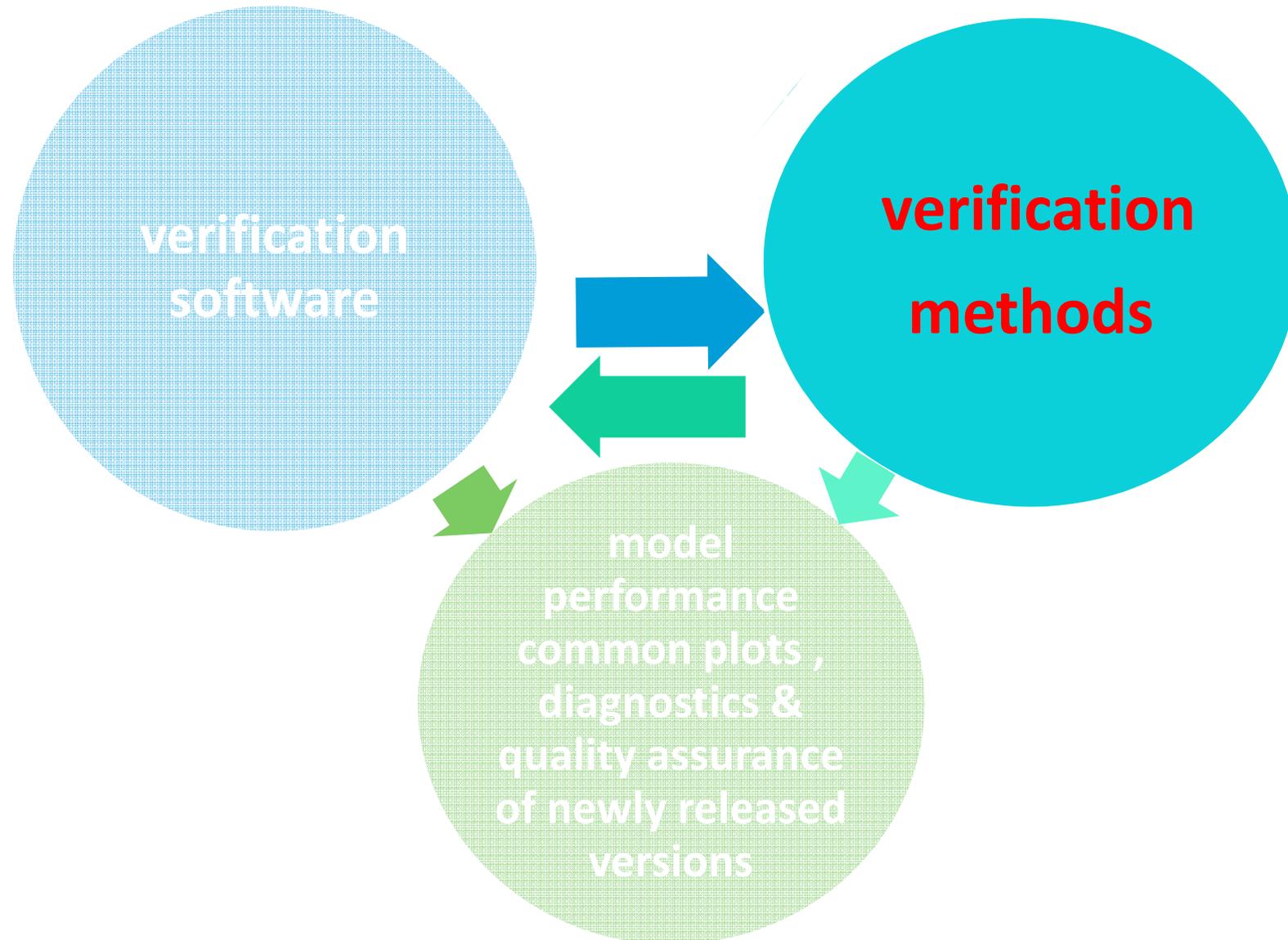
from
-2

to
2

Point size
0.75 1 1.25 1.5 1.75 2 2.25 2.5 2.75 3



main activities



Mesoscale Verification Inter-Comparison over Complex Terrain (<https://www.ral.ucar.edu/projects/icp/>)

Prerequisite for verification method inter-comparison:

-use of same data (Obs and FC) on the same grid and over the same area (Alpine area)

From MAP D-PHASE COPS archive

- Deterministic 2 km COSMO-2 Init-time: Initialised 06 UTC FC-range: 24h
- Deterministic 2 km CMC-GEM-H Init-time: Initialised 06 UTC FC-range: 18h
- Ensemble 10 km COSMO-LEPS Init-time: Initialised 12 UTC FC-range:132h

MCH

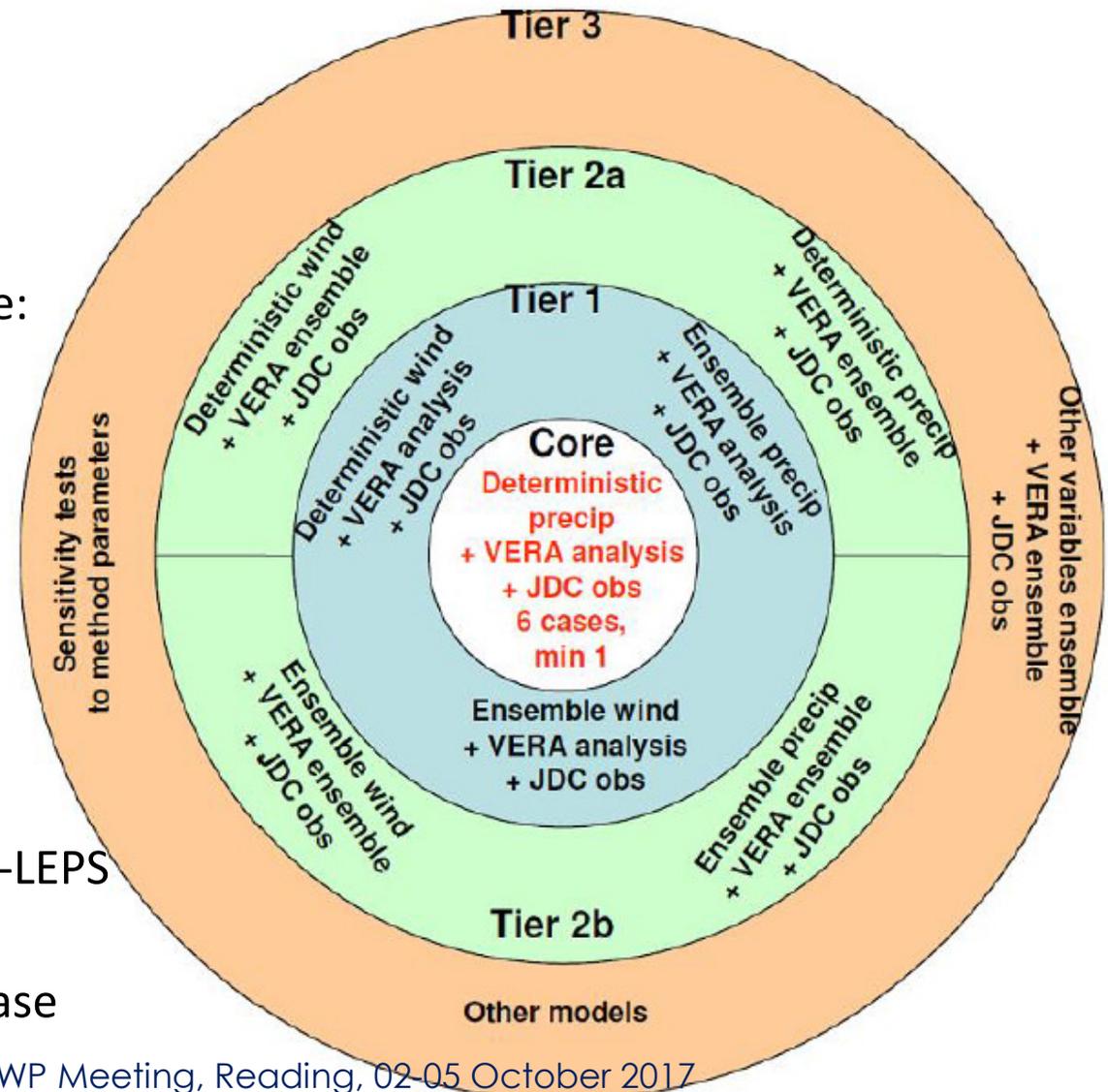
- Reruns COSMO-1 models for 4 cases

ARPAE

- ECMWF-IFS reruns for cases 1,2
- to provide boundary conditions for COSMO-LEPS

•Roshydromet

- COSMO-Ru2-EPS: rerun for 1st MesoVICT case



PP-INSPECT: Intercomparison of Spatial Verification methods for COSMO Terrain



- summarizes COSMO experience of applying spatial verification methods able to capture the relative skill of very high-resolution systems
- INSPECT runs in parallel to MesoVICT (INSPECT tasks involve reruns of COSMO models (determ and EPS) for MesoVICT test cases over **complex terrain** and analysis of them)
- Same as MesoVICT, INSPECT focuses also on the ensembles and variables besides precipitation
- In addition to targeting the goals of MesoVICT, INSPECT is providing COSMO users more choice of verification domains and reference data – more recent and longer periods, two complex terrains (the Alps and the Caucasus)
- Sharing of software/scripts/methods
- **Finally, INSPECT provides criteria for deciding which methods are best suited to particular applications**

Tasks involving development of routines for neighborhood, CRA, SAL, and MODE applications

- For the most part, the software is based on **free R SpatialVx package** (developed by E. Gilleland). For SAL and Neighborhood methods comparisons are made with alternative packages -> bug fixing of SpatialVx (scripts available in WG5 code repository)
- VAST software development (for neighborhood methods only): *inclusion of time dimension and the possibility to operate with other variables besides precipitation, primarily TCC*

New spatial methods



- Many ✓
 ✓
- **Neighborhood**
(Ebert, 2008)
 - **Scale Decomposition**
- DIST**

- **Features-based**
 - ✓ Contiguous Rain Area (CRA)
(Ebert and McBride, 2000)
 - ✓ Method for Object-based
Diagnostic Evaluation (MODE)
(Davis et al., 2006)
 - ✓ SAL technique
(Wernli et al., 2008)
- **Field Deformation**

Almost all the categories of spatial methods are applied by PPINSPECT participants.

Compact visualization of total precipitation FSS: to focus on the useful scale for a given lead-time and threshold (MCH)

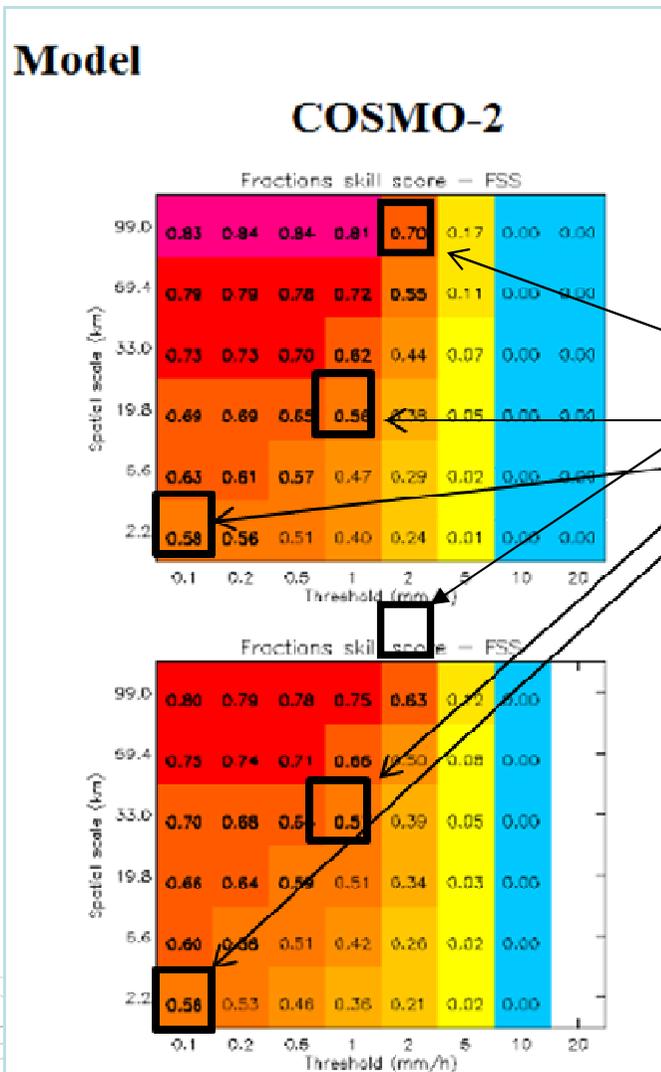
$$FSS = 1 - \frac{\frac{1}{N} \sum (P_f - P_o)^2}{\frac{1}{N} \left[\sum P_f^2 + \sum P_o^2 \right]}$$

FSS: Fractions Skill Score

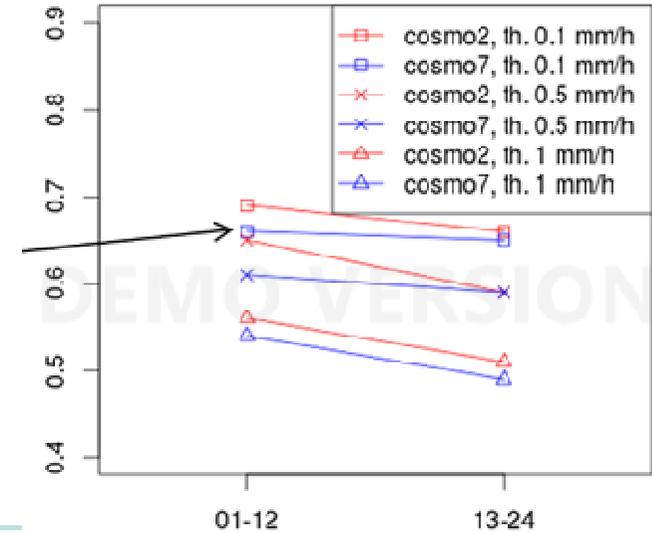
P is the event fraction in the neighborhood.

Lead time

MAM2015



Score as a function of leadtime for a single meaningful scale



Useful scale as a function of lead time

Threshold	0.1 mm/h	1mm/h	2 mm/h
01-12	2.2 km	19.8 km	59.4 km
13-24	2.2 km	33.0 km	99.0 km

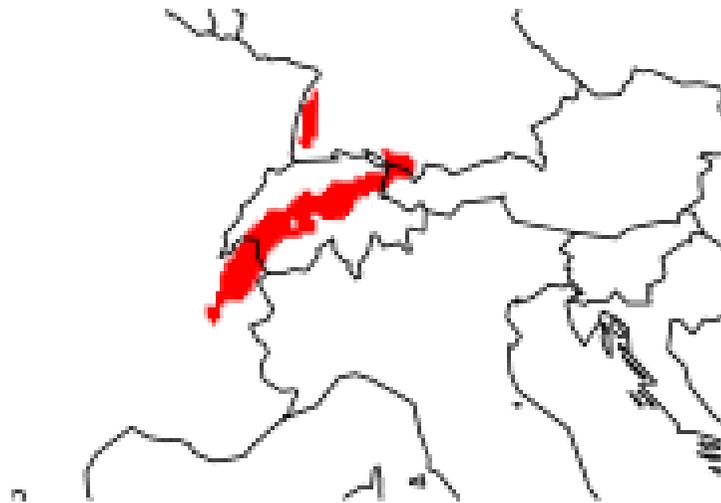
Object-based methods

- MODE, CRA, SAL

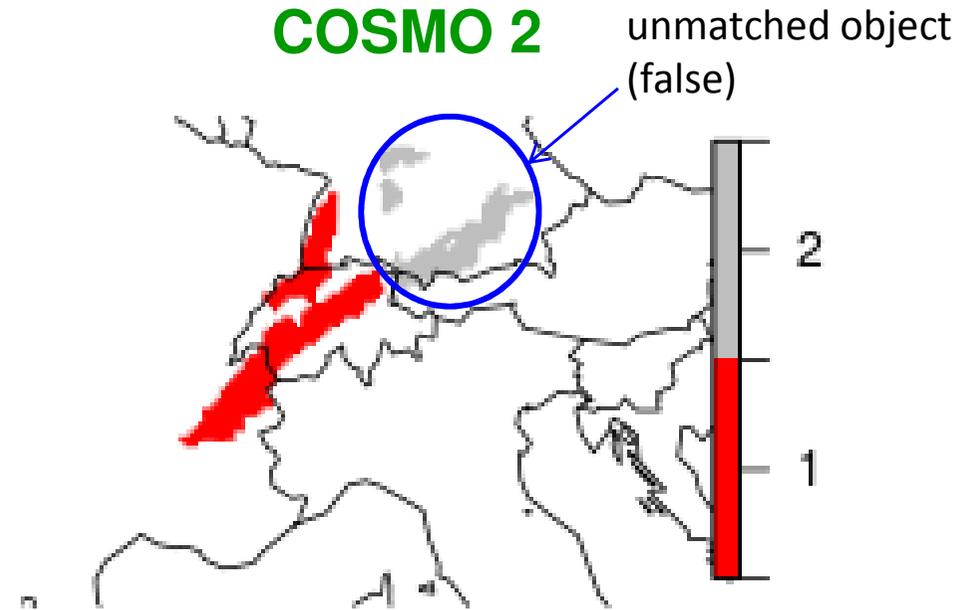
Case study

2007.09.25.06, 6h precipitation, threshold $\geq 5\text{mm}$

VERA



COSMO 2



Selected feature pairings based on total interest

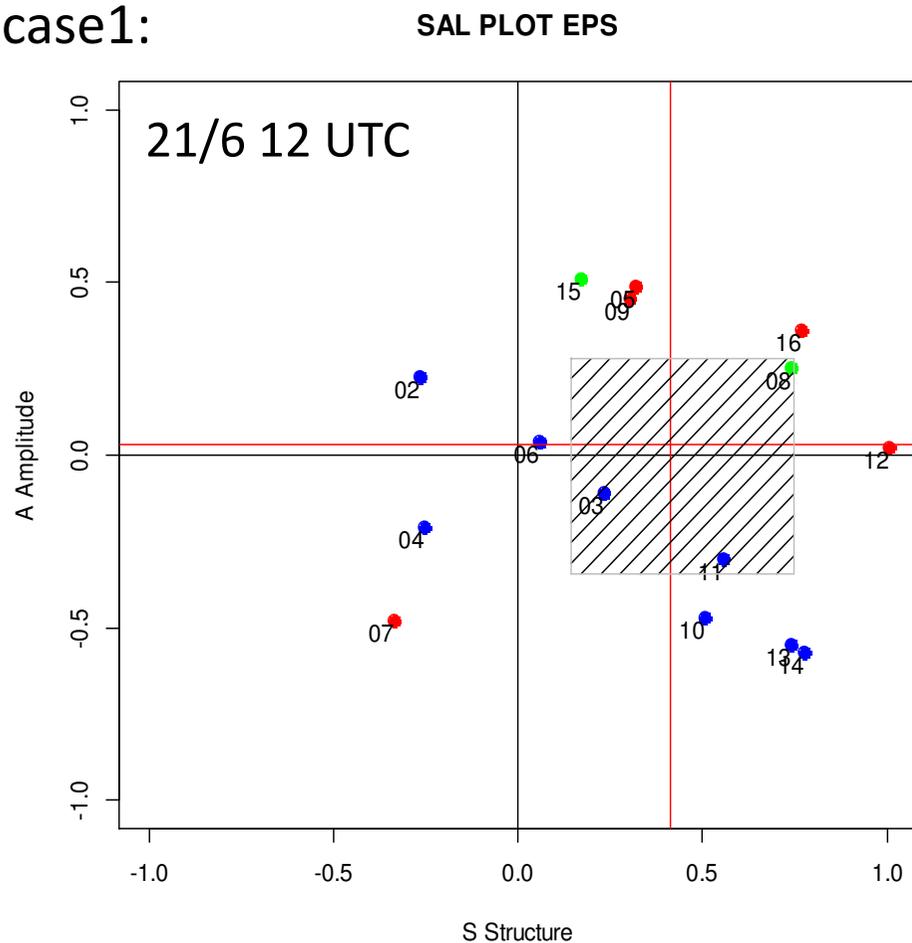
obs feature	mod feature	total interest
1	1	0.898

Methods of calculating SAL for EPS evaluation

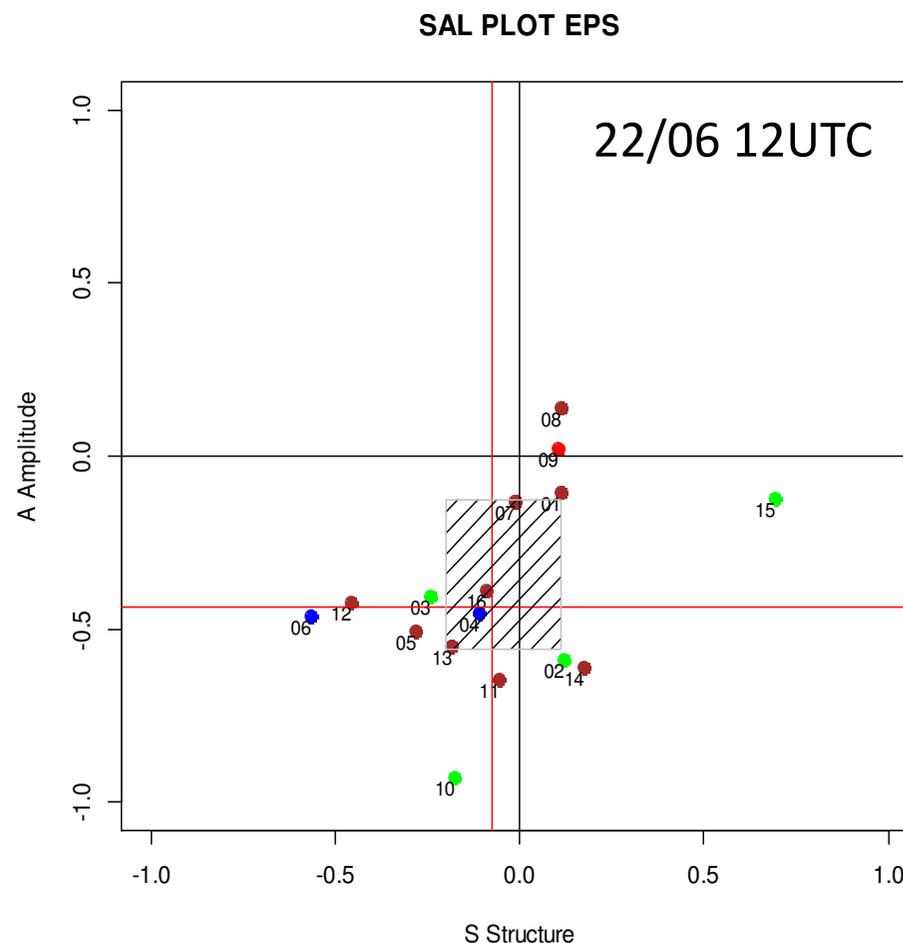
1) For one specific timestep:

SAL method can be applied in order to estimate the performance of an ensemble forecast (**Barrett et al. 2015**). Each point of the SAL plot represents one member, and an ensemble performance can be estimated.

An example case shown here used data for COSMO2 LEPS 16 members MesoVICT case1:



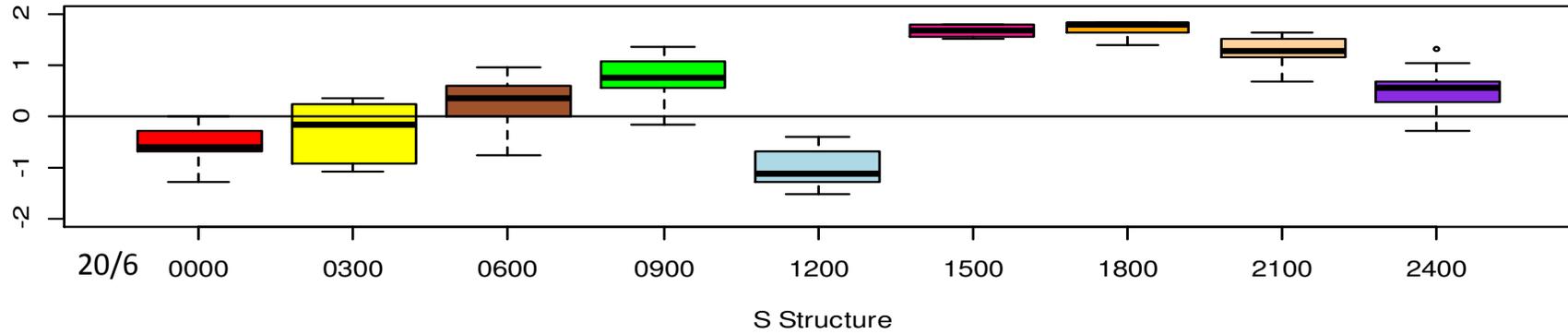
S > 0 larger objects predicted



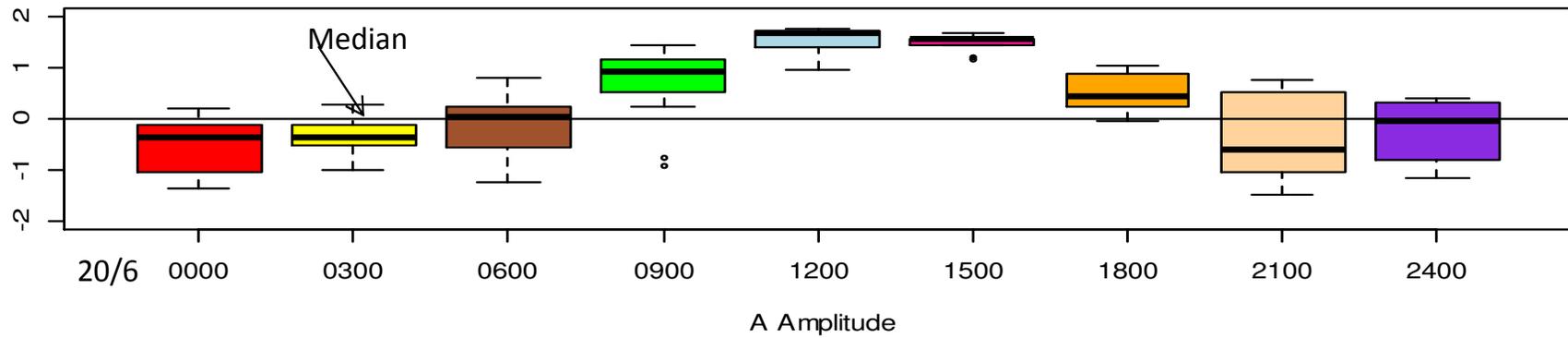
A < 0 domain values underestimated

2) For a series of time leads in one plot boxplots can be used

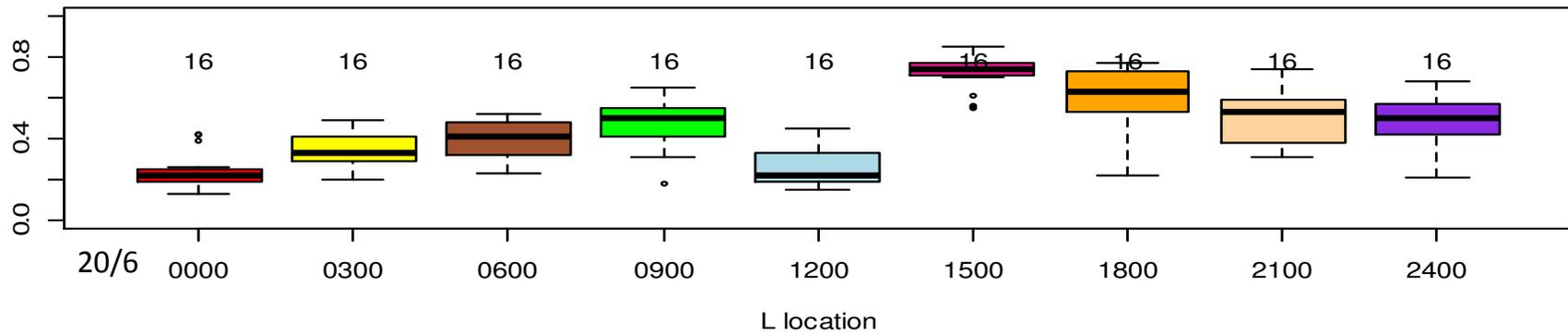
S EPS PLOT



A EPS PLOT

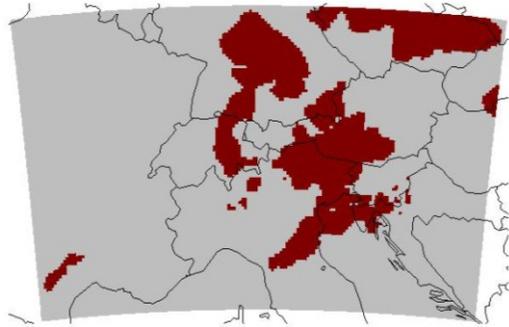


L EPS PLOT

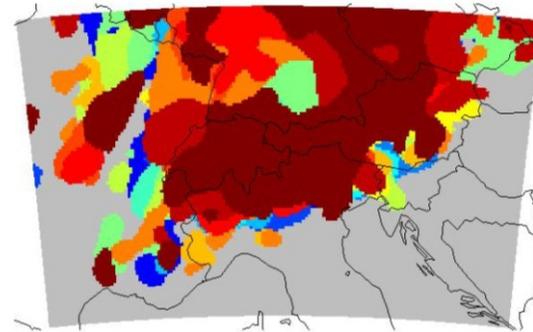


A new approach into investigation : EPS Probability Objects

- Without observation uncertainty



Observations



LEPS



Probability



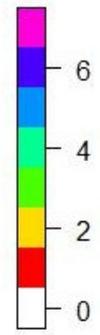
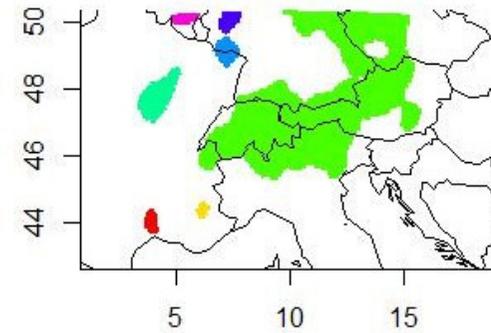
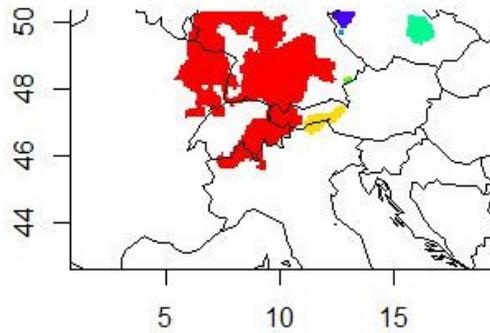
Probability

3 h Prec. 21/6 12 UTC

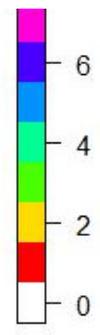
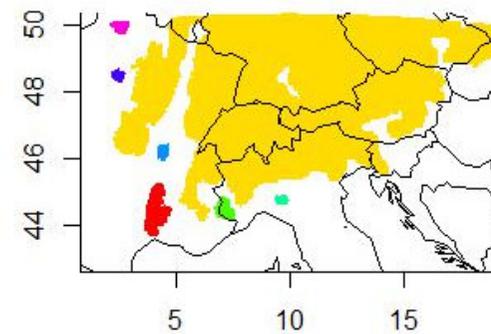
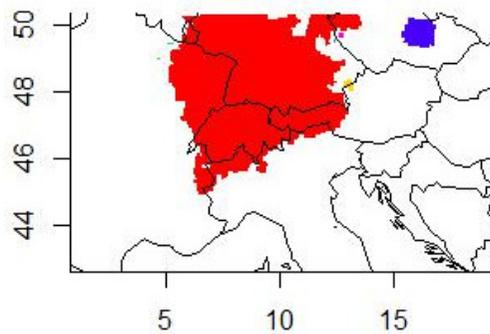
Constant threshold : Probability that precipitation ≥ 2 mm

- What is the fraction of model members that predicts precipitation ≥ 2 mm ?
- Observation certainty ($P=1$) compared with probability objects of 16 members exceeding the threshold (the brown objects)

Objects comparison for probability of precipitation $\geq 2\text{mm}$



Prob (Preci $\geq 2\text{mm}$) = 1 (All models and all VERA ens predict ≥ 0.2)
S=0.37, A=0.6, L=0.1

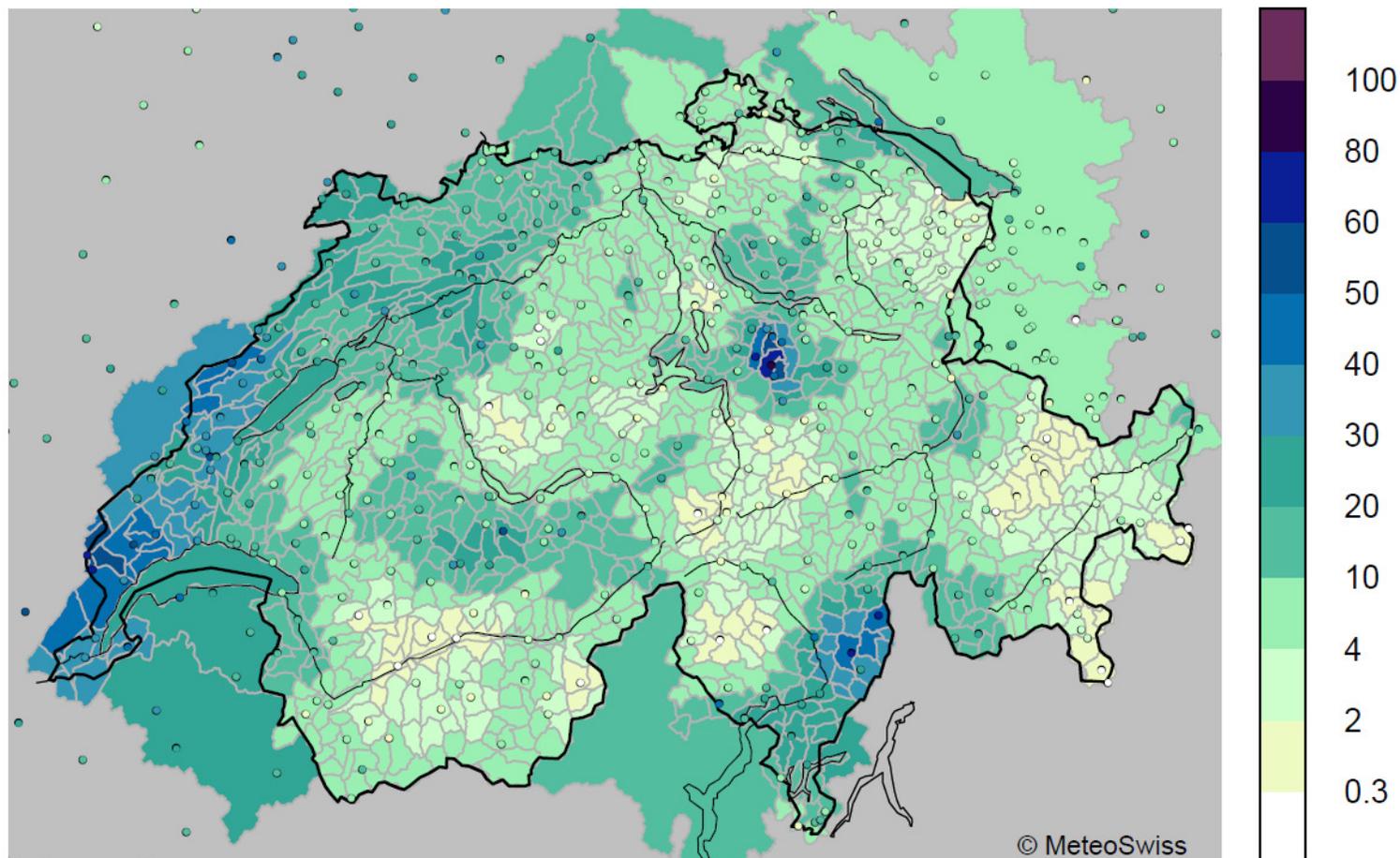


Prob (Preci $\geq 2\text{mm}$) ≥ 0.5 (At least half of the members)
S=0.65, A=0.6, L=0.05

New ensemble precipitation observation product by MeteoSwiss

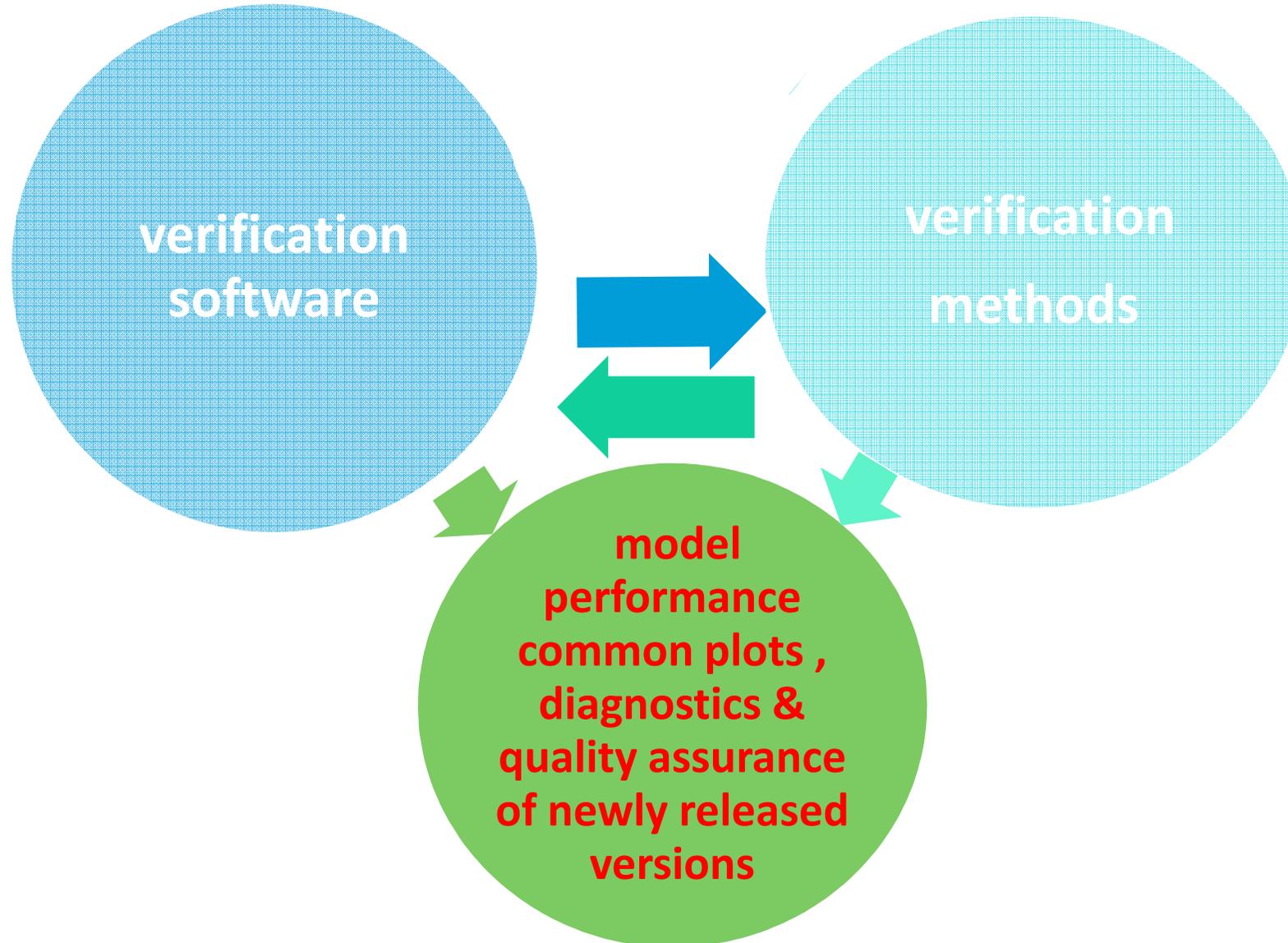
- Available for the past data (e.g. for Mesovict cases)
- Available for Swiss + whole alpine domain for daily accumulation

Precipitation (mm) 2007-06-20 (Ens. Median, basis)

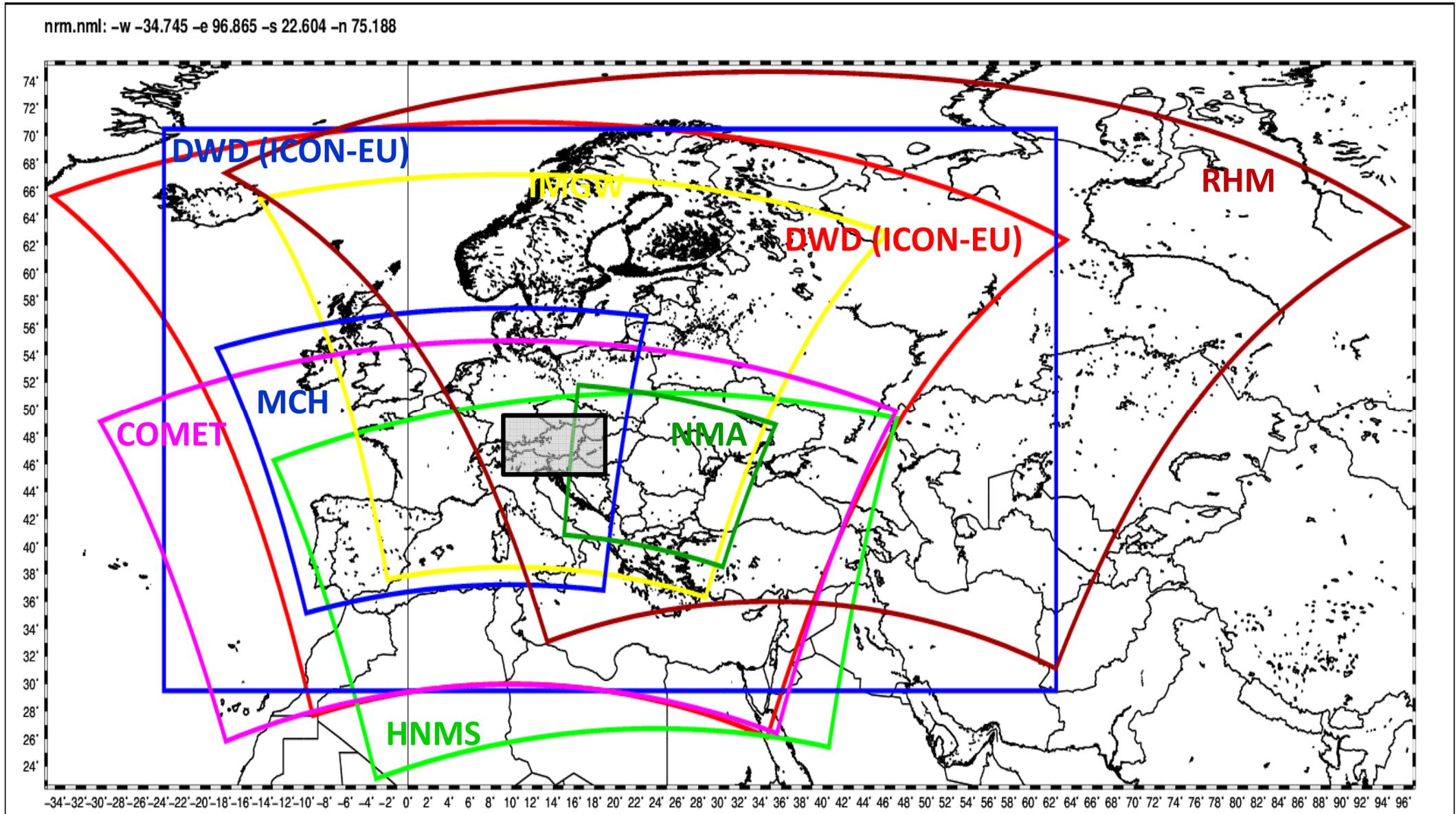


RhydchprobD v1.0, 2017-08-30 18:38

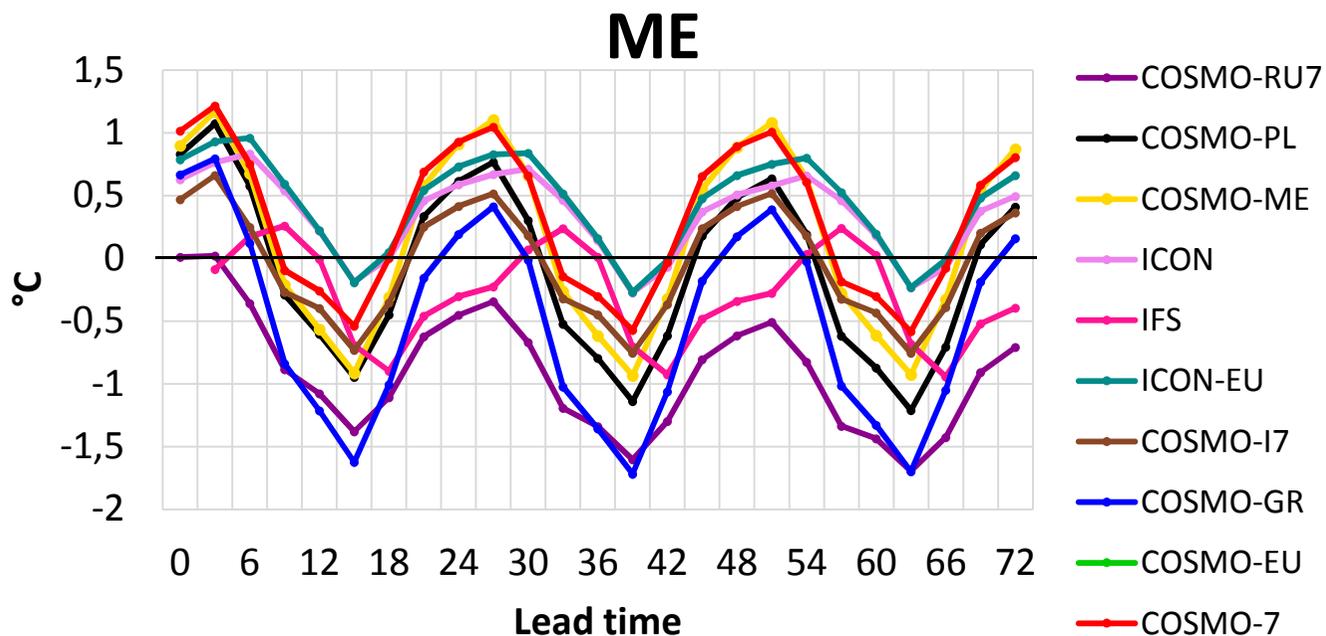
main activities



operational coarse res models

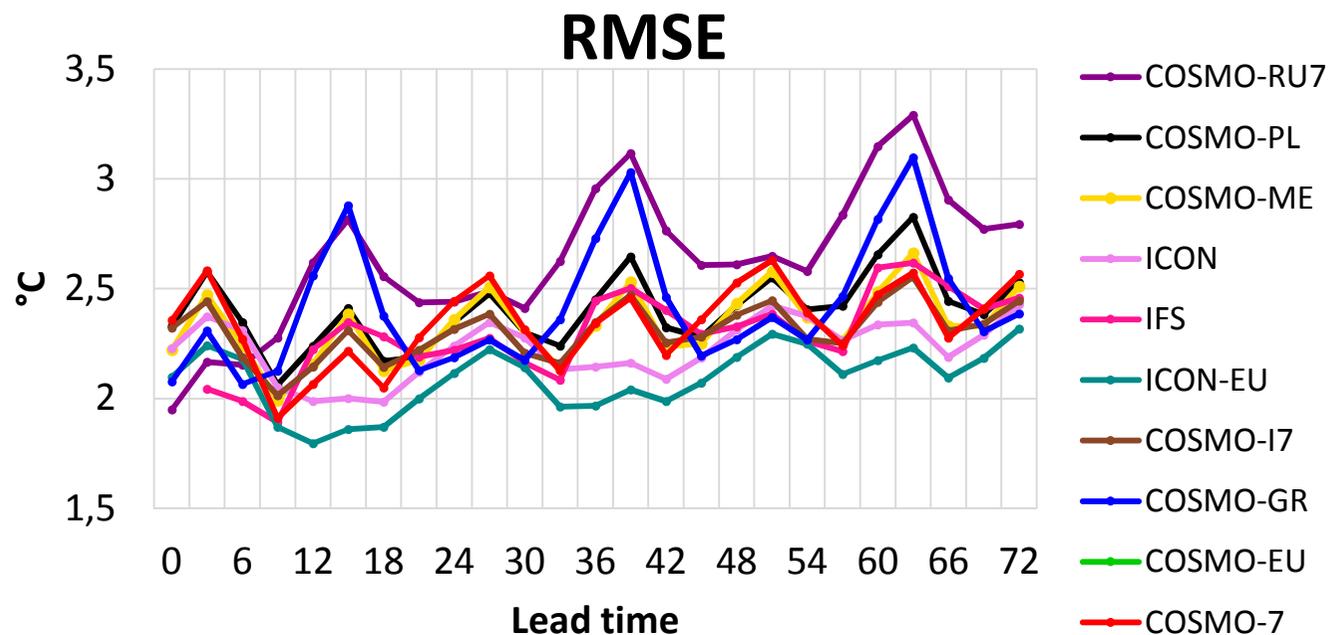


TEMPERATURE AT 2M From: 2016-06-01 To: 2017-05-31



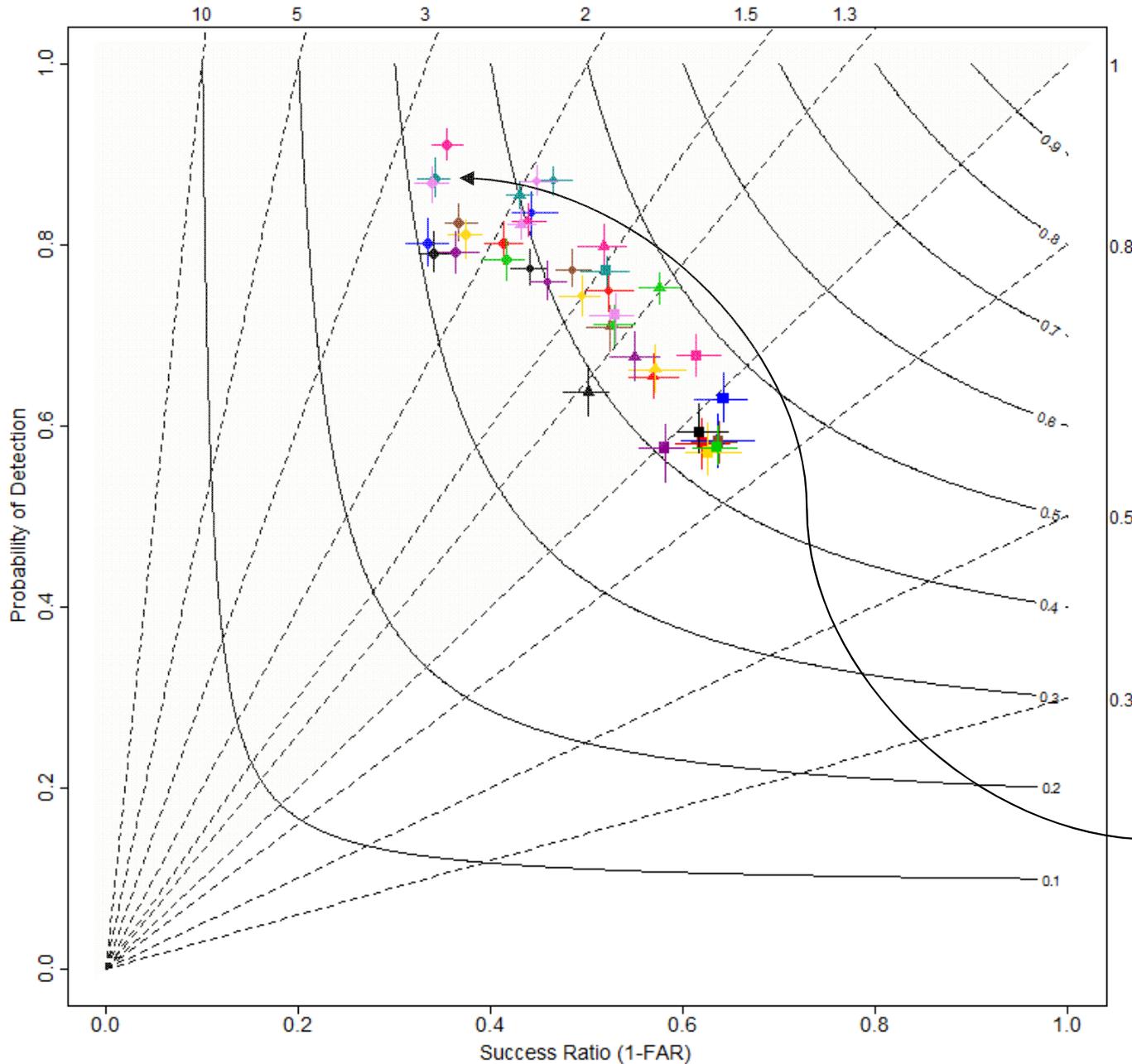
Diurnal variability underestimation:
negative ME during day, positive ME during night

ICON-EU/ICON:
the lowest RMSE and diurnal score variations



Last year tendency: JJA RMSE ↓ SON RMSE ↓ DJF RMSE ↑ MAM RMSE ↑

Precipitation in 6h 0.2 mm threshold From: 2016-06-01 To: 2016-08-31

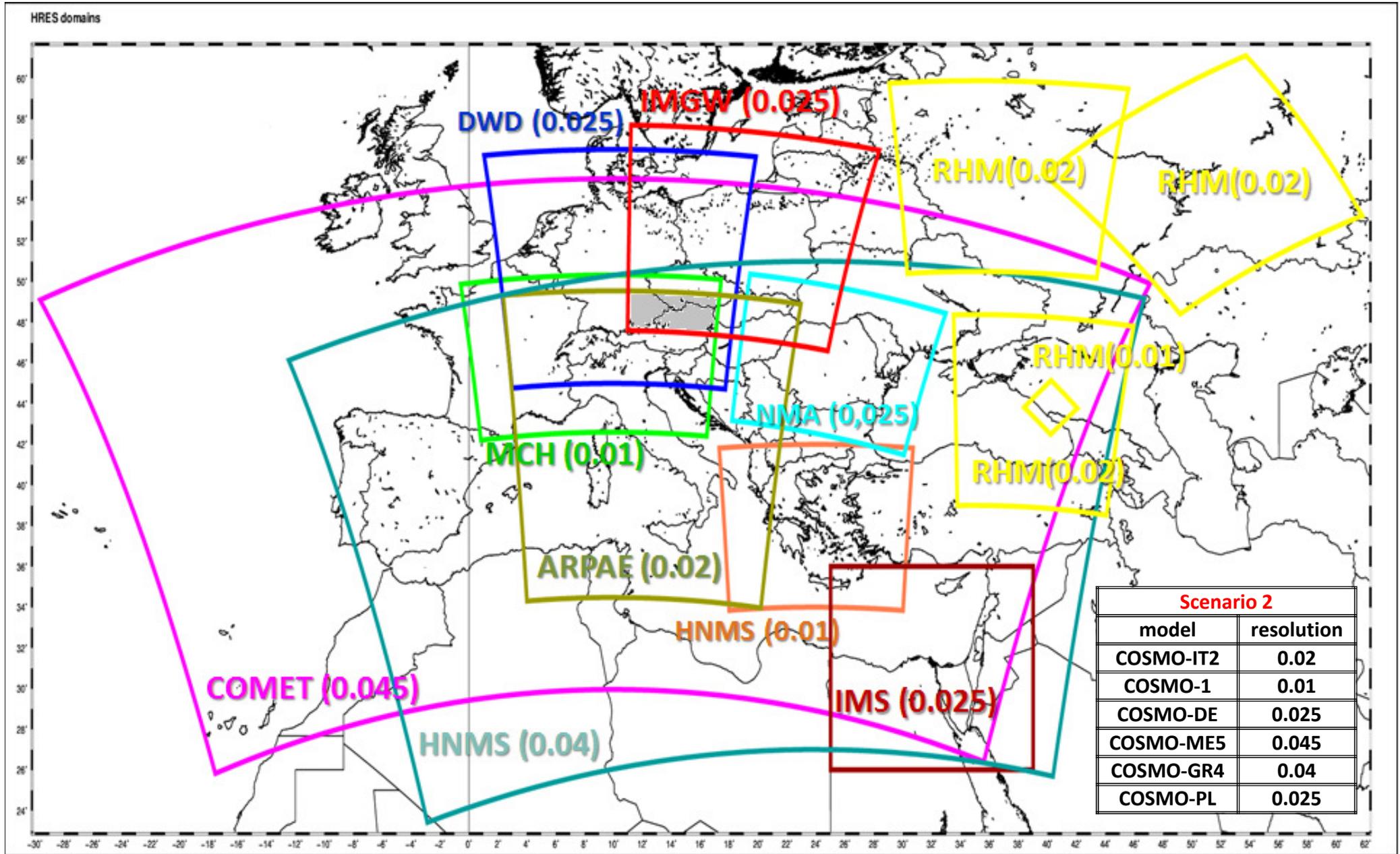


The most significant diurnal variation and the **lowest** Threat Score compared to other seasons

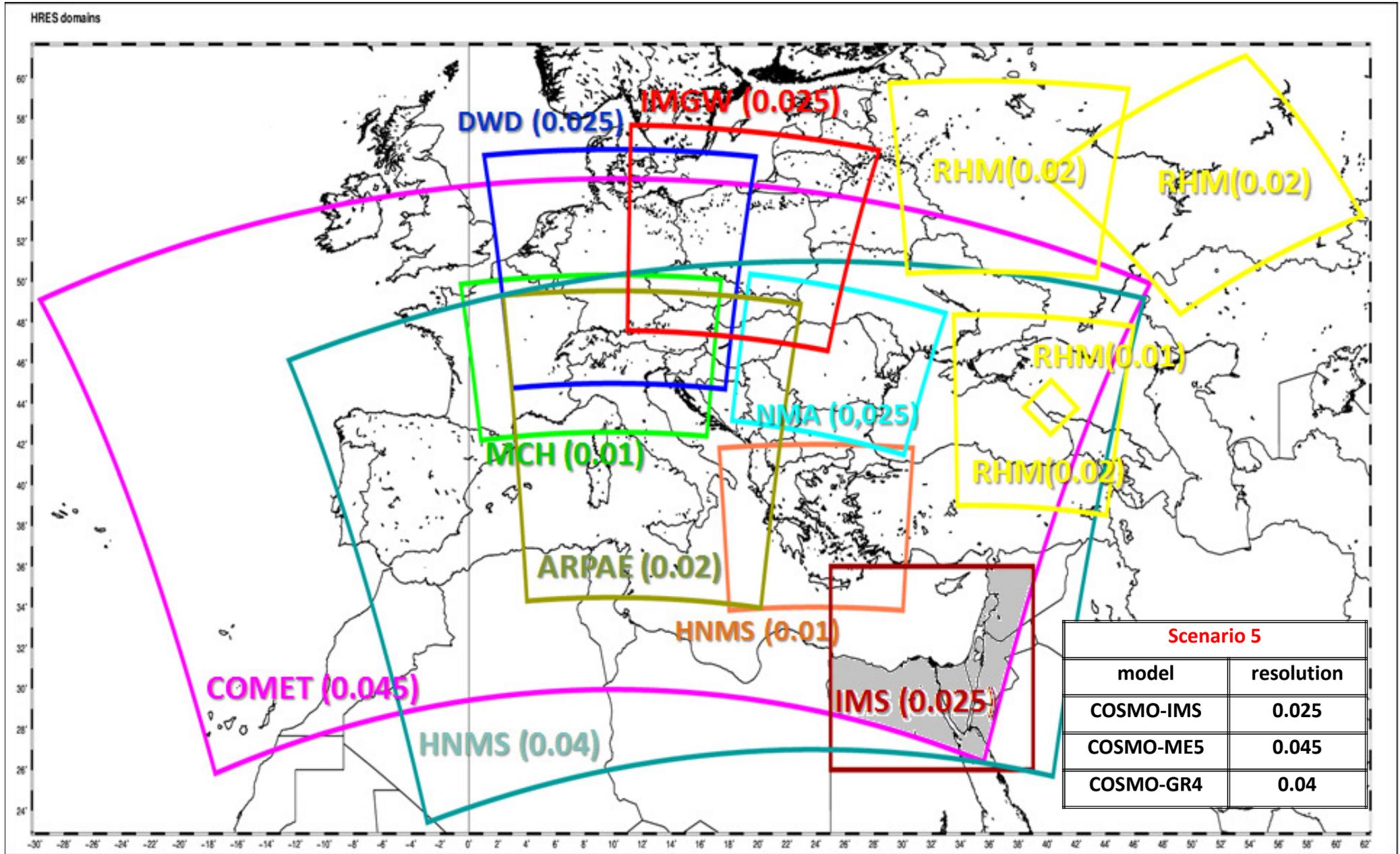
ICON, ICON-EU, and IFS overestimate low precipitation more than **COSMO** models: Presumably, IFS/ICON convection scheme impact?

- ▲ +06 lead time ■ COSMO-RU7 ■ COSMO-PL
- +12 lead time ■ COSMO-ME ■ ICON
- ◆ +18 lead time ■ IFS ■ ICON-EU
- +24 lead time ■ COSMO-I7 ■ COSMO-GR
- COSMO-EU ■ COSMO-7

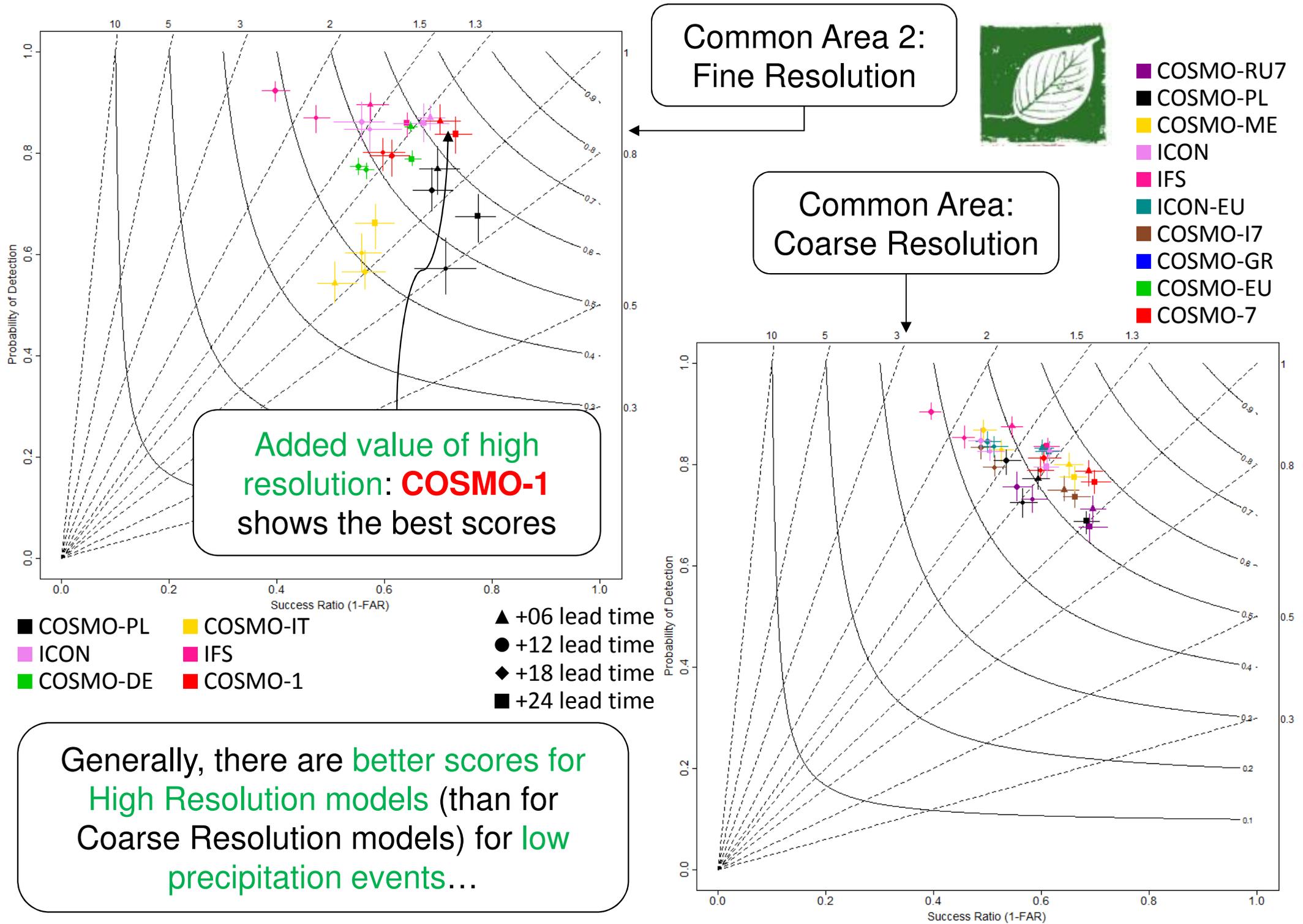
CP HRES scenario 2



CP HRES scenario 5

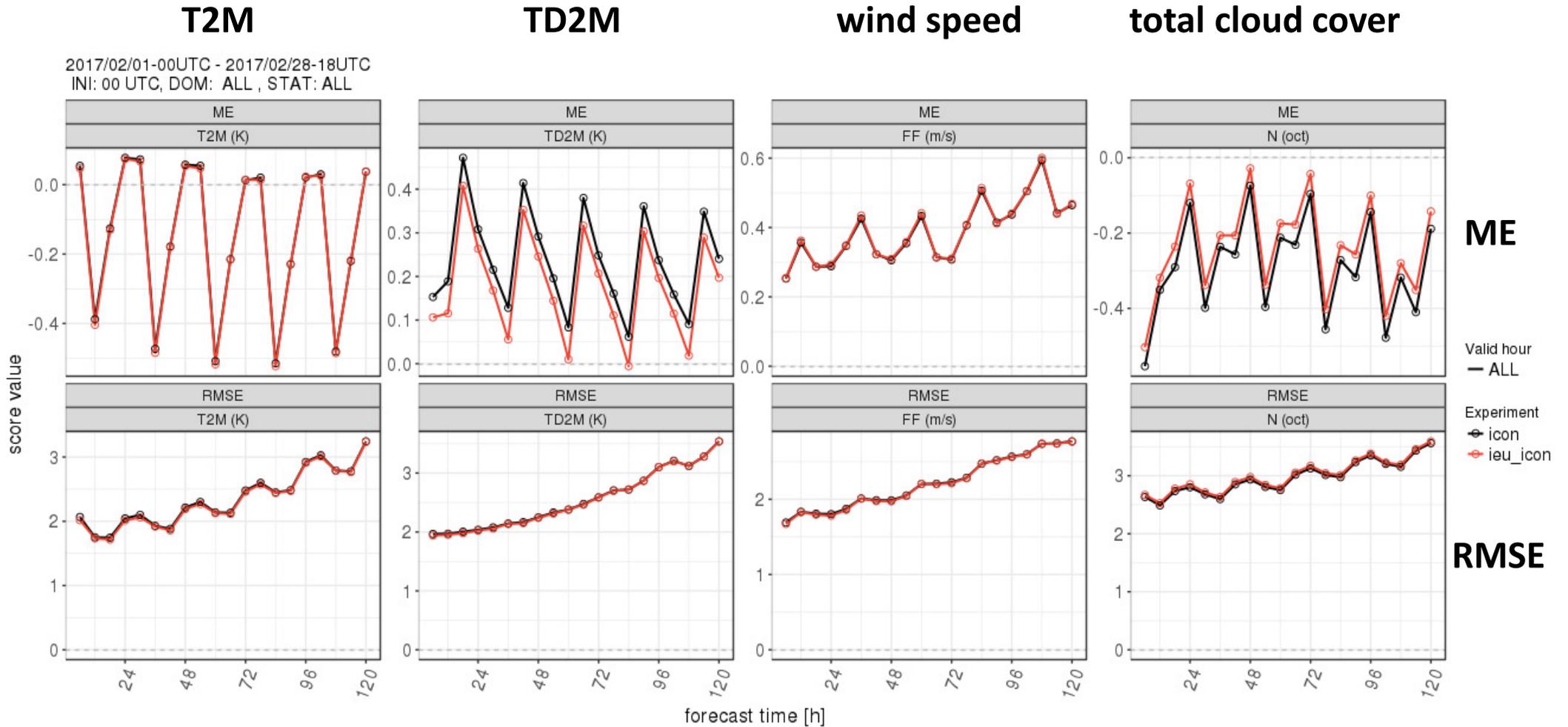


Precipitation in 6h 0.2 mm threshold From: 2017-03-01 To: 2017-05-31



ICON-EU vs ICON

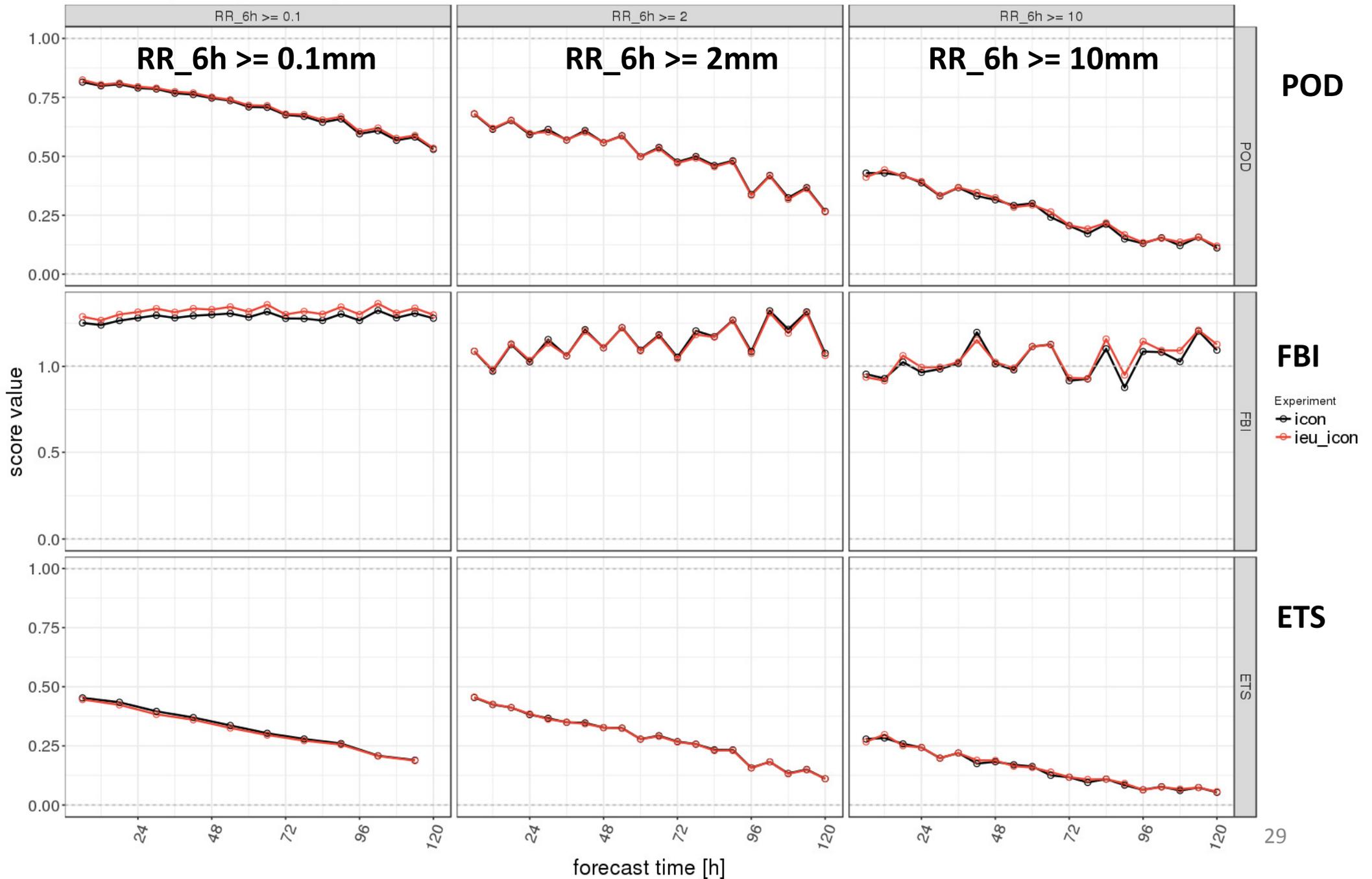
00 UTC runs, continuous verification, SYNOP, Feb 2017



ICON-EU vs ICON

All runs, categorical verification, SYNOP, Feb 2017

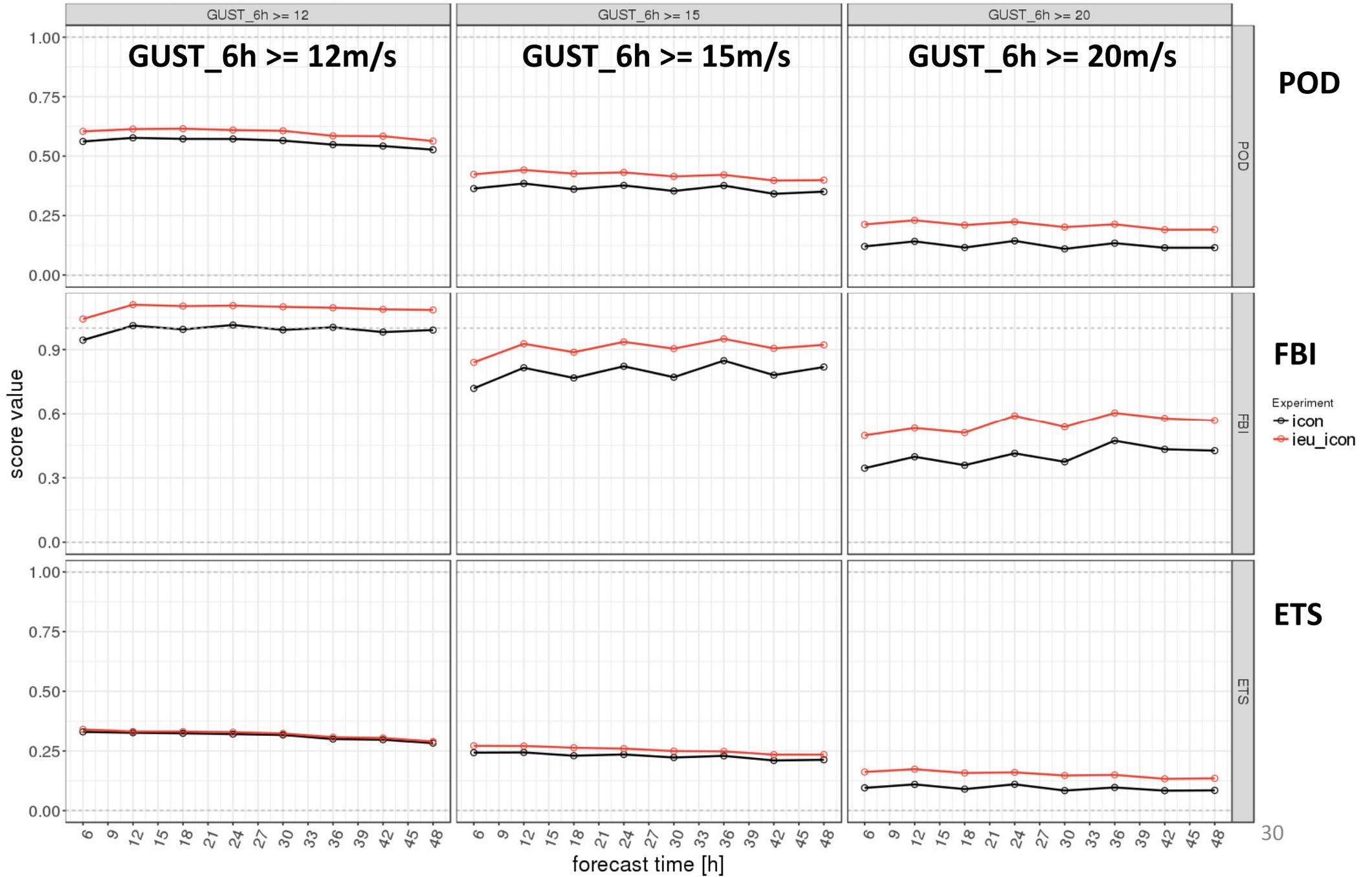
2017.02.01-00UTC - 2017.02.28-18UTC
VAL: ALL UTC, INI: ALL, STAT: ALL, DOM: ALL



ICON-EU vs ICON

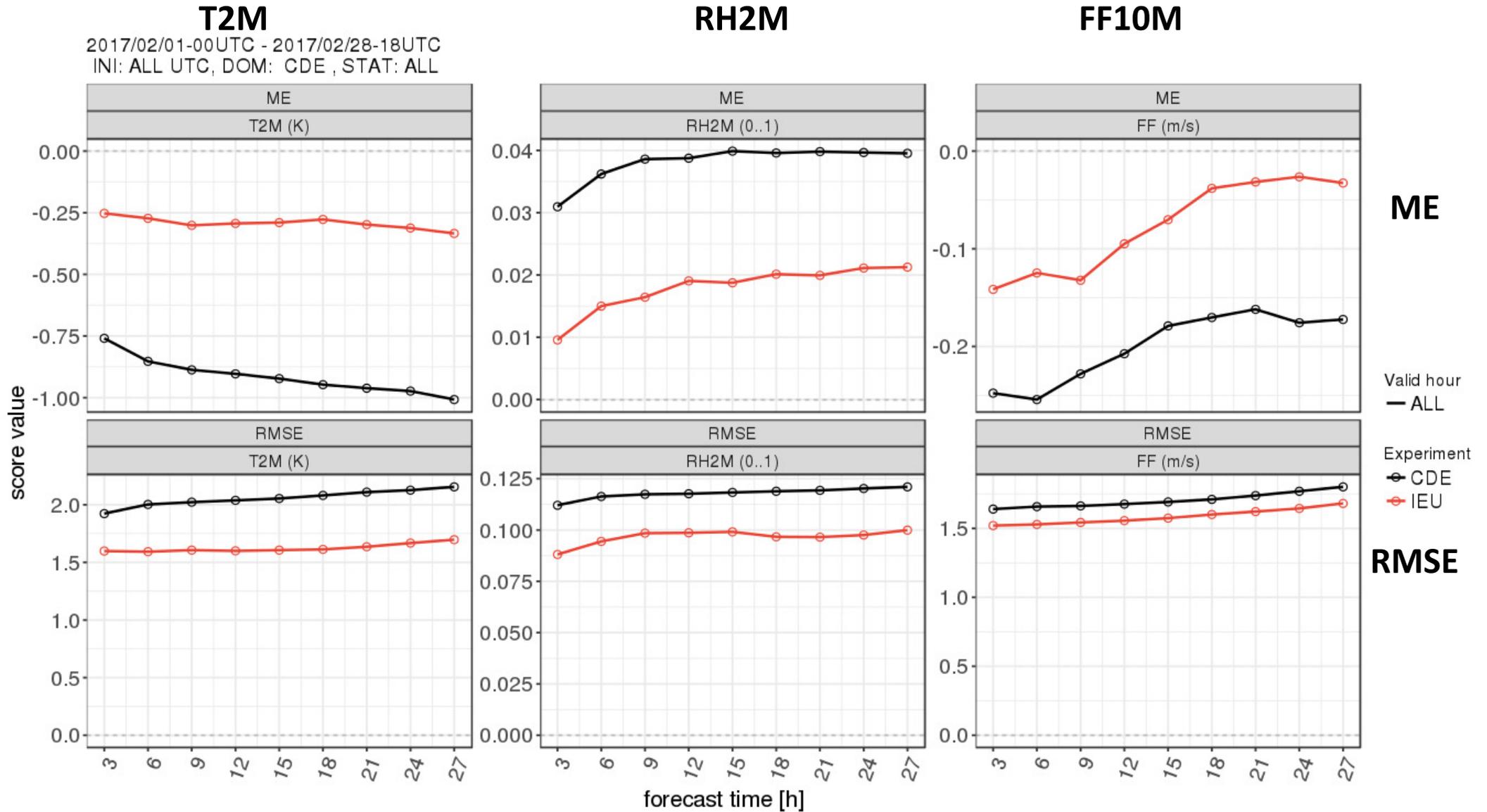
All runs, categorical verification, SYNOP, July 2017

2017.07.01-00UTC - 2017.07.31-18UTC
VAL: ALL UTC, INI: ALL, STAT: ALL, DOM: ALL



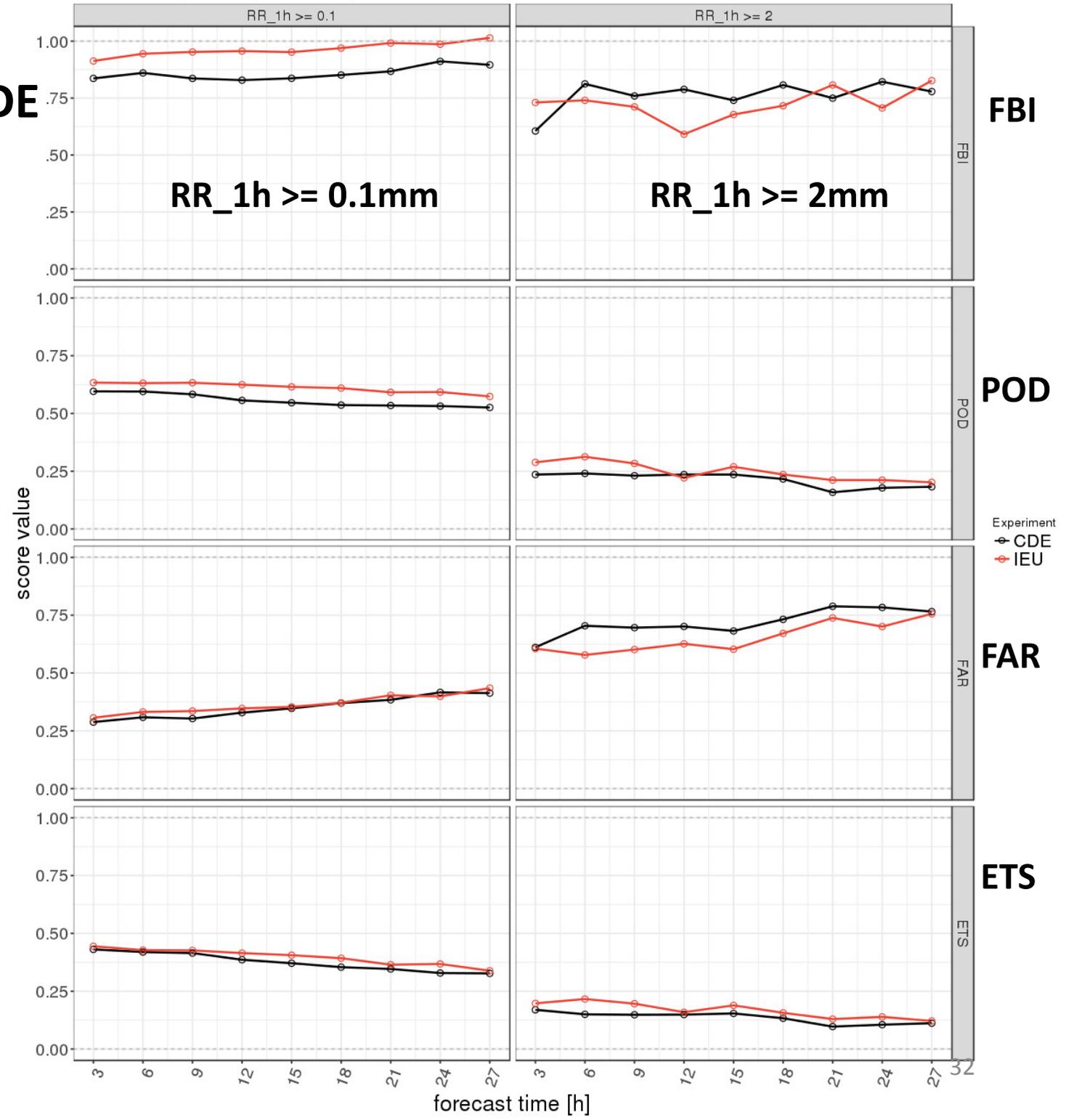
ICON-EU vs COSMO-DE

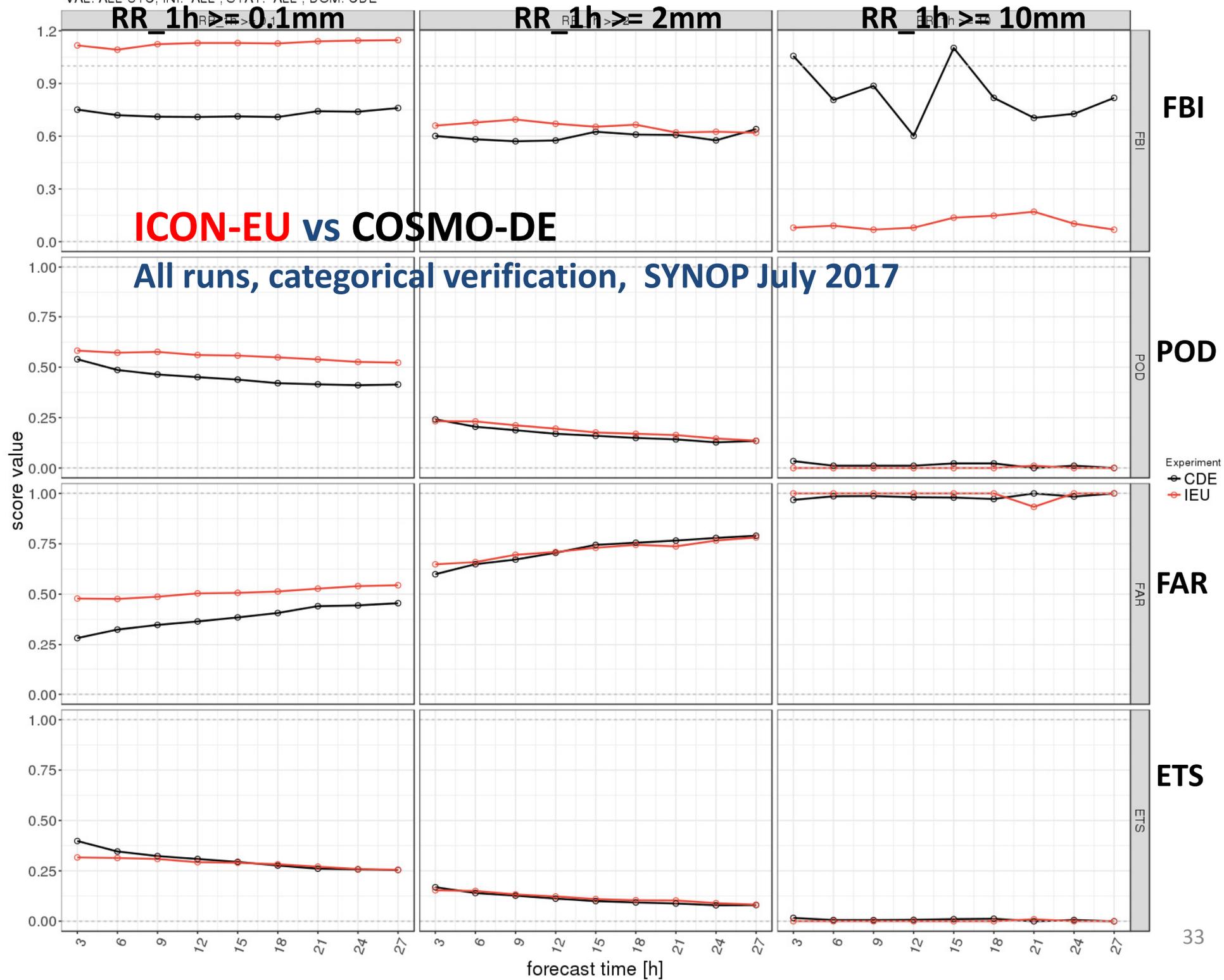
All runs, continues verification, SYNOP Feb 2017



ICON-EU vs COSMO-DE
 all runs
 categorical verification
 SYNOP Feb 2017

2017.02.01-00UTC - 2017.02.28-18UTC
 VAL: ALL UTC, INI: ALL, STAT: ALL, DOM: CDE





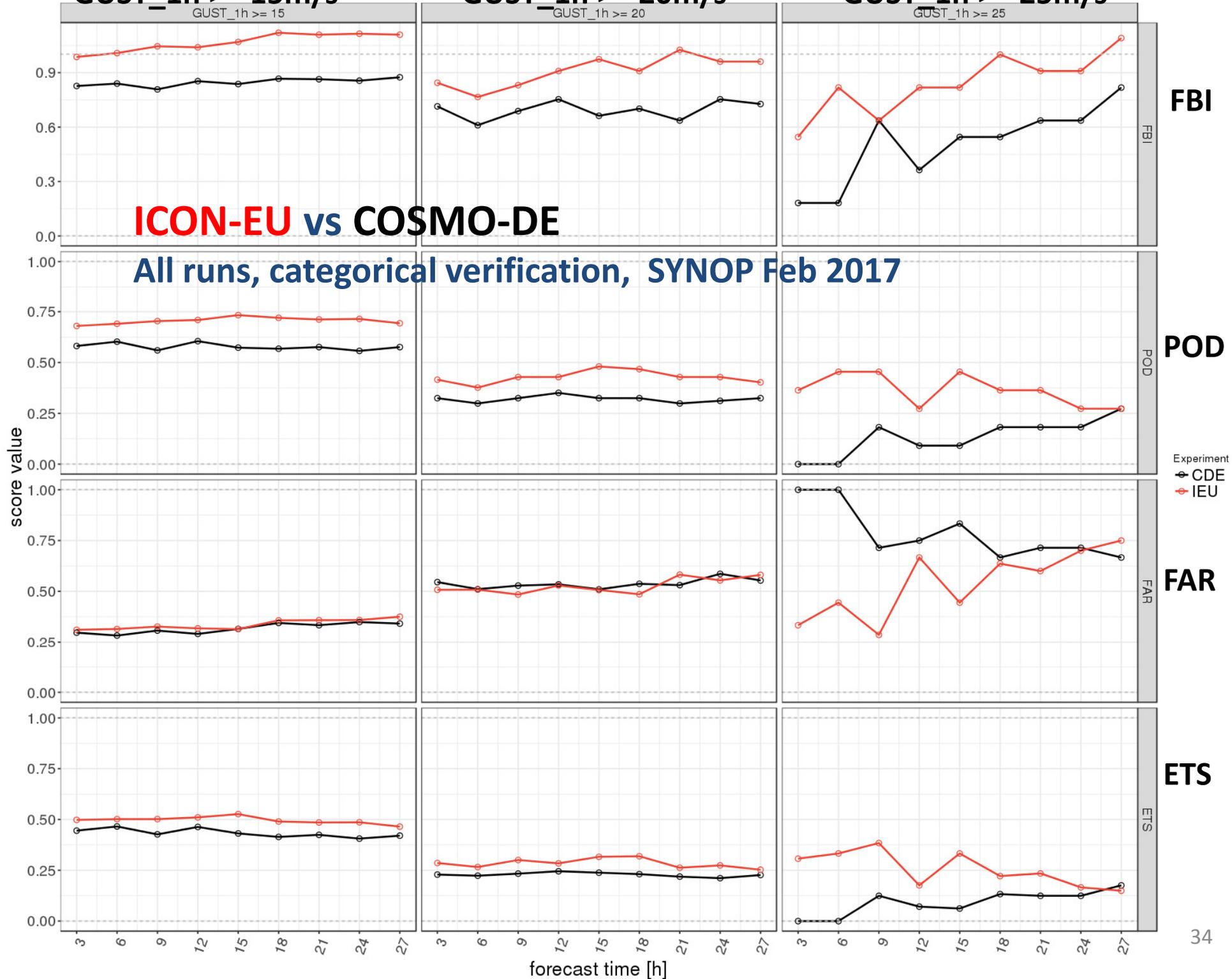
GUST_1h >= 15m/s

GUST_1h >= 20m/s

GUST_1h >= 25m/s

ICON-EU vs COSMO-DE

All runs, categorical verification, SYNOP Feb 2017



contributing team

Dimitra Boucouvala, HNMS

Roberto Bove, COMET

Anastasia Bundel, RHM

Rodica Dumitrache, NMA

Felix Fundel, DWD

Flora Gofa, HNMS

Amalia Iriza, NMA

Pirmin Kaufmann, MCH

Alexander Kirsanov, RHM

Xavier Lapillonne, MCH

Joanna Linkowska, IMGW

Elena Oberto, ARPA-PT

Ulrich Pflüger, DWD

Maria Stefania Tesini, ARPAE

Naima Vela, ARPA-PT

Antonio Vocino, COMET

Yiftach Ziv, IMS

Ευχαριστώ

Thank you