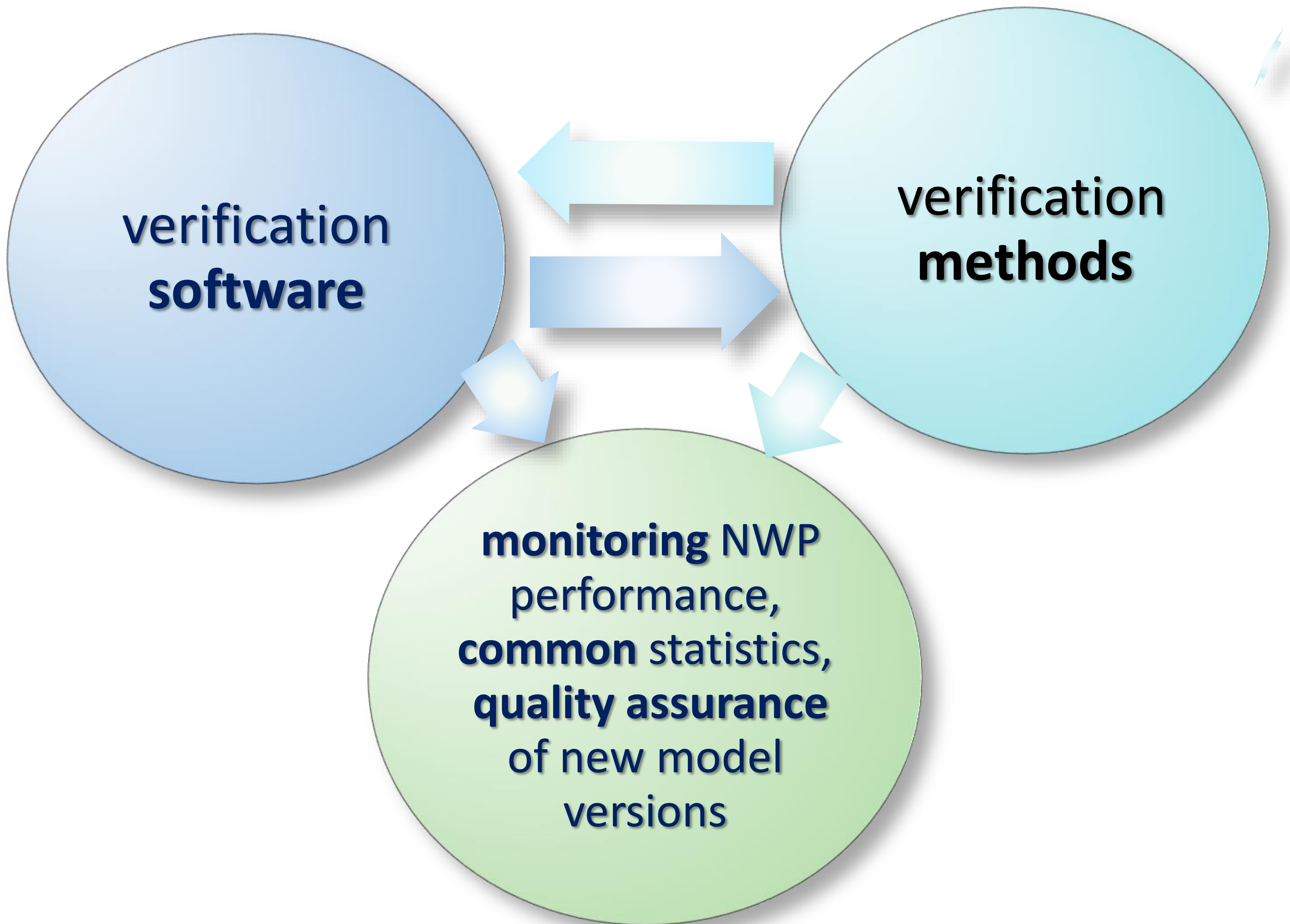
Overview of verification activities in COSMO

Flora Gofa

F. Batignani (CoMET), D. Boucouvala (HNMS), A. Bundel (RHM), R. Dumitrache (NMA), F. Fundel (DWD), A. Iriza-Burca (NMA), P. Kaufmann (MCH), A. Kirsanov (RHM), X. Lapillonne (MCH), J. Linkowska (IMGW), B. Maco (NMA), U. Pflüger (DWD), M. S. Tesini (ARPAE), N. Vela (ARPA-PT), A. Shtivelman (IMS)

41st EWGLAM and 26th SRNWP Meeting, Sofia, 30 Sep-03 Oct 2019

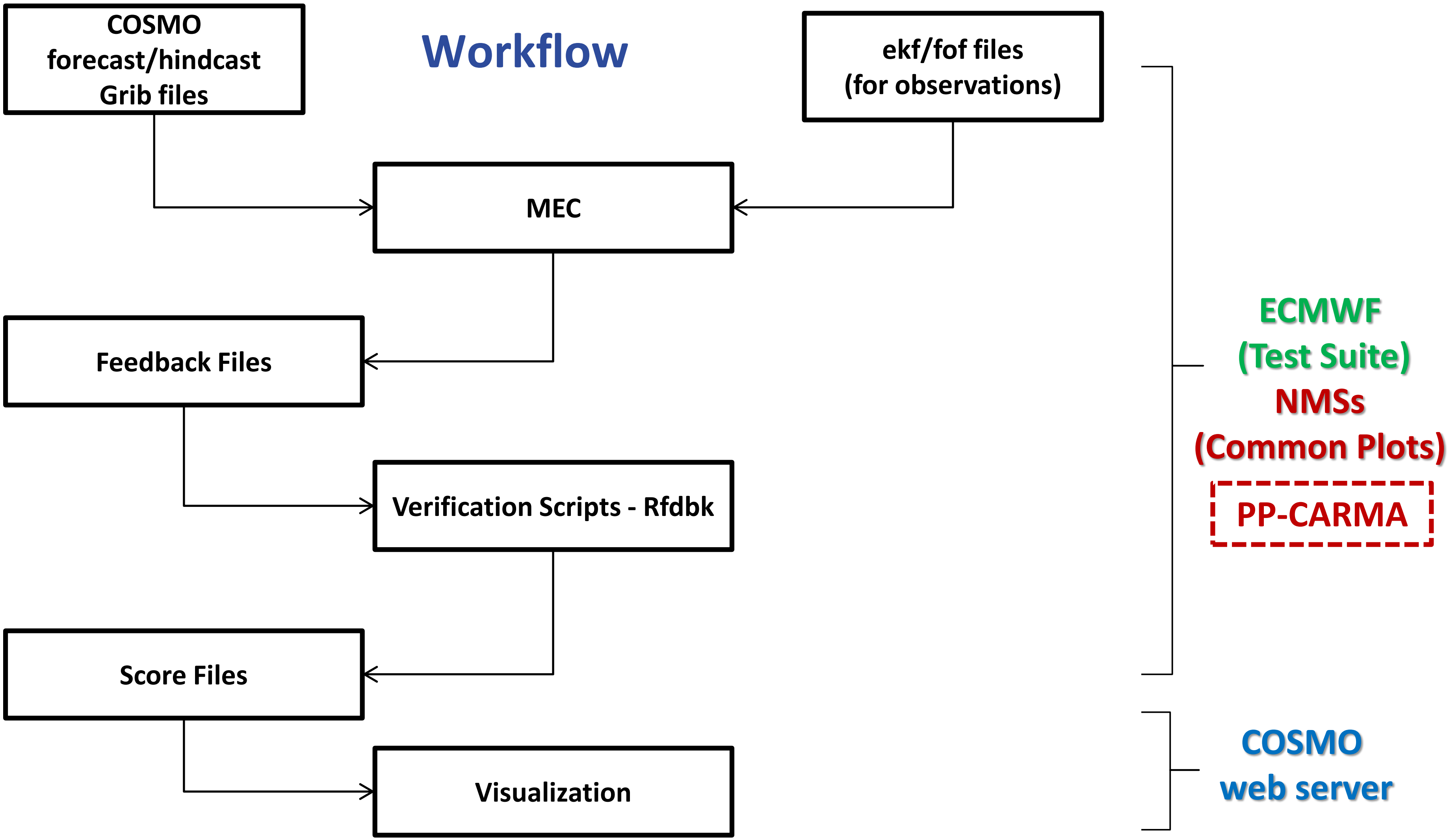
main activities



Feedback File based Verification

Felix Fundel
Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Workflow



Feedback File Verification: Namelist

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

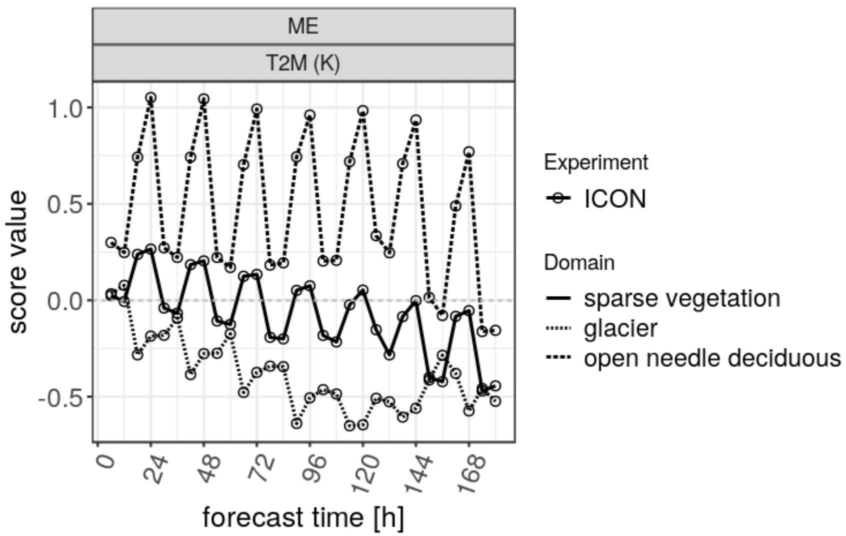
Optional namelist options

NAME	*VALUE*	*DESCRIPTION*
IdentList	'/path/to/your/identlist'	# Use only station(s) given in list file (integer)
statidList	'/path/to/your/statidlist'	# Use only station(s) given in list file ()
dateList	'/path/to/your/datelist'	# Verify dates in list separately (YYYYMMDD)
lonlims/latlims	'0,30'	# Restrict verification domain (faster)
iniTimes	'0,12'	# Use only runs given in argument
inclEnsMean	'TRUE'	# Include EPS mean in det. verification
mimicVersus	\`TRUE\`	# Uses VERSUS quality check only
sigTest	'TRUE'	# Perform sign. test on differences in score mean
conditionN	\`*R code defining condition*\`	# Perform conditional verification
alignObs	\`TRUE\`| \`FALSE\`|\`REDUCED\`	# full/no/reduced observation alignment
insType	\`1,2,3..\`	# Select txpe of instrument

Essentially any observation/forecast characteristic contained in feedback files (~50) can be used to refine the verification via namelist.

II Feedback File Verification

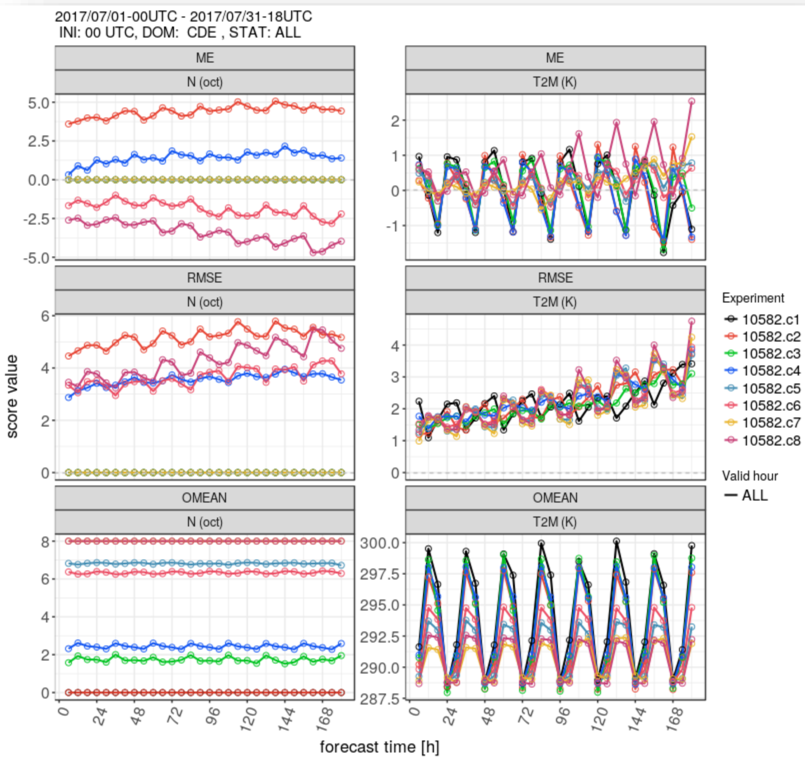
- User defined stratification of the verification domain
- Station or polygon based
- Initiated via namelist
- ASCII File with domain specification has to be provided by the user
- Only condition: Domains must not overlap!



\# Example polygon domain table
name lon lat
NORD 8 50.001
NORD 15 50.001
NORD 15 55
NORD 8 55
SUED 8 45
SUED 15 45
SUED 15 50
SUED 8 50

\# Example station domain table
name id DE Q887
DE 10837
DE 10184
CH 06670
CH 06612
CH 06610

Conditional Verification



- Using observation properties to define conditions
- Several properties can be combined
- Arbitrary number of conditions is possible
- Conditions are set in namelist
- Model name is extended by number of class
- Stations that do not report an observation used in a condition statement are not used

Example namelist

```
condition1    "list(N='obs==0',N='abs(veri_data-obs)<1') "
condition2    "list(N='obs==0',N='abs(veri_data-obs)>=1') "
condition3    "list(N='obs%between%c(1,4)',N='abs(veri_data-obs)<1') "
condition4    "list(N='obs%between%c(1,4)',N='abs(veri_data-obs)>=1') "
condition5    "list(N='obs%between%c(5,7)',N='abs(veri_data-obs)<1') "
condition6    "list(N='obs%between%c(5,7)',N='abs(veri_data-obs)>=1') "
condition7    "list(N='obs==8',N='abs(veri_data-obs)<1') "
condition8    "list(N='obs==8',N='abs(veri_data-obs)>=1')"
```

With this implementation conditions need to relate to the observation (i.e. not possible is lon%between%c(0,20))

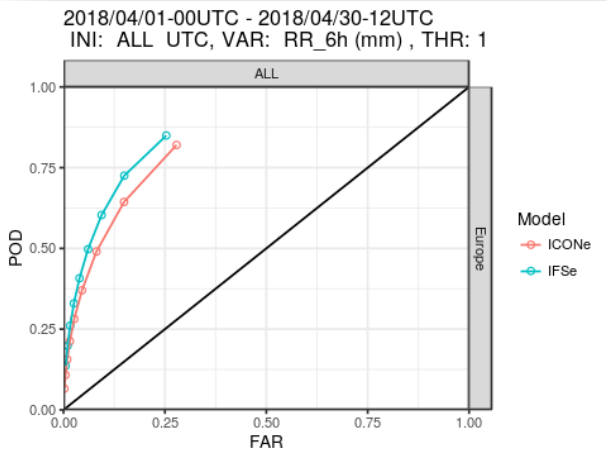
Feedback File Verification - EPS

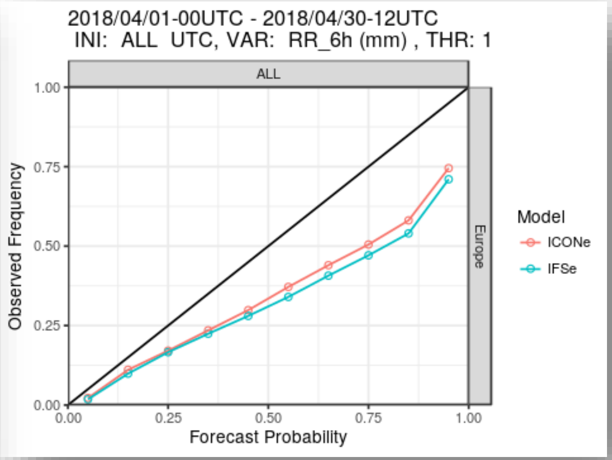
Based on „probability files“ that need to be produced from feedback files that hold information on probability of an EPS forecast to exceed a threshold

Arbitrary probability files can be aggregated to calculate e.g. Brier Scores (and decomposition), ROC-Area, ROC curve, reliability diagram

Revised EPS Verification

- Now: keeping only domain average scores as in deterministic verification
- Additional efficiency plus from fdbk_wide function in Rfdbk
- Time series and significance test for ensemble scores are now possible
- Low memory usage allows for high degree of parallelization
- Verification results in a single score file, and one app was written to show ensemble (e.g. CRPS) and probabilistic (e.g. ROC) scores
- **All verification scripts can now be run on multiple cores**

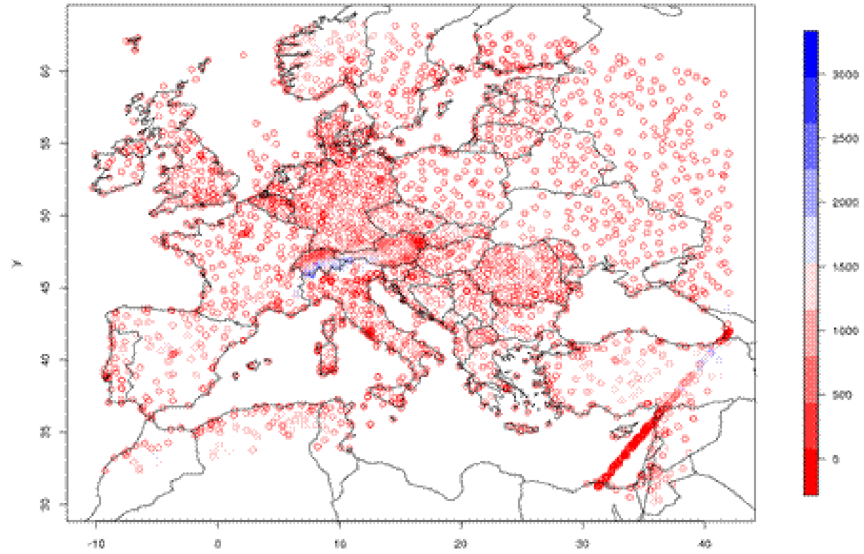


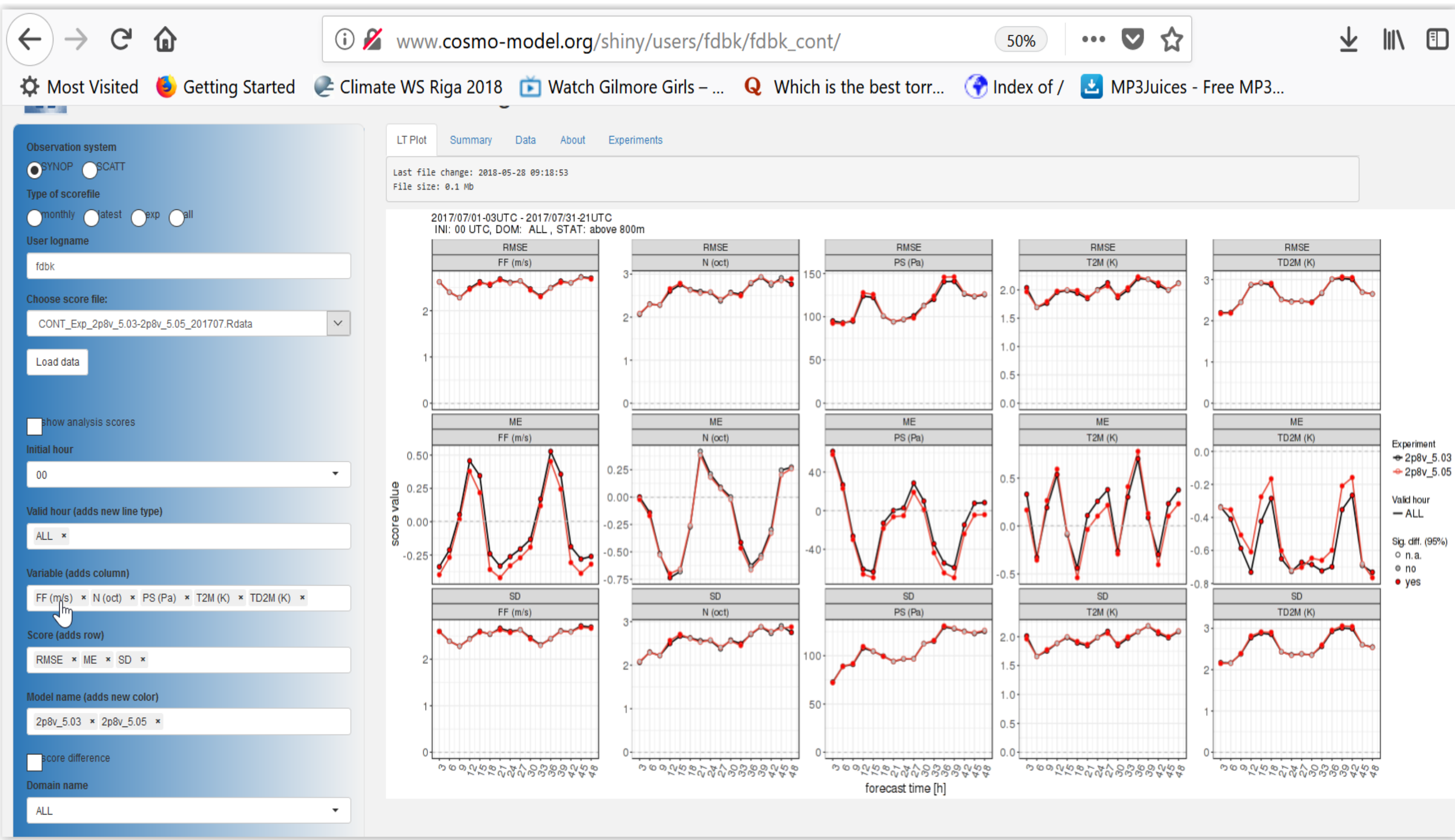


NWP Meteorological Test Suite @ ECMWF

Hindcast mode experiments
Winter and summer period
Coarse and high resolution model implementations

http://www.cosmo-model.org/shiny/users/fdbk/





Feedback File Verification

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

DWD

http://www.cosmo-model.org/shiny/users/fdbk/RfdbkVeriDoku.html

www.cosmo-model.org/shiny/users/fdbk/RfdbkVeriDoku.html

Feedback File Verification Suite at DWD

Felix Fundel

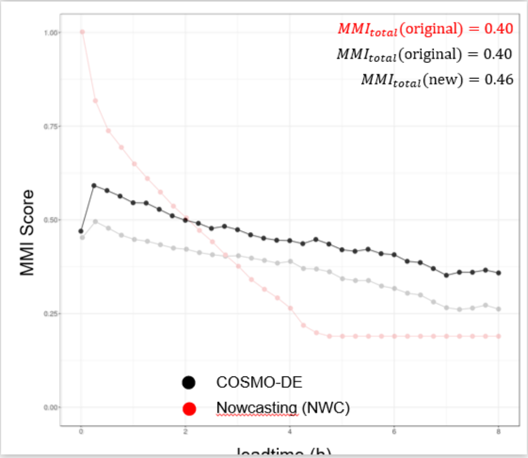
13 Juni, 2019

Introduction

Spatial Verification Efforts for SINFONY

- **Review of existing neighborhood/spatial verification methods for deterministic and ensemble forecasts**
  - Deterministic
    - Neighborhood: methods & scores from Ebert 2008
    - Object-based: Focus on Total Interest (TI) & Median of Maximum Interest (MMI)
    - Object: Konrad3D objects
  - Ensemble
    - Neighborhood.:NEP (Schwartz et al. 2010); NEP + time fuzzyness (Duc et al. 2012,2013)
    - Object: Konrad3D, clustering of EPS objects

- **Developing R functions (eventually resulting in a package)**
  - Namelist control
  - Reading capability for most common data formats (grib, Rdata, binary (Radolan), XML/HDF5 (Konrad3D))
  - Aggregation functionality (important for routine verification)
  - Alignment observation/forecast data from different experiments
  - Interactive visualization of scores interactively (shiny-server)
  - *No pre-processing (e.g. regridding, restructuring) provided*



PP-AWARE project: Appraisal of "Challenging WeAther" FoREcasts

<u>Joint Project</u>: Verification and Case Studies (WG5) ***& Interpretation and Applications (WG4) with collaboration with Predictability and Ensemble Methods (WG7)***

The goal of the PP is to provide COSMO Community with an overview of forecast methods and forecast evaluation approaches that are linked to high impact weather (not necessarily considered extreme to all users).

WMO: HI-Weather definition of HIW

Extreme in amplitude (intense winds, or heavy convective precipitation)

Rare tail of climatological distribution for a particular location

Prolonged ‘regimes’ (droughts, heat-waves or cold-spells)

Challenging if society is vulnerable to them (e.g. impact of fog on transportation).

CONSORTIUM FOR SMALL SCALE MODELING COSMO

Proposed Tasks

Task 1: Challenges in observing CW/HIW (WG5 and WG4 related)

Question: How well high-impact weather is represented in the observations, including biases and random errors, and their sensitivity to observation density?

HIW phenomena studied: visibility range, thunderstorms (w. lightning), intense precipitation, extreme temperatures and winds.

Task 2: Overview of appropriate verification measures for HIW (WG5 related)

Question: How well high-impact weather forecast quality is represented with commonly used verification measures? What is the most appropriate verification approach? Extreme Value Theory application on HIW?

HIW phenomena studied: intense precipitation, thunderstorm (lightning activity, visibility range (fog).

Task 3: Verification applications (with a focus on spatial methods) to HIW (WG5 and WG7 related). This task will make use of the findings of Task 2 and is connected with and continued from PP-INSPECT and MesoVICT projects analysis of intense precipitation patterns.

Question: Can spatial verification methods contribute to the proper evaluation of HIW phenomena and in what way?

HIW phenomena studied: intense precipitation, thunderstorm (lightning activity LPI, visibility range (fog).

Task 4. Overview of forecast methods, representation and user-oriented products linked to HIW (WG4 related)

Question: How well is HIW is represented in postprocessing? What are the pros/cons of DMO vs. PostPro with respect to HIW phenomena predictions? What is the user’s interpretation of forecast value in high-impact weather situations?

HIW phenomena studied: fog/visibility, convection related CW (thunderstorms, lightning, hail, squalls, showers, flash floods)

High-impact weather verification

- The newly developed products used in operations for the forecast of high-impact weather need to be verified
  - complement the traditional verification of the meteorological parameters involved in the occurrence of a high-impact weather phenomenon (precipitation, temperature, wind) with a specific verification of these products
- Observations: conventional meteorological obs, remote sensing datasets, datasets from telecommunication systems (e.g. cell phones), data collected from citizens, reports of impact and claim/damage reports from insurance companies.
- The verification of these products require a different approach to the objective verification process

Creating the pair

the predicted quantity:

- define the quantity or object to be verified, which is selected as representative of the phenomenon
- e.g. for thunderstorm: not the accumulated precipitation itself, but precipitation can be an ingredient for the definition
- a quantity which can be either directly observed, or for which an "observable" exists, being highly correlated to it

the observed quantity:

- measurements which permit to observe a quantity really representative of the high-impact weather phenomenon
- should have a usable spatial and temporal coverage and a documentation of the quality
- include the observation uncertainty: e.g. use observed data coming from different sources

Matching between prediction and observation

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

DWD

obs: lightning

0 1 10 50 100 1000

Number of lightnings

obs: radar

1 10 50 100 1000

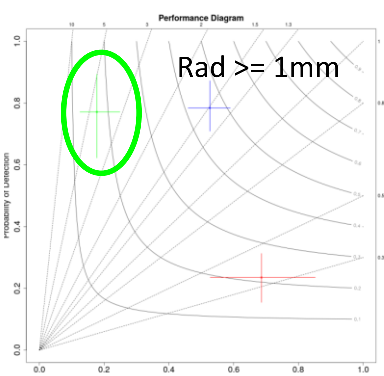
Precipitation (mm)

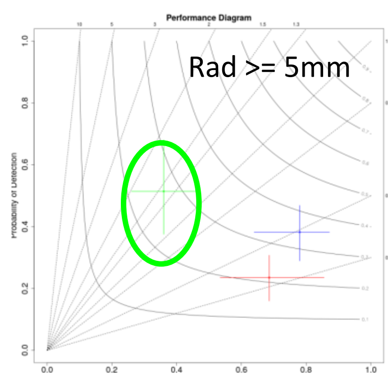
forecast: ensemble

0 10 30 50 100

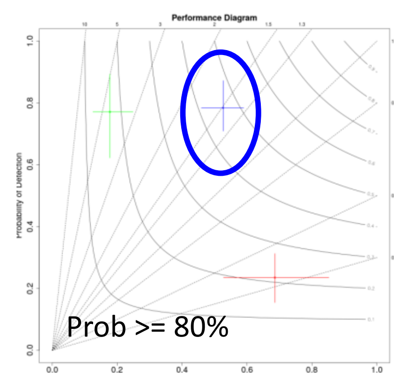
Probability of rain > 1mm

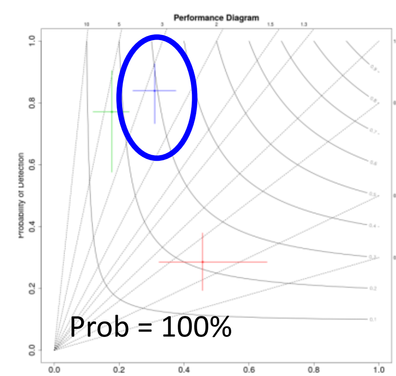
radar vs lightning





ensemble vs radar





Creating the pair

matching between prediction and observation:

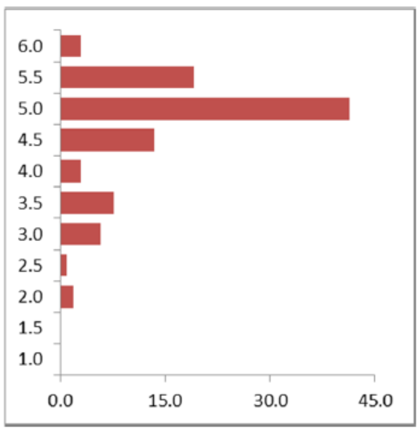
- the matching in the prediction-observation pair should be ensured!
- e.g.: is “thunderstorm cell” - “at least one lightning” a suitable pair?
- a preparatory step is needed:
  - if climatologically (statistically over a long period) there is a good correlation between them, it can be assumed that one can provide the reference for the other and objective verification can be performed
  - this may involve the definition of thresholds (of both quantities) to be used to identify the objects to be compared
  - an important part of this process is to assess spatial and temporal representativeness and to suitably average or re-grid forecasts and/or observations

Applications of **thunderstorm** verification:
subjective evaluation of testers

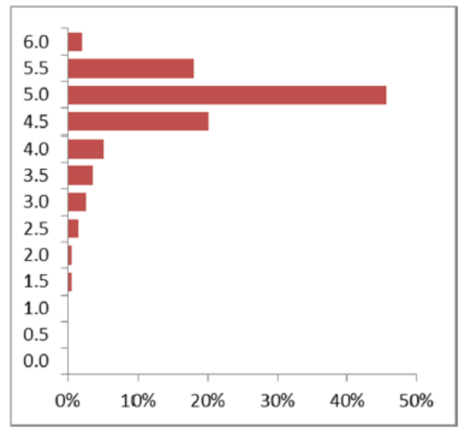
Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Simulated: thunderstorm warning from nowcasting sent to the testers
Observed: testers subjective evaluation

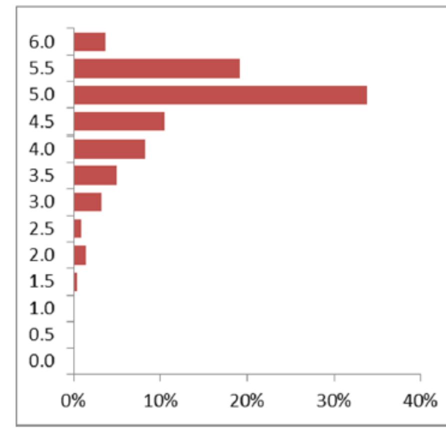
Experimental Thunderstorm Information from MeteoSwiss:
a ***developing*** */* ***moderate*** */* ***severe*** */* ***very severe*** *thunderstorm*
is expected in the next XX min. in (NAME of municipality).



2014 (104 responses)



2015 (195 responses)



2016 (219 responses)

2014	2015	2016
POD: 75 %	POD: 88 %	POD: 67 %
FAR: 25 %	FAR: 33 %	FAR: 29 %

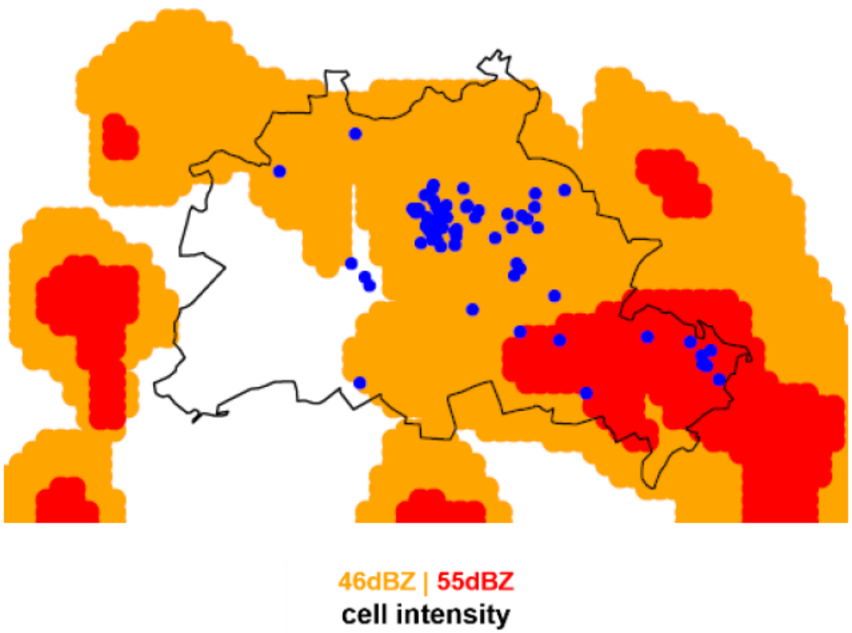
subjective evaluation of the beta tester (feedback after each warning)

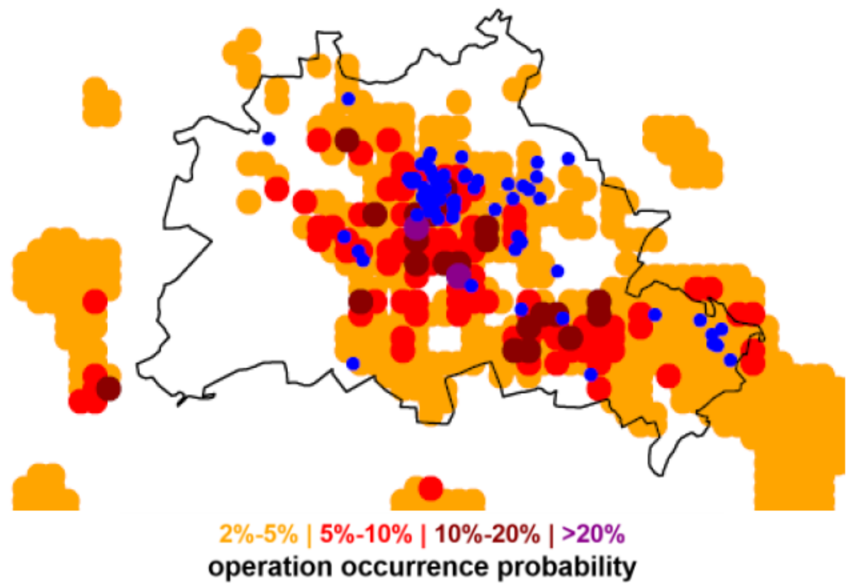
Gaia et al., 2017

Applications of **thunderstorm** verification: fire brigade operations

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

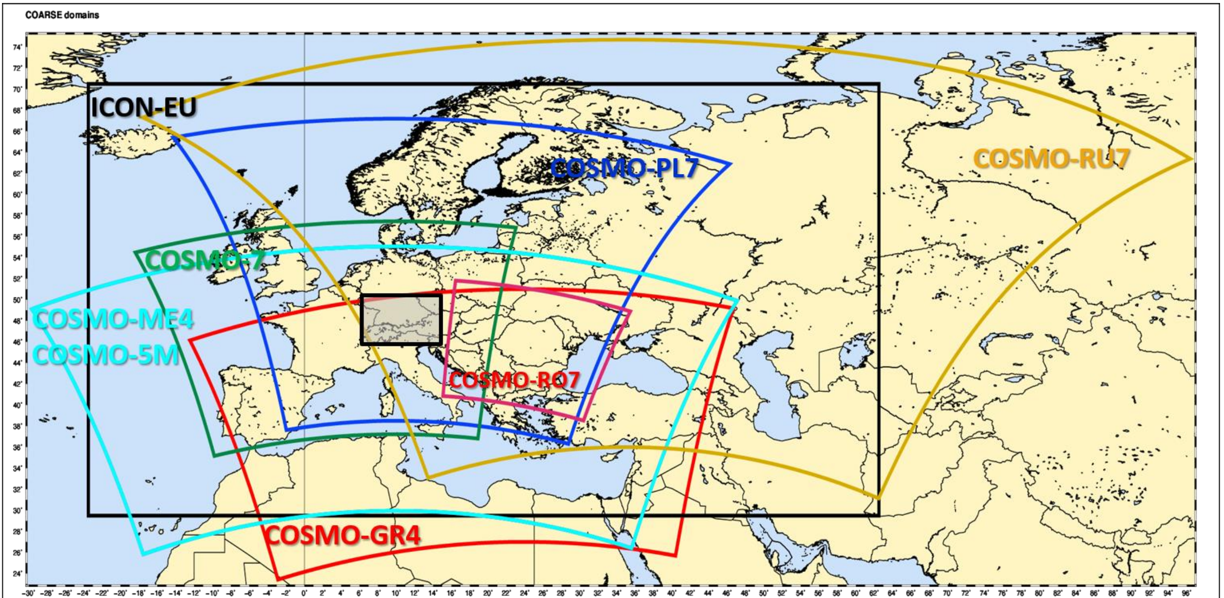
Simulated: „footprint“ of convective cell detected by a nowcasting algorithm
Observed: fire brigade operations (water related)





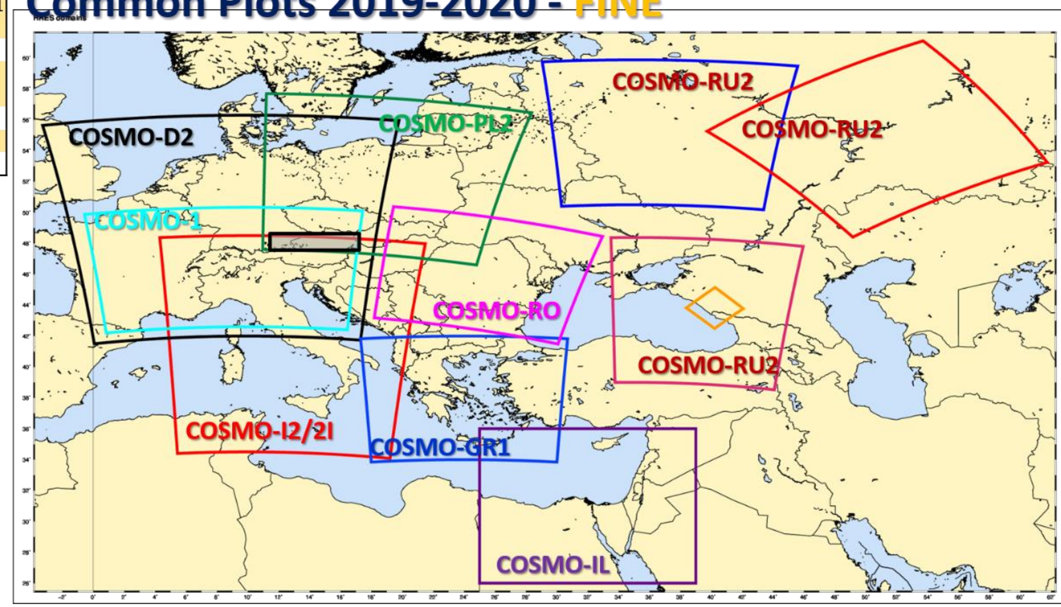
strong dependence also on exposure and vulnerability

Pardowitz and Göber, 2017

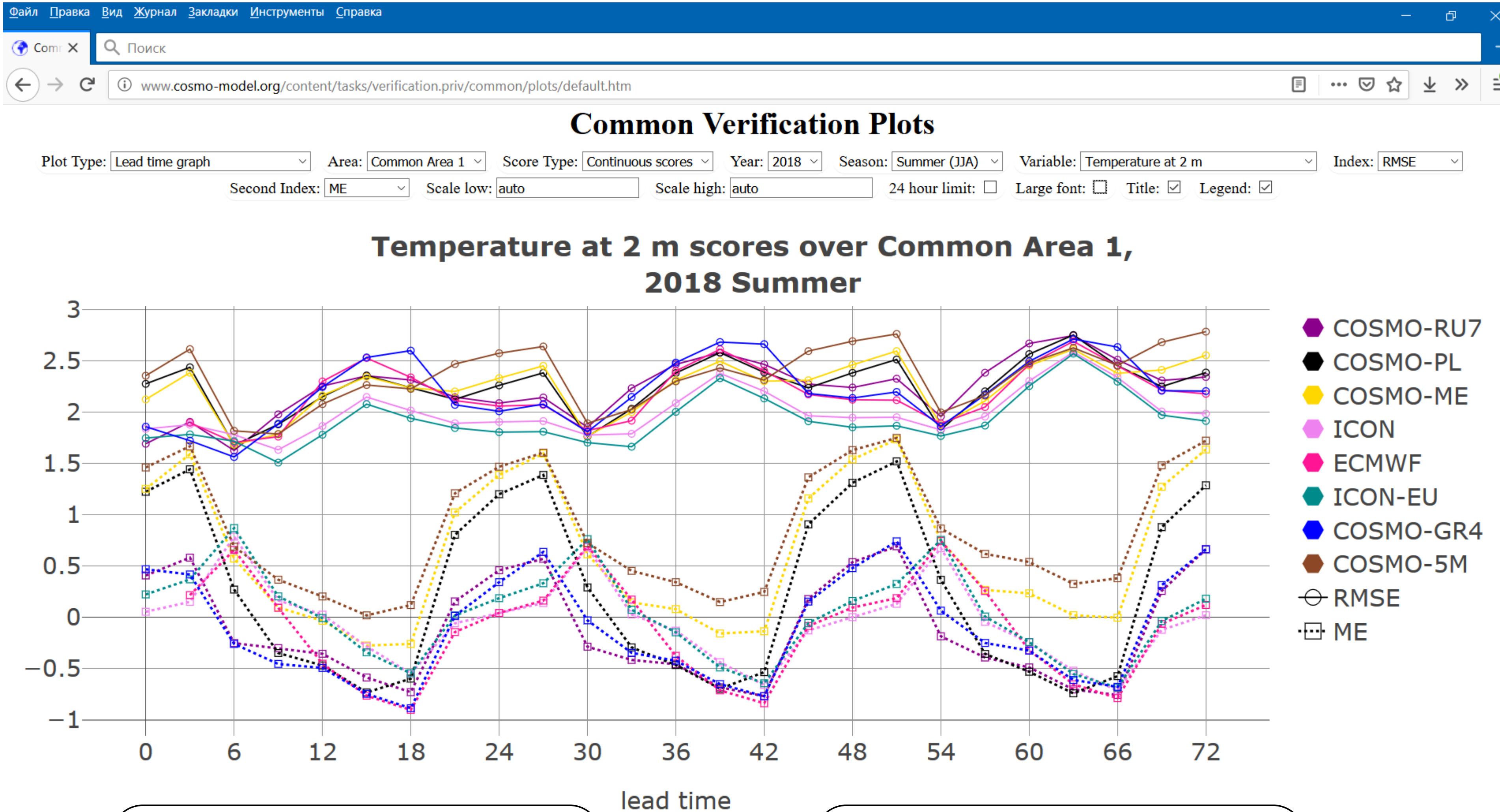


COARSE		FINE	
COSMO	*ICON*	*COSMO*	*ICON*
COSMO-5M		COSMO-2I/2IRUC	
COSMO-GR4		COSMO-GR1	ICON-GR2 or 1
		COSMO-IL	ICON-IL
COSMO-RU7	ICON-RU	COSMO-RU++	
	ICON-EU	COSMO-DE	ICON-D2
COSMO-PL7		COSMO-PL2.8	ICON-PL2.5
COSMO-7		COSMO-1	
COSMO-ME	ICON-ME	COSMO-IT	ICON-IT
COSMO-RO7		COSMO-RO1	

Common Plots 2019-2020 - FINE



Common Plots Interactive view



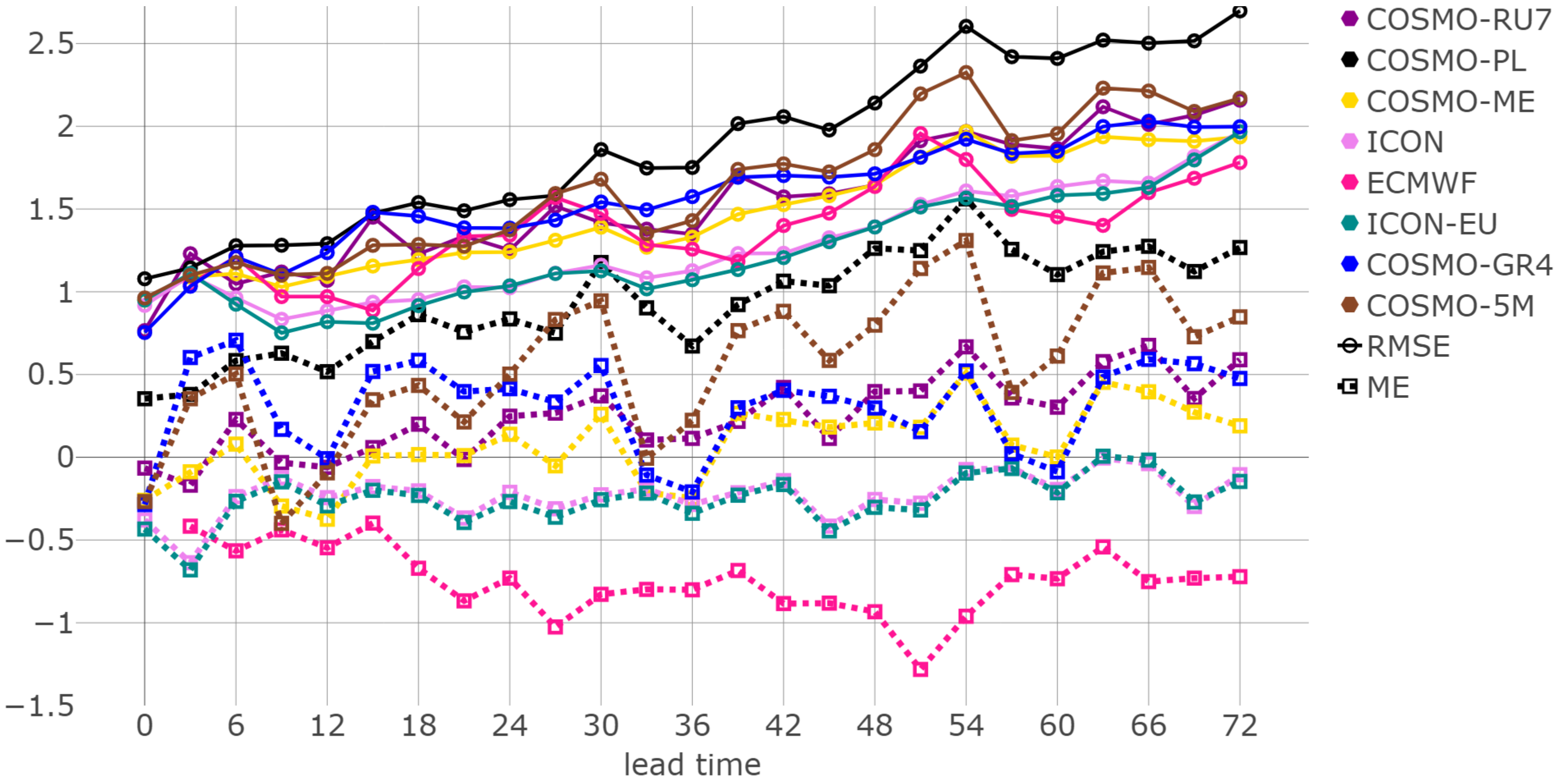
Diurnal variability underestimation persists

ICON-EU has the lowest RMSE after 6 h lead time

All the scores you can view at the website:

http://cosmo-model.org/content/tasks/verification.priv/common/plots/default.htm

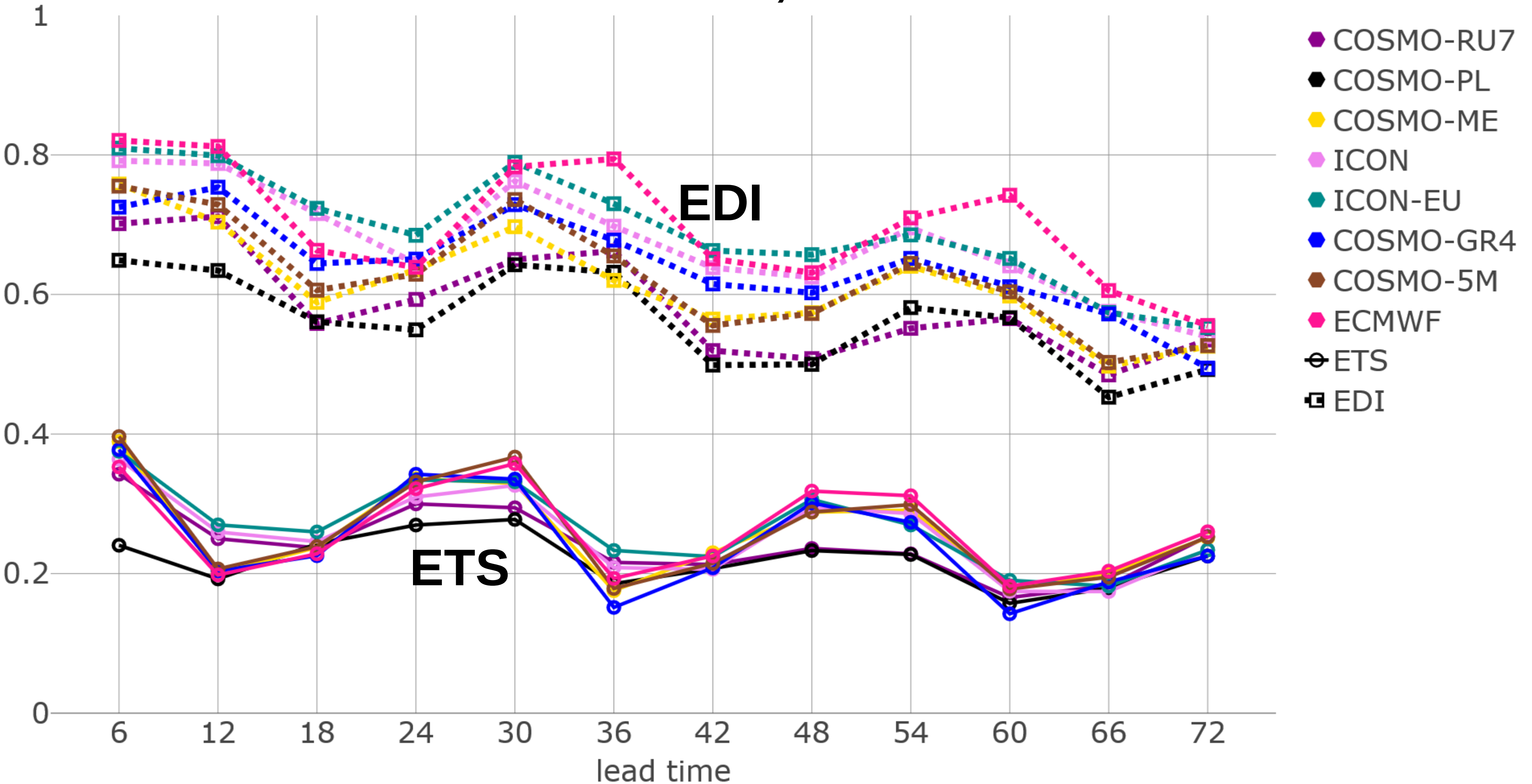
Pressure reduced to Mean Sea Level scores
Common Area 1, Winter 2018-2019



<u>**NO**</u> RMSE
maximum during
<u>afternoon</u>

ICON and **IFS** have
better RMSE
than **COSMO**

>0.2 mm Total precipitation in 6 hours scores
Common Area 1, Summer 2018



While having better FBI **COSMO** does not have better ETS than ICON

ETS higher at winter
worst 12-18 UTC
best 0-6 UTC

THANK YOU!

БЛАГОДАРЯ!