Verification and Case studies

Overview of activities

Flora Gofa

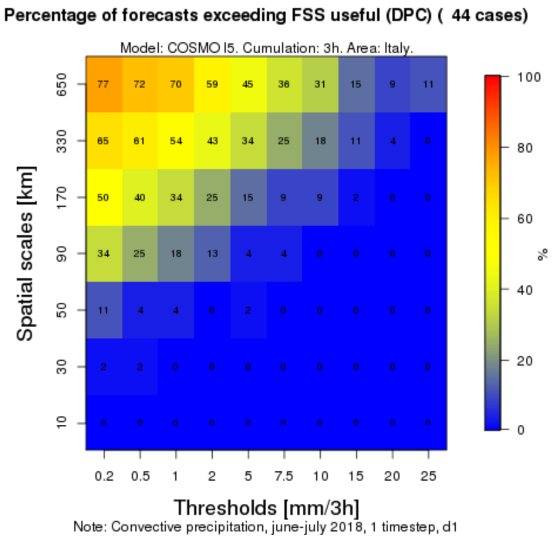

Verif
Software
WG5
activities

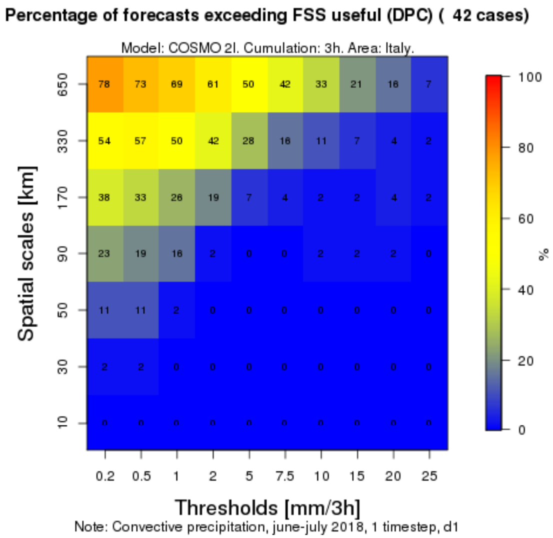
VAST improvements: time windows

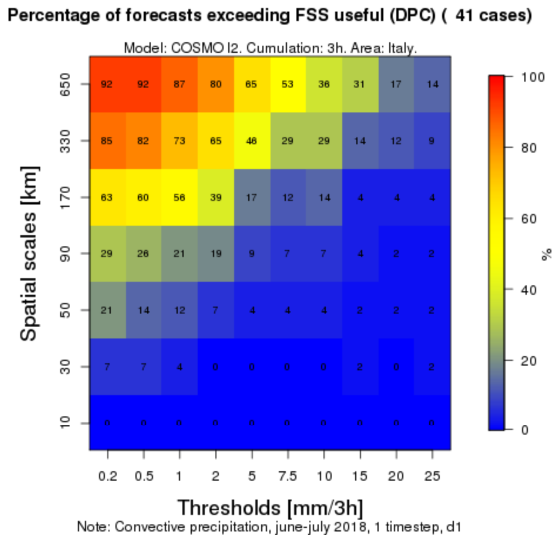
Naima Vella

- Implemented in VAST version 2.0
- The user has the possibility to produce the verification using more than one observation/forecast timestep
- At the moment all the timesteps must be available (no time gaps in data)
  - If not, wrong scores will be produced
- The number of timesteps used for the verification can be set in «input_fuzzy.nml», variable «n_timesteps»
  - The value must be an odd number:
    - 1 => only current timestep (2D verification, same as previous VAST versions)
    - 3 => previous one, current, following one
    - 5 => previous two, current, following two
    - …
  - ATTENTION: the higher the number, the slower the process! Naima Velia

JJ 2018 – D1 – 1 timestep – 10x10km grid

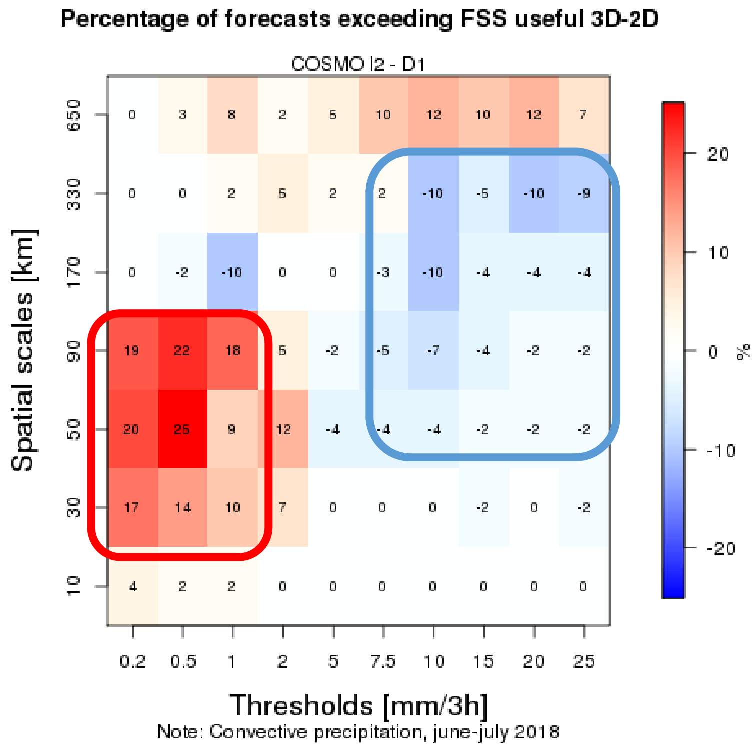






Best performances for COSMO I2 for all spatial scales and all precipitation thresholds .
Very similar performance for COSMO I5 and COSMO 2I.

JJ 2018 – D1 – (3-1) timesteps – COSMO I2



Spatial scales [km]	0.2	0.5	1	2	5	7.5	10	15	20	25
650	0	3	8	2	5	10	12	10	12	7
330	0	0	2	5	2	2	-10	-5	-10	-9
170	0	-2	-10	0	0	-3	-10	-4	-4	-4
90	19	22	18	5	-2	-5	-7	-4	-2	-2
50	20	25	9	12	-4	-4	-4	-2	-2	-2
30	17	14	10	7	0	0	0	-2	0	-2
10	4	2	2	0	0	0	0	0	0	0

Advances in Rfdbk based Verification at DWD

Felix Fundel

Deutscher Wetterdienst

FE 15 – Predictability & Verification

Tel.:+49 (69) 8062 2422

Email: Felix.Fundel@dwd.de

Rfdbk Verification

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

DWD

Optional namelist options

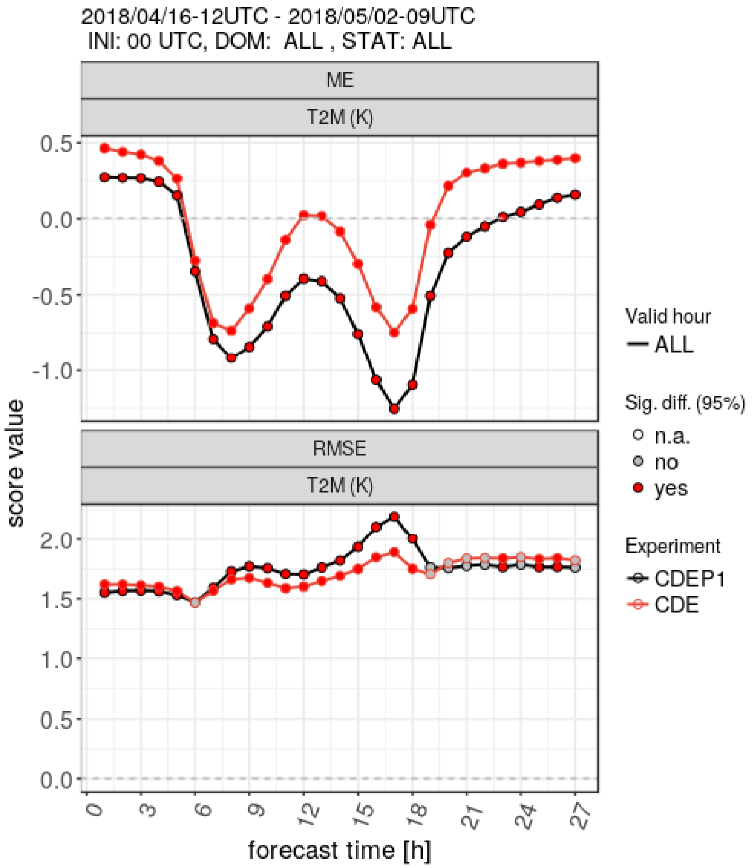
NAME	*VALUE*	*DESCRIPTION*
IdentList	'/path/to/your/identlist'	# Use only station(s) given in list file (integer)
statidList	'/path/to/your/statidlist'	# Use only station(s) given in list file ()
dateList	'/path/to/your/datelist'	# Verify dates in list separately (YYYYMMDD)
lonlims/latlims	'0,30'	# Restrict 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

If not given, DWD standard settings are used! Essentially any observation/forecast characteristic contained in feedback files (~50) can be used to refine the verification via namelist.

Significance Test

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Included in TEMP, SYNOP, det. and ensemble verification



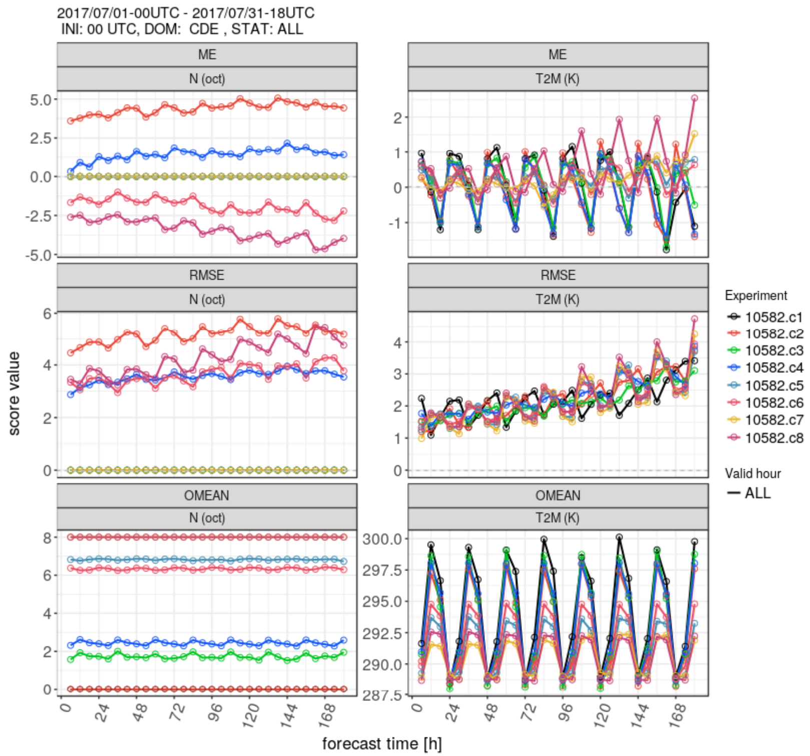
- t-test for significant difference from 0 of difference of scores from 2 experiments
- Implemented for area mean scores (not station based)
- 24 hours between score validity time needed (e.g. test for each initial time separately)
- t-tests requirements (normally dist. measurements, iid) is approximately valid for daily scores
- Point colors indicate significance
- Only visible if runs are not aggregates (e.g. Ini Time 00 or 12)
- Visible also when score difference is plotted
- Works also in hindcast mode

As for now, not possible with conditional verification

Conditional Verifikation II

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Implemented for SYNOP deterministic 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

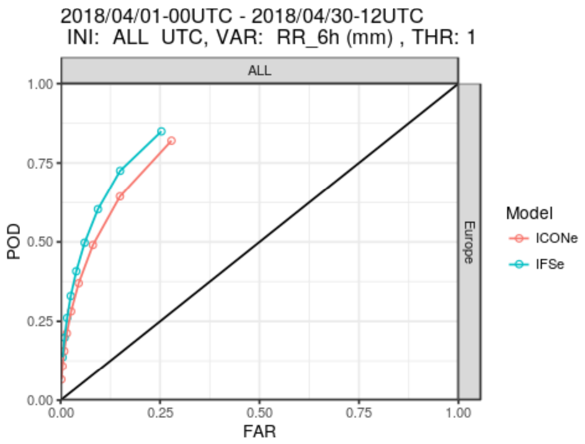
Example namelist

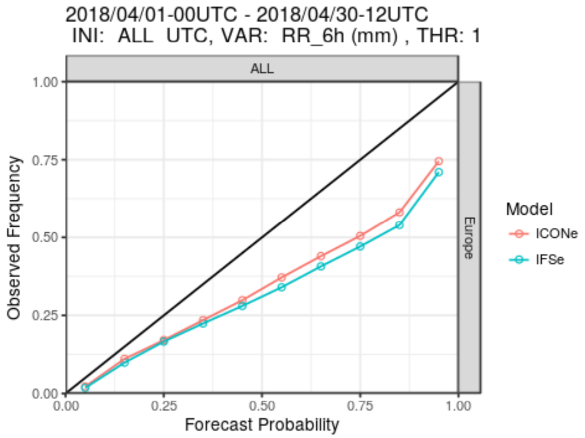
```
condition1    "list(N='obs==0',N='abs(veri_data-obs)<1') "
condition2    "list(N='obs==0',N='abs(veri_data-obs)>=1') "
condition3    "list(N='obs%between%c(1,4)',N='abs(veri_data-obs)<1') "
condition4    "list(N='obs%between%c(1,4)',N='abs(veri_data-obs)>=1') "
condition5    "list(N='obs%between%c(5,7)',N='abs(veri_data-obs)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

Probabilistic Verification

Deutscher Wetterdienst
Wetter und Klima aus einer Hand





- Based on „probability files“ that need to be produced from feedback files.
- „probability files“ 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
- This approach is time consuming and not very flexible
- Working on verification script to combine ensemble and probabilistic verification

Plans

Rfdbk Verification

- Conditional verification for TEMP
- Unify probabilistic and ensemble verification
- Add flexibility concerning bias correction (allow to use/discard correction)
- Significance test for categorical scores (not good idea yet)
- Allow for single member verification
- Maybe separate verification functionality from Rfdbk and create a extra R package for that

Spatial Verification Efforts at DWD

Felix Fundel

Deutscher Wetterdienst

FE 15 – Predictability & Verification

Phone:+49 (69) 8062 2422

Email: Felix.Fundel@dwd.de

II. Approach

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

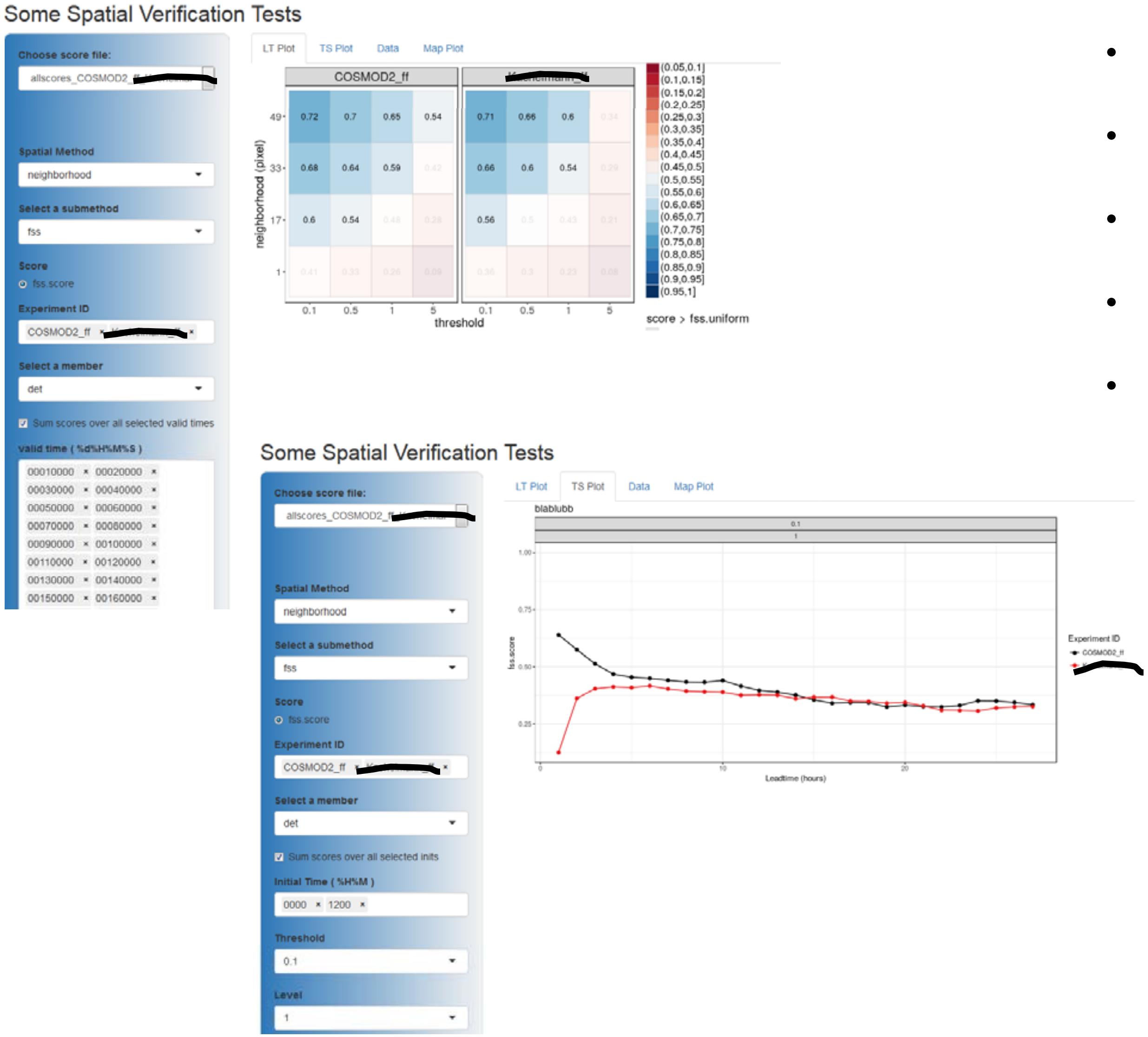
- **Review of existing neighborhood/spatial verification methods for deterministic and ensemble forecasts**
  - Deterministic
    - Neighborhood: methods & scores from Ebert 2008 (incl. single member verification)
    - Object: Konrad3D objects, scores t.b.d.
  - Ensemble
    - Neighborhood.:NEP (Schwartz et al. 2010); NEP + time fuzzyness (Duc et al. 2012,2013)
    - Object: Konrad3D, clustering of EPS objects?, scores t.b.d.

- **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 (too complex)*

III. First Results

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

DWD



- Interactive score visualization via shiny-server
- Methods, score, experiments, lead-times, ini-times etc. selectable
- In app aggregation over selected lead-times & ini-times
- Prepared for det., EPS & single member verification
- Export of data possible


Verif
Software
NWP test
suite
WG5
activities

NWP METEOROLOGICAL TEST SUITE

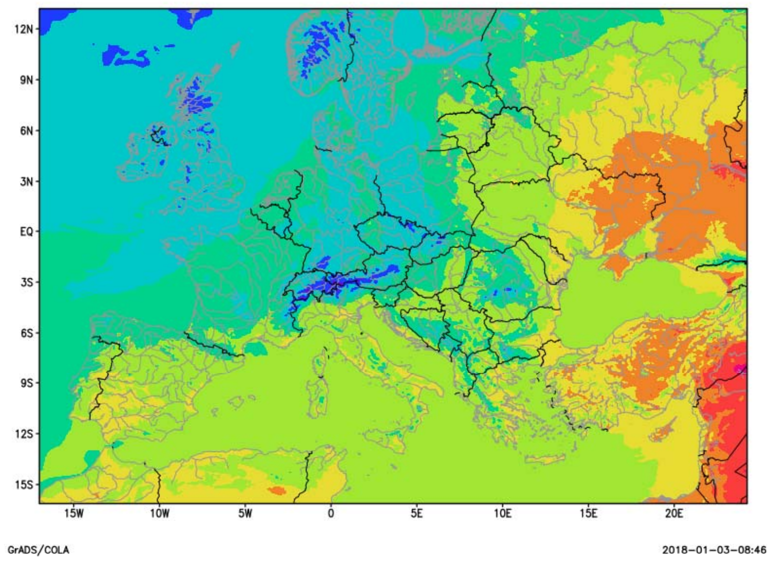
VERSUS2RFDBK MIGRATION

Amalia IRIZA-BURCA (NMA), Felix FUNDEL (DWD), Flora Gofa (HNMS)

NWP Meteorological Test Suite @ ECMWF – Current Status

- The suite is implemented to test the present version of COSMO (e.g v5.03) and the experimental one (e.g. v5.05) for:

2 months (currently July 2017 and December 2017)
COSMO@7p0: ie_tot=661, je_tot=471, 40ML; dlon=dlat=0.0625; fc+72h;
COSMO@2p8: ie_tot=1587, je_tot=1147, 50ML; dlon=dlat=0.125; fc+48h.



Roadmap to MEC-Rfdbk for NWP Test Suite

1. implemetation of MEC-Rfdbk for the NWP Test Suite

2. verification of v5.05 against v5.03 forecast, DP (comparison of results from VERSUS and MEC-Rfdbk)

3. verification of SP runs for v5.05, 7km only with MEC-Rfdbk

+ visualization of results on the COSMO shiny web-server

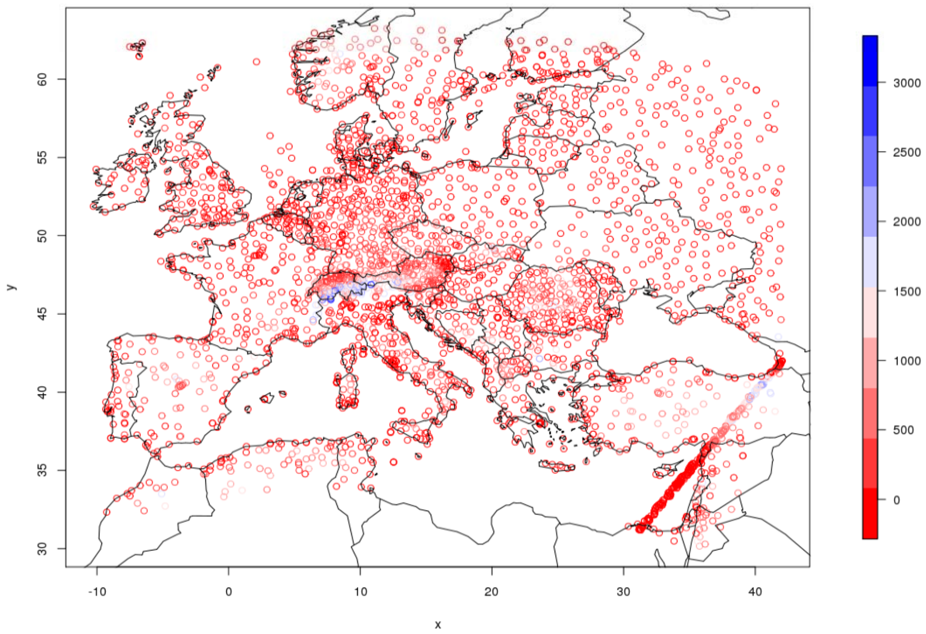
1. Implementation for the NWP Test Suite

- MEC already existing @ ECMWF (by DWD) – run on cca
- Rfdbk already existing @ ECMWF (by F. Fundel) – run on ecgate
- implementation on a user account with access to the ECMWF SP resources
    - access to the MEC executable on cca
    - installation on Rfdbk software package
    - run scripts for MEC and Rfdbk provided by F. Fundel
    - documentation
- first tests (duplicate the same verification on both accounts, to check installation)

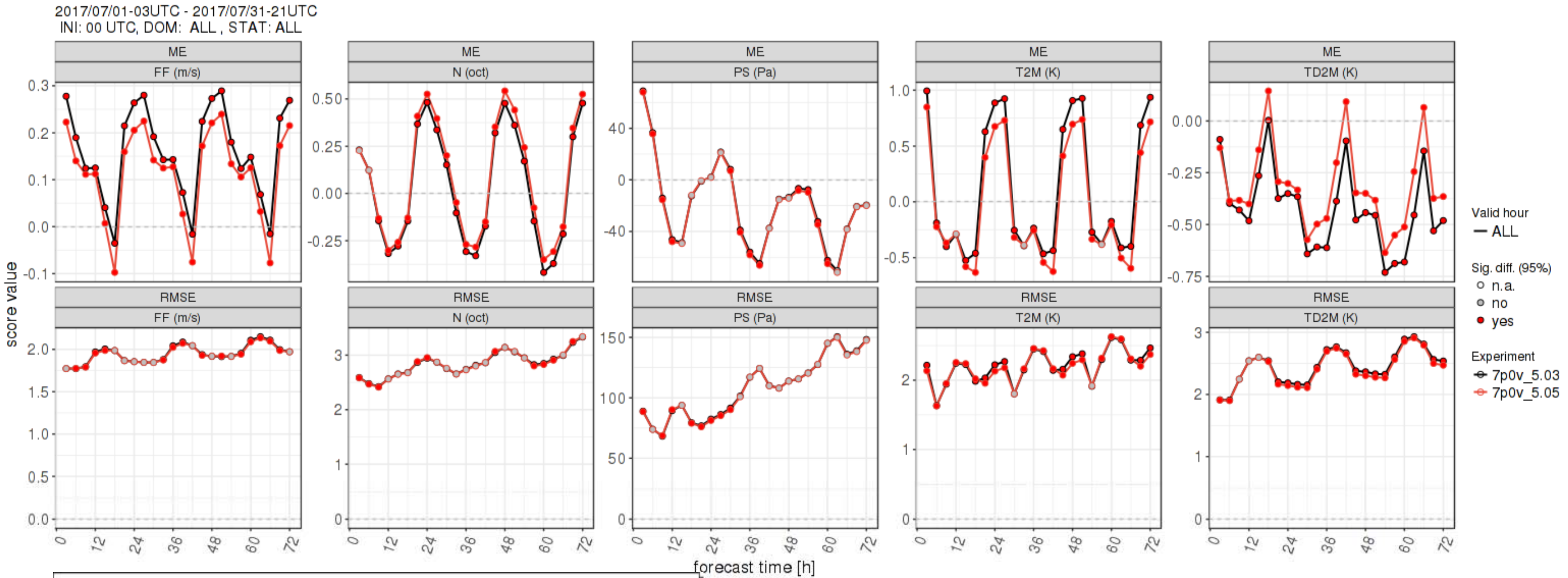
2. Verification of v5.05 against v5.03 forecast, DP (comparison of results from VERSUS and MEC-Rfdbk) F. Fundel (DWD), A. Iriza-Burca (NMA)

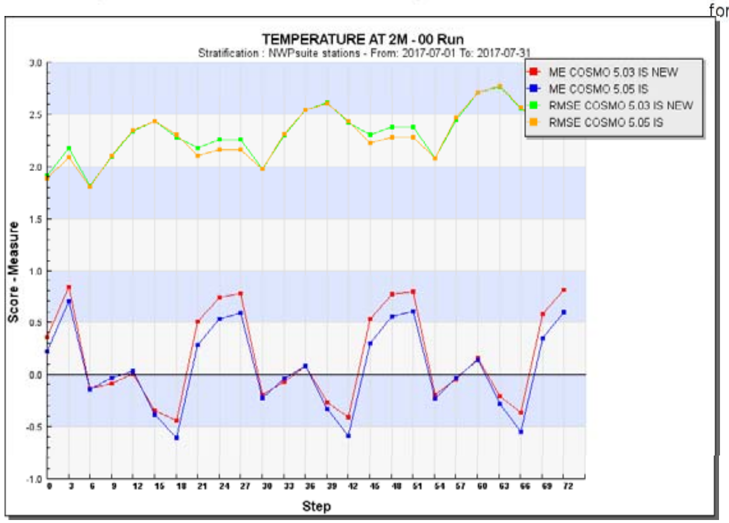
- COSMO@7km: forecast mode (00 UTC + 72 hours), DP
- COSMO@2.8km: forecast mode (00 UTC + 48 hours), DP
- Surface parameters (T2m, TD2M, MSLP, WS10M, TCC + precipitation)
- Upper air parameters (TEMP, RH, WS)
- December and July 2017
- Same observations (~3600 stations)
- Verification duplicated, to test implementation

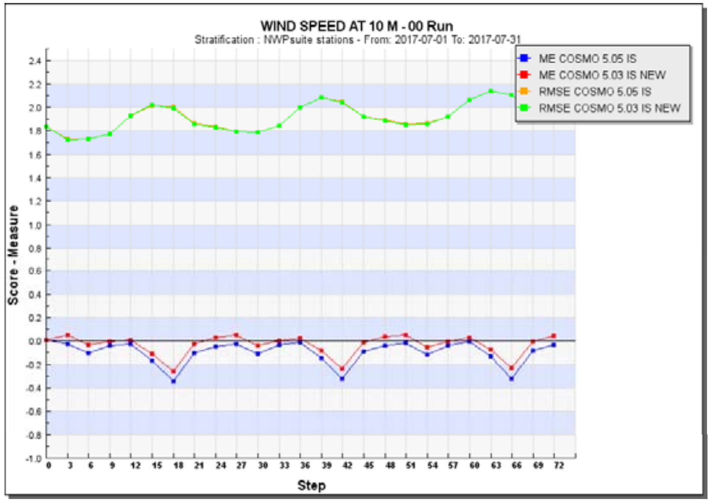
Similar results from VERSUS and MEC -Rfdbk



NWP METEOROLOGICAL TEST SUITE

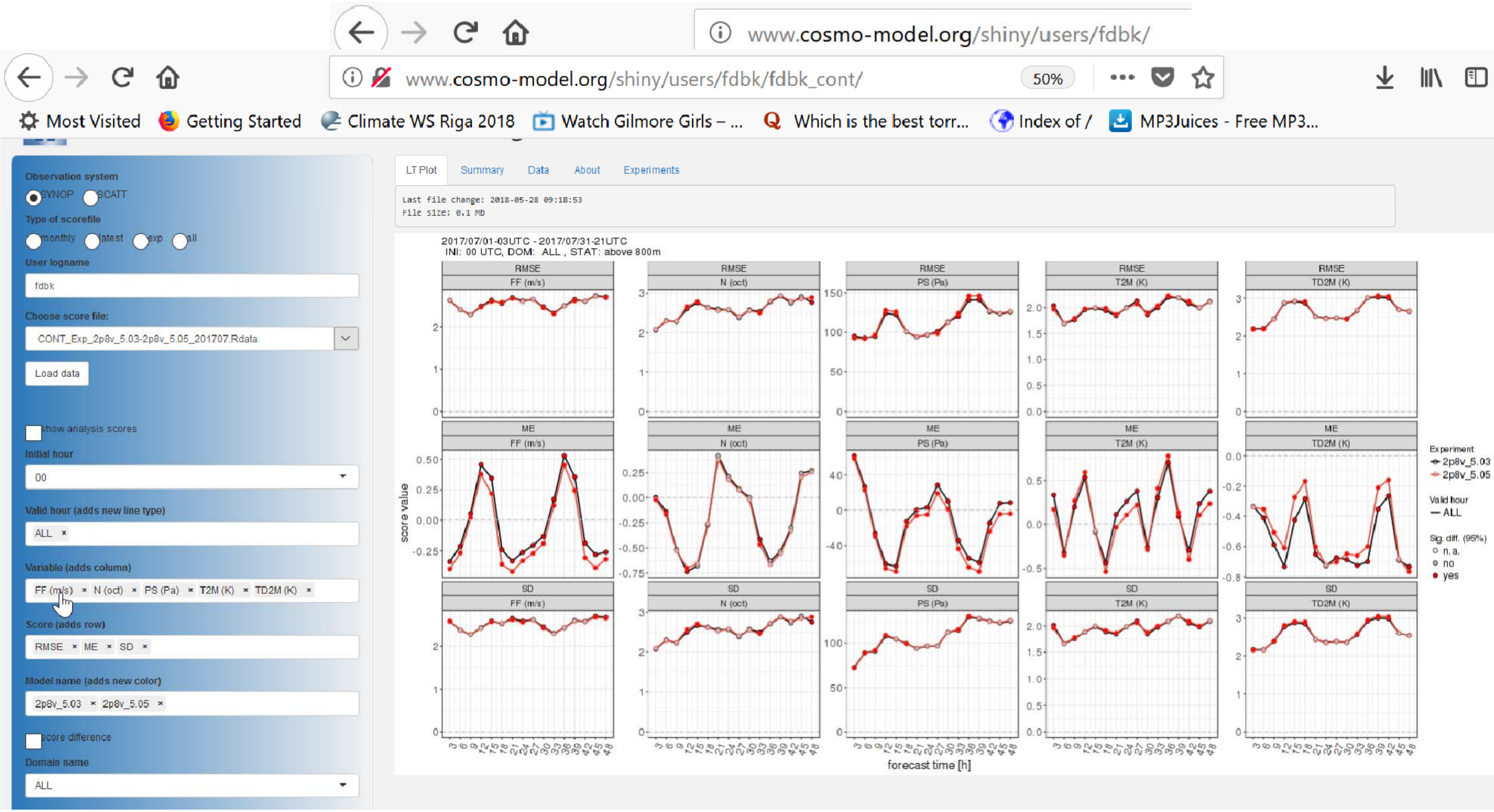






NWP METEOROLOGICAL TEST SUITE

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




Verif
Software
NWP Test
suite
WG5
activities
Applications

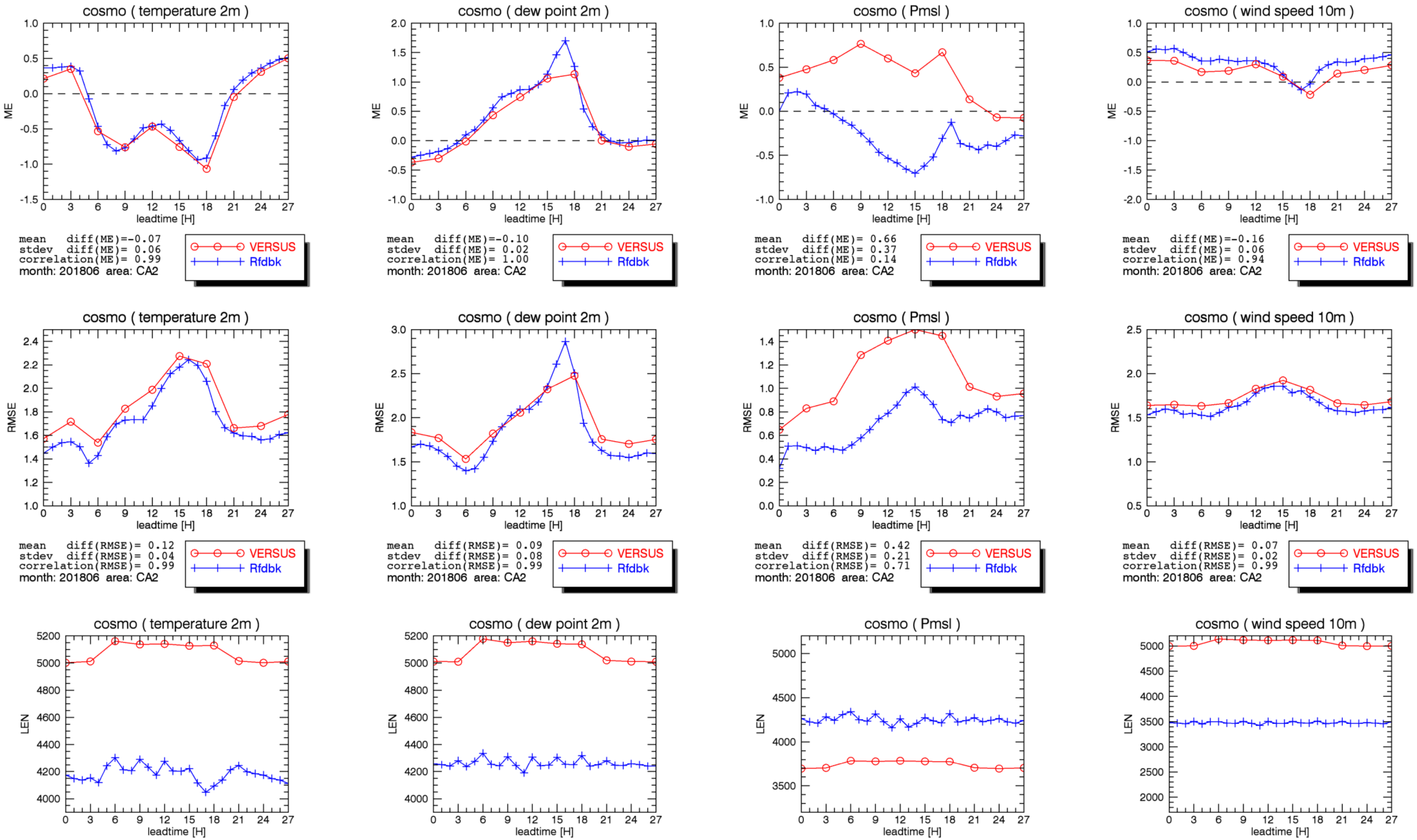
A comparison of Rfdbk and VERSUS verification results

Ulrich Pflüger
Deutscher Wetterdienst

COSMO-D2

Rfdbk and VERSUS (CA2, 201806)

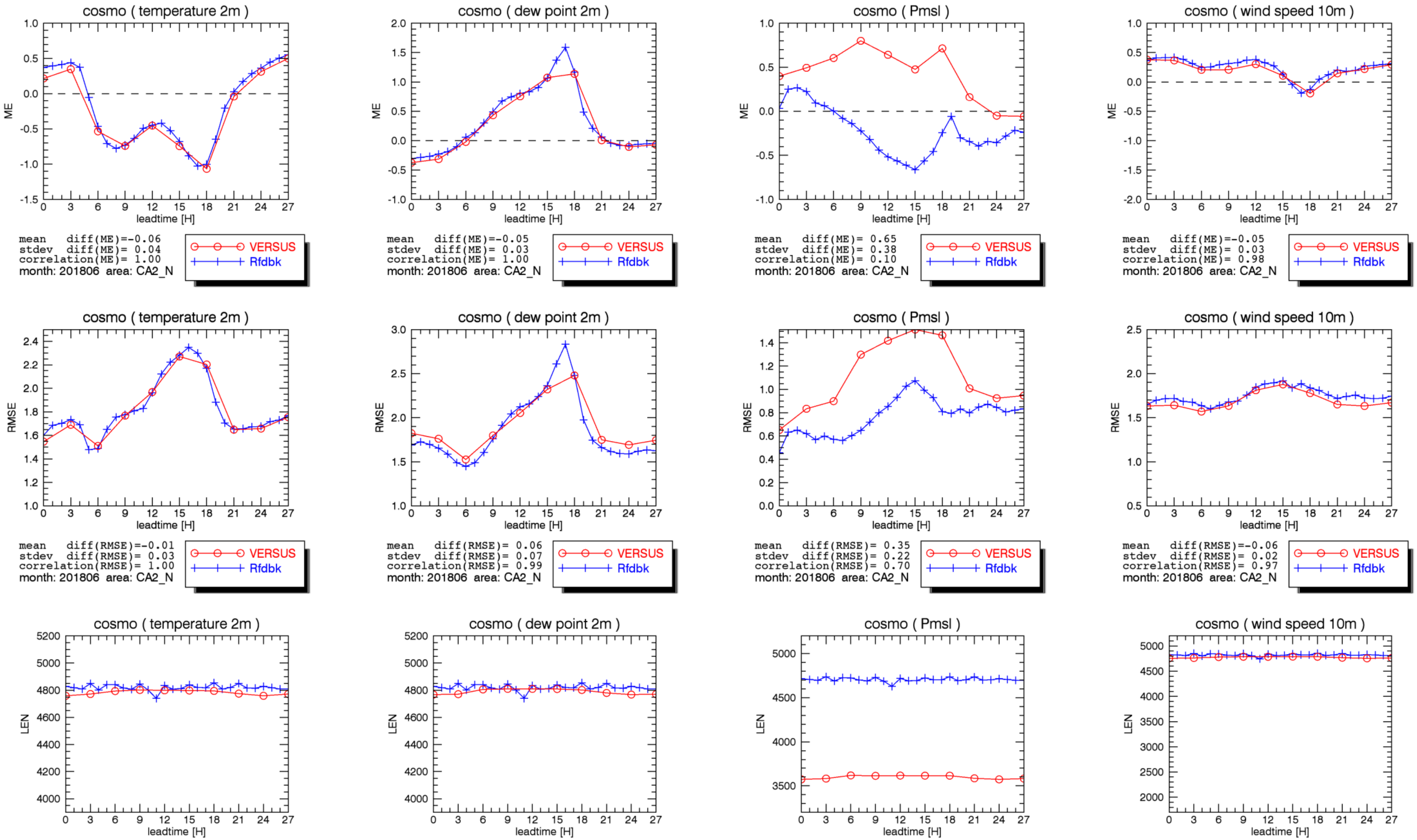
Deutscher Wetterdienst
Wetter und Klima aus einer Hand
DWD



COSMO-D2

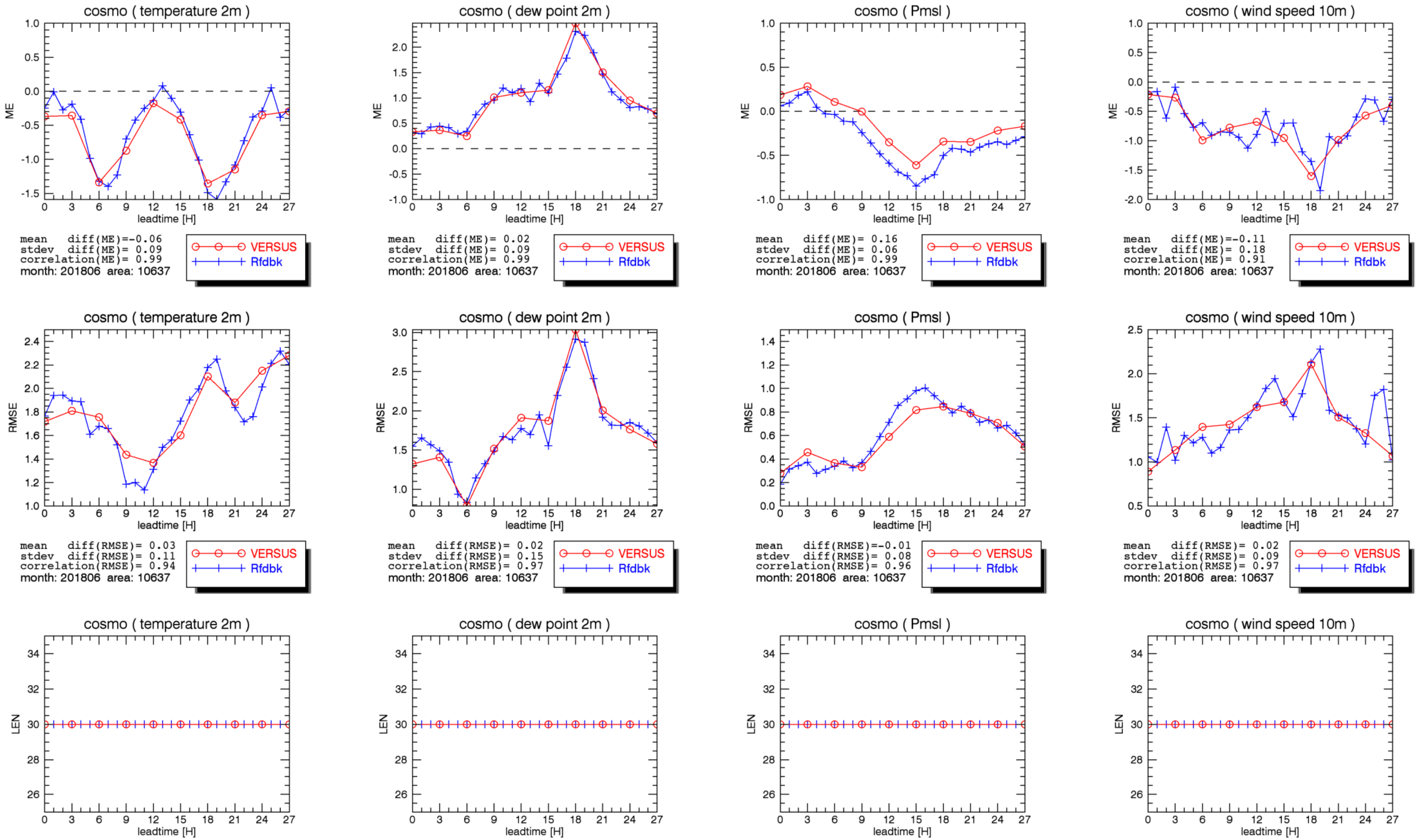
Rfdbk and VERSUS (CA2_N, 201806)

Deutscher Wetterdienst
Wetter und Klima aus einer Hand



COSMO-D2
Rfdbk and VERSUS (10637, 201806)

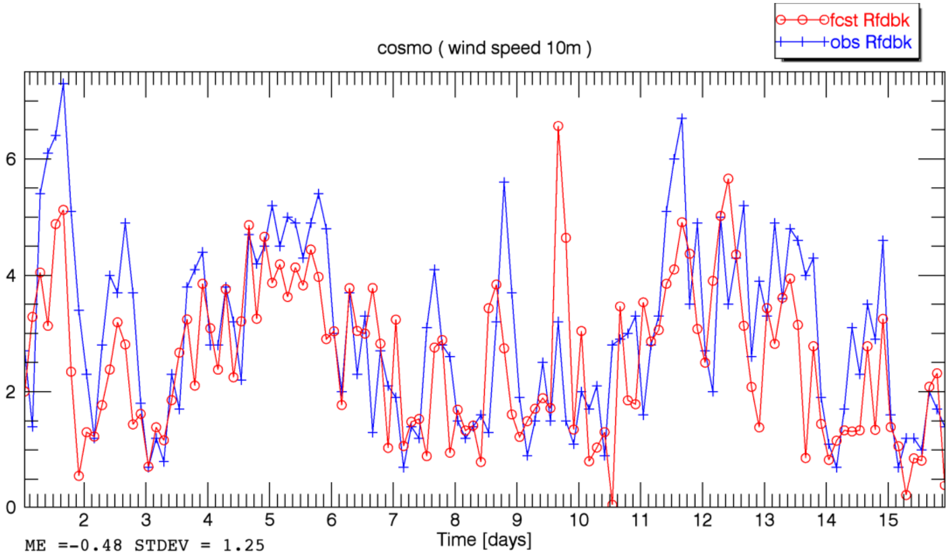
Deutscher Wetterdienst
Wetter und Klima aus einer Hand
DWD

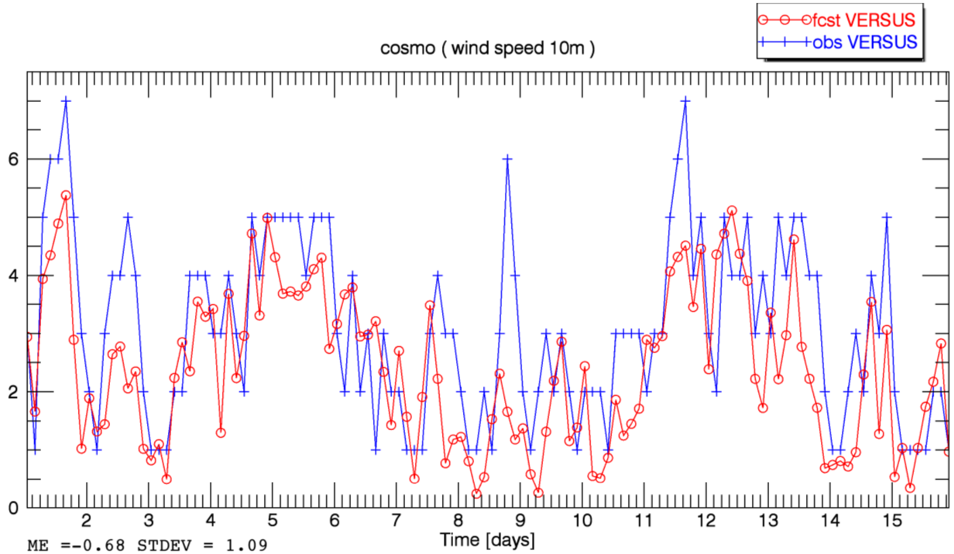


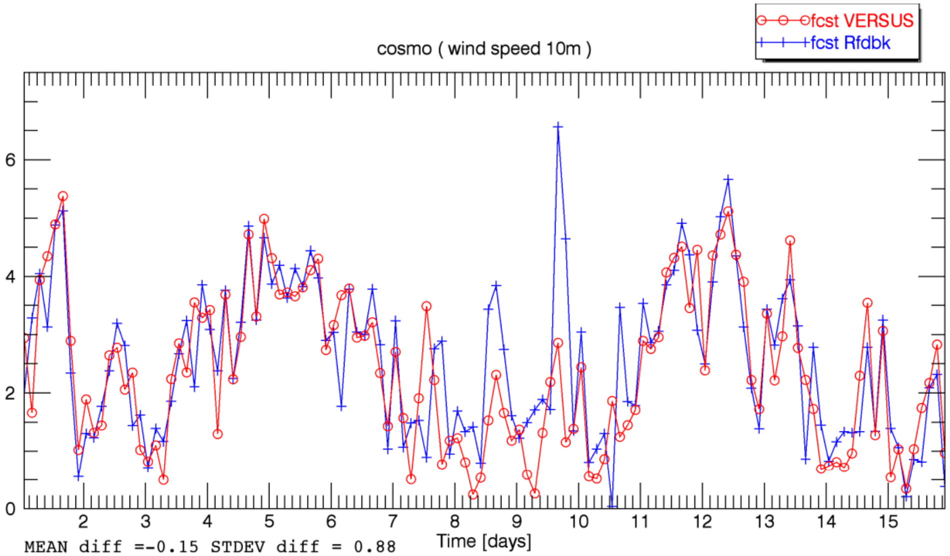
COSMO-D2 (wind speed)

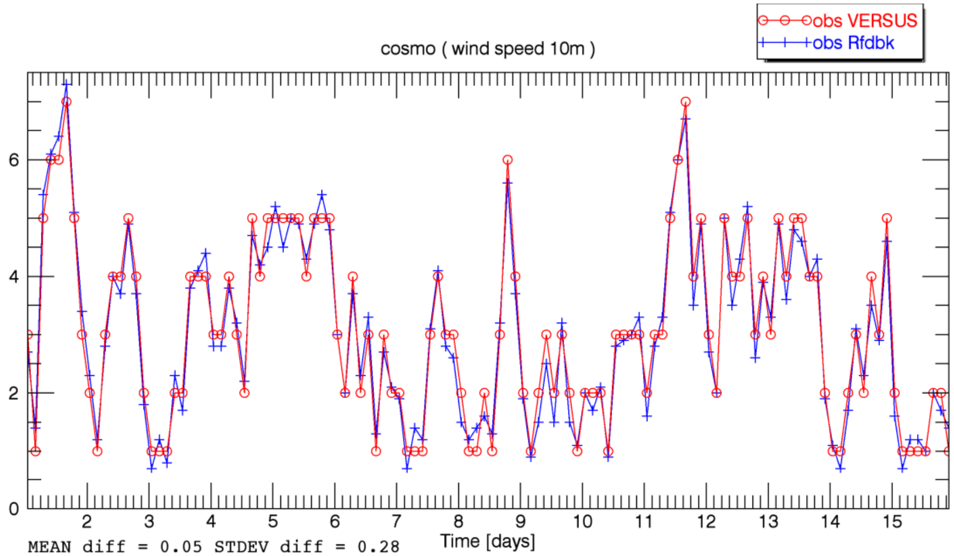
fcst Rfdbk < --- > obs Rfdbk	fcst Rfdbk < --- > obs VERSUS
fcst VERSUS < --- > fcst Rfdbk	obs VERSUS < --- > obs Rfdbk

Deutscher Wetterdienst
Wetter und Klima aus einer Hand









Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Home Affairs FDHA
Federal Office of Meteorology and Climatology MeteoSwiss

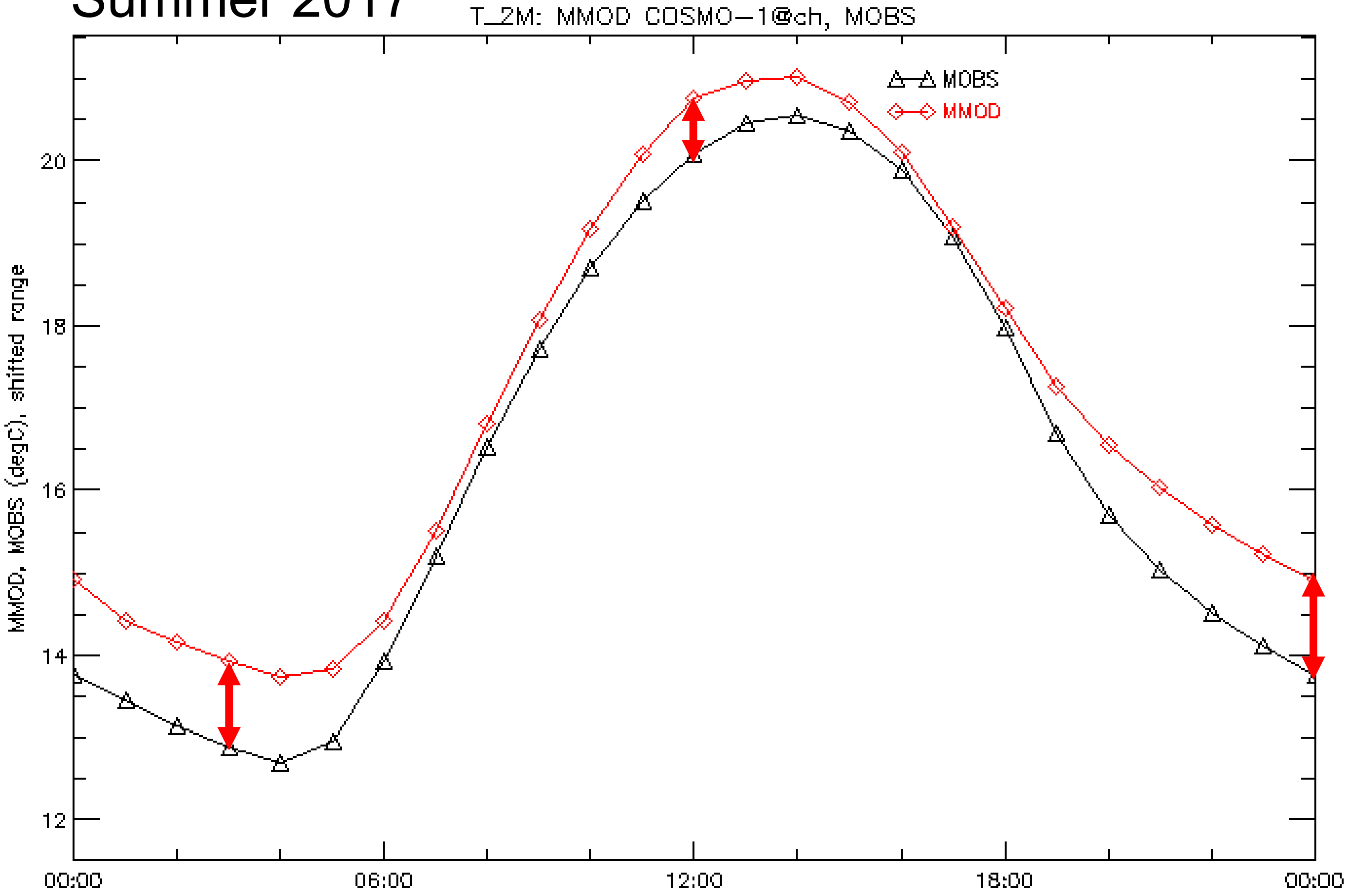
Some Verification Highlights and Issues in Precipitation Verification

Saint Petersburg, 3 September 2018

Pirmin Kaufmann, MeteoSwiss

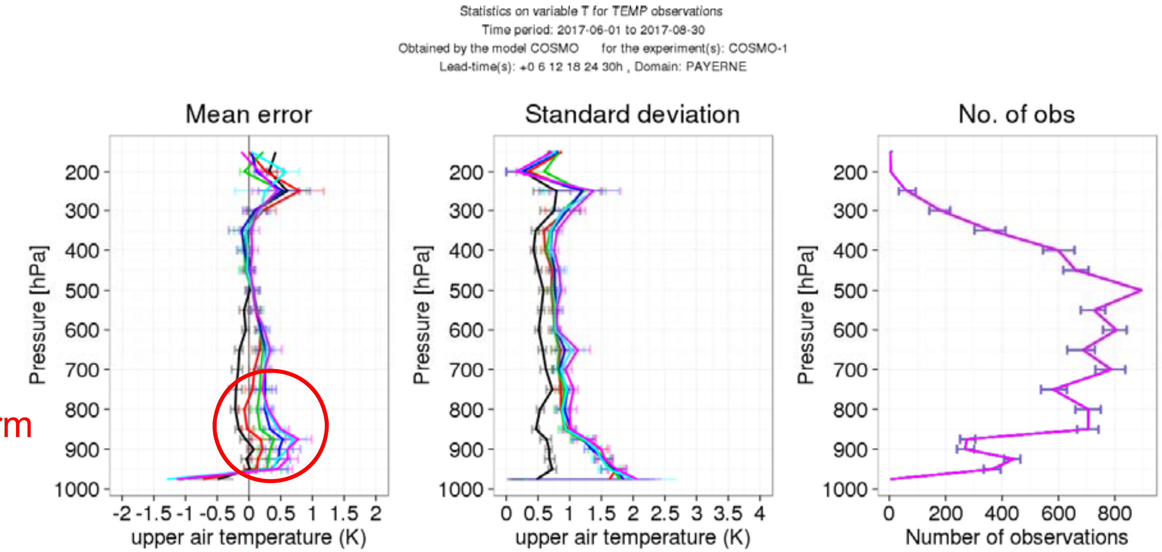
T2m Bias: Strongest During the Night

Summer 2017



daytime

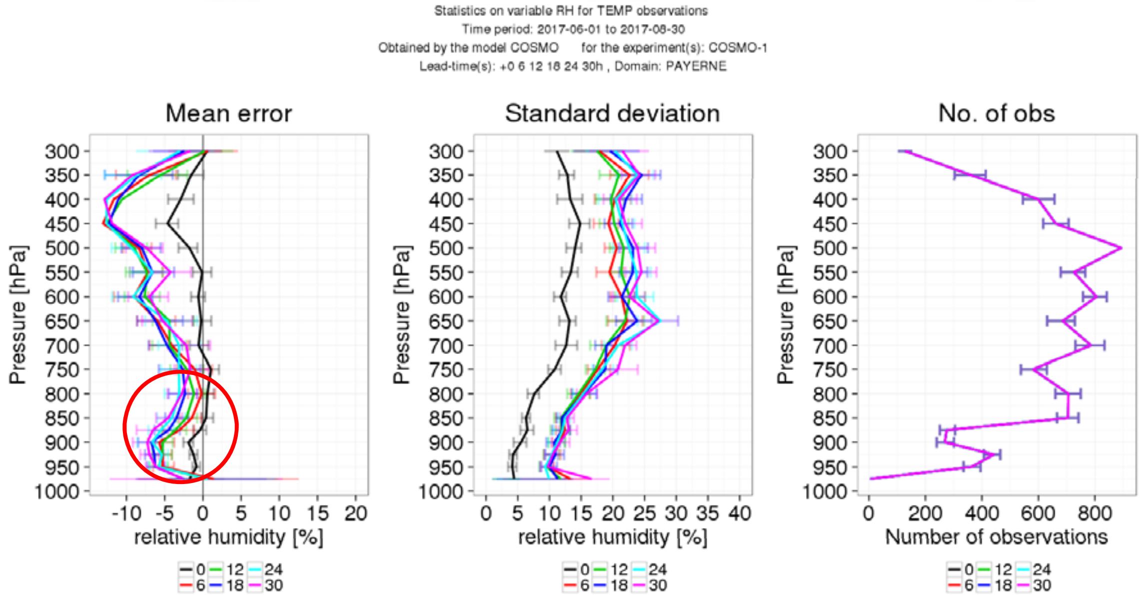
Upper-Air Verification Summer



BL too warm

COSMO-1
Summer 2017
Payerne

Temp.



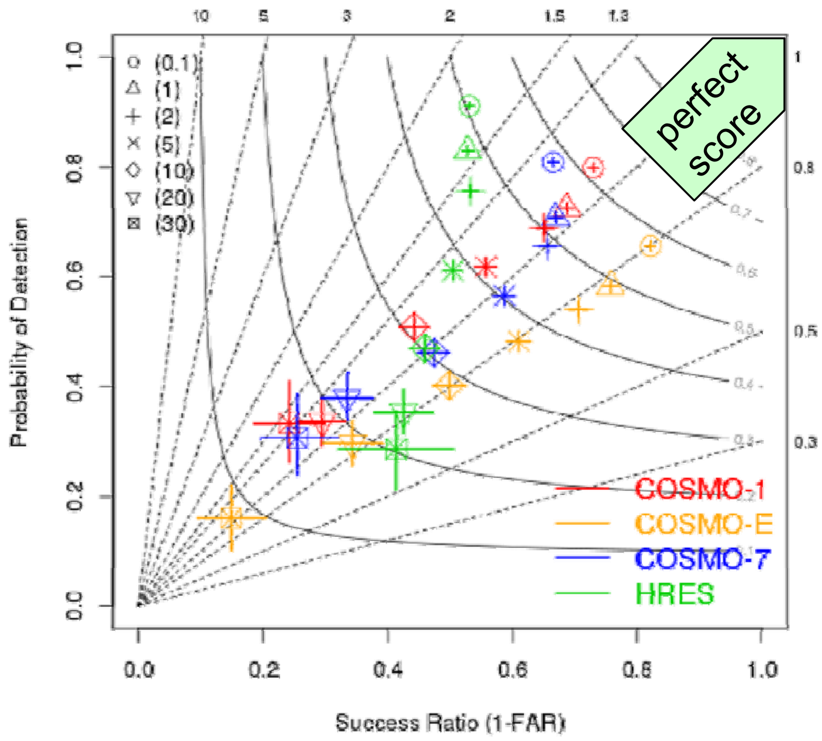
BL too dry

Rel. Hum.

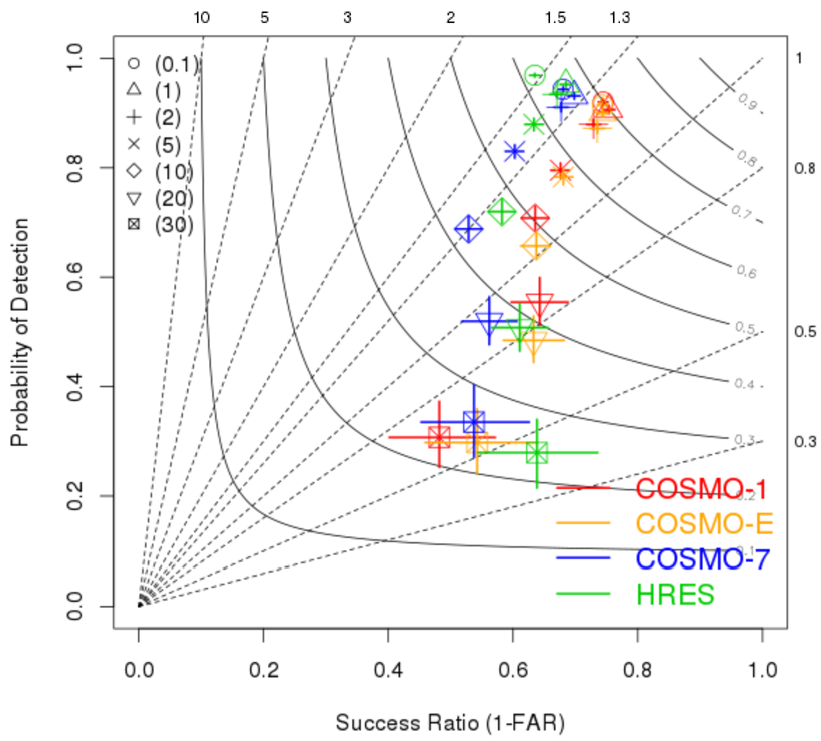
Precipitation Comparison Oper. Models

COSMO-1 best model in summer (12h-sum +12 to +24)
COSMO-1 and -E together best in winter

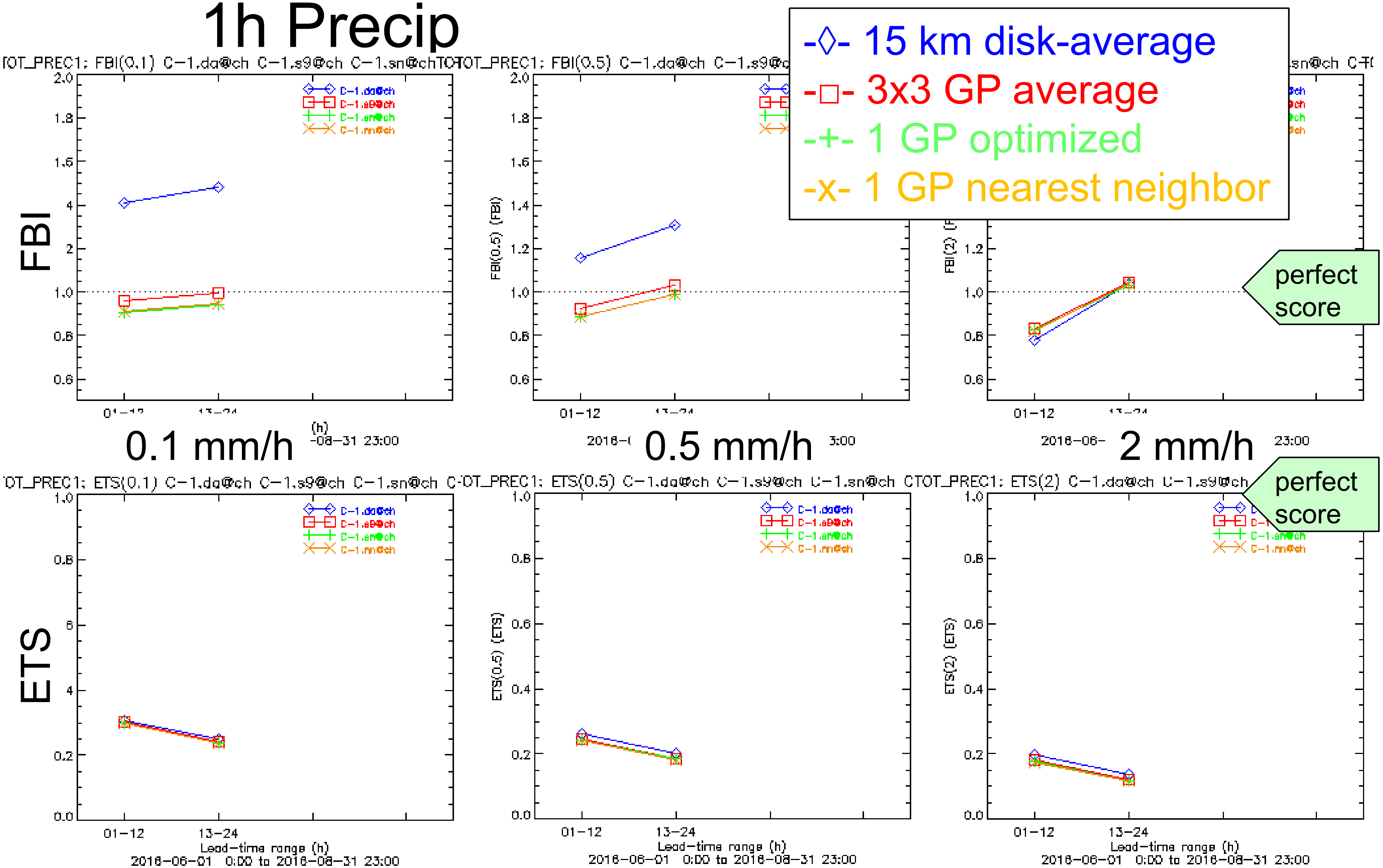
Summer 2017



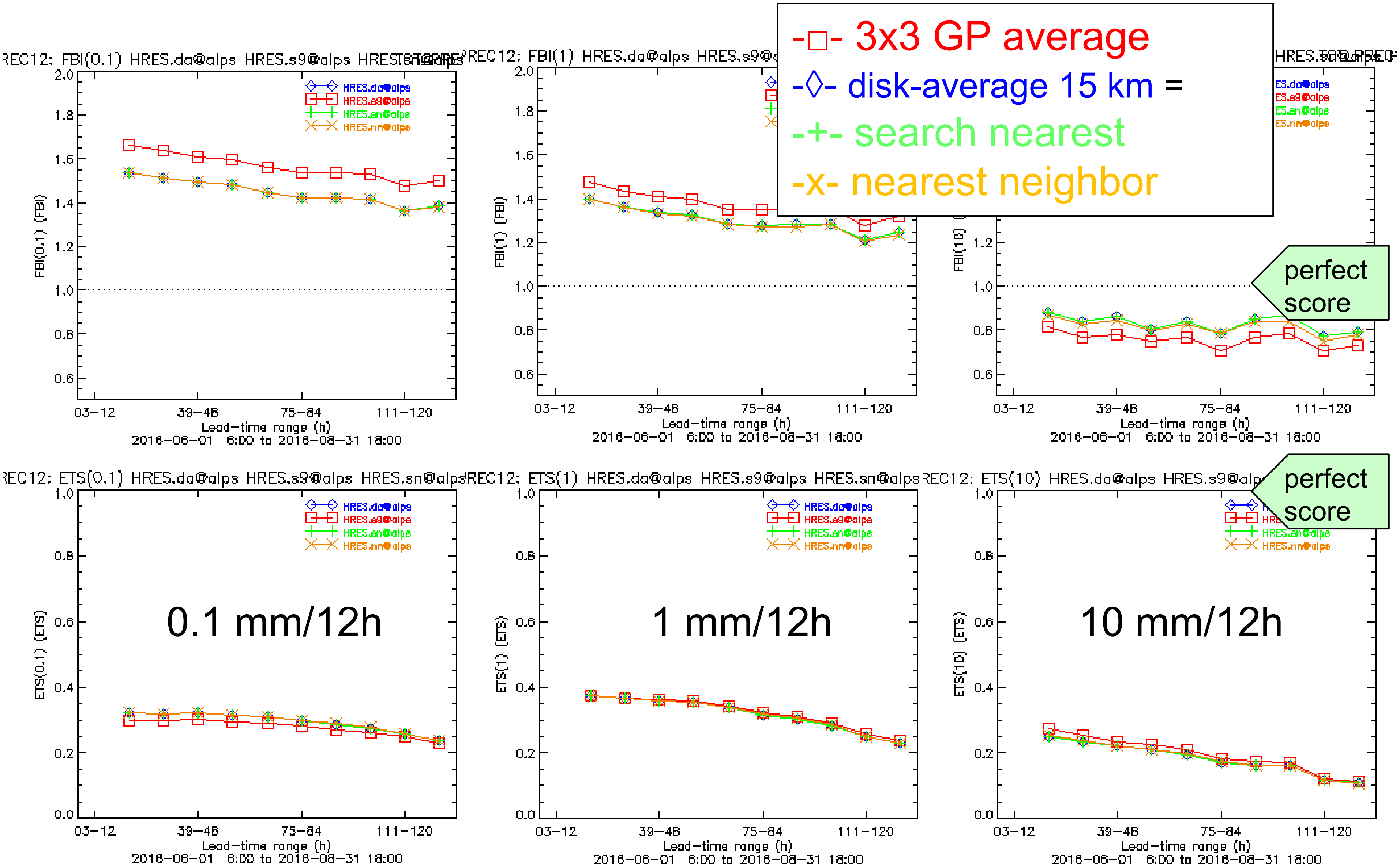
Winter 2017/18



COSMO-1 Summer 2016 (Switzerland) 1h Precip



HRES summer 2016, “alps” domain




Verif
Software
NWP
Test Suite
WG5
activities
Application
Verification
Methods

PP INSPECT is finished, but the work is going on

Anastasia Bundel

Final report is prepared and sent to the SMC members for eventual comments, critics, or suggestions

MesoVICT: Regular participation of COSMO people in monthly MesoVICT webinars. Final meeting in 2019, possibility(?) during to hold it around COSMO GM in Rome to acknowledge COSMO contribution

WG5

Processing large volumes of data: CRA for STEPS nowcasting at Roshydromet (Anatoly Muraviev)

- ***The core of the system is the statistical STEPS scheme (Short Term Ensemble Prediction System) (Bowler N. et al., 2006)***
- ***Verification period: May-September 2017***
- ***9 radars in Central Russia***
- ***Forecasts for intense precipitation areas only were analyzed (169 situations)***
- ***10 min time step until 3 h***
- ***Grid size of about 2 km***
- ***SpatialVx was used to indentify objects and to calculate CRA scores.***
- ***The objects with areas less than 35*35 grid points (about 70x70 km) and lager than 128x128 grid points (about 250x250 km) were excluded from analysis using min.size and max.size option in FeatureFinder function. The radius of averaging for convolution smoothing was chosen empirically as 9 grid points (18 km)***
- ***Precipitation threshold for object recognition showed here is 1 mm/h***

WG5

Statistics of CRA object centroid longitude shift: mean, median, quartiles, max, and min.

RAKU: CRA x_shift statistics OVER SITUATIONS

lead	min	q25	med	mean	q75	max
1	-15.280	-7.620	-2.680	-0.609	3.990	16.730
2	-23.480	-9.408	-2.505	-1.519	5.140	18.240
3	-25.920	-9.505	-5.345	-3.125	0.588	30.640
4	-33.920	-9.970	-0.940	-0.467	6.923	32.660
5	-50.030	-17.730	-5.530	-6.011	7.145	32.350
6	-43.530	-19.230	-4.490	-1.933	18.115	30.640
7	-44.900	-19.790	0.800	-2.318	13.760	41.780
8	-44.210	-13.613	-2.990	-2.133	10.455	41.160
9	-48.090	-12.640	-8.150	-5.302	10.428	27.000

Critical shift value empirically defined as 35 grid points (radius of a smallest round object)

Red : shifts for all objects don't exceed the critical value

Green : shifts for not less than 50% of objects don't exceed the critical value (until 90 min lead time)

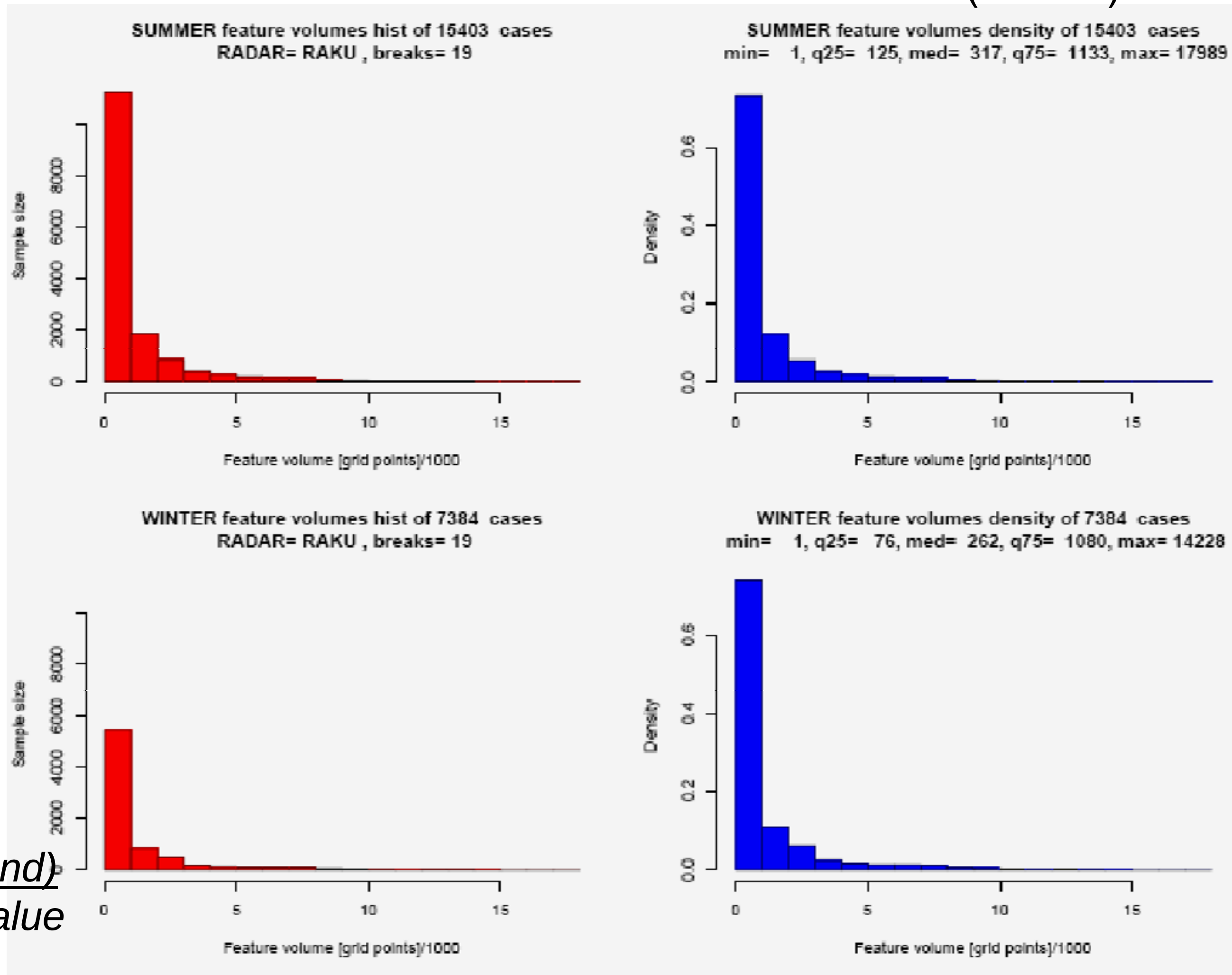
Systematic shifts and other CRA error components can be determined in such a way.

Distribution histograms of object sizes

One radar (RAKU) data

Number of objects (>=1 mm) in winter is twice less than in summer, but the relative frequencies of sizes are comparable

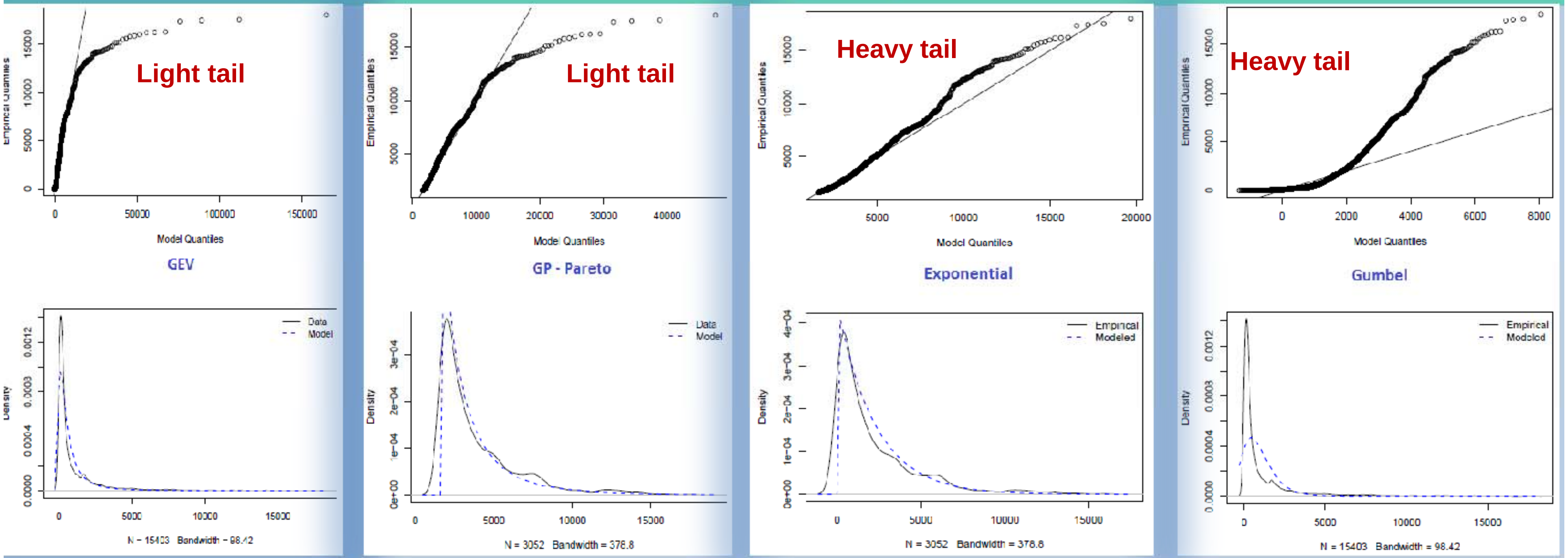
Upper row for warm period,
Lower row, cold period
Red : number of objects,
Blue: relative frequencies



R extRemes package(E.Gilleland)
Module fevd – fitting extreme value distributions to data, plotting histograms, parameter estimation, distribution densities, fitting of tails, quantile-quantile (QQ) plots, statistical significance, …

OBJECT SIZE in GRID POINTS*1000

Fitting object size distribution using *fevd* : QQ-plots and probability density functions (pdf) for warm period (RAKU radar) for a threshold of 1600 grid points



Gumbel distribution doesn't fit well, as well as Poisson (not shown)
Q-Q plots: GEV and GP-Pareto, acceptable up to ~13000 grid points area, were chosen for further analysis
Here, the serial correlation (development of the same object in time) was not taken into account. Now, the study is ongoing to decimate the series of forecasts based on mesoscale time-space structure

WG5

Collaboration Terrain: High Impact Weather verification
WG4, WG5, WG7

Basic Idea: Understanding the forecast quality in high impact weather situations. Important is also to **verify** such weather forecasts in a **meaningful** way to the end users by exploring metrics of the value of forecasts in decision making.
Goal: Provide COSMO Community with an overview of forecast methods and forecast evaluation approaches that are linked with high impact weather (not necessarily considered extreme to all users).

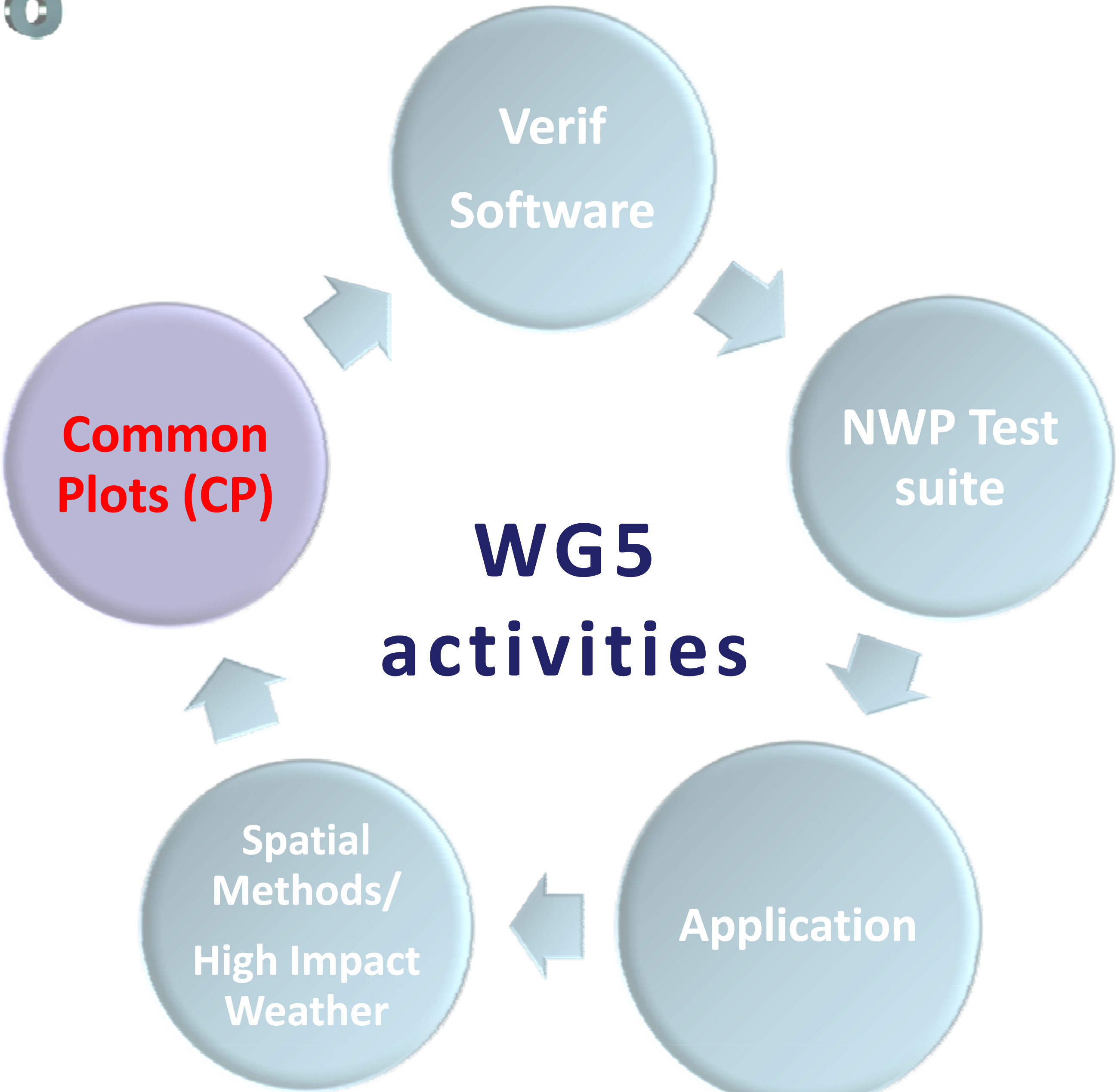
In line with **WMO** focus of research through HIWeather project

Weather parameters of interest: precipitation (intensity, thunderstorms), wind (+gusts), min-max temperature (persistence), visibility (fog)

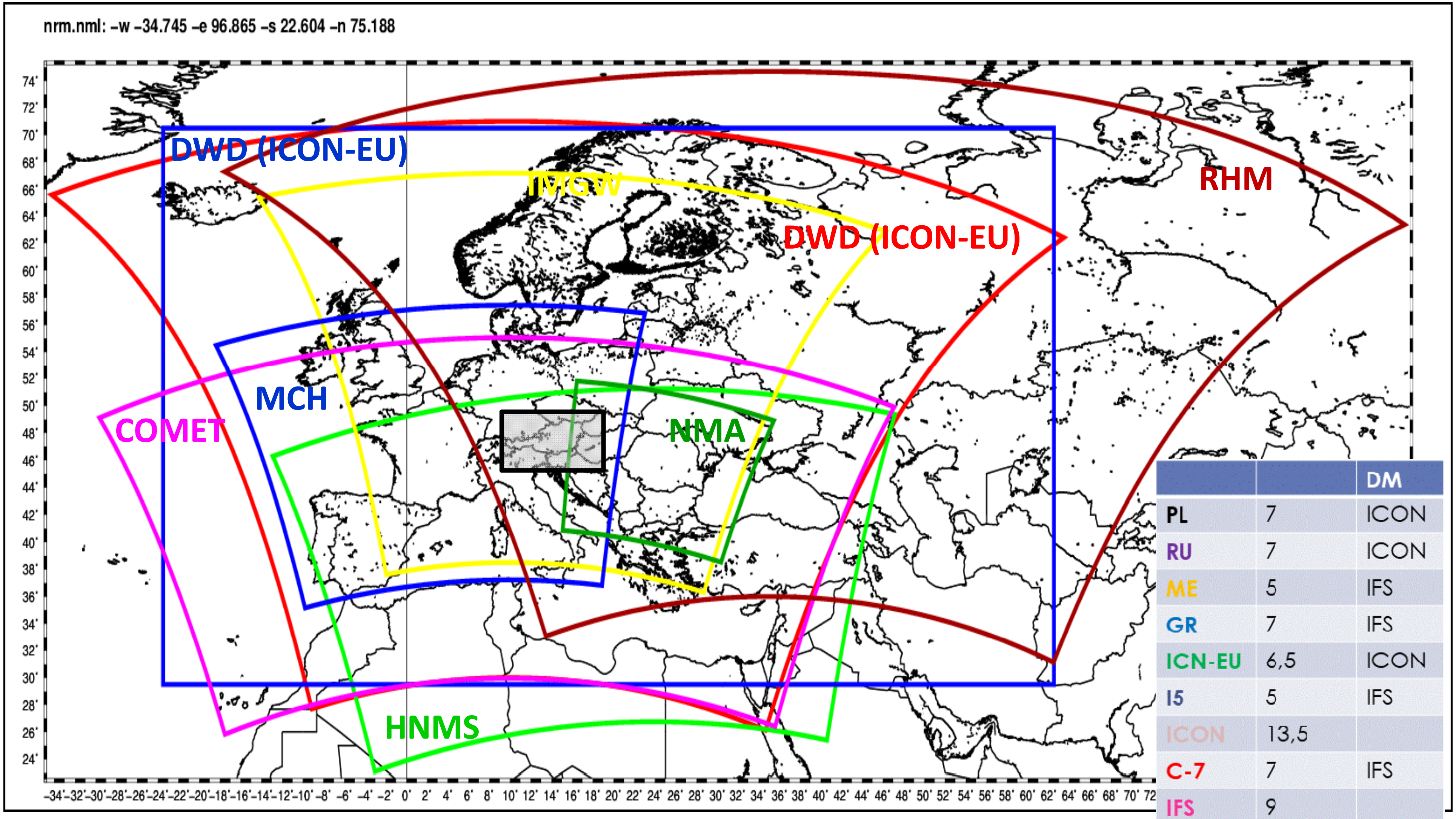
Impact/prediction relation: it can be a mismatch between what models can provide and what information warnings need to be made for (lightning, hail, wind gusts, fog, …)

Analyze the possibility to formalize working Tasks between WGs or even define a PP
Joint session: Tuesday 5pm, Small Conf Hall **(bring your ideas!)**

COSMO
CONSORTIUM FOR SMALL SCALE MODELING


Verif
Software
NWP Test
suite
Application
Spatial
Methods/
High Impact
Weather
Common
Plots (CP)
WG5
activities

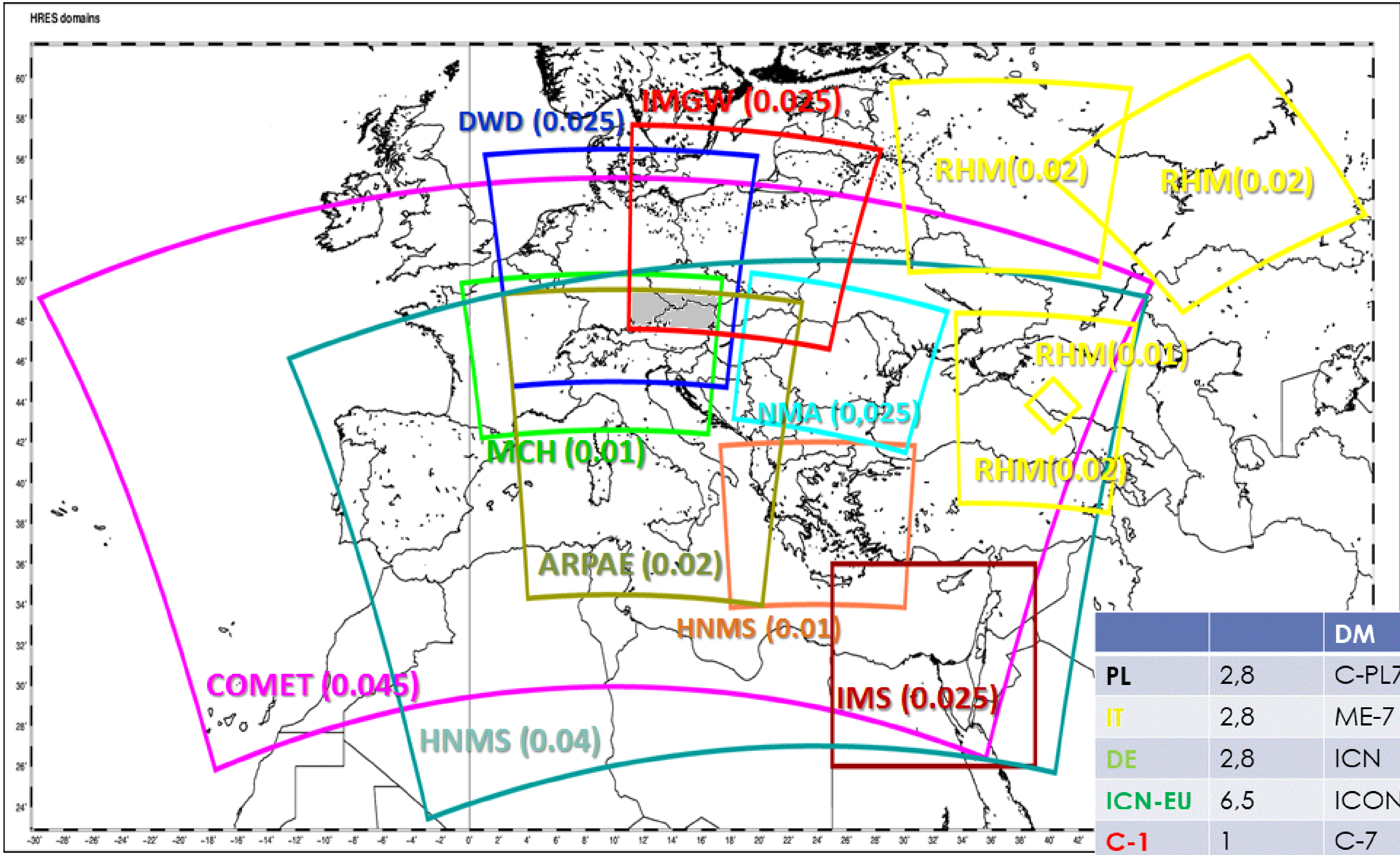
Common Area 1: coarse resolution



		DM
PL	7	ICON
RU	7	ICON
ME	5	IFS
GR	7	IFS
ICN-EU	6,5	ICON
I5	5	IFS
ICON	13,5	
C-7	7	IFS
IFS	9	

COSMO
CONSORTIUM FOR SMALL SCALE MODELING

Common Area 2



		DM
PL	2,8	C-PL7
IT	2,8	ME-7
DE	2,8	ICN
ICN-EU	6,5	ICON
C-1	1	C-7

Common Plots activity for 2018-2019

- Keep the standard parameter scores for CM1 and CM2 Ito maintain long term trends
- Focus on to dedicated verification analysis: model performance of various weather regimes/climate (CM3), performance rose for wind sp/dir (M.S.Tesini)
- Include FSS for CA2 aggregated over cases that warnings were issued (possibly by Arpa-PT – radar obs in gridded format need to become available)
- Abort VERSUS as mandatory for Common Plots production when is necessary. Delivery of txt files with statistical results following the guidelines in the any possible degree

PP CARMA (proposal)

Common Area with Rfdbk/MEC Application

Amalia IRIZA-BURCA (NMA), Flora GOFA (HNMS)

PP CARMA

Goal

- replace the existing VERSUS verification **software environment** with the MEC-Rfdbk software developed by DWD, as a Common Verification Software (CVS) to perform part of the verification activities in the consortium
- main use of the new CVS - production of the Common Plot (CP) verification
- EPS, spatial and other verification - with MEC-Rfdbk (not the purpose of this project) or with other available tools (VERSUS, VAST, etc.) in each service
- addresses the need to perform traditional point verification both for the surface and the upper air
- centralized transfer and visualization of CP statistics on the COSMO wed server (*following NWP Test suite example*)

WG5 Contributing Scientists

Francesco Batignani (CoMET)
Dimitra Boucouvala, HNMS
Anastasia Bundel, RHM
Rodica Dumitrache, NMA
Felix Fundel, DWD
Flora Gofa, HNMS
Amalia Iriza-Burca, NMA
Pirmin Kaufmann, MCH
Alexander Kirsanov, RHM
Xavier Lapillonne, MCH
Joanna Linkowska, IMGW
B. Maco (NMA)
Elena Oberto, ARPA-PT
Ulrich Pflüger, DWD
Maria Stefania Tesini, ARPAE
Naima Vela, ARPA-PT
Alon Shtivelman, IMS