WG5 Verification and Case studies

Overview of activities

Flora Gofa

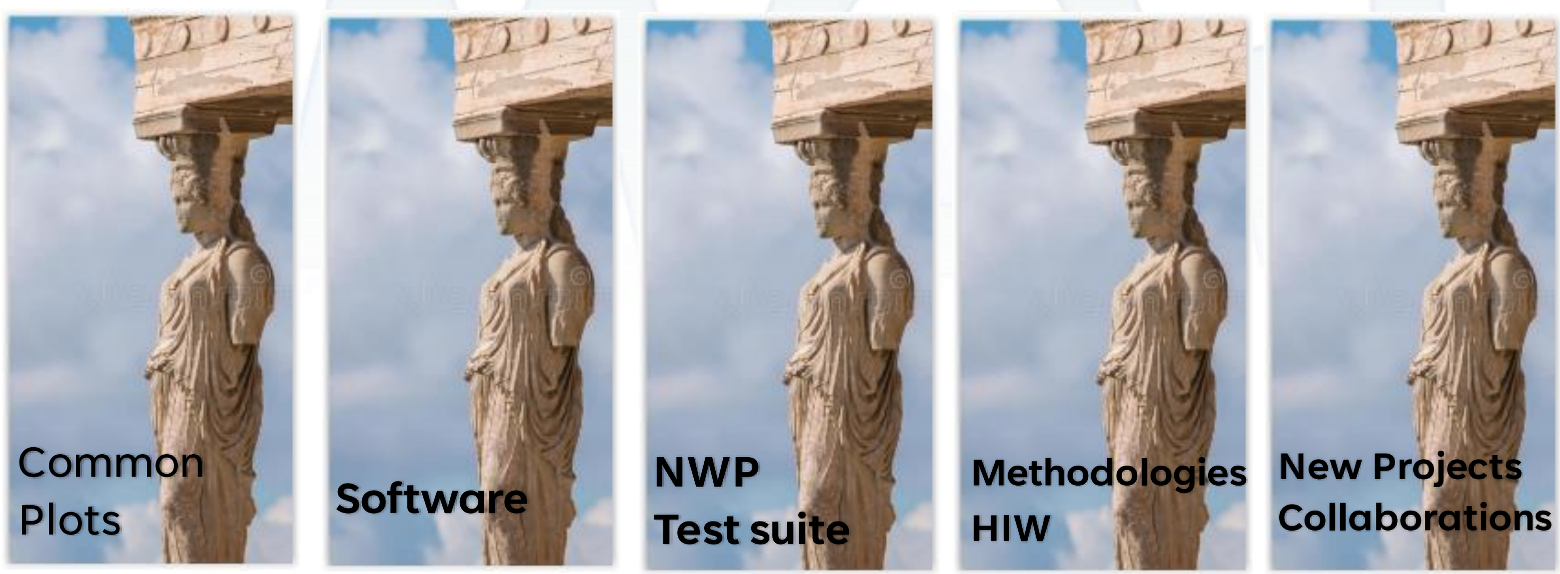
CONSORTIUM FOR SMALL SCALE MODELING COSMO

ΕΘΝΙΚΗ ΜΕΤΕΩΡΟΛΟΓΙΚΗ ΥΠΗΡΕΣΙΑ
HELLENIC NATIONAL METEOROLOGICAL SERVICE

WG5 Guidelines

https://www.cosmo-model.org/content/consortium/reports/WG5_Guidelines_2021.pdf

- ❑ Common Verification framework: developments concerning EPS verification with MEC-Rfdbk and its conditional verification capabilities. *PP-CARMA, PP-CARMENs*
- ❑ Exploitation of spatial verification techniques: Analyse how methods relate to one another, how each method works, what information could be gleaned from each method, and whether a given method actually conveys any useful information *PP-INSPECT, PP-AWARE*
- ❑ Severe and High Impact Weather. Forecast methods and verification are important aspects of any HIW consideration. *PP-AWARE addresses issues* such the representation in the observations of HIW, importance of observation uncertainty, systematic and stochastic errors of HIW forecasts and their sensitivity to model resolution.
- ❑ Utilization of non-conventional observational datasets: obs often do not permit characterization of the phenomenon of interest for objective verification. Discussion on new PT on crowdsource data potential atNWP



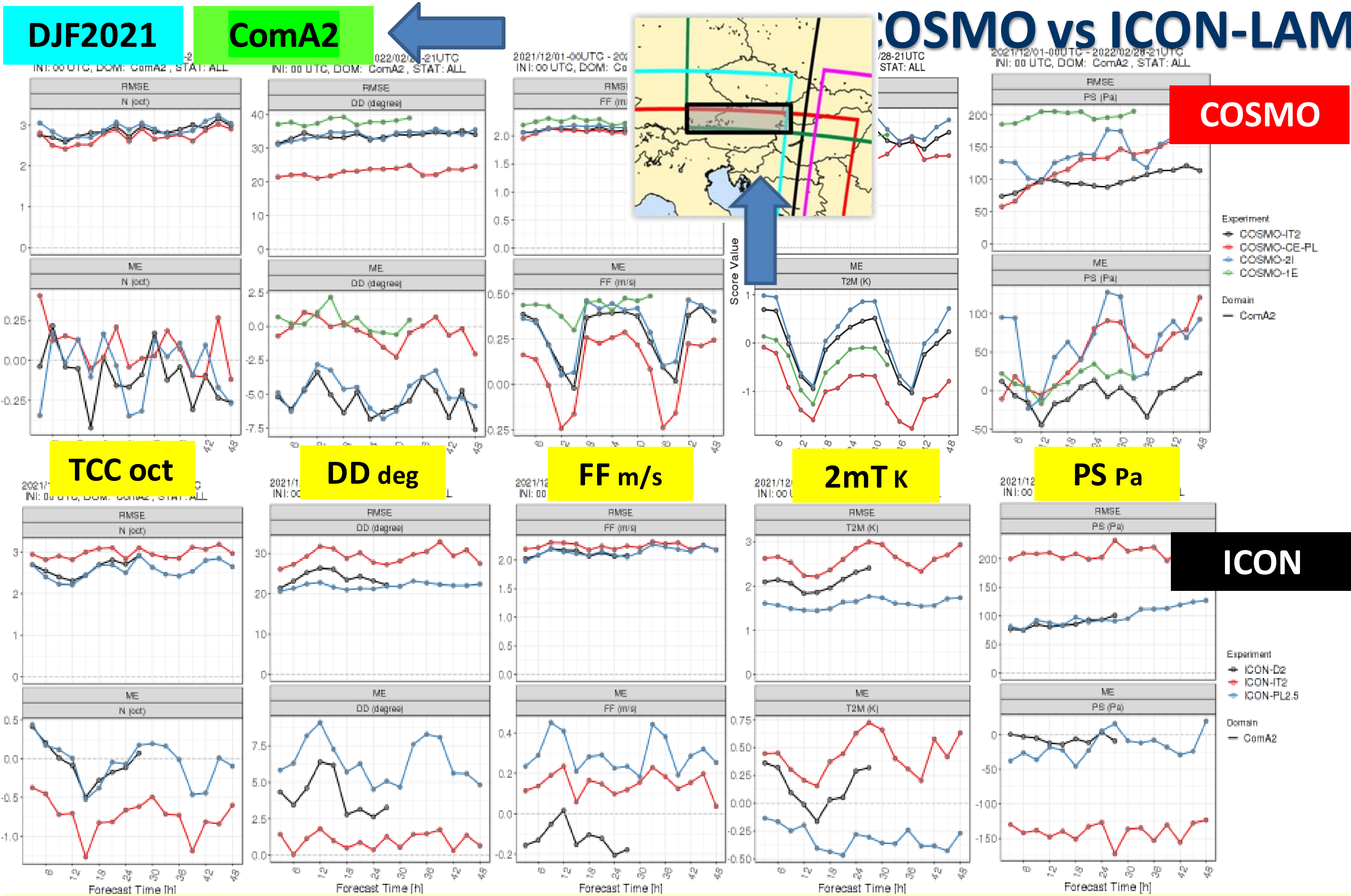
***calculation and representation of verification results** of statistical indices derived using **operational ICON-LAM and/or COSMO model** in each service.*

*Domain, resolution, statistical scores/methods, frequency and graphical representation, are decided on an **annual basis***

CP activity: operational models

COARSE | FINE

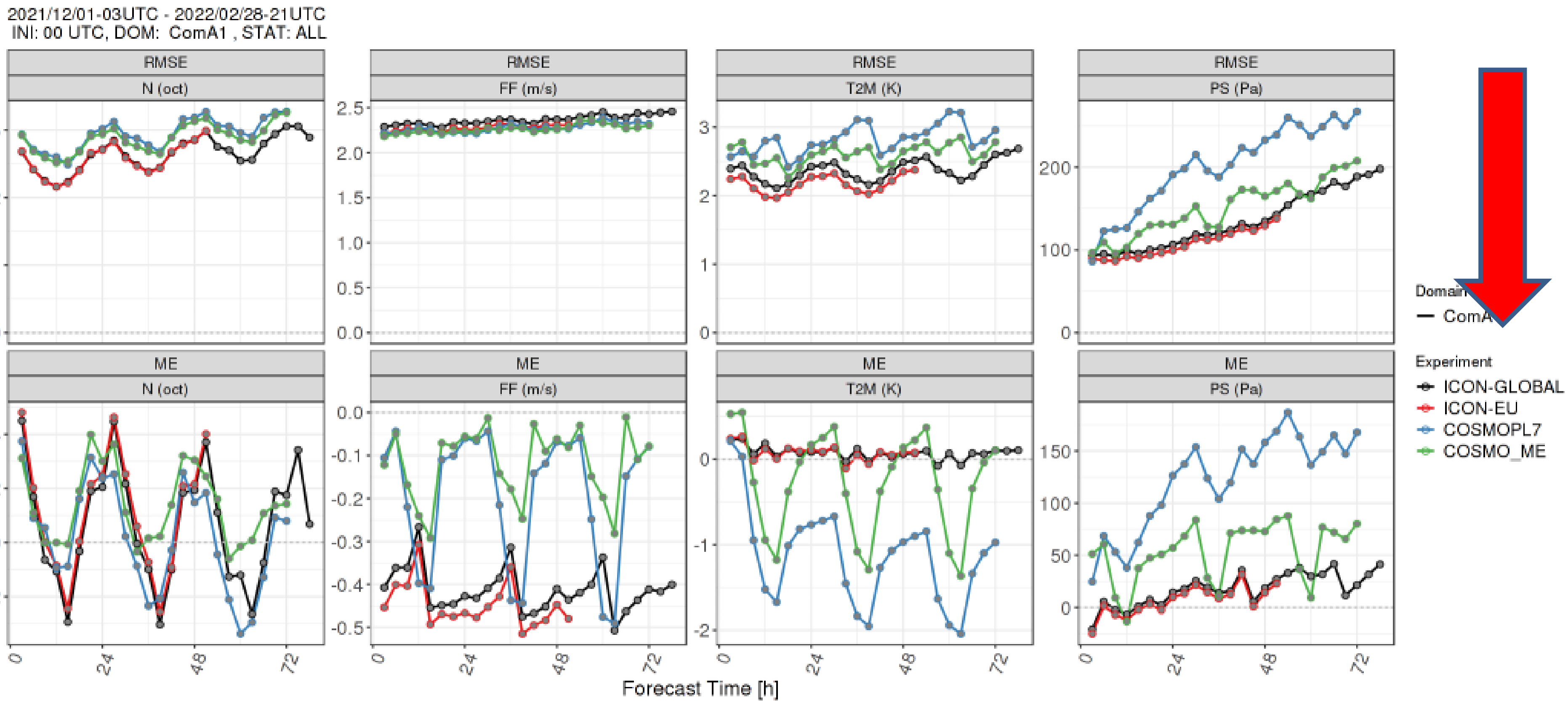
- **DWD:** **ICON-EU (0.0625)**, **ICON-D2 (0.02)**, **ICON-D2-EPS (0.02)**
- **COMET:** **COSMO-ME (0.045), COSMO-IT (0.02), ICON-IT (0.02)**, **COSMO-ME-EPS (0.0625), COSMOIT-EPS (0.02)**
- **IMGW-PIB:** **COSMO-PL7 (0.0625), COSMO-CE-PL2k8 (0.025), ICON-PL (0.025), COSMO-PL2.8-eps (0.025)**
- **HNMS:** **COSMO-GR4 (0.04), ICON-GR (0.025)**
- **MCH:** **COSMO-1E (0.01), COSMO-2E (0.02),** ICON-1, ICON-2 in preoperational phase
- **IMS:** **ICON-IL (0.025), ICON-IL-EPS (0.025)**
- **NMA:** **COSMO-RO7 (0.0625), COSMO-RO3 (0.025), ICON-RO2p8 (0.025)**
- **ARPAE-SIMC:** **COSMO-5M (0.045), COSMO-2I (0.02), COSMO-2I-EPS (0.02), ICON-2I,** in preoperational phase



OVERALL: Smaller amplitudes of BIAS diurnal cycle. reduced RMSEs in ICON-LAMs;
Reduction of **T2m** error, FF with no smaller changes but with reduction in error in DD partially associated with Pa error reduction. TCC performance not clearly improved. Spread among ICON-LAMs in performance

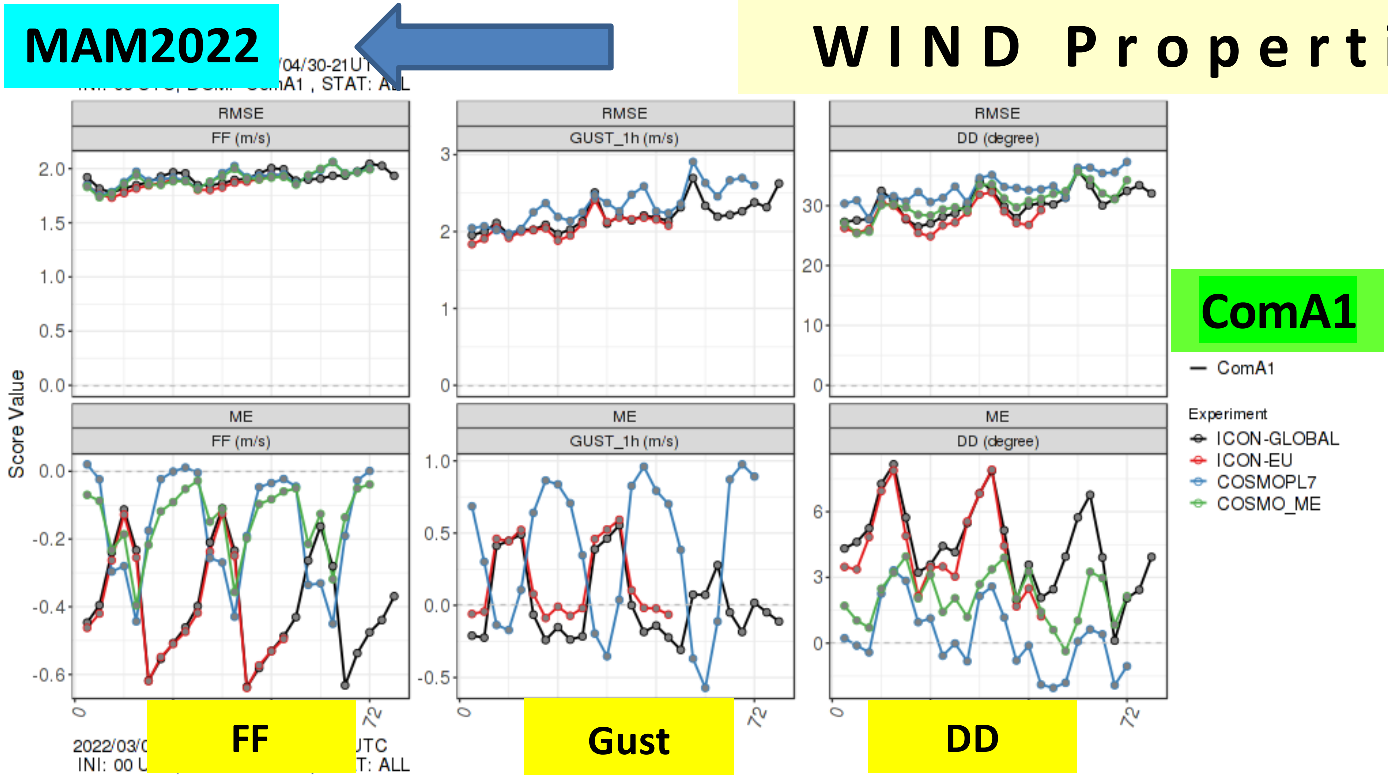
DJF2022 ComA1

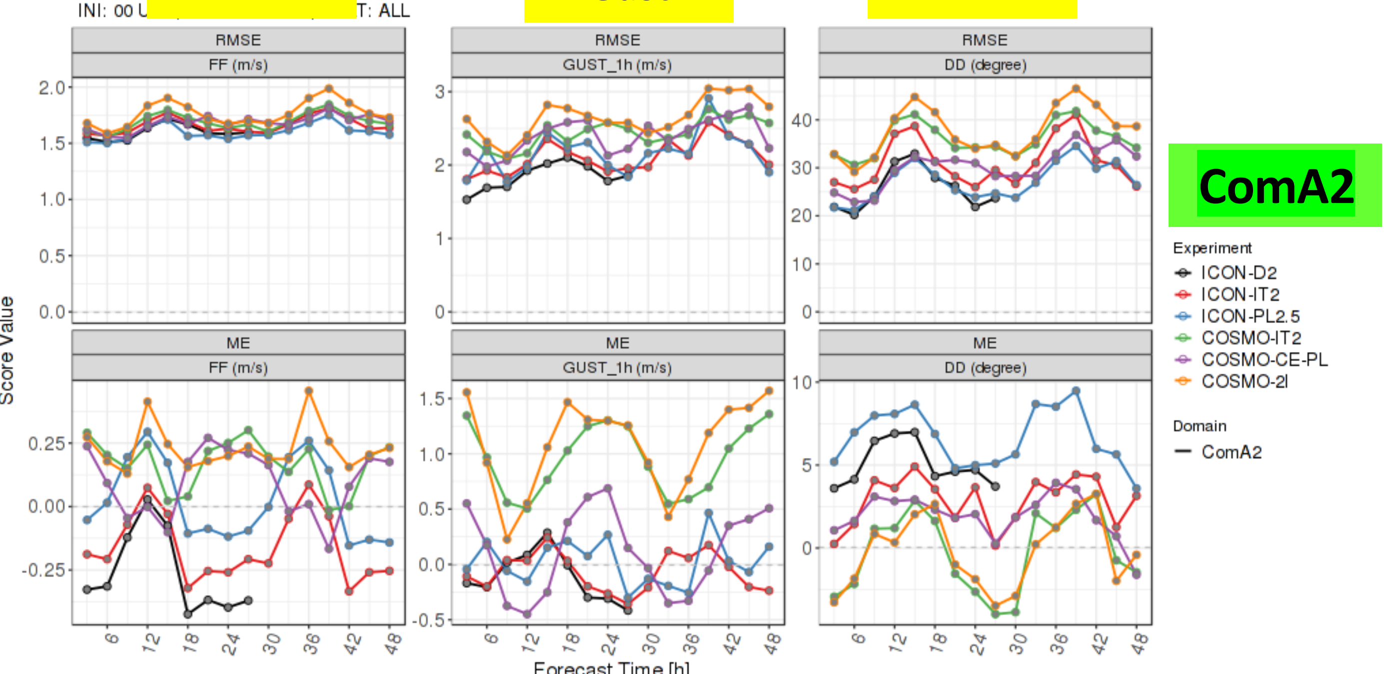
TCC oct FF m/s 2mT K PS Pa



Coarser models: 2mT and Pa clearly improved, FF change in error phase, TCC large biases for all models

WIND Properties





Resolution-dependence and model improvement effect on wind properties

Gust: resolution effect is smaller but improvement in ICON performance more clear

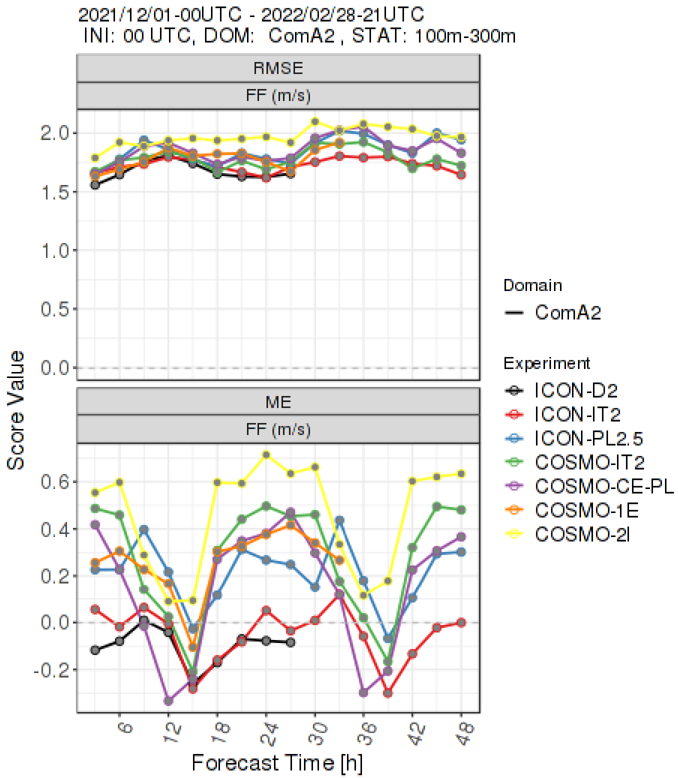
At higher resolution tendency of ICON-LAMs to underestimate wind more.

Error has diurnal cycle with differences in ME maxima phase among COSMO/ICONs

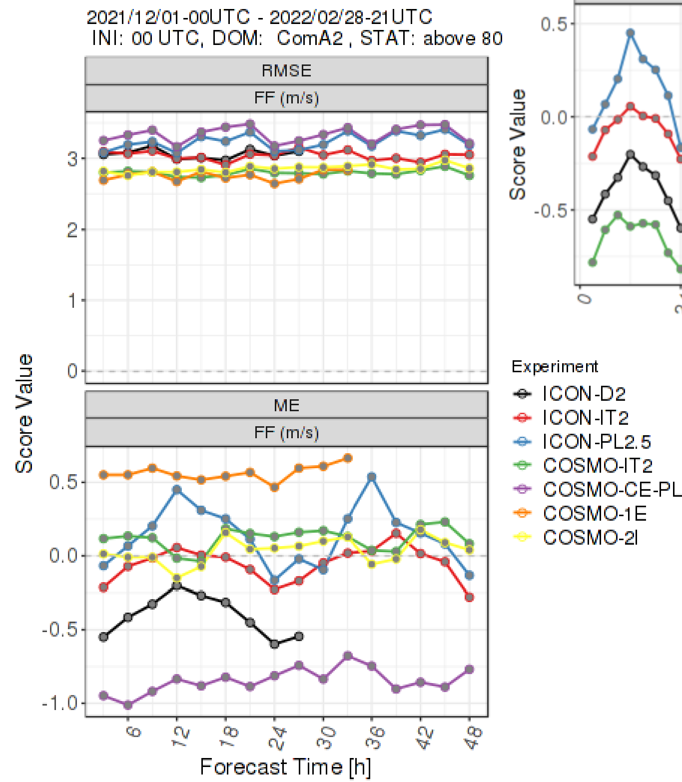
ComA2 MAM2022

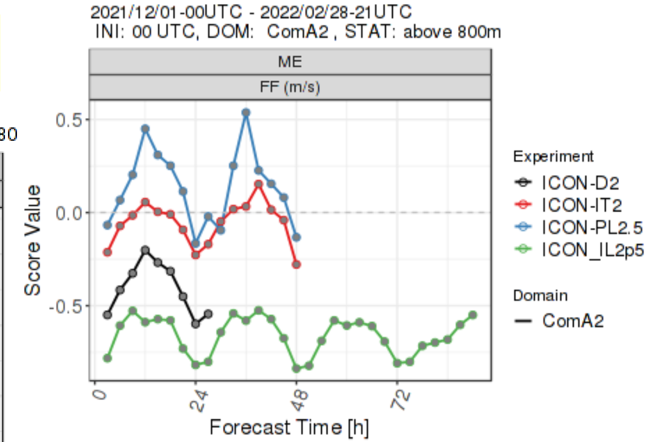
Wind Speed – Elevation

Lower Elevation



Higher Elevation

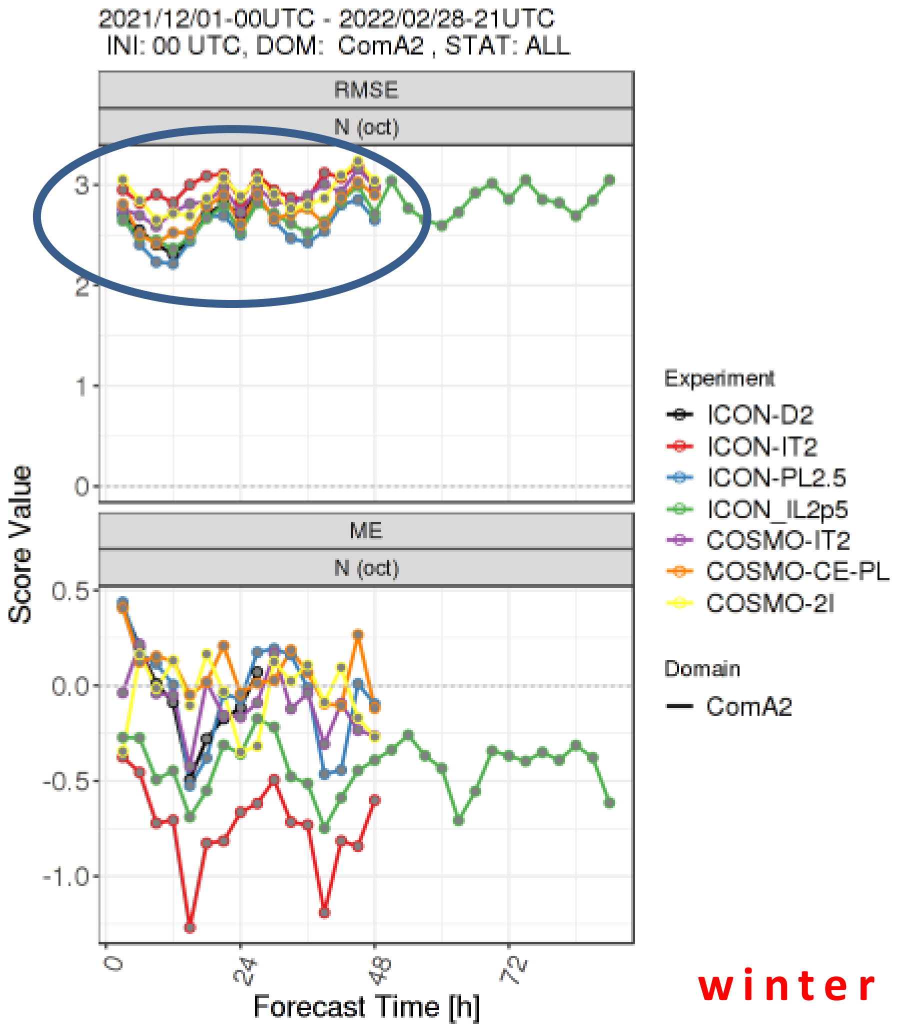




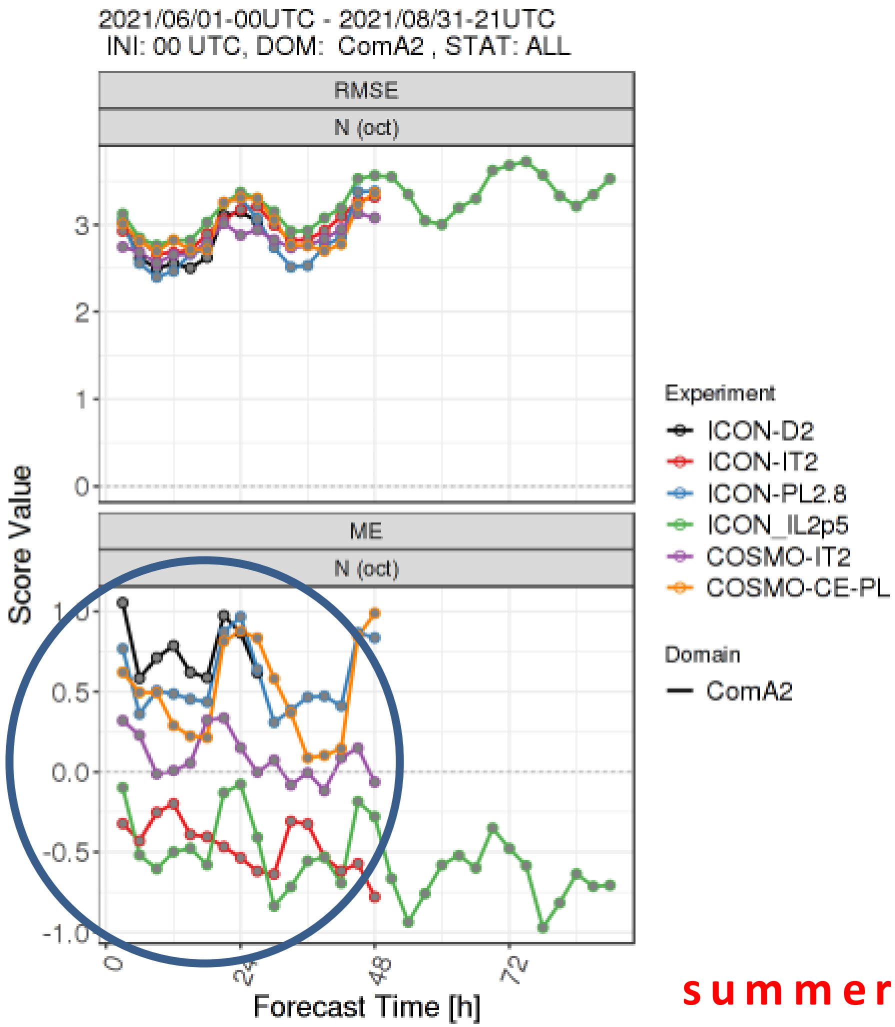
Clear altitude dependence in performance (RMSE) in all models.
FF **RMSE grows in higher elevation points**, with a general tendency to be underestimated
While RMSE error is similar among models in each range (low or high altitude), ICON-LAMs more consistently underestimate FF **in stations above 800m**

Total Cloud Cover (SYNOP)

ComA2



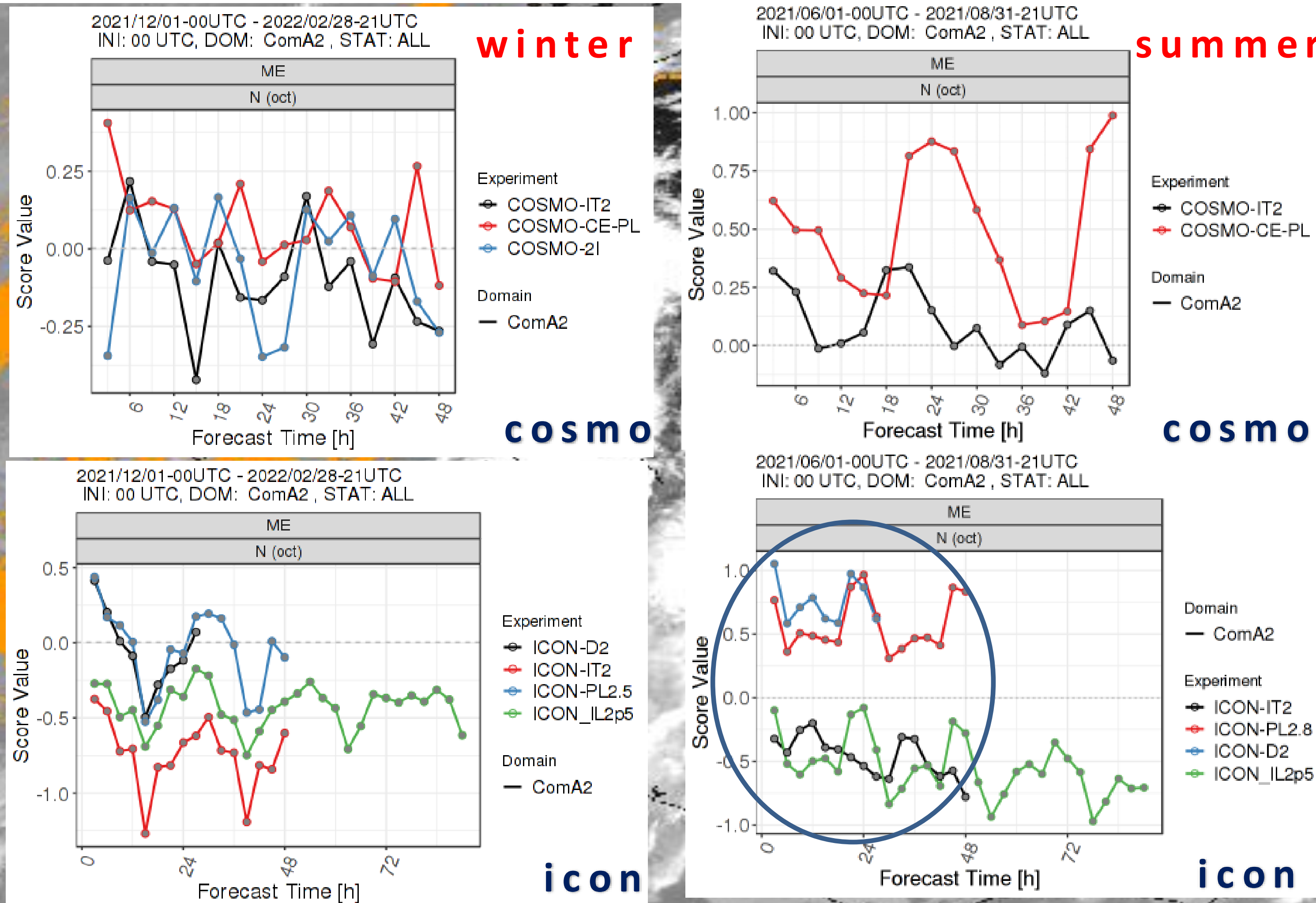
winter



summer

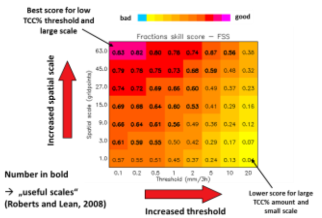
- Large error, already in forecast day 1 (2.5-3oct)
- Underestimation of cloudiness by most models mainly during afternoon in winter, partial overestimation in summer mainly during night hours

Verification against SYNOP: COSMO/ICON

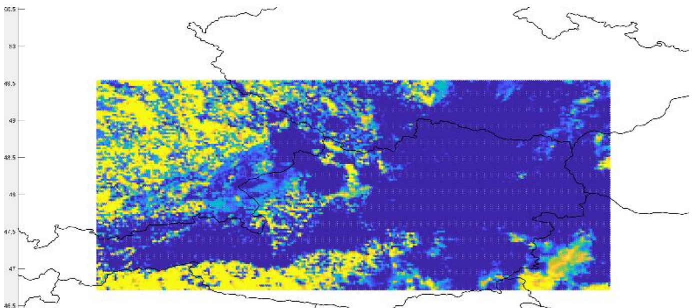


- During winter, large **underestimation** of ICON models systematically seems to appear during afternoon
- In summer, strong **overestimation** by COSMO models, while for ICON-LAMs behaviour is ambiguous

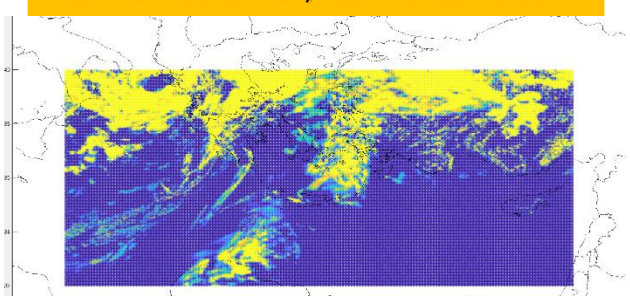
ComPlot: FSS for cloudiness



- Reasoning: Investigate Cloudiness performance over certain areas
- Models:
  - COSMO2I, COSMO I2, ICON-PL2.5, ICON-IL-2p5, ICOND2, ICONEU, ICONGR2.5, COSMOGR4
- Period: more organized from Feb-Jun 2022
- Scores: FSS (more scores could follow in next phase)
- Cumulation: 3h
- Areas: ComA2, Mediterranea (large, mainly over sea)

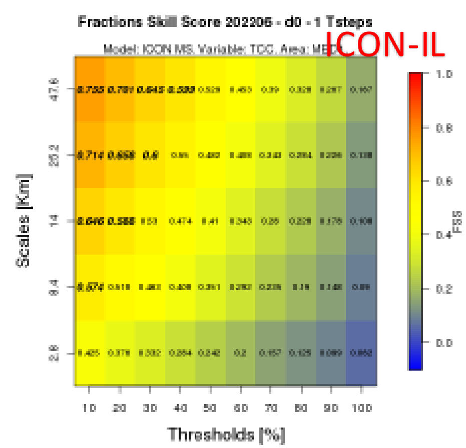


ComA2: *restricted, mountainous*
Lon: 16.000-17.424, Lat: 46.725-49.550

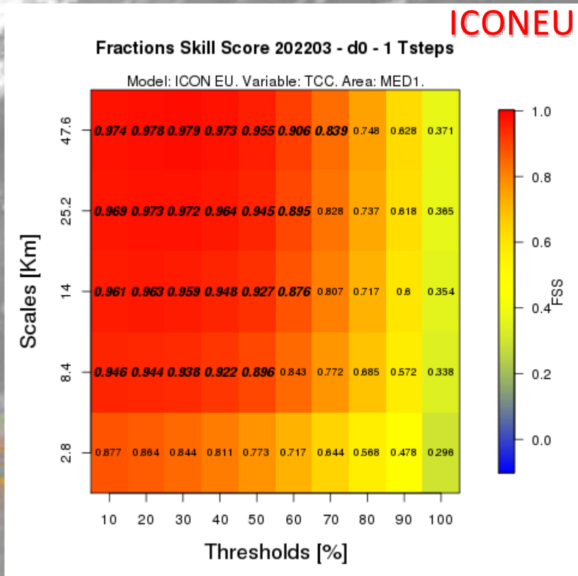


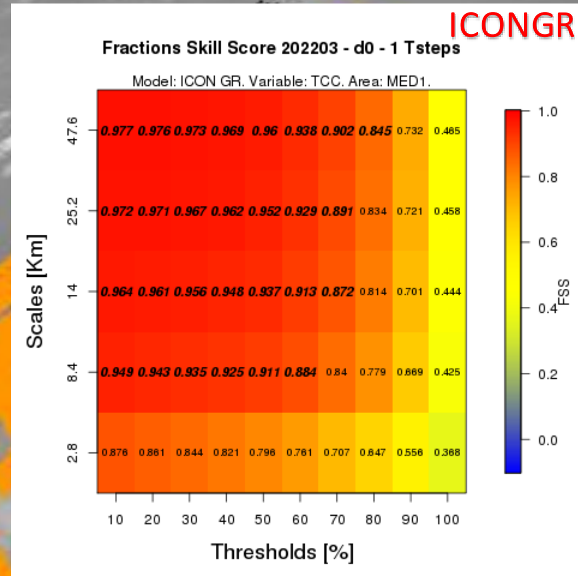
Med1: *Extended, over water*
Lon: 16.00-35.00, Lat: 32.00-40.00

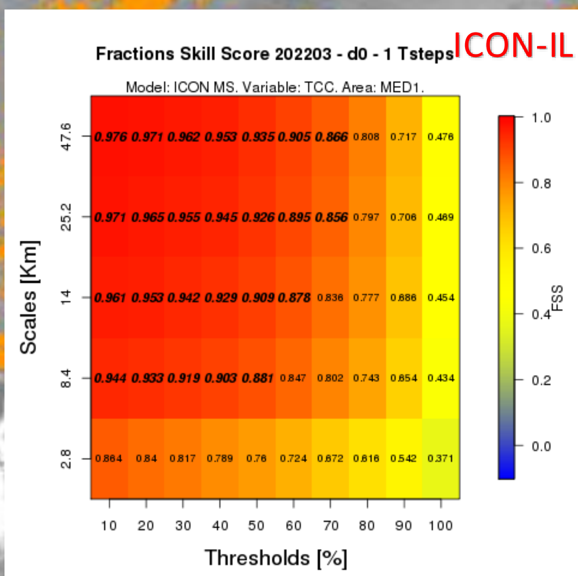
Domain: lon1=-12; lon2=39; lat1=26; lat2=55;
Interpolated resolution: 0.025 degrees.
Adaptation Method: *4km 15min CMA fields average 3 time steps: -15min, 0, +15min multiply by 8 to get an estimation to the cloud cover in octas. Calculated TCC fields provided by P.Khain (thanks)*

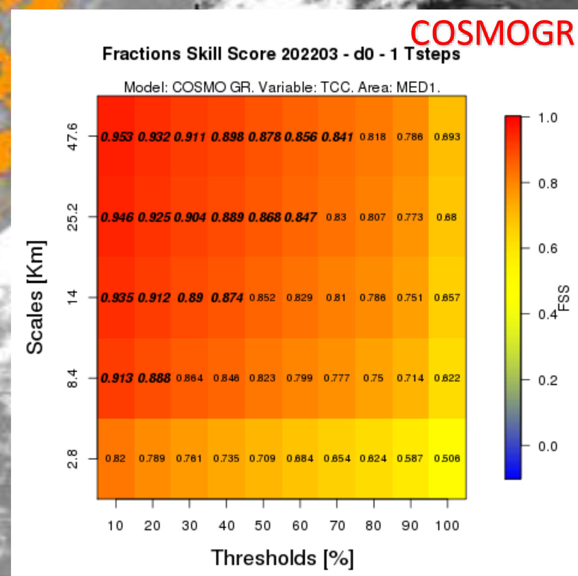


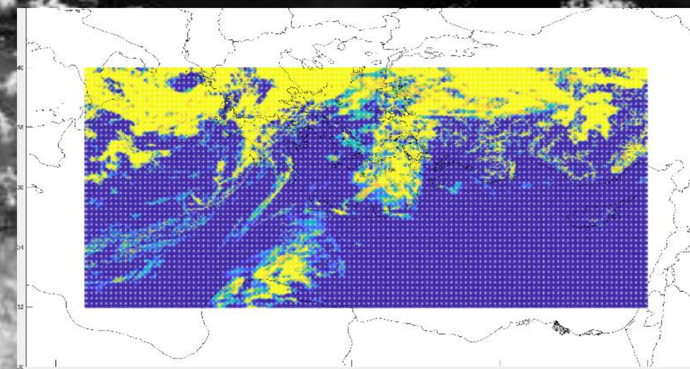
Period: 01-31 Mar 2022
Area: Mediterranean
Sample: 3h timesteps/all
Index: FSS
Models: ICONGR, COSMOGR
ICONEU, ICON-IL2p5



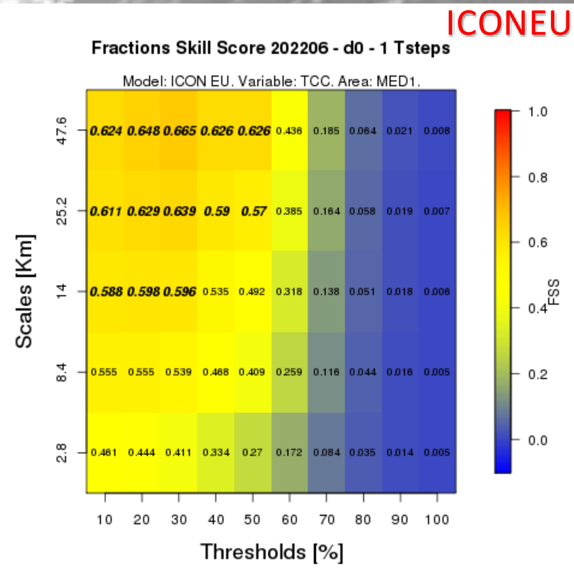


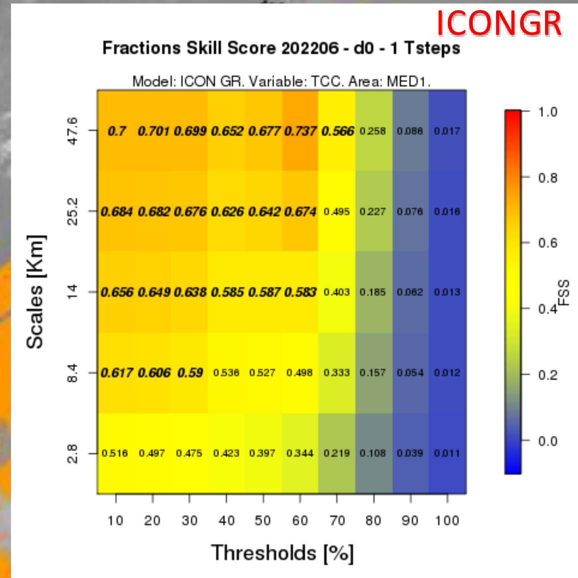






For scales higher than 8km and for lower thresholds, performance is very good for all models ICON-LAMs perform clearly better than COSMOGR for smaller thresholds while COSMOGR gives higher scores than all ICON models when observed we have almost total cloudiness





Verification agst NWCSAF

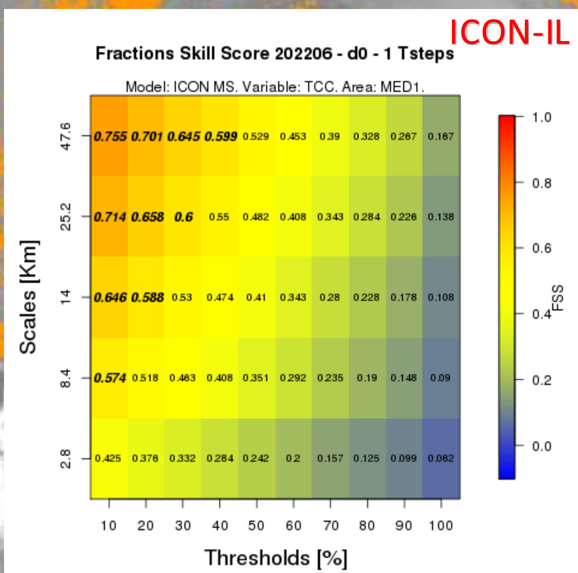
Period: 01-31 Jun 2022

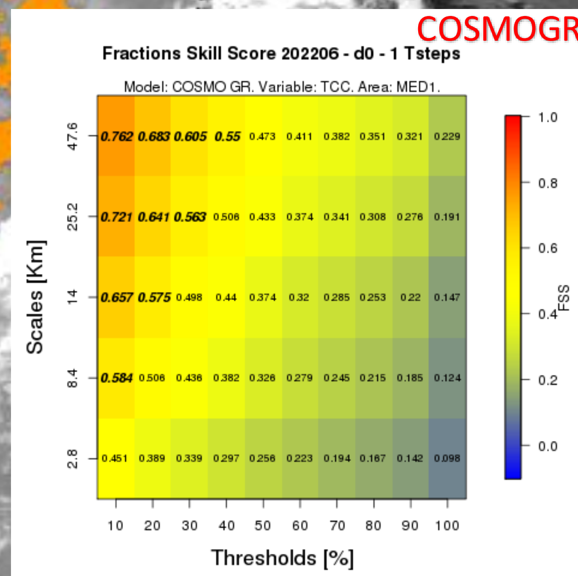
Area: Mediterranean

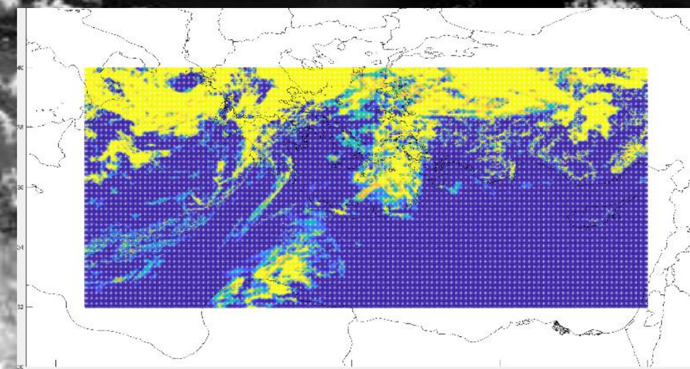
Sample: 3h timesteps

Index: FSS

Models: ICONGR, COSMOGR, ICONEU, ICON-IL2p5

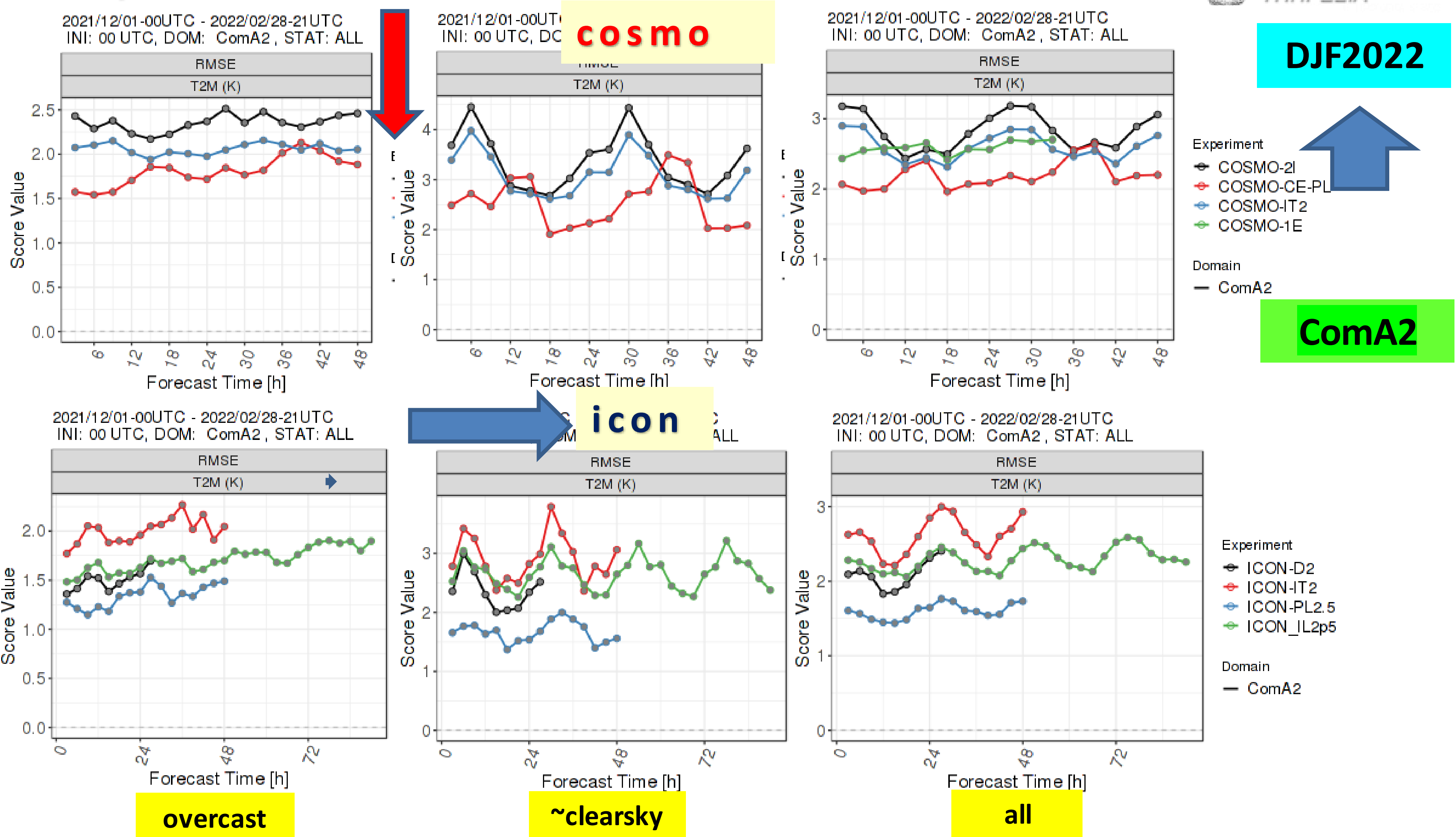






Performance of all models poorer. Useful scales for windows averaged higher than 8km and for less than 30% cloudiness (near clear sky). COSMO at higher TCC% outperforms ICON-LAMs.

Temperature w.r.t. Cloudiness - RMSE



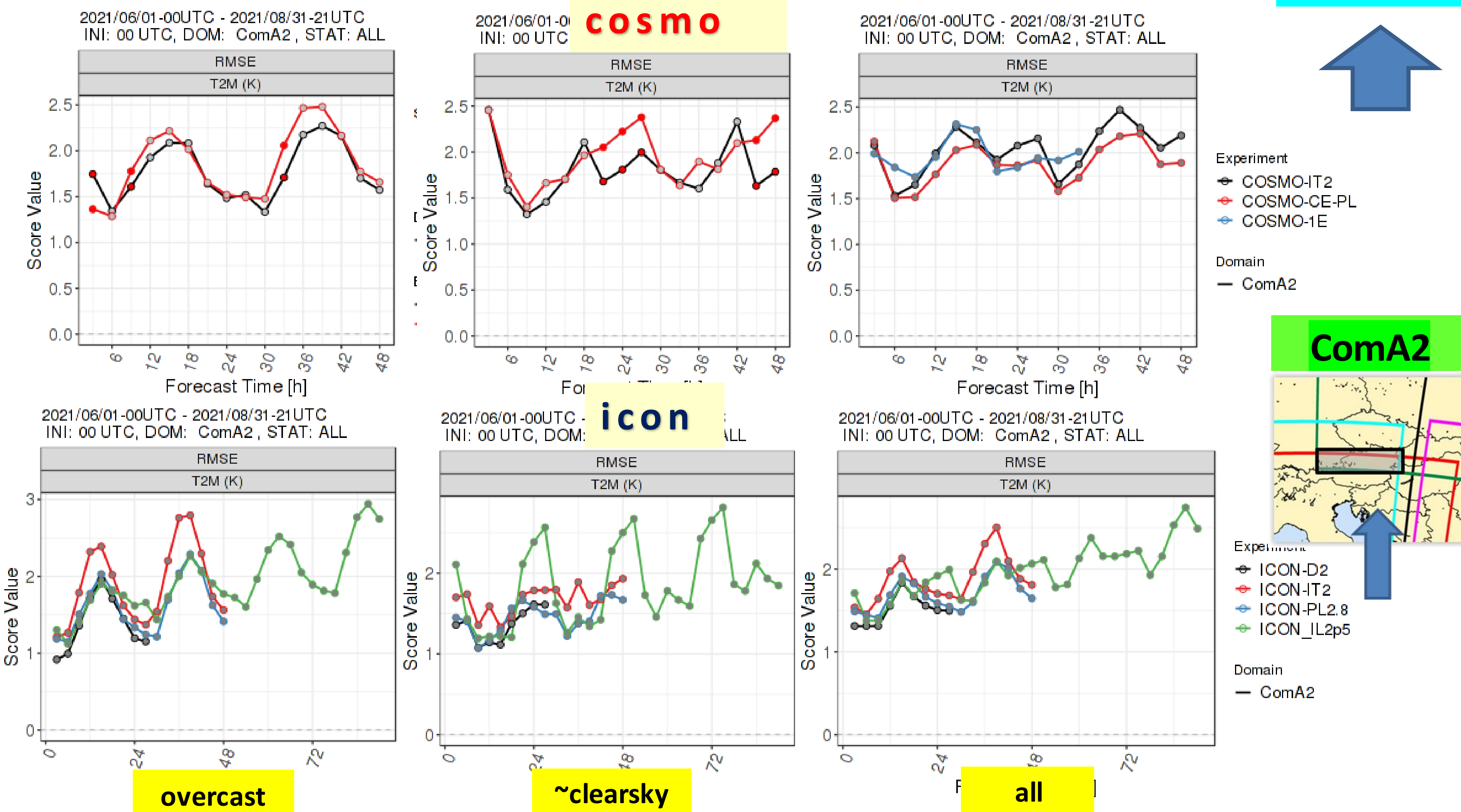
Winter: Higher errors in 2mT in clear sky conditions, and lower errors when overcast conditions only.
Stronger diurnal variability of error with COSMO models in days with few clouds
Significantly improved performance of 2mT with ICON models in the winter in all cases.

Temperature w.r.t. Cloudiness - RMSE

C1. 2mT verification when: (condition based on obs) Total cloud cover observation>= 75% (i.e. 6 in octa)

C2 2mT verification when: (condition based on obs) Total cloud cover observation <= 25% (i.e. 2 in octa)

JJA2021



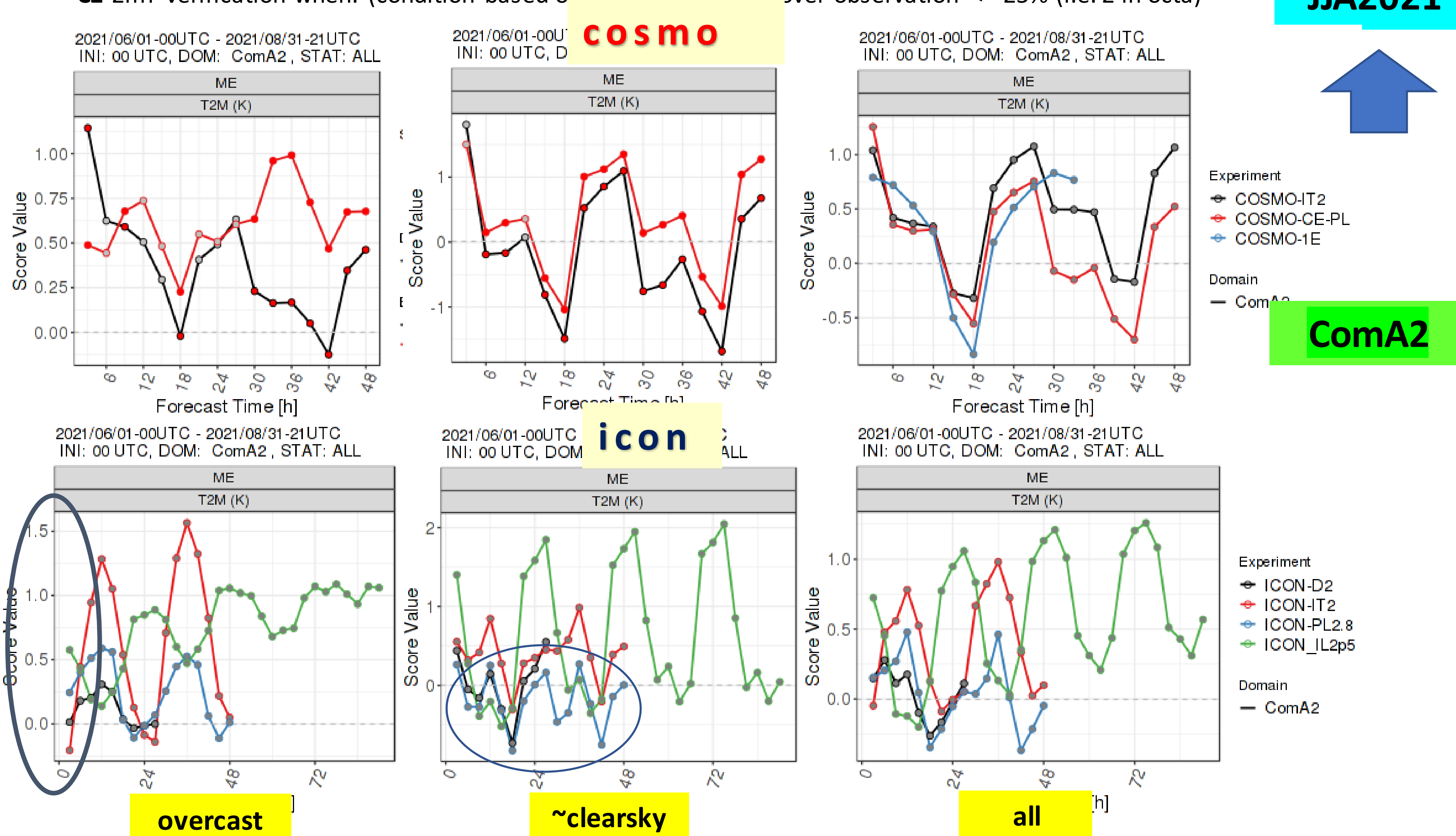
Summer: Smaller impact of cloudiness in 2mT error compared to DJF with larger errors during the afternoon hours. Significant improvement in 2mT error in no cloud conditions mainly with ICON-LAM but not for all implementations (ICON-IL)

Temperature w.r.t. Cloudiness – Mean Error

C1. 2mT verification when: (condition based on obs) Total cloud cover observation>= 75% (i.e. 6 in octa)

C2 2mT verification when: (condition based on obs) Total cloud cover observation <= 25% (i.e. 2 in octa)

JJA2021



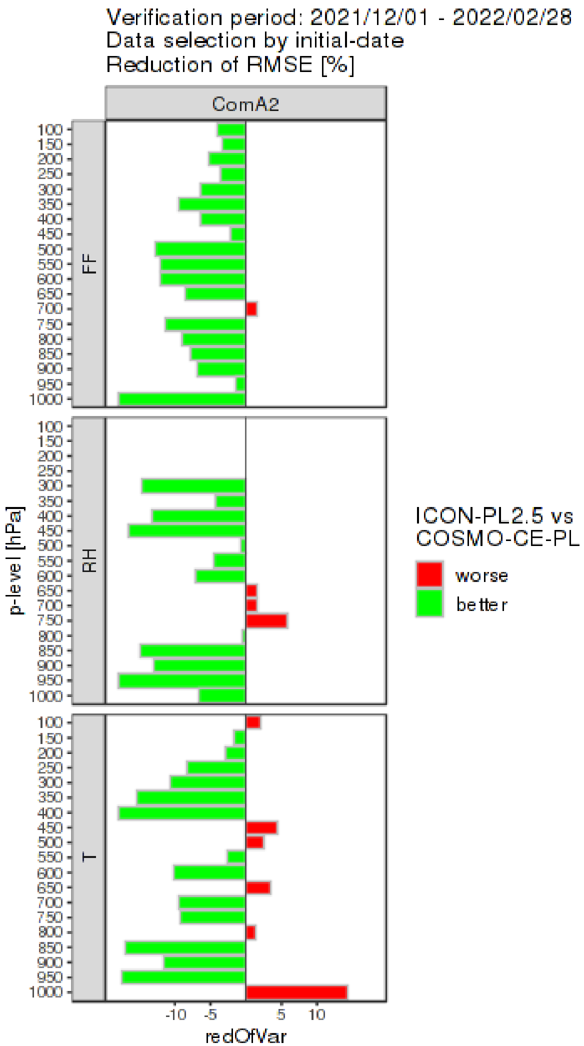
Summer: Overestimation of 2mT, more on cloudy days, which seems to be higher in some ICON-LAMs, consistent with errors in cloudiness for these models too. Worst warming in midday while at night the effect is reverse in clear days with cooler models

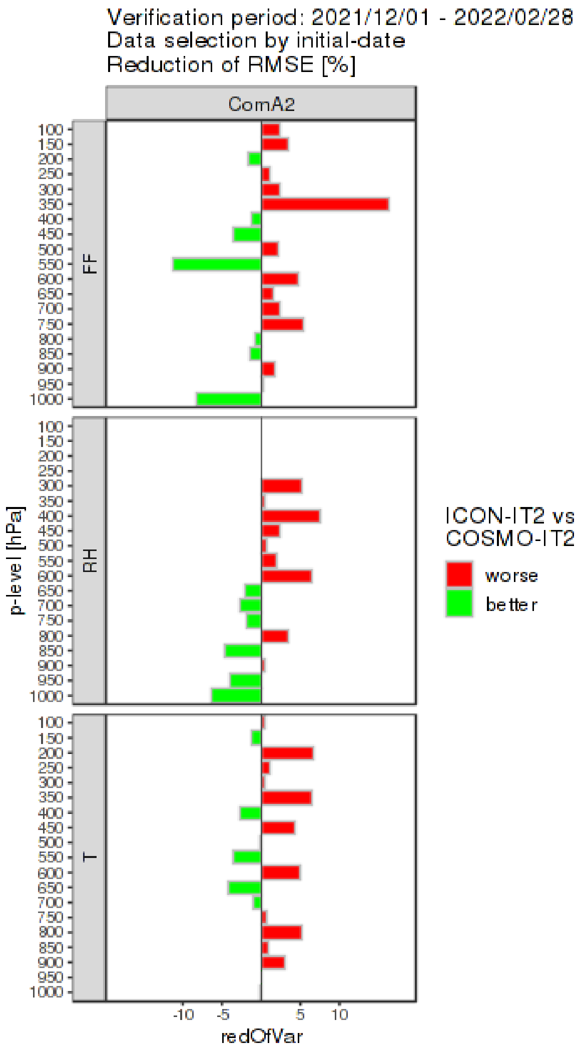
UPPER AIR VERIFICATION

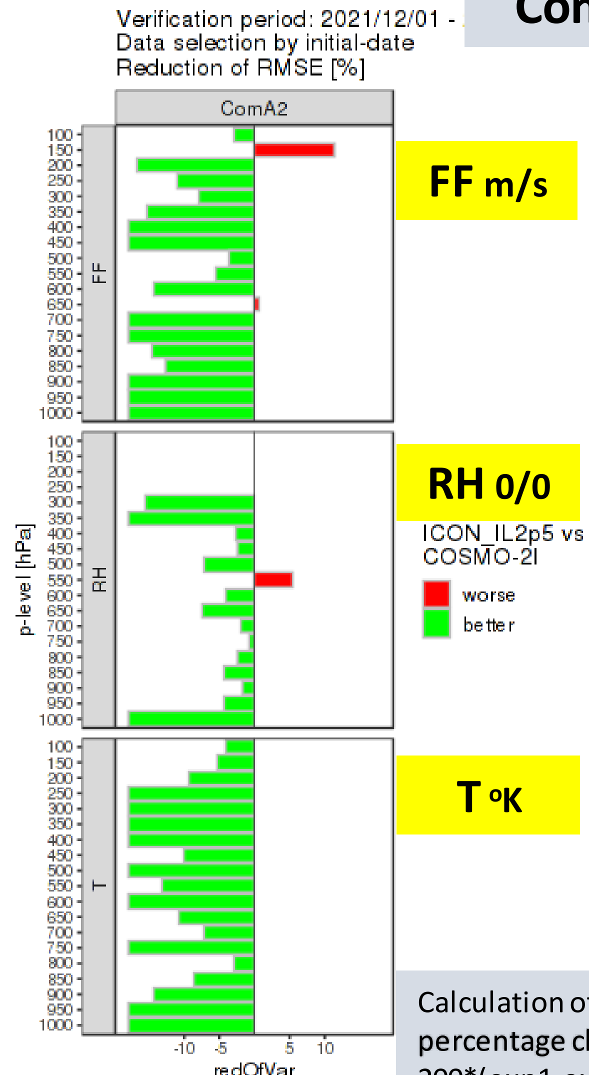
COSMO | ICON

DJF2022

ComA2







FF m/s

RH 0/0

T ºK

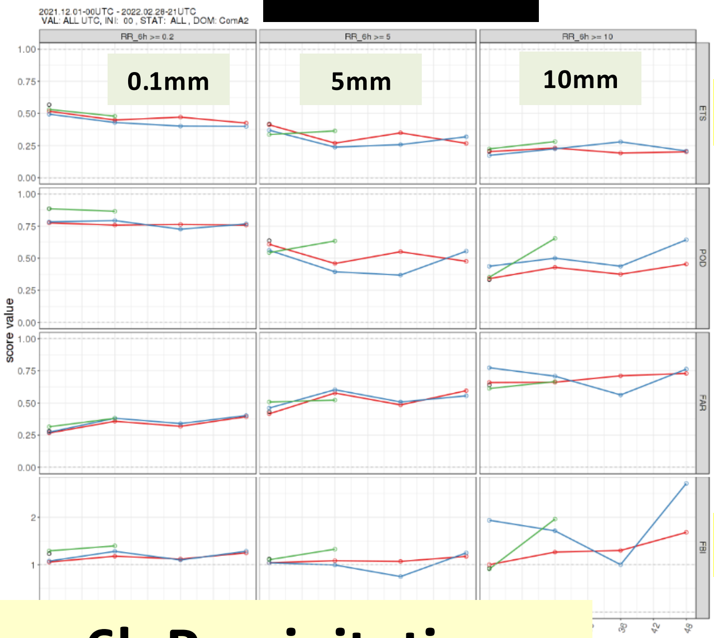
Calculation of percentage change 200*(exp1-exp2) / (exp1+exp2) in RMSE. The scores are aggregated over all initial times and all forecast ranges > 0h.

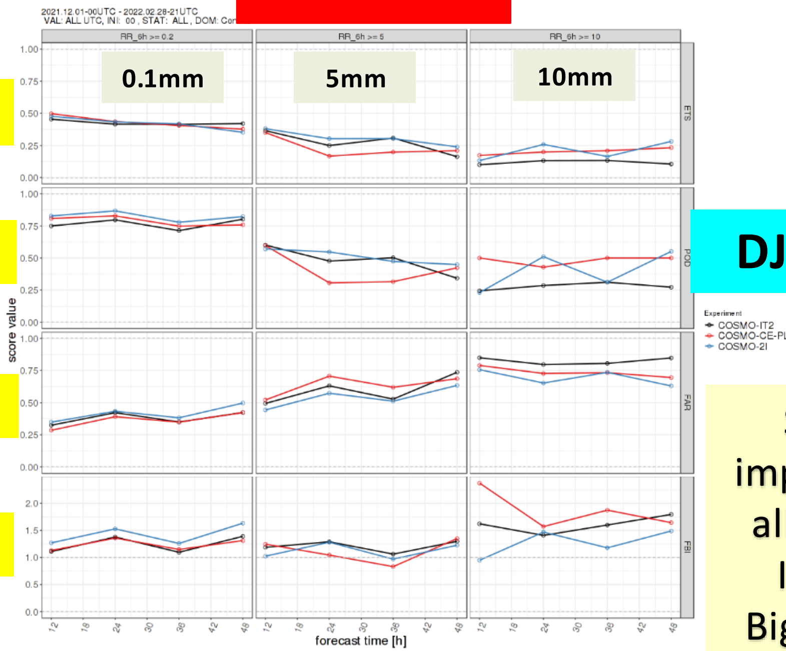
ICON-LAMs overall performs better than COSMO with reduced or similar RMSE compared to COSMO models for T and FF in all seasons and lead times.
For RH, less clear improvement, but ICONs has reduced RMSE values at lower troposphere and similar values with COSMO at other levels.

ICON

COSMO

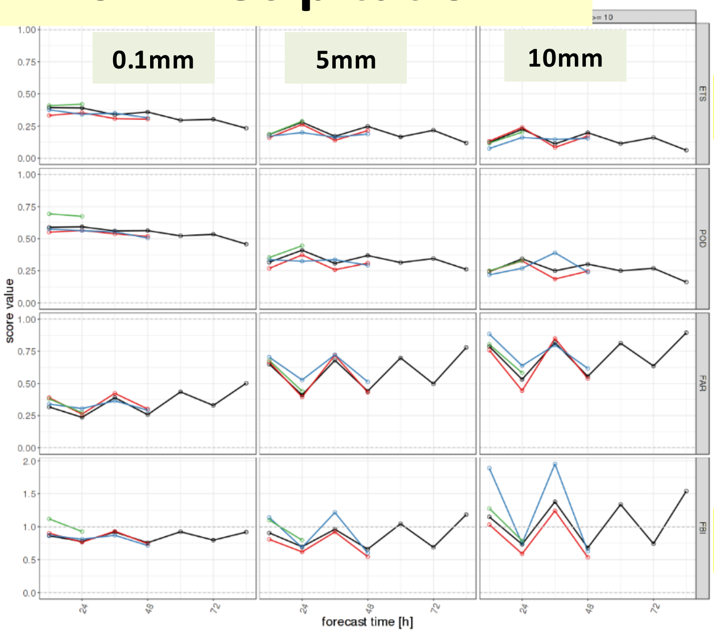
ComA2

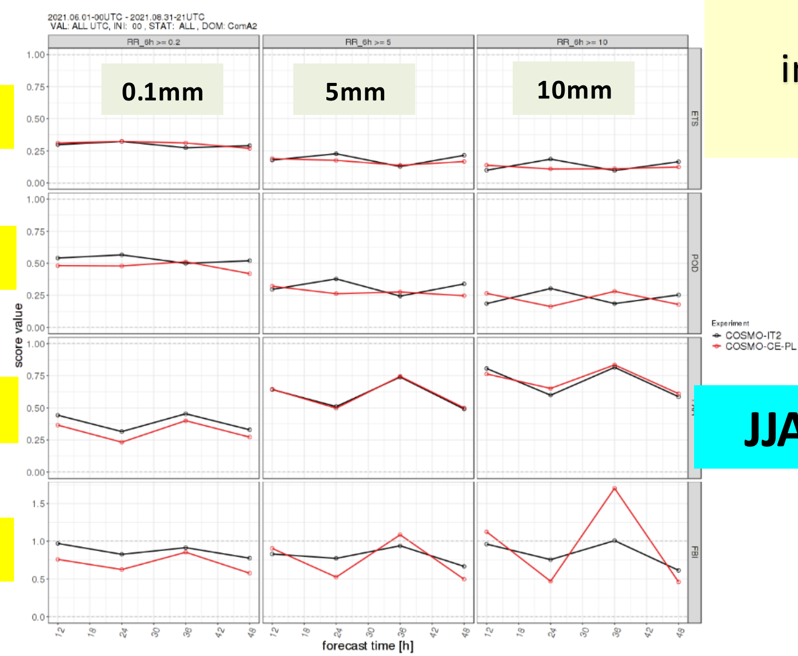




DJF2022

6h Precipitation

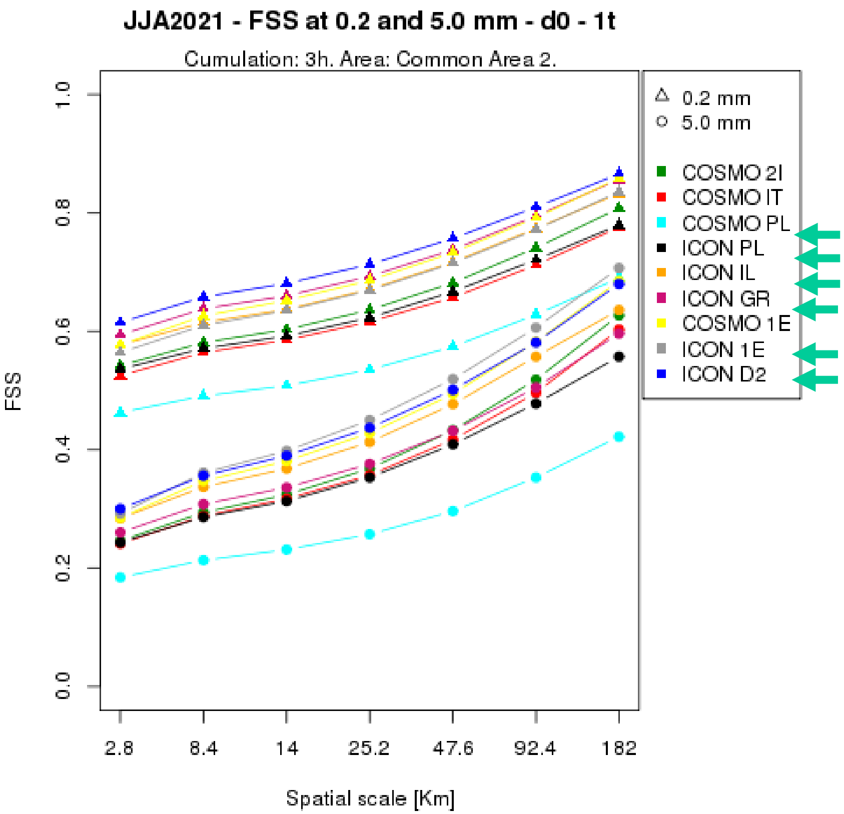




Significant improvement in all scores with ICON-LAMs
Bigger positive changes in summer period

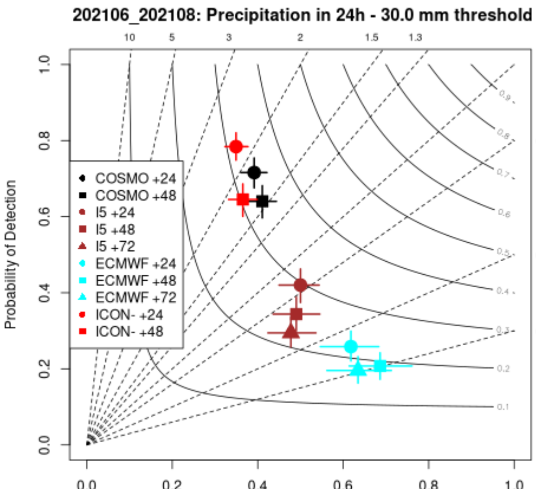
JJA2021

FSS – D0 - 1T

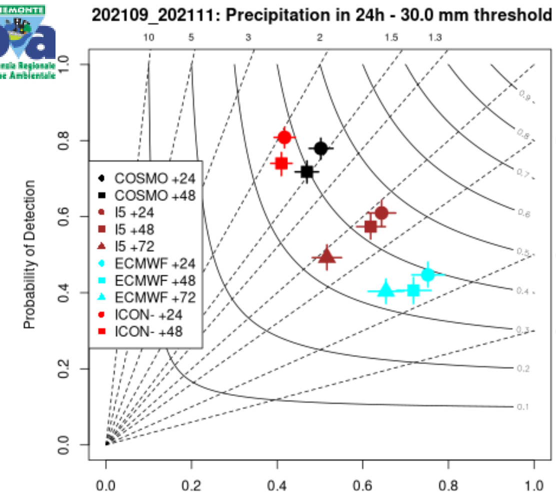


- 0.2 mm: All the ICON models (except for ICON PL) have very good performances when compared to the COSMO ones (apart from COSMO 1E that performs similarly to the ICON ones).
- 5.0 mm: same behaviour as 0.2 mm/3h.

Performance diagram Maximum Values 30mm/24h

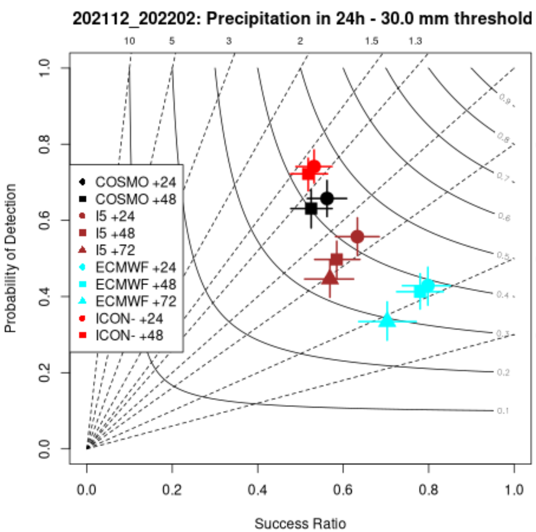


JJA 2021

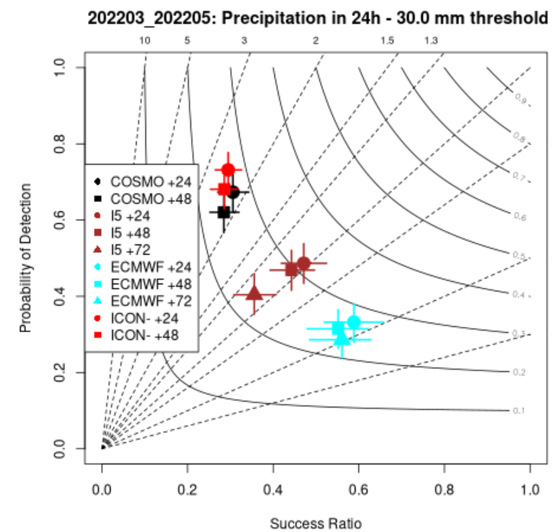


SON 2021

For the issuance of civil protection alerts, information on precipitation maxima is of fundamental importance:
the COSMO and ICON models at 2Km resolution, even if at the expense of many false alarms, are able to provide a valid support compared to what ECMWF does, in particular at the highest thresholds (30mm)



DJF 2022



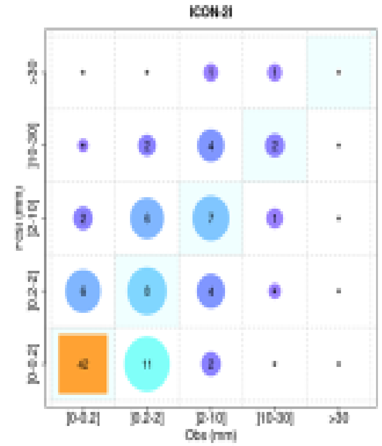
MAM 2022

ICON= ICON-IT run from CNMCA
COSMO= COSMO-2I run from Arpae
I5= COSMO-5M run from Arpae
ECMWF= IFS High resolution run from ECMWF

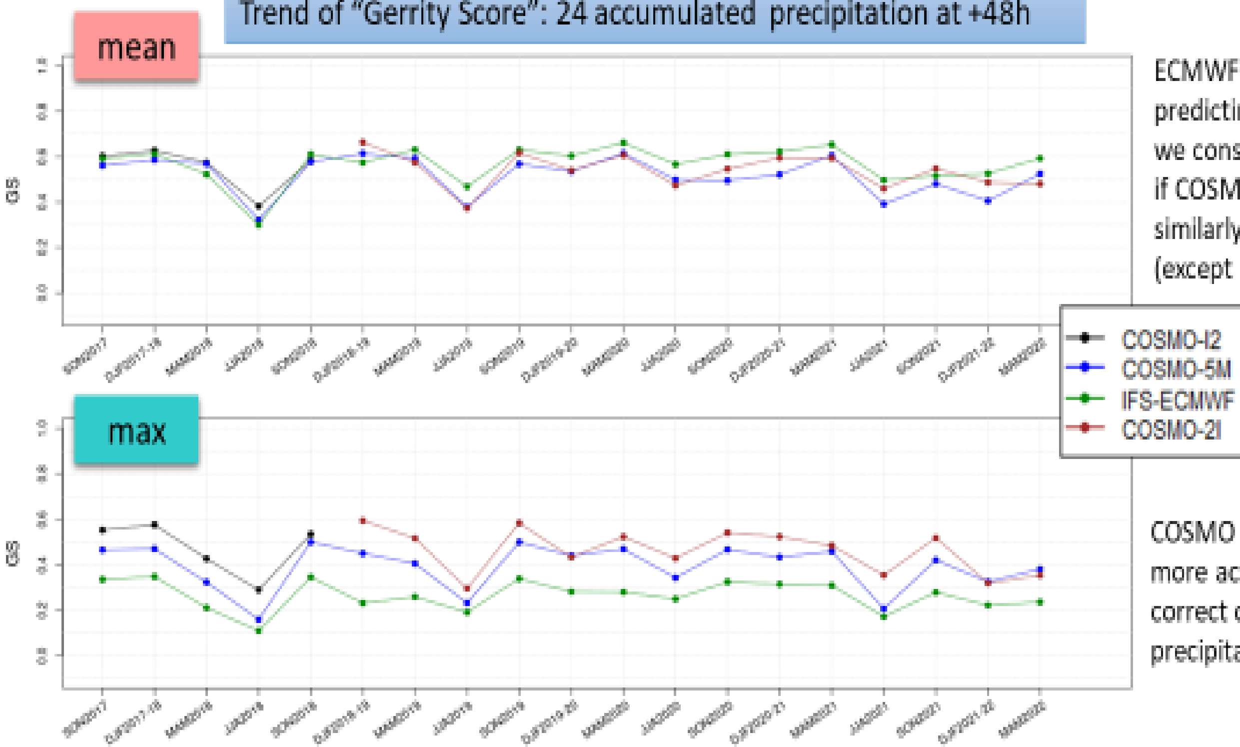
Visual verification with "bubble plots" can be useful to evaluate the behavior of model over a single area

For an objective summary, the use of **Gerrity-Score** allows to evaluate the ability of the model to correctly separate the various classes/category.

Multi-category verification shows that high resolution models (COSMO-2I or ICON-IT/2I) are able to reproduce the precipitation spectrum within the alert areas, distinguishing well especially the precipitation maxima.



Trend of "Gerrity Score": 24 accumulated precipitation at +48h



Range: -1 to 1, 0 indicates no skill. Perfect score:1

ECMWF seems more accurate in predicting the correct category if we consider the mean value, even if COSMO-2I performed very similarly in the last year (except MAM)

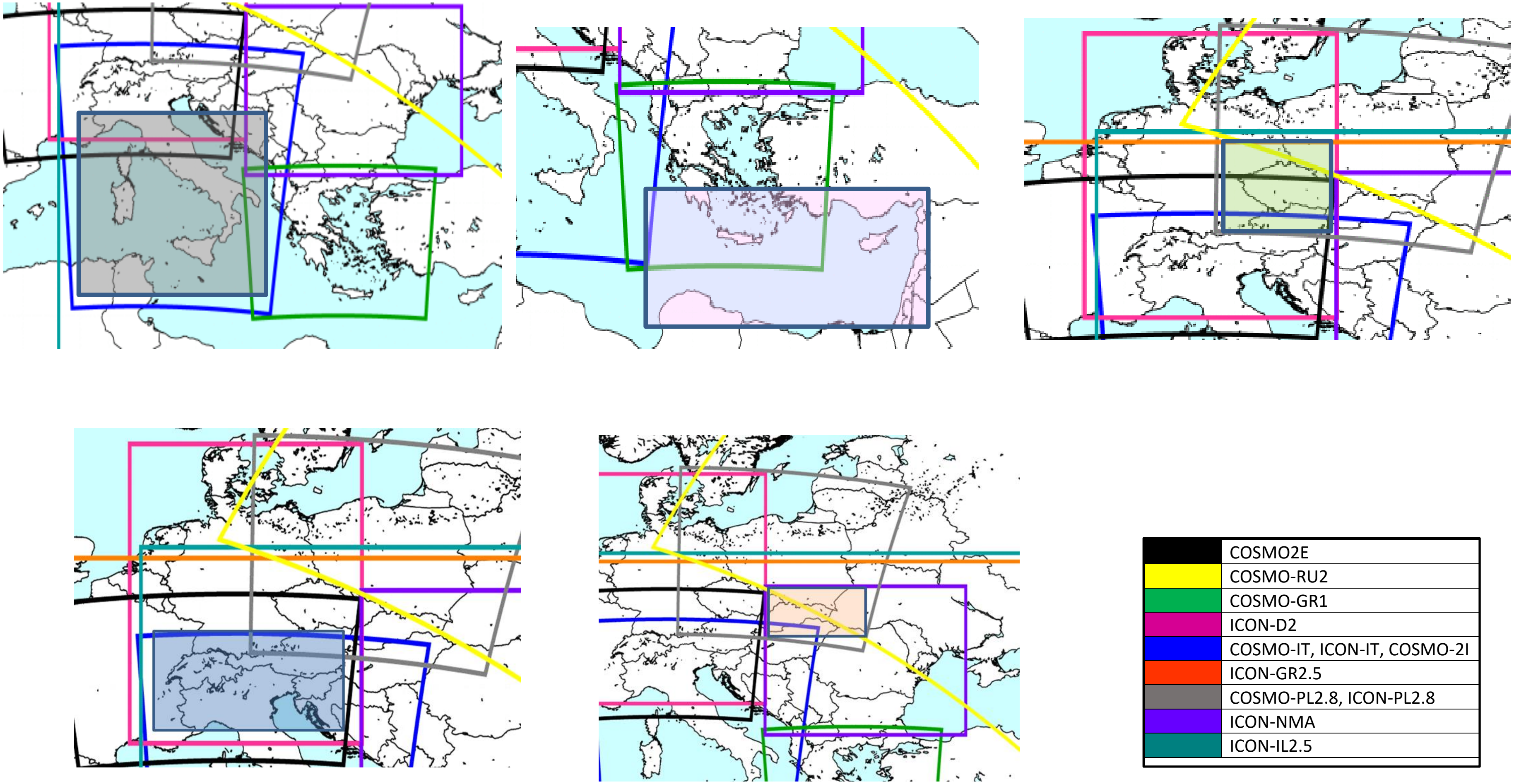
COSMO models (in particular2I) are more accurate in representing the correct category for the maximum precipitation

CLASSES FOR MEAN PRECIPITATION	
mm/24h	**mm/3h**
0-0.2	0-0.2
0.2-5	0.2-1
5-20	1-5
20-45	5-10
>45	>10

CLASSES FOR MAXIMUM PRECIPITATION	
mm/24h	**mm/3h**
0.2 -5	0-0.2
5-25	0.2-2
25-50	2-10
50-75	10-30
75-100	>30
100-150	
>150	

Overall…..

Improvement in performance in most cases/parameters analyzed with ICON-LAMs.
Seems there are components that further model development is needed, as long-term biases are still present
The deviation among model performance is greater in ICON-LAMs than in COSMO models, revealing the need for further model tuning especially in high resolution scales



- MEC/Rfdbk system allows for more flexibility to areas analyzed with no additional effort
- Preparation of a newsletter based on 2021-2022 activity and contributions to the Final PPC2I report from verification analyses over various domains

WG5

Advances in Rfdbk and Feedback File Verification at DWD

Felix Fundel

Old (FFV)
- One script for each observation system (~8)
- One script for each forecast system (x2)
- One script producing scores by date and one for aggregation (x2)
- One extra script for station wise verification (+16)
- Based on the separate R package Rfdbk

New (FFV2)
- Modular structure
- Functions for each task that work with all forecast and observation systems
- Rfdbk package is integrated in FFV2 package so no longer needed

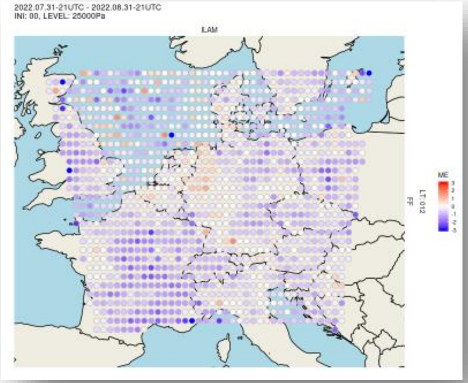
- Easier to maintain
- New features can be made available to all verification tasks by modifying functions or writing new functions
- All verification jobs are technically on the same level

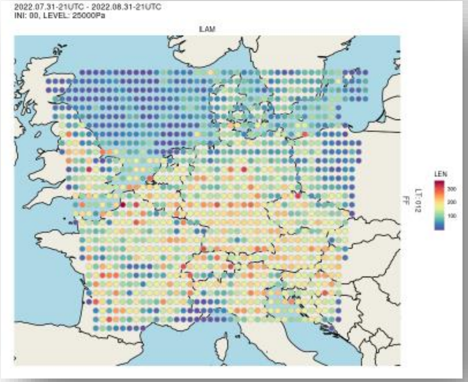
Felix Fundel

New features

Verification of non local observation systems

- Moving observation systems do not allow for a station based verification.
- Score for one location would be supported by one observation only.
- FFV2 offers option to aggregate scores on a user defined lat-lon grid.
- Namelist entry e.g.: rasterLatLon "0.5"





Figures:
ICON-D2 AIREP 250hPa wind-speed bias and number of observations on a 0.5° lat-lon grid

Felix Fundel

New features

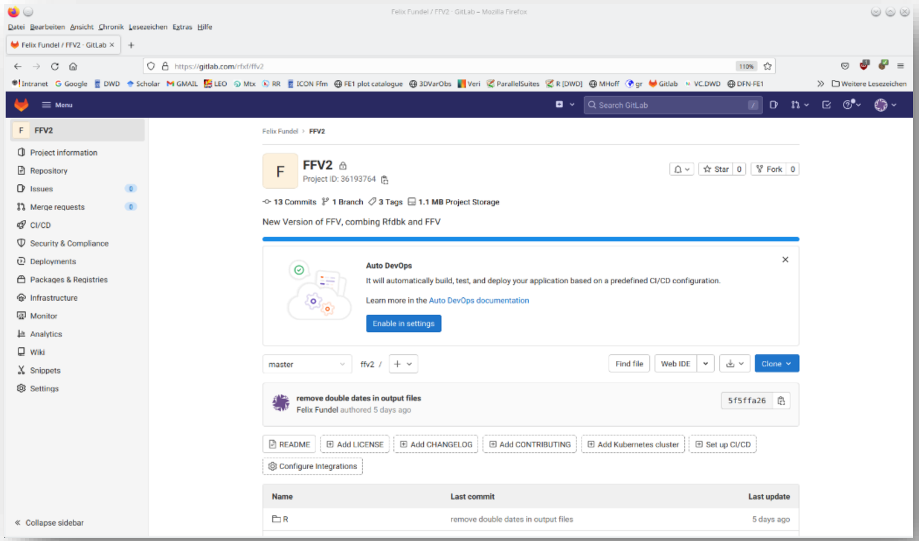
Conditional Verification based on external data

- Conditional verification required data to make the decision to be contained in the feedback file (e.g. T2M score based on TCC threshold).
- FFV2 allows to read external data on model grid.
- This data can than be used to make conditions.
- So far it covers data in NetCDF on native ICON grid.

Felix Fundel

Package

Package
https://gitlab.com/rfxf/ffv2



Install
git clone git@gitlab.com:rfxf/FFV2.git
R CMD INSTALL FFV2

Run (example)
Rscript ../Rlib/FFV2/demo/starter_scores_by_date.R namelist.nl SYNOP DET 6
Rscript ../Rlib/FFV2/demo/starter_aggregate.R namelist.nl SYNOP DET 6

NWP TS Update *(Sept. 2022)*

- "COSMO and ICON numerical weather prediction test suite"
- *2021 used: 3 932 198.65 alloc: 5000000.00* (**78%**)
- COSMO **v5.08** vs **v5.06** and **ICON v2.6.1** Report online in the usual format; **editing to Tech Report format**
- **Update of Rfdbk system and scripts on-going**
- Migration to Atos **on-going**
- Used resources for this year until now: **0%**

AWARE: Appraisal of "Challenging WeAther" FoREcasts

DWD: C. Marsigli, M. Hoff, G.Pante, **MCH:** D. Cattani, **HNMS:** F. Gofa, D. Boucouvala, **IMGW-PIB:** A. Mazur, J. Linkowska, G. Duniec, **RHM**: A. Bundel, A.Muraviev, E.Tatarinovich, **ARPAE:** M.S. Tesini

- Prolongation to complete Tasks and provide the related deliverables until Dec 2021.
- Problems with RHM participation in COSMO caused additional delay in some deliverables

Deliverable Reports at:

http://www.cosmo-model.org/content/tasks/priorityProjects/aware/default.htm

- Task 1.2. Approaches to introduce observation uncertainty

Delayed due to limited human resources. The overview is under preparation.

- Task 2.3. Extreme Value Theory (EVT) approach, Fitting precipitation object characteristics to different distributions

Task finished, report is ready

- Task 3.3. CRA and FSS analysis on intense precipitation

Task finished. Report is under revision

- Task 4.4. Representing and communicating HIW forecast for decision making

Cancelled within PPAWARE

Final Project Report is to be submitted by the end of 2022

ideas for PP-AWARE continuation

I. Stressing of observations role in HIW

✓ **new obs types use in the evaluation of forecasted phenomena (severe convection, fog).**

Obs Types:

Remote sensing. Use of satellite products (e.g. cloud optical thickness, brightness temp, LWR, SWR) to evaluate characteristics of convection, NWC-SAF products for fog verification

Crowd-sourced data: third party and citizen met stations, smart phones, web & social media etc. usefulness for NWP predictions and evaluation - **Included in New PP idea presented by IMGW-PIB**

✓ **observation uncertainty and impact on scores**

II. Verification scheme for convection permitting ensemble forecasts

✓ **object-based approaches:** methodology and criteria for reduction/summarizing of object information, metrics for performance evaluation, visualisation

Long term activity of DWD though SINFONY project

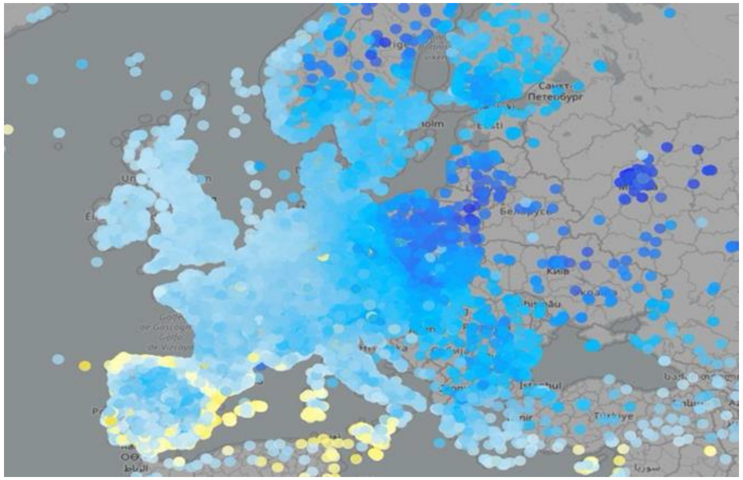
✓ **build of a robust common verification framework for sensitivity tests**

Lack of participation does not allow in the present time for an organized PP on verification schemes, still a permanent activity within WG5. Special focus on EPS applications

WG5

CONSORTIUM FOR SMALL SCALE MODELING
COSMO

Priority Task Idea: EPOCS (Evaluate Personal Weather Station and Opportunistic Sensor Data CrowdSourcing)



IMGW-PIB: Joanna Linkowska, Jan Szturc, Anna Jurczyk, katarzyna Ośródka, Marcin Grzelczyk, Radosław Doździoł
CIMA: Massimo Milelli, Umberto Pellegrini
CNMCA: Francesco Sudati

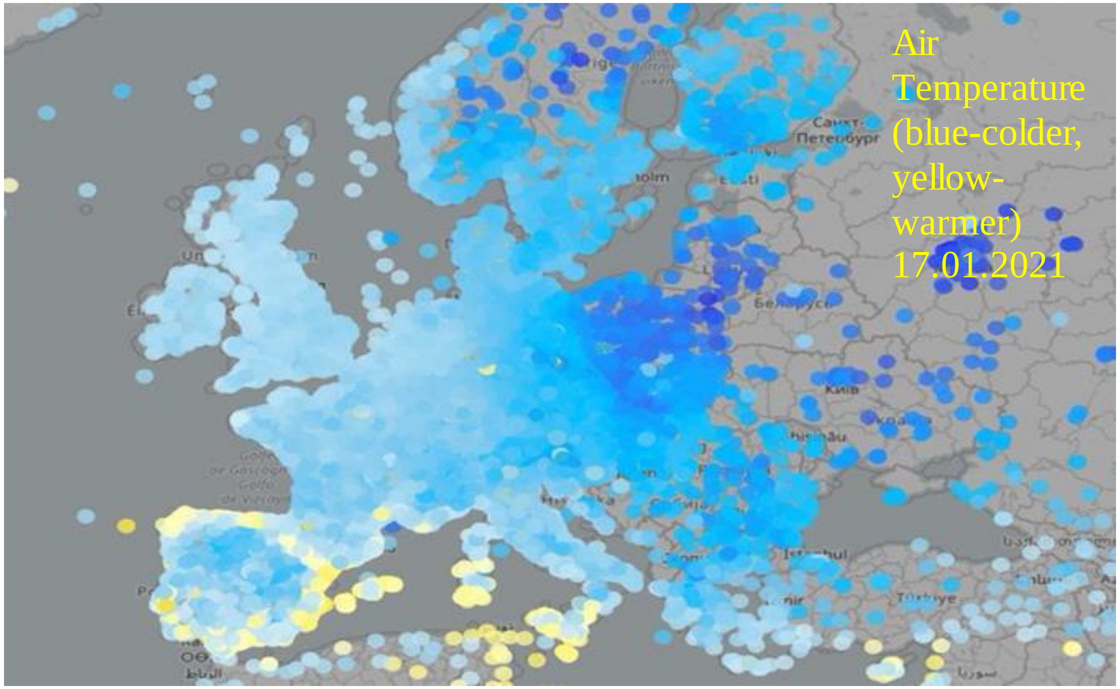
METEO IMGW-PIB meteo.imgw.pl

cima RESEARCH FOUNDATION

PROTEZIONE CIVILE
Presidenza del Consiglio dei Ministri
Dipartimento della Protezione Civile

CNMCA

Motivation



Networks of personal weather stations (PWS)

https://wow.metoffice.gov.uk/

- weather measuring instruments that you can install at your own home or business
- dense network of observations possess a potential to capture high-resolution meteorological information

- PWS sensors are maintained and operated by owners
- prerequisite for ensure data credibility and sustainability
- development and testing quality control (QC) methods and software
- QC assessment of a test set of data, poor quality data removal

Potential PT-EPOCS Content

1. **Survey on PWS data availability within different networks**
   - comprehensive survey of available data platforms at the European and Global level
   - create storage for PWS opertaed by IMGW-PIB employees
   - testing integrity and correctness of stored data, external projects (CENAGIS)

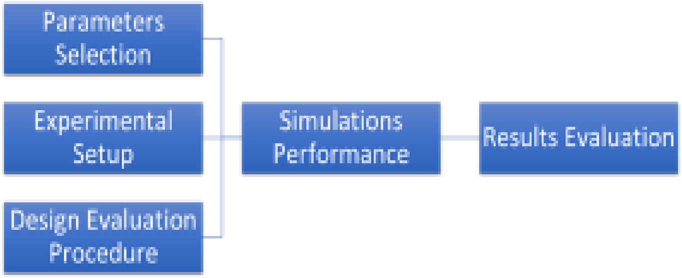
2. **Data quality control (QC) of PWS**
   - survey on QC algorithms and processing software (e.g. TITAN from Norway Met Services, IMGW-PIB's software, COST-OPENSENSE developments, etc.)
   - development/tuning/testing of RainGaugeQC and TITANLIB algorithms
   - PWS QC assesment : Netatmo, Meteonetwork, Centro Meteo Lombardo, Meteotracker

3. **QC of rainfall estimates (RainGRS+)**
   - processing different rainfall data sources (private rain gauges, commercial microwave links, sewer/water service stations, etc.) combine them with other standard data (telemetry, radar, satellite) into new a **enhanced rainfall estimates** (RainGRS+)
   - surwey QC independent data and spatial/object based verification methods

4. **Local variability of precipitation based on the testing PWS stations**
   - **potential of using PWS to monitor extreme events**
   - QC of PWSs precipitation depending on different meteorological conditions

Collaboration with CLM community WG EVAL

Project COPAT2 provides users with "optimal" model configurations for Europe for the new model versions COSMO-CLM 6.0 and ICON-CLM while the comparison and evaluation of model results against observations is central to the project. At the moment in the project there is a lot of discussion on the metrics to consider for the evaluation, and on the consideration of different uncertainty sources.



ETH zürich

The Coordinated Parameter Testing 2 (COPAT2) initiative of the CLM-Community: towards a recommended configuration of COSMO-CLM and ICON-CLM new model versions

Emmanuele Russo[1], Christian Steger[2], Beate Geyer[3], Ronny Petrik[3], Klaus Keuler[4], Burkhardt Rockel[3], Klaus Goergen[5], Patrick Ludwig[6], Hendrik Feldmann[6], Mauro Sulis[7], Bijan Fallah[8], HeimoTruhetz[9], Ha Thi Minh Ho-Hagemann[3], Jan-Peter Schulz[2], and Praveen Pothapakula[6]

1 ETH Zurich, Institute for Atmospheric and Climate Science, Zurich, Switzerland
2 Deutscher Wetterdienst (DWD), Offenbach, Germany
3 Helmholtz-Zentrum Hereon, Institute of Coastal Systems-Analysis and Modeling, Geesthacht, Germany
4 Brandenburg University of Technology, Cottbus, Germany
5 Research Centre Jülich (FZJ), Institute of Bio- and Geosciences (Agrosphere, IBG-3), Jülich, Germany
6 Karlsruhe Institute of Technology (KIT), Institute of Meteorology and Climate Research, Karlsruhe, Germany
7 Luxembourg Institute of Science and Technology, Environmental Research and Innovation Department, Esch-sur-Alzette, Luxembourg
8 Potsdam Institute for Climate Impact Research (PIK), Potsdam, Germany
9 University of Graz, Wegener Center for Climate and Global Change, Graz, Austria

1

PP CARMENs

Cosmo Application of Rfdbk/MEC on ENS

Amalia IRIZA-BURCĂ (NMA)

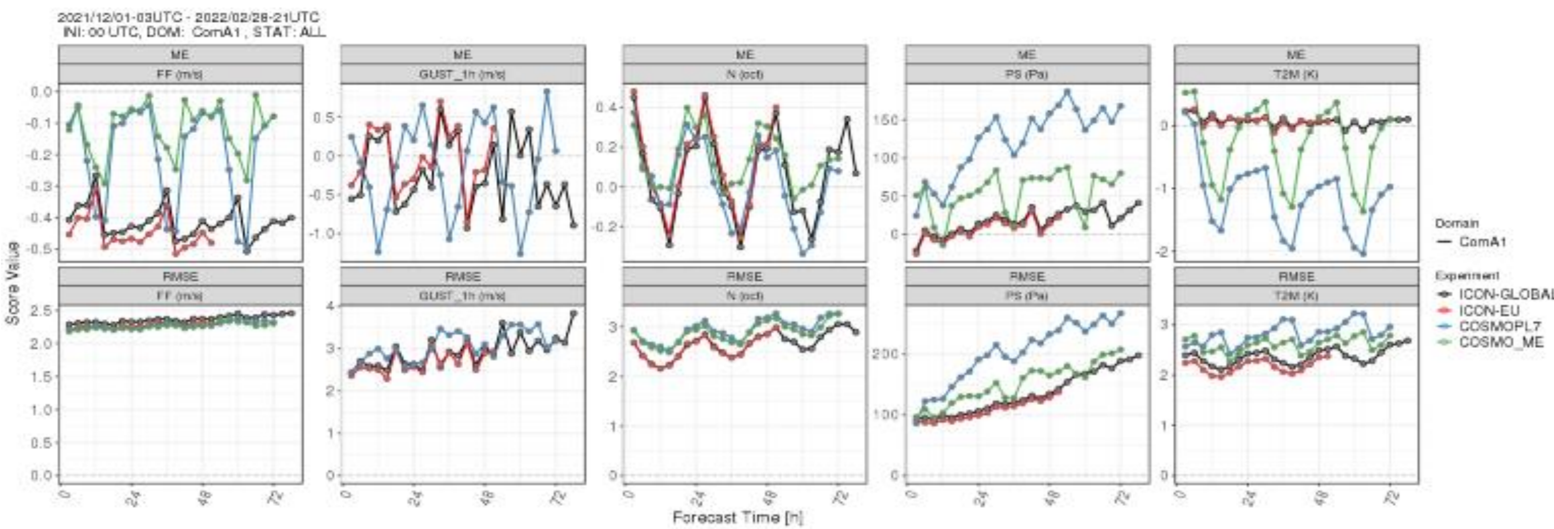
Goal

- 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**

- centralized transfer and visualization of CP statistics on the COSMO wed server (*following NWP Test suite example*)



Current Status

MEC-Rfdbk system implemented and running operationally in most countries of the consortium

products obtained for CP activities:

- Categorical scores for Gust, RR_6h and N;
- Scores for continuous parameters;
- Scores for upper air parameters;
- Comparison between two models showing the trend in various scores;
- Domain average and station based verification;
- Common Area and national domain stratification.

Remaining open issues performed regularly through WG6 SPRT Common Plot activity

Documentation and templates for the use of the MEC-Rfdbk system available (deterministic features).

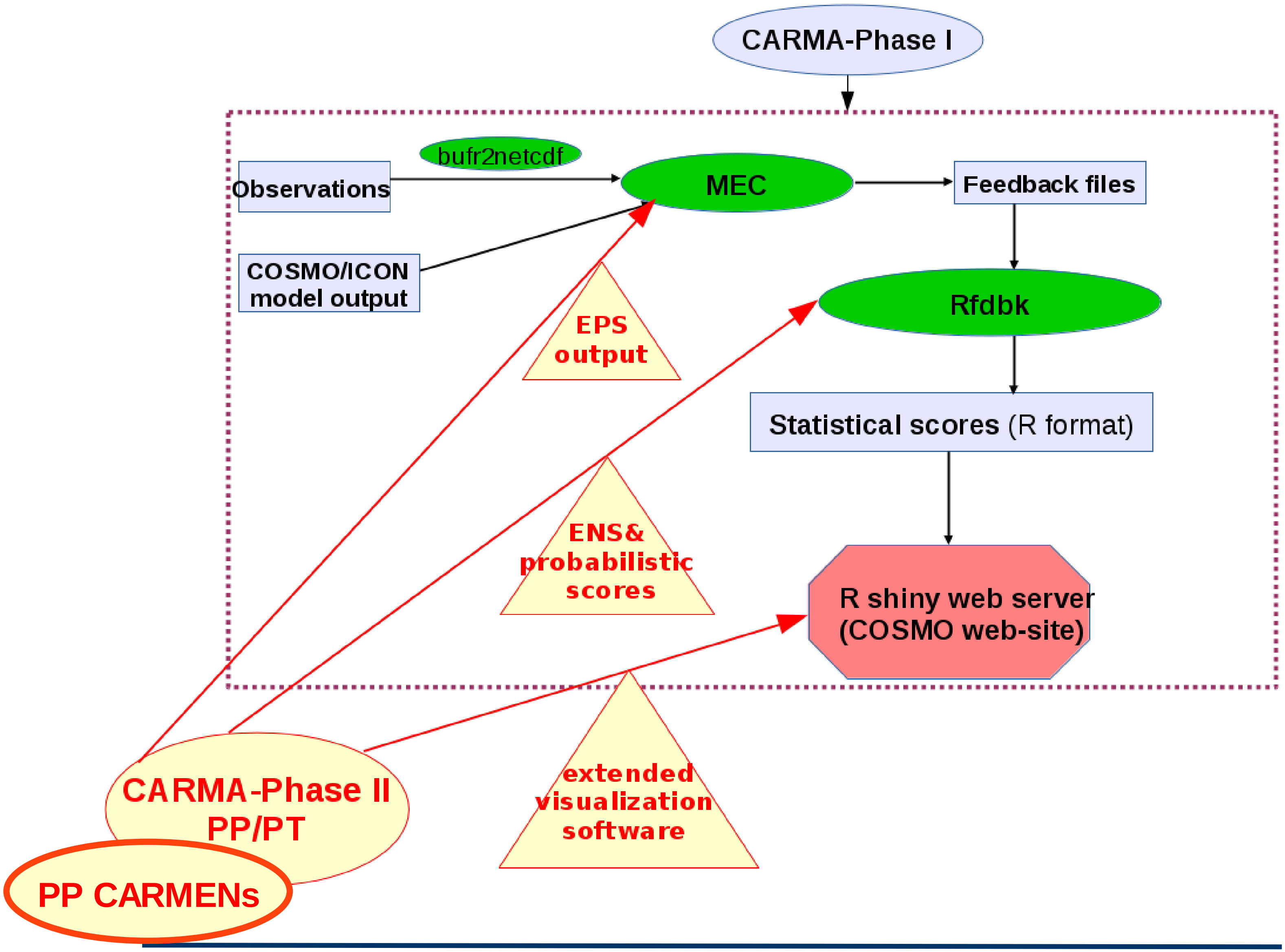
Goal

➔**extend the implementation and usage of the MEC-Rfdbk system to the evaluation of EPS model outputs**

➔available statistical results for selected time periods of ensemble COSMO and ICON-LAM based systems over national domains to be produced and published on the COSMO Verification web page

➔the possibility of an extension of CP activities to EPS (selectively over common areas) will be assessed

PP CARMENs


CARMA-Phase I
bufr2netcdf
Observations
MEC
Feedback files
COSMO/ICON
model output
Rfdbk
EPS
output
Statistical scores (R format)
ENS&
probabilistic
scores
R shiny web server
(COSMO web-site)
CARMA-Phase II
PP/PT
extended
visualization
software
PP CARMENs

WG5 activities contributors

F. Gofa, A. Iriza-Burca, N. Vela, F. Fundel, J. Linkowska, P. Khain, F. Batignani, F. Sudati, D. Boucouvala, A. Pauling, M.S.Tesini, E. Oberto, and more….

Ευχαριστώ πολύ!