WG5 Verification and Case studies

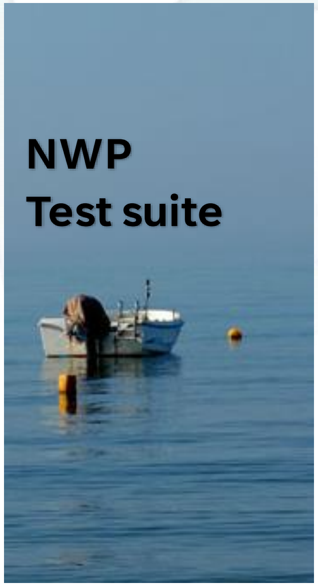
Overview of activities

Flora Gofa

WG5 Guidelines

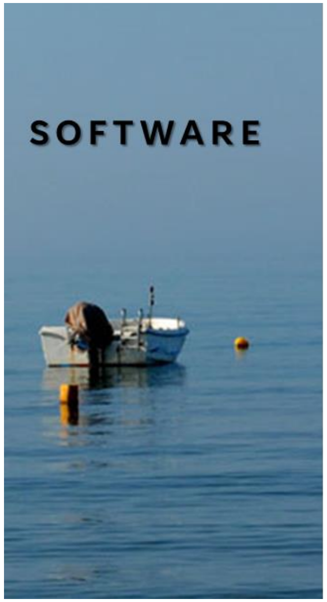
https://www.cosmo-model.org/content/consortium/reports/WG5_Guidelines_2021.pdf

- 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



Deutscher Wetterdienst
Wetter und Klima aus einer Hand

User defined subdomains



- Verification can be performed on user defined model sub-domains
- Subdomains can be defined by polygons or station lists
- Old: overlapping sub-domains were not allowed due to unclear observation handling
- New: overlap is allowed

Application example:

- Calculate and show scores for common area 1 and 2 in one run

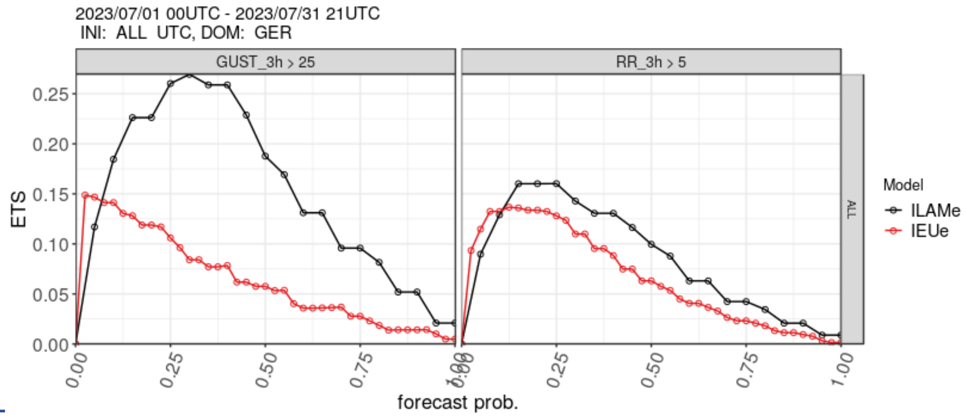
Advances in Feedback File based Verification

Felix Fundel

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Categorical Scores in Ensemble Verification

- Motivation: finding the best probability threshold for decision making based on ensemble forecasts
- Simple but effective verification solution is to give a categorical score for all possible probabilities
- I.e. the ensemble forecast for an event is treated as hit if the probability to exceed the event exceeds a threshold
- This feature is implemented as additional plot type in the EPS verification shiny application
- For more extreme events scores like the SEDI are included
- Now you only have to decide which score to use…



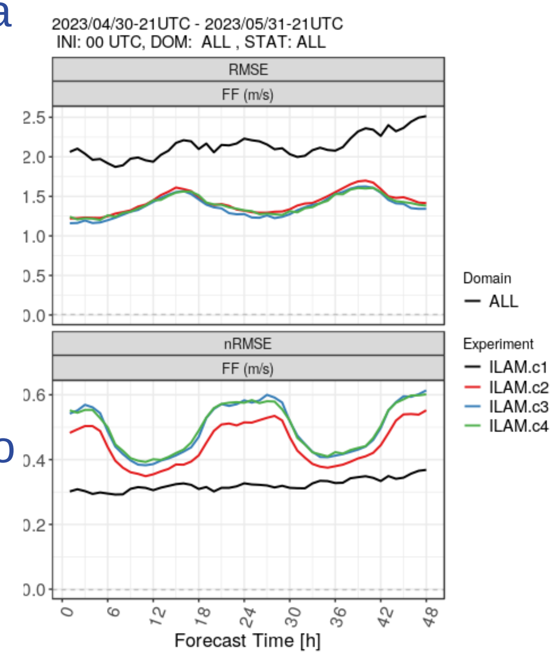
Advances in Feedback File based Verification

Felix Fundel

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Conditional verification I

- Conditional verification capabilities of FFV2 are restricted to data contained in feedback files (e.g. verifying T2M as function of TCC)
- Sometimes additional external input is needed for a conditional verification
- Example of scores depending on AOD (given as external fields) was presented last year, external field had to be provided for each time step.
- Now also a single external field (e.g. extpar data) can be used to define conditions (e.g. scores for classes of roughness length)
- External data and category thresholds must be specified by namelist
- Grid structure of model and external data must be the same!
- Comparison of models on different grids is not advisable.
- Supporting only icosahedral grids, lat/lon support could be implemented on request.
- For better comparison between conditions a RMSE normalized by observation mean (nRMSE) was introduced.



C1 Z0<1.1e-6
C2 1.1e-6<Z0<0.154
C3 0.154<Z0<0.47
C4 0.47<Z0

VAST - COSMO software for fuzzy verification
New released planned: September 2024
N.Vela, Arpa Piemonte

A new version of the VAST software to be delivered with capabilities that are already used for CP.

- **MAIN IMPROVEMENTS:**

- No need of LIBSIM to preprocess the data.
- GRIB files accepted as input data both for the observation and forecast.
- Possibility to calculate the desired score over long periods (months or seasons) for FSS, POD, FAR, ETS and FBI.
- Documentation

- **POSSIBLE IMPROVEMENTS (not guaranteed):**

- Produce verification differentiated by lead time

Methods - HIW

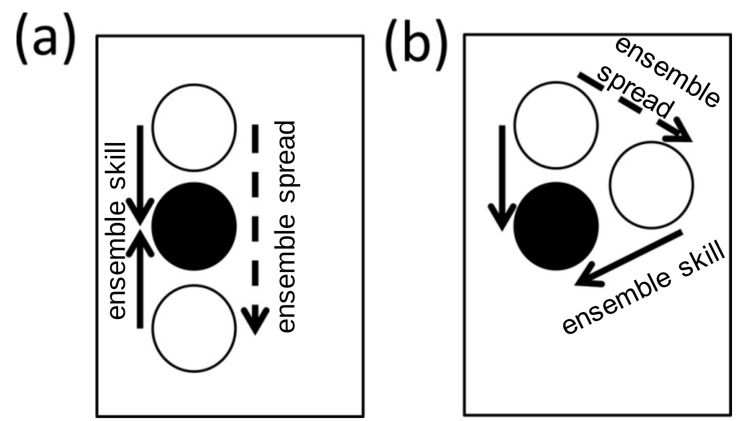
Michael Hoff

- → Methods have been developed to evaluate differences in ensemble forecasts with focus on spatial structures and location errors → spatial ensemble spread-skill at different scales
- → Neighborhood-based spread-skill methods implemented at DWD
  - → Dispersive Fractions Skill Score
  - → Agreement Scales
  - → Central Member

Work performed for SINFONY project

Dispersive (and error) FSS

- It is a method for characterizing an ensemble forecast spatially based on Fractions Skill score
- build mean and std. dev. of FSS values over all member-member pairs at different spatial scales and different forecast times to evaluate the spatial differences of the forecast
- is a „global“ measure over the entire domain, no local information
- Gives information about: identification of useful scales, temporal evolution of spatial ensemble spread, model spin-up times, error growth and saturation of forecast differences
  - **Spatial ensemble spread** is characterized by calculation of FSS for all independent member-member pairs ($N_p(N) = N \times (N-1)/2$, e.g.: $N_p(N=20) = 190$ independent combinations)
  - **Spatial ensemble skill** is characterized by the calculation of FSS for all member-radar pairs (**eFSS**)
  - **Ensemble spread-skill** is dependent on spatial scale and threshold (**dFSS**)



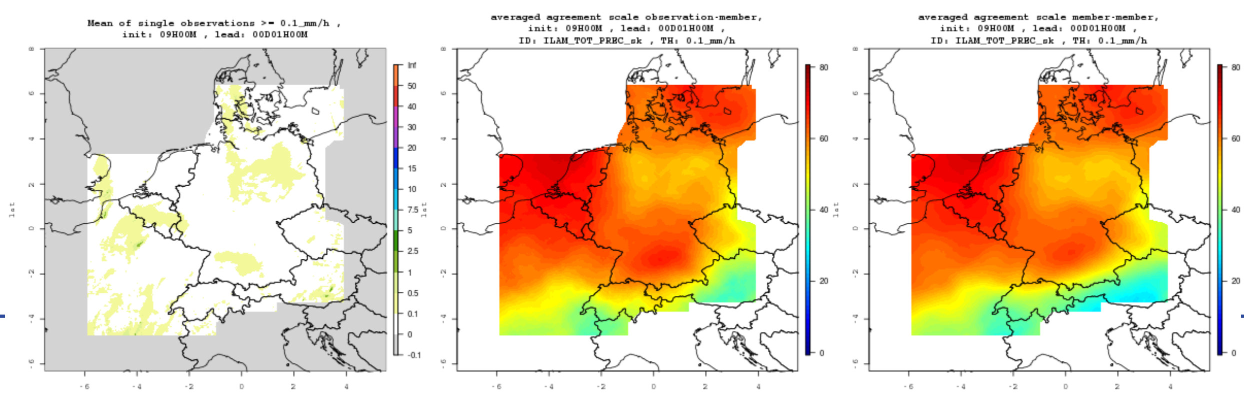
Ensemble spread twice as large as skill

Ensemble spread equal to skill

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Agreement Scale

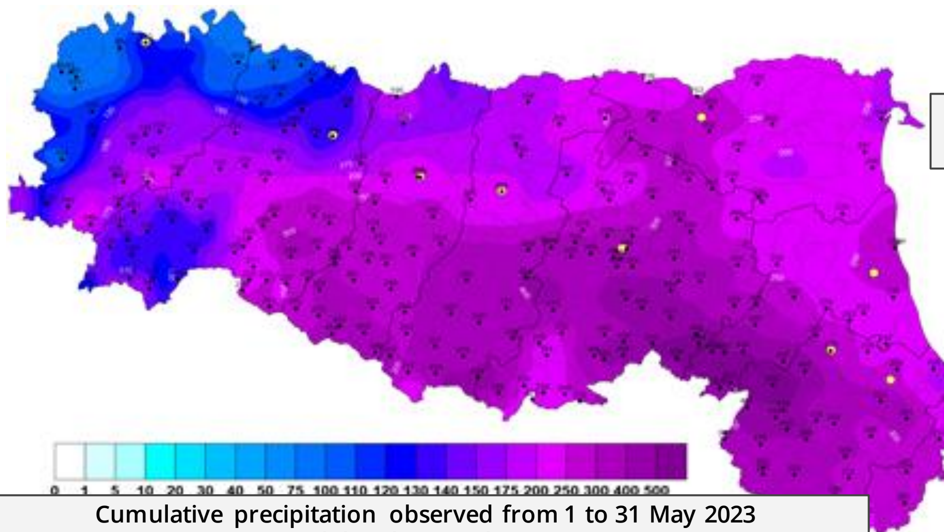
- ➔ Is based on neighborhood approach for differences between 2D forecast fields (e.g. mean precipitation over all grid points in the surrounding)
- ➔ Evaluation and characterization of local variations of predictability of ensembles, e.g. precipitation at convective scale (different locations in domain host different physical regimes)
- ➔ Spatial scales over which ensembles „agree“ (agreement scales, S^A) are calculated at each grid point resulting in a map of spatial agreement between forecasts
- ➔ $S^{A(\overline{mm})}$ is the agreement scale averaged over all independent member-member pairs
- ➔ $S^{A(\overline{mo})}$ is the agreement scale averaged over all independent member-radar pairs
- ➔ A comparison of both gives a location-dependent spatial spread-skill relationship



SINFONY

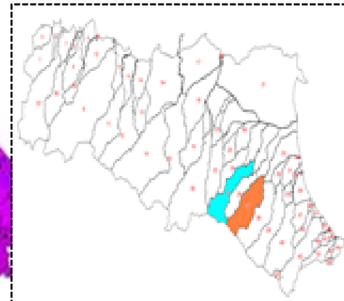
Models perfomance during the flood events of May 2023 in Emilia-Romagna region

Maria Stefania Tesini



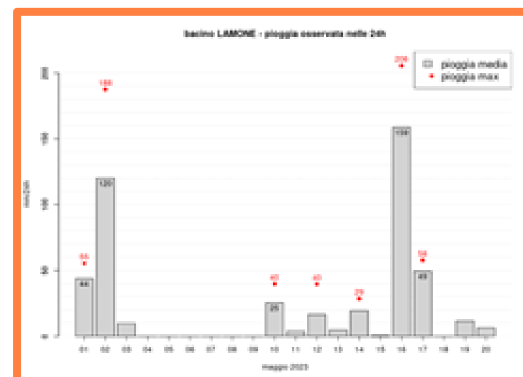
Cumulative precipitation observed from 1 to 31 May 2023
(validated ERG5 dataset)

Hydrological basins
(400-500 Km^2)



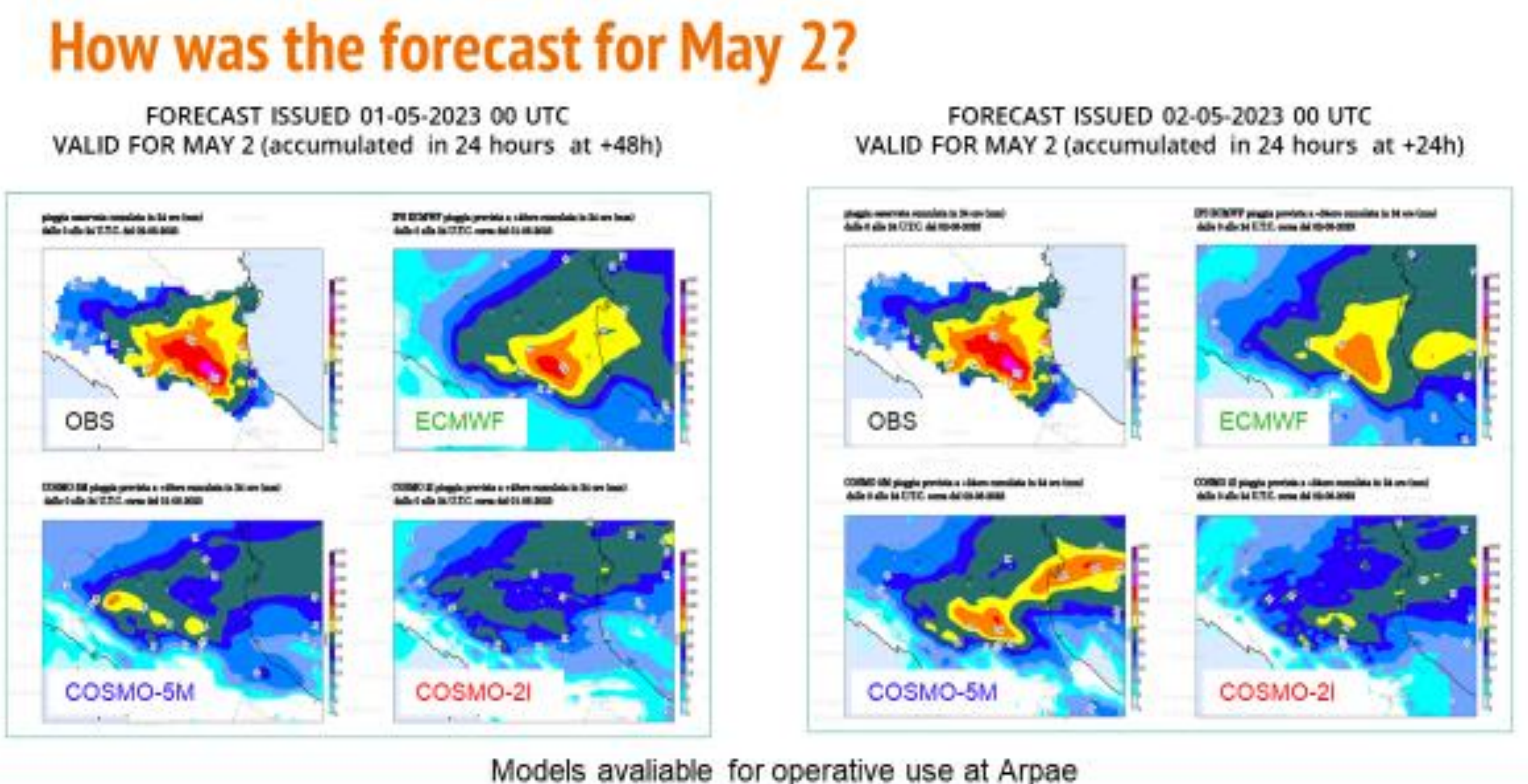
Maximum amounts in the period 1-17 May:

- Trebbio (Lamone basin) 609 mm
- Le Taverne (Santerno basin) 563 mm
- Historical records for most rain gauges in the central-eastern sector with values over 300-400 mm (some with 100 or more years of data)
- The rain that has fallen in these areas over the entire period represents about a quarter of the annual cumulative climate value, while in each of the two main events (1-3 May and 16-17 May) it clearly exceeded the monthly cumulative climate value.



Models perfomance during the flood events of May 2023 in Emilia-Romagna region

Maria Stefania Tesini



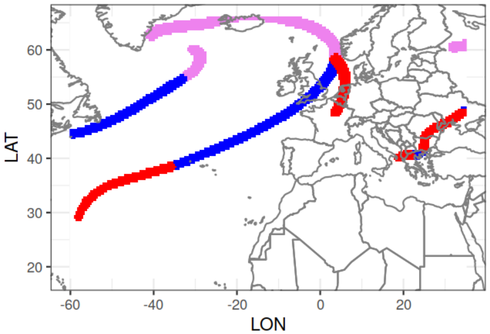
ML Weather Fronts Prediction and Application in Verification

Felix Fundel

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Motivation

- Making use of ML in verification
- Making use of weather front analysis from DWD forecasters (human)
- Train ML model to identify fronts in forecasts
- Verify in the vicinity of fronts
- Verify properties of front predictions as objects



Started March 2021, total of >1800 cases and growing

Literature:

- ML superior to classical methods (Niebler et al., WCD, 2021)
- Methodology inspired by Biard and Kunkel, ASCMO, 2019

Advantages:

- Front analysis and prediction based on (ICON) model only (consistent in time)
- AI should learn robust front properties and (hopefully) discard personal forecaster preferences
- AI can predict fronts for every model date, not just 00 & 12 UTC and some selected forecast times
- AI might help quantify and understand model errors
- AI might support forecasters for drawing fronts or fast identification of areas of relevant weather

ML Weather Fronts Prediction and Application in Verification

Felix Fundel

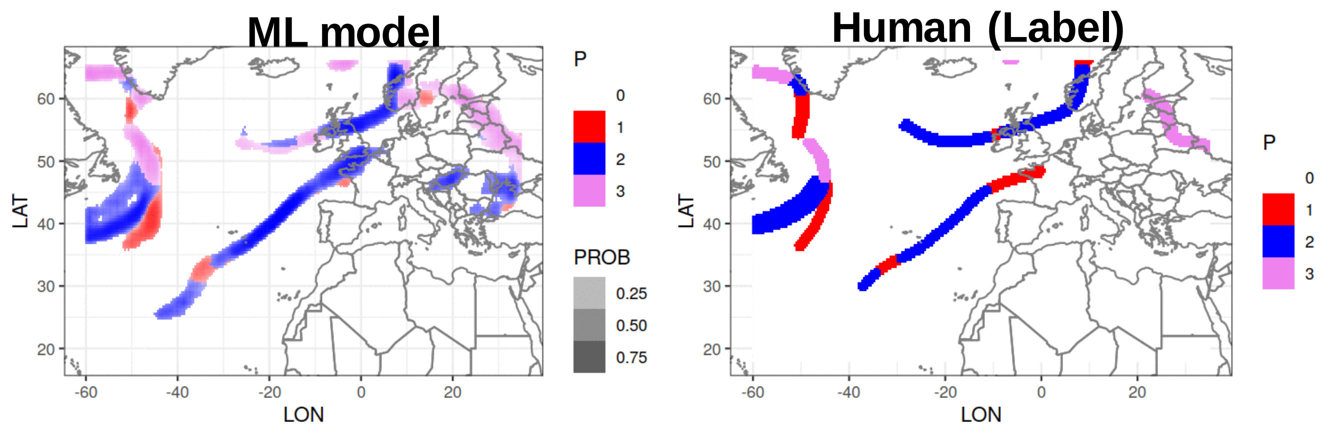
Deutscher Wetterdienst
Wetter und Klima aus einer Hand

ML Model

- Written with the Tensorflow library with Keras functionality
- 2d convolutional Encoder-Decoder NN with concatenation of features from the encoder path (Unet)

Prediction

- Some examples from the validation data set (unseen during training)
- red: warm front, blue: cold front, violet: occlusion
- Transparency ~ class probability
- No front is drawn if “no-front” probability is >60% (random choice)



- Realistic front prediction.
- Discarding lower probabilities will lead to sharper fronts and less artefacts but some true features might be lost.

ML Weather Fronts Prediction and Application in Verification

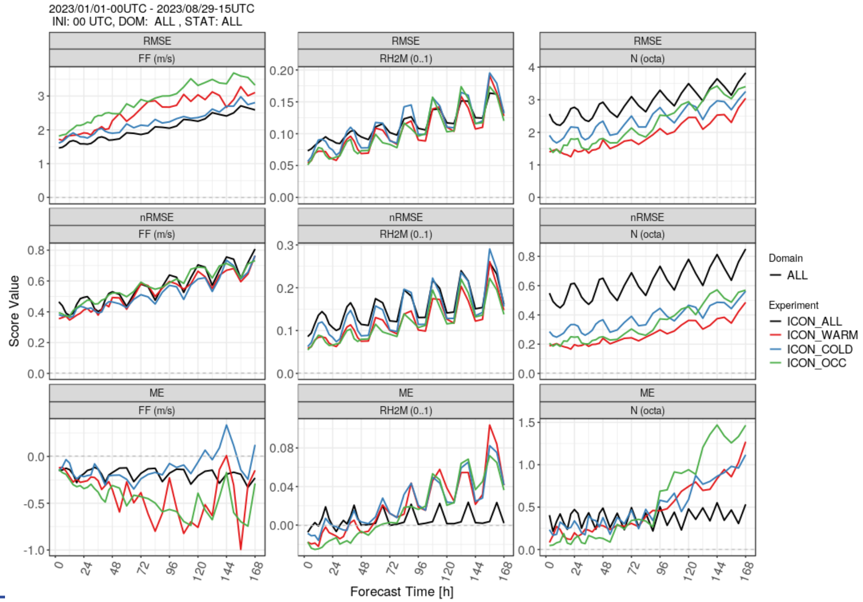
Felix Fundel

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

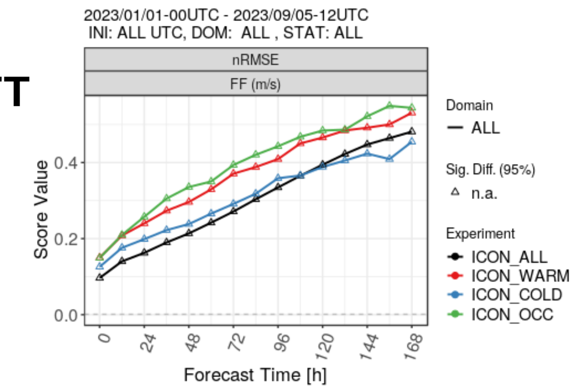
Station Based Verification

- Observation based verification has been upgraded to ingest varying polygons as mask for a conditional verification
- Forecast errors are calculated for observations in the vicinity of cold- and warm fronts and occlusions
- RMSE can be misleading, normalized RMSE is more appropriate
- SYNOP: Wind forecasts equally good in all classes, humidity and cloud cover probably influenced by range of data.
- SCATT: Higher forecast errors in front classes, especially warm fronts and occlusions.

SYNOP



SCATT



PP / PT COLLABORATIONS

PP CARMENs

Cosmo Application of Rfdbk/MEC on ENS

Amalia IRIZA-BURCĂ (NMA)

Stefan Ionu Odărian

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

Task 1. Administrative Tasks and Technical Support - Start 09.2022 - End 09.2024

1.1 Administrative Tasks – on going

- **Meeting May 31, 2023**
  - establishing the details related to the delegation of Amalia Iriza's tasks to the other colleagues in the NMA
  - establishing the new coordinator of the CARMENs priority project
- **Quick update on the status of the activities that have been transferred from Amalia to NMA colleagues for**

sent to WG5 leader for the SMC in 26.06.2023

- **Meeting August 9, 2023 - discussions related to the status of the tasks within the CARMENs priority project**

Start 09.2022 - End 09.2024

1.2 Technical Support– on going – this task should start together with task 3 – on going

Start 09.2022 - End 09.2024

Total Resources Task 1: 0.65 FTEs

Task 2. MEC and Rfdbk system adaptations for EPS systems – delayed 6 months

2.1 Instructions for use and adaptation of MEC software for Feedback Files based on EPS model output (LAM based EPS, IFS ENS) - Problems solving. - delayed

- **discussion with Andreas Pauling and Daniel Leuenberg from MeteoSwiss regarding the production of FF for Swiss ensemble forecast. The swiss colleagues provided us a data set and MEC namelist example for producing FF from ensemble forecast.**

- **setting the MEC on NMA machine for producing the FF based on swiss example – everything worked properly**

- **discussion with Enrico Minguzzi for providing complete data sets (containing all parameters requested by MEC) of COSMO-LEPS forecast. He uploaded a few days of COSMO-LEPS forecast on NMA ftp server.**

- **setting the MEC on NMA machine for producing the FF based on COSMO-LEPS data provided by Enrico. – everything worked properly**

Start 09.2022 - End 05.2023
task reschedule period proposal 10.2023 – 02.2024

2.2 Instructions for adaptation of Rfdbk/FFV2 verification system for production of ENS and probabilistic scores (LAM based EPS, IFS ENS) - Problems solving – delayed

Start 12.2022 - End 05.2023
task reschedule period proposal 10.2023 – 02.2024

2.3 EPS evaluation guidelines as part of Common Plot activity– delayed

Start 11.2022 - End 05.2023
task reschedule period proposal 10.2023 – 02.2024

2.4 Scripts to produce verification scores for EPS forecasts available to project participants through the common WG5 repository - delayed

Start 12.2022 - End 05.2023
task reschedule period proposal 10.2023 – 02.2024

Total Resources Task 2: 1.00 FTEs

Task 3.Semi-automatic use of the Rfdbk for ENS production and probabilistic scores

3.1 Installation and adaptation of MEC-Rfdbk system for EPS over national domains by all participants – delayed
Start 05.2023 - End 12.2023

3.2 Preparation of Seasonal FF and test with Rfdbk – on going
Start 09.2023 - End 09.2024

3.3 Adaptation of Shiny visualisation libraries for EPS properties and scores - delayed
Start 06.2023 - End 09.2023
task reschedule period proposal 10.2023 – 02.2024

3.4 Test of system features on Seasonal EPS forecasts over Common Area (if applicable) – on going
Start 05.2024 - End 09.2024

Total Resources Task 3: ~1.6 FTEs

Open Questions

- **Which COSMO members will be able to provide ensemble forecast Feedback Files and/or Rdata for centralized verification/visualization of scores?**
- **Test the functionality of the system using COSMO-LEPS model output**
- **What SCORES should the EPS verification produce?**

  EPS verification scores mentioned in the project:
  - Ensemble & probabilistic scores such as the rank-histogram, reliability diagram, Brier score with decomposition, ROC curve, economic value, fair CRPS, outlier statistics, spread/skill and deterministic scores based on the ensemble mean.
  - Time series of ensemble scores and station-based ensemble scores.
  - Evaluation of a subset of ensemble members only. Conditional verification capabilities.

Main Approaches	SCORES
Verification of the ensemble distribution	• Rank histograms (Talagrand diagrams) • Continuous rank probability score and its related skill score(a measure for the difference between the cumulative distribution function of the forecasts and observations)
Verification measures for the pdf of a generic probability forecast	• Linear probability • Ignorance score
Verification of the probability of an event.	• Brier score and its decomposition • Brier skill score • Reliability diagrams • ROC curves and area • Rank probability score

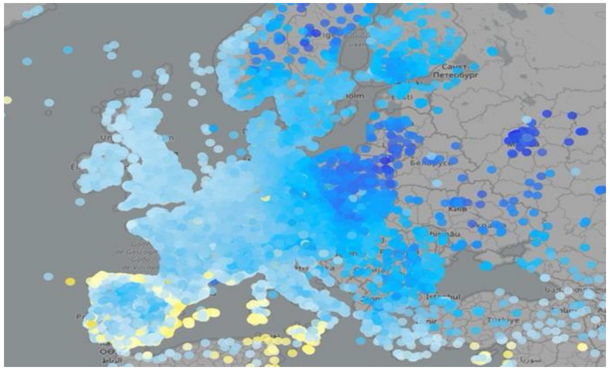
Priority Task: EPOCS (Evaluate Personal Weather Station and Opportunistic Sensor Data CrowdSourcing)

PL: Joanna Linkowska, IMGW-IB

The aim of PT EPOCS is to assess the use of weather data from Personal Weather Stations (PWS) and other Opportunistic Sensors (OS).

The main scientific aims of this PP are:

1. The development and testing of data Quality Control (QC) algorithms.
2. The evaluation of quality and usefulness of this data for potential applications (nowcasting, NWP and model forecast verification)



IMGW-PIB: Joanna Linkowska, Jan Szturc, Anna Jurczyk, Katarzyna Ośródka, Marcin Grzelczyk, +Radosław Droździoł
CIMA: Massimo Milelli, Elena Obert, Umberto Pellegrini
CNMCA: Francesco Sudati

COSMO GENERAL MEETING, Gdańsk, 11-14.09.2023

CONSORTIUM FOR SMALL SCALE MODELING
COSMO

COSMO PT EPOCS Tasks

1. PWS databases survey and exploitation

1.1 Comprehensive survey of available data platforms at the European and Global level.

Participants: Marcin Grzelczyk, Francesco Sudati, Massimo Milelli (March 2023 - June 2023: **completed**)

1.2 Testing the process of collection of a real-time PWS data (from IMGW-PIB employees that are using their own stations) by starting new internal database server.

Participant: Marcin Grzelczyk (March 2023 – October 2023: **ongoing**)

1.3 Testing integrity and correctness of stored data, assess usefulness of external databases/projects (CENAGIS).

Participant: Marcin Grzelczyk (November 2023 – February 2024, **started earlier – ongoing**)

1.4 Analysis of the mobile PWS sensors: testing QC proprieties of a new mobile weather sensors from Meteotracker.

Participants: Francesco Sudati, Massimo Milelli (March 2023 - February 2024 – **ongoing**)

PWS SERVICES

METEO IMGW-PIB meteo.imgw.pl

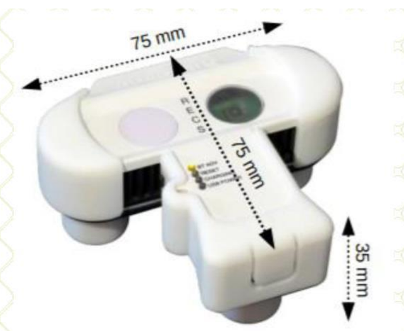
No	Web service	Stations in EU	API	Privacy Policy (Terms of service)	API Data
1	Wunderground (www.wunderground.com)	EU >1000	Yes, professional	Not commercial or science usage; only personal use; commercial need to pay	All data (observation, historical above 5min steps)
2	Aeris weather (www.pwsweather.com/)	EU >1000 (DE, GE, IT, PL ~60, RO, CH)	Yes, professional	Attribution required for public usage,	1,000 accesses/day (rate limit: 100 accesses/minute), Observations, Daily Forecasts (7 days), Hourly Forecast (24 Hours), Sun & Moon, Places, Alerts, Observations/Summary, Observations/Archive , Air Quality
3	Awekas (www.awekas.at)	EU >1000 (DE, GE, IT, PL ~90, RO, CH)	No, only text information without download	Not commercial or science usage; only personal use after pay annual fee (https://www.awekas.at/wp/shop/licenses/awekas-stationsweb-annual-fee/?lang=en)	No API, but have primitive export option
4	Meteomatics (https://www.meteomatics.com/), WMO (https://public.wmo.int/en/programmes/public-weather-services-programme)	EU = 13154 station of WMO	Yes, professional	free basic package for non-commercial use, but this has only forecast data and 24 hours of historical data (no observations).	Wind speed, Maximum temperaturę, Minimum temperaturę, Precipitation amount, Sunrise, Sunset
5	Weathercloud (weathercloud.net)	EU >30.000 (**Germany**: 20,000, **Greece**: 600, **Italy**: 5,000, **Poland**: 2,000, **Romania**: 300, **Switzerland**: 1,200)	Yes, basic	Personal, worldwide, non-assignable and non-exclusive license to use Weathercloud. This license is for the sole purpose of enabling you to use and enjoy Weathercloud in the manner permitted by these terms. This license is revocable at any time. This license does not include: The distribution of Weathercloud's content., Modifying or otherwise making any derivative uses of Weathercloud or any portion thereof., Use of any scraping, data mining, robots or similar data gathering or extraction methods., Downloading, other than page caching, any portion of Weathercloud, except as expressly permitted., Accessing Weathercloud's API with an unauthorized client., Any use of Weathercloud other than for their intended purposes.	3 Devices (basic plan), **10 Minute** update interval, **12 Month** cloud database, Current weather & evolution graphs, Data export to CSV, Custom plots, Daily & monthly reports

Subtask 1.1
Provide comprehensive survey of available data platforms at the European and Global level with the analysis of the API data access.
In relation to research activities in Italy (a doctorate project under way for the specific validation of the ICON model), a special attention will be made on PWS data related to the Meteonetwork and Centro Meteo Lombardo associations.
Especially, the second one is a very dense network that covers the area including and around Lombardy, subject to constant strict control by the members, characterized by the presence outside the inhabited centres, especially in very complex orography terrain.
The other aspect of the survey analysis are the legal limitations for data usage which may be a key issue in data applicationfor the research and operational use.

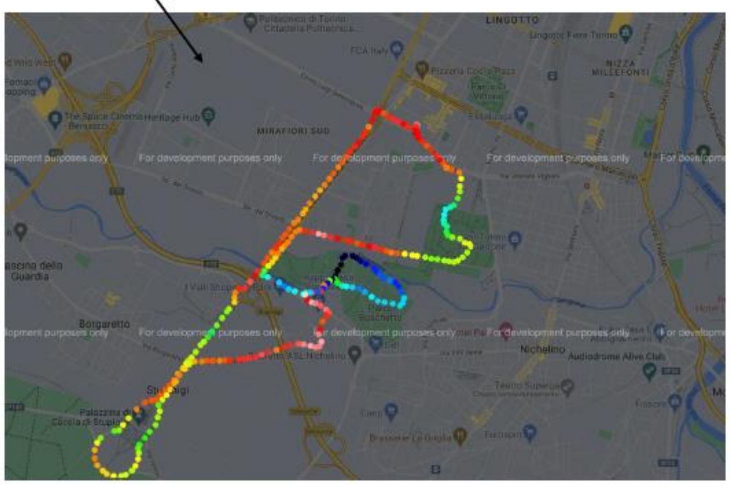
1.4 Analysis of the mobile PWS sensors: testing QC proprieties of a new mobile weather sensors from Meteotracker

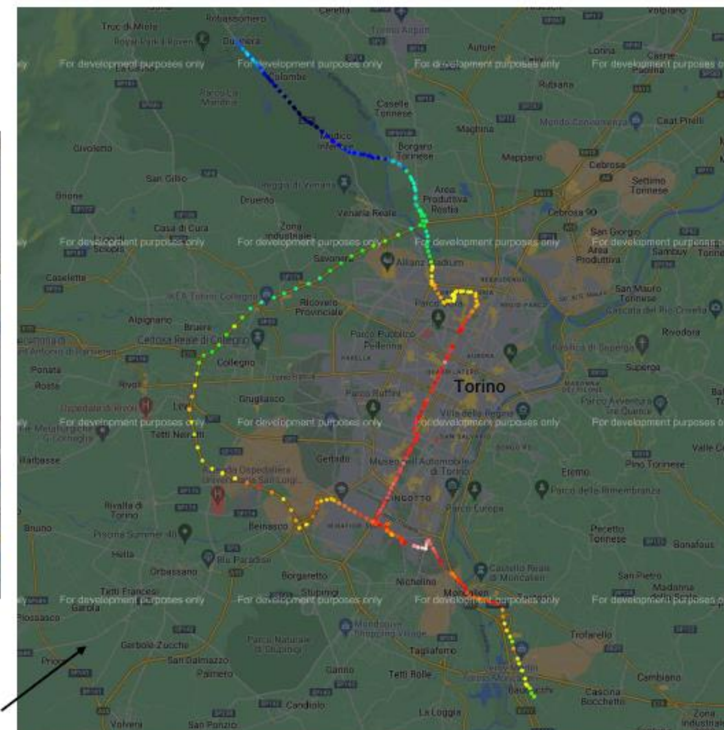
MeteoTracker is a mini weather station specifically designed and patented for measurements taken on the move: ✓ Air temperature, ✓ Relative Humidity, ✓ Pressure.

Derived parameters: Dew Point, Altitude, Vertical Temperature Gradient, Solar Radiation Intensity, Humidex Index (thermal comfort), Vehicle Velocity.



Day time: evidence of different urban texture



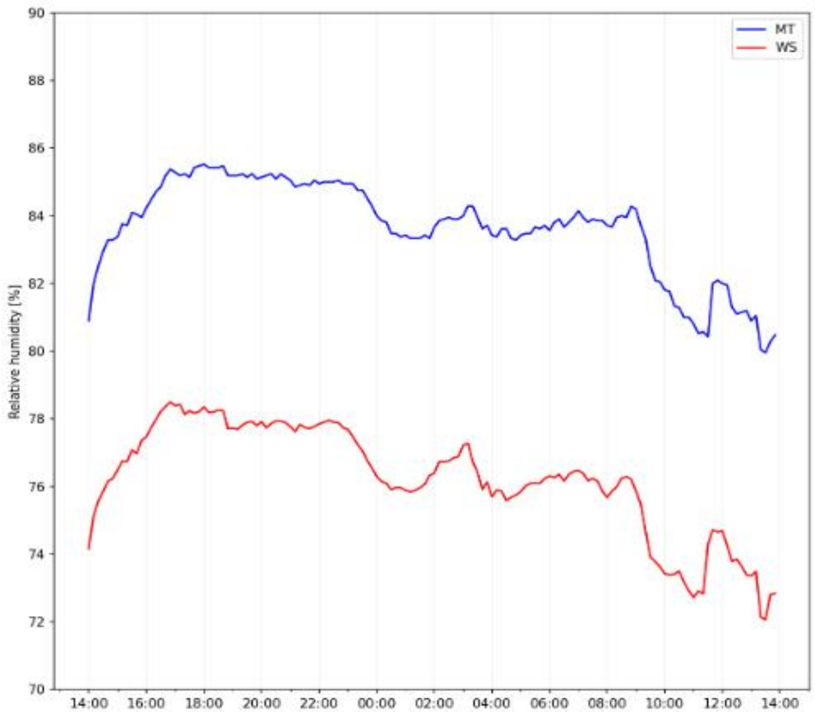


Night time: Urban Heat Island

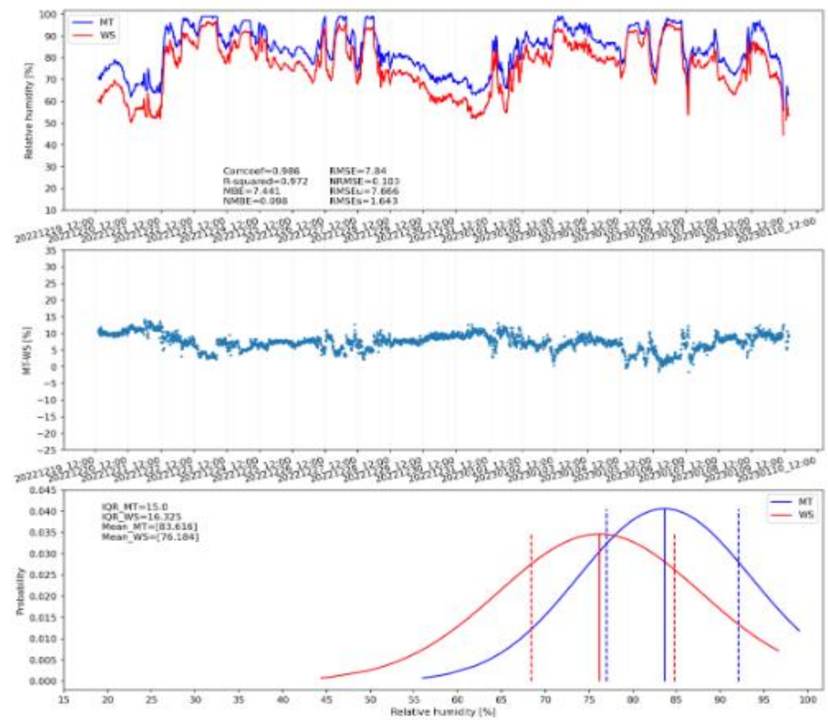
Slide courtesy
of Massimo Milelli

1.4 Analysis of the mobile PWS sensors: testing QC proprieties of a new mobile weather sensors from Meteotracker

Comparison between a MT in a Stevenson box and a WMO sensor (WS) on the roof of CIMA.
Relative Humidity, Winter period (19/12/2022 - 10/01/2023)



Mean daily cycle



Slide courtesy of Massimo Milelli

CONSORTIUM FOR SMALL SCALE MODELING COSMO www.cimafoundation.org

RH accuracy is poorer and MT tends to overestimate RH (~10% difference in winter and ~20% difference in summer).

COSMO PT EPOCS Tasks

CONSORTIUM FOR SMALL SCALE MODELING COSMO

2. QC algorithms for precipitation

2.1 Development and testing automatic QC methods based on the RainGaugeQC algorithms developed at IMGW-PIB.

Participants: Katarzyna Ośródka, Jan Szturc, Anna Jurczyk (March 2023 - February 2024: **ongoing**)

2.2 Testing and application of the open-source software package TITAN for a quality control of ground data.

Participants: Elena Oberto, Umberto Pellegrini (March 2023 - February 2024: **ongoing**)

Purpose: potential of further use of PWS for data assimilation in NWP/Nowcasting models or verification of model forecast.

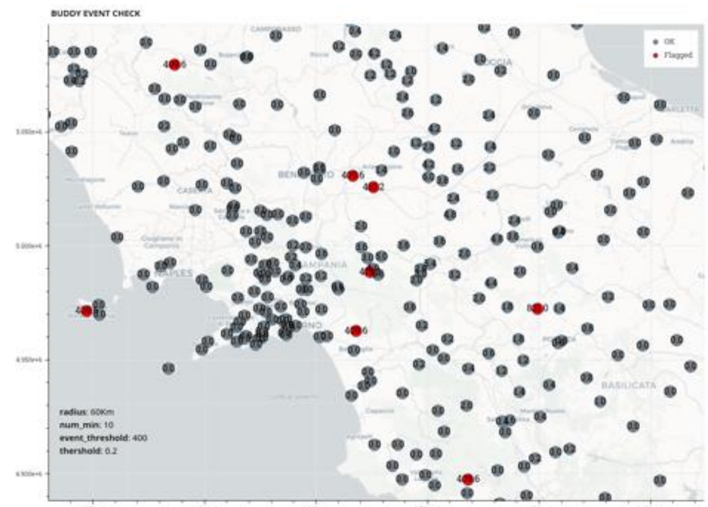
2.2 Testing and application of the open-source software package TITAN for a quality control of ground data

TITANLIB is an open source software developed at the Norwegian meteorological institute. It is a library of automatic quality control routines for weather observations. The main goal of the task is to prepare a "clean" precipitation field for model verification, for instance. Starting point – test on the official Italian network, not on personal weather stations (PWS).

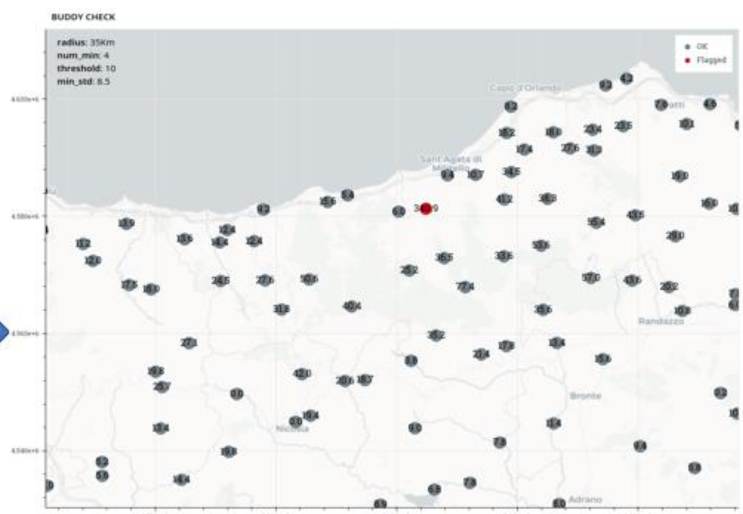
TITANLIB main tests

Buddy event check →

The observations are converted into yes/no values of exceeding a specified threshold (**event_threshold**).
The **threshold** argument in this test is minimum fraction of other observations in the neighbourhood that must agree with the observation being inspected.



Buddy check →



The buddy check compares an observation against its neighbours (i.e. buddies) and flags outliers in a radius specified by the user.
The buddy check flags observations if the (absolute value of the) difference between the observations and the average of the neighbours normalized by the standard deviation in the circle is greater than a predefined **threshold**.

Slide courtesy of Elena Oberto

COSMO PT EPOCS Tasks

3. Analysis of PWS based gridded rainfall products

3.1 Processing different rainfall data sources (private rain gauges, commercial microwave links, sewer/water service stations, etc.)

Participants: Katarzyna Ośródka, Jan Szturc, Anna Jurczyk (March 2023 - February 2024: **ongoing**)

Combine PWS with other standard data (telemetry, radar, satellite) into new enhanced rainfall estimates (RainGRS+)

3.2 Reliability analysis of a gridded RainGRS+ high-resolution estimates of precipitation.

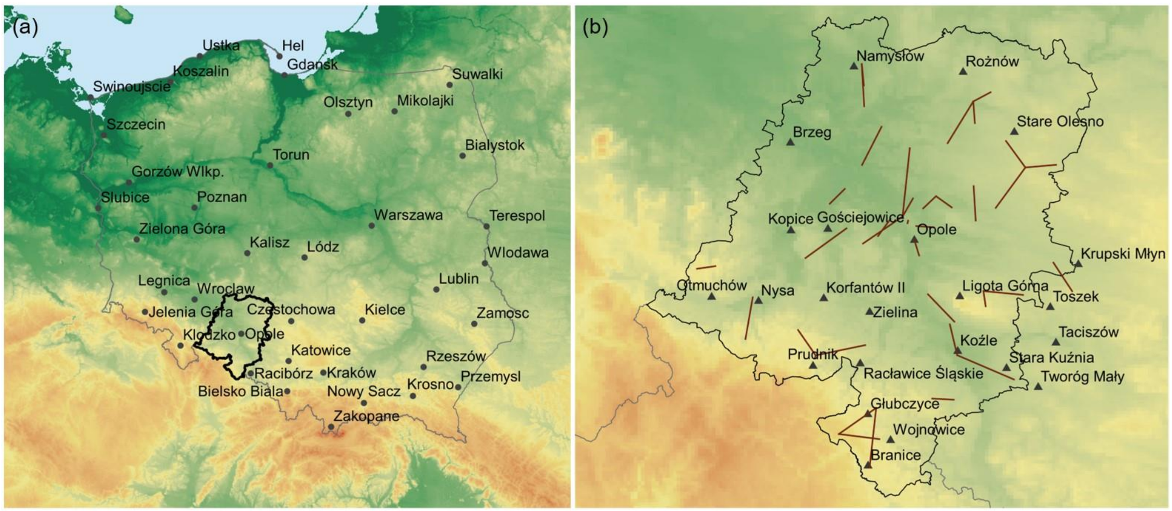
Participant: Joanna Linkowska (September 2023 - February 2024 **ongoing**)

Sample data of RAINGRS+ and RAINGRS (without PWS) fields will be verified against chosen independent precipitation data.

3.1. Multi-source precipitation (RainGRS+) CML-based data

METEO IMGW-PIB
meteo.imgw.pl

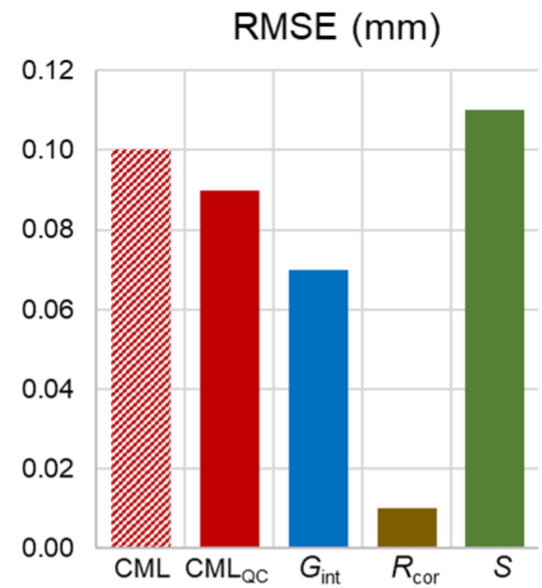
CMLs (commercial microwave links) provide precipitation estimates based on attenuation on the links.

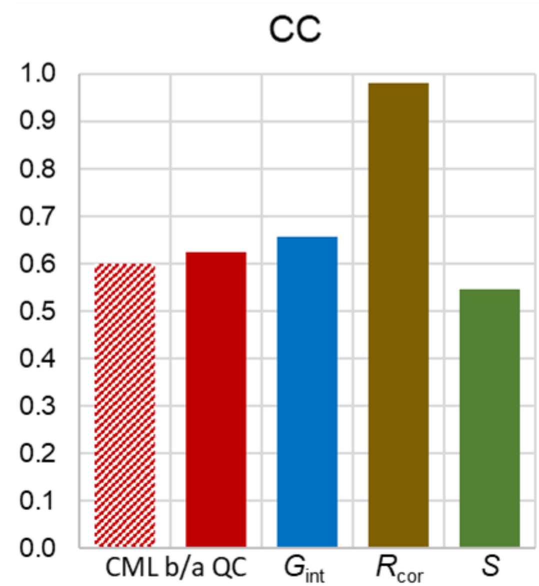


RMSE and correlation coefficient (CC) of CML-based precipitation determined before and after quality control, telemetric rain gauges (G_{int}), corrected radar (R_{cor}), and satellite (S) data.

RainGRS as a reference.

Half-hourly accumulations (July 19 – August 18, 2022)





Slide courtesy of Jan Szturc

COSMO PT EPOCS Tasks

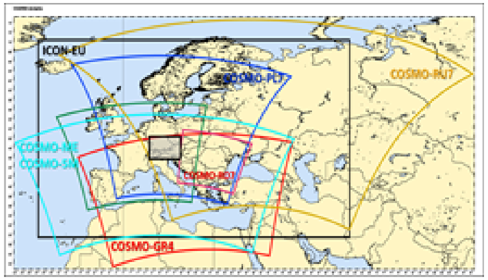
4. Suggestions for the follow up activities (ongoing)

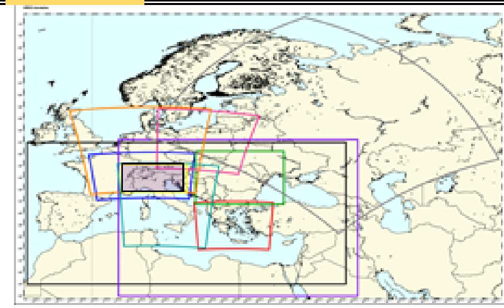
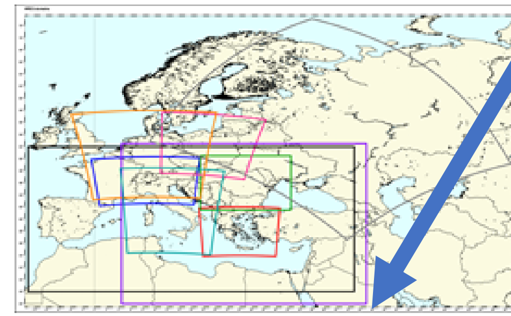
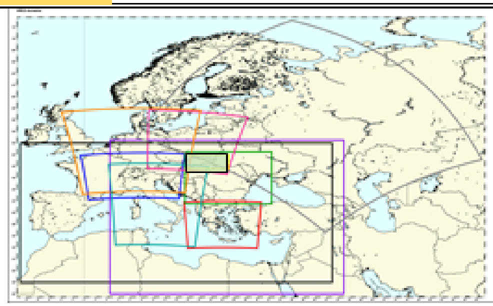
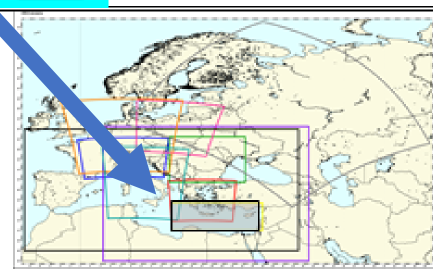
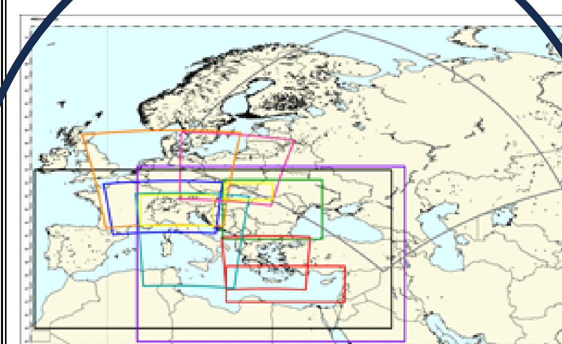
- Planning a collaboration for a longer Priority Project (PP)

- Assessment for application of project results in supporting other COSMO R&D activities.

  - forecast verification,
  - data assimilation of NWP/Nowcasting models,
  - postprocessing (machine learning), etc.

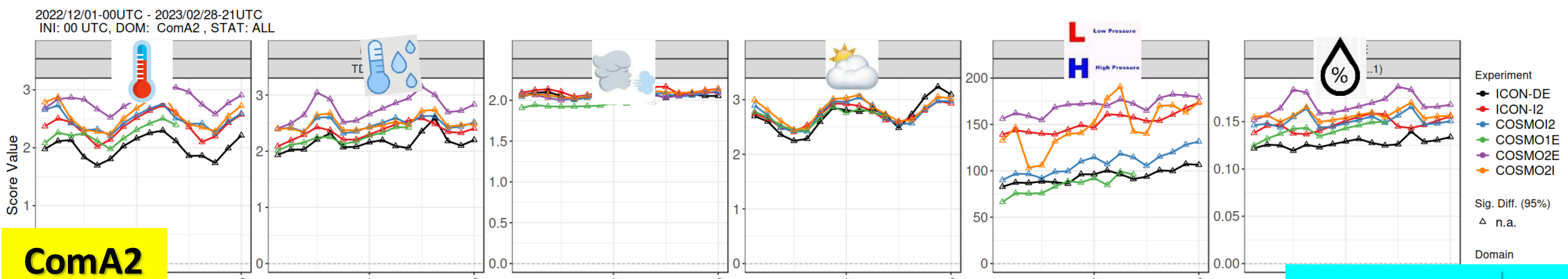
WG5

<table>
<tr><th colspan="4">Common Plots: redefined areas</th></tr>
<tr><td></td><td>COARSE ComA-1</td><td>FINE ComA-2</td><td>MIX NoComA (National Domains)</td></tr>
<tr><td>Specs</td><td>

Forecast run: 00UTC
Forecast Horizon: 48h
Seasons: JJA22, SON22, DJF23, MAM23</td><td>
Forecast run: 00UTC
Forecast Horizon: 48h
Seasons: JJA22, SON22, DJF23, MAM23
Area: 43.5/5.0/48.2/16.0</td><td>
Area: national domains
Forecast Horizon: variable
Seasons: JJA22, SON22, DJF23, MAM23</td></tr>
<tr><td>Models</td><td>Global: ICON, IFS
LAMS: DWD: ICON-EU, COMET: COSMO-ME
IMGW-PIB: COSMO-PL7</td><td>Driving models: ICON-EU, IFS, ICON
LAMs: DWD: ICON-D2, MCH: COSMO-1E (control), COSMO-2E, HNMS: ICON-GR
COMET: COSMO-I2, ICON-I2, ARPA-E: COSMO-2I
IMS: ICON_IL2p5</td><td>COSMO and ICON-LAM
DWD, MCH, COMET, HNMS, IMGW-PIB, NMA, RHM, IMS, ARPA-E</td></tr>
<tr><td></td><td>FINE ComA-3</td><td>MIX ComA-TCC</td><td>Optional ComA-OnDemand</td></tr>
<tr><td>Specs</td><td>
Forecast run: 00UTC
Forecast Horizon: 48h
Seasons: JJA22, SON22, DJF23, MAM23
Area: 47.5/17.7/50.0/25.0</td><td>
Forecast run: 00UTC
Forecast Horizon: 48h
Period: June2022, Dec2022, Apr2023
Area: 31.0/18.0/37.4/32.0</td><td></td></tr>
<tr><td>Models</td><td>Driving models: ICON-EU, IFS, ICON
LAMs: HNMS: ICONGR2.5, NMA: COSMO-NMA/ICON NMA. IMGW-PIB: ICON-PL, IMS: ICON–IL2.5</td><td>Driving models: IFS, ICON
LAMs: HNMS: ICONGR2.5, IMS: ICON_IL2p5, ICON-EU</td><td>Area and LAMs, Specs based on specific experiment</td></tr>
</table>

Common Plots: verification specifications

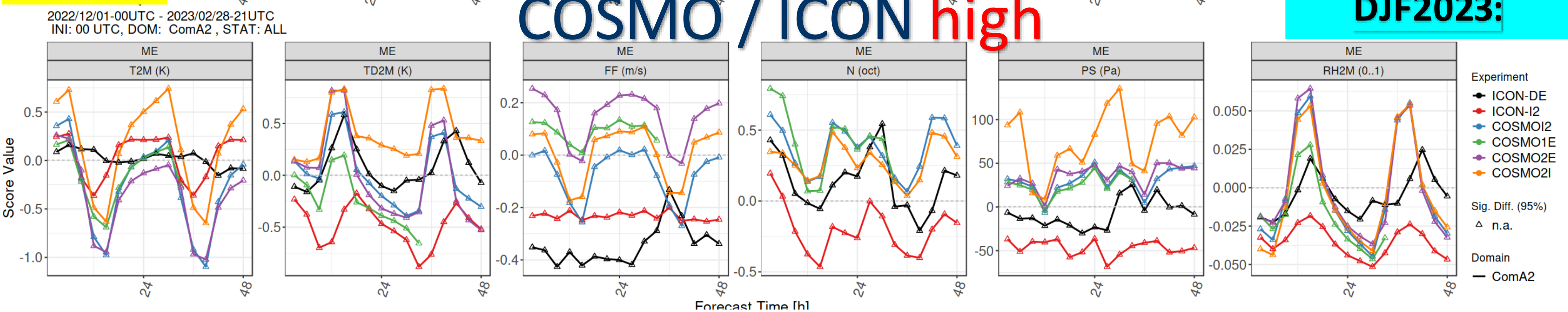
Parameters	**Surface: Continuous** T2m, SurfPressure, Td, WSpeed, TCC **Surface: Dichotomic** 6h Precip (Thresholds):0.2,0.8,1,5,8,10,15,20 mm TCC (Intervals): [0,25], (25,75), [75,100] Wind gust(thresholds):12.5,15,20 m/sec **Upper Air**: Temperature, RH, Wind Speed **Areas:** ComA-1, ComA-2, ComA-3, NoComA
Stratification	**-100m, 100m-300m, 300m-800m and >800m** **Areas:** ComA-1, ComA-2, ComA-3, NoComA
PointVerif Indices	**Surface - Continuous/UpperAir**: ME, RMSE, StdDv **Surface – Dichotomic**: Contingency table attributes: FBI, ETS, CSI **Areas:** ComA-1, ComA-2, ComA-3, NoComA
SpatialVerif Precipitation	**Obs:** OPERA composite **Indices:** FSS, POD, FAR, FBI, TS **Spatial windows:** 2.8, 8.4, 14, 25.2, 47.6, 92.4km **Areas:** ComA-2, ComA-3 **Resolution:** 0.025
SpatialVerif Total Cloud Cover	**Obs:** NWC-SAF Cloud Mask **Indices:** FSS, FBI, TS **Spatial windows:** 2.8, 8.4, 14, 25.2, 47.6, 92.4km **Thresholds:** 0, 20, 40, 60, 80, 100% **Areas:** ComA-TCC **Resolution:** 0.025
Conditional	**Critical Choices: conditions imposed on the observations, alignment is ON** **A.2mT verification when:** Total cloud cover >= 75% Total cloud cover <= 25% **B. Wind Speed when:** *(roughness length as constant field)* Roughness length<0.2m Roughness length>1m **Areas:** ComA-2, ComA-3

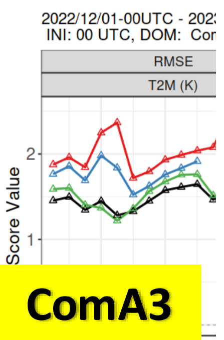


ComA2

COSMO / ICON high

DJF2023:





ComA3

2mT: Strong diurnal cycle, underestimation at noon, reduced with ICON models

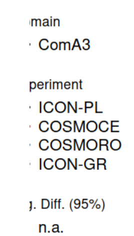
2mTd: Ovestimation during midday

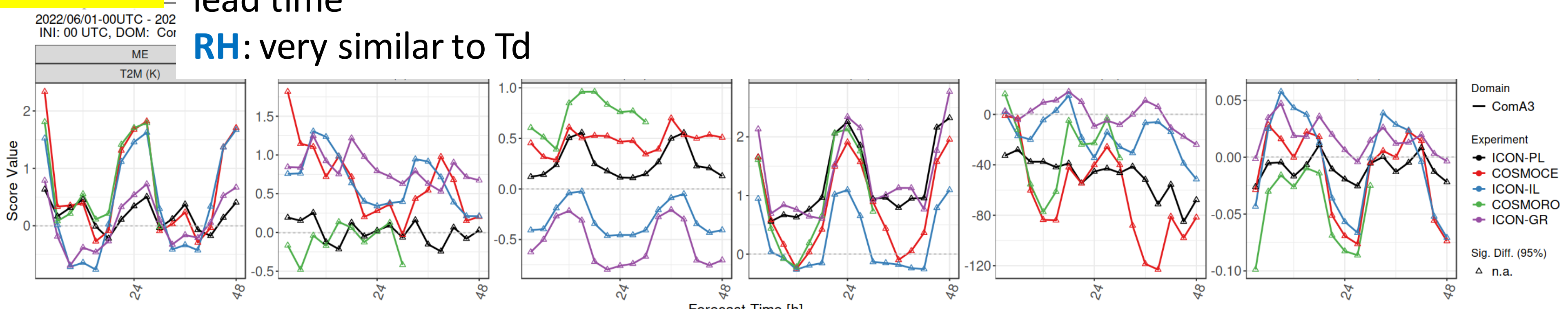
Windsp: Tendency ICONs to underestimate relative to COSMO models

TCC: Larger errors at night, overpredicting presence of clouds

Pressure: Clear improvement with ICON in error and increasing tendency with lead time

RH: very similar to Td

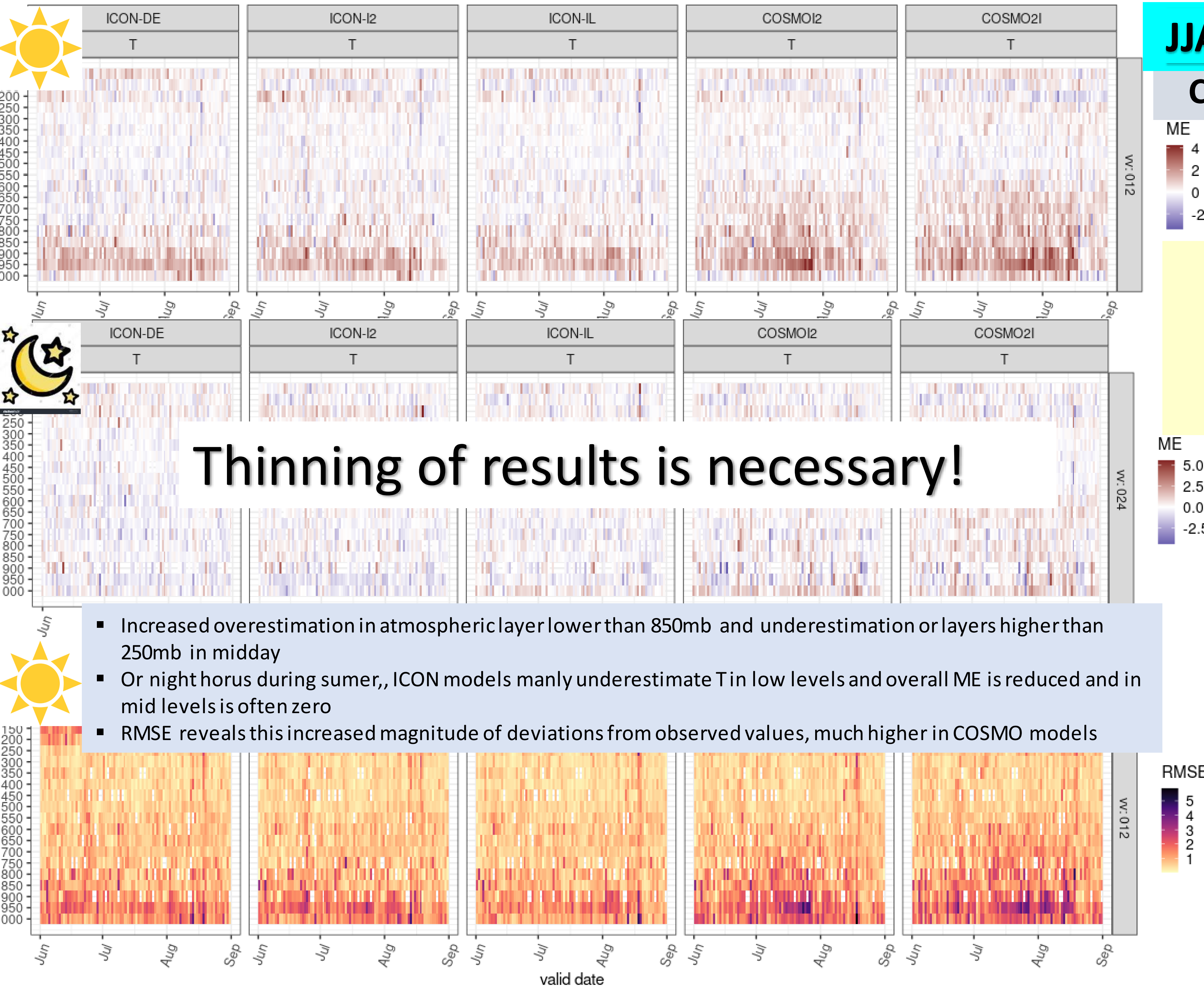




JJA2022

ComA2

UPPER AIR



Thinning of results is necessary!

- Increased overestimation in atmospheric layer lower than 850mb and underestimation or layers higher than 250mb in midday
- Or night horus during sumer,, ICON models manly underestimate T in low levels and overall ME is reduced and in mid levels is often zero
- RMSE reveals this increased magnitude of deviations from observed values, much higher in COSMO models

Meeting on: MODEL ERRORs Identification

Videoconference: November 2023 (TBD)

Based on Common Area and National Domain verification results
Standard and Conditional Verification
Fuzzy on precipitation and TCC

Aims:
Relative performance of COSMO/ICON implementations
Reporting of systematic errors of ICON-LAMs (dependence on: season, hour, geographical location, weather, other parameters)
Tuning on systematic model errors

Summary to be included in COSMO newsletter: verification report

ICON-CH1 versus COSMO-1E Control

Dichotomic Verification **Spring** 2023

Deterministic Verification ICON-CH1 versus COSMO-1E Control

Parameter	THS				FBI			
	ch		ch-sp		ch		ch-sp	
	01-12	13-24	01-12	13-24	01-12	13-24	01-12	13-24
TOT_PREC6 ≥ 0.2	-2 %	-3 %	-1 %	-2 %	-12 %	-15 %	-7 %	-8 %
TOT_PREC6 ≥ 1	-1 %	-4 %	-1 %	-2 %	-10 %	-15 %	-4 %	-6 %
TOT_PREC6 ≥ 5	-1 %	-3 %	+2 %	-3 %	-1 %	-9 %	-9 %	+4 %
CLCT ≥ 2.5	-1 %	-0 %	-0 %	+0 %	-3 %	+0 %	-0 %	+3 %
CLCT ≥ 6.5	+2 %	+2 %	+1 %	+1 %	-17 %	-13 %	-9 %	-6 %
VMAX_10M6 ≥ 12.5	-0 %	-0 %	-4 %	-5 %	-10 %	+1 %	-60 %	-66 %
VMAX_10M6 ≥ 20	-3 %	-2 %	-6 %	-5 %	-31 %	-15 %	-59 %	-69 %

✓

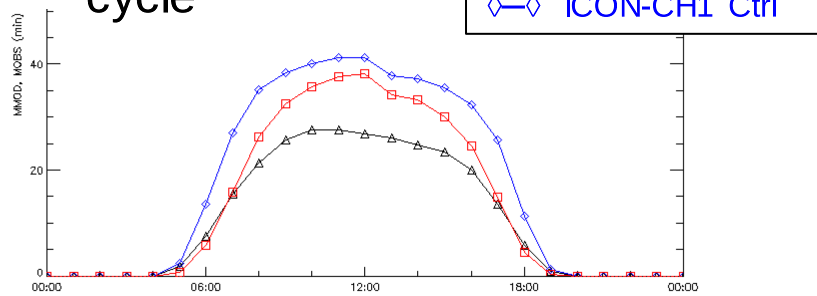
pirmin.kaufmann@meteoswiss.ch

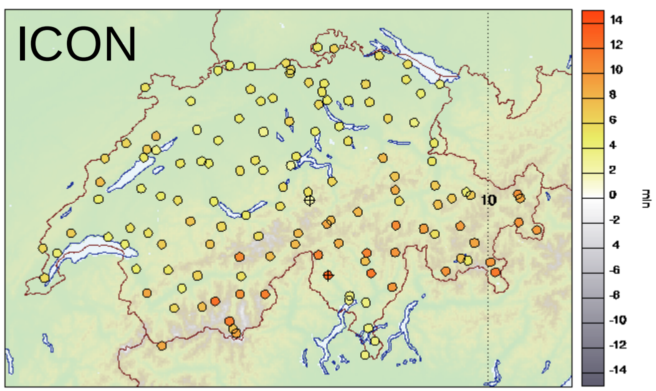
Sunshine Duration

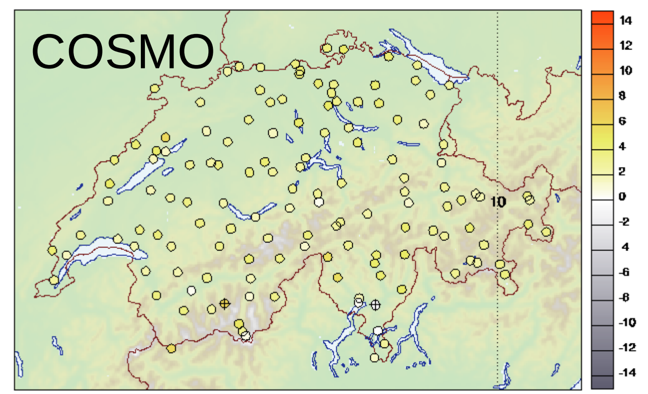
Sunshine = direct shortwave radiation > 200 W m^{-2}

- ICON stronger overestimation
- Mostly south side of Alps

Diurnal cycle

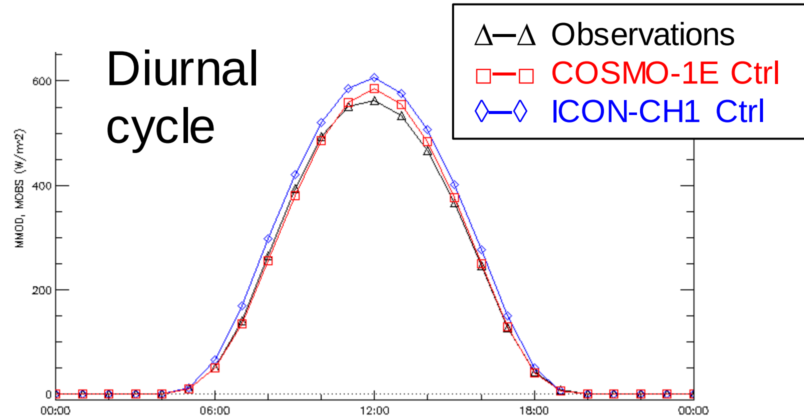


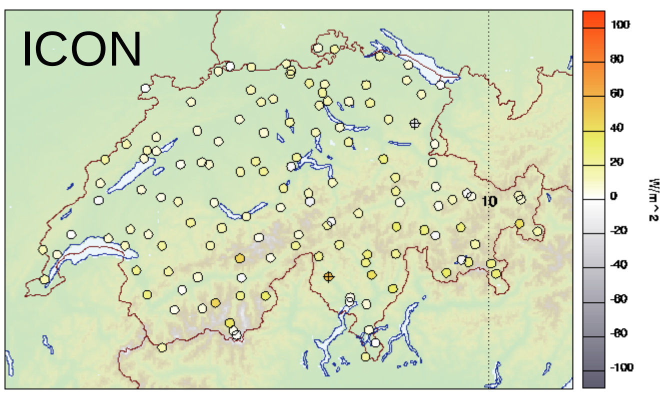




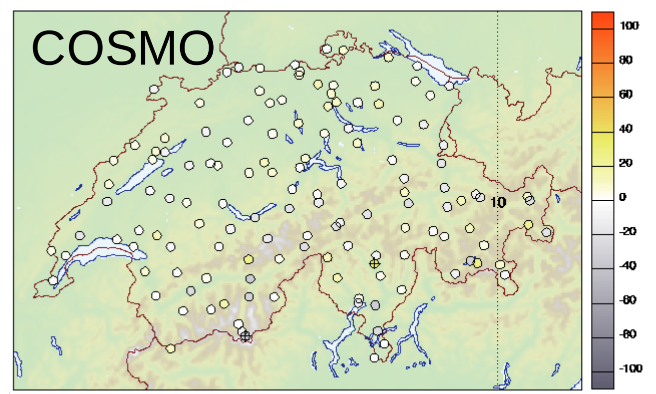
Shortwave (Global) Radiation

- ICON only slightly stronger overestimation
- → Problem due to partitioning between direct and diffuse



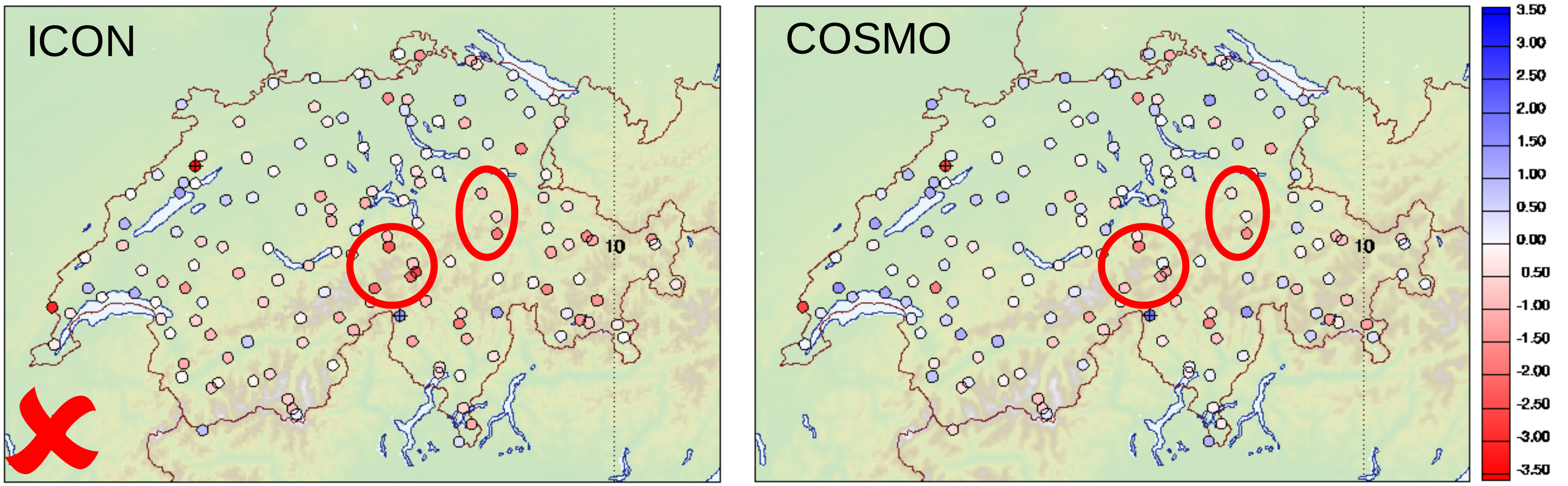


≈



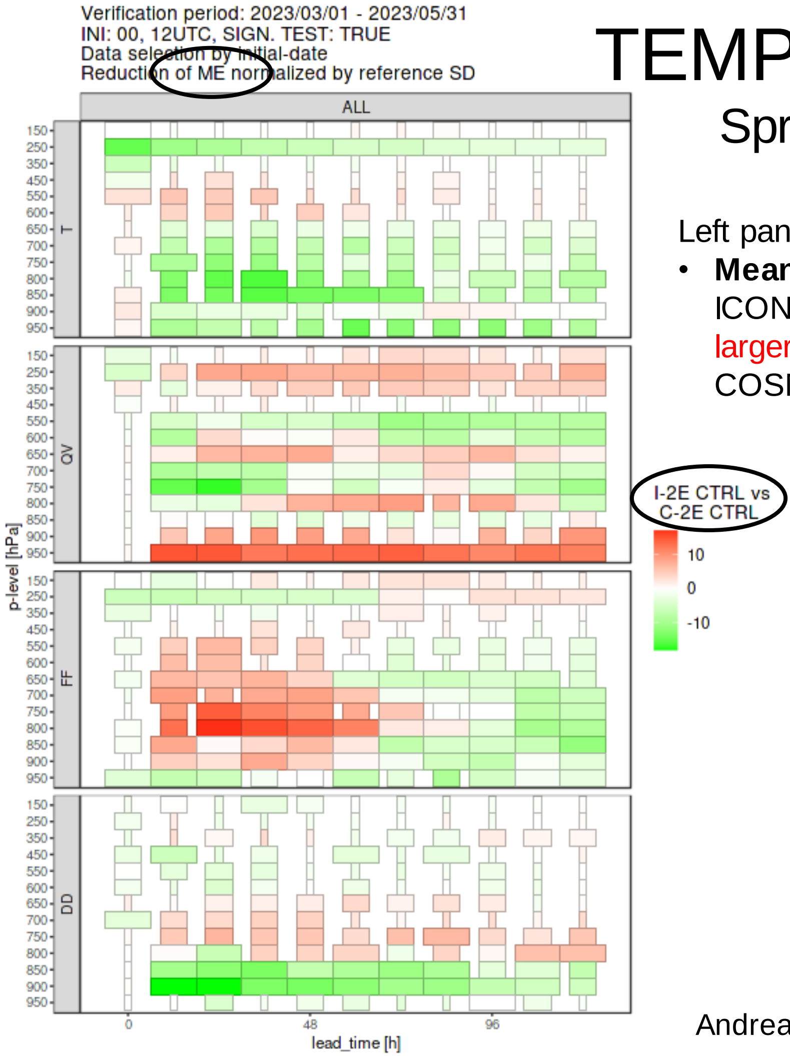
Wind Speed

ICON underestimates wind speeds in complex terrain (mountain tops, narrow valleys) more than COSMO



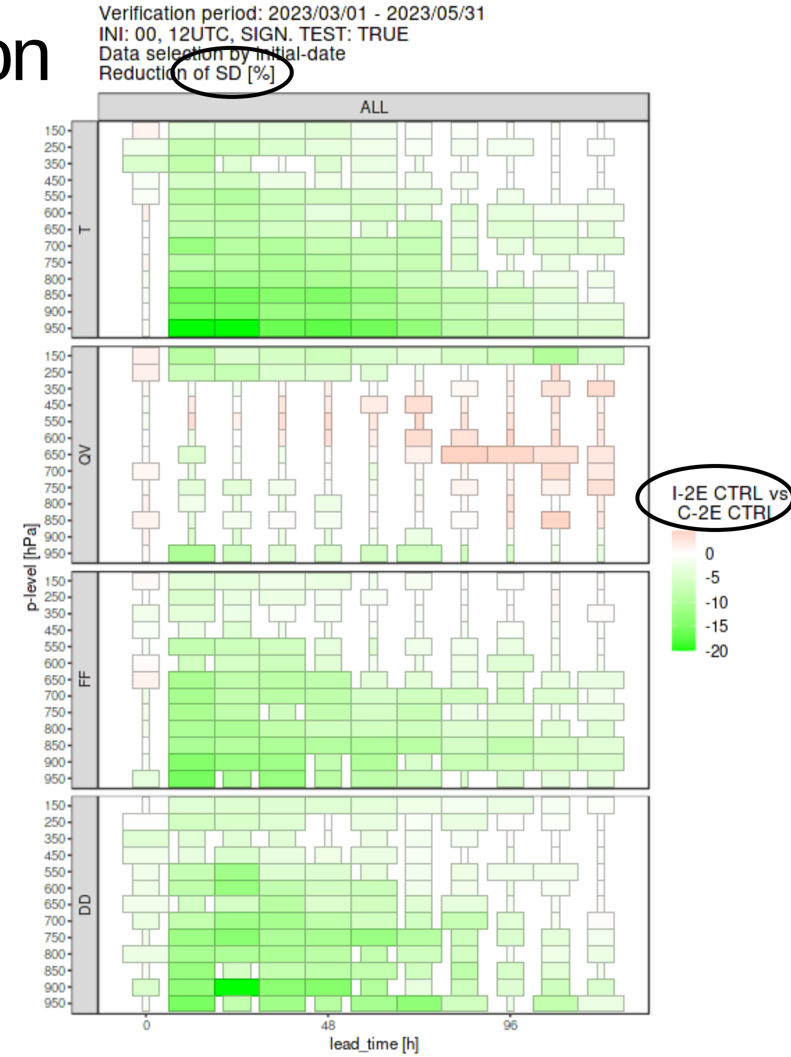
TEMP Verification

Spring 2023



Left panels:

- **Mean error** (Bias) of ICON-CH2 partially larger than bias of COSMO-2E CTRL



Right panels:

- **SD** (standard deviation of the error) of ICON-CH2 much smaller than SD of COSMO-2E CTRL

Andreas Pauling, MeteoSwiss

WG5 activities contributors

F. Fundel, A. Iriza-Burca, J. Linkowska, N. Vela, S. Gabrian, S. Dinicila, M.S.Tesini, F. Sudati, D. Boucouvala, P. Kaufmann, E. Oberto, P. Khain, M. Bogdan, B. Maco, F.Gofa,….

CLM-Community EVAL Group meeting

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

ICCARUS 2023, 09 March 2023

TOPICS

- WG5 current activities - enable R Shiny also for global and regional climate ICON simulations ("seamless verification"), especially for testing of new ICON versions
- Tuning ICON-Seamless-Atmosphere
- COPAT2
- Discussion on future collaboration between WG5 and WG EVAL
  - focus on topics of common interest.
  - prepare a list with points of common interest
  - preliminary meeting in a smaller group to discuss a priority list of common interest