COSMO Verification Overview

Flora Gofa

(on behalf of)

Working Group on Verification and Case studies

37th EWGLAM and 22th SRNWP Meeting, Belgrade, 05-08 October 2015

Definition of Common Verification Activities

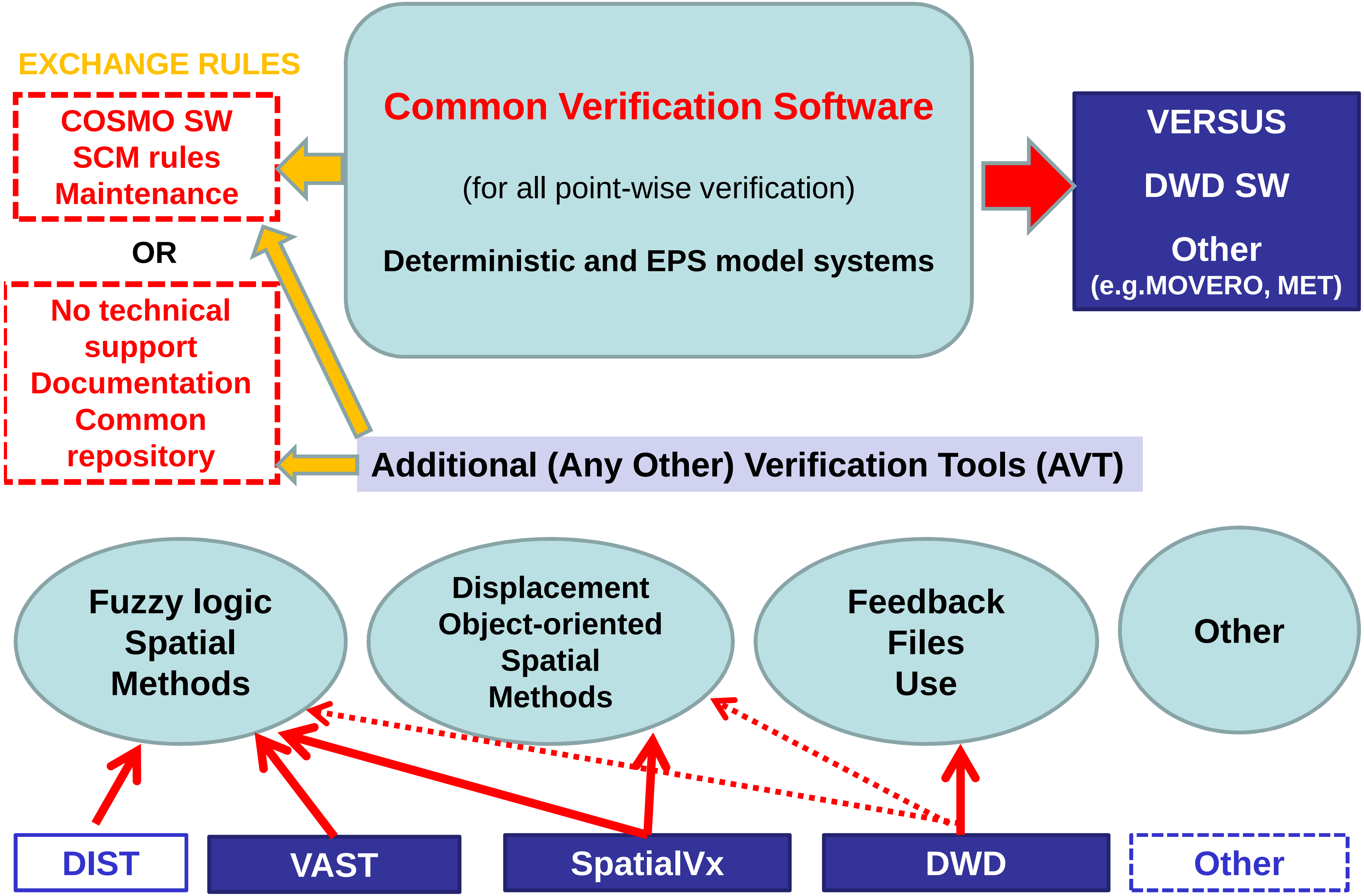
Science Plan Priorities: Investigation on statistical methods to identify the skill of convection-permitting and near convection-resolving model configurations, probabilistic and ensemble forecast verification, severe and high impact weather verification.

Common Plot Seasonal Reports: Verification results of statistical indices for main weather parameters derived using the operational COSMO model implementations in each service. The domain (common or custom), resolution, statistical scores/methods, frequency and graphical representation, are decided on an annual basis from WG5. The main findings of this organized analysis is presented during the GM plenary session together with the long term trend of them, providing a basis to track the performance of COSMO model

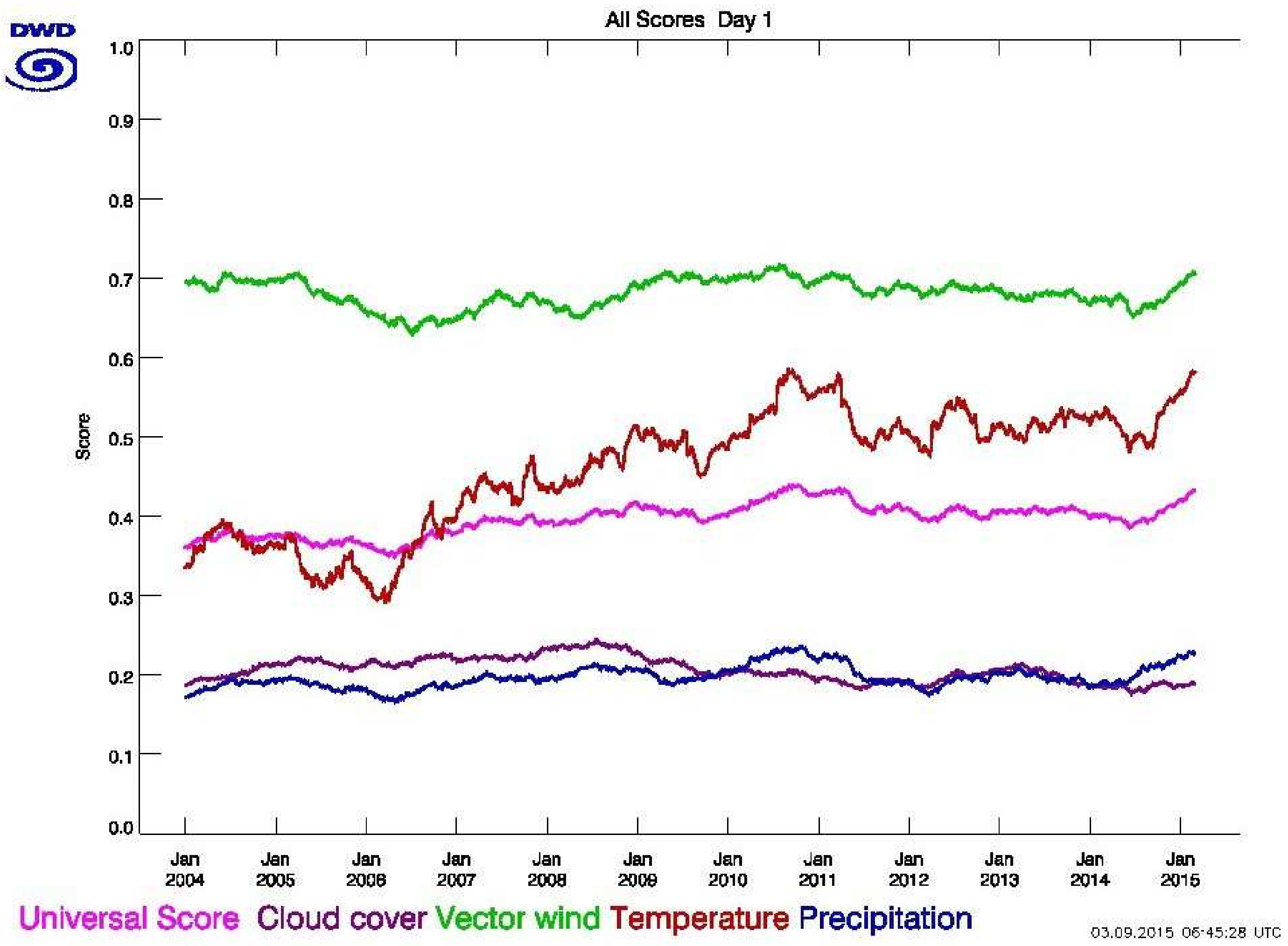
Conditional Verification Tests: Methodical evaluation of model performance in order to reveal the typical shortcomings of a model and to provide information to the model developers as well as to the forecasters with regard to model reliability.

Strategy on verification tools

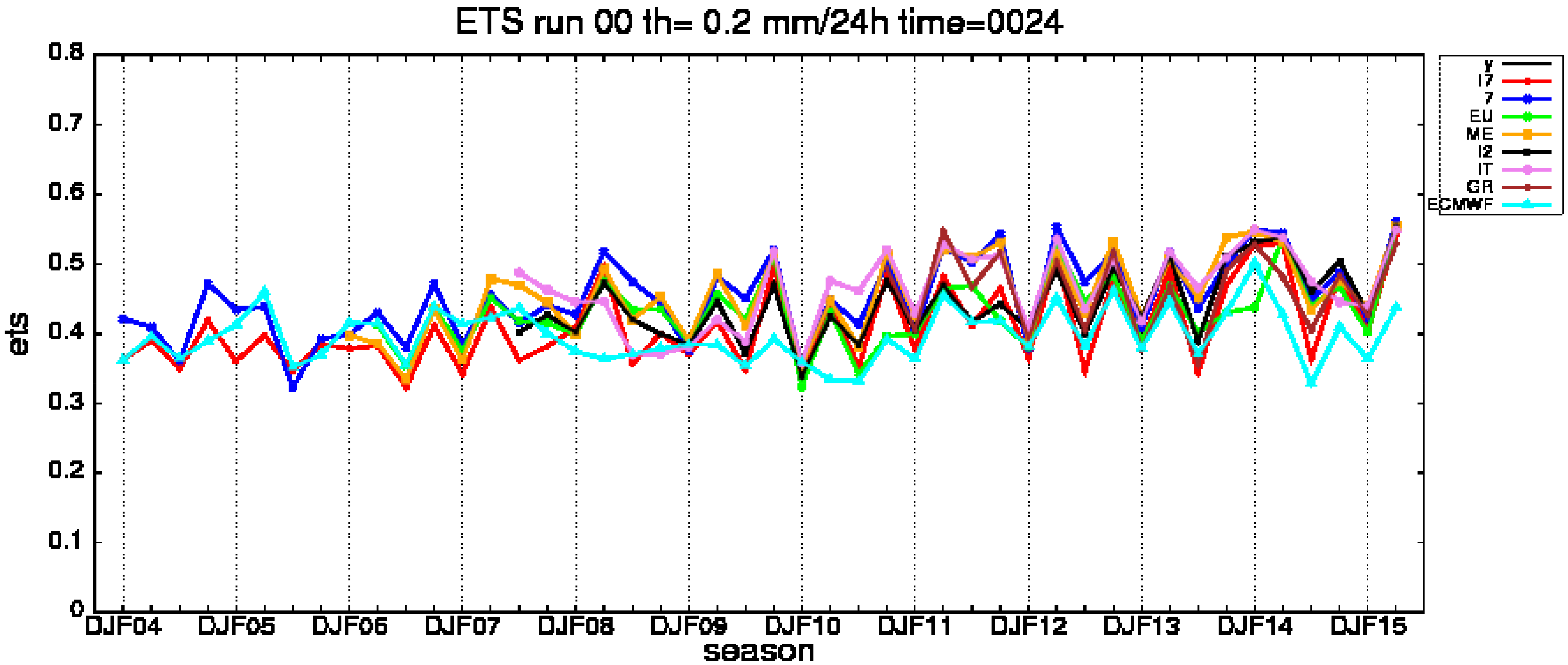
C V S +

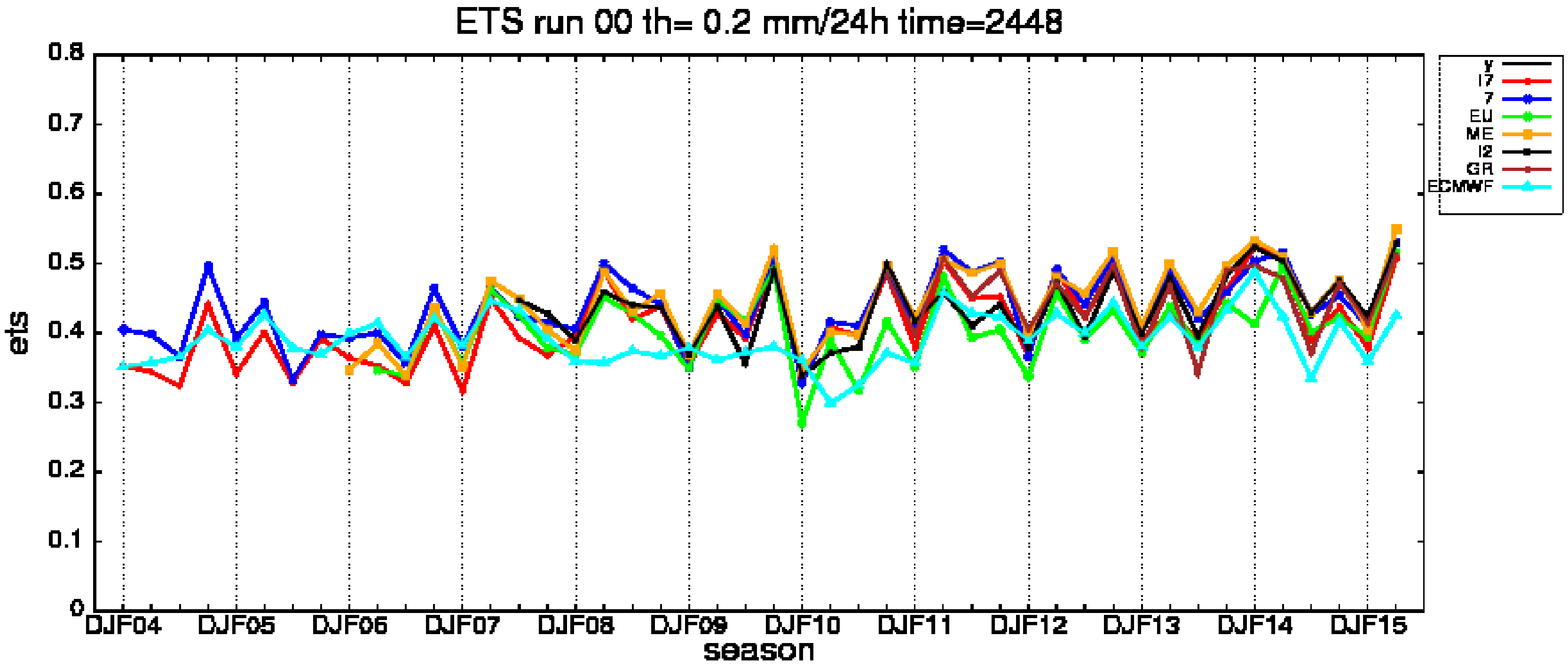

EXCHANGE RULES
COSMO SW
SCM rules
Maintenance
OR
No technical support
Documentation
Common repository
Common Verification Software
(for all point-wise verification)
Deterministic and EPS model systems
VERSUS
DWD SW
Other
(e.g.MOVERO, MET)
Additional (Any Other) Verification Tools (AVT)
Fuzzy logic
Spatial
Methods
Displacement
Object-oriented
Spatial
Methods
Feedback
Files
Use
Other
DIST
VAST
SpatialVx
DWD
Other

The COSMO-Index
Day 1



LONG TREND PRECIPITATION with high resolution stations

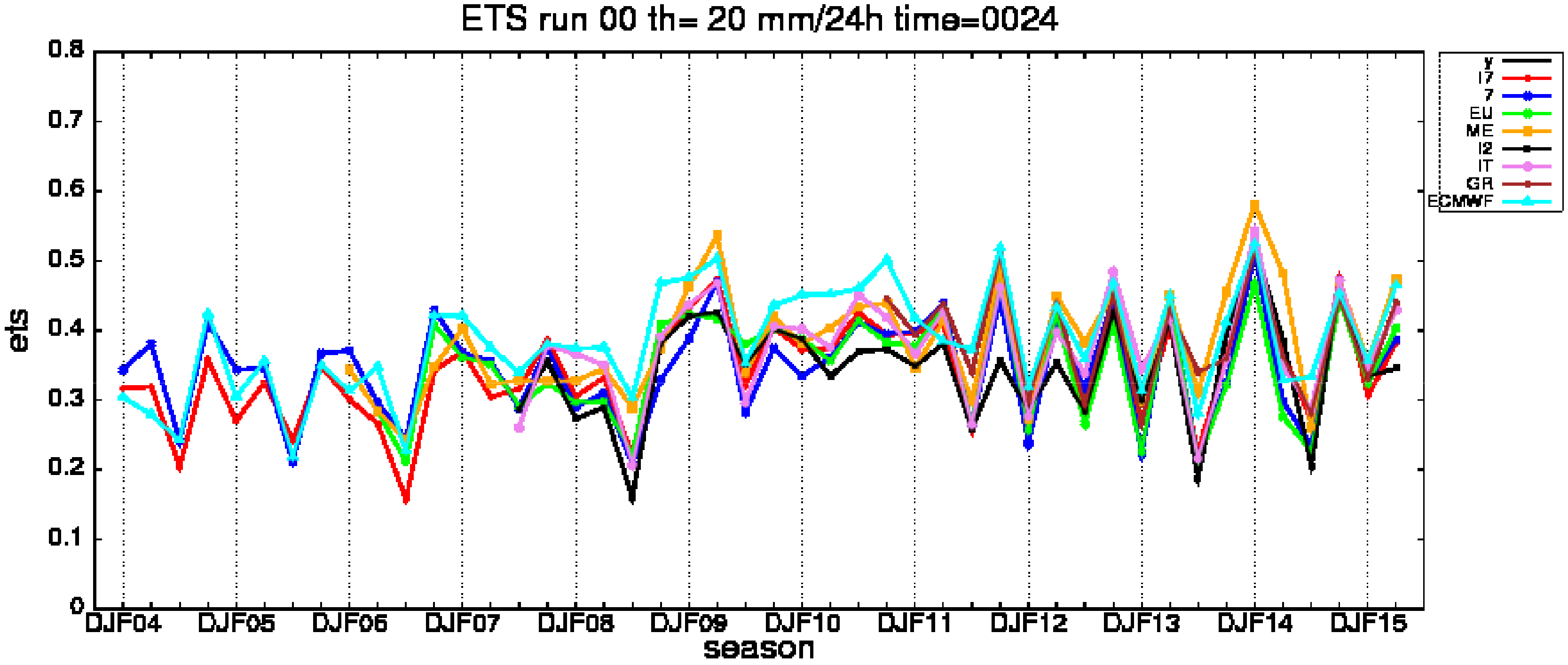


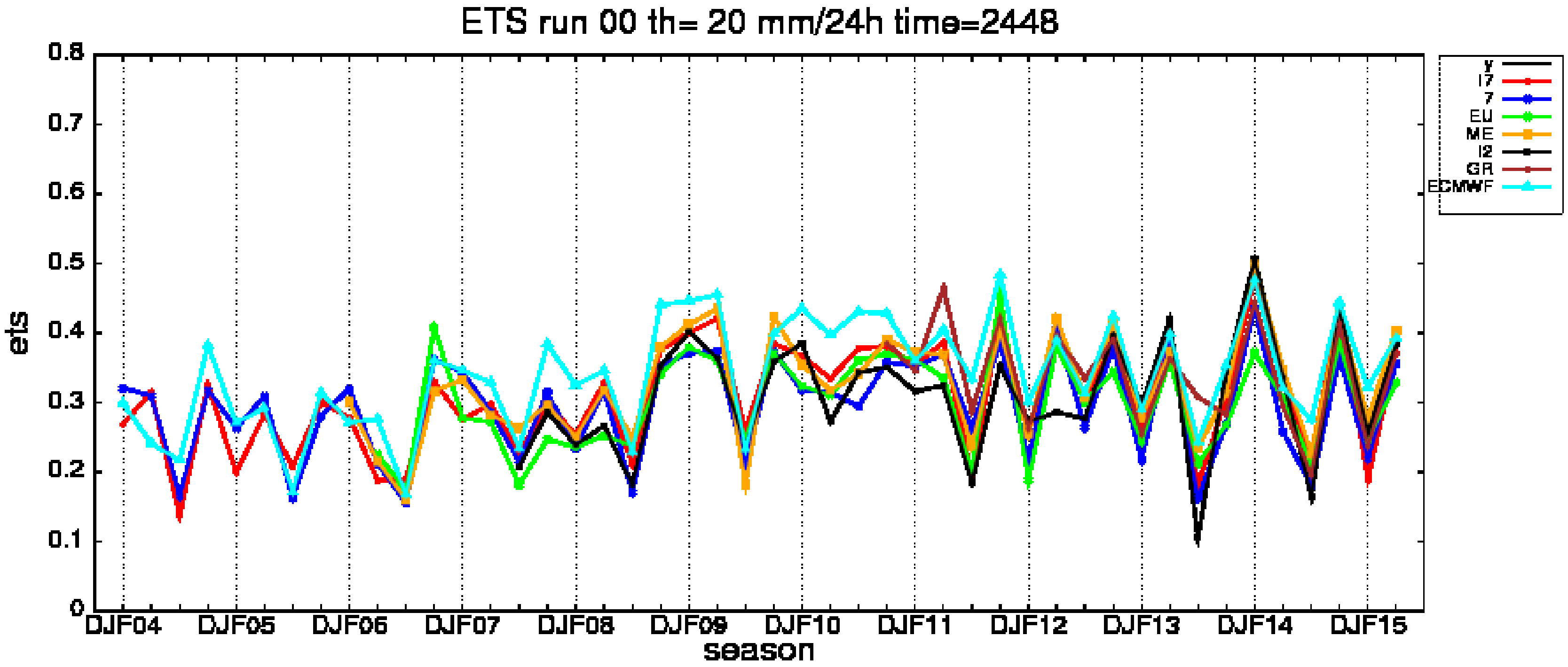


LOW THRESHOLDS

LONG TREND PRECIPITATION with high resolution stations

HIGH THRESHOLDS





Different methods have different aims

Scale separation and neighbourhood methods

- Focus on accuracy quantification
- What is the forecast accuracy at small scales? Large scales? Low / high intensities?
- What spatial scales and intensities have reasonable accuracy?
- Different methods emphasize different aspects of accuracy

Feature-based methods

- Focus on describing the error
- What is the error in this forecast?
- What is the cause of this error (wrong location, wrong size, wrong intensity, etc.)?

Field deformation methods (morphing)

- Focus on describing phase errors
- Does the shape/placement of the forecast resemble the observations?

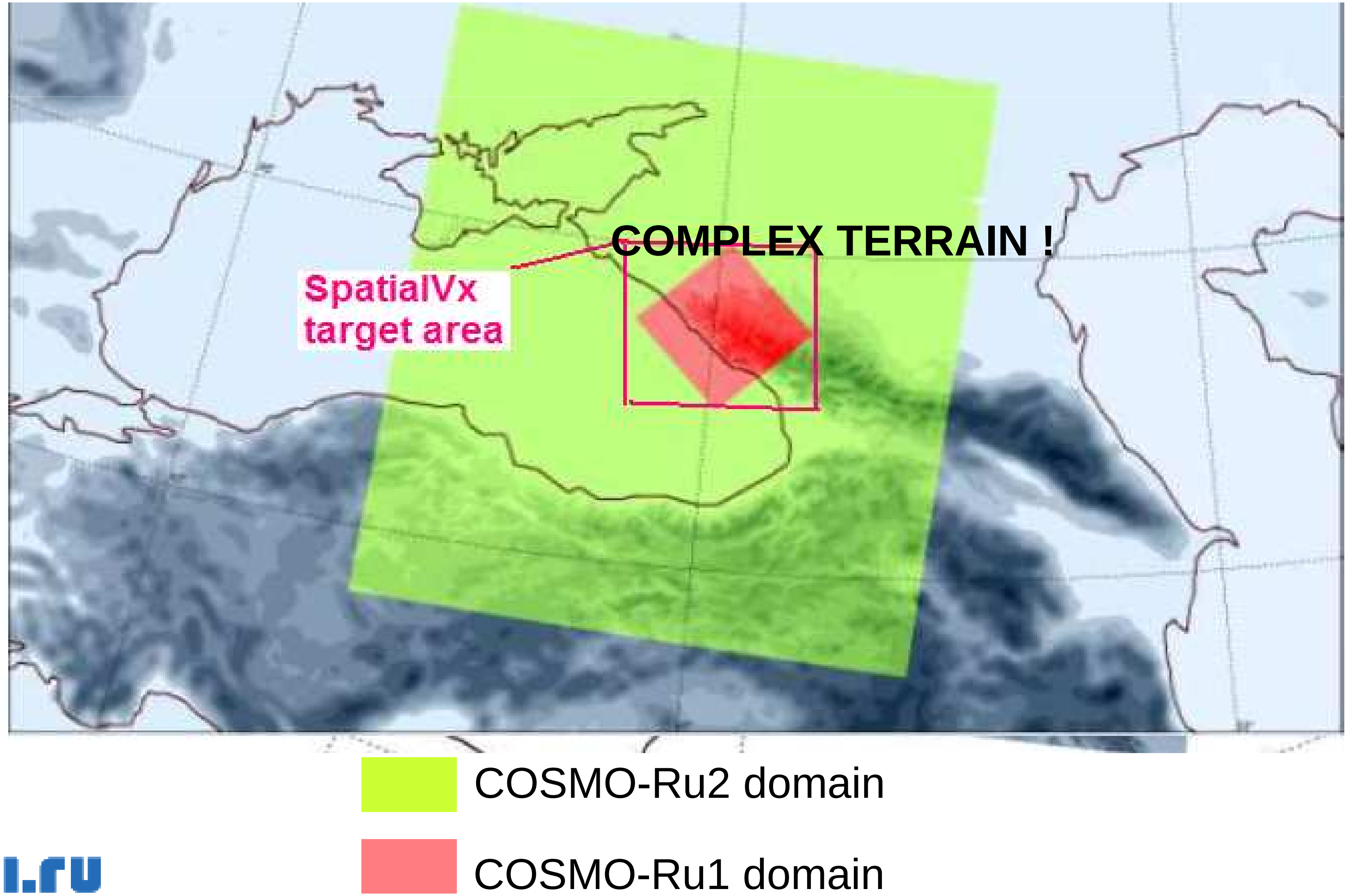
COSMO Priority Project

INSPECT: INtercomparison of SPatial vErification methods for COSMO Terrain

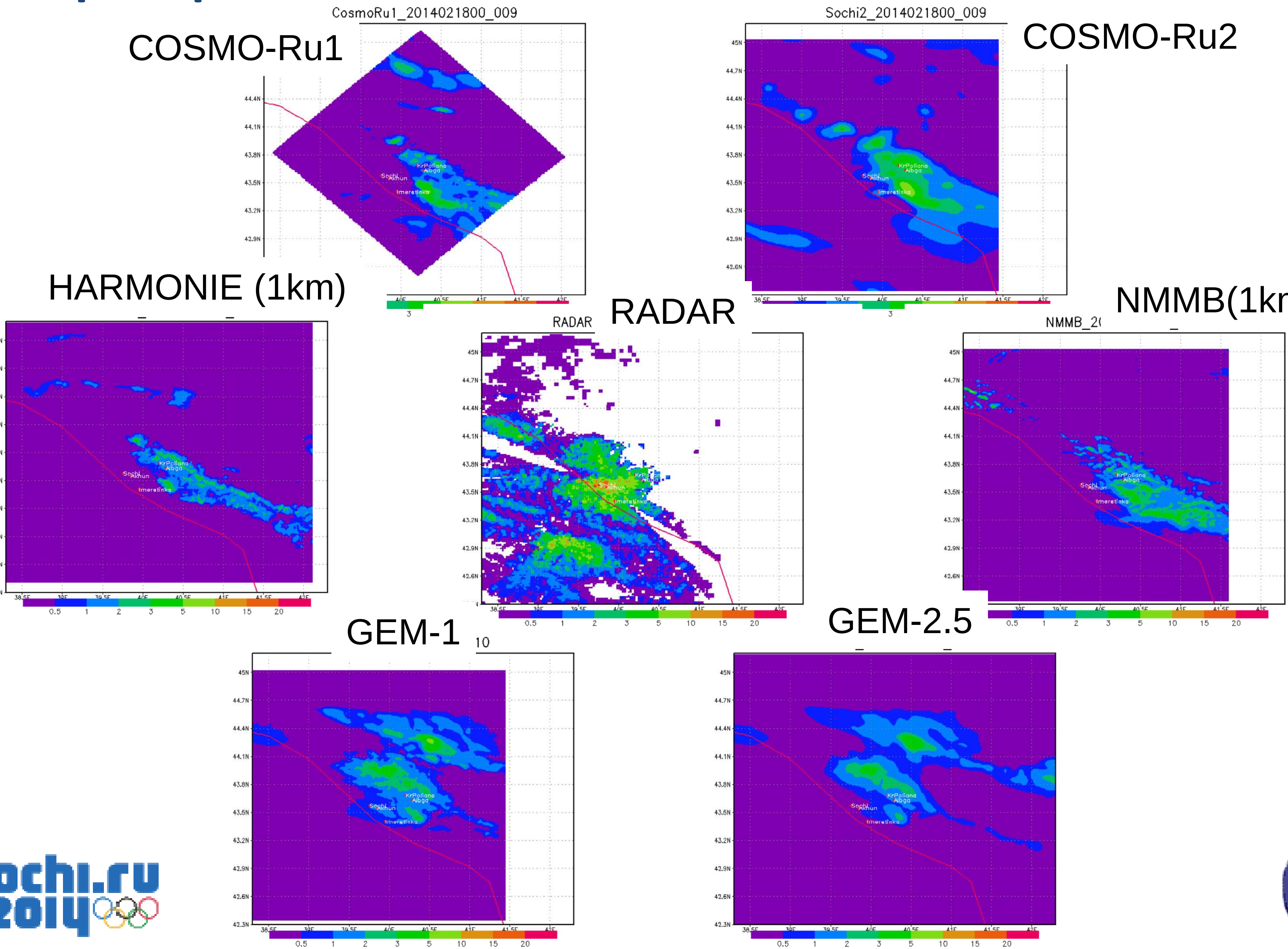
- runs in parallel to MesoVICT (Mesocale Verification Intercomparison in Complex Terrain, community project)
- summarizes the COSMO experience of applying spatial verification methods to high and very-high-resolution systems
- *a wider range of **spatial** verification **methods** will become **commonly used** within the COSMO community and **Guidelines** will be proposed to ensure the correct interpretation of results of these methods.*
- Same as MesoVICT, INSPECT focuses on EPS *forecasts* and *variables besides precipitation*
- In addition to targeting the goals of MesoVICT, INSPECT provides more choice of verification domains and reference data - newer and longer periods, two complex terrains (the Alps and the Caucasus)
- Share the tools that will be developed or adapted for common use

Area of the study

349 lon points * 481 lat points with **0.00833** lat-lon increments.
1 grid size by **longitude** = 111*0.00833 = **930 m**,
1 grid size by **latitude** = cos(43°35')*930 m = 0.72*930 = **~ 670 m**

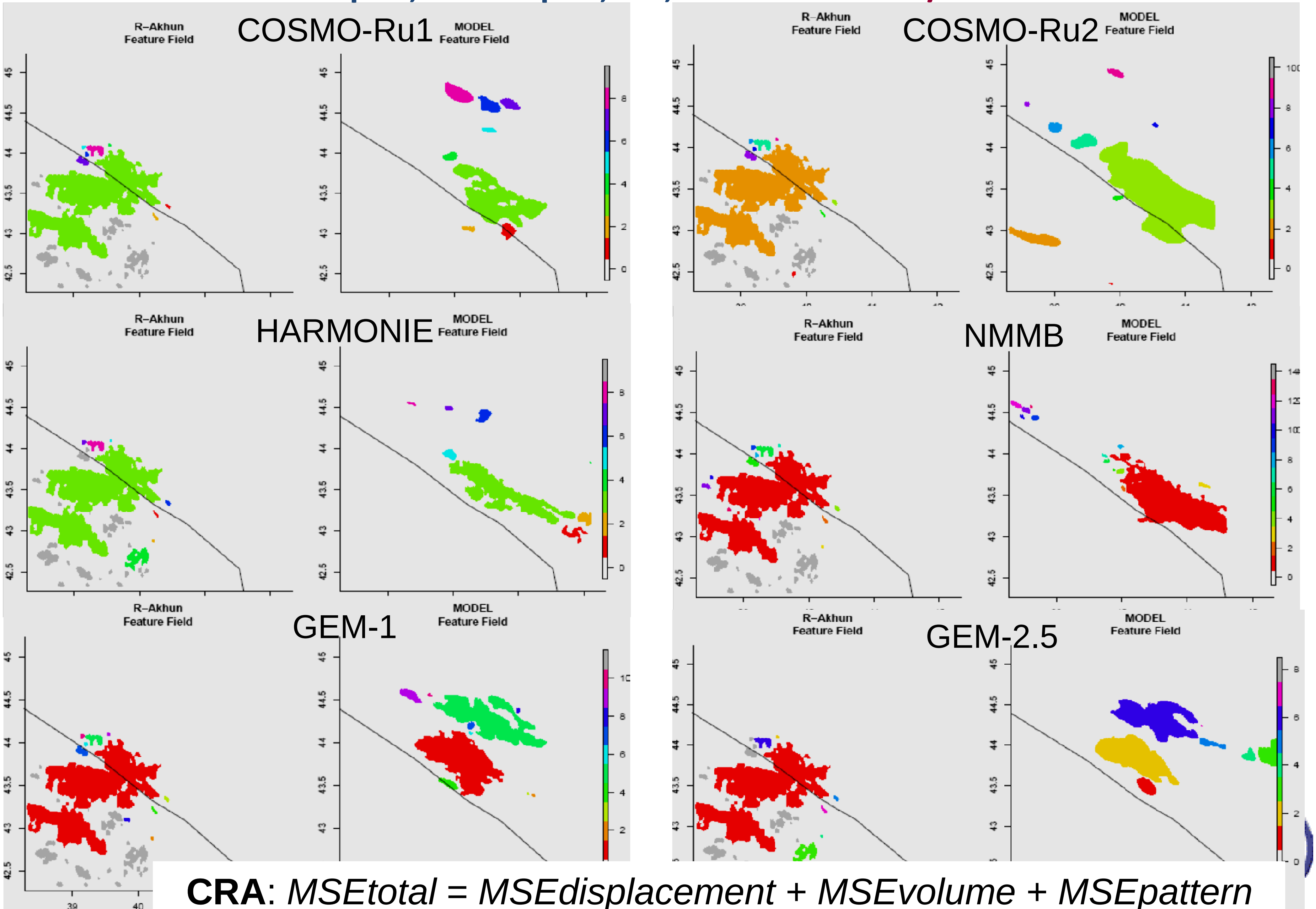


18 Feb 2014, 09 UTC, cold front: All models underestimated max precip and didn't give precip over the sea.



Pairs of matched objects from *craer*, 18 Feb 2014, 09 UTC

Colors indicate the 1st pair, the 2nd pair, etc, threshold: 1mm/h

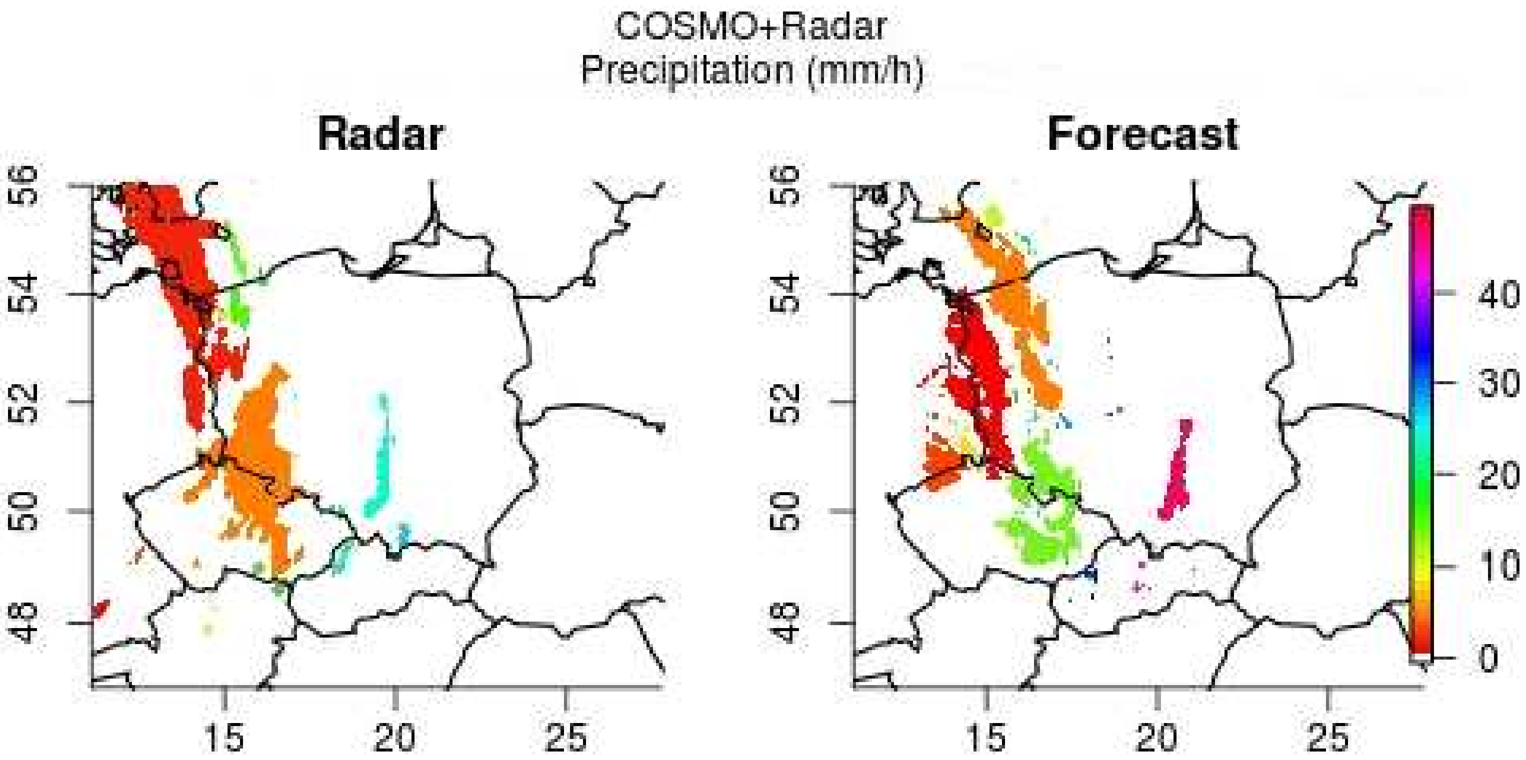


CRA: *MSEtotal = MSEdisplacement + MSEvolume + MSEpattern*

MODE Application in COSMO-PL

Identification of features

COSMO PL 2.8, 24.08.2015 18 UTC +T13
(25.08.2015, 07 UTC)

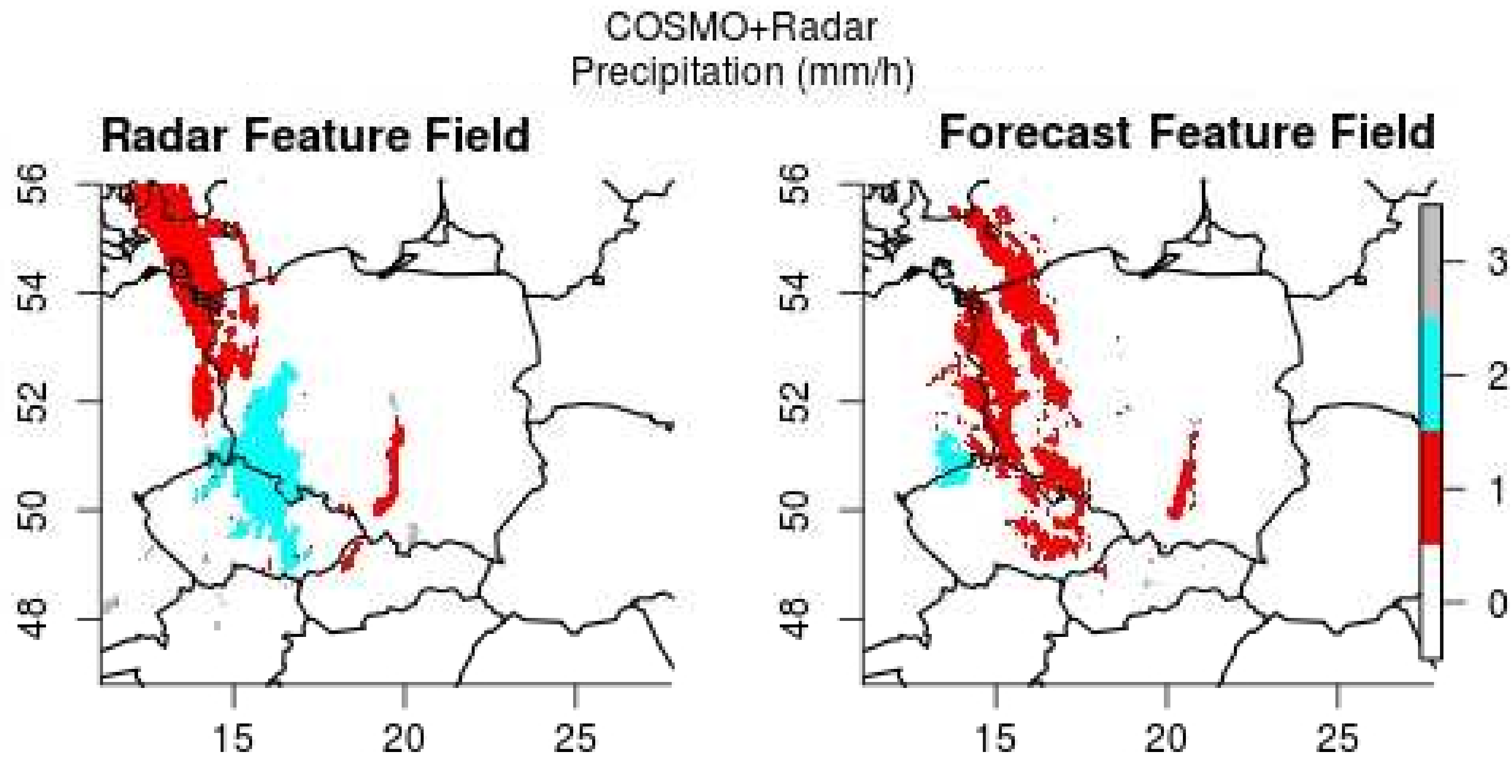


Smoothpar=1.5 , thresh=0.15,
24 radar features, 47 forecast features

Joanna Linkowska
Andrzej Wyszogrodzki
IMGW-PIB

Merging and/or matching features

COSMO PL 2.8, 24.08.2015 18 UTC +T13 (25.08.2015, 07 UTC)



2 implicitly defined merges groups (red, blue)

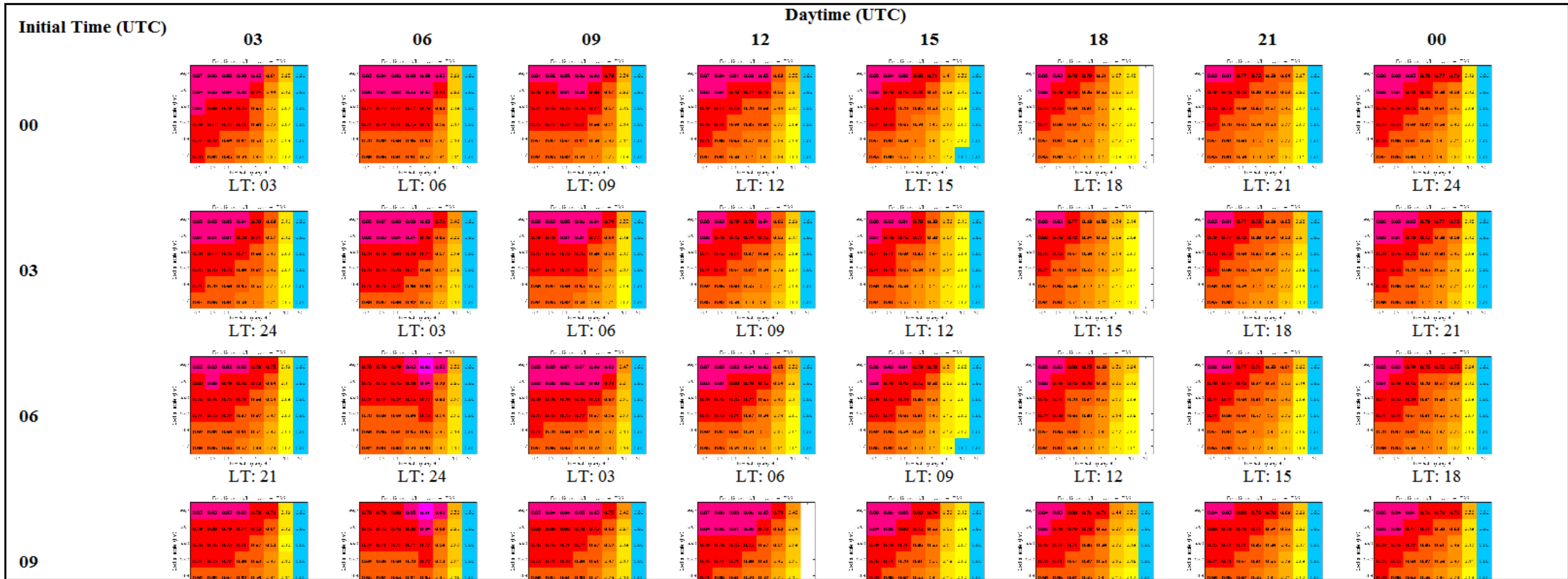
white - zero values, grey - unmatched features

D (centroid distance) < sum of the sizes of the two features

Joanna Linkowska
Andrzej Wyszogrodzki
IMGW-PIB

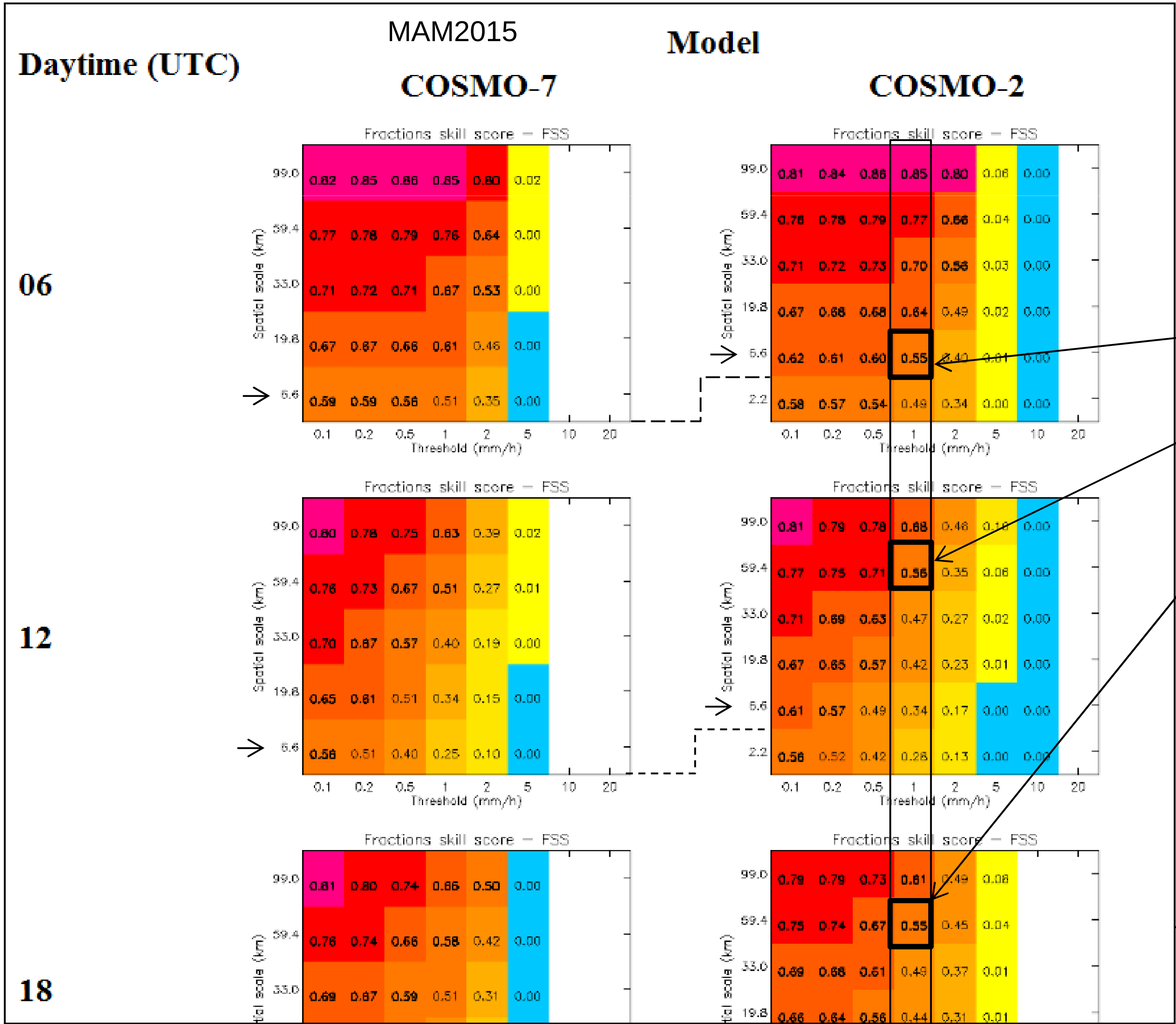
Evaluating model quality with spatial data

- Use available gridded data (radar, satellite)
- High resolution
- Applications : monitoring, verifications
- Produces potentially a lot of new data -> how to get useful/usable information ? for which applications ?



X. Lapillonne, D. Leuenberger, P. Kaufmann

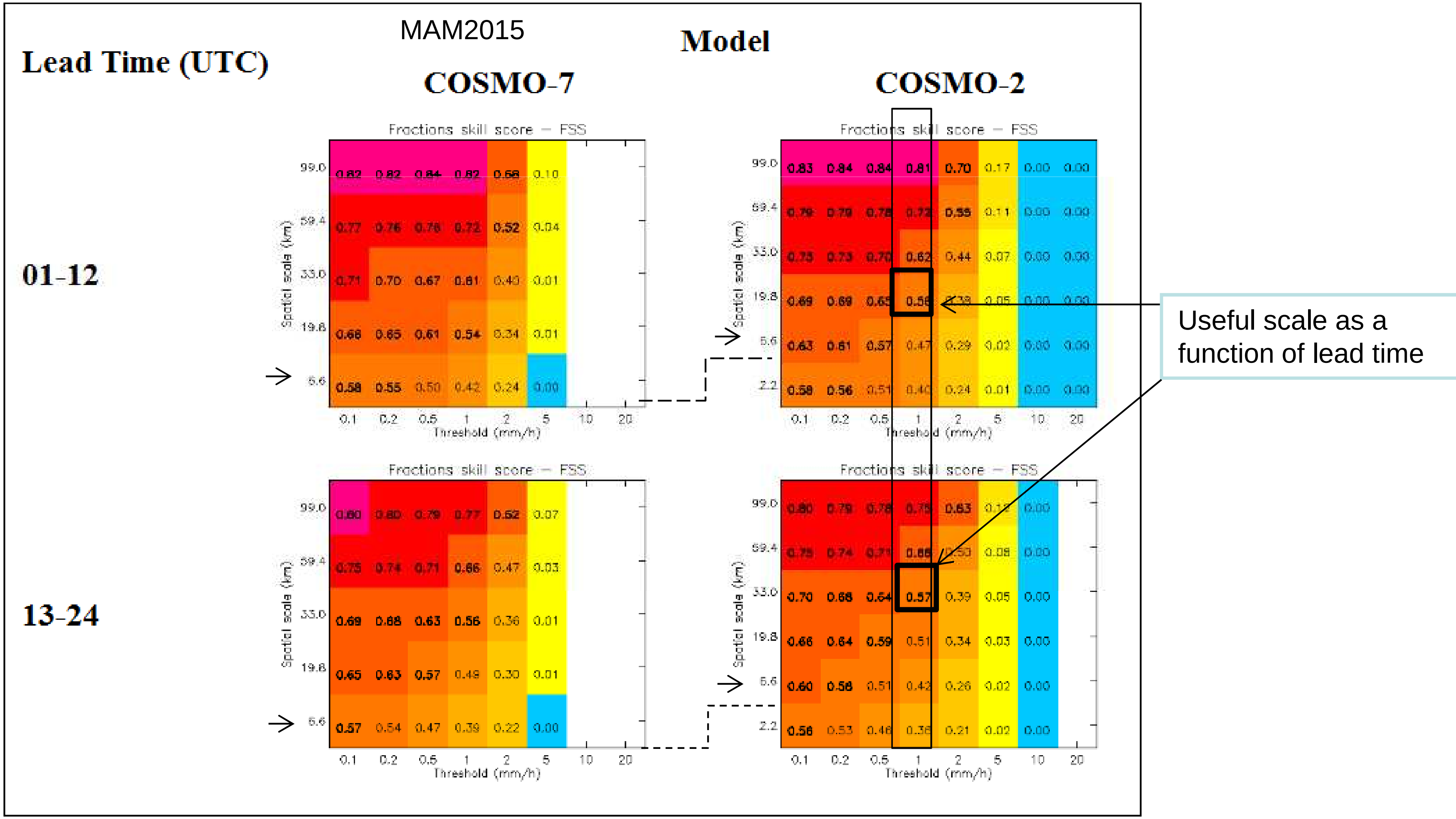
Aggregate over initial time : diurnal cycle



Useful scale at a given threshold (here 1 mm/h) for different day time

X. Lapillonne, D. Leuenberger, P. Kaufmann

Aggregate runs : lead time scores

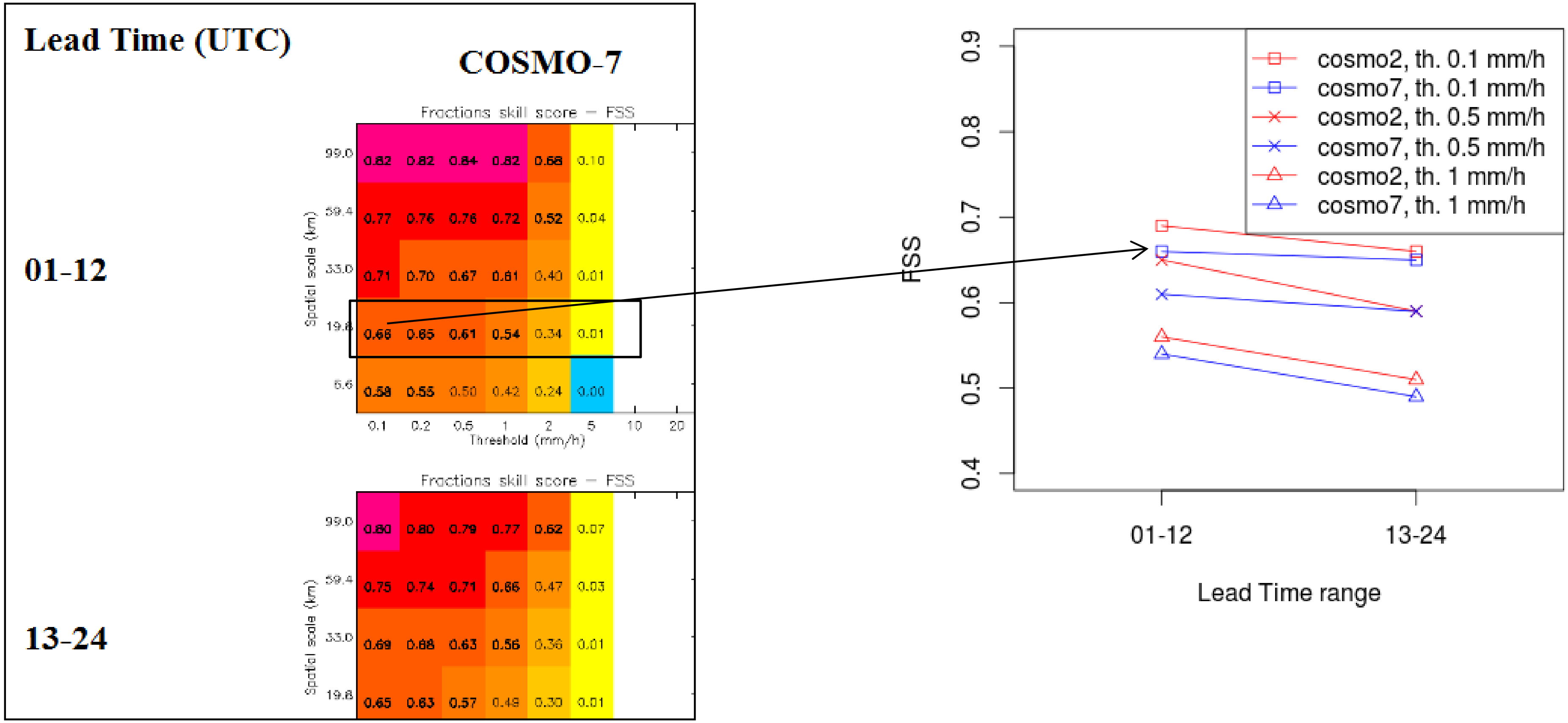


X. Lapillonne, D. Leuenberger, P. Kaufmann

Focus on a particular scale

- Generate similar plot as synop verification but using the gridded observation
- We choose here 19.8 km, close to warning region size

MAM2015

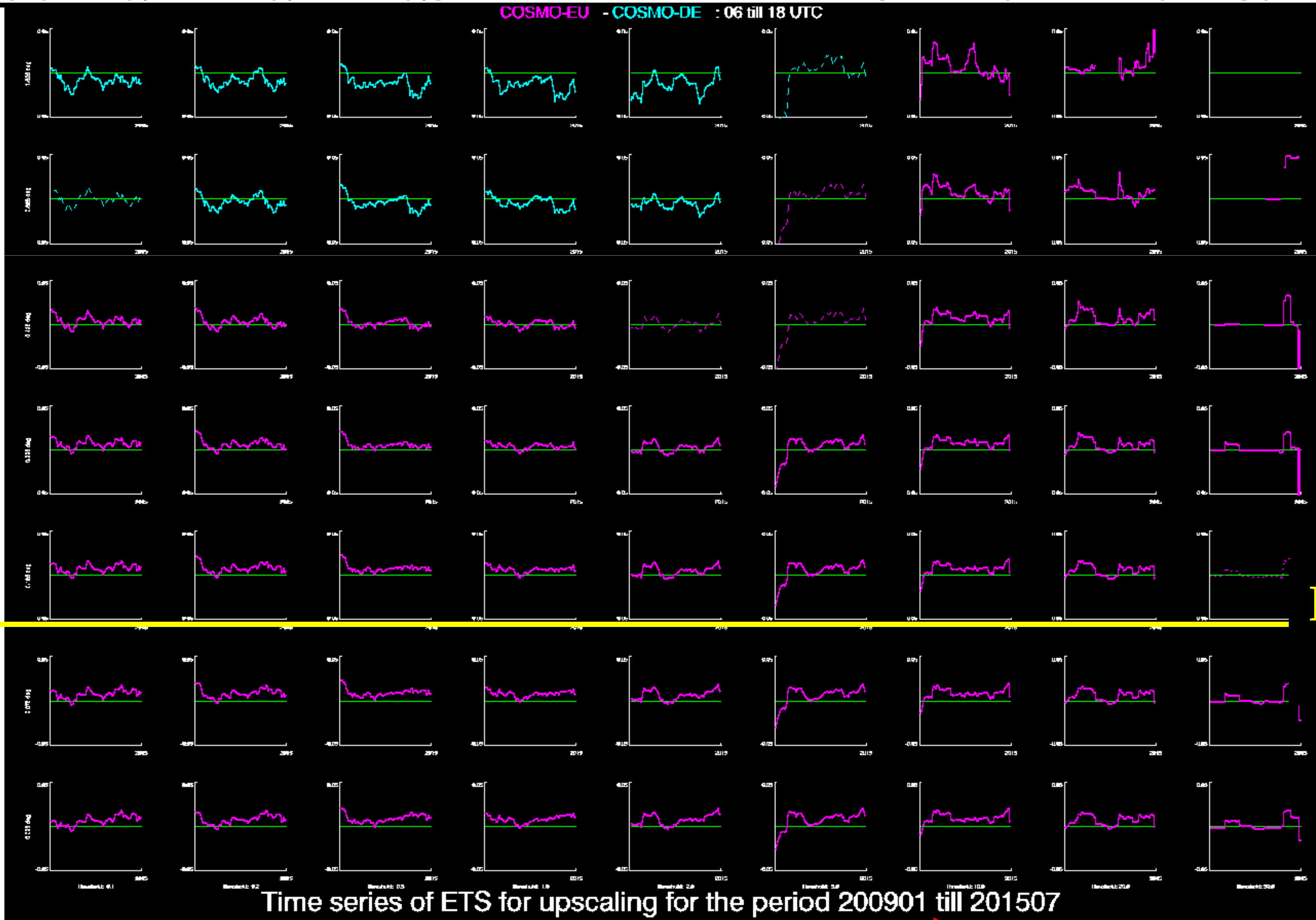


Comparison of COSMO-EU to COSMO-DE – upscaling ETS (differences)

DWD
Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Threshold 0.1 0.2 0.5 1 2 5 10 20 50

Window size: 1.625, 0.825, 0.425, 0.225, 0.125, 0.075, 0.025



Mesh width of COSMO-EU

Precipitation amount

Common Plot Reports

Verification | Display | Interpretation and Applications | Home

Common Verification Reports

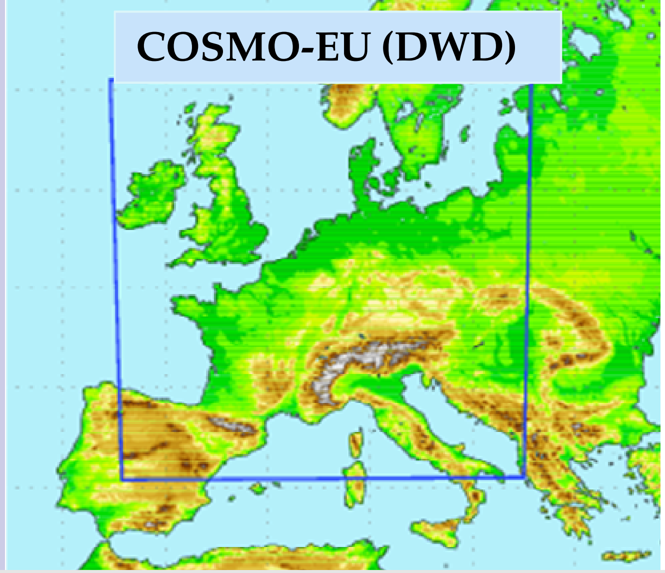
Last updated: at most recent report

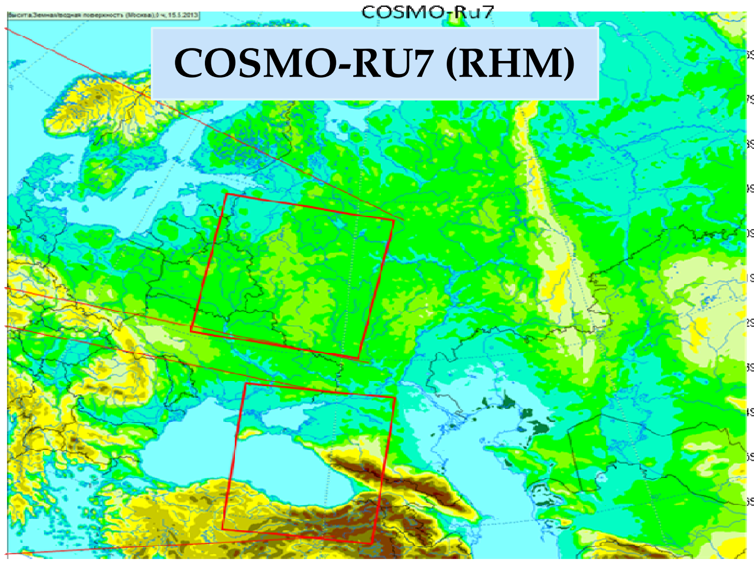
Edit

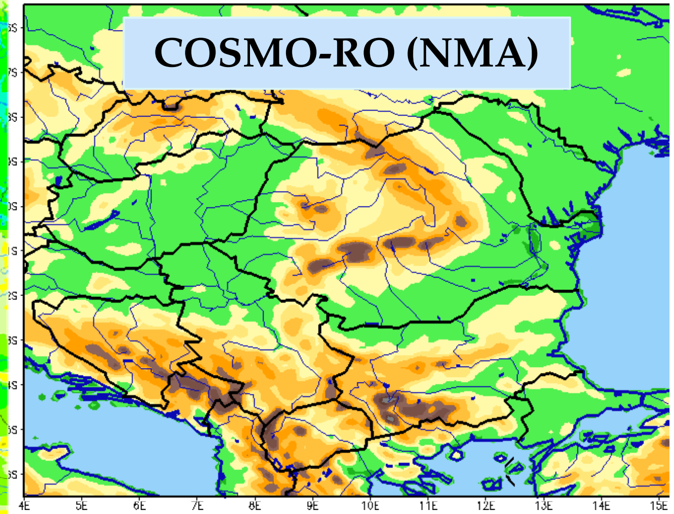
In the framework of COSMO verification activities, statistical scores extracted from CVS (common verification suite) or other packages, are presented for all COSMO countries with the use of a common graphic package.

See the guidelines of the verification reports (pdf, since Oct 2014)

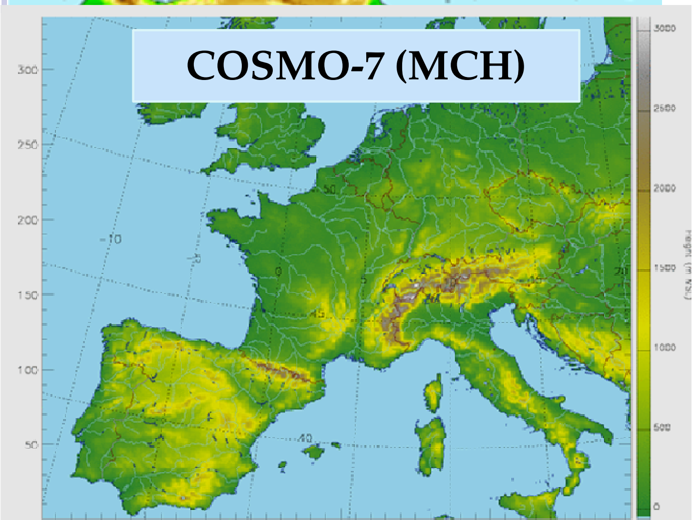
year	Dec-Jan-Feb	Mar-Apr-May	Jun-Jul-Aug	Sep-Oct-Nov
2015	get pdf	N/A	N/A	N/A
2014	get pdf	get pdf	get pdf	get pdf
2013	get pdf	get pdf	get pdf	get pdf
2012	get pdf	get pdf	get pdf	get pdf
2011	N/A	N/A	get pdf (for both seasons)	

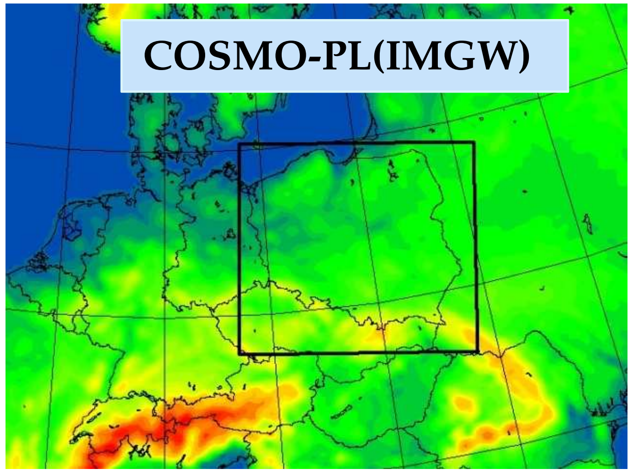

COSMO-EU (DWD)

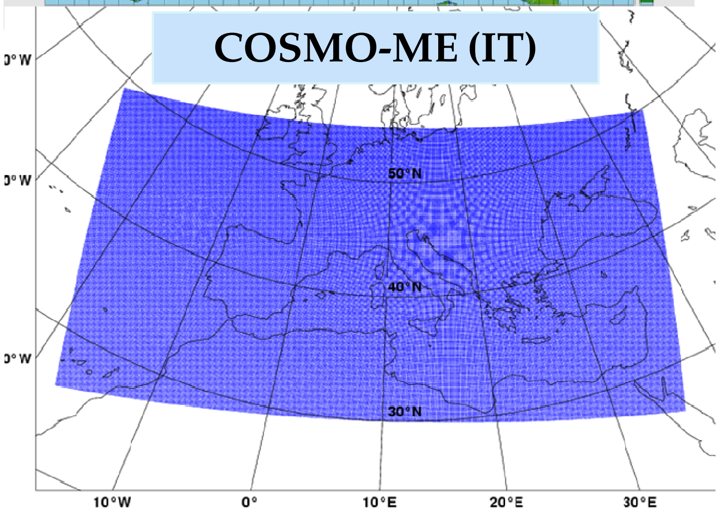

COSMO-Ru7
COSMO-RU7 (RHM)

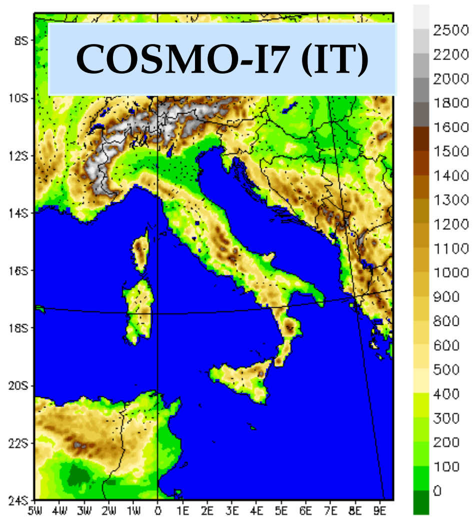

COSMO-RO (NMA)

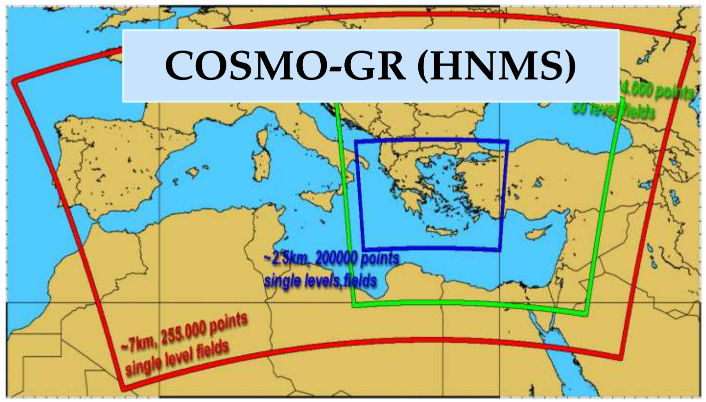
THE MODELS


COSMO-7 (MCH)

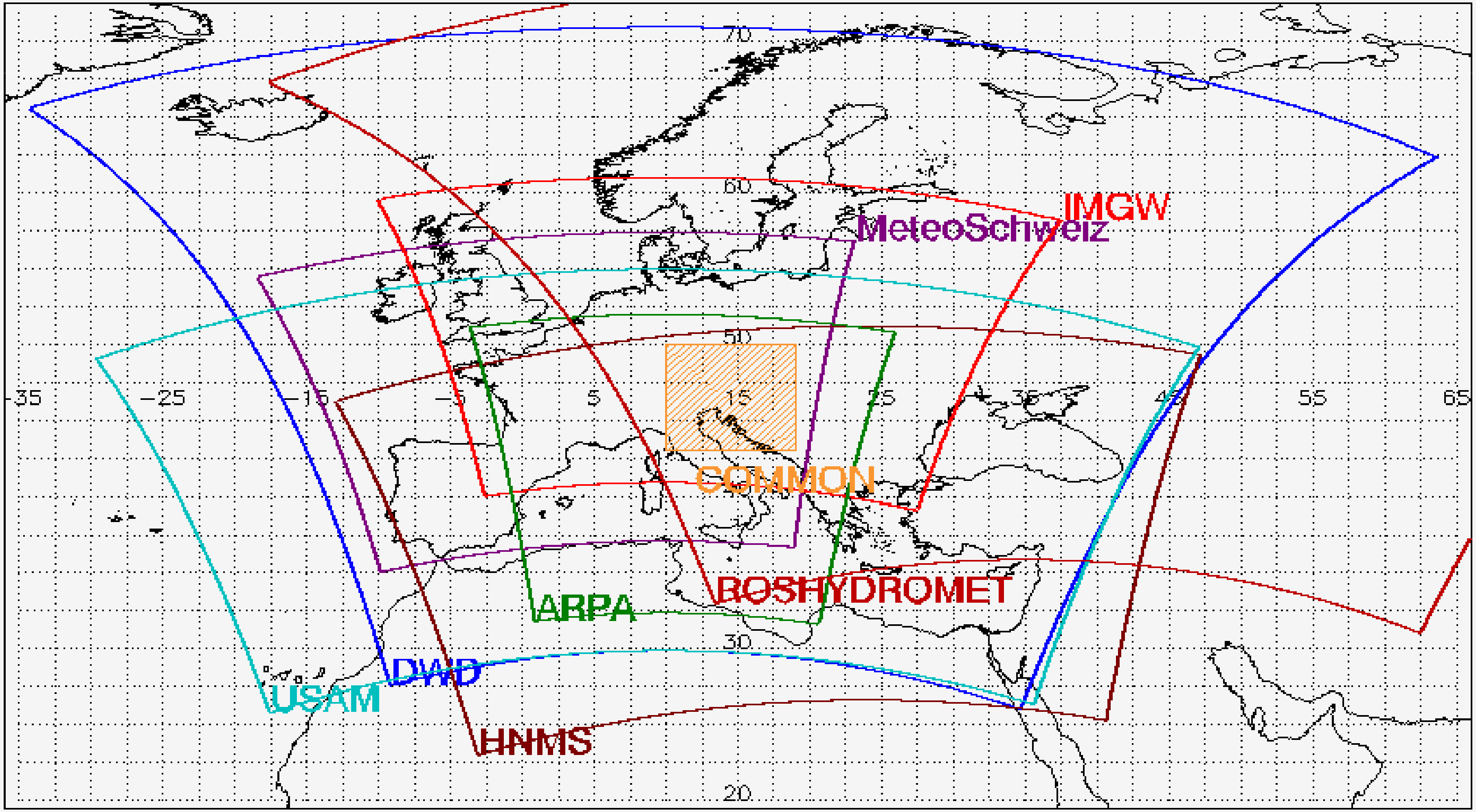

COSMO-PL(IMGW)


COSMO-ME (IT)
50°N
40°N
30°N
10°W
0°
10°E
20°E
30°E


COSMO-I7 (IT)
2500
2200
2000
1800
1600
1500
1400
1300
1200
1100
1000
900
800
600
500
400
300
200
100
0


COSMO-GR (HNMS)
~7km, 255.000 points
single level fields

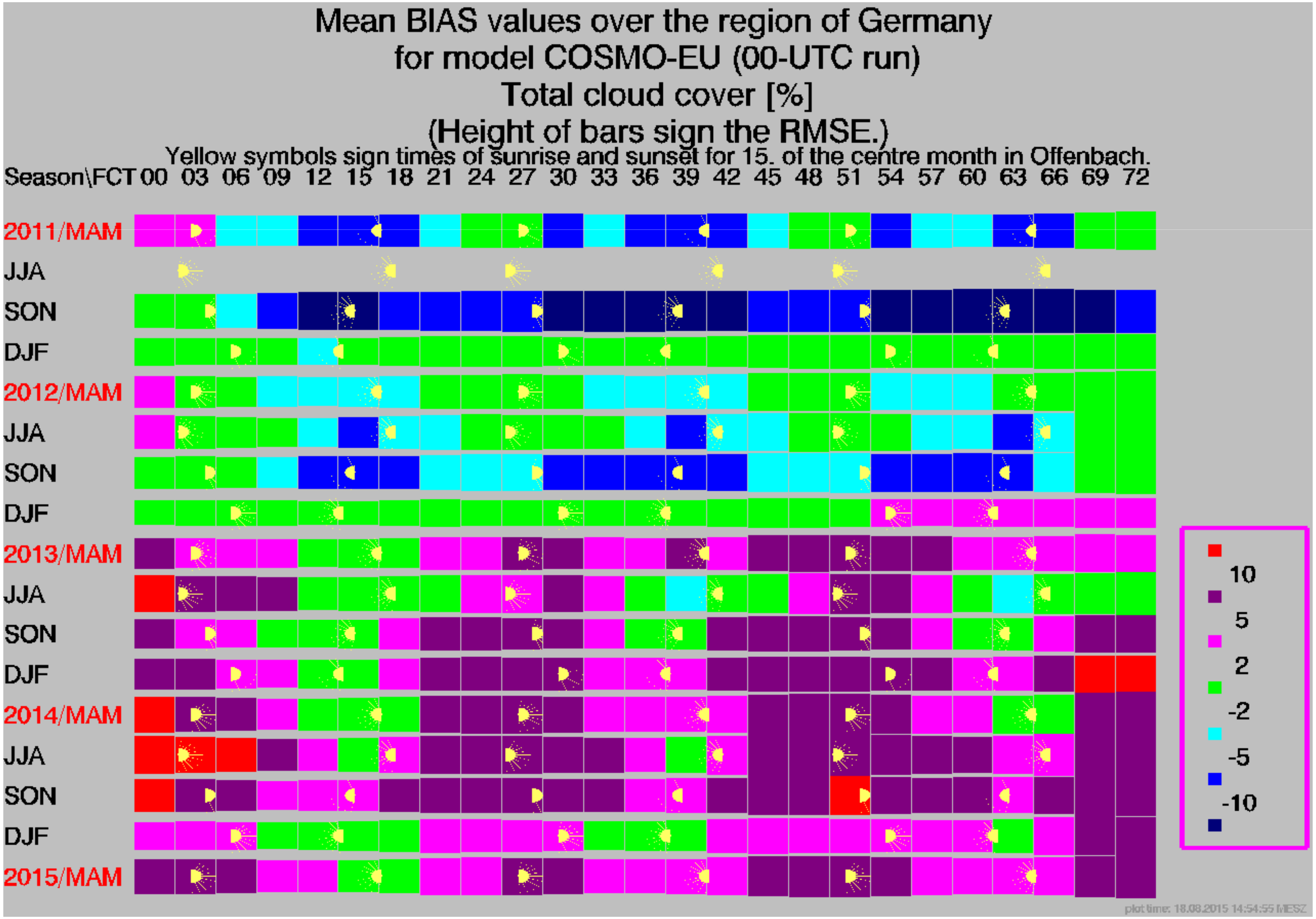
Standard Verification on Common Area



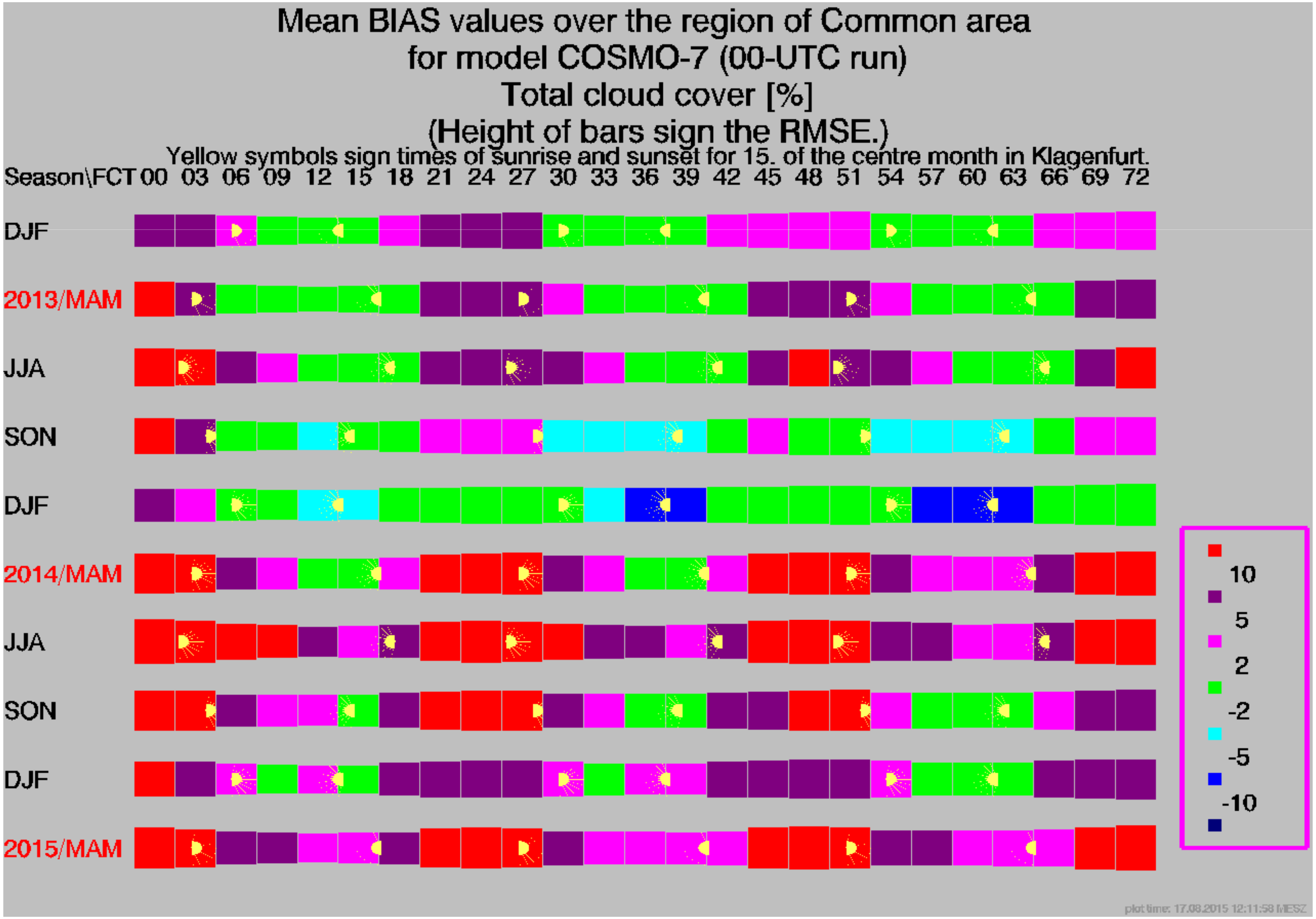
Based on Common Plots, one can.....

- Capture more important verification results on seasonal basis
  - **What are common forecast errors for different model setups?**
  - **What are differerent forecast errors for different model setups?**
  - Trends of errors
- Long term trends in verification results for surface weather elements
  - General trend
    - **The COSMO-Index – trend and single event ranges**
    - **and its components**
- Special consideration on quantitative precipitation forecast
    - **Observation data base: SYNOP**
    - **Observation data base: high resolution networks**
    - **Observation data base: radar data**

Common error behaviour
Total cloud cover COSMOEU (Germany)



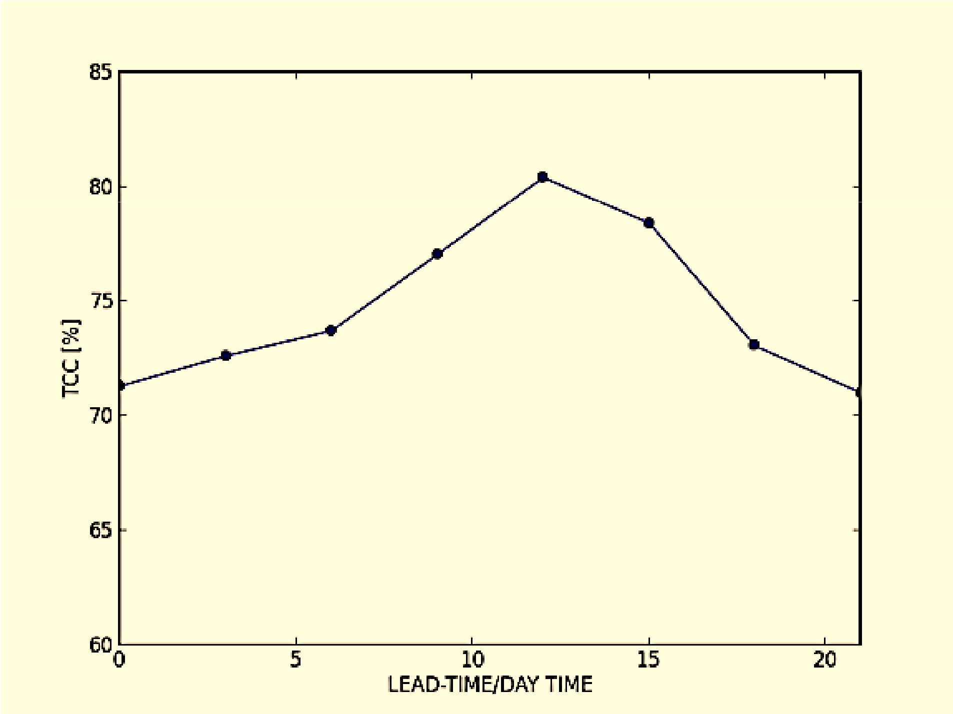
Common error behaviour
Total cloud cover COSMO7 (Meteoswiss)



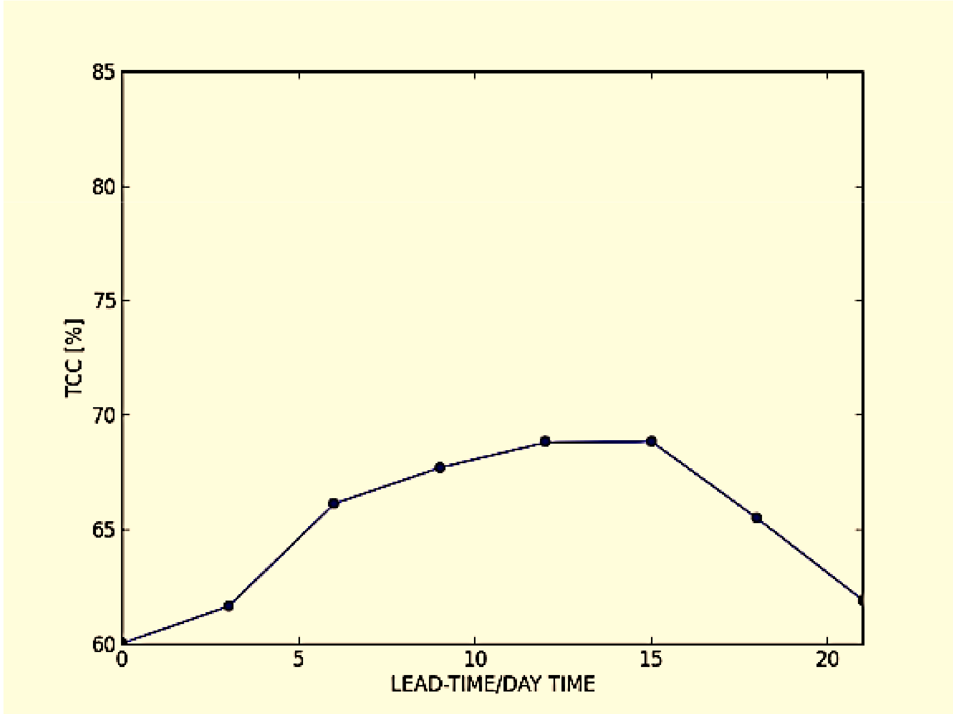
CONSORTIUM FOR SMALL SCALE MODELING COSMO

Diurnal Cycle in TCC
05/2012 – 05/2013

MSG/SEVIRI CM



SYNOP TCC



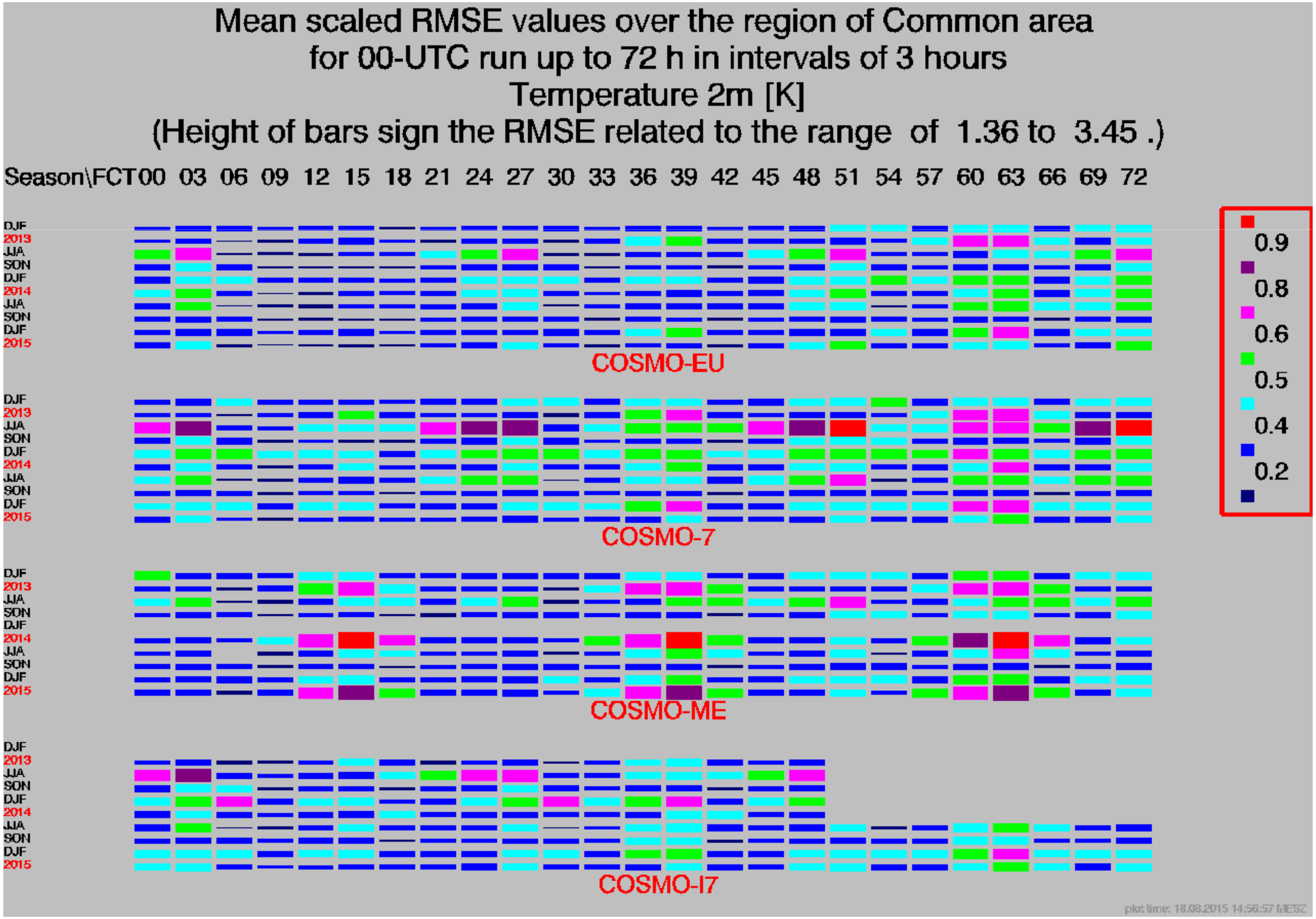
Mean difference: ~+10%

Mean Bias now: ~ +10%

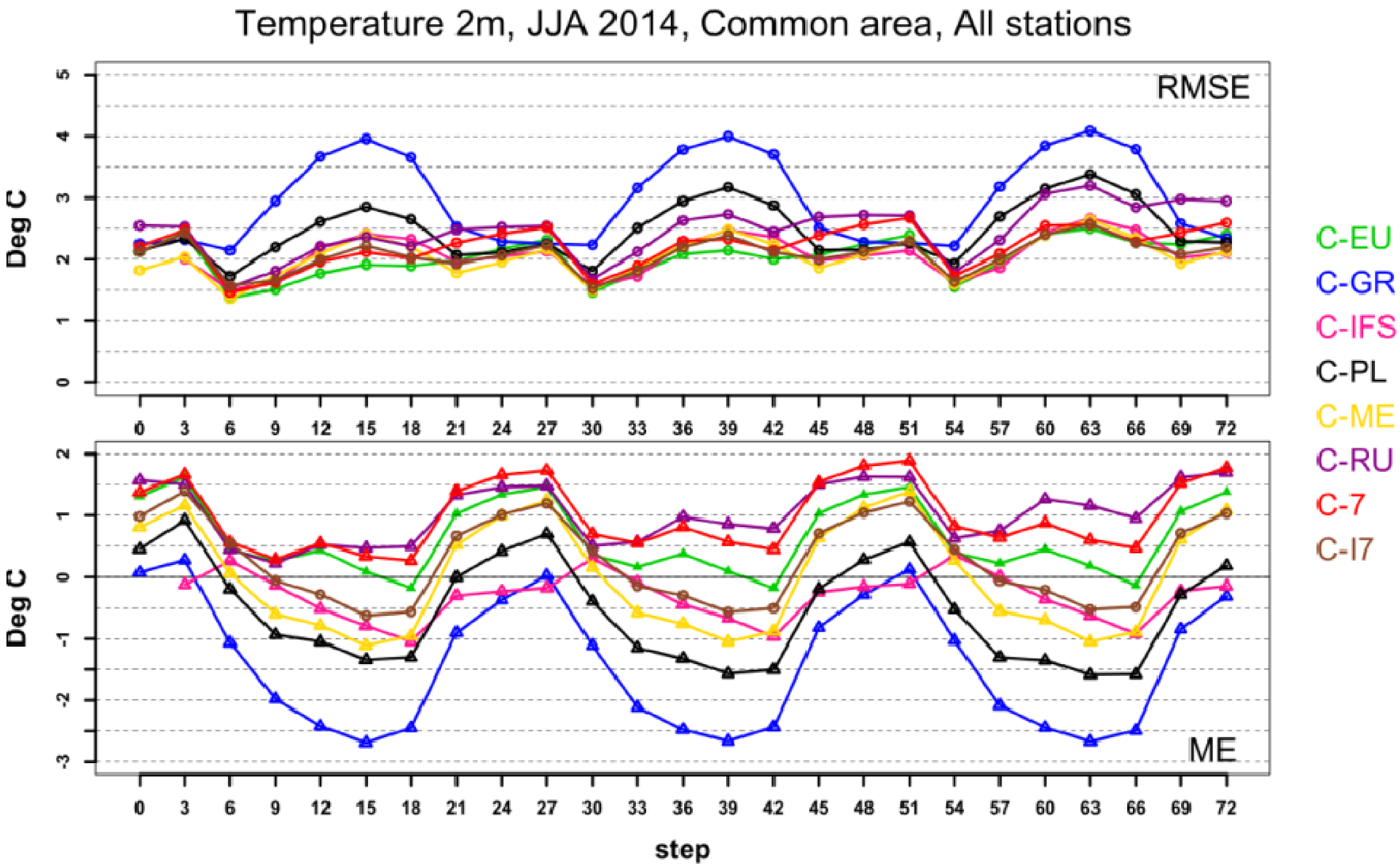
Is there a real BIAS in the TCC forecast?

By courtesy of Felix Fundel

Common error behaviour
Mean scaled RMSE for T2m



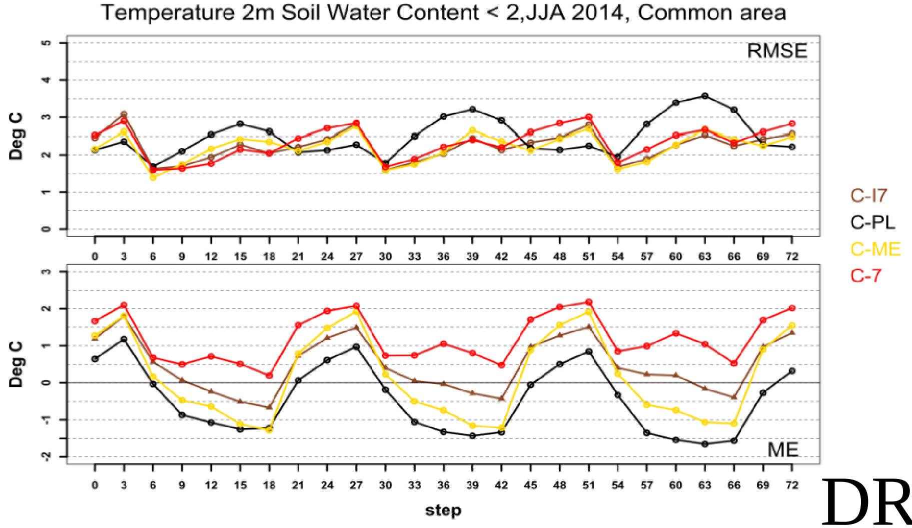
Common error behaviour
T2m JJA 2014



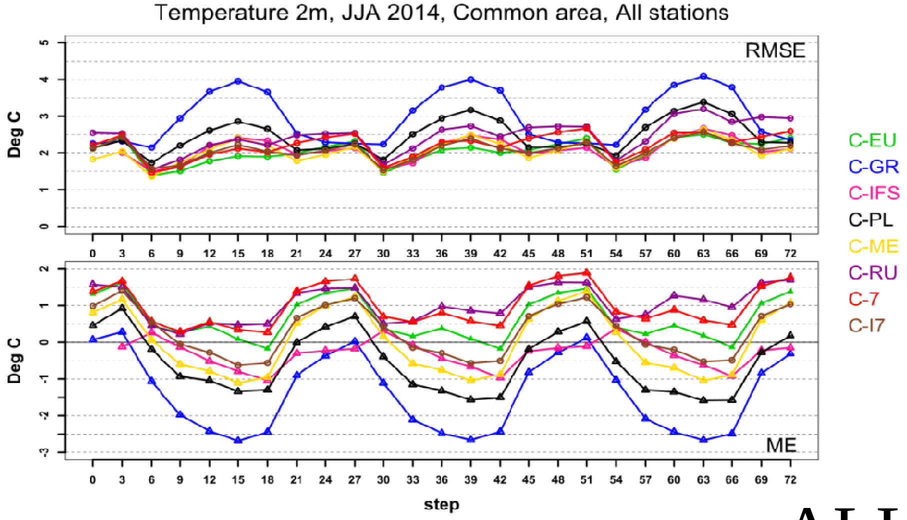
Conditional Verification on Common Area (All seasons)

- **2mT verification with the following criteria (1 condition):**
  - Soil Water Content >= 4 (moist condition) (condition based on forecasts)
  - Soil Water Content < 2 (dry condition) (condition based on forecasts)
- **Wind speed verification with the following criteria (1 condition):**
  - Roughness length >= 1m (rough cases) (condition based on forecasts)
  - Roughness length < 0.2 m (smooth cases) (condition based on forecasts)

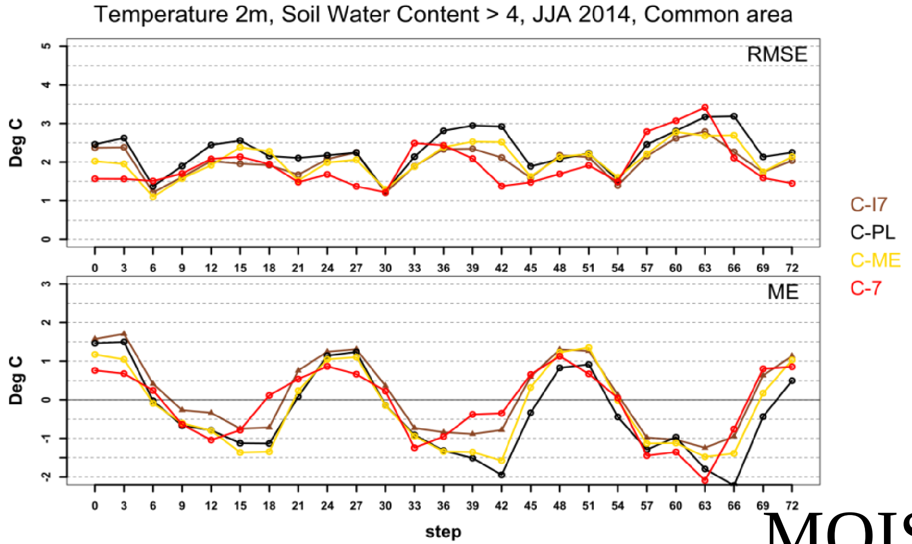
Conditional verification for T2m JJA 2014 depending on soil moisture



DRY



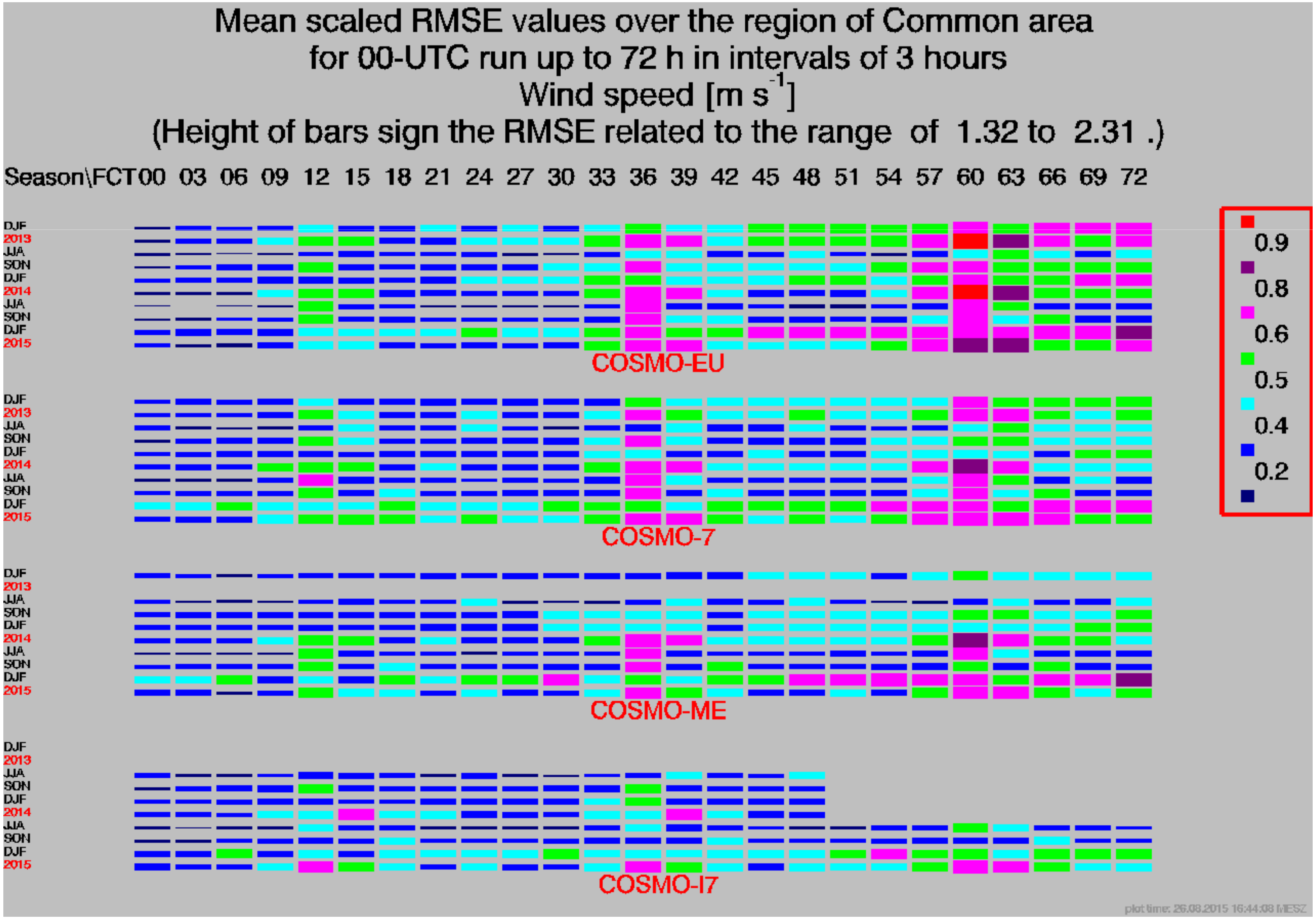
ALL



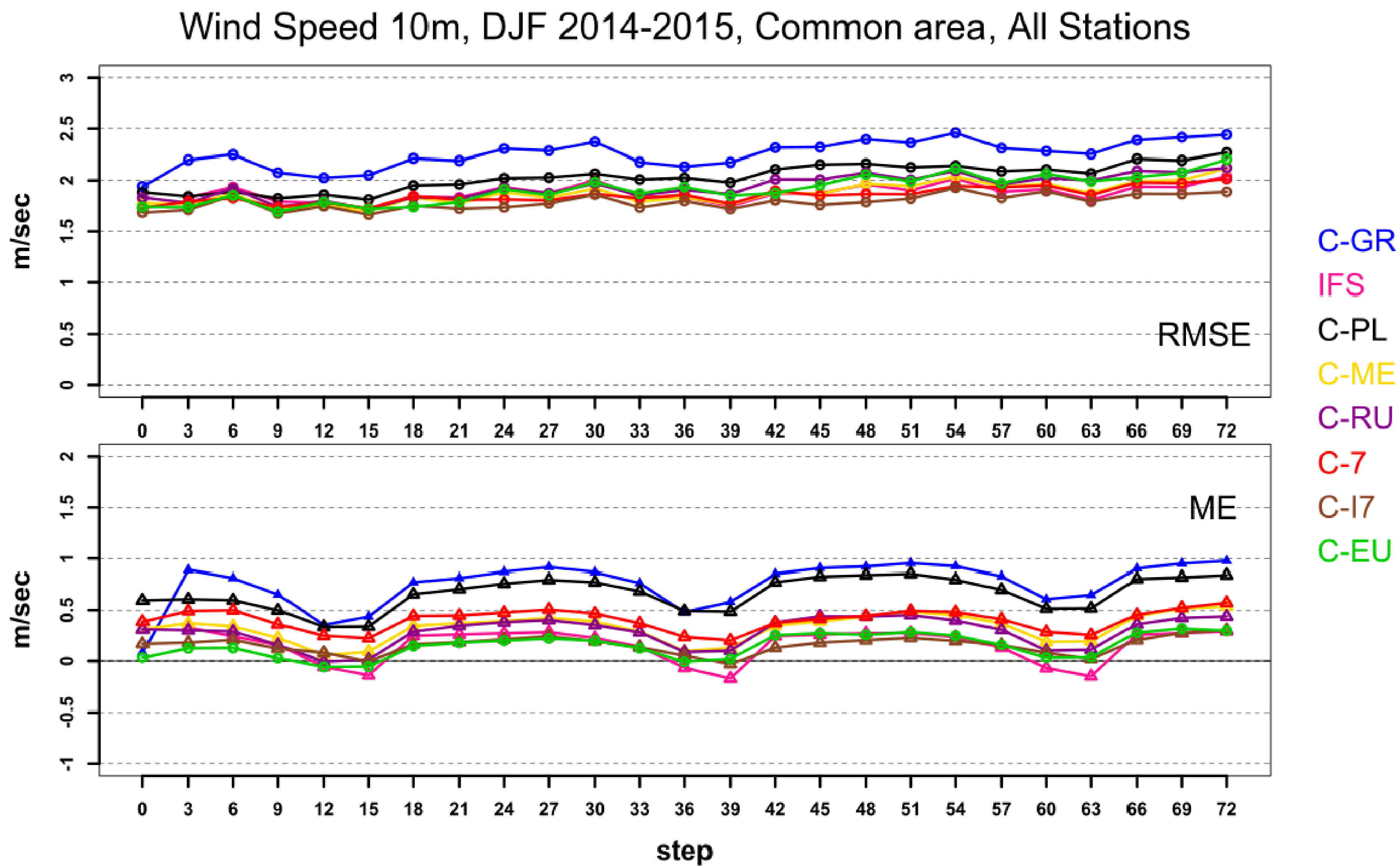
MOIST

DRY and MOIST have
Similar diurnal variation, except for C-7, MOIST models grouped together .
Are there any common trends ?

Common error behaviour
Mean scaled RMSE for WS 10m



Different error behaviour
WS 10m DJF 2014/2015

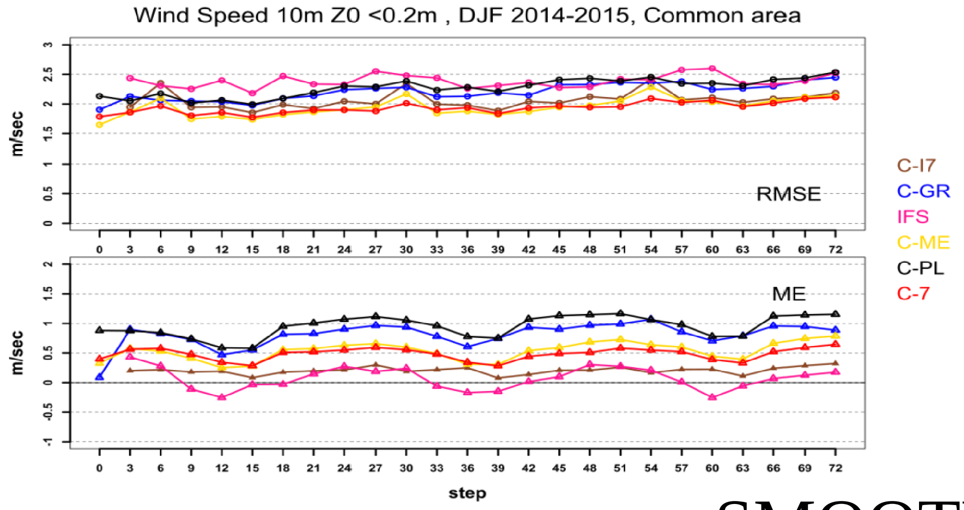


Conditional Verification on Common Area (All seasons)

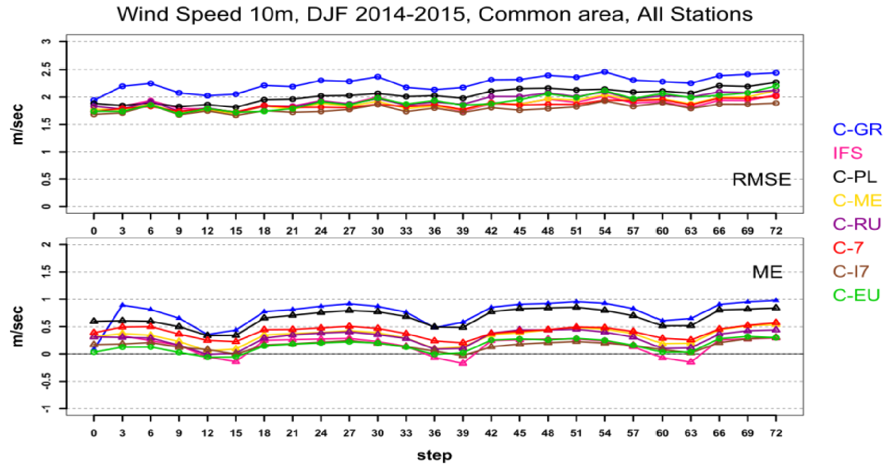
- **<u>2mT verification with the following criteria (1 condition)</u>:**
  - ***Soil Water Content >= 4 (moist condition) (condition based on forecasts)***
  - ***Soil Water Content < 2 (dry condition) (condition based on forecasts)***

- **<u>Wind speed verification with the following criteria (1 condition)</u>:**
  - ***Roughness length >= 1m (rough cases) (condition based on forecasts)***
  - ***Roughness length < 0.2 m (smooth cases) (condition based on forecasts)***

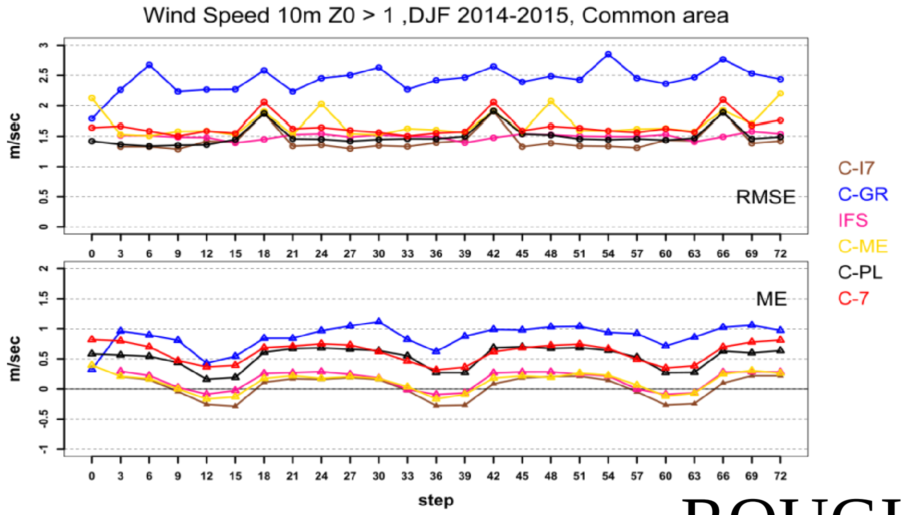
Conditional verification for WS 10m DJF 2014/2015 depending on roughness length



SMOOTH

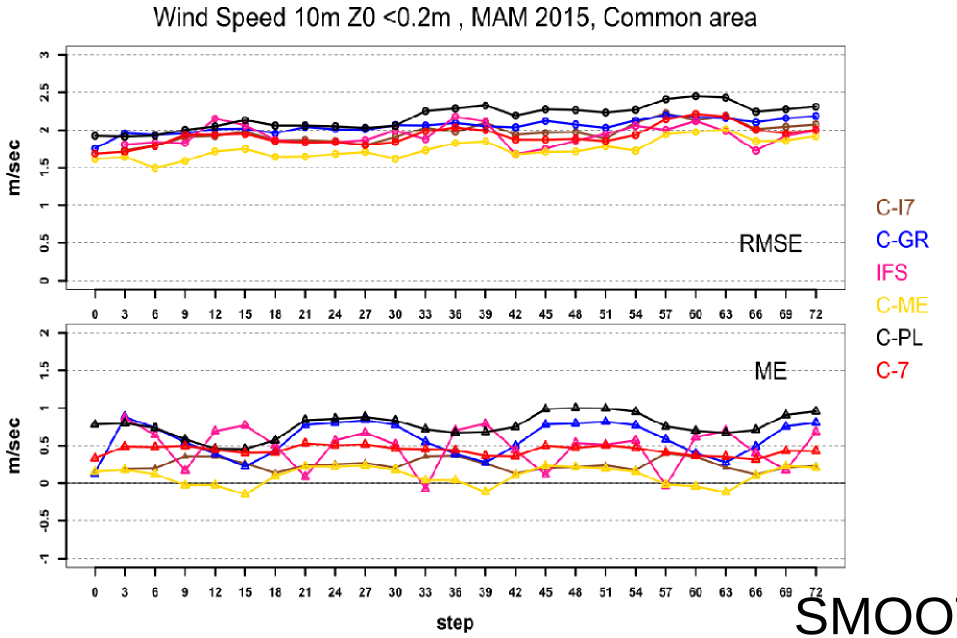


ALL

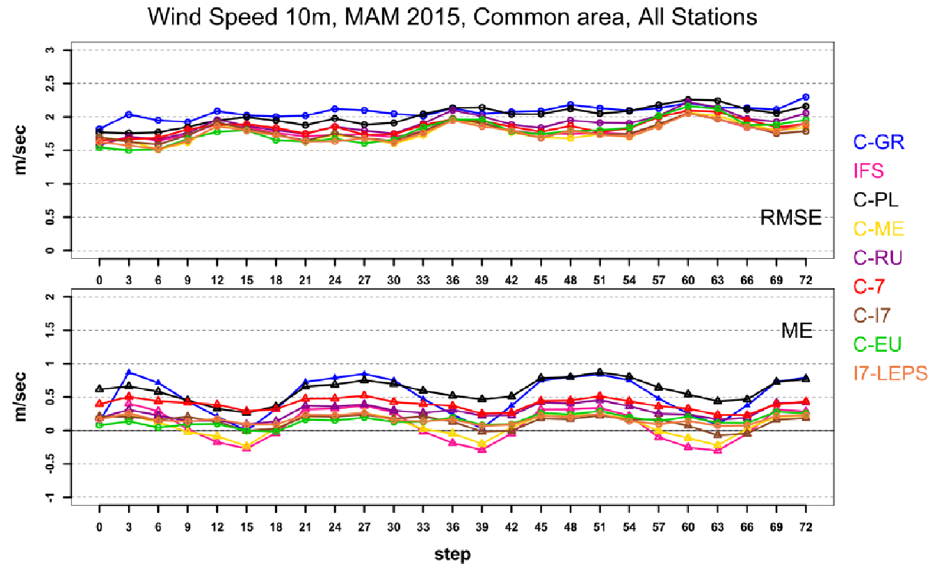


ROUGH

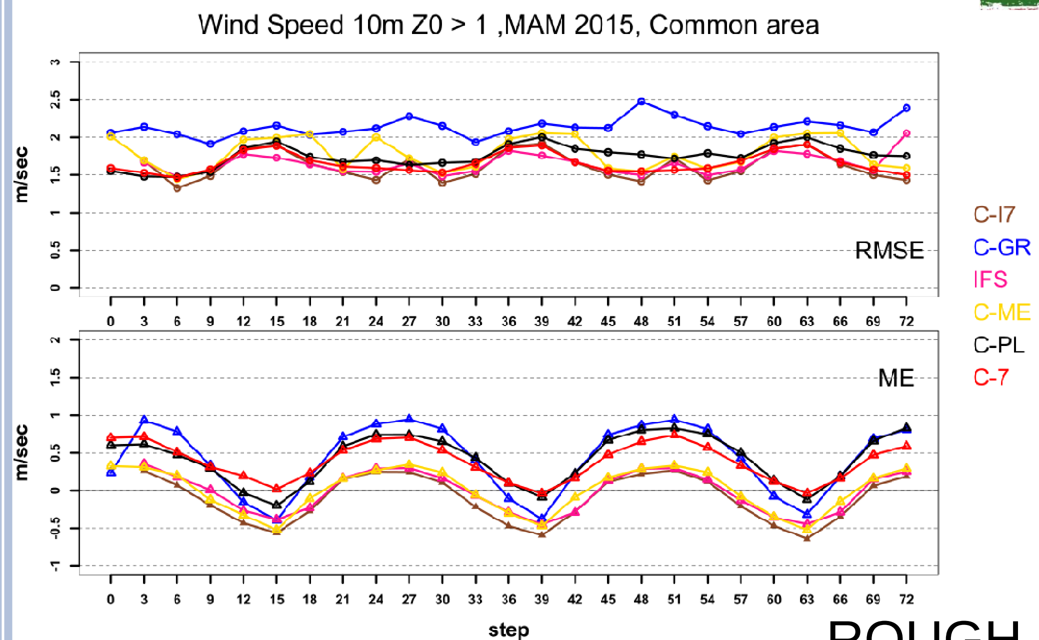
ROUGH and SMOOTH ME cycles are now similar to ALL



SMOOTH



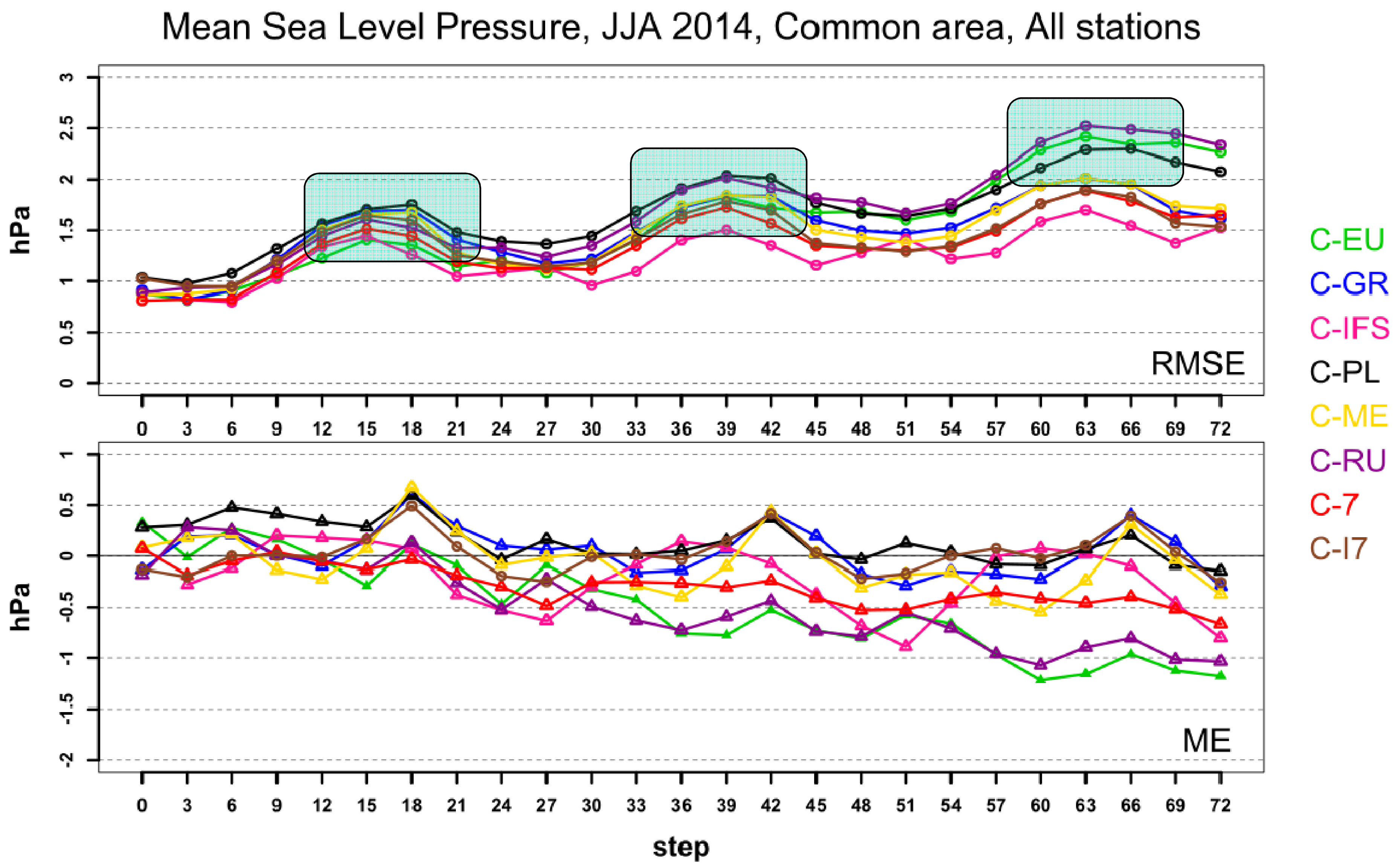
ALL



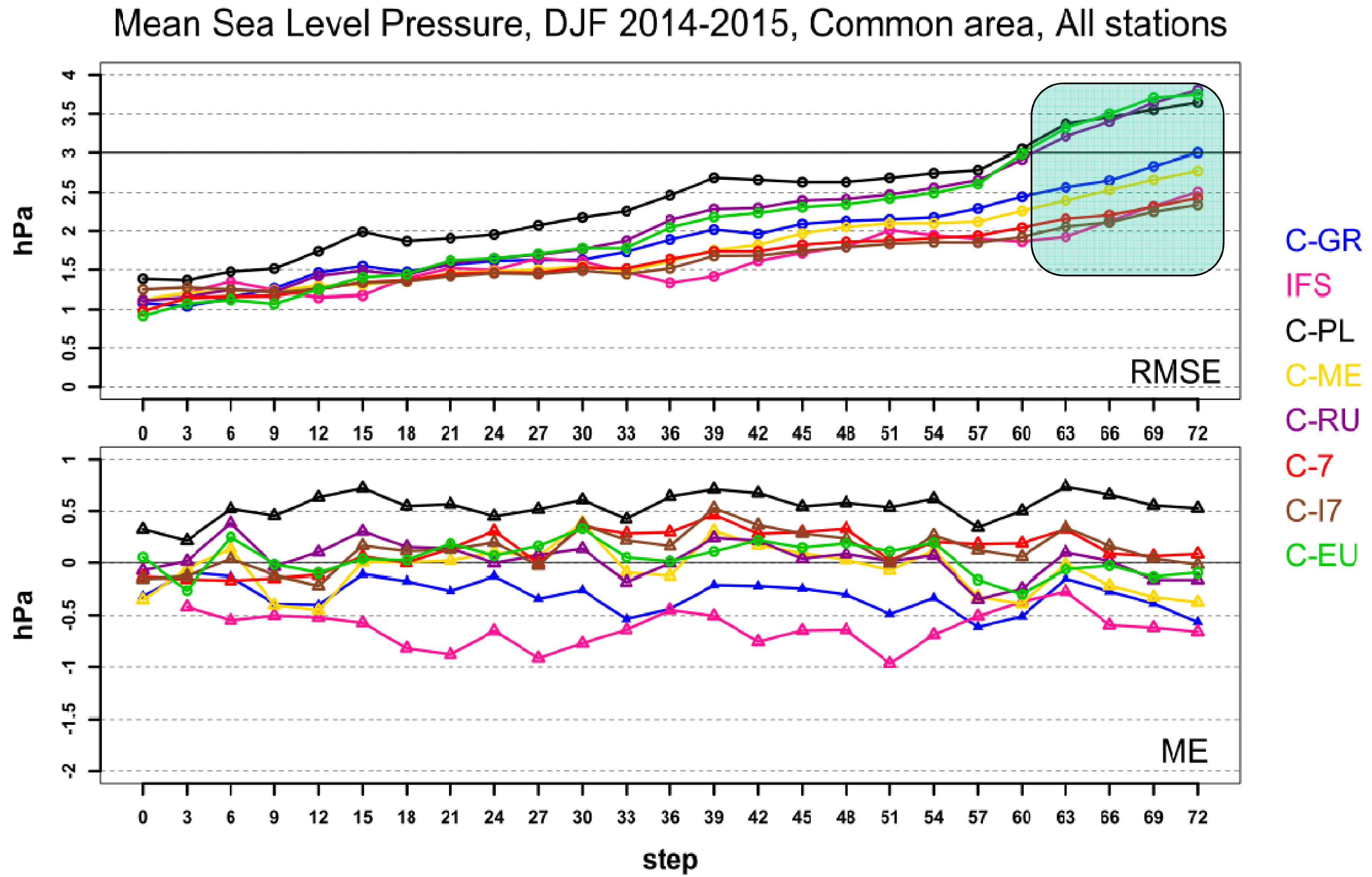
ROUGH

ROUGH ME cycle is sharper.
SMOOTH weaker cycle

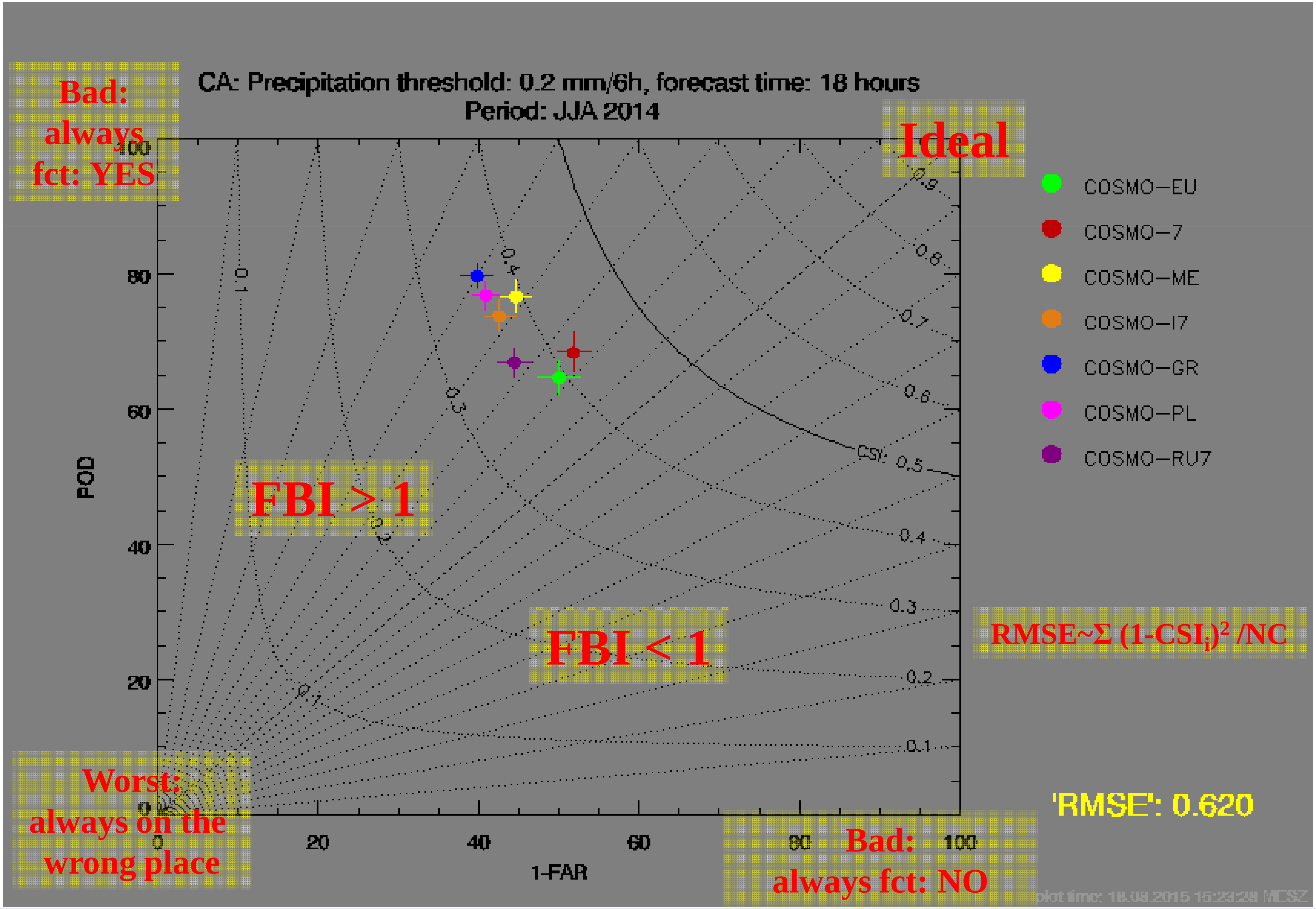
Different error behaviour
MSLP   JJA 2014



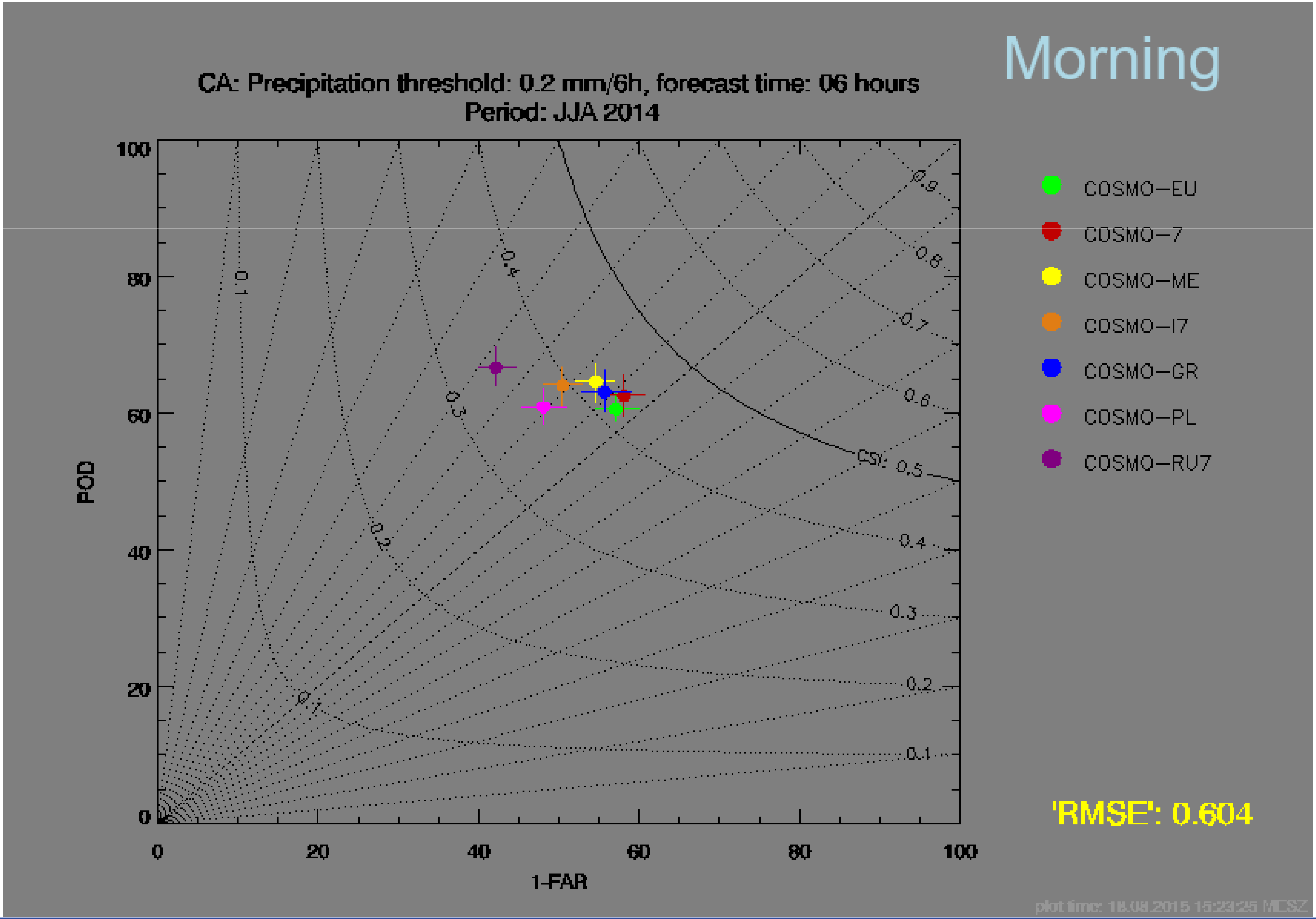
Different error behaviour
MSLP DJF 2014/2015



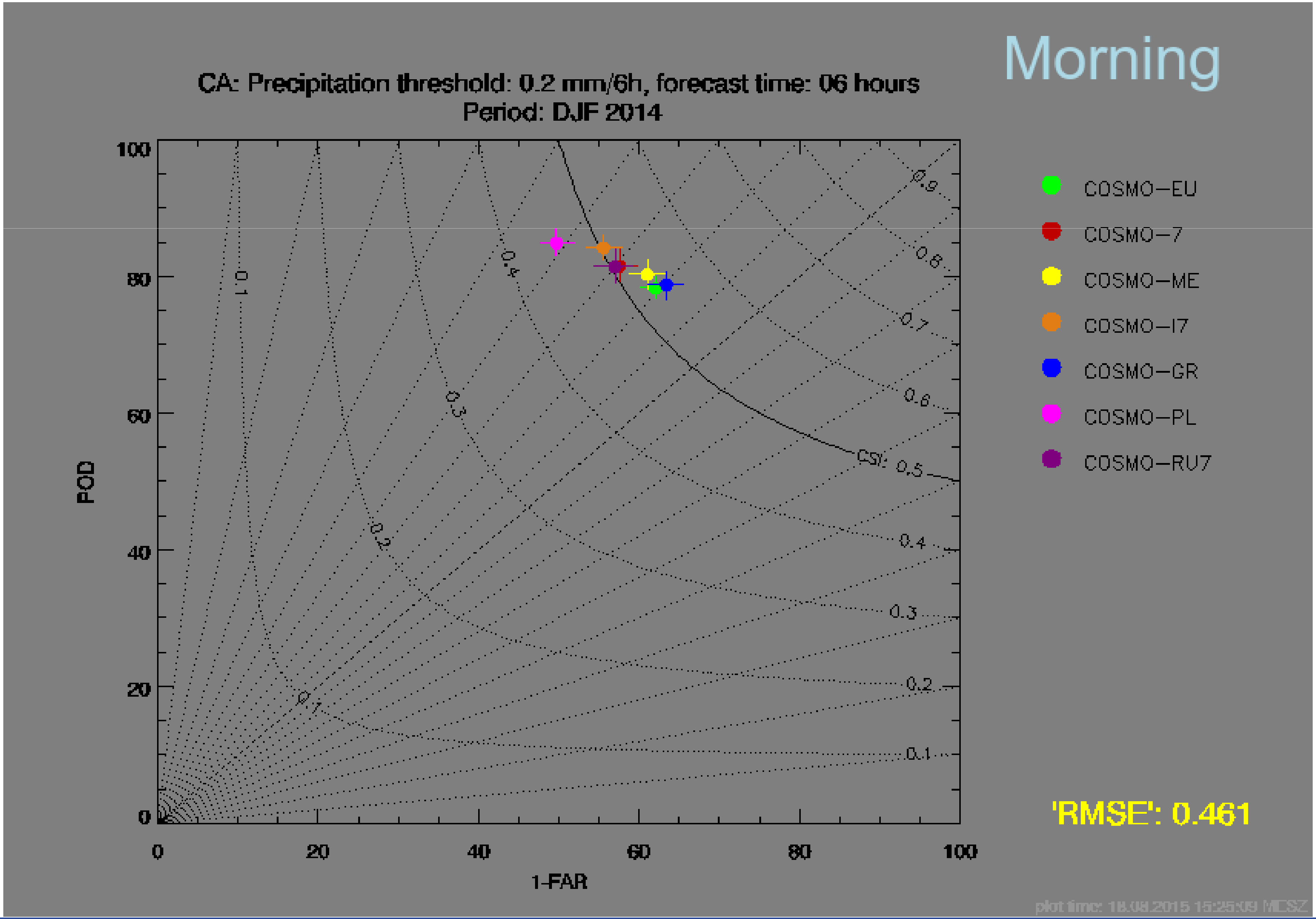
Performance diagram precipitation (example)



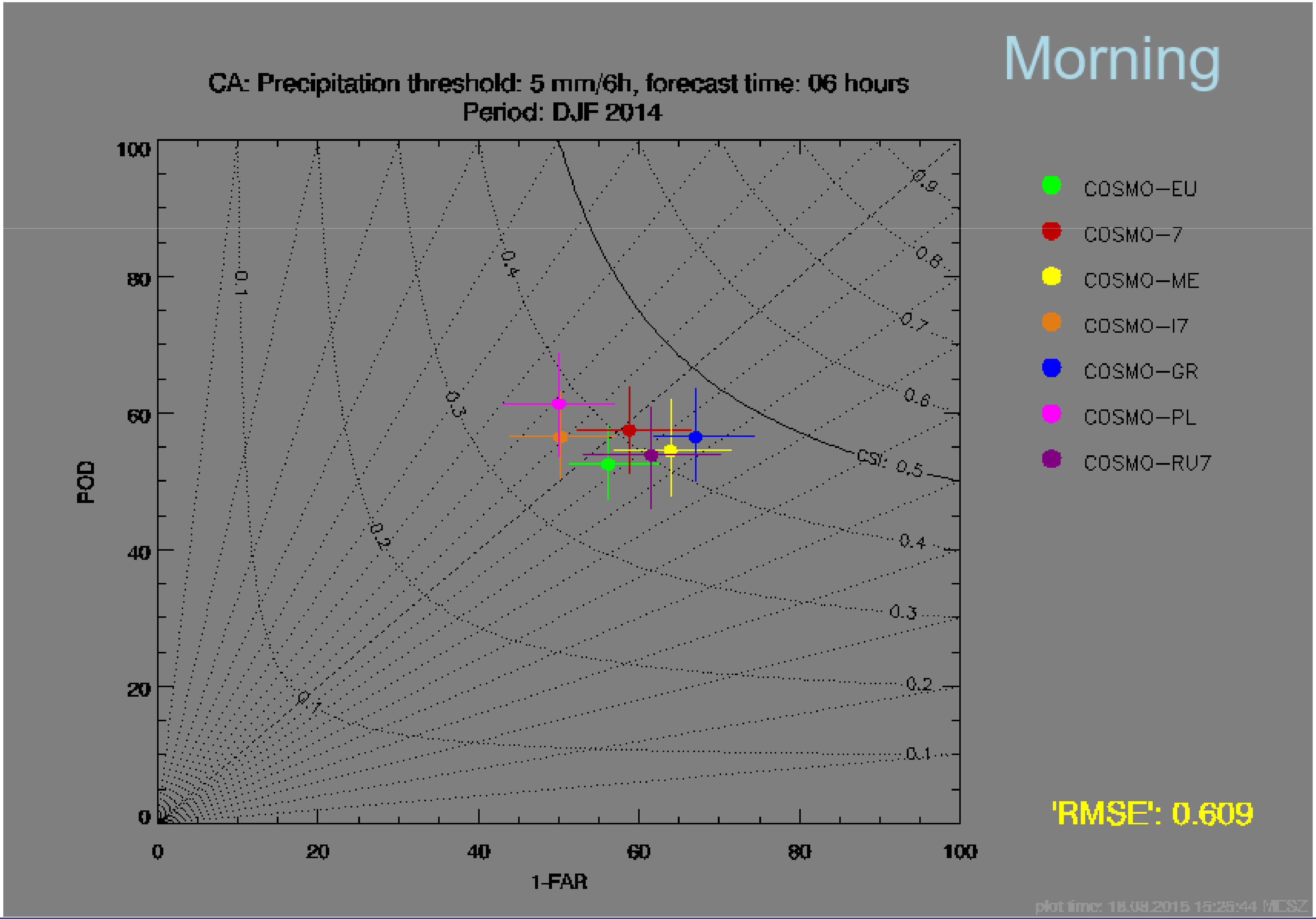
Performance diagram precipitation JJA 2014 0.2mm/6h



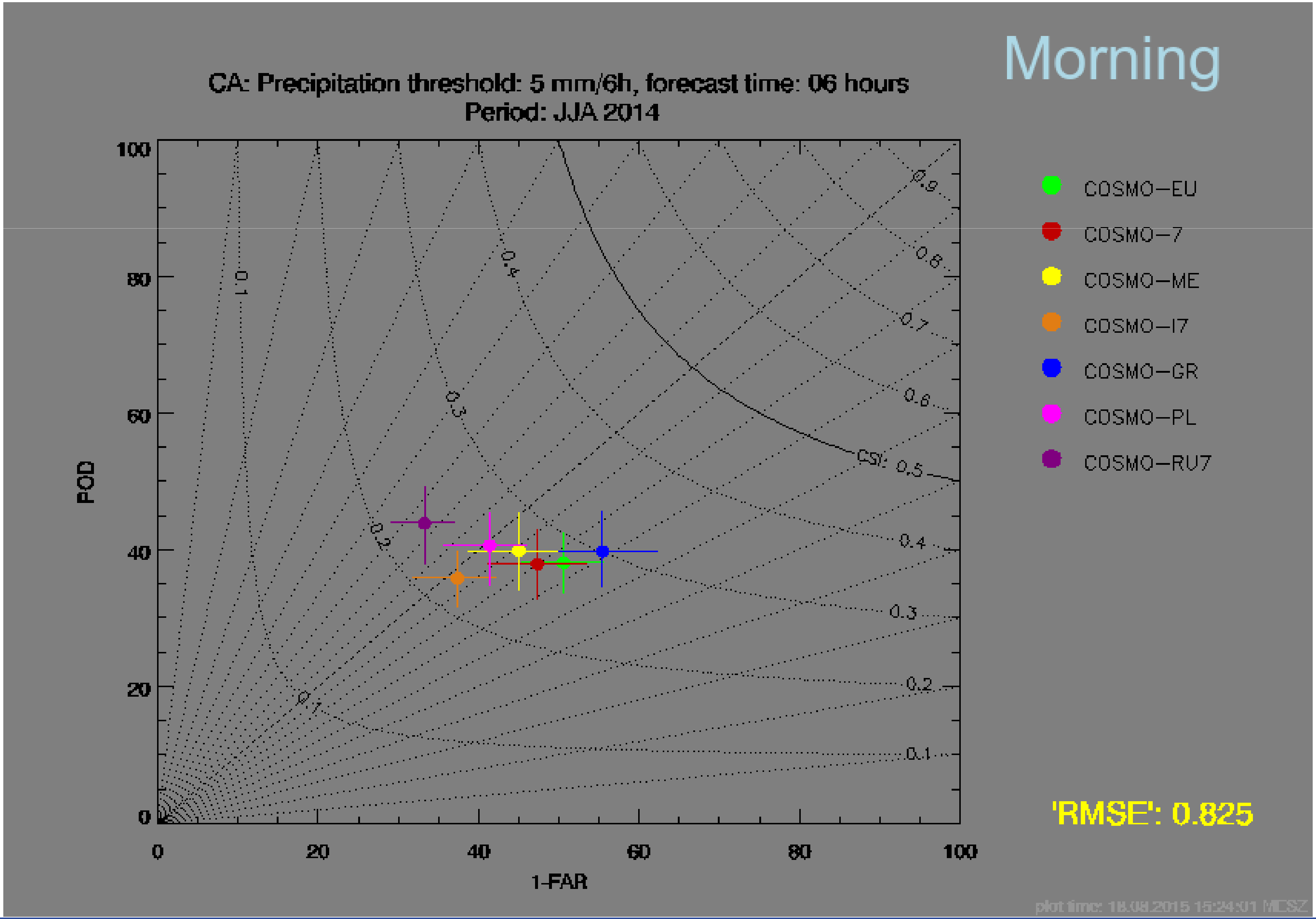
Performance diagram precipitation DJF 2014/2015 0.2mm/6h



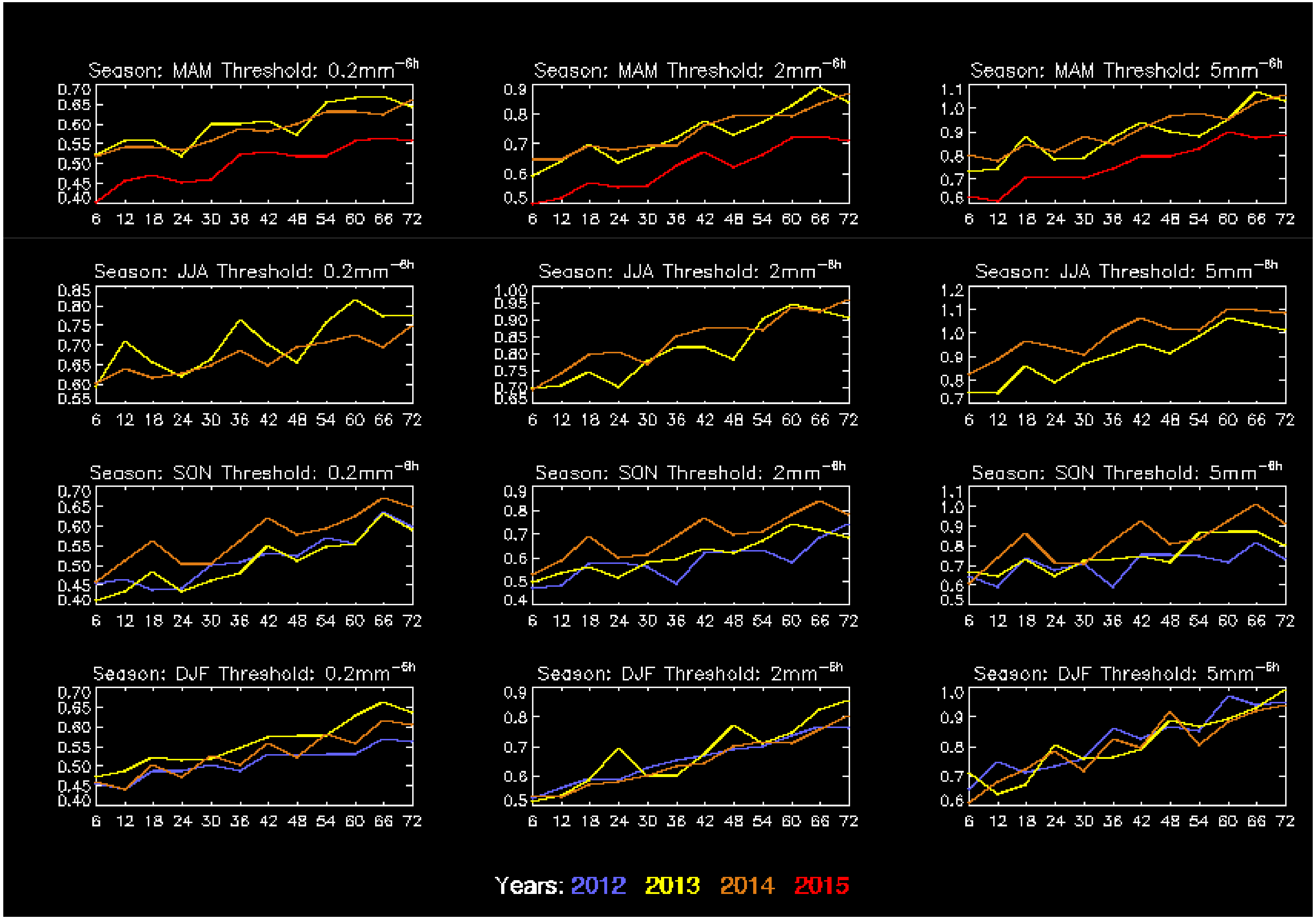
Performance diagram precipitation
DJF 2014/2015 5 mm/6h



Performance diagram precipitation
JJA 2014 5 mm/6h



‚RMSE‘ precipitation depending on forecast time, threshold and season: Is there any trend?



Goal of Verification activities

CONSORTIUM FOR SMALL SCALE MODELING

- Assess COSMO model performance and trends through organized and methodical way: Important to follow the configuration differences for each model version and check the complete history (at least for the last three years) of model changes

- to identify the relative skill of high resolution model implementations (and the scales or applications that are more useful)

- give feedback to modelers:
  - ✓Contribute to COSMO model development
  - ✓Improve the understanding of forecast errors
  - ✓Identify possible sources of errors in COSMO
- give hints for a better understanding of COSMO model to the users (e.g. Forecasters in the daily operational activity) and contribute to guidelines on how to use COSMO forecasts)

➢ **Is there any need to revive the SRNWP expert team on Diagnostics with more interconsortia activitiies?**

➢**Can the area of verification <u>tools development/spatial method applications</u> be one of these area of focus?**

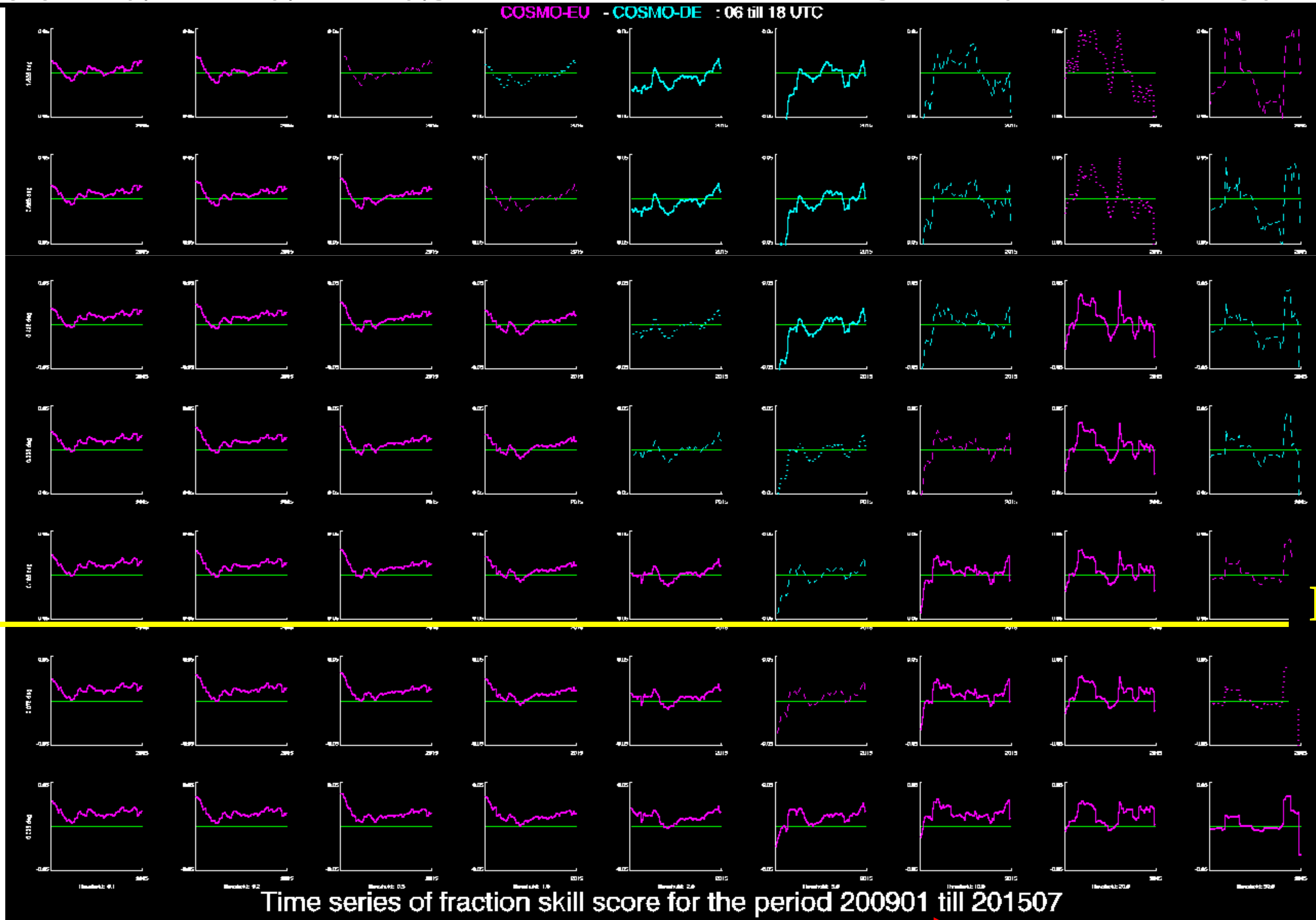
Contributing Scientists

Ulrich Damrath, DWD
Ulrich Pflunger, DWD
Pirmin Kaufmann, MCH
Xavier Lapillonne, MCH
Angela Celozzi, USAM
Antonio Vocino, USAM
Flora Gofa, HNMS
Dimitra Boucouvala, HNMS
Joanna Linkowska, IMGW
Rodica Dumitrache, NMA
Amalia Iriza, NMA
Anastasia Bundel, RHM
Alexander Kirsanov, RHM
Maria Stefania Tesini, ARPA-SIM
Elena Oberto, ARPA-PT
Naima Vela, ARPA-PT
Alon Stivelman, IMS

Comparison of COSMO-EU to COSMO-DE – FSS (differences)

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Threshold 0.1 0.2 0.5 1 2 5 10 20 50



1.625
0.825
0.425
0.225
0.125
0.075
0.025

Window size

Mesh width of COSMO-EU

Precipitation amount