Verification and Case studies

Overview of activities

Flora Gofa

Strategy on verification tools

WG5 recommendations Report

Definition of Common Verification Activities (use of a CVS can be advantageous)

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 - ***CVS and possible AVT requirement***

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. Verification software that allows for CVS applications is necessary – ***CVS requirement***

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. The application of these is closely related to existence of the necessary verification tools – ***CVS and AVT requirement***

The necessity to have Common Verification Software and Tools (CVS+) for Common Verification Activities was underlined by the vast majority (but to be followed by scientific and technical standards).

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

CVS

Scenario 1: VERSUS+VAST

Scenario 2: VERSUS+VAST with gradual migration to other CVS (DWD-SW) (2 year period)

Scenario 3: VERSUS+VAST+DWD SW (parallel use)

Prerequisites/Implementation

- Successful completion of all **VERSUS2** Tasks and improvement of software performance
- 3 month intensive evaluation period after **VERSUS2** Task work completion
- Extension of **VERSUS** Maintenance Plan (I,III)
- **DWD SW** as CVS candidate, conditions and rules of exchange
- **DWD SW** test phase within WG5 (available in 2016)

AVT

Scenario 1: Unrestricted exchange of tools to a common repository

Scenario 2: Development or adaptation of tools or verification code as requested from WG5. This can be realized as part of COSMO PP/PT(s)

For AVT both scenarios can coexist

Prerequisites/Implementation

- New methods once evaluated and requested by users, exchangeable in common repository (definition)
- **Scenario I**, no SCM rules apply or any obligation for technical support, I/O format (only documentation)
- **Scenario II**, AVT follows certain standards (I/O common interface, scripting for adaptation to existing systems, installation on major linux distributions).
- Application and adjustment on common I/O interface of already available software packages can be performed as subTasks of COSMO PPs (e.g. SpatialVx in INSPECT)

Prerequisites and Implementation Plan for CVS

✓ **Successful completion** of all **VERSUS** Tasks included in phase7 of VERSUS2 project.

✓ **Improvement in VERSUS software performance** as described in the optimization processes to be included in the preceding software version (final test results available September 2015).

✓Three (3) month **intensive evaluation period** from WG5 members after the completion of all Tasks of VERSUS2 project (September-November 2015).

✓**Extension of VERSUS Maintenance Plan** in the case of Scenario I, III. Additional minor required developments, to be included in the Maintenance Plan.

✓ For **DWD software as CVS candidate, conditions and rules should be defined for being exchangeable.** As it is strongly dependant on MEC software for input Feedback Files preparation, the adaptation as COSMO CVS will be relied on the decision of each service to move towards this approach for verification applications. This can introduce delays for WG5 activities and a parallel CVS usage should be considered for a considerable time period (MEC software just released, DWD software available end of 2016).

✓An evaluation testing phase of DWD-SW within WG5 is recommended.

CVS

Scenario 1: VERSUS+VAST

Scenario 2: VERSUS+VAST with gradual migration to other CVS (DWD-SW) (2 year period)

Scenario 3: VERSUS+VAST+DWD SW (parallel use)

Prerequisites/Implementation

- Successful completion of all **VERSUS2** Tasks and improvement of software performance
- 3 month intensive evaluation period after **VERSUS2** Task work completion
- Extension of **VERSUS** Maintenance Plan (I,III)
- **DWD SW** as CVS candidate, conditions and rules of exchange
- **DWD SW** test phase within WG5 (available in 2016)

AVT

Scenario 1: Unrestricted exchange of tools to a common repository

Scenario 2: Development or adaptation of tools or verification code as requested from WG5. This can be realized as part of COSMO PP/PT(s)

For AVT both scenarios can coexist

Prerequisites/Implementation

- New methods once evaluated and requested by users, exchangeable in common repository (definition)
- **Scenario I**, no SCM rules apply or any obligation for technical support, I/O format (only documentation)
- **Scenario II**, AVT follows certain standards (I/O common interface, scripting for adaptation to existing systems, installation on major linux distributions).
- Application and adjustment on common I/O interface of already available software packages can be performed as subTasks of COSMO PPs (e.g. SpatialVx in INSPECT)

Prerequisites and Implementation Plan:

✓**Scenario I and II can be complementary**

✓AVT products following new methods will be presented during WG5 annual meetings and once evaluated and requested by the users, necessary software will be exchanged in a common repository.

✓**For scenario I, no SCM rules apply to AVT exchange** neither any obligation for technical support. Only commitment is for AVT to be accompanied by adequate documentation.

✓Any software can become part of AVT, developed or adapted in the framework of a PP or PT if it follows certain standards (I/O common interface, scripting for adaptation to existing systems, installation on major linux distributions - Scenario II). This can include any application or adjustment of already available software packages (e.g. SpatialVx applications in INSPECT Task 2).

✓Necessary a **common repository definition** for all software included in the AVT concept.

I/O data format

Following the experience of other consortia, as well as estimating the effort that is usually devoted to the adjustment of verification tools to local DB and I/O formats, any future verification strategy should be closely related to this aspect.

• **Uniform I/O format standards for exchangeable data** (required from verification tools) should be agreed at least **for the AVT** (CVS can have already a non flexible architecture: e.g. VERSUS).

• Any required (by verification software) **data format adaptation** will have to be performed **externally**, separating in this way the **main focus of the verification tools that is the correct application of statistical methods from the proprocessing of data or the graphical representation, that is the focus of other tools**.

• Fieldextra is the official COSMO software and its use is suggested for the postprocessing of any gridded forecast or observation field, while LIBSIM and other tools can be utilized for other forms of observations (e.g. BUFR).

• **AVT software** suggested in the scenario II, **should be followed by I/O data format standards**, making in this way easier the adaptation of each tool from most users and permitting the exchange of data for common experiments (e.g. Task2 INSPECT).

INSPECT: INtercomparison of SPatial vErification methods for COSMO Terrain

Priority Project

Started:April 2015

Presentation by A.Bundel (to follow)

PP VERSUS2 Phase 7 – Final

Major Tasks
Grib2, Feedback Files implementation
Operation-ability issues in VERSUS use
Maintenance Plan

Presentation to follow (PL A.Celozzi)

Overview of activities

WG5

WG5

COSMO CONSORTIUM FOR SMALL-SCALE MODELING

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

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

Verification With Averaged Model Output

Problem

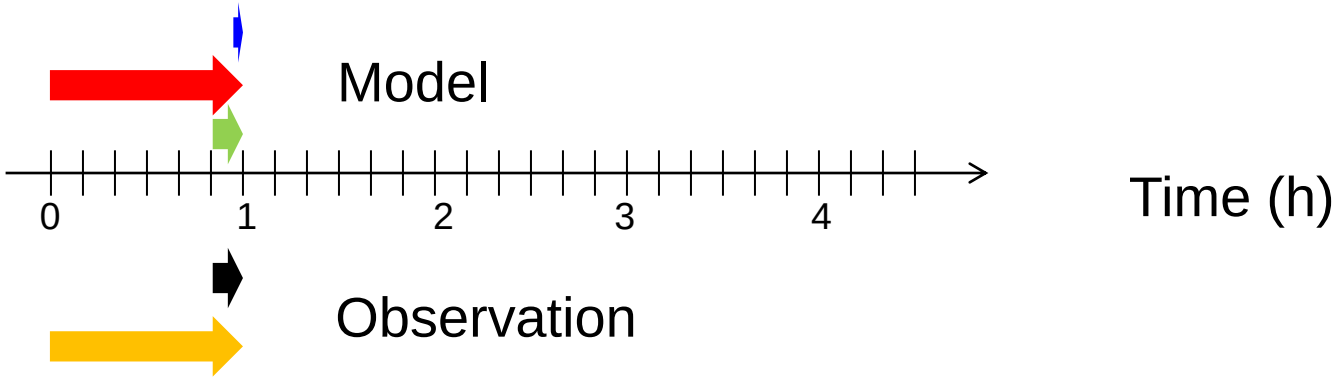
Initial problem:

- COSMO-1 shows at some stations larger spread of error (STDE) than COSMO-2
- Hypothesis: Higher variability in COSMO-1 could be the cause

General considerations:

- We have in fact a time-representativeness mismatch between model and observation:
  - SYNOP Observations are 10 min averages
  - Standard COSMO model output are instantaneous values
- How does this influence our verification results?

Averaging Intervals



SYNOP: 10 min Average

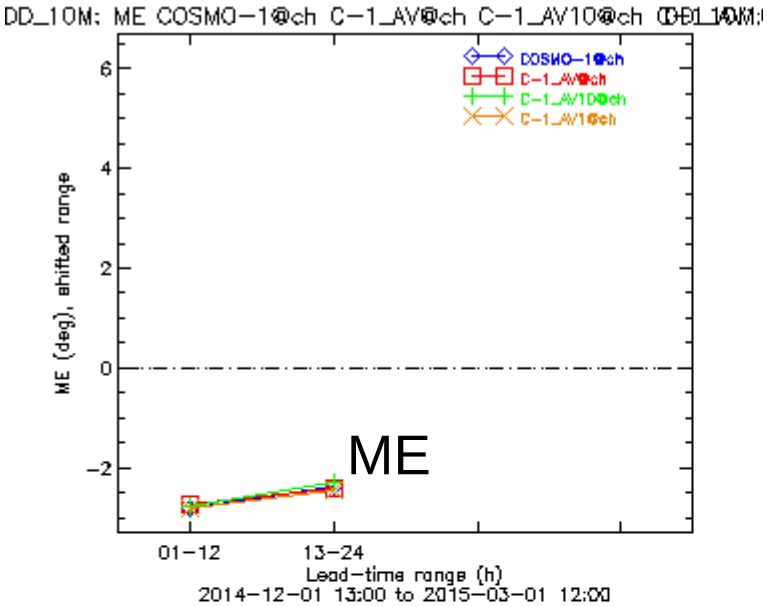
Model: last time step (COSMO standard output)

Model 1 h average over all COSMO time steps

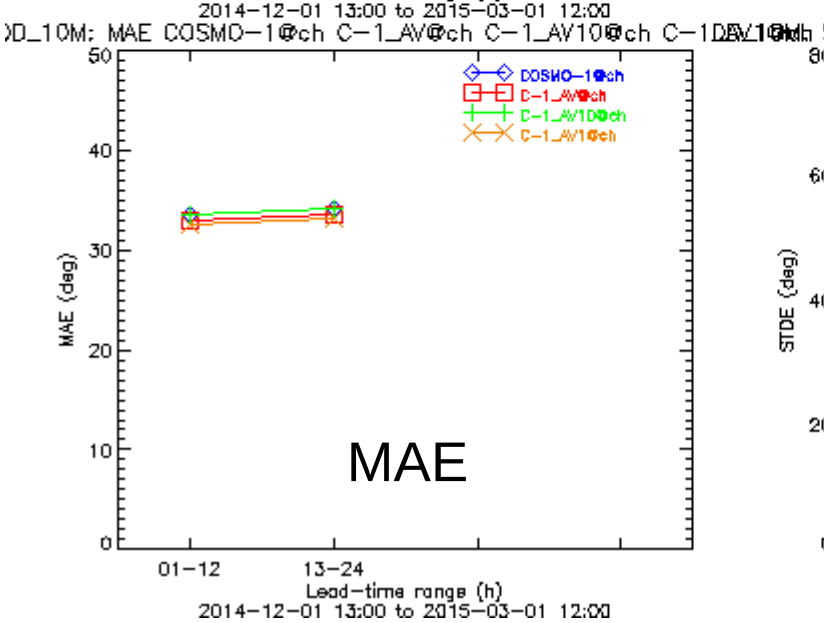
Model 10 min average over all COSMO time steps

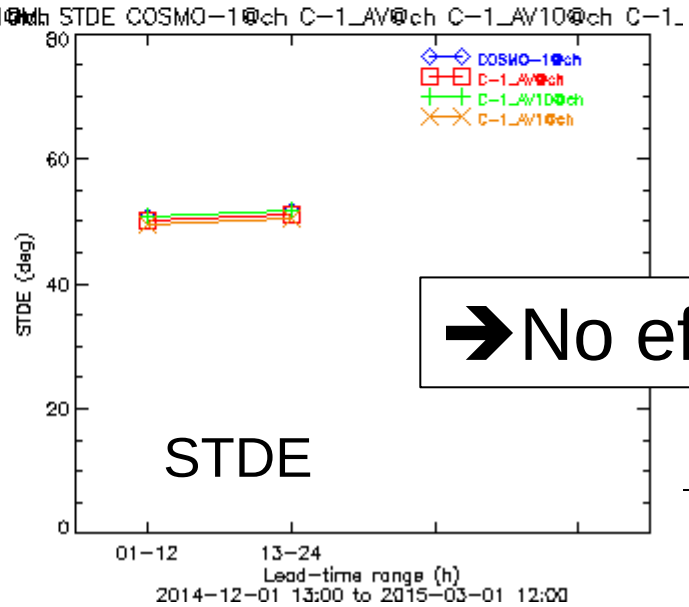
Model 1 h average compared with obs 1 h average

Wind Direction



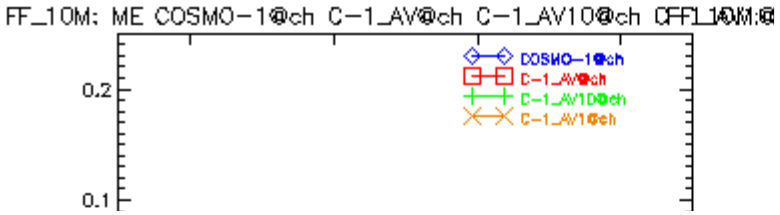
Model instantaneous
Model 1 h
Model 10 min
Model 1 h, Obs 1 h



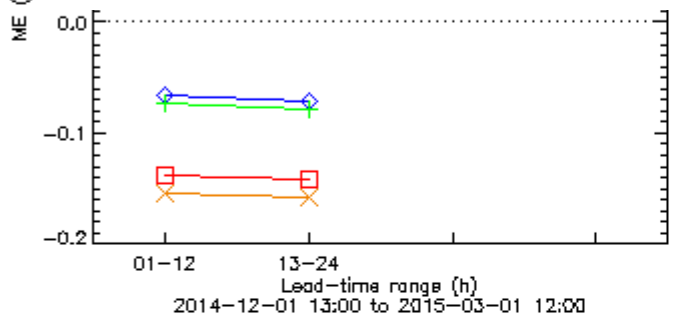


➔No effect!

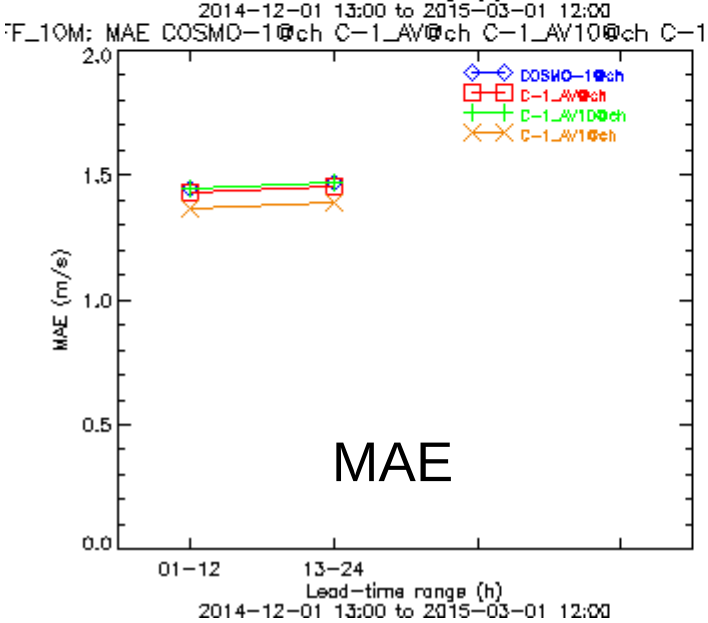
Wind Speed



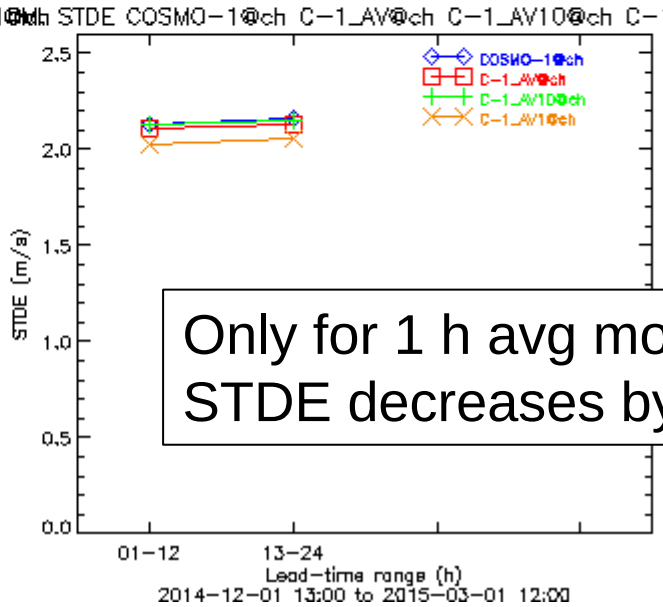
ME differs – unclear why



Model instantaneous
Model 1 h
Model 10 min
Model 1 h, Obs 1 h

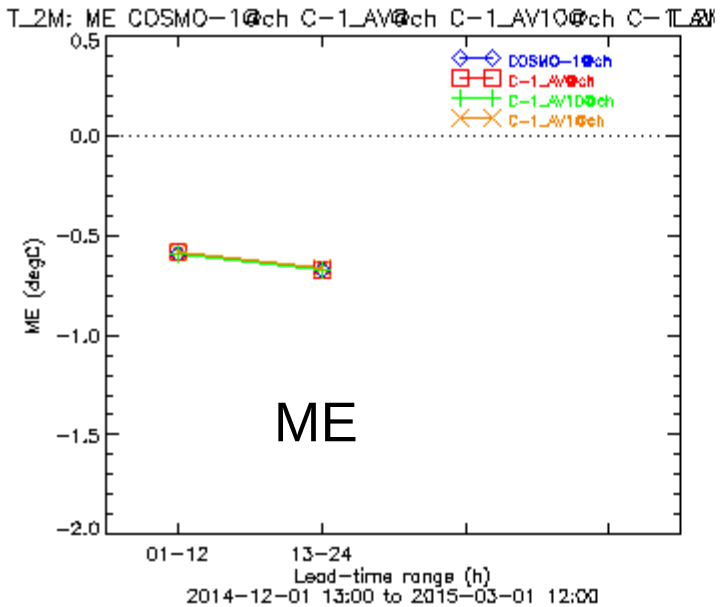


MAE



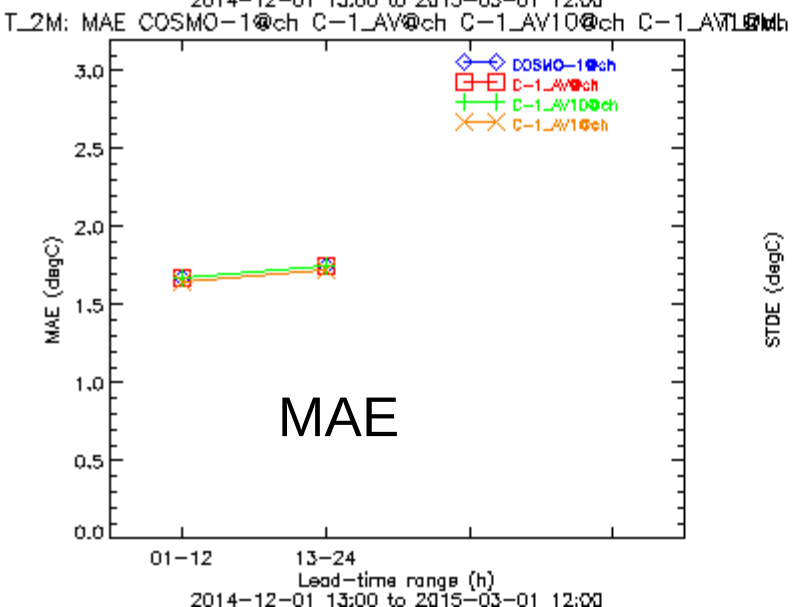
Only for 1 h avg mod/obs:
STDE decreases by 5%

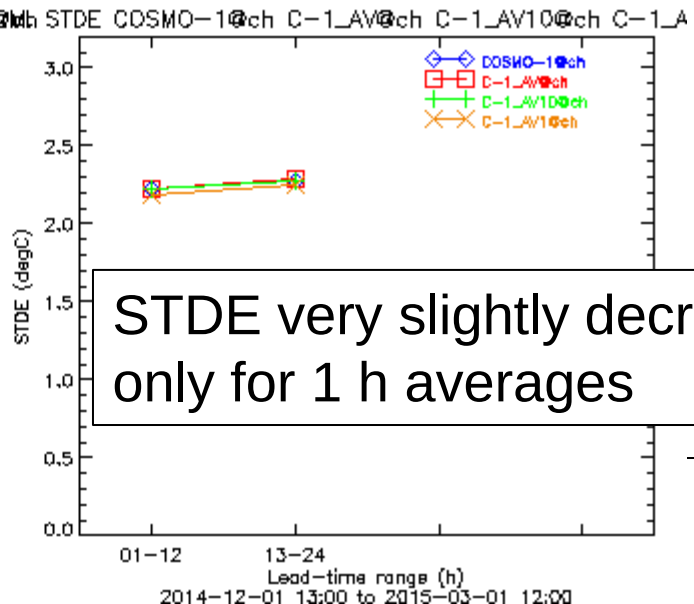
2 m Temperature



➔No effect

Model instantaneous
Model 1 h
Model 10 min
Model 1 h, Obs 1 h





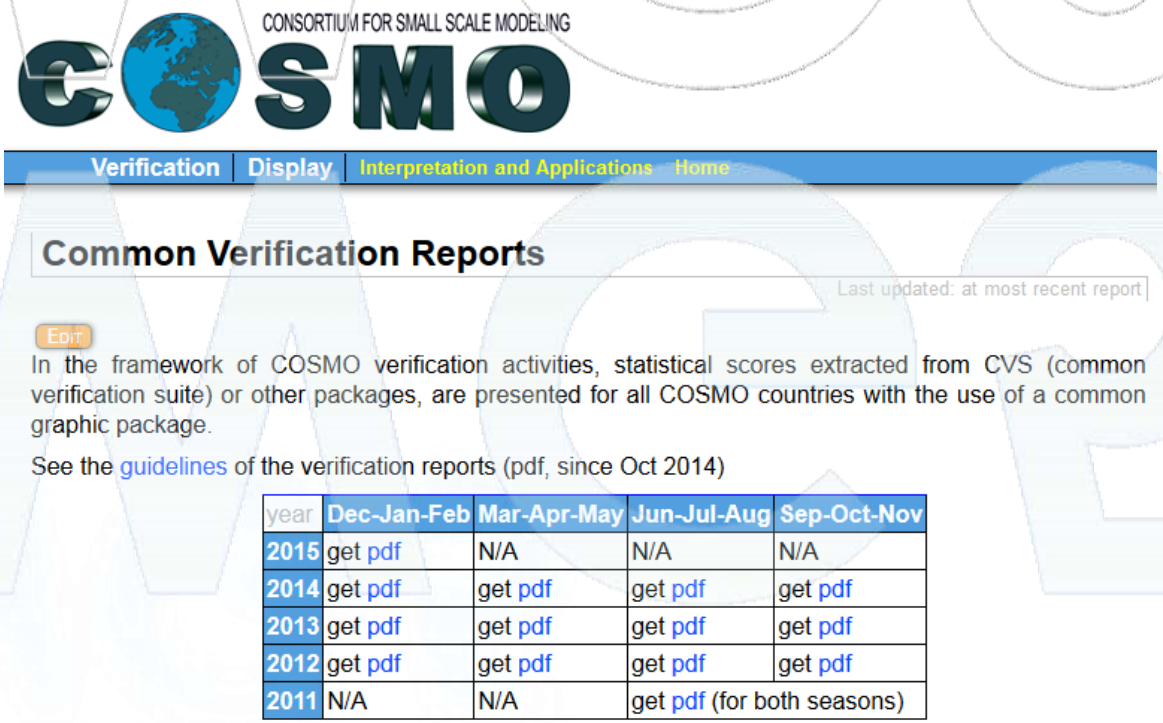
STDE very slightly decreased, only for 1 h averages

Summary

- Using 10 min model averages instead of the standard instantaneous COSMO output does hardly change the verification results
- → We do not artificially increase the error when we compare instantaneous model output instead of 10 min averages to 10 min averaged SYNOP observations
- Comparing hourly averages (of model and obs) does decrease the spread of the error (STDE) but only by a very small amount and only for few parameters
- → Largest effect for wind speed, reduction of STDE by 5%
- → Sub-hourly “noise” is not a significant problem
- Averaging observation reduces STDE more than averaging model

Item 4: Common Plot Reports 2015-2016

Presentation of Verification Overview (U.Damrath



Task description

Responsible: D. Boukouvala (HNMS), U.Damrath (DWD)

Score Production

Preparation of input data and calculation of seasonal statistics over a common area according to the guidelines derived on an annual basis from WG5 (http://www.cosmo-model.org/content/tasks /verification.priv/common/guidelines.pdf) for each participating model. This Task includes conditional verification tests performed over this area. ***IFS driving model statistics*** has also been added this year.

Seasons: JJA 2014, SON 2014, DJF 2015, MAM 2015

Reporting

Processing of data from all models for each parameter and conditional verification test in appropriate format

R scripting for production of graphs (cross model representation)

Preparation of report for each season

Commenting of significant errors or discrepancies between models

Preparation of web graphics based on DWD representation regime

Long term trend calculations

Interaction Subgrid scale Orography-Turbulence in COSMO

Results from case study and common plots

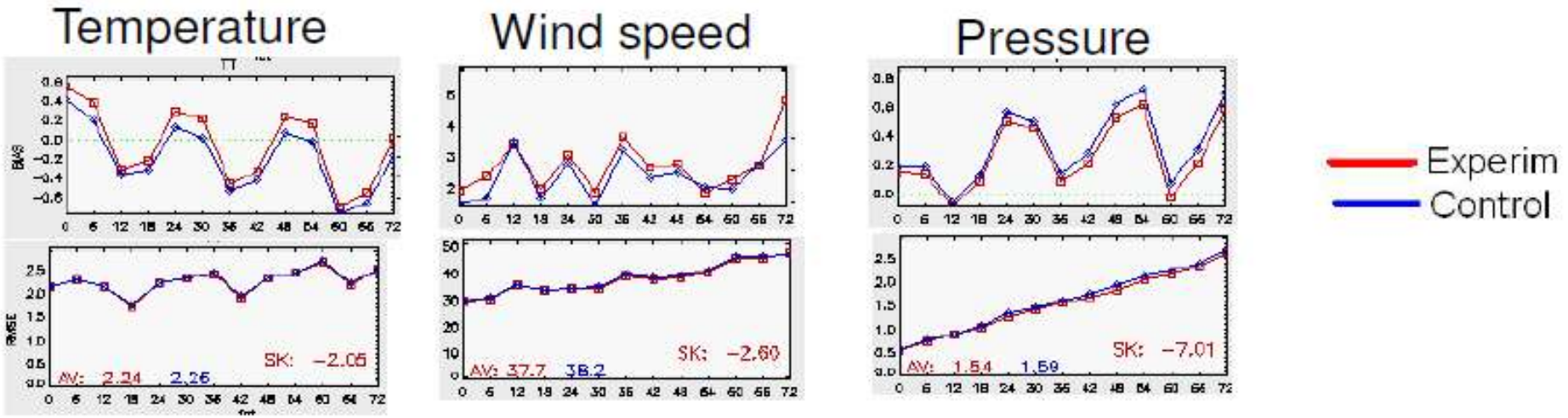
I. Cerenzia[1,2], MS. Tesini [2], D. Boucouvala [3], M. Raschendorfer [4], F. Gofa[3]

[1] University of Bologna
[2] Arpa-EMR SIMC
[3] Hellenic National Meteorological Service
[4] Deutscher Wetterdienst

09-2015

Background

ltkesso tested in a parallel test over COSMO-EU domain (2 months in 2011):



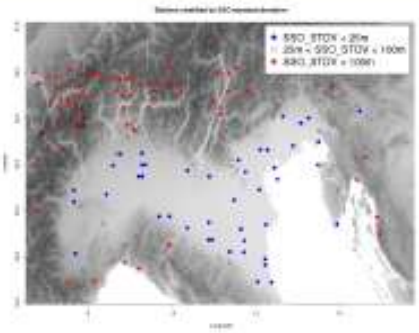
(Raschendorfer, 2011)

What is the effect of the parameterization on turbulence and dynamical variables in points with high and low SSO?

Case studies

Common plots

Case study verification



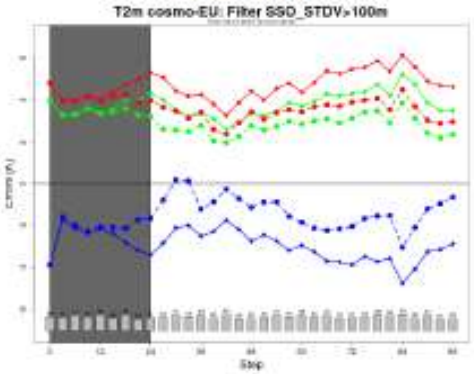
Case 1) Anticyclonic regime

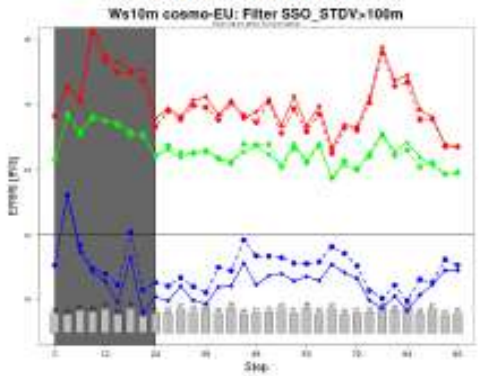
T2m

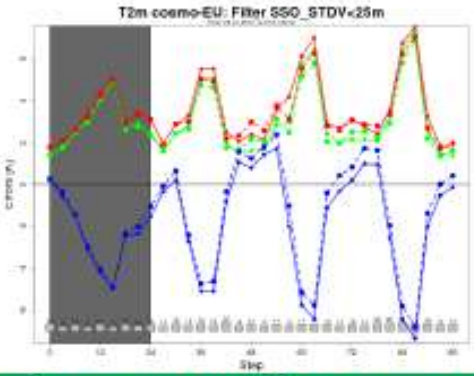
Ws10m

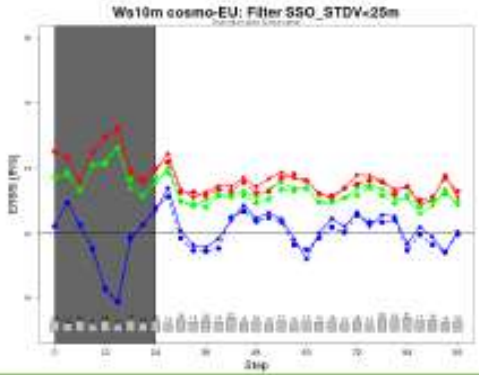
HIGH SSO

LOW SSO





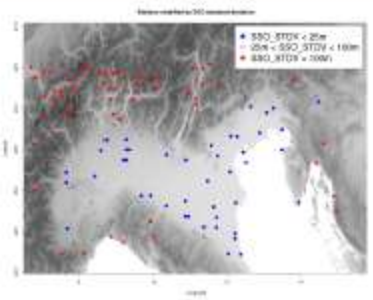




Control
ME
MAE
RMSE

Experim
ME
MAE
RMSE

Case study verification

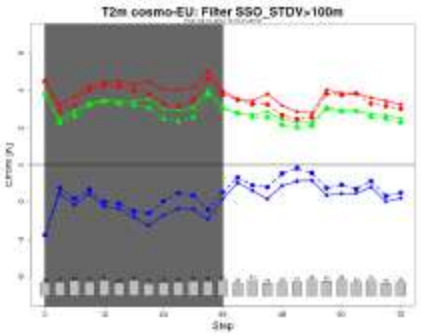


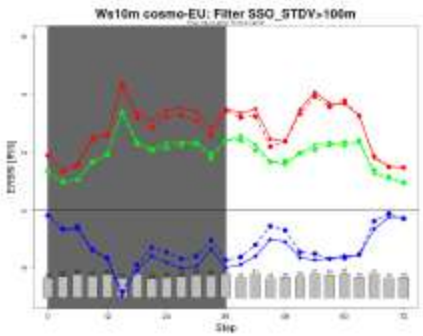
Case 2) Frontal passage

T2m

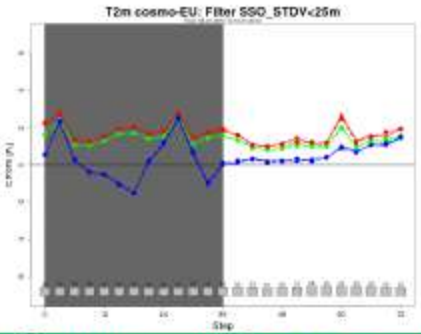
Ws10m

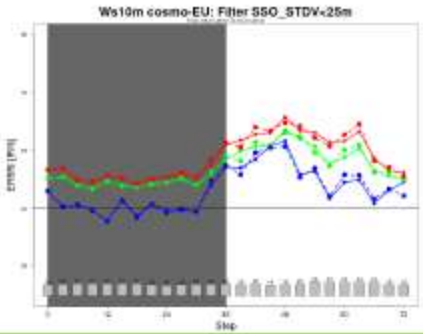
HIGH SSO





LOW SSO





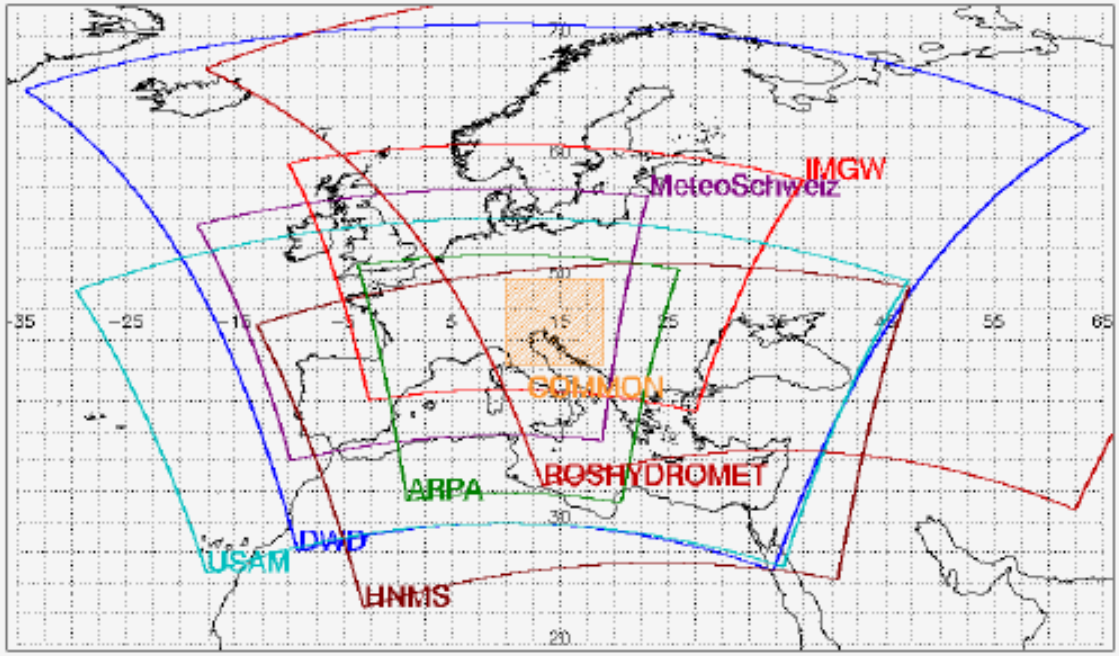
Control
ME
MAE
RMSE

Experim
ME
MAE
RMSE

Common plots

Aim: confirmation of the results on a longer statistical base.

- *ltkesso* is activated in some COSMO members
- Filter the stations on SSO standard deviation base (approx. same number)
- periods: DJF and MAM 2015
- domain: common area



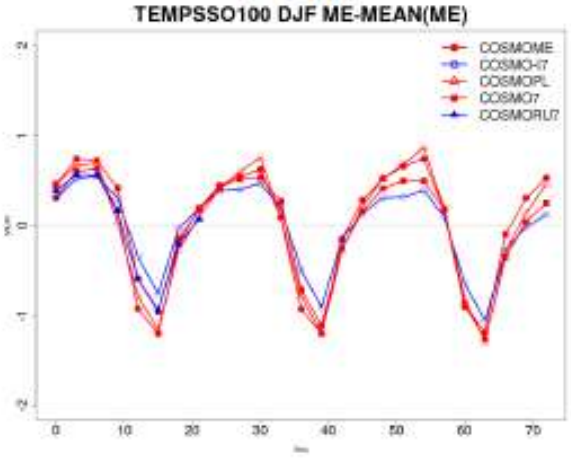
Common plots: DJF

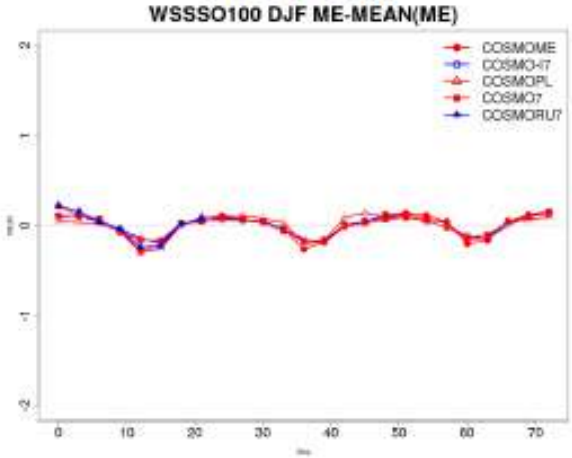
T2m

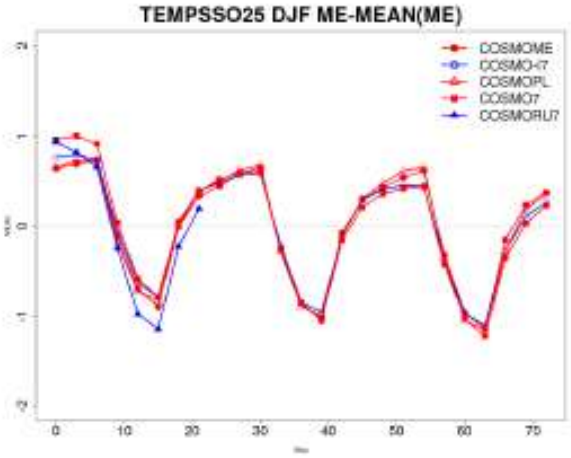
Ws10m

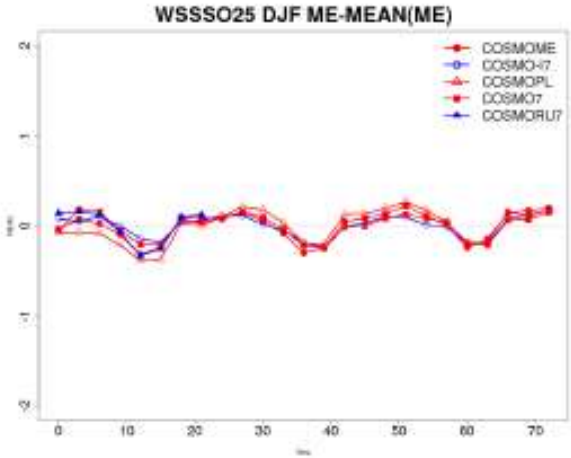
HIGH SSO

LOW SSO









LTKESSO
TRUE
FALSE

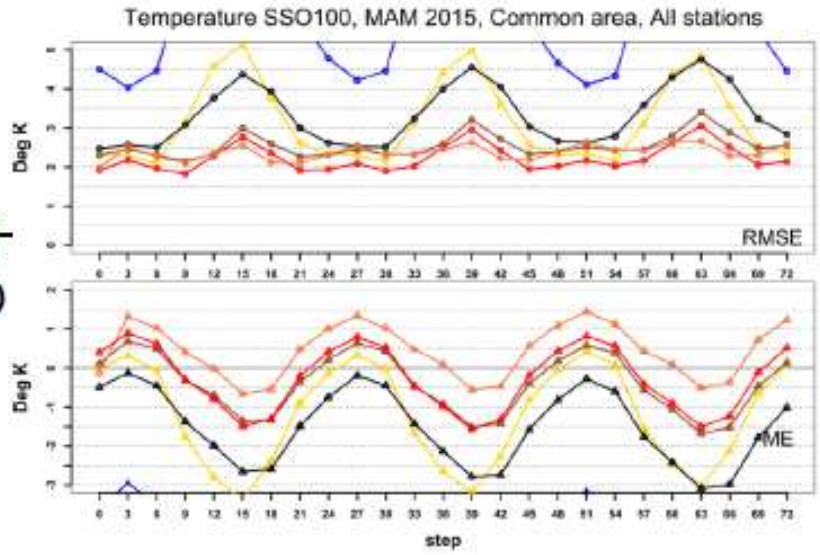
- sensitivity in ME during night (stabler stratification) for T2m in points with high SSO

Common plots: MAM

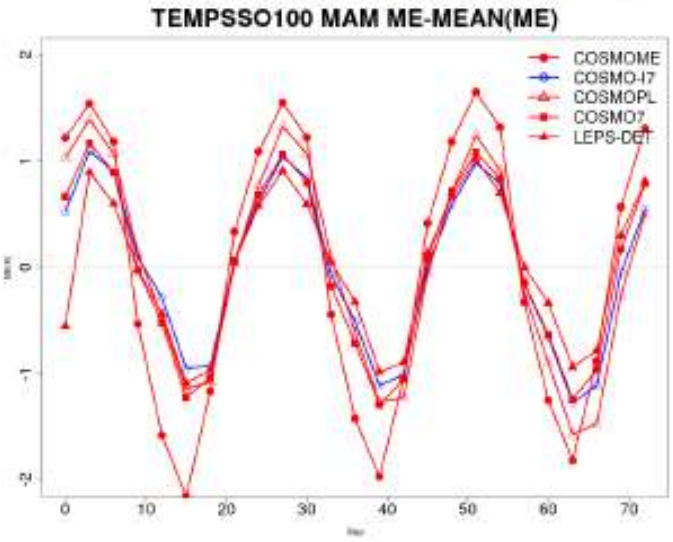
$T2m$

$T2m(ME - Mean(ME))$

HIGH SSO



C-GR
C-ME
C-I7
C-PL
C-7
I7 LEPS



In MAM the signal in ME is lost: synoptic circulation induces less frequently a stable stratification in points with high SSO.

Conclusions

Not visible a clear signal from the Common Plots analysis

What do we learn?

- long statistics smooths out the signal visible in the case studies (only $\Delta T2m$ as big as 2C in the case studies can be detected in the Common Plots)
- ME is more sensitive than RMSE
- some help may come by filtering for meteorological conditions (e.g. stable stratification)

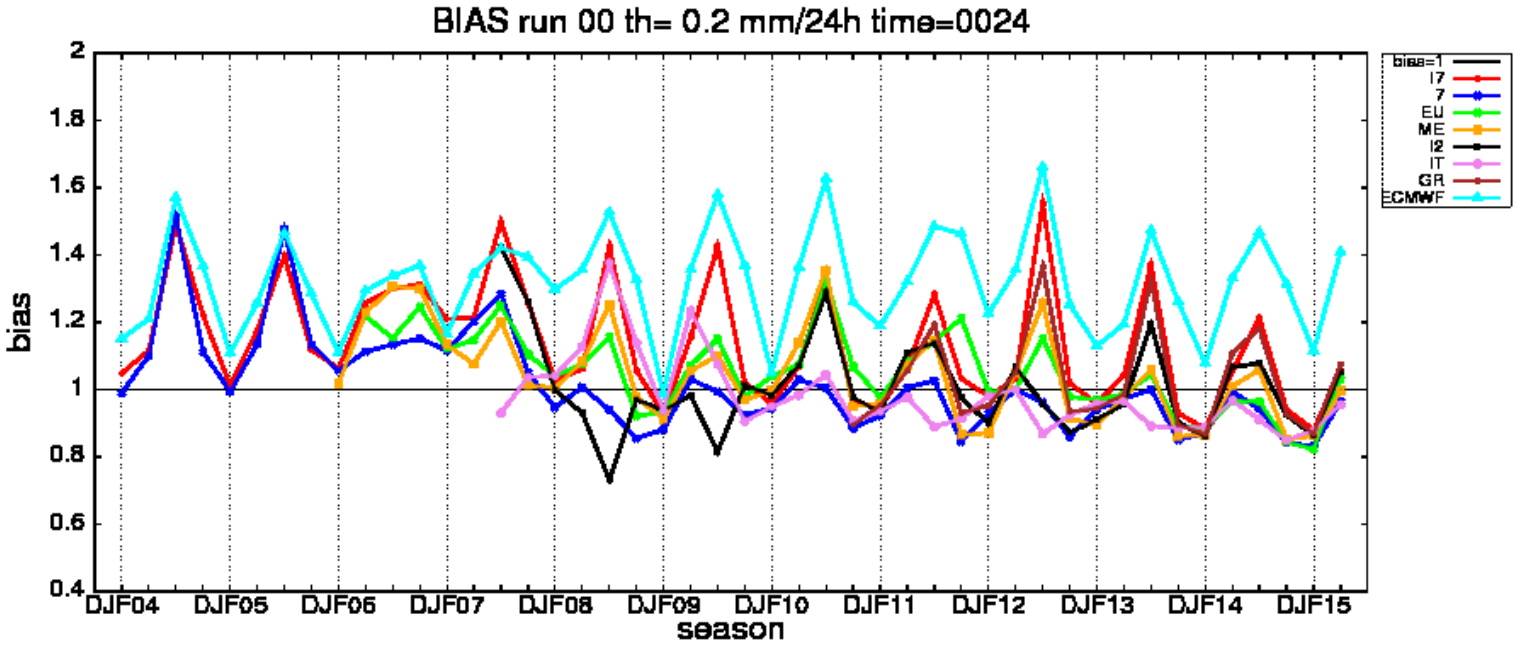
COSMO models (8) comparison over Italian alert areas: long trends and last year review

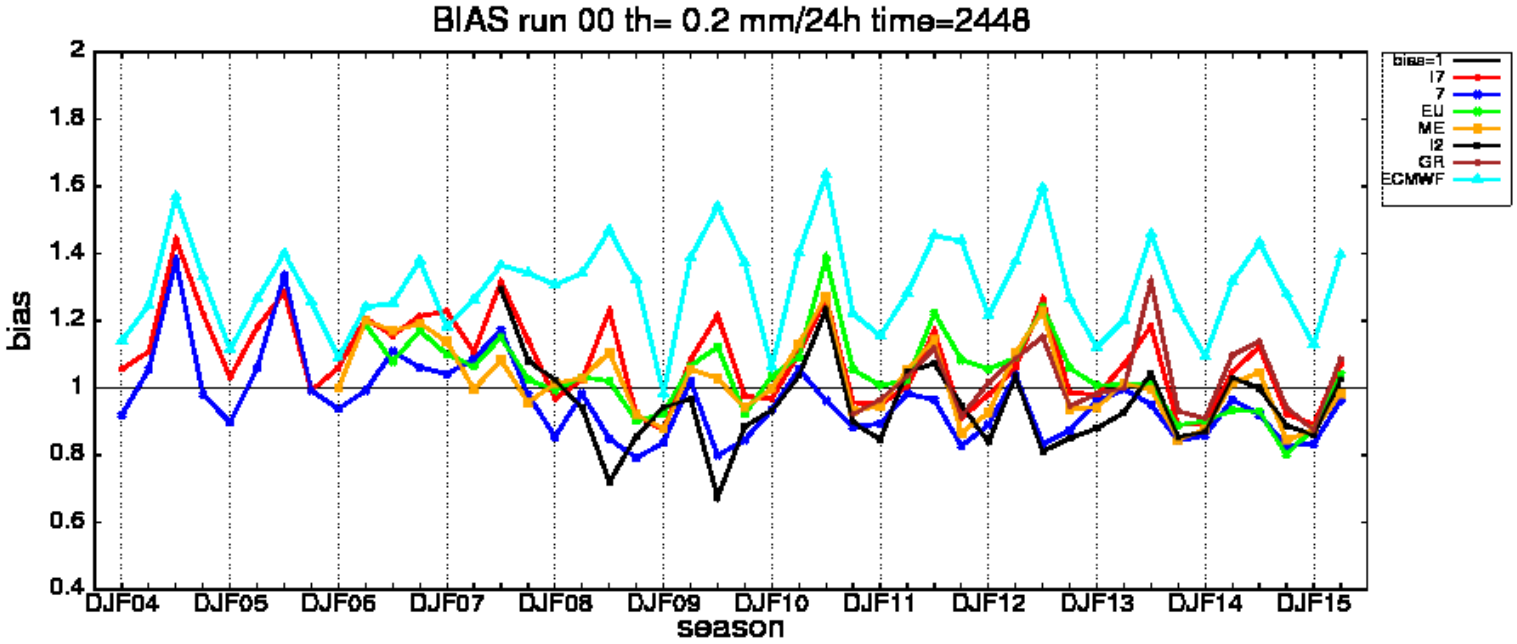
Elena Oberto
Naima Vela

This work has been done with the collaboration and the funds of Civil Protection Department. Furthermore thanks again to Civil Protection Department for making available the high resolution rain gauges dataset usefull for the verification tools.

Cosmo General Meeting 2015 –
Wroclaw (Poland)

LONG TREND PRECIPITATION with high resolution stations

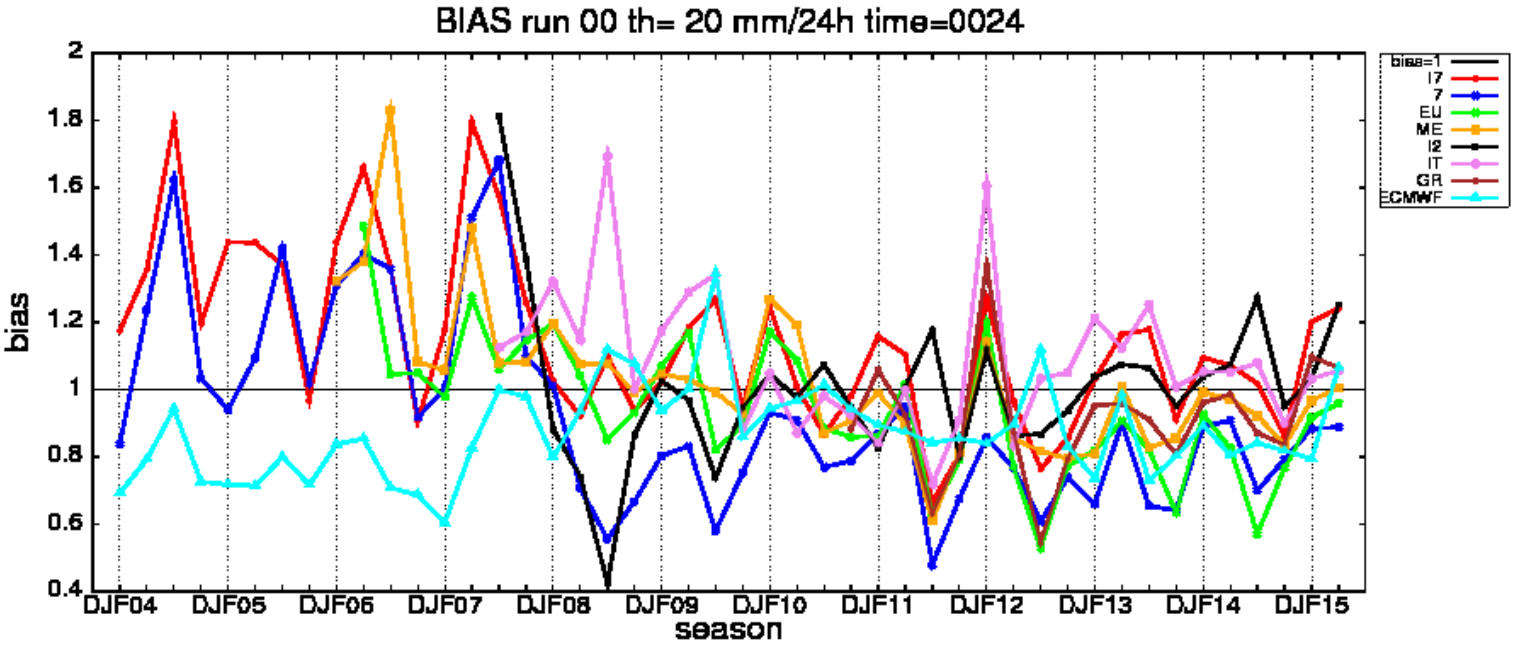




LOW THRESHOLDS

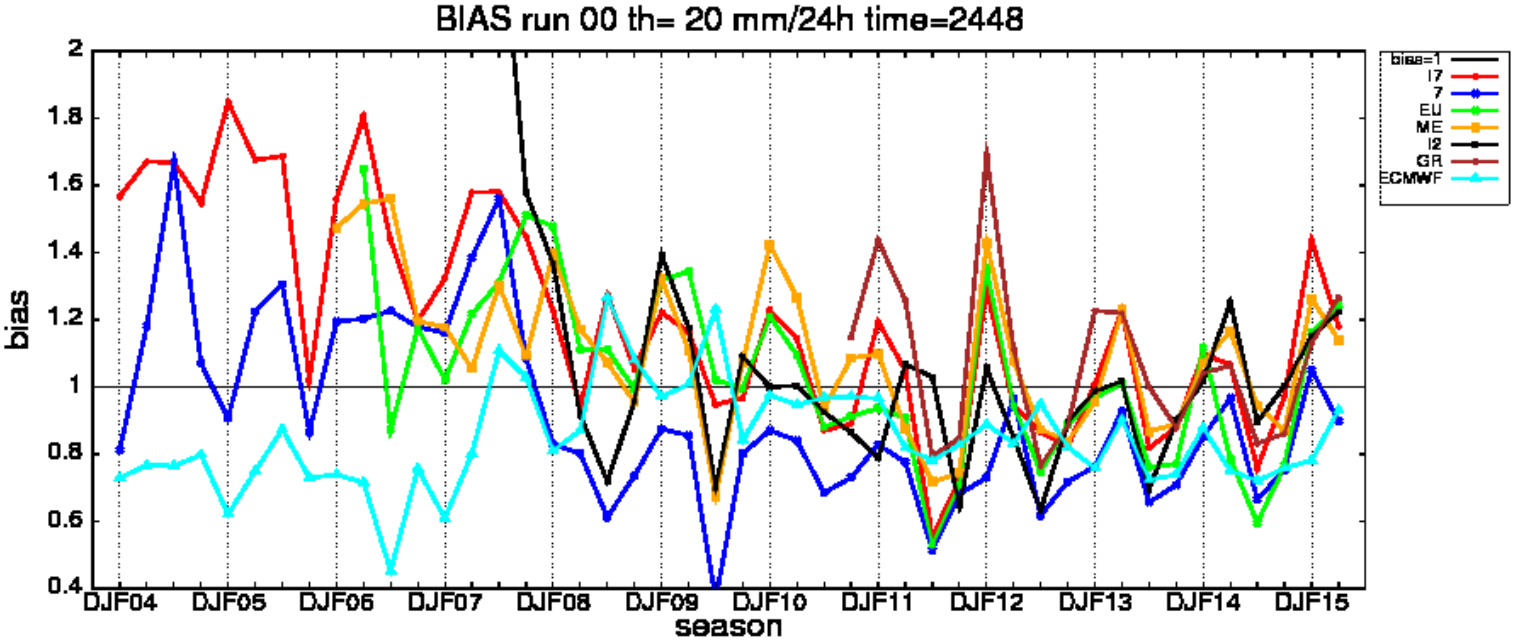
- Ecmwf overestimation
- Summer overestimation

LONG TREND PRECIPITATION with high resolution stations

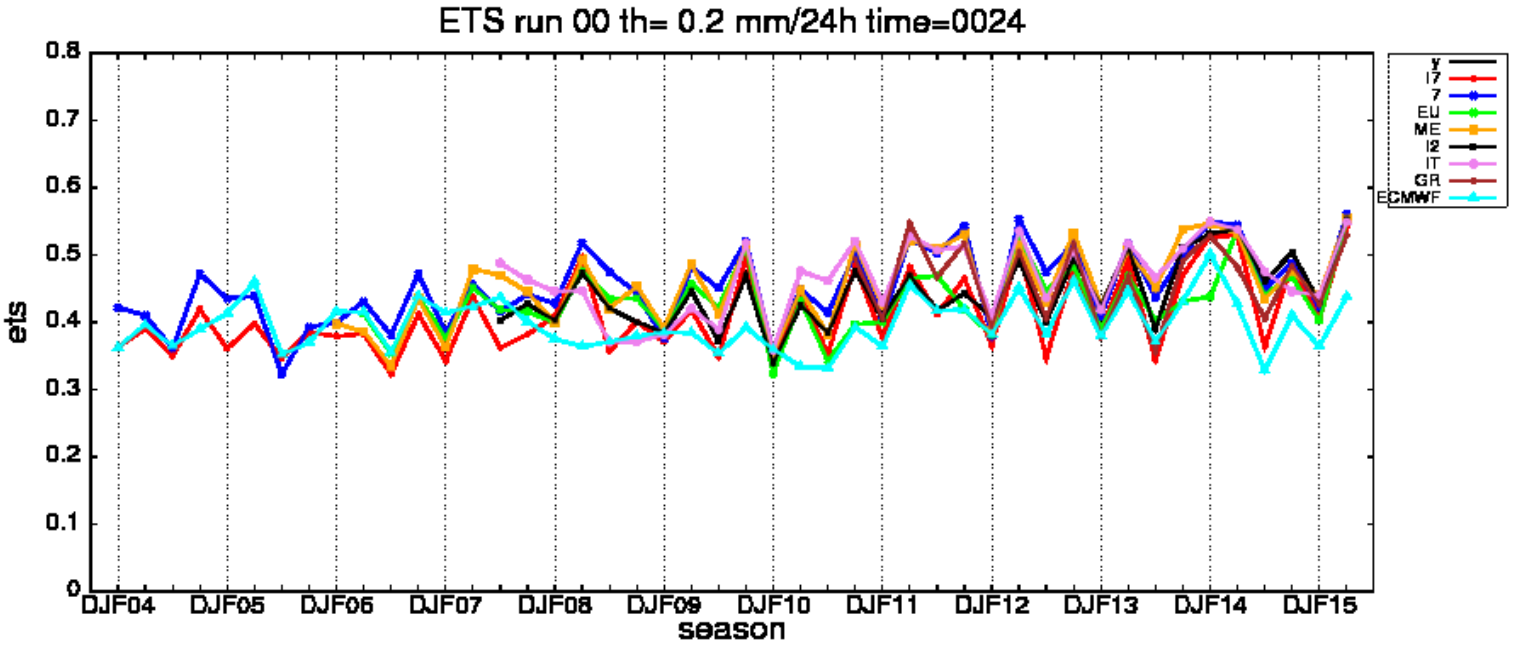


HIGH THRESHOLDS

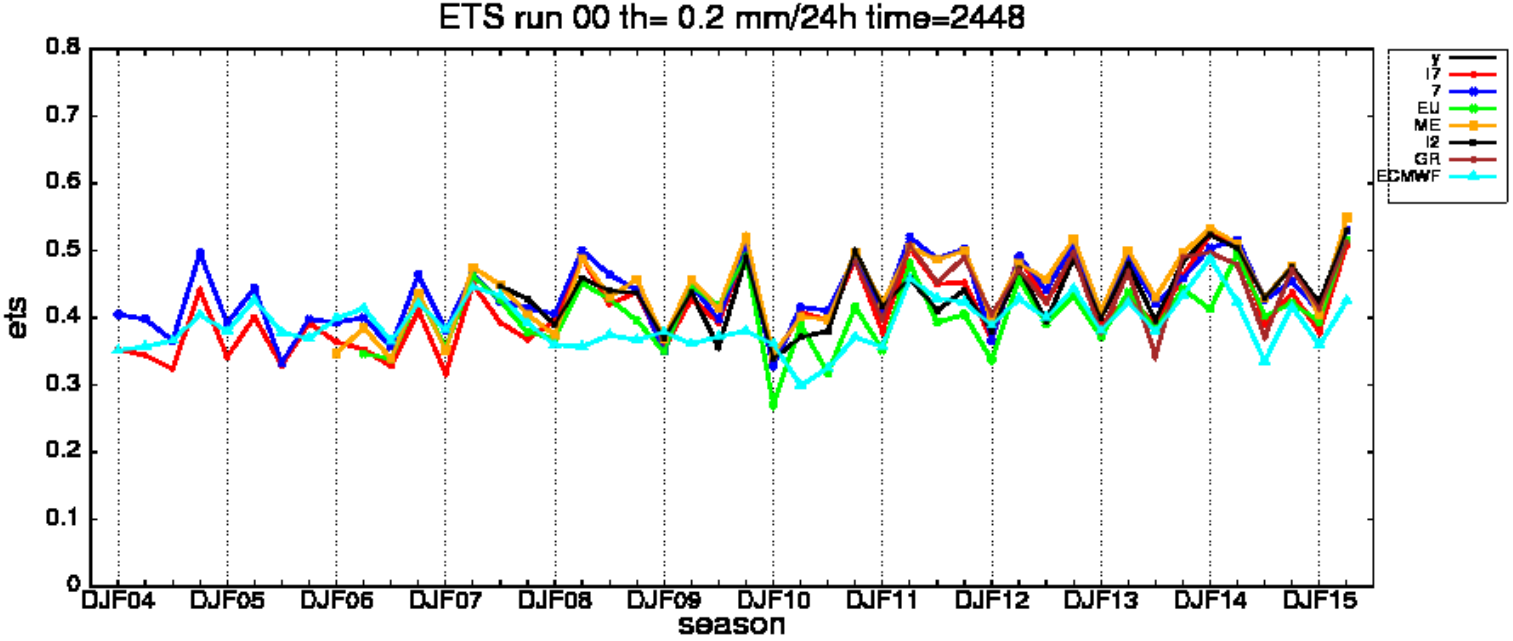
- General underestimation, especially 7, EU



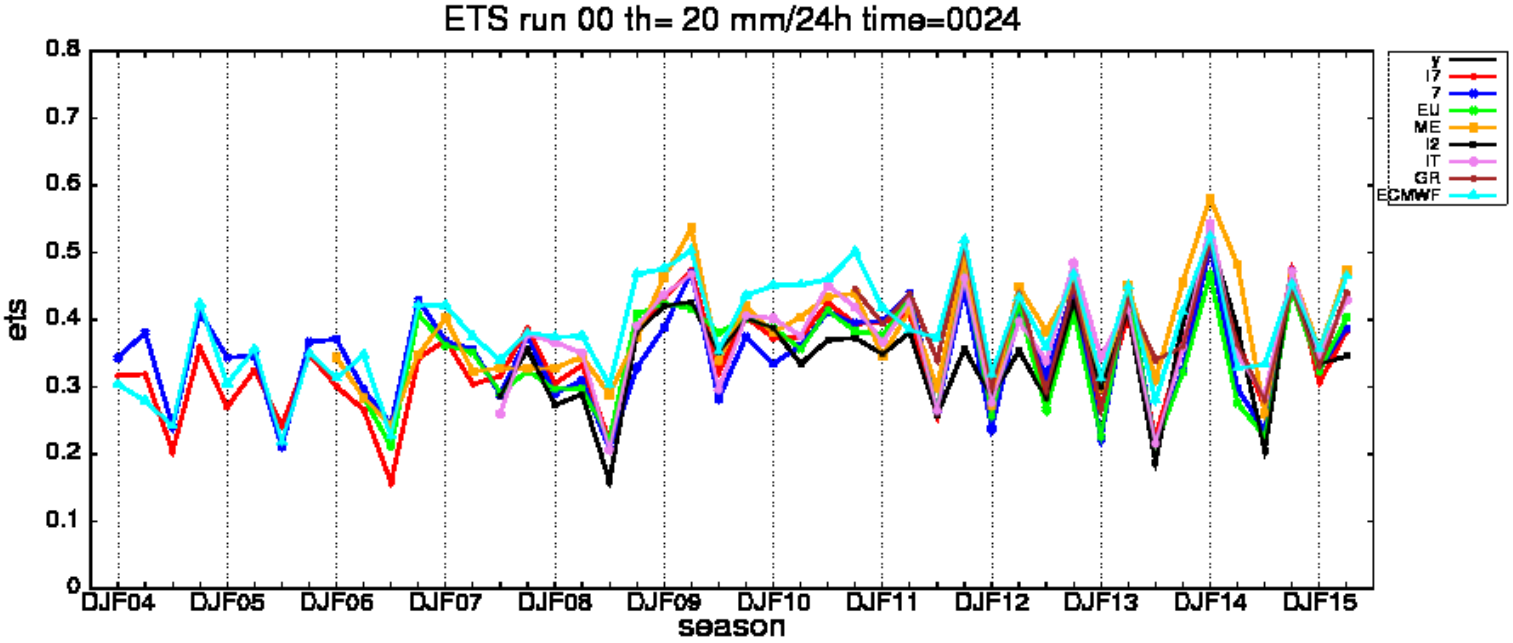
LONG TREND PRECIPITATION with high resolution stations



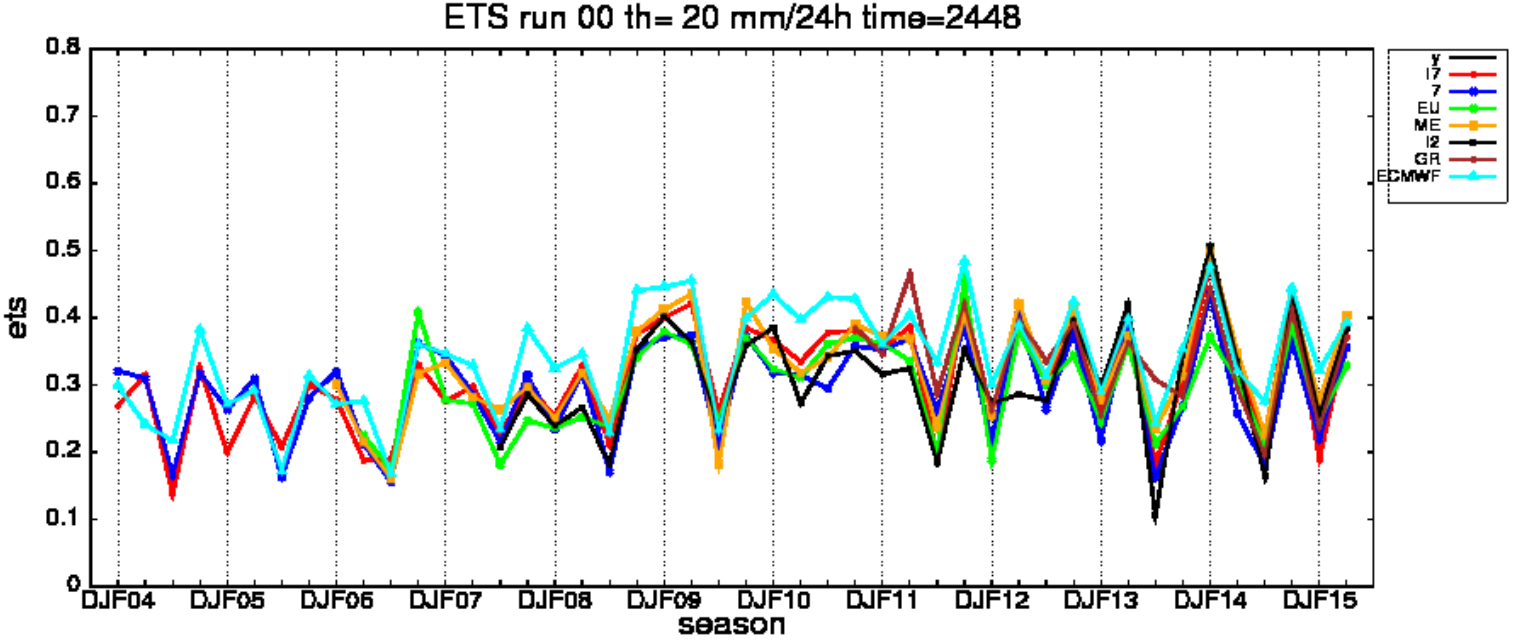
LOW THRESHOLDS



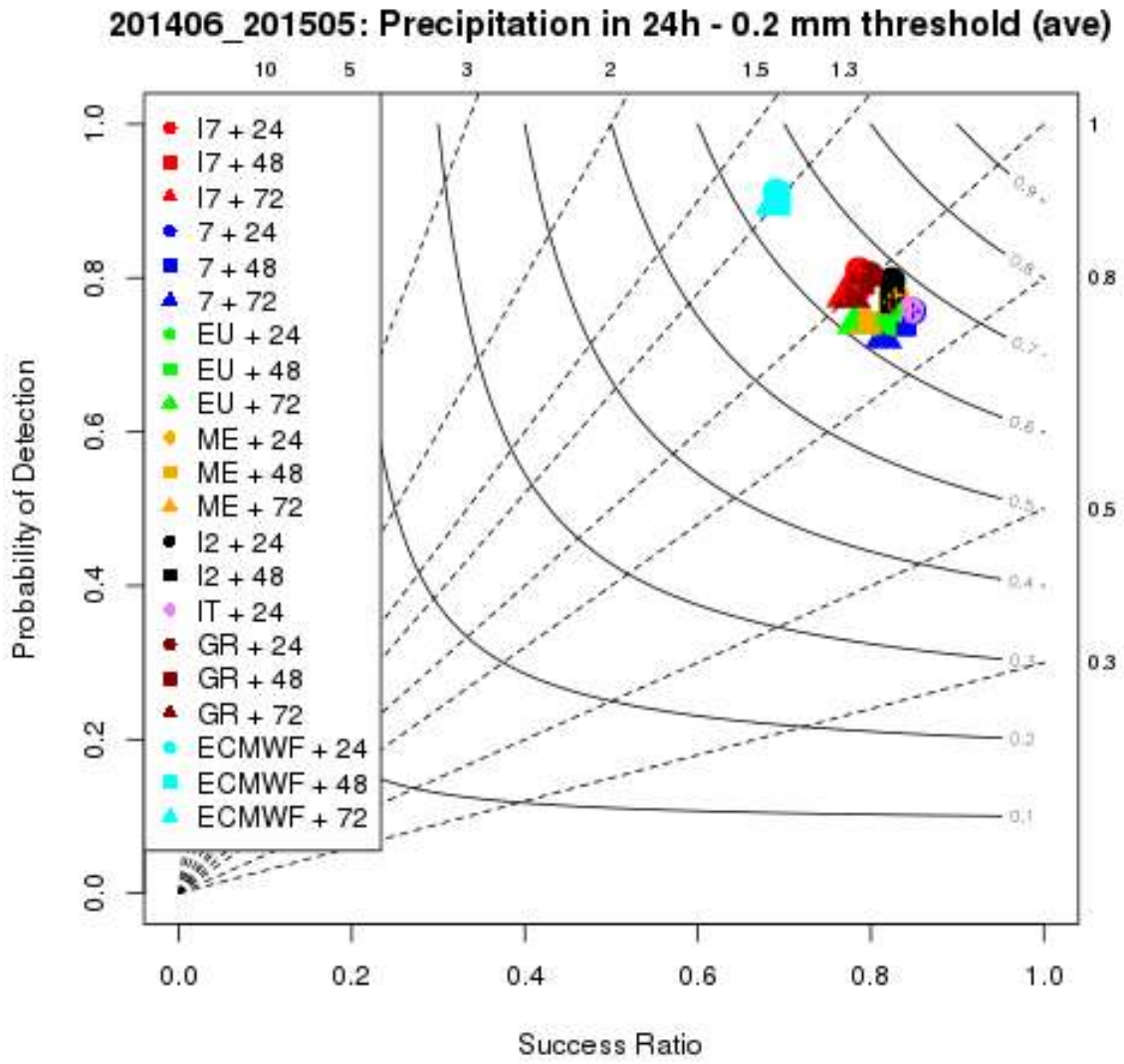
LONG TREND PRECIPITATION with high resolution stations



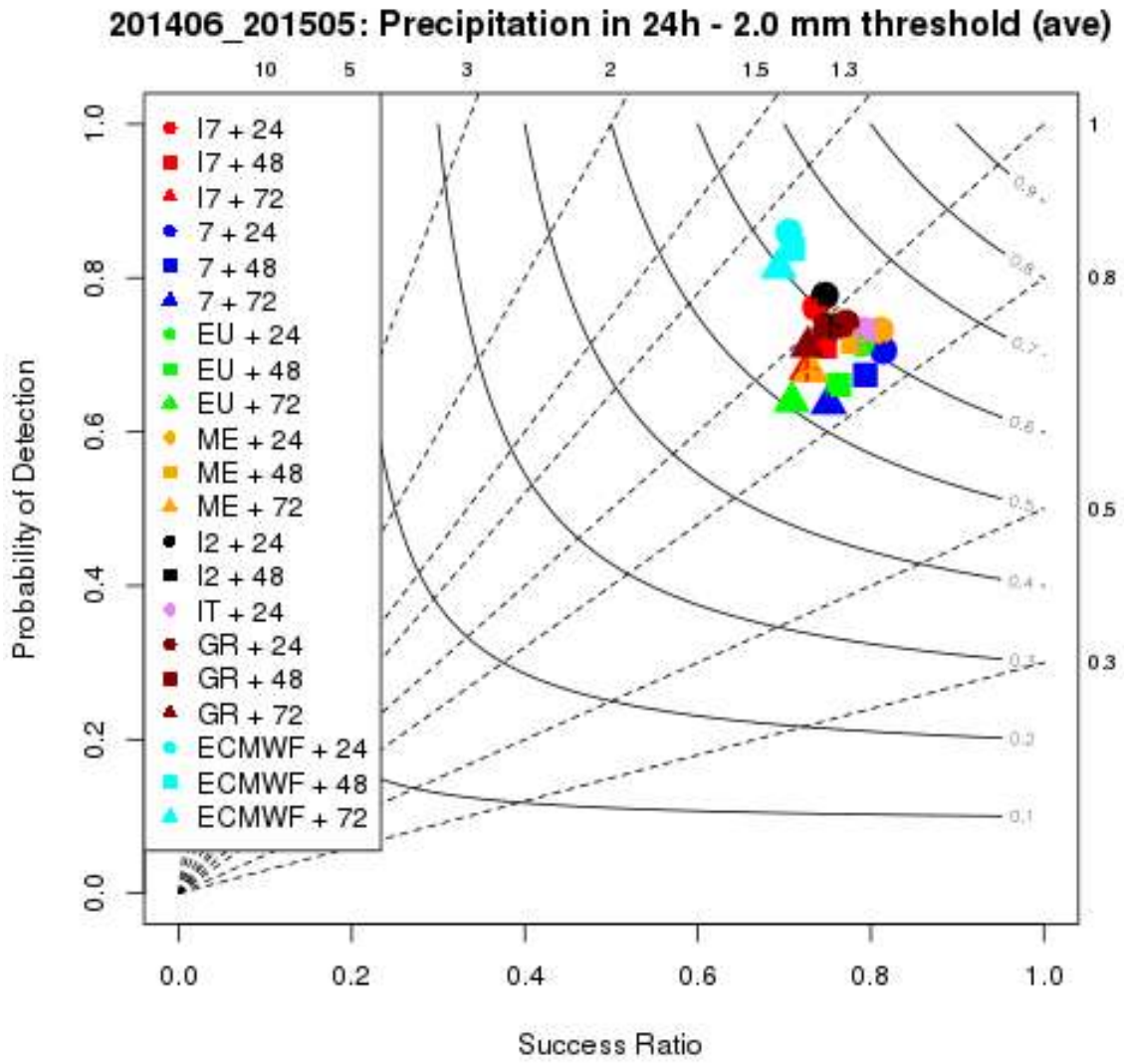
HIGH THRESHOLDS



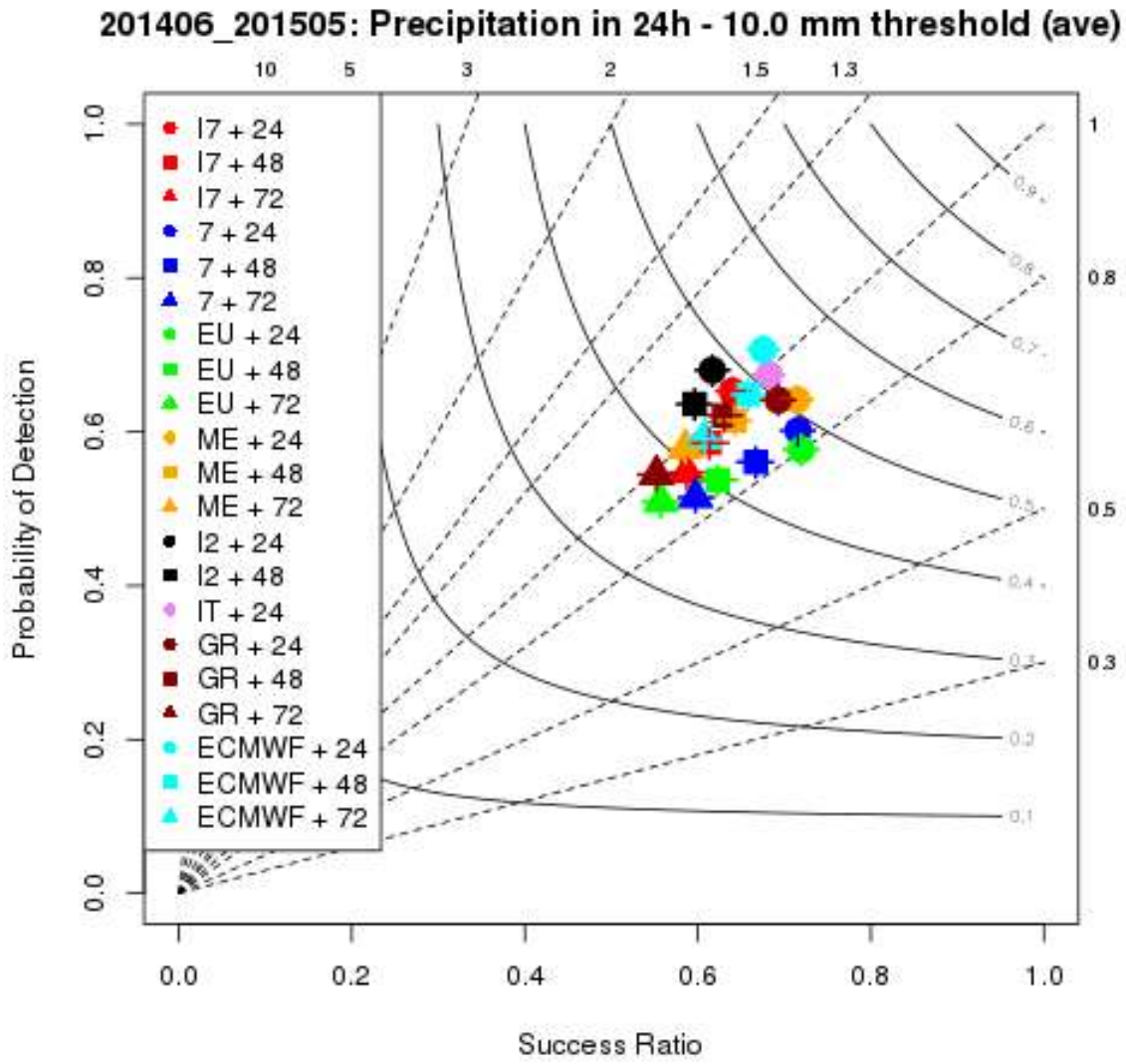
201406-201505: Average over area > 0.2 mm/24h



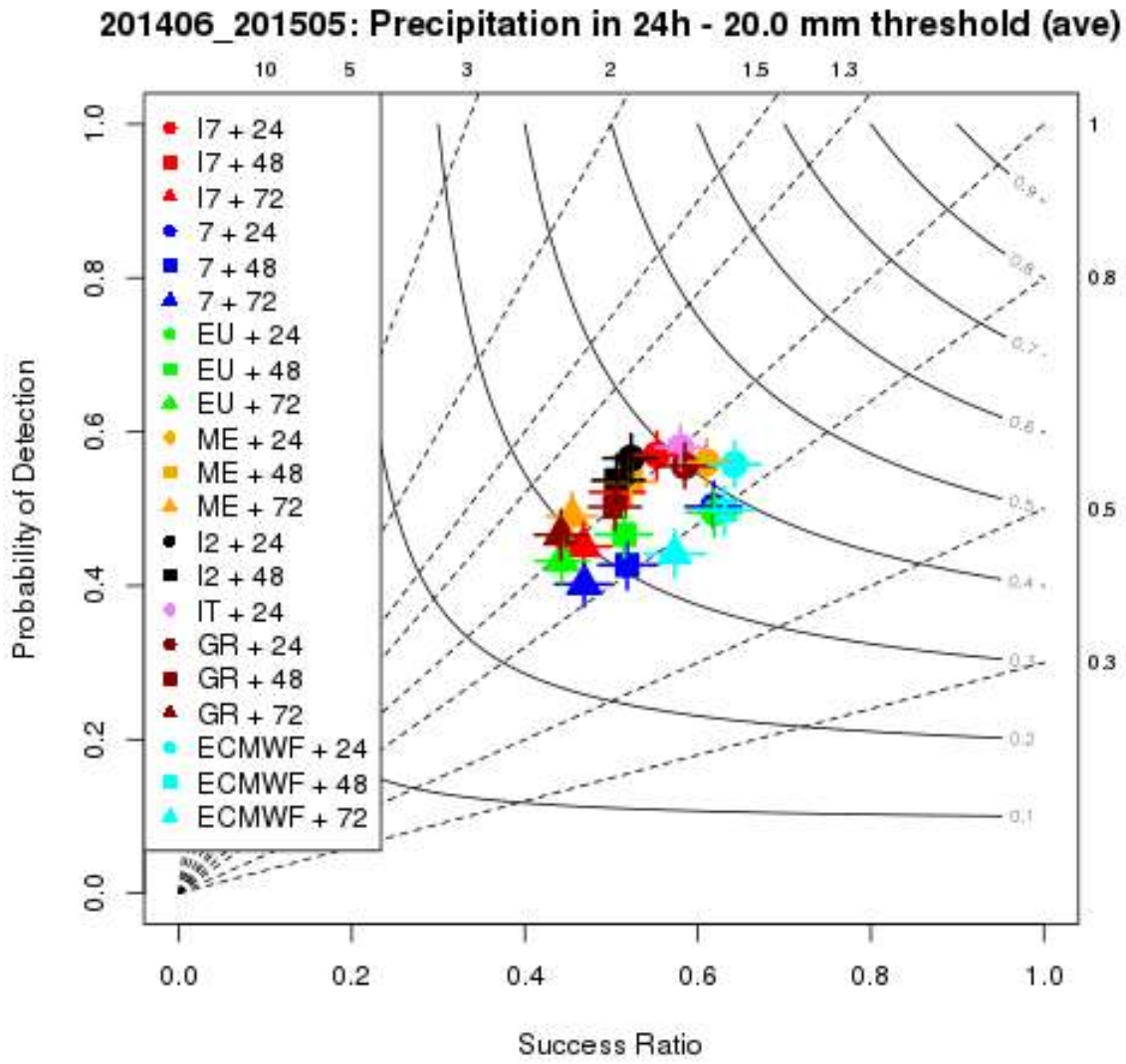
201406-201505: Average over area > 2 mm/24h



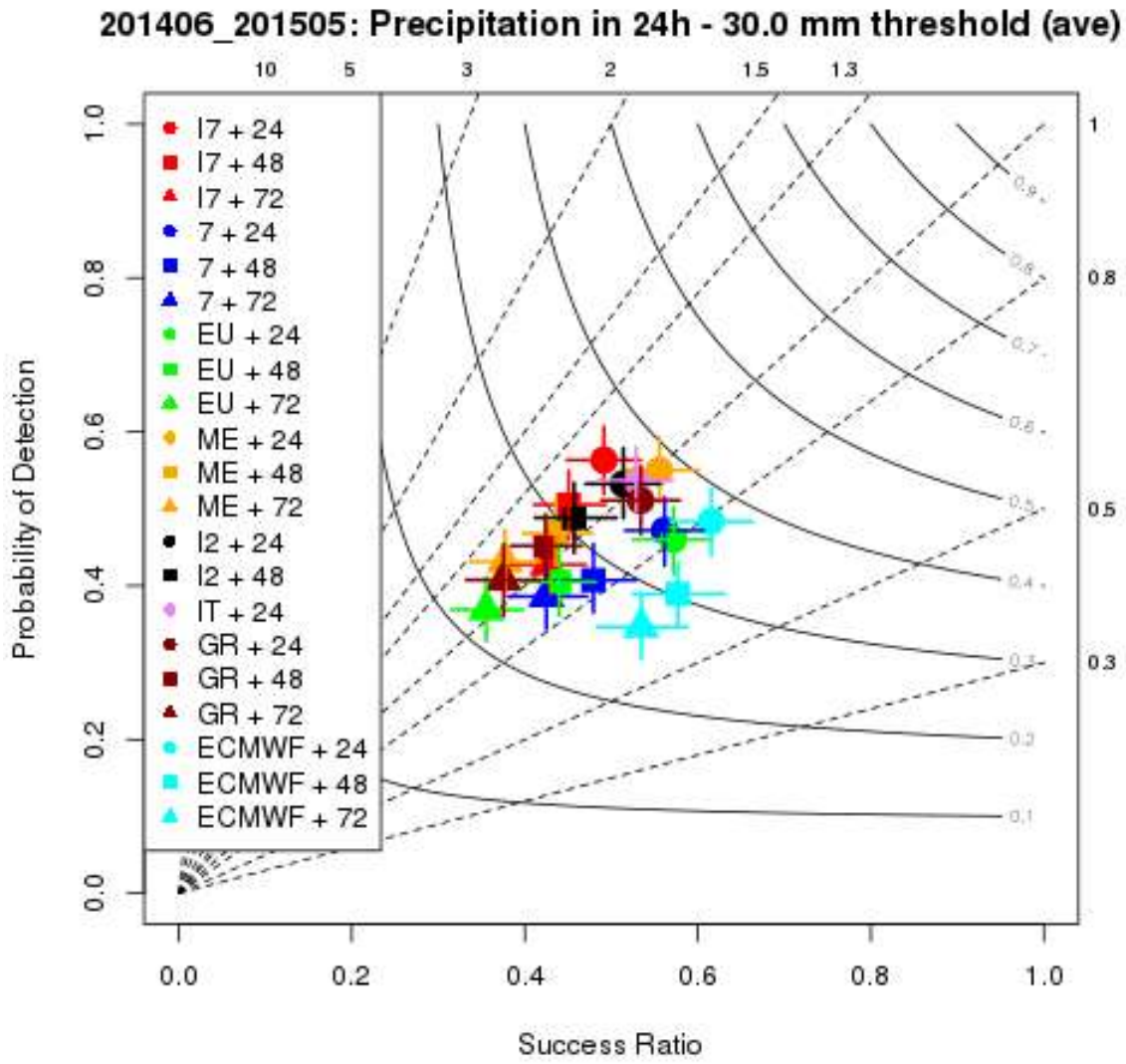
201406-201505: Average over area > 10 mm/24h



201406-201505: Average over area > 20 mm/24h



201406-201505: Average over area > 30 mm/24h



Operational verification using DIST:

Comparison of models with different resolution

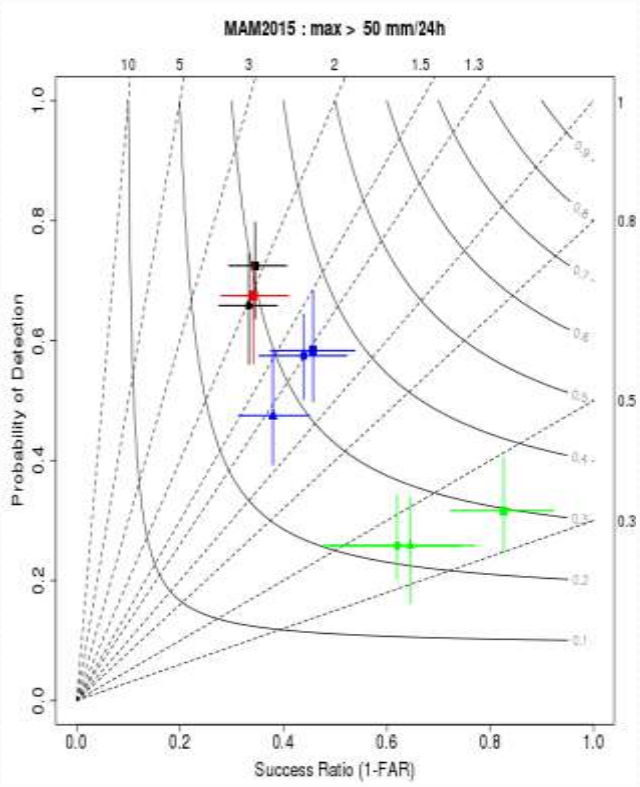
(COSMO1, COSMO-I2, COSMO-I7, IFS-ECMWF)

Maria Stefania Tesini
ARPA-SIMC Emilia-Romagna

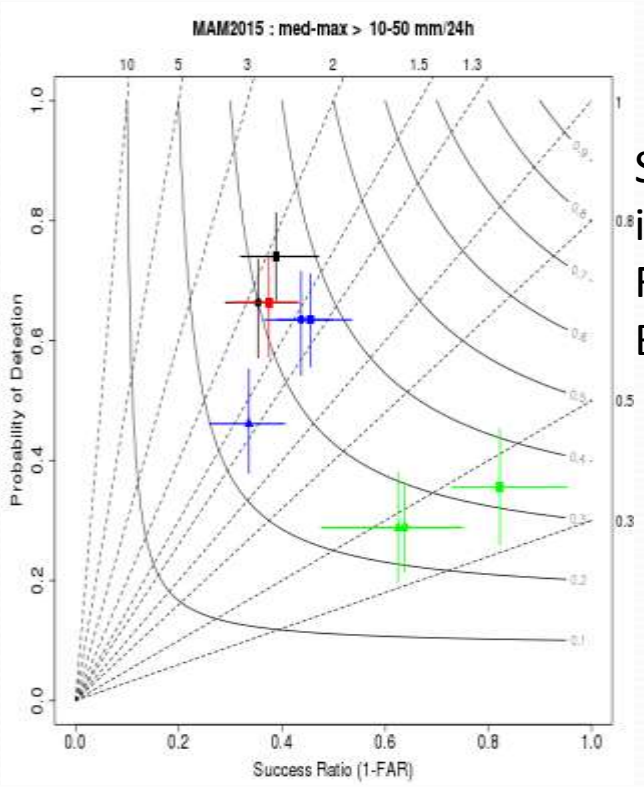
17th COSMO General Meeting
07/09/2015 Wrocław (Poland)

MAM 2015 – Median & Max

MAX > 50 mm/24h



MED > 10 mm/24h & MAX > 50 mm/24h



Slight improvement FAR reduced, Better TS

COSMO-I2_00 - fc + 24
COSMO-I2_00 - fc + 48
COSMO-I7_00 - fc + 24
COSMO-I7_00 - fc + 48
COSMO-I7_00 - fc + 72
COSMO1CH_00 - fc + 24
ECMWF_00 - fc + 24
ECMWF_00 - fc + 48
ECMWF_00 - fc + 72

arpa ER

WG5 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