WG5: Verification and Case studies

Overview of activities

Flora Gofa

COSMO GM Plenary session, 8-11 Sept 2014, Eretria, Greece

SP: WG5 related activities (validation and diagnostics)

•*Tackling model performance improvement issues through the use of conditional verification (CV)*

As model errors should be related to specific inaccurately simulated processes, verification under specific conditions (CV) have to be chosen in order to extract selected model uncertainties due to simulation errors, isolating single processes or uncertainties responsible for measured simulation errors. This procedure is based on the selection of forecast products and "mask variables" (model variables, observations or external variables) and application of arbitrary thresholds (conditions) to produce verification.

Cross cutting issue section: Processing verification feedback on model development

•*Statistical methods to identify the skill of convection permitting and near-convection-resolving model configurations*

Increasing of models resolution can lead, especially for precipitation but also for continuous surface parameters, to forecasts detail more realistic but inaccurate, the so-called double penalty effect. For this reason neighborhood methods were employed to compare forecasts in appropriate selected size neighborhoods with the gridded radar data for precipitation. For this reason a verification framework needs to be defined (even probabilistic).

Statistical methods proposed should lead to the estimation of the relative skill gained using higher resolution, to the assistance in the decision-making process for model upgrades for similar horizontal resolution and to the comparison between the determinist forecasts with ensemble ones.

•***Exploitation of available observational dataset for operational and scientific purposes***

For model-oriented verification, processing of the observation data needs to be done to match the spatial and temporal scales resolvable by the model. This requires the availability of high spatial resolution observations (satellite or radar post-processed data) to be used to produce vertical profiles or gridded surface analysis. Furthermore particularly important is the exploitation of controlled and possibly homogenous set of surface observations, concerning fluxes, radiation and soil characteristics, such as those available from SRWNP Data Pool Exchange.

•***Development of tools for probabilistic and ensemble forecast verification***

The challenges in verifying "convection-permitting" ensembles are basically the same as in mesoscale "convection-parameterisation" ensembles, with some **added complexities**. Due to their nature, convection-permitting ensembles focus on the shortest range (0-24h) and large error growth in such systems which are correlated strongly to the highly non-linear physical processes of convection, thus verification measures must focus on the relevant gain of the use of such systems toward better representation of convection-based parameters. As for deterministic forecasts, neighborhood methods are proposed to be employed to account for the spatial mismatches between forecasts and observations, especially for precipitation, even though ensemble forecasts can address uncertainties of small-scale processes more adequately.

•***Severe and High Impact Weather***

As there is an increased demand that meteorological services provide accurate forecasts of extreme weather, it is therefore important to be able to objectively evaluate the model performance in these cases. Severe events are rare and this is the reason that standard skill scores are not useful as they depend on base rate. Dependency scores like SEDS and SEDI have been extensively used by the NWP community for some time, but the use of other scores and methods will also be evaluated.
The SEEPS is not designed for extremes but does provide a very useful, visual way of displaying forecast issues to local biases because it utilises a climatology.

•***User-oriented Verification products***

With increasing model resolution, the number of products the users will ask, as well as their objective performance in terms of their expected quality is only going to rise. Different users might have needs for different verification information (e.g. administrative decisions may depend on model performance), so different verification strategies have to be chosen. It will be necessary to diversify verification methodologies to match the different needs and to this end, the scientific community will have to work more closely with the user community in the design of such verification strategies.

These main activities could be reviewed and updated in the light of future developments in the main fields of model improvements concerning physics and data assimilation, in order to respond to the actual needs of developers and users alike.

Main Reviewer Comments

- "Clear plan, and good to see so much attention devoted to diagnostics! "(JO)
- "The conditional verification approach is a very interesting one. An element in the validation which is not really mentioned is the (routine) use of a set of well-defined, well-observed case studies and associated forcing or observational information. "(JO)
- "The feature-based, fuzzy, probabilistic and high impact weather verification techniques are sound, but they all depend very much on gridded data. What I find lacking a bit is a strategy how to get the observations you need for this, other than radar. "(JO)
- "As the information content of satellite observations increases, effort must be given for further exploitation of the data as radiation, cloudiness, vertically integrated water vapour content, for verification purposes." (MM)
- "To detect severe weather events it might help to compare the forecast with climatological probabilities of occurrences of these kind of events from hindcasts to account for the model shortcomings. "(BF)

Resources

The experience gained over the past several years indicates skilled but limited resources in the COSMO community regarding operational verification activities and implementation of new approaches and methodologies.

As the lack of resources is a common problem to other European Consortia, in order to optimize them, a recommended strategy would be to monitor the efforts of the various European Consortia and International Programmes in the field of verification, namely to use or adapt what has already been developed and encourage knowledge sharing amongst the scientific and operational communities regarding new methodologies, research results and approaches to verification issues.

MesoVICT:Mesoscale Verification Inter-Comparison over Complex Terrain

WMO Joint Working Group on Forecast Verification Research (JWGFVR)

The aims of the project can be summarised as follows:

- To investigate the ability of existing or newly developed spatial verification methods to verify fields **other** than **deterministic precipitation** forecasts, e.g., wind forecasts and ensemble forecasts.
- To demonstrate the capability of spatial verification methods over **complex terrain**, and gain an understanding of the issues that arise from this more challenging situation.
- To encourage community participation in the **development** and improvement of **spatial verification methods**, especially for evaluating high resolution numerical forecasts.
- To provide a community testbed where common data sets are available, but also for the sharing of data and code to assist in developing and testing spatial verification methods

Link of MesoVICT project to a future COSMO PT/PP for the adaptation of strategy, methods and sofware to COSMO community

New spatial methods

Filtering methods

- **Neighborhood**

  (Ebert, 2008)
- **Scale Decomposition**

Displacement methods

- **Features-based**
  - ✓Contiguous Rain Area (CRA) (Ebert and McBride, 2000)
  - ✓Method for Object-based Diagnostic Evaluation (MODE) (Davis et al., 2006)
  - ✓ SAL technique (Wernli et al., 2008)
- **Field Deformation**

PP VERSUS2 Phase6

Major Tasks accomplished:

EPS refinements-new scores implementation

Activity Proposal to support spatial methods development (VAST)

PP VERSUS2 Phase7 (Last?)

Major Tasks planned:

Grib2, Feedback Files implementation

Operationality issues in VERSUS use

Importance of extension in LT development and Maintenance Plan

PT NWP Test Suite (completed)

Build up a software environment to perform carefully-controlled and rigorous testing of each new COSMO model version at ECMWF capabilities

Provide the COSMO community with standards against which the impacts of new developments in the model should be evaluated

PL Amalia IRIZA (NMA)

Andrea Montani (ARPA – SIMC)

Flora Gofa (HNMS)

Rodica Claudia DUMITRACHE (NMA)

Adriano Raspanti (USAM)

Overview of verification activities

Authors: ALL

WG5

Neighbourhood verification at MCH for precipitation and brightness temperature

COSMO GM – WG5 Session
8 September 2014, Eretria (GR)

Francis Schubiger, Daniel Leuenberger, Thomas Leutert
MeteoSwiss

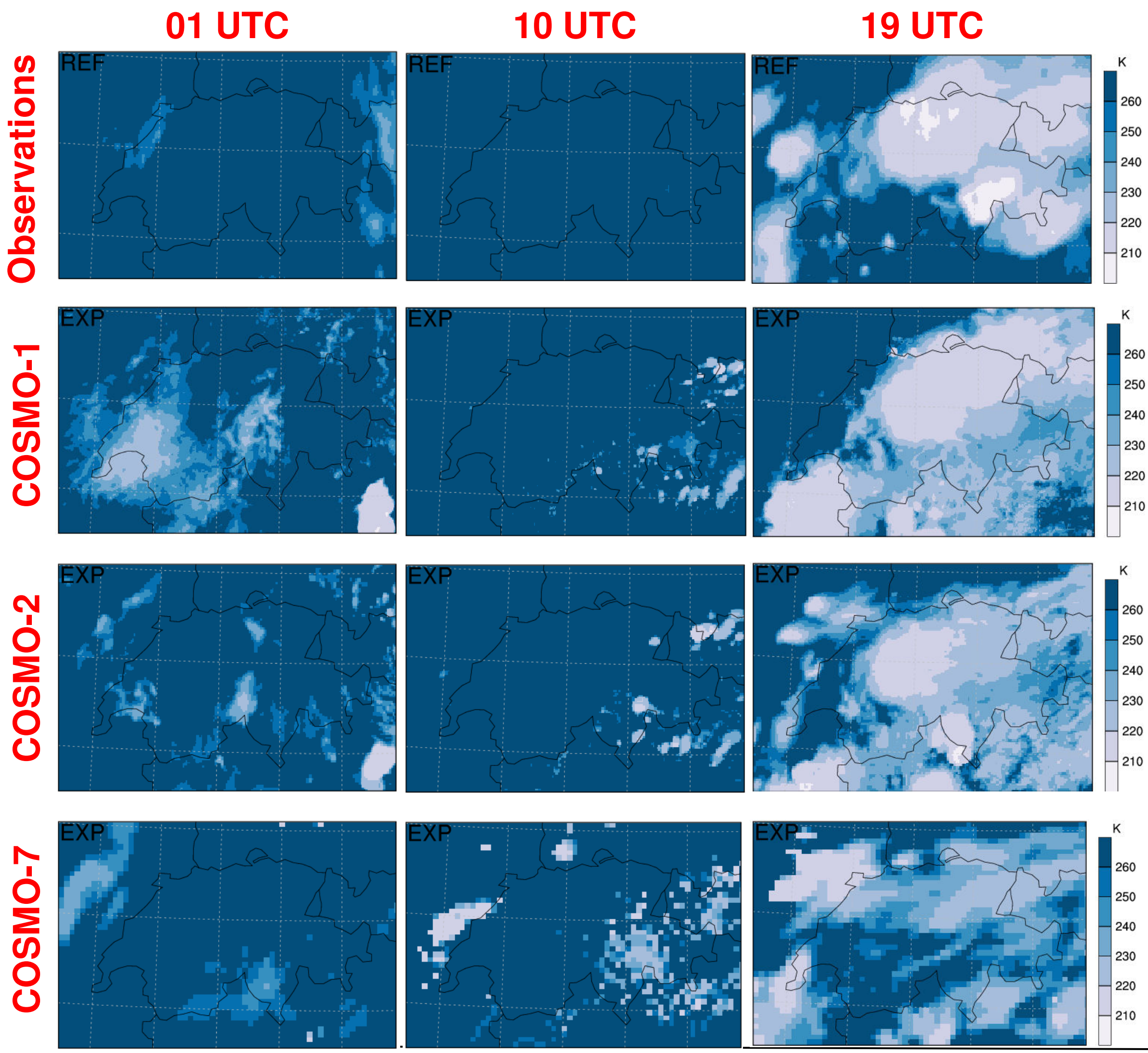
Investigation of the daily cycle of convection in Summer 2014

Goal: insight of the diurnal convection in high spatial and temporal resolution over the Alps with neighbourhood verification

- **Observations (measurements):**
  - **METEOSAT-8 data:** infrared 10.8μm channel of MSG SEVIRI -> **brightness temperature** (BT): detection of clouds in contrast to warm emission by the earth surface
    pixel resolution: 5 km

- **Models (COSMO-1/ -2, /-7): 00 UTC forecasts up to +24h**
  - **Brightness temperature**: **LMSynSat product** that produces synthetic satellite images (from NWP-SAF; RTTOV version 7)

Case study: Brightness temperature



12 June 2014

strong convective activity in the evening

FSS 12 June 2014 19 UTC COSMO-1

precipitation

Fractions skill score – FSS

Spatial scale (km)	0.1	0.2	0.5	1	2	5	10	20
99.0	0.60	0.57	0.51	0.47	**0.55**	**0.54**	0.34	0.06
59.4	0.53	0.50	0.43	0.37	0.40	0.31	0.14	0.02
33.0	0.44	0.41	0.34	0.26	0.22	0.11	0.02	0.00
19.8	0.40	0.36	0.30	0.22	0.16	0.05	0.00	0.00
6.6	0.33	0.30	0.25	0.16	0.09	0.00	0.00	0.00
2.2	0.30	0.27	0.22	0.14	0.07	0.00	0.00	0.00
1.1	0.28	0.26	0.21	0.13	0.07	0.00	0.00	0.00

threshold [mm/h]

brightness temperature

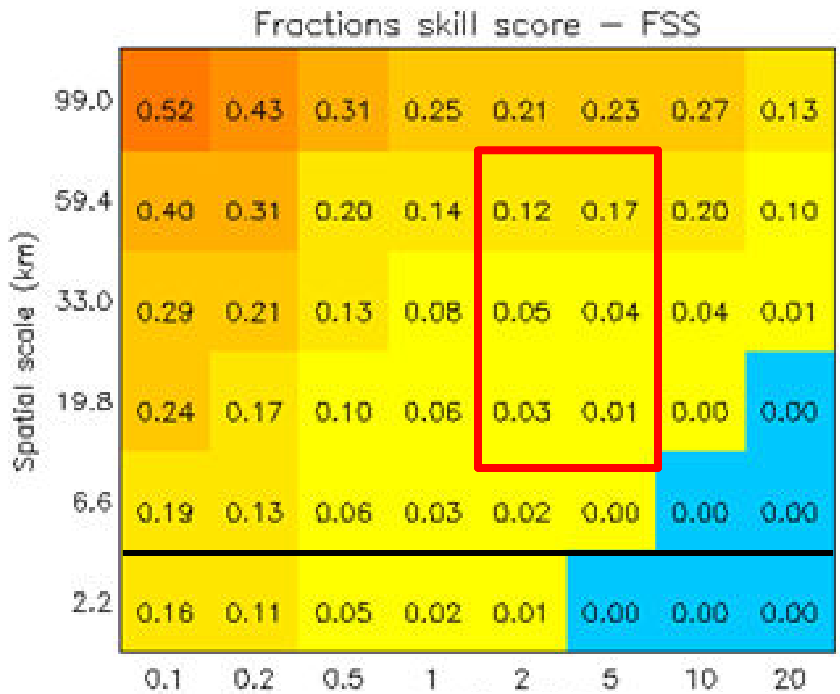
Fractions skill score – FSS

Spatial scale (km)	210	220	230	240	250	260
99.0	0.00	**0.87**	**0.86**	**0.84**	**0.86**	**0.87**
59.4	0.00	0.78	**0.79**	**0.79**	**0.81**	**0.83**
33.0	0.00	0.69	0.73	**0.76**	**0.78**	**0.80**
19.8	0.00	0.63	0.69	0.73	**0.76**	**0.78**
6.6	0.00	0.56	0.64	0.70	**0.73**	**0.75**
2.2	0.00	0.54	0.62	0.68	0.71	**0.74**
1.1	0.00	0.53	0.61	0.67	0.71	**0.73**

threshold [K]

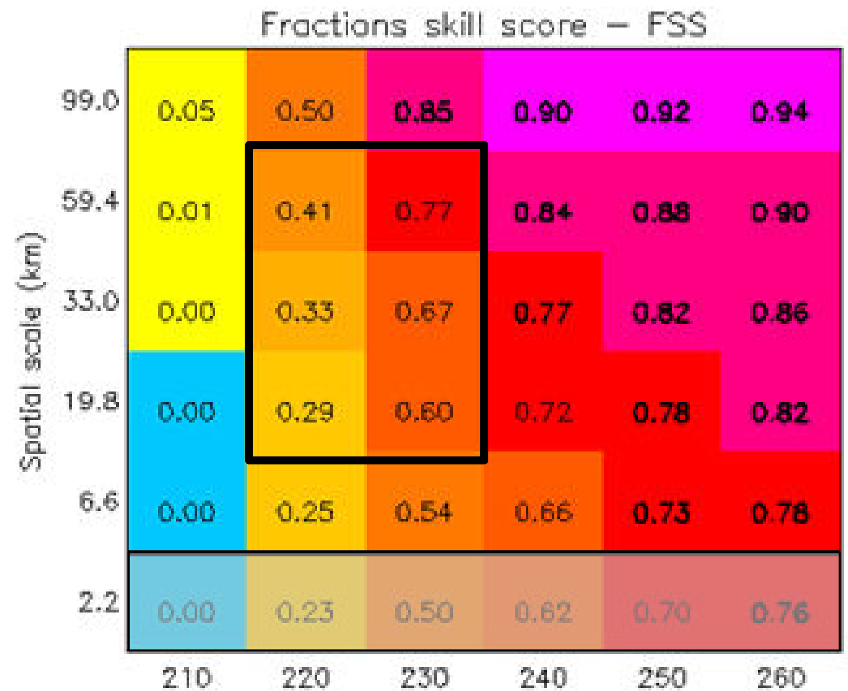
FSS 12 June 2014 19 UTC COSMO-2

precipitation



threshold [mm/h]

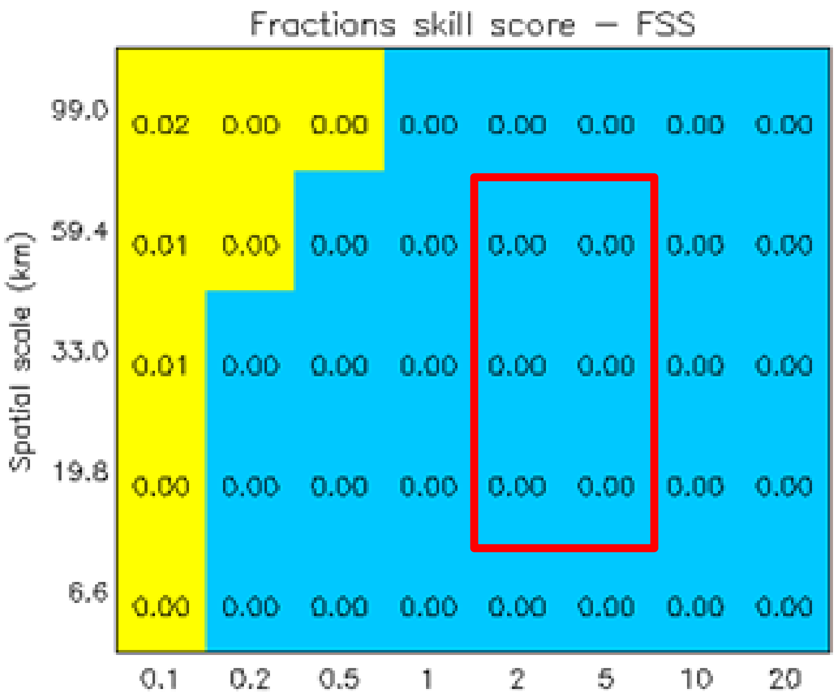
brightness temperature



threshold [K]

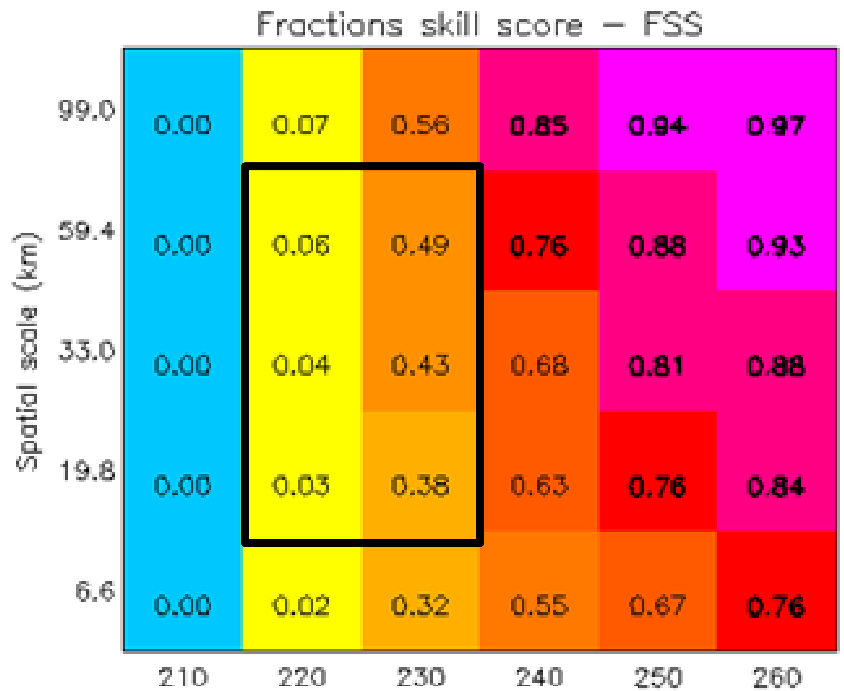
FSS 12 June 2014 19 UTC COSMO-7

precipitation



threshold [mm/h]

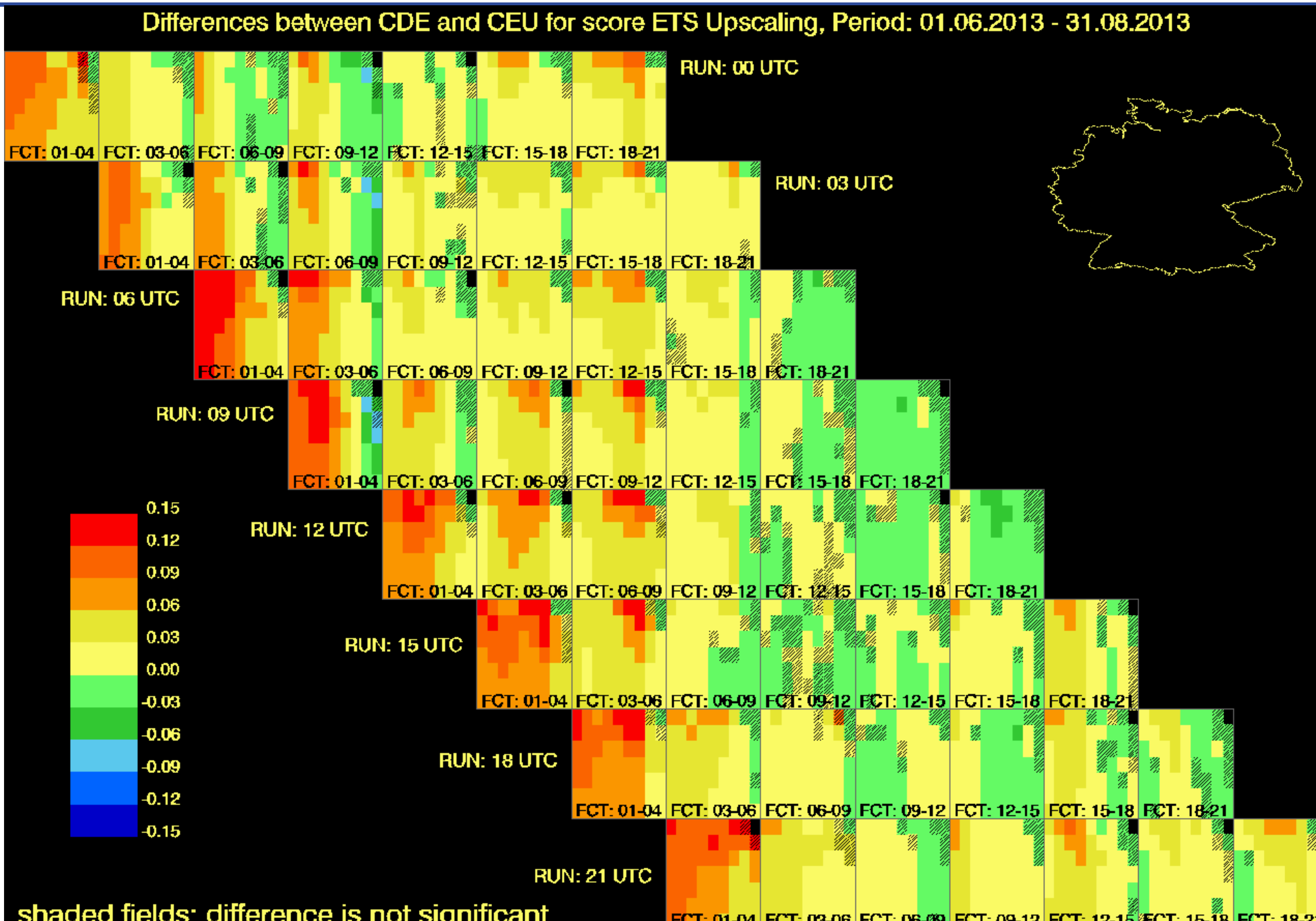
brightness temperature



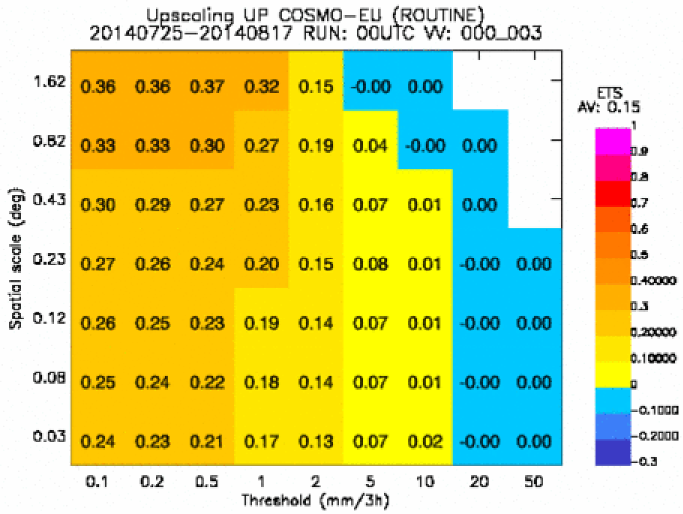
threshold [K]

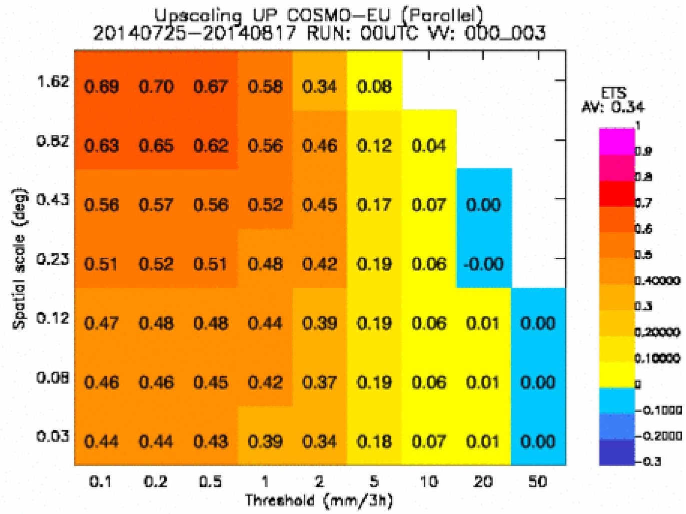
Fuzzy verification: CDE against CEU (Summer 2013)

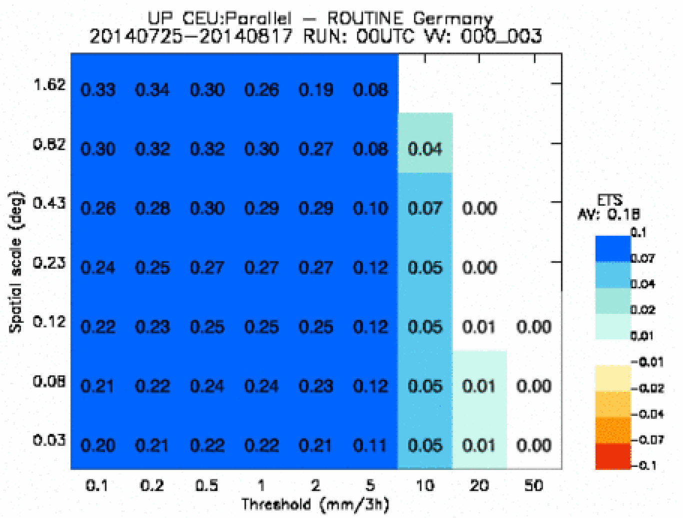
Deutscher Wetterdienst
Wetter und Klima aus einer Hand



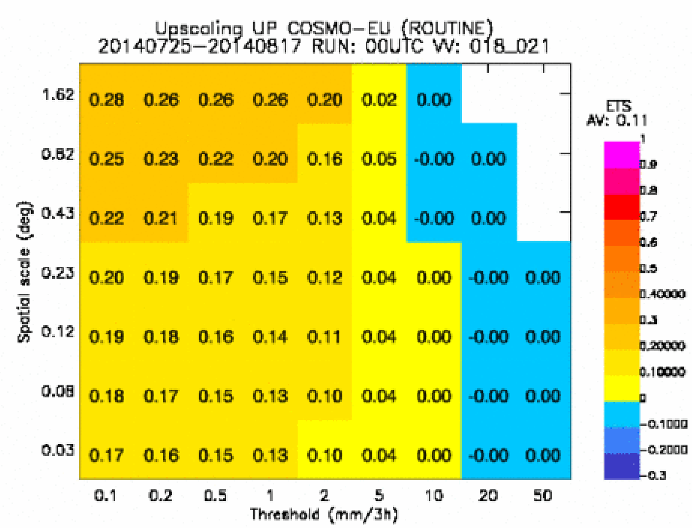
Fuzzy verification: CEU(with OPERA data) against CEU (operational) (vv=00-03)

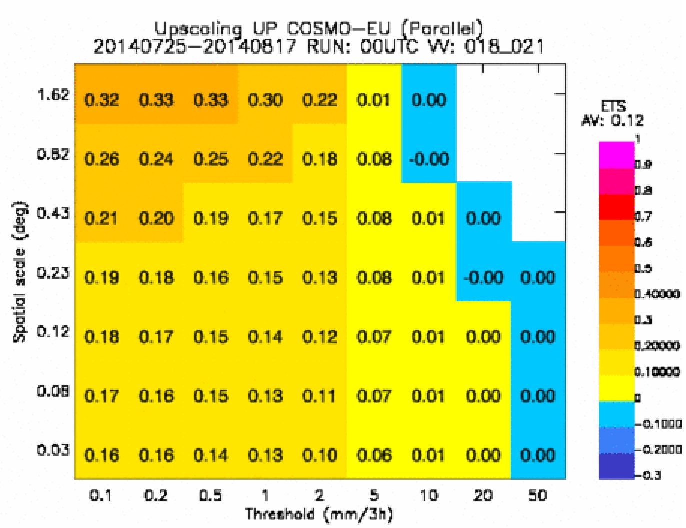


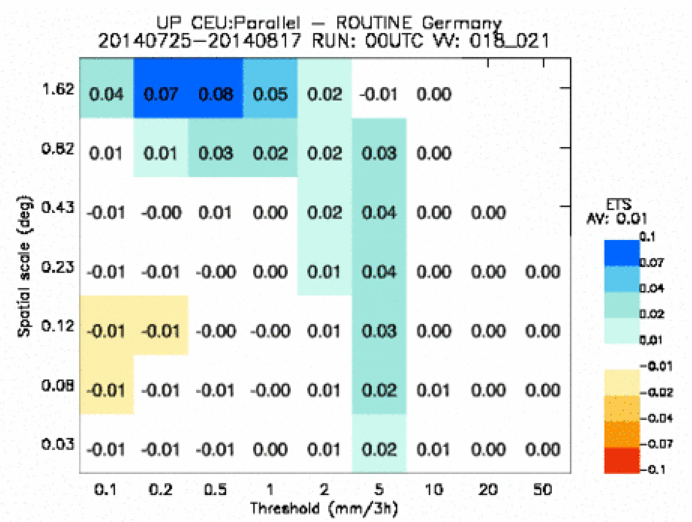




Fuzzy verification: CEU(with OPERA data) against CEU (operational) (vv=18-21)







VAST project - status

Fuzzy verification toolbox development

Naima Vela, Elena Oberto, Maria Stefania Tesini

September 8, 2014

COSMO PP VERSUS - Project Plan - Task 4: VAST

Overview

- Introduction of additional statistical techniques in VERSUS
- Needed for high resolution forecast and observation data (neighborhood methods)
- The main goal of the activity is the integration or adaptation of pre-existing packages
  - Beth Ebert Fuzzy Verification Toolbox
- The pre-processing operations will be performed by the LIBSIM software
  - So the system will be able to receive GRIB (1 and 2) as input both for observation and forecast

Results: FSS (with the indication of the skilful scales)

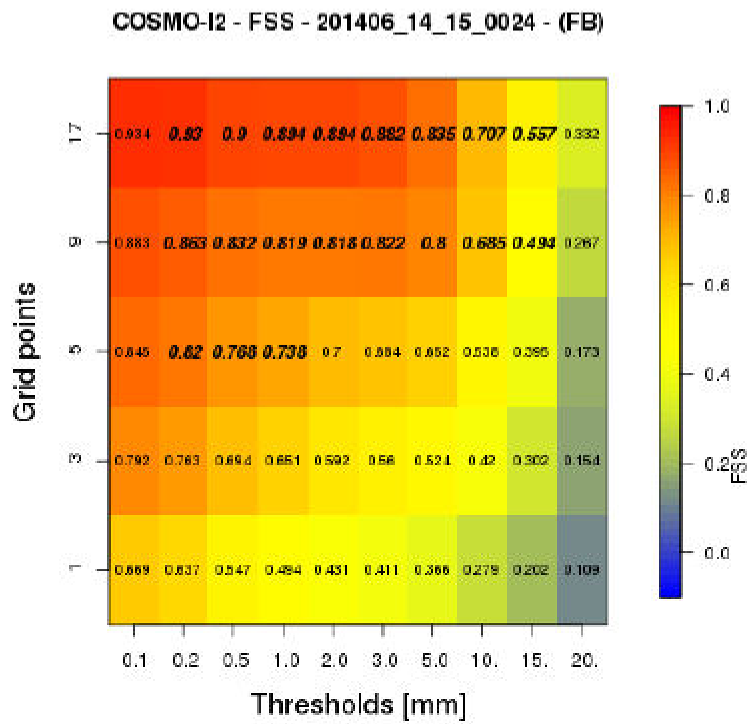


Figure: FSS, first 24 hours of forecast
.

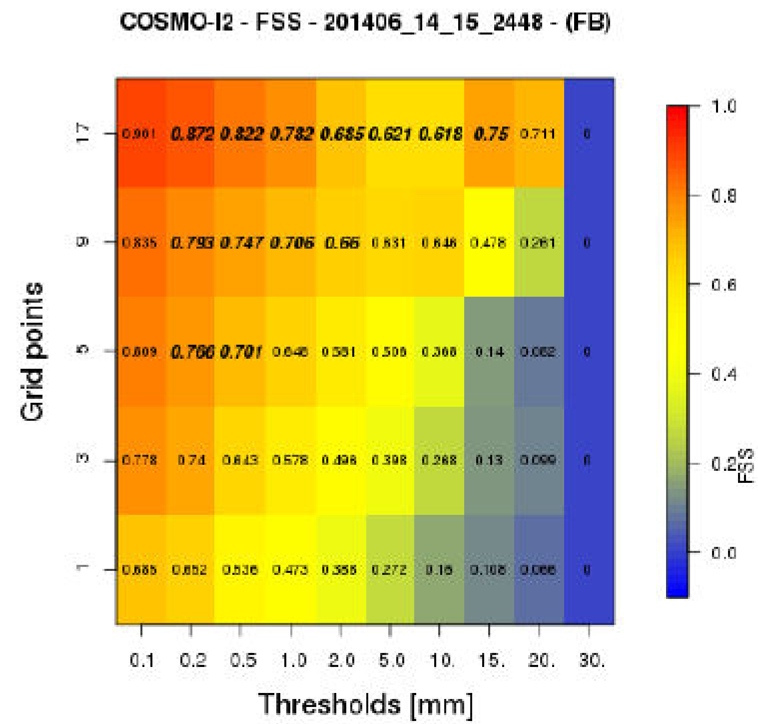


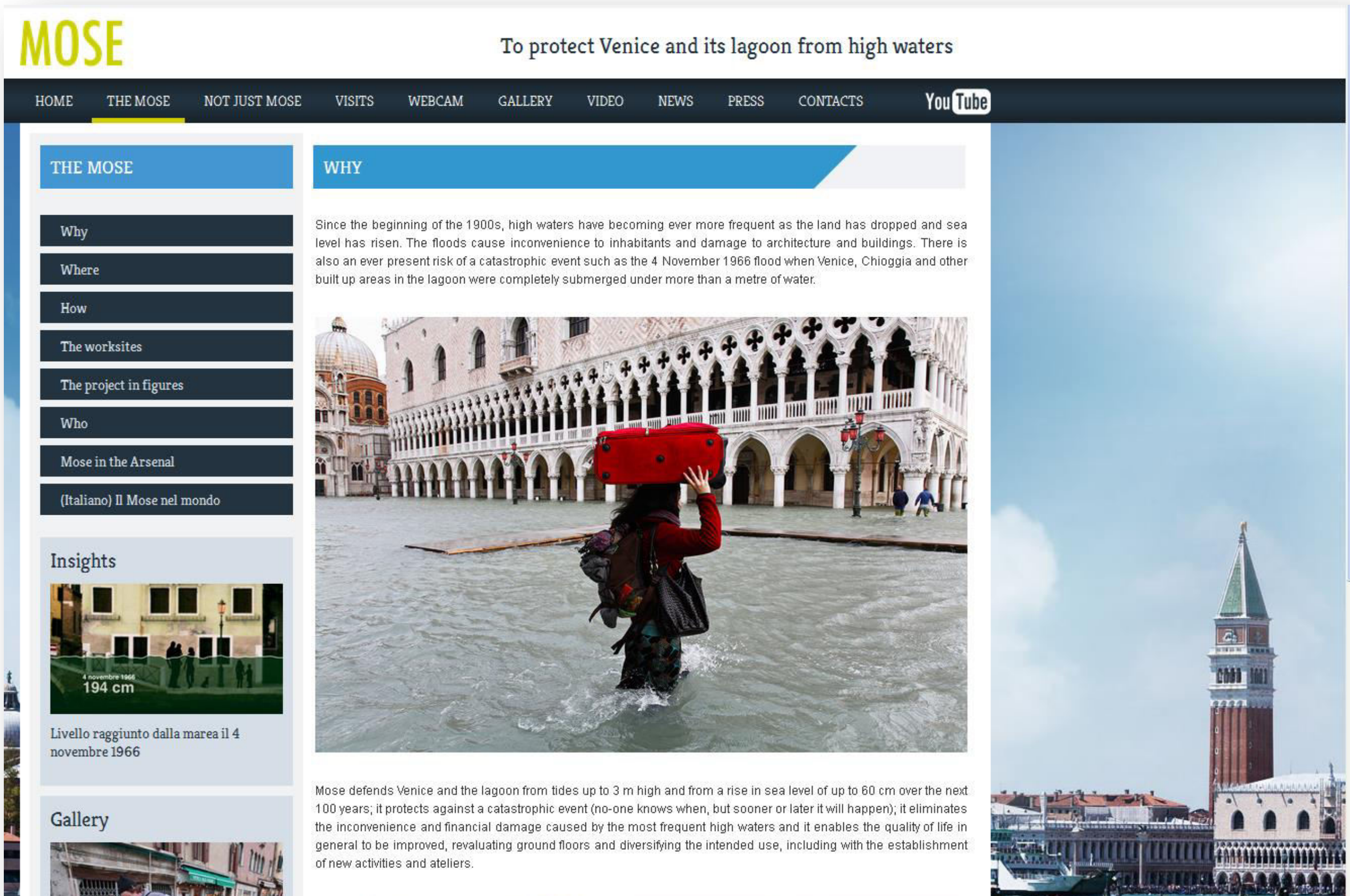
Figure: FSS, second 24 hours of forecast

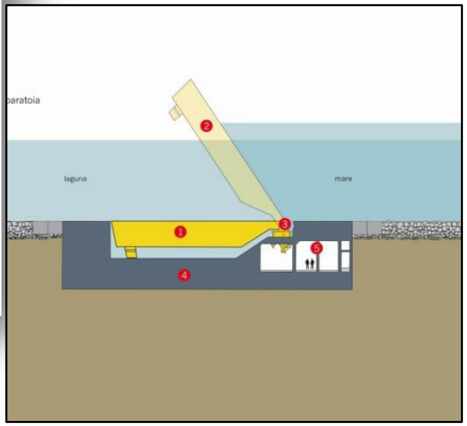
SPECIAL VERIFICATION APPLICATION: OPERATIONAL WIND VERIFICATION OVER NORTH ADRIATIC SEA

Maria Stefania Tesini

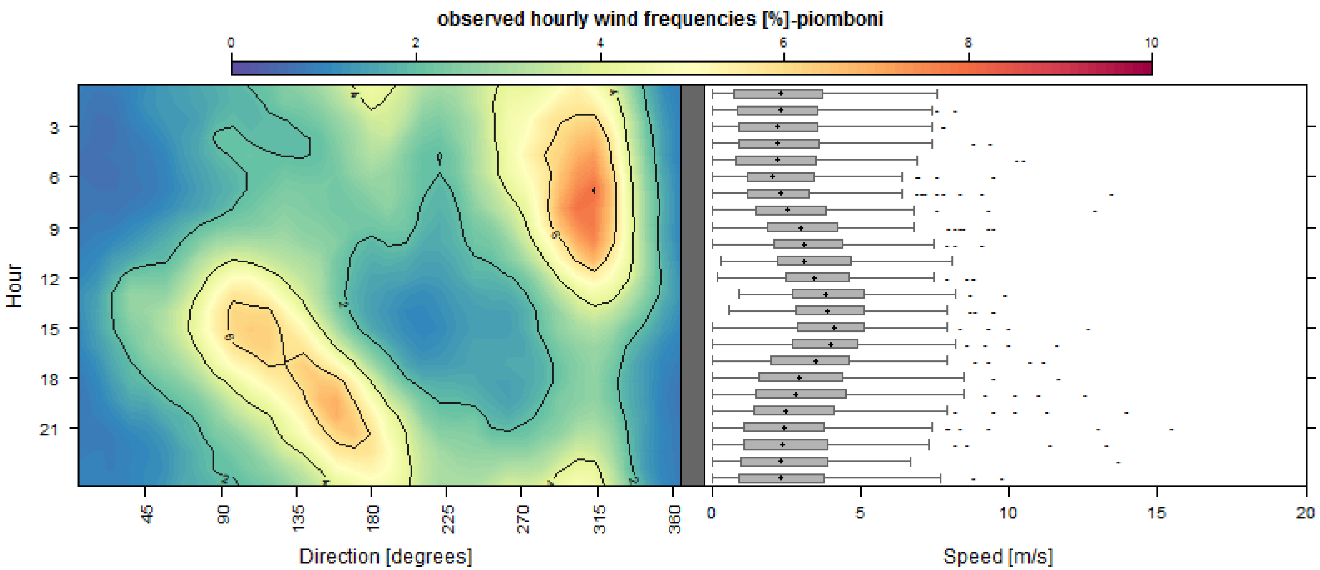
16th COSMO General Meeting
8-11 September 2014, Eretria (Greece)

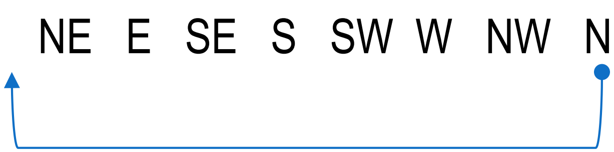
One of the applications: meteorological support for MOSE

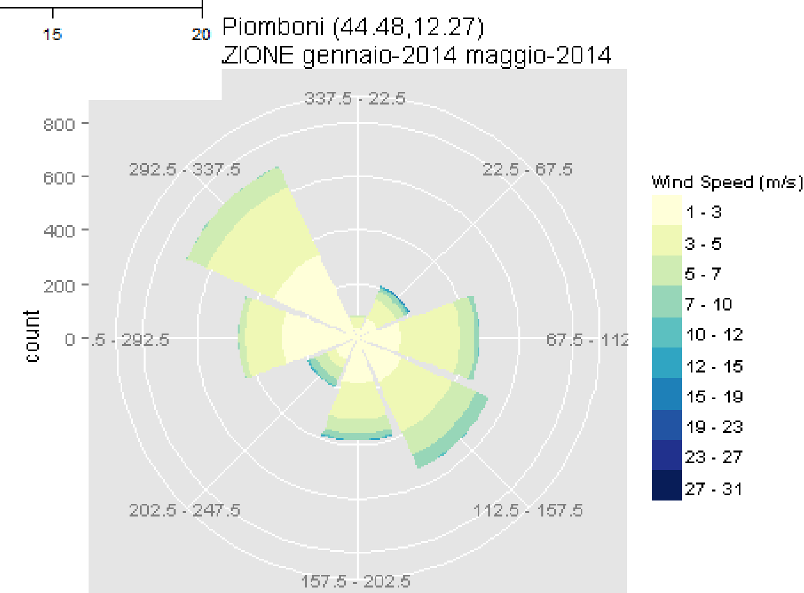




Visualizing diurnal wind climatology







In the next graphs some examples of these plots are presented to illustrate some of the properties that can be deduced from this type of plot. Results are not complete to define the quality of the models

Overview of Italian verification

Elena Oberto
Maria Stefania Tesini
Naima Vela
Antonio Troisi
Angela Celozzi

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 2014 – Eretria (Greece)

C/L ratio model (Richardson)

You are a decision maker

Expense matrix

	Event occurs	Event does not occur
Action taken	*C*	*C*
Action not taken	*L*	*0*

No forecast info

- $E_{always} = C$
- $E_{never} = sL$, s= climatological base rate

$E_{always} < E_{never}$ → action

$E_{always} > E_{never}$ → no action

Optimal strategy=mean expense=minimise losses

$E_{climate} = \min(C, sL)$

Perfect forecast

$E_{perfect} = sC$

V of forecast system $= (E_{climate} - E_{forecast})/(E_{climate} - E_{perfect})$

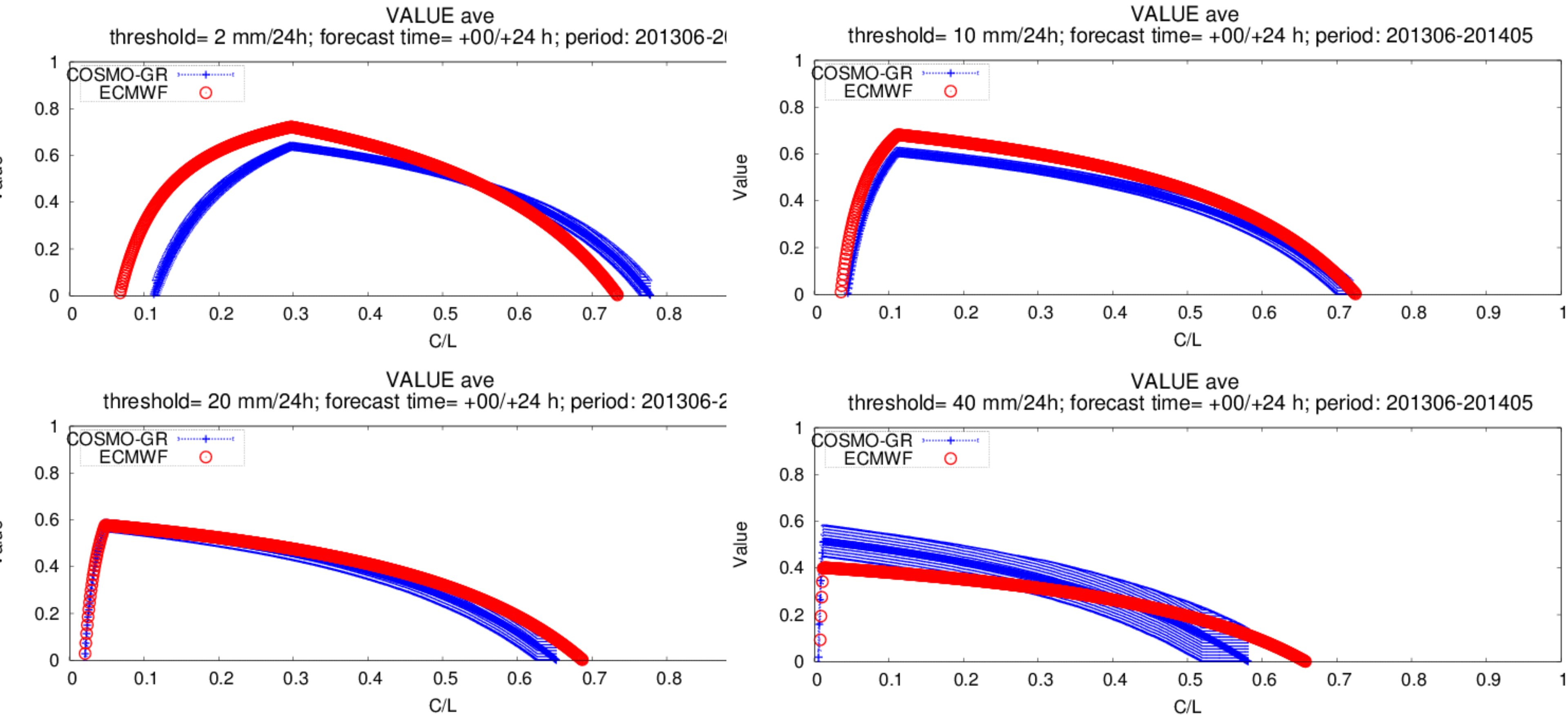
A maximum value is when the system perfectly forecasts the future. If V >0 the decision maker will gain economic benefit by using forecast info in addition to climatology.

$V_{relative} = [\min(C/L,s) - F(1-s)C/L + Hs(1 - C/L) - s]/[\min(C/L,s) - sC/L]$, s=a+c (base rate)

V relative depends on quality of system, observed base rate and user's C/L

ECMWF/COSMOGR

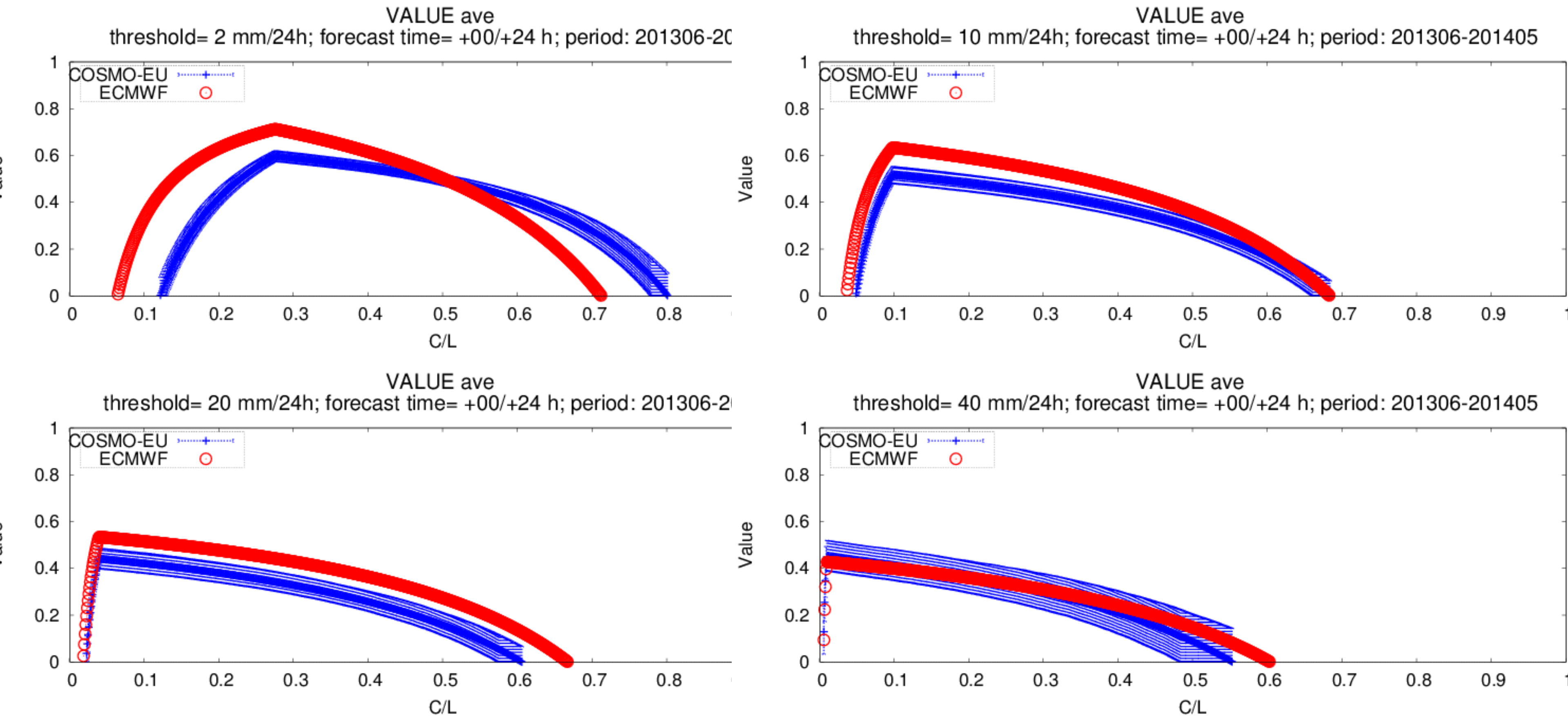
- For low thresholds → higher value for ecm for low C/L ratio, higher value COSMOGR for high C/L ratio
- For medium thresholds → higher value for ecm
- For high thresholds → equivalent or best COSMOGR



Thanks to M.Milelli for the graphs

ECMWF/COSMOEU

- For low thresholds → higher value for ecm for low C/L ratio, higher value COSMOEU for high C/L ratio
- For medium thresholds → higher value for ecm
- For high thresholds → equivalent



Thanks to M.Milelli for the graphs

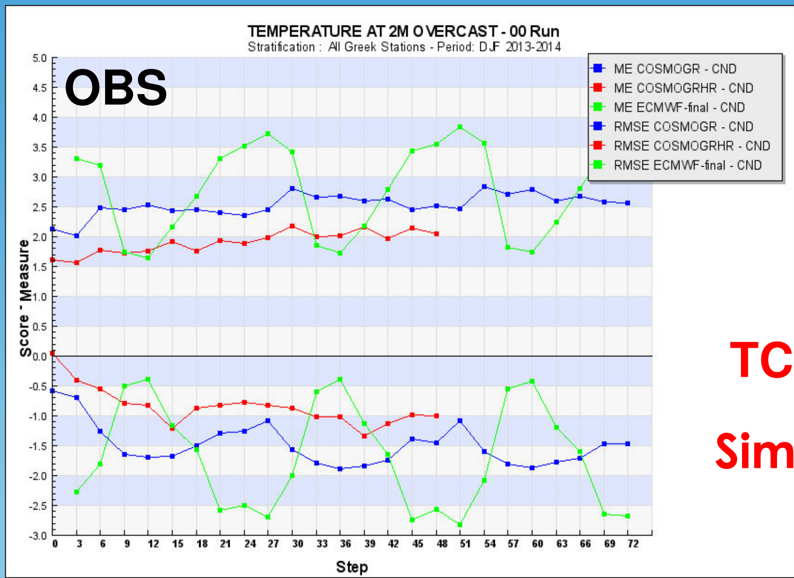
COSMO Verification for the region of Sochi-2014 Olympics

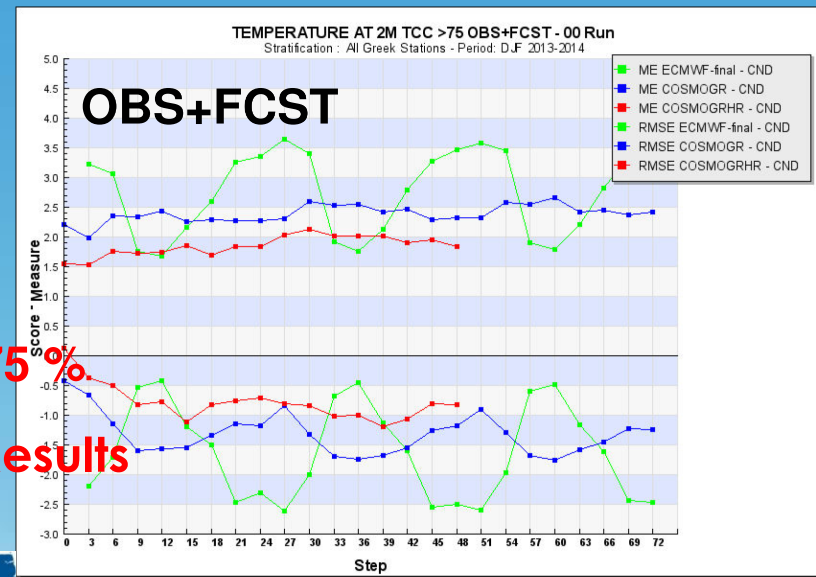
Comparison of forecasts from COSMO versions of different scales

A.Bundel, A. Kirsanov, A. Muraviev, G. Rivin, I. Rozinkina, M. Shatunova, D. Kiktev, M. Tsyrulnikov, D. Blinov, and many others

Roshydromet

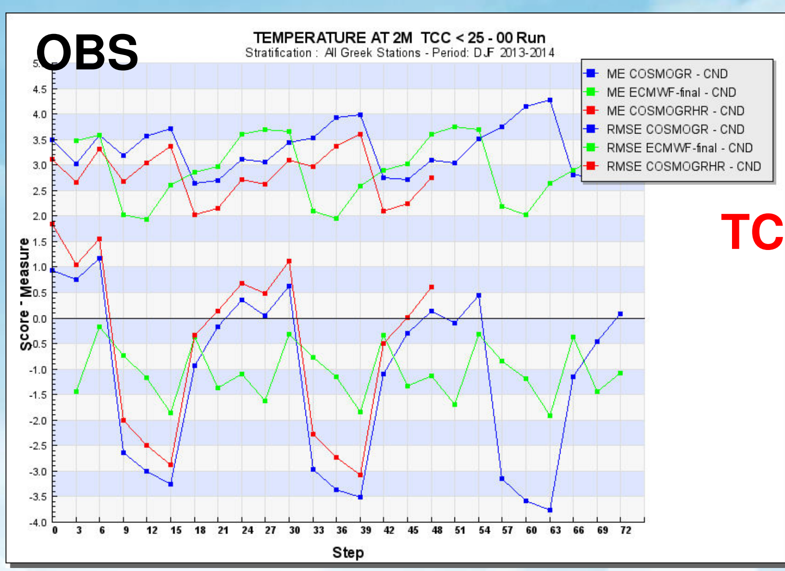
Conditional T2m based on OBS and OBS+FCST COSMO/COSMOHR/ ECMWF



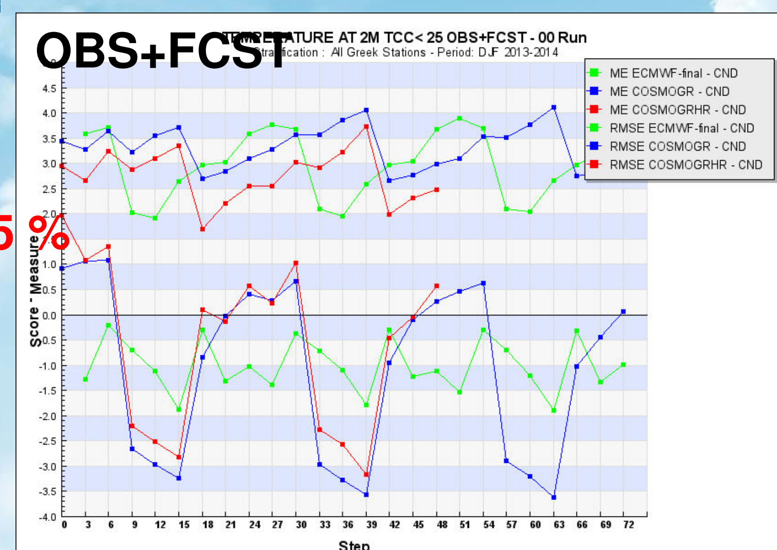


TCC > 75 %

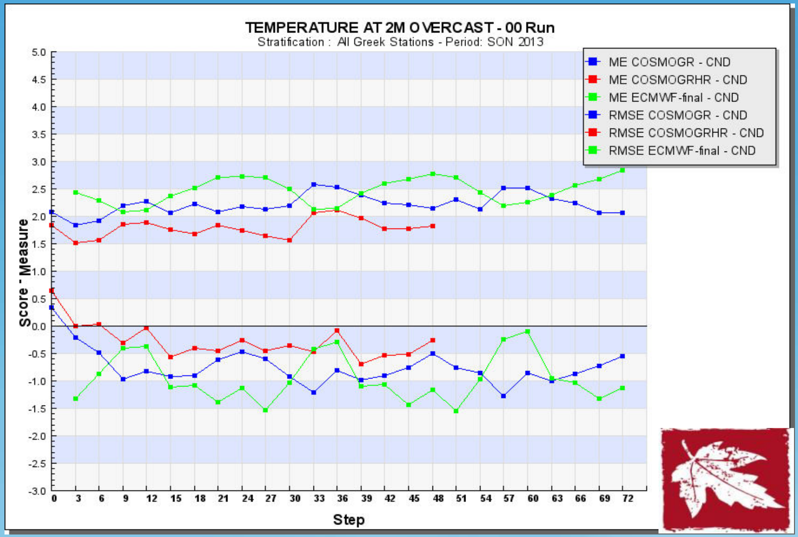
Similar Results

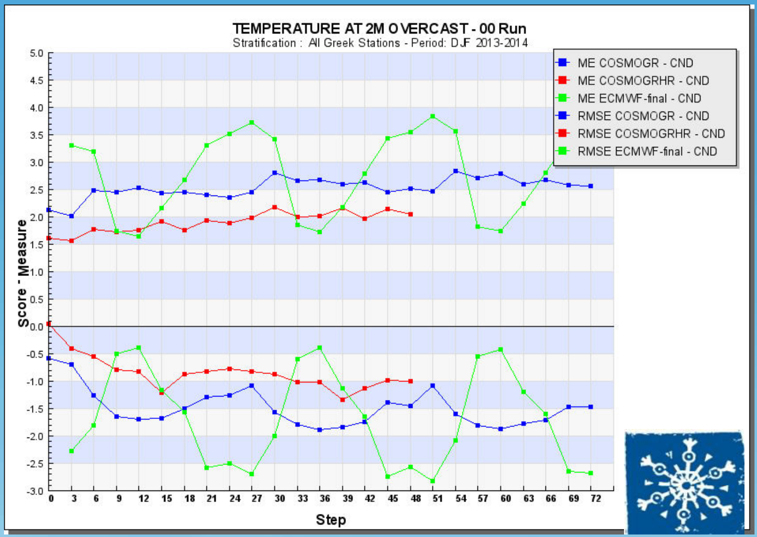


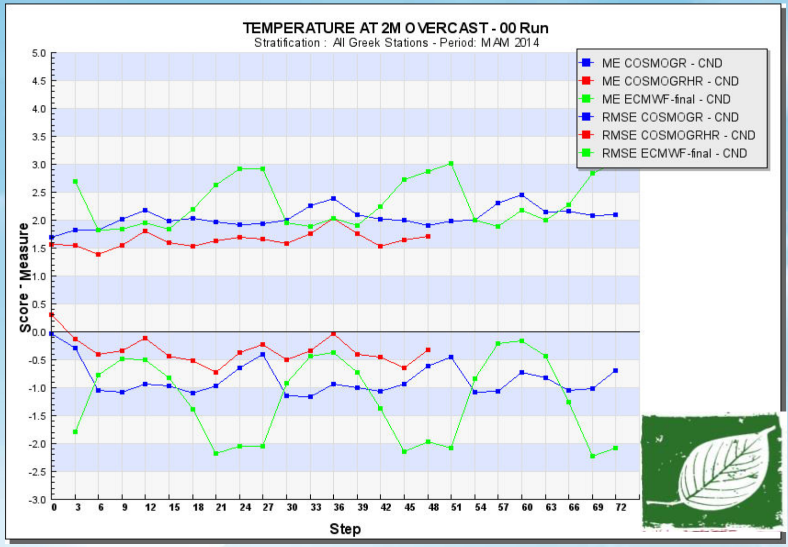
TCC < 25 %



Conditional T2m CC > 75% (obs) ME/RMSE COSMO/COSMOHR/ ECMWF

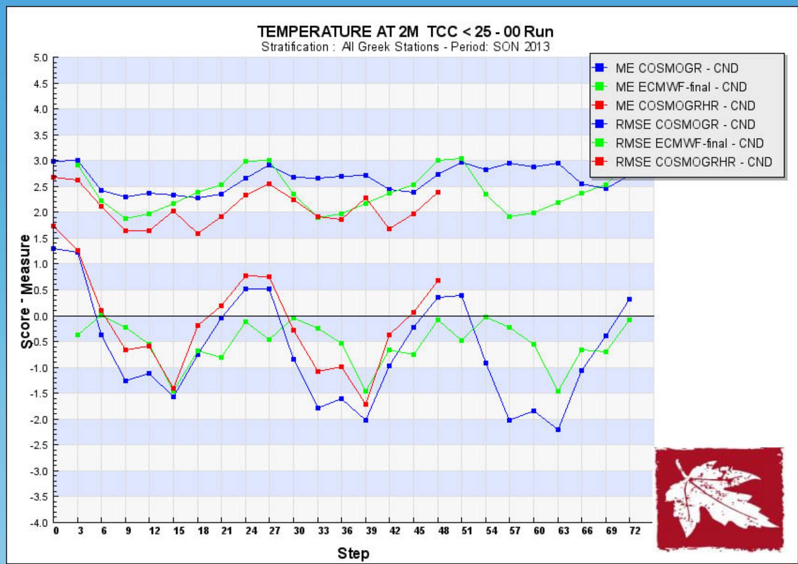


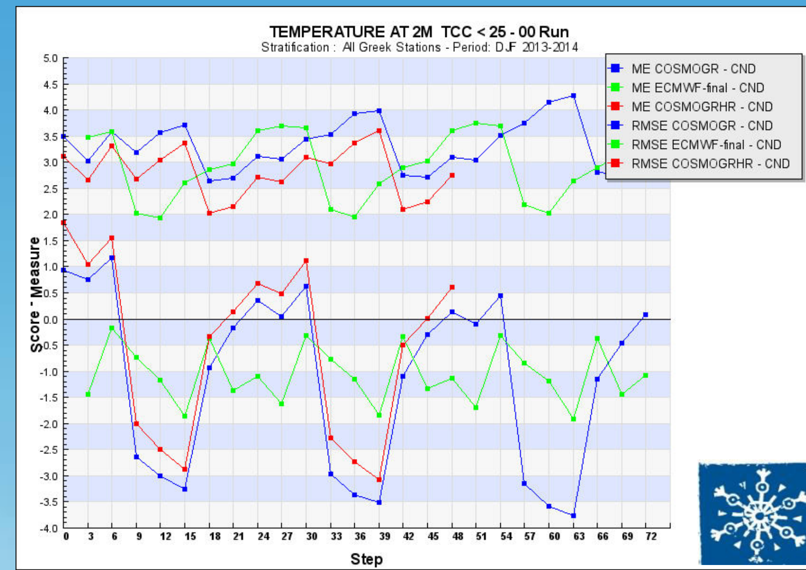


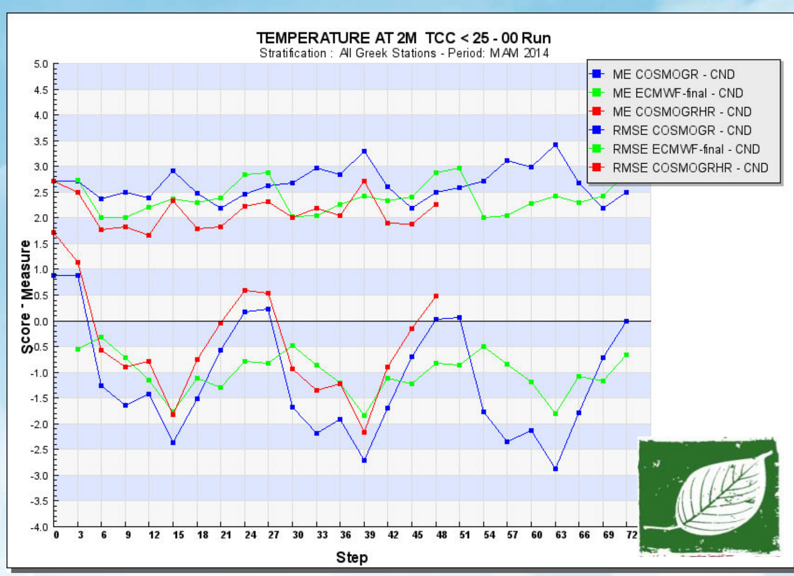


Overcast: Better performance for COSMOHR. Small diurnal variation for COSMO, ECMWF diurnal variation and underestimation at night. (winter and spring)

Conditional T2m CC < 25% (obs) ME/RMSE COSMO/COSMOHR/ ECMWF







Sky clear: COSMO ME diurnal variation, with daytime underestimation. Winter Hysteresis RMSE similar to T2m.

The methodologies

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Weather elements – standard verification

Using VERSUS
(operative verification)
•Upper air
•Surface Parameters
Precipitation:
dataset → synop stations
Method →6h/12h/24h averaged cumulated forecasted precipitation values over 15 km radius, 6h/12h/24h cumulated observed precipitation values over station point

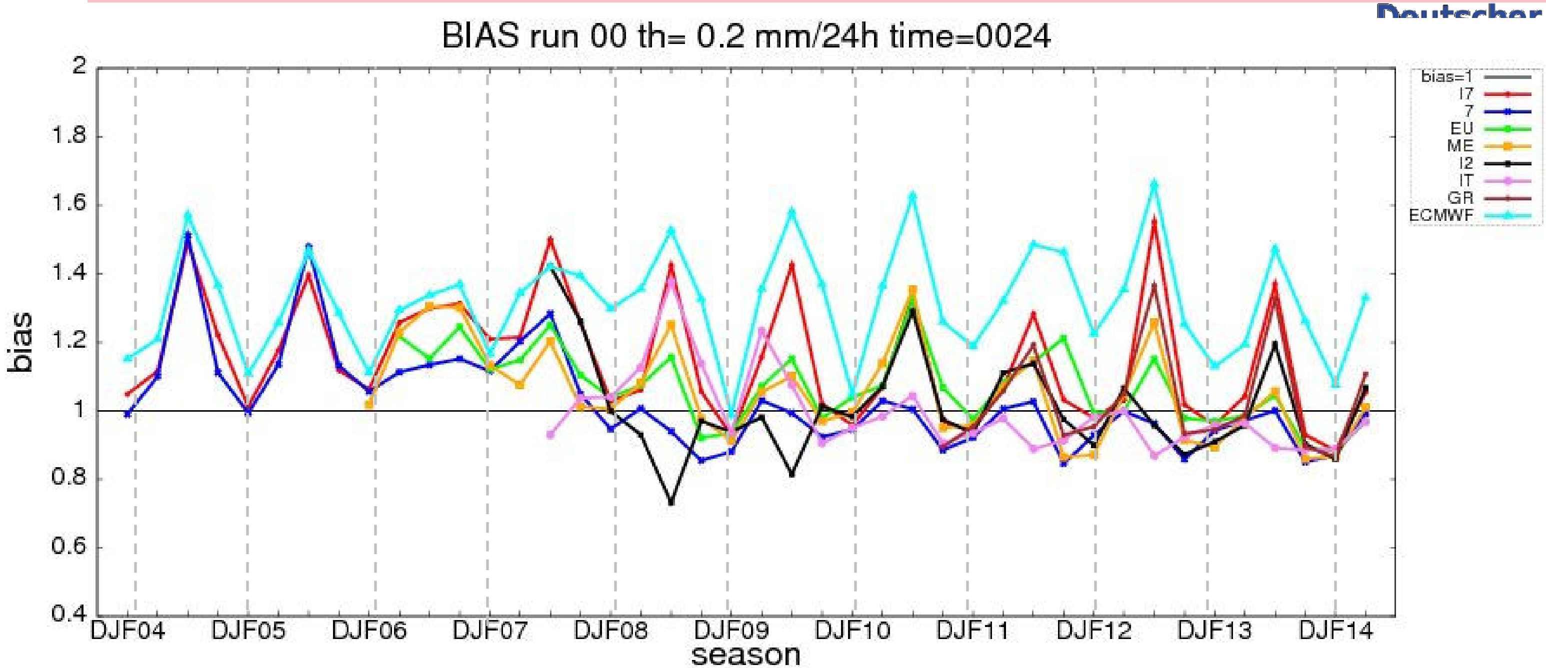
Precipitation- high resolution network

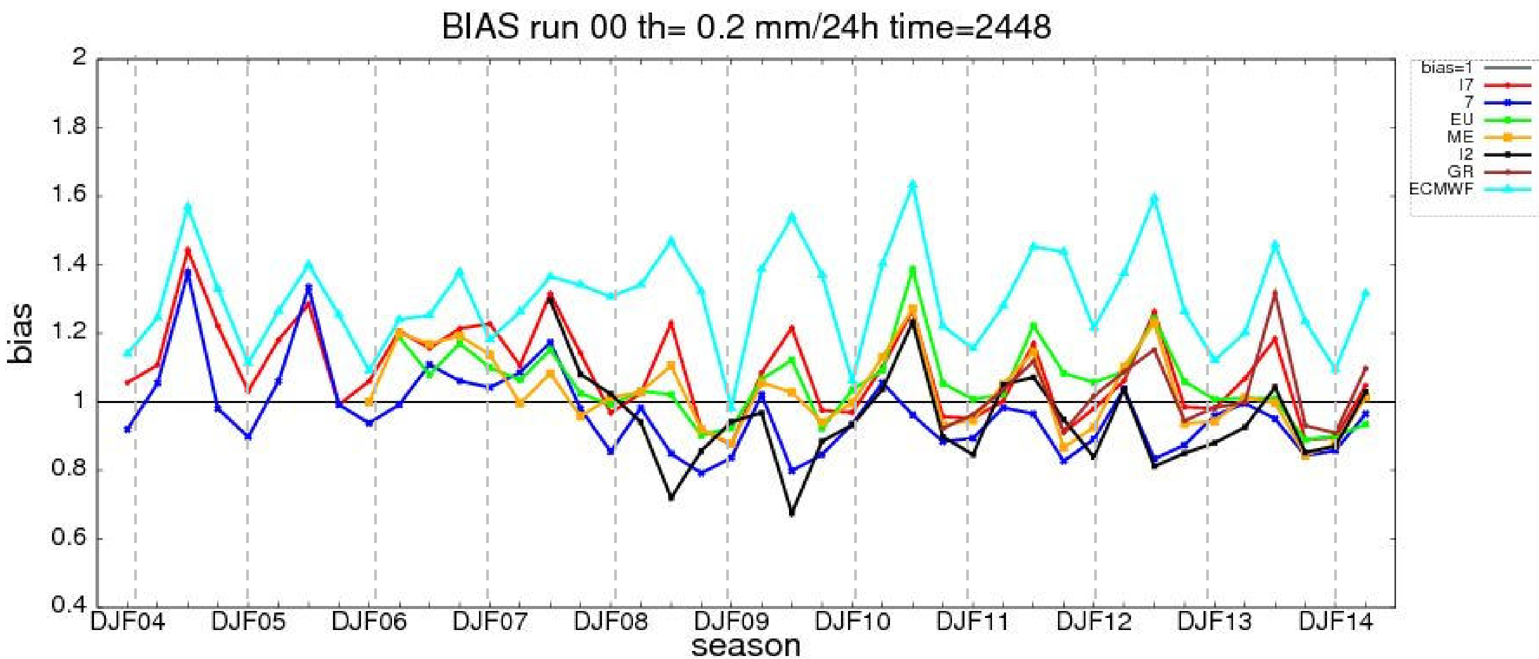
•Common area → Italy
•Dataset → high res raingauges
•Method → 24h/6h averaged cumulated precipitation or maximum values (both observed and forecasted) over 90 meteo-hydrological basins

LONG TREND PRECIPITATION with high resolution stations

Deutscher Wetterdienst
...ma aus einer Hand

LOW THRESHOLDS





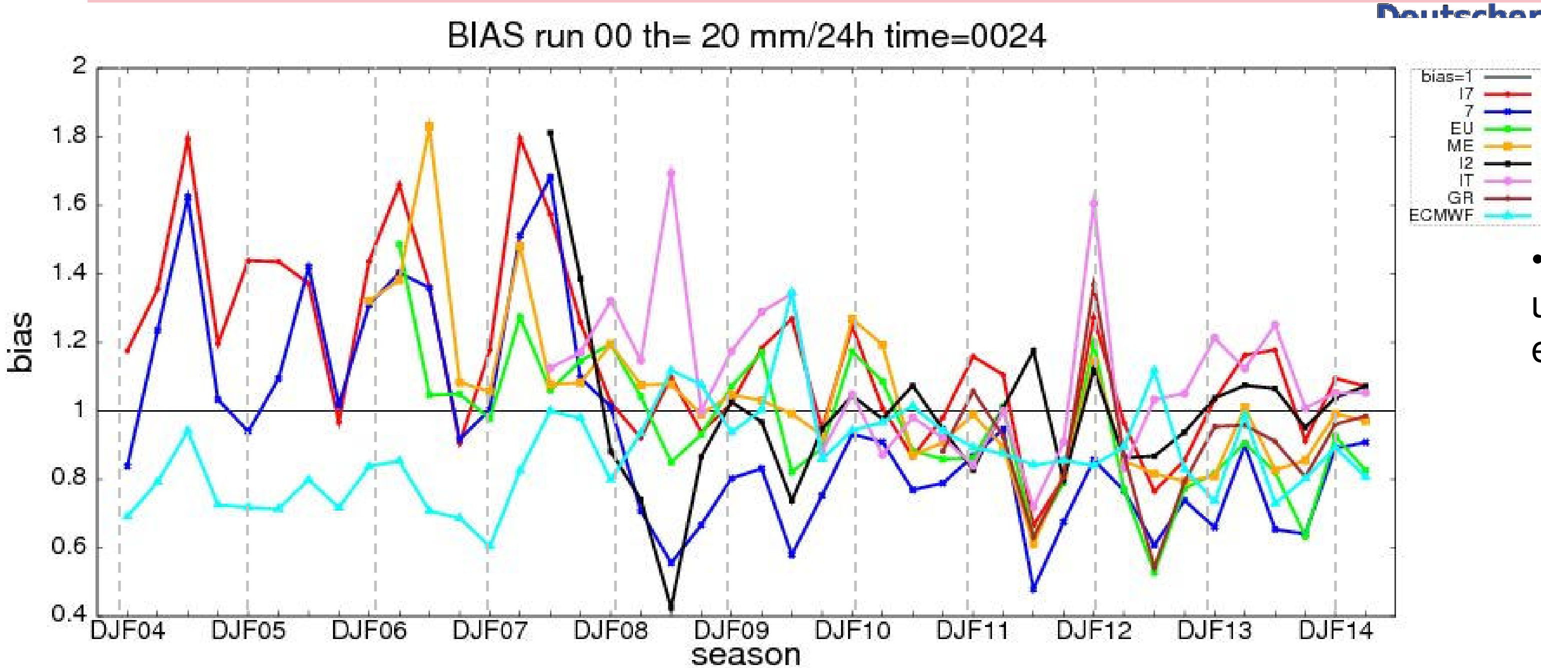
- Ecmwf overestimation
- Summer overestimation

LONG TREND PRECIPITATION with high resolution stations

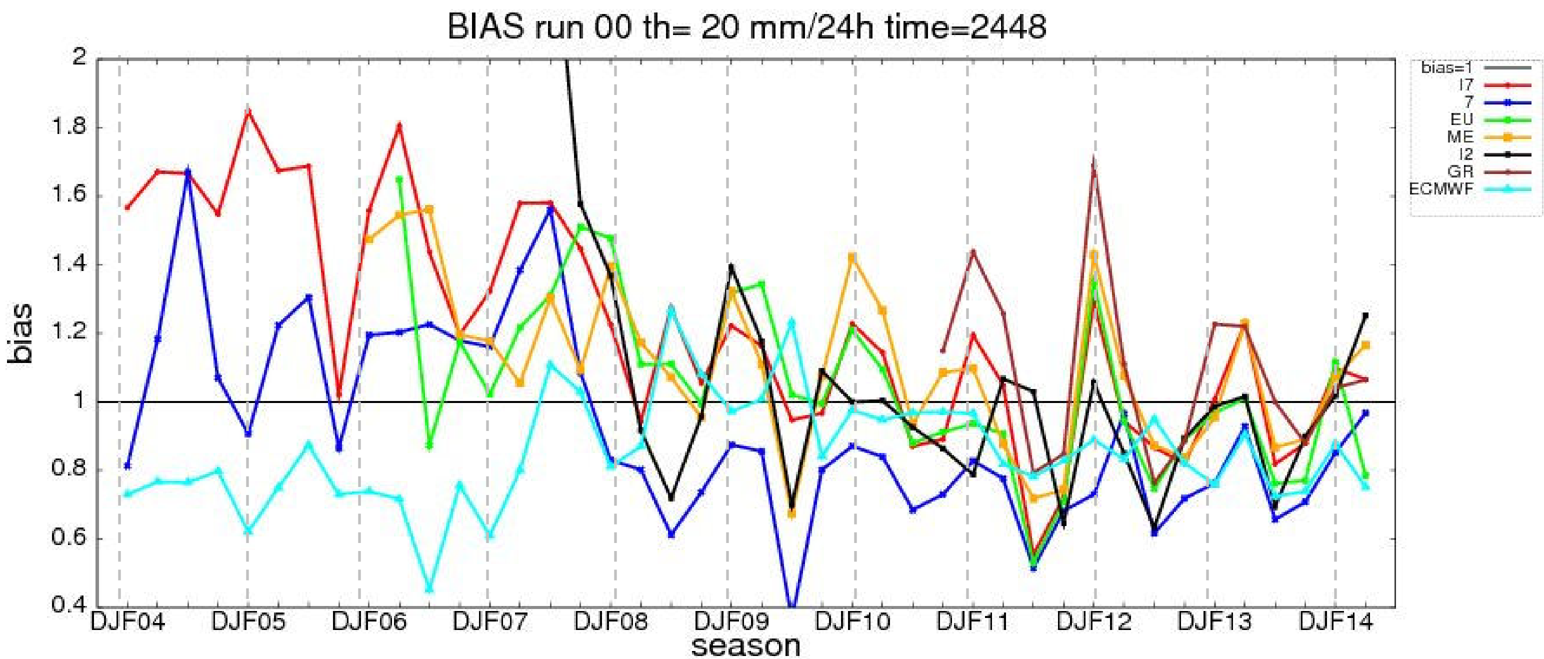
Deutscher Wetterdienst

...ma aus einer Hand

HIGH THRESHOLDS



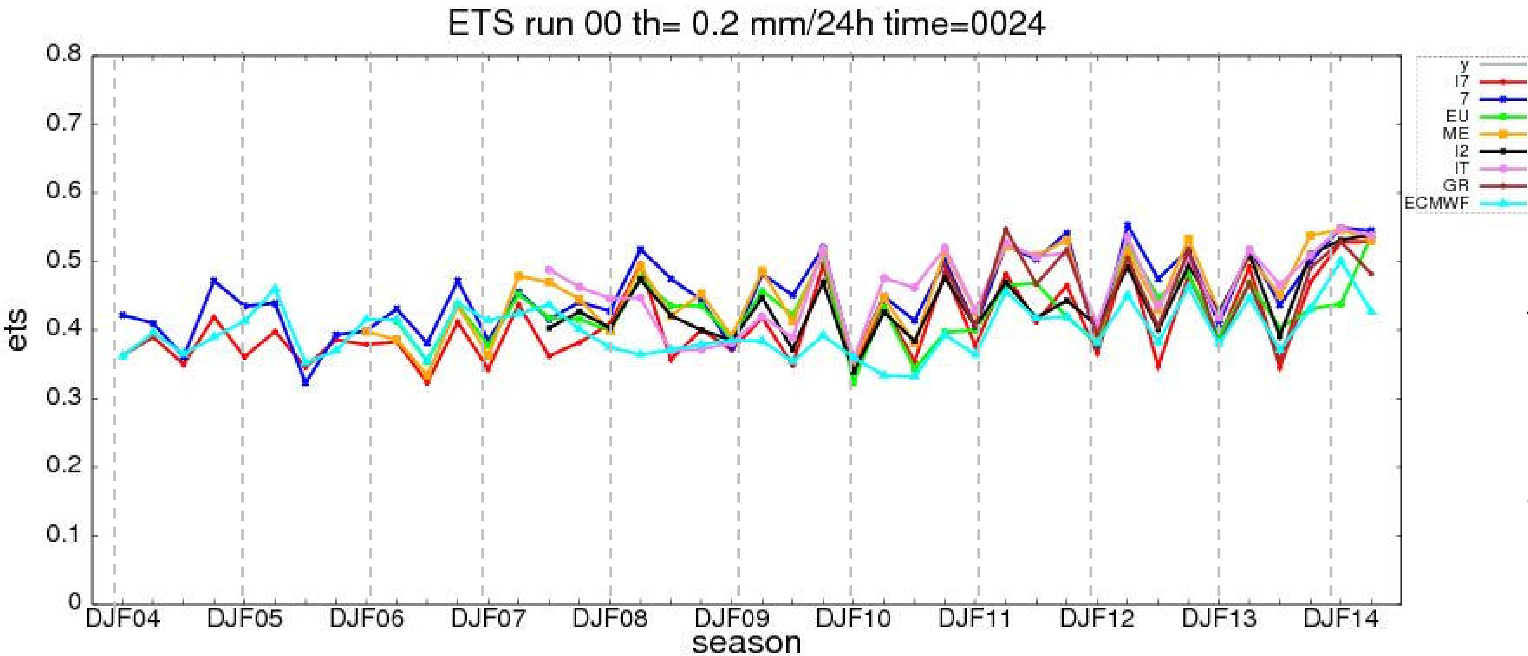
- General underestimation, especially 7, EU

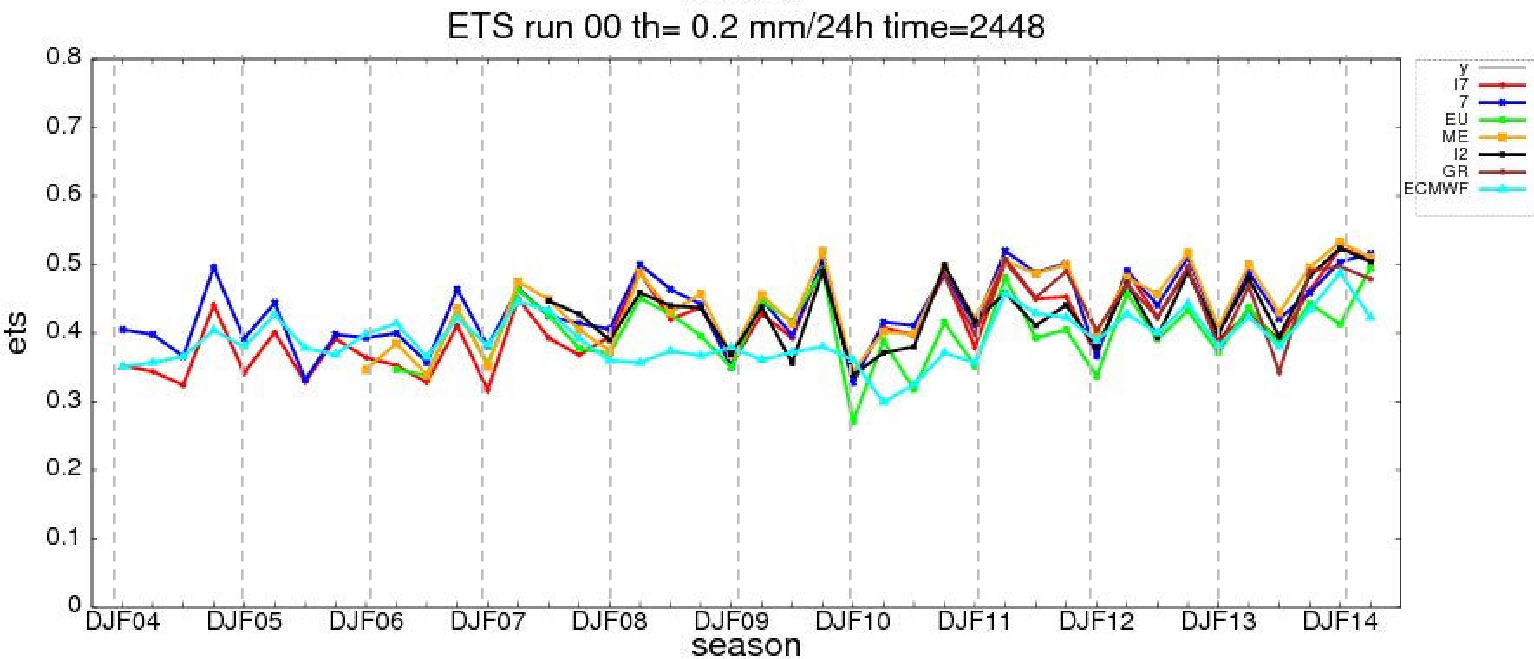


LONG TREND PRECIPITATION with high resolution stations

Deutscher Wetterdienst
...ma aus einer Hand

LOW THRESHOLDS



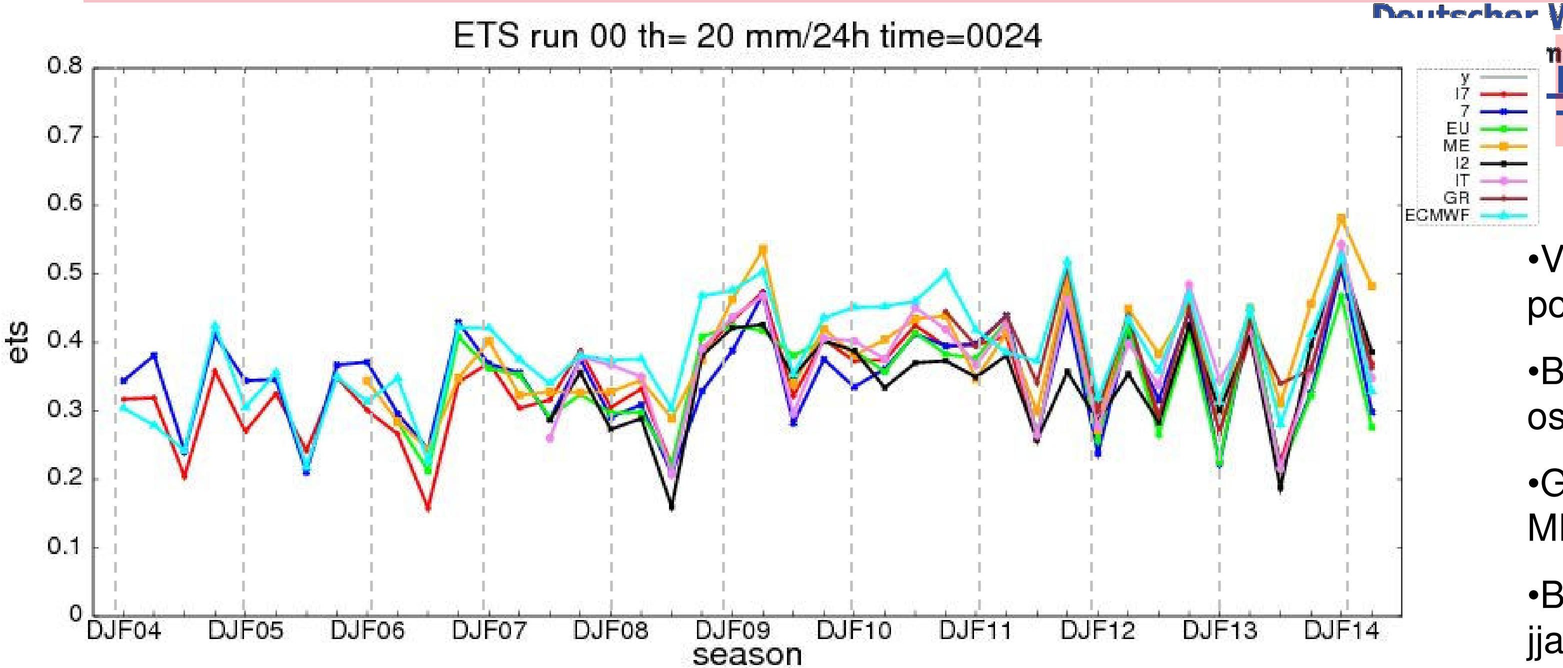


- Very slightly positive/steady trend
- Good ME,7
- Big seasonal oscillation

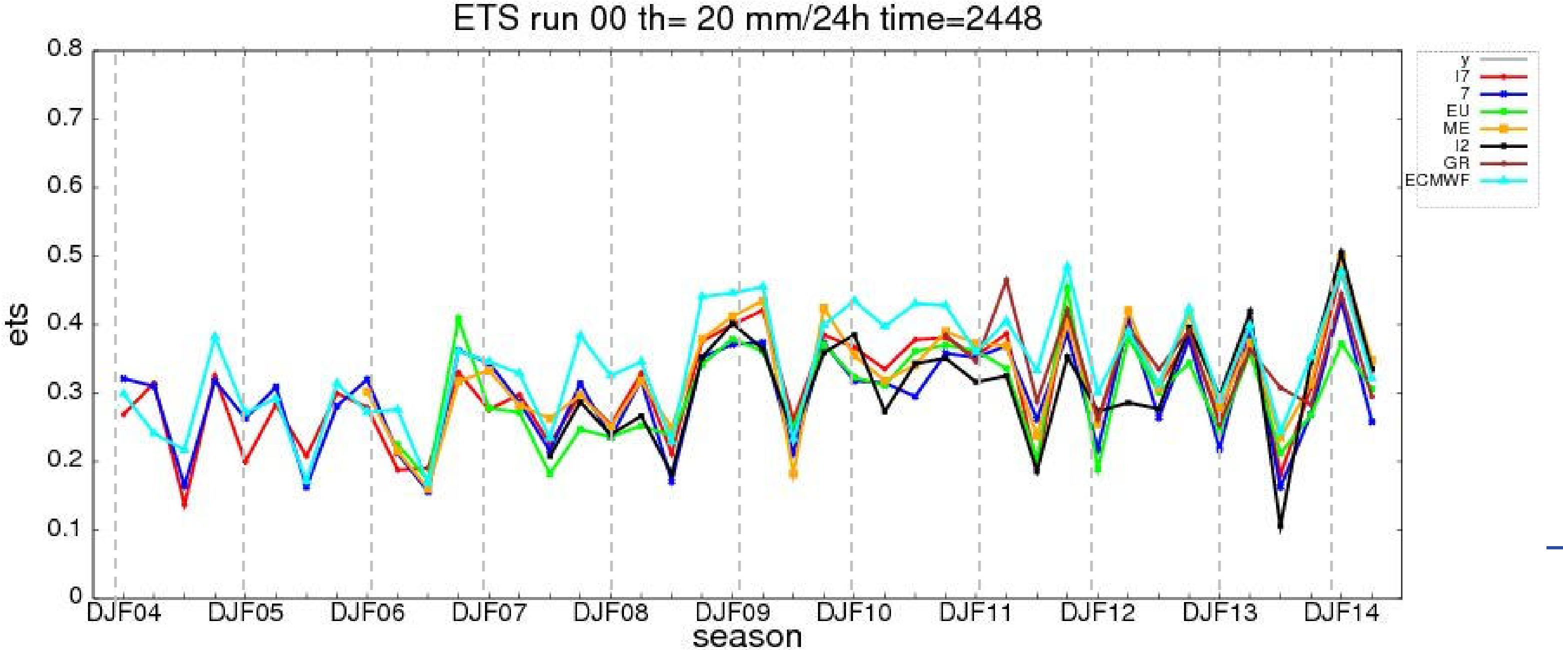
LONG TREND PRECIPITATION with high resolution stations

Deutscher Wetterdienst
…na aus einer Hand

HIGH THRESHOLDS



- •Very slightly positive trend
- •Big seasonal oscillation
- •Good ets for ME
- •Big worsening jja13



Common Plot Report Preparation :WG5 Task 1.2

www.cosmo-model.org/verification/tasks/

Data provided seasonally by all countries (when available)

Responsible for Report Preparation (2013-14):

Joanna Linkowska

Analysis of Trend of last years was performed and will be presented based on CP data by Ulrich Damrath

WG5 Contributing Scientists

Ulrich Damrath, DWD
Francis Schubiger , MCH
Pirmin Kaufmann, MCH
Angela Celozzi, USAM
Adriano Raspanti, USAM
Antonio Troisi, 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
Pavel Khain, IMS
Alon Stivelman

Thank you all!