# Verification Overview

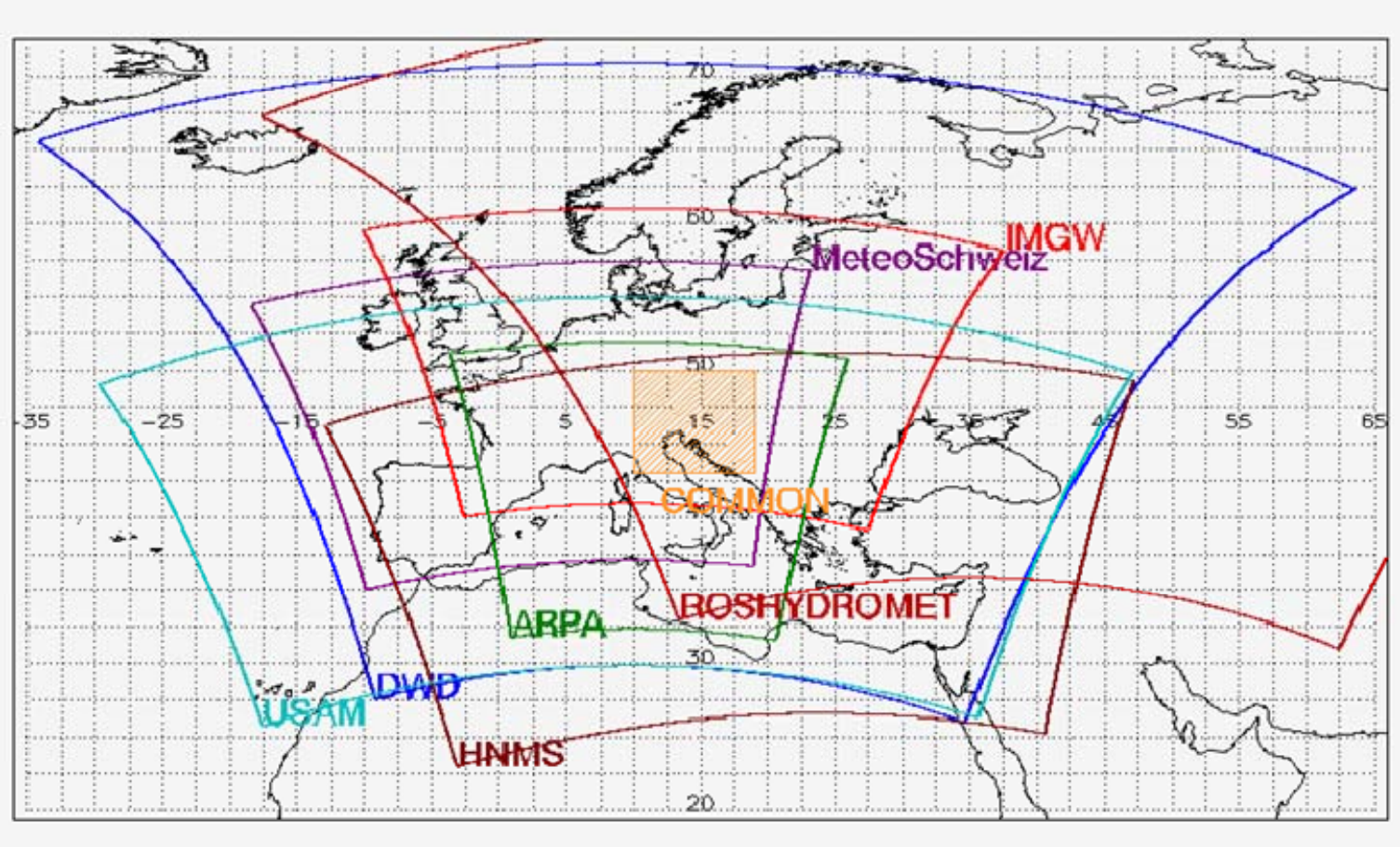
## (based on CP activity)

WG5

Dimitra Boucouvala & WG5

# VERIFICATION OVER THE COMMON AREA

COSMO-7(7km), COSMO-GR(4km), COSMO-I5(5km), COSMO-ME(5km), COSMO-PL(7km), COSMO-RU7(7km)
ICON-EU (6.5km), IFS (9km), ICON(13.5)




Annual reports and Seasonal analytics are located at the website:
http://cosmo-model.org/content/tasks/verification.priv/

# Scores for 00 UTC and 12 UTC Runs

## Continuous parameters (3h)

*Temperature at 2 m*
*Dew point temperature at 2 m*
*Pressure reduced to Mean Sea Level*
*Wind speed at 10 m*
*Total cloud cover*

$$\mathrm{ME} = \frac{1}{n}\sum_{k=1}^{n}(f_k - o_k) \qquad \mathrm{RMSE} = \sqrt{\left(\frac{\sum_{k=1}^{n}(f_k - o_k)^2}{n}\right)}$$

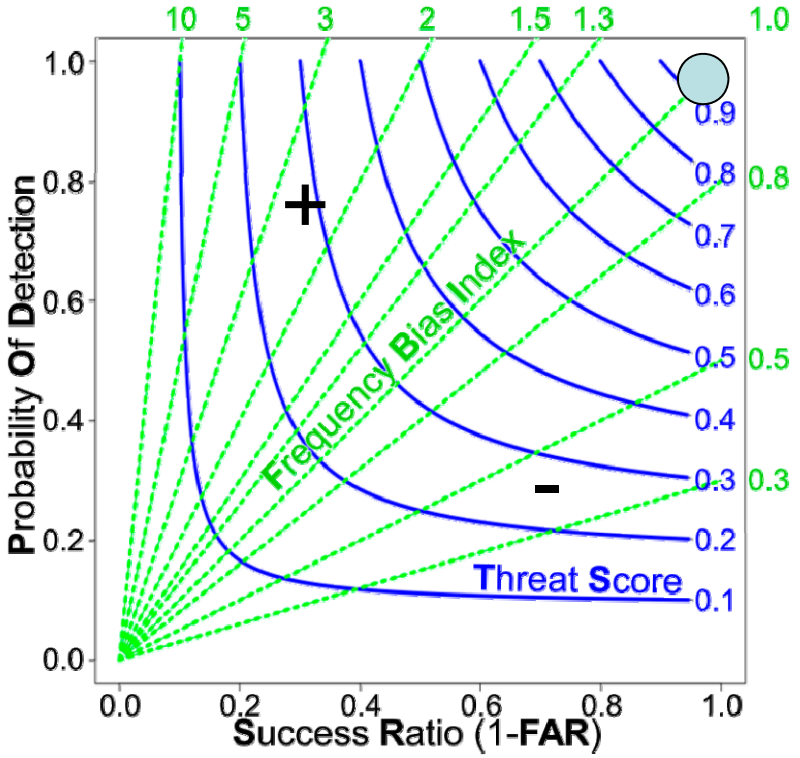
## Dichotomic parameters

*Total Precipitation in 6 hours*
*Total Precipitation in 24 hours*
*Total Cloud Cover (0-25, 25-75, 75-100%)*
*Wind Gust ( Threshold > 2.5, 15, 20)*

$$FBI = \frac{a+b}{a+c} \qquad TS = CSI = \frac{a}{(a+b+c)}$$

$$FAR = \frac{b}{(a+b)} \qquad POD = \frac{a}{(a+c)}$$

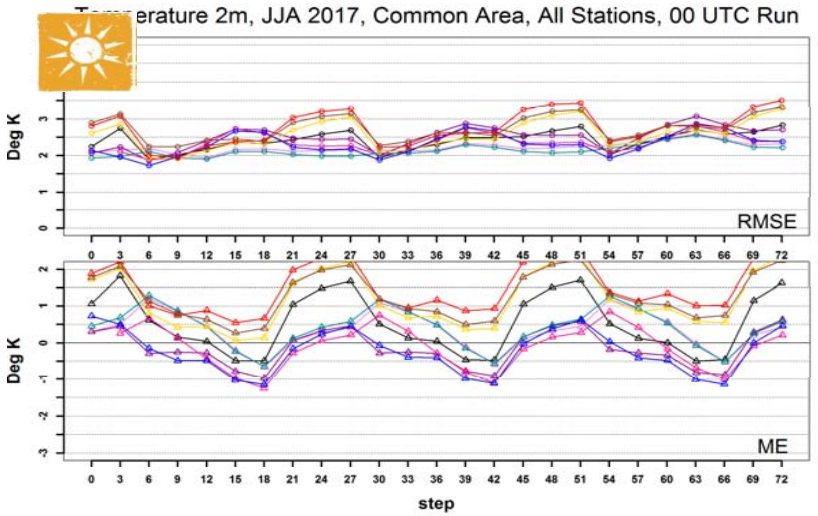
**Performance Diagram**

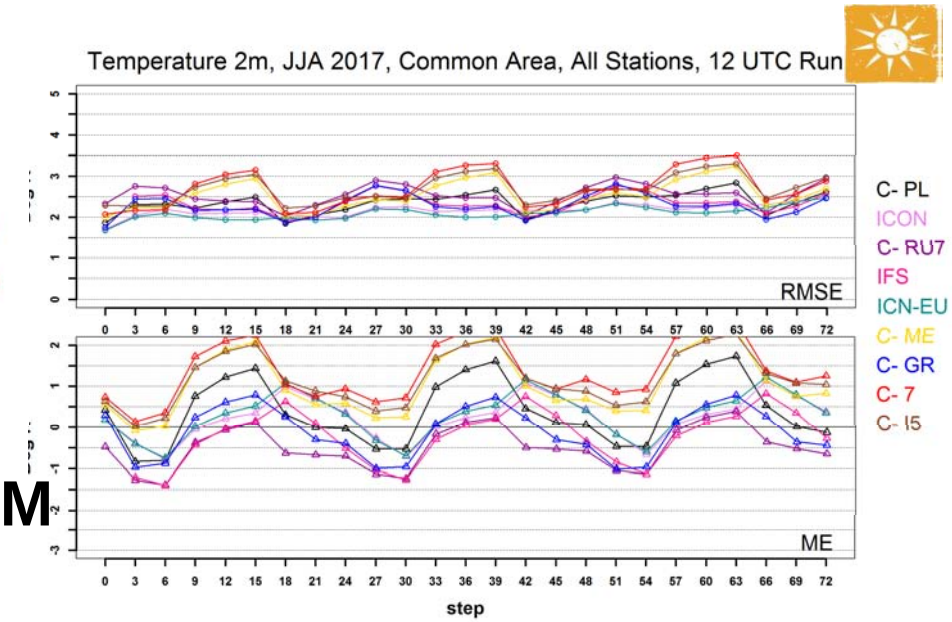



**Extremal Dependence Indice EDI**
**6h -24h Preci**

$$H = \frac{a}{a+c} \qquad F = \frac{b}{b+d} \qquad EDI = \frac{\log F - \log H}{\log F + \log H}$$

Thresholds: 1, 5, 10, 15, 20, 25, 30



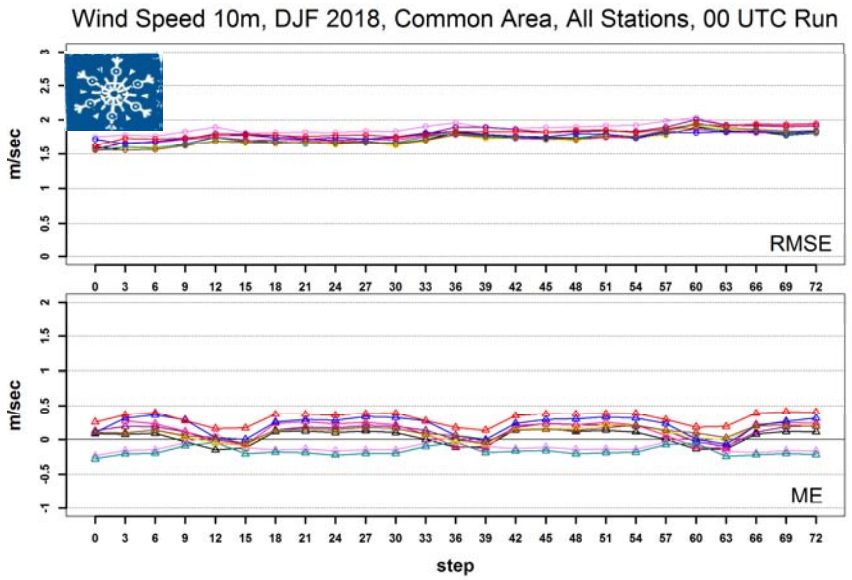



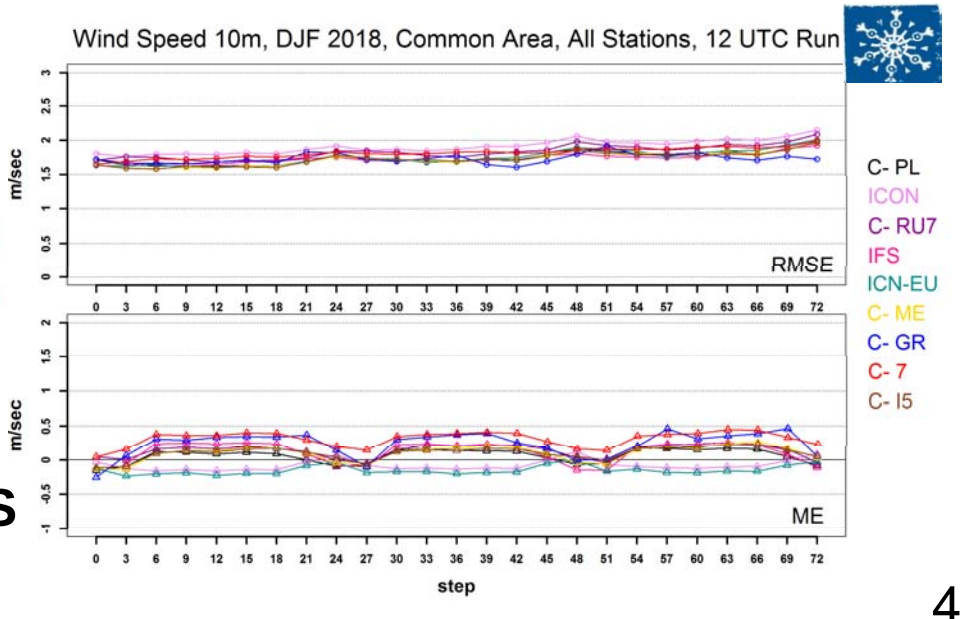

**T2M**

**00 UTC Run**

00 UTC Run and 12 UTC Run score difference is not significant

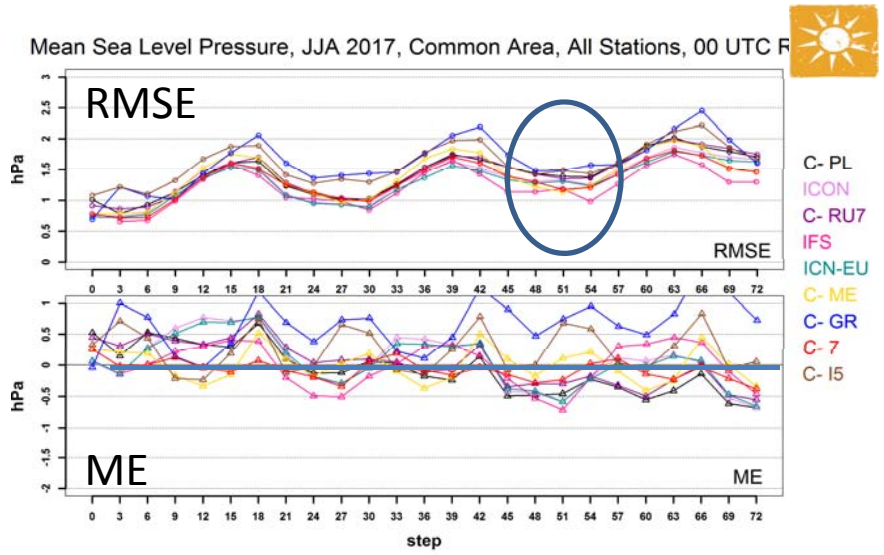
**12 UTC Run**

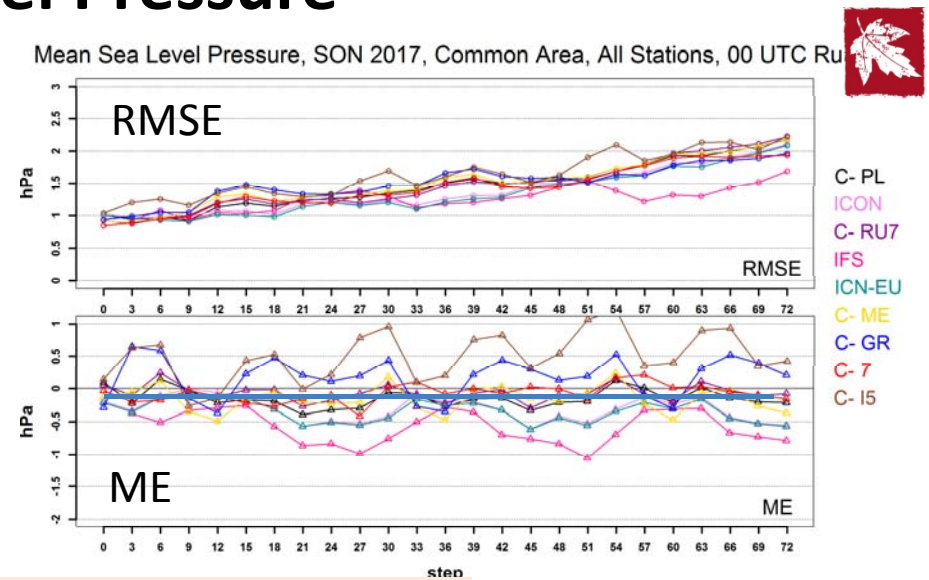


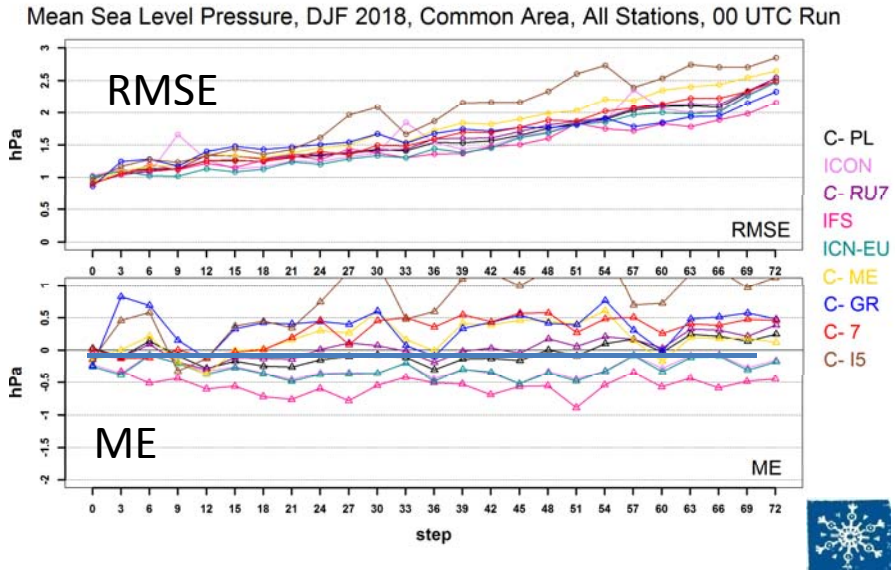

**WS**

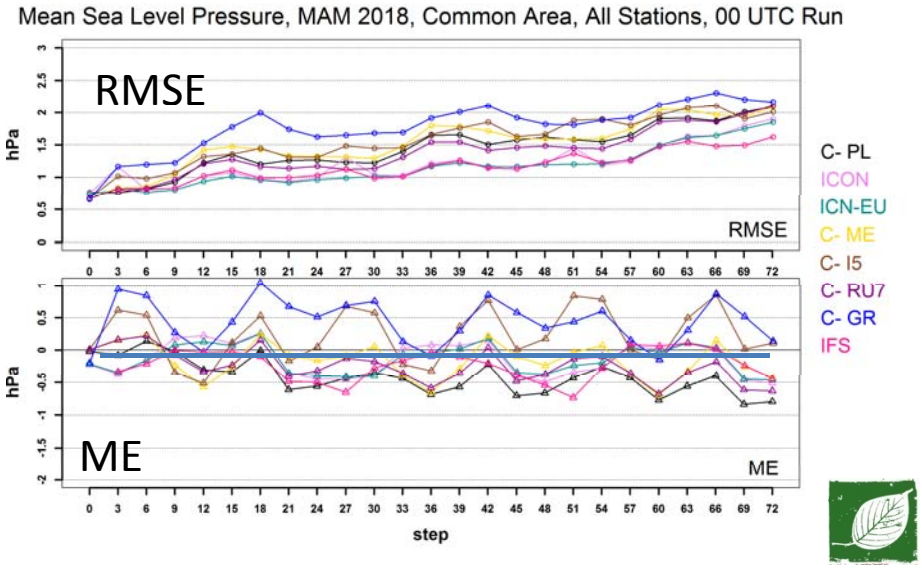
# Mean Sea Level Pressure



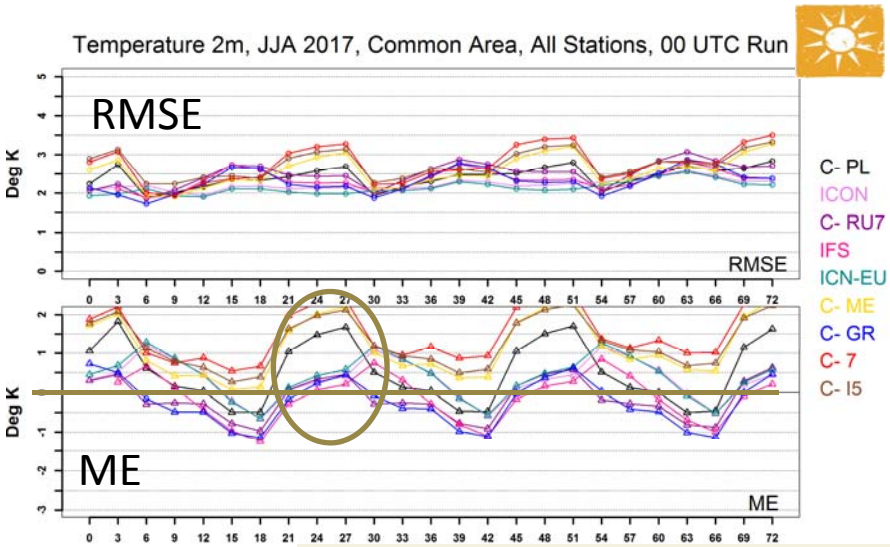



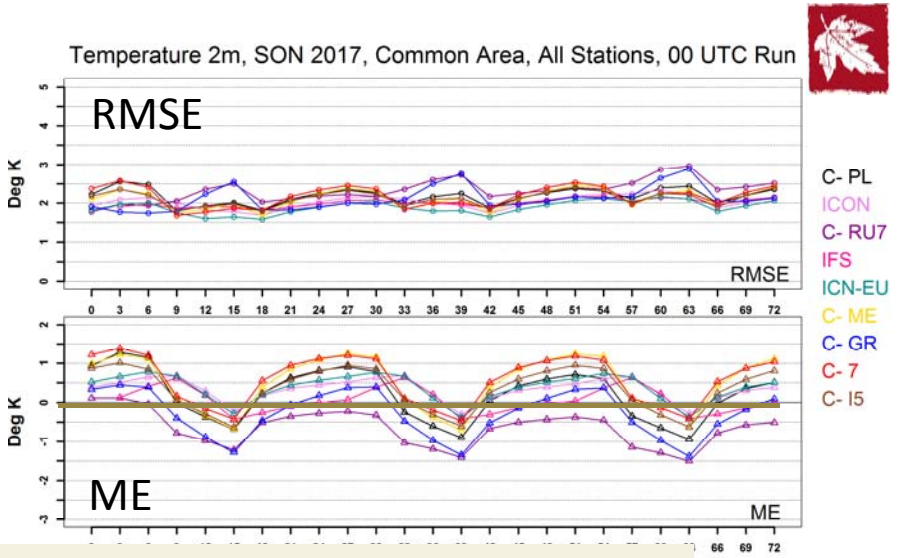

**JJA RMSE clear diurnal variability.and increases with time lead in all seasons.**



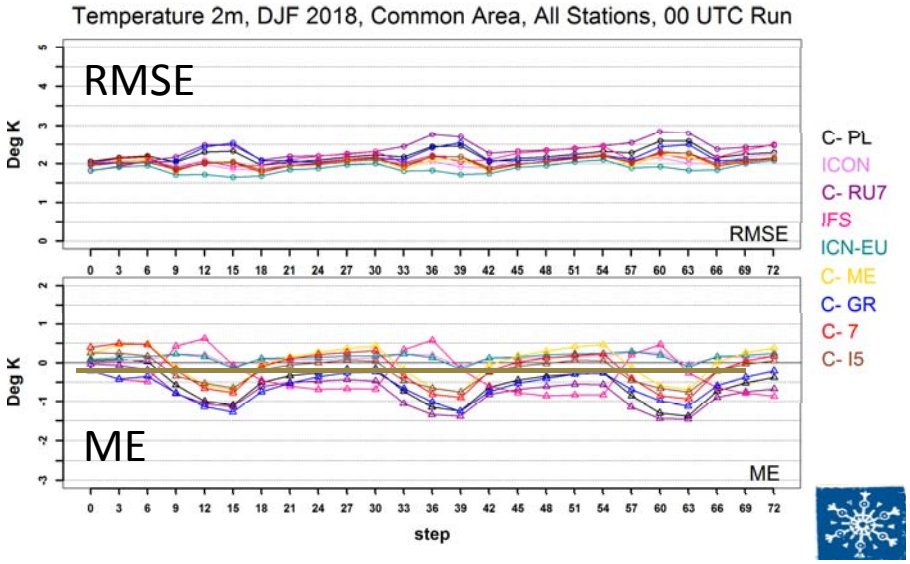



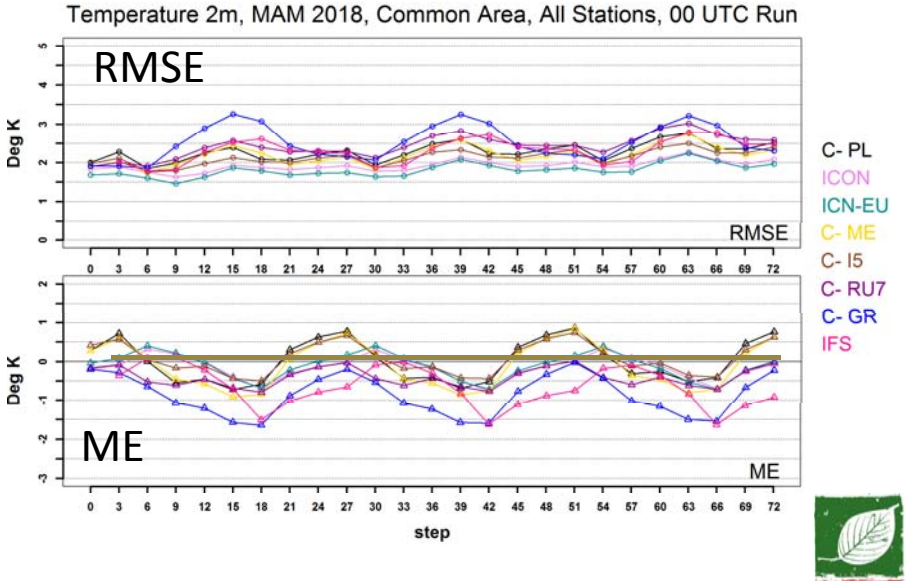
# Temperature 2m




**Clear bias diurnal variability with overstimation at night and understimation in the day. In JJA bias is high at night .**

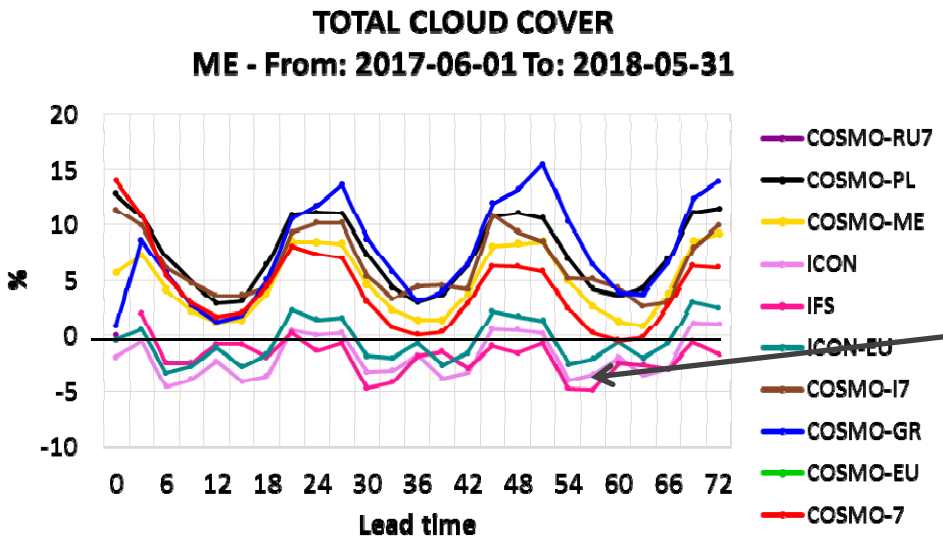


# TOTAL CLOUD COVER From: 2017-06-01 To: 2018-05-31

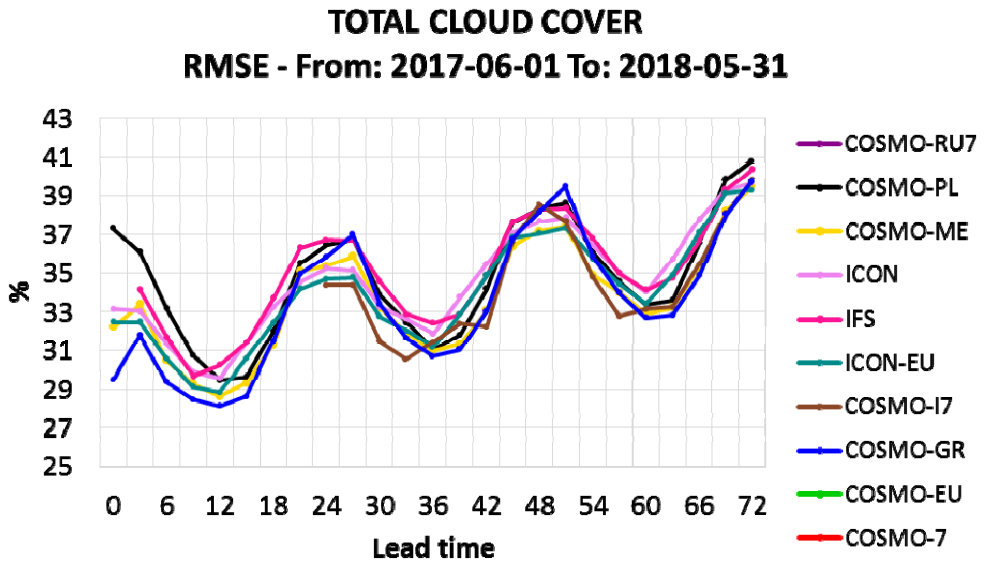
## ME

**Mean Annual Daily Cycle Values**



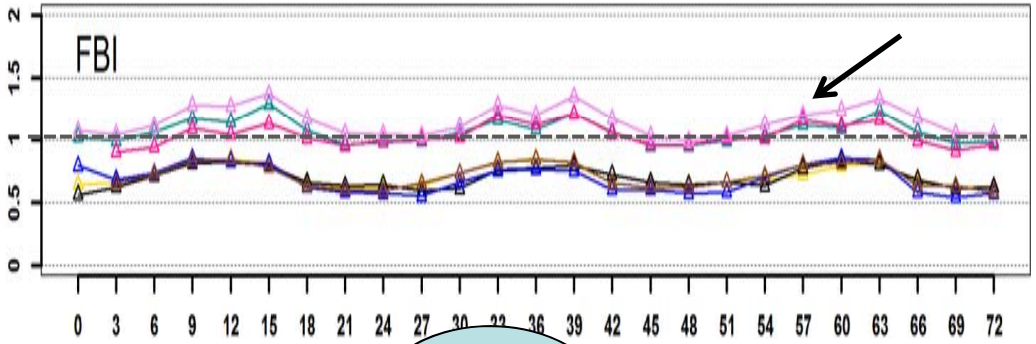

**COSMO** → ME>0
**ICON-EU/ICON** → ME~0
**IFS** → ME<0

## RMSE

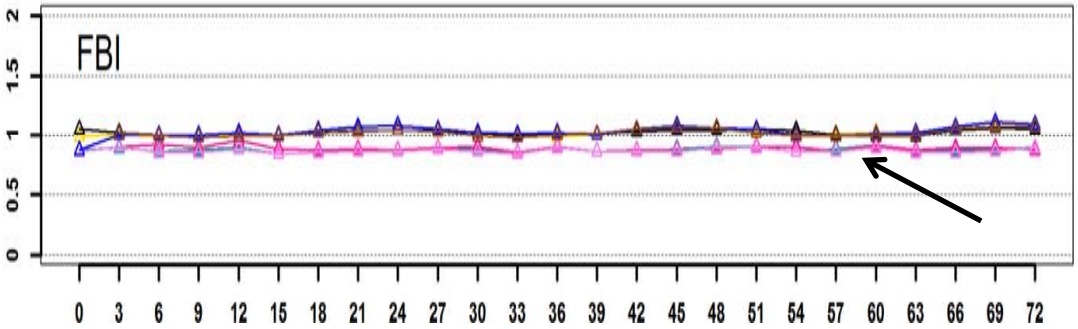



ME and RMSE maximum during nighttime

# Total Cloud Dichotomic FBI (0-25,25-75,75-100%)




0-25%



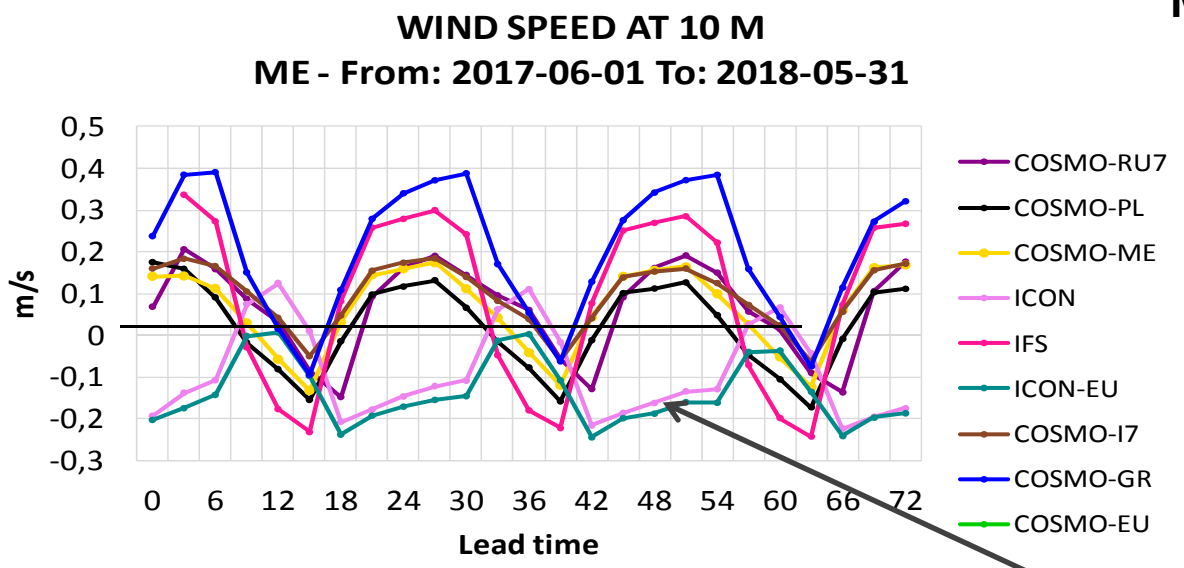

75-100%

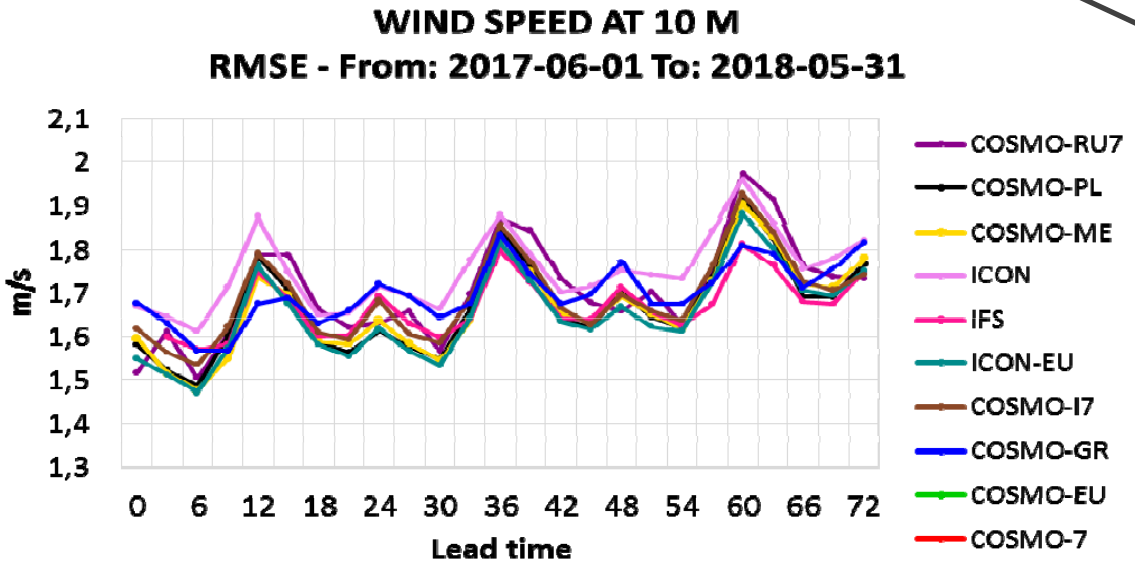
**IFS, ICON, ICON-EU**
**FBI > 1 for 0-25%**
**FBI < 1 for 75-100%**

**Low**
**Cloud cover occurrences are slightly overestimated and high cloud cover underestimated**

# WIND SPEED AT 10 M From: 2017-06-01 To: 2018-05-31

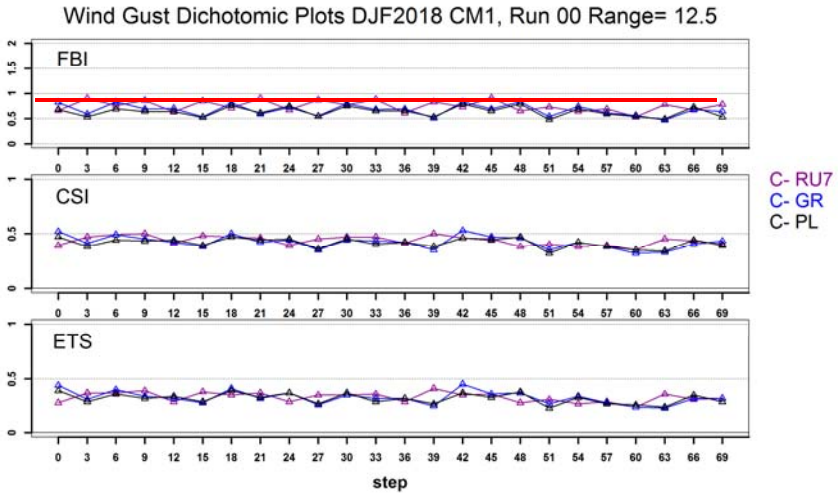
**Mean Annual Daily Cycle Values**



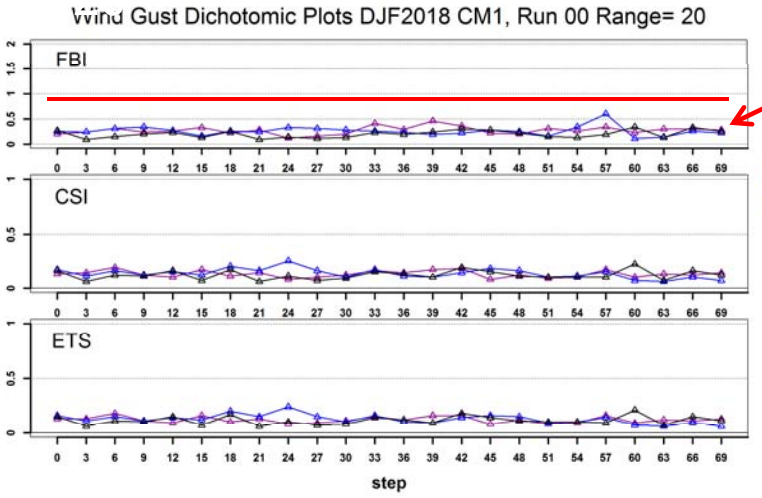




**ICON-EU/ICON** have negative Bias, less diurnal bias variation, maximum in daytime in contrast to other models.

# WIND GUST Dichotomic 10m 00 UTC Run DJF



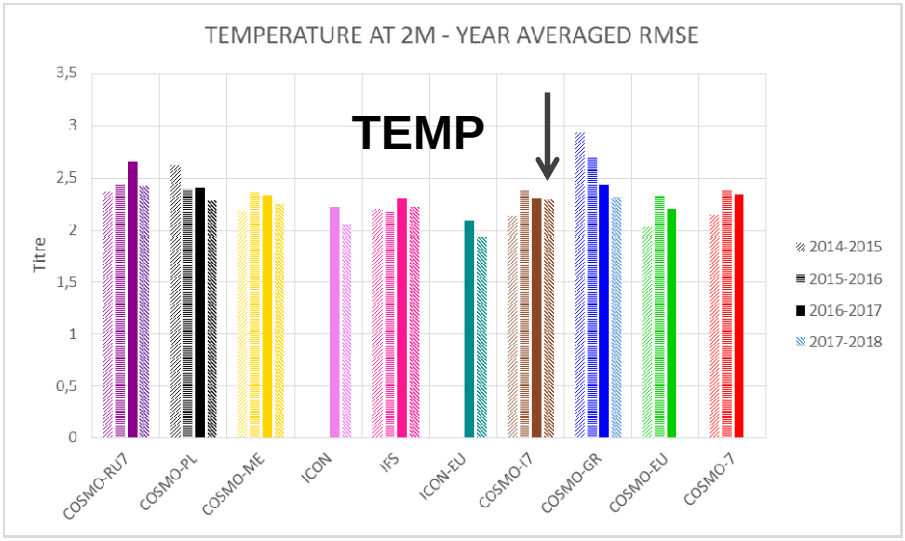

**> 12.5**

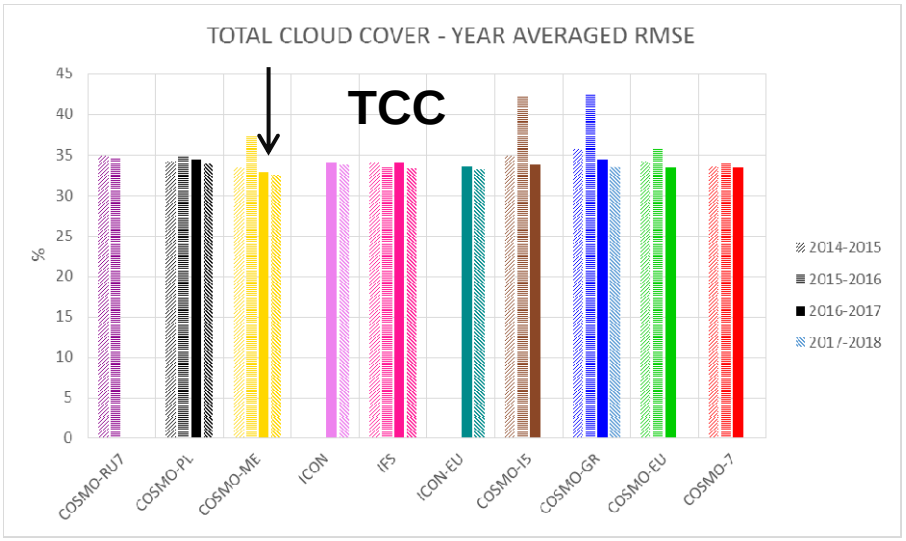


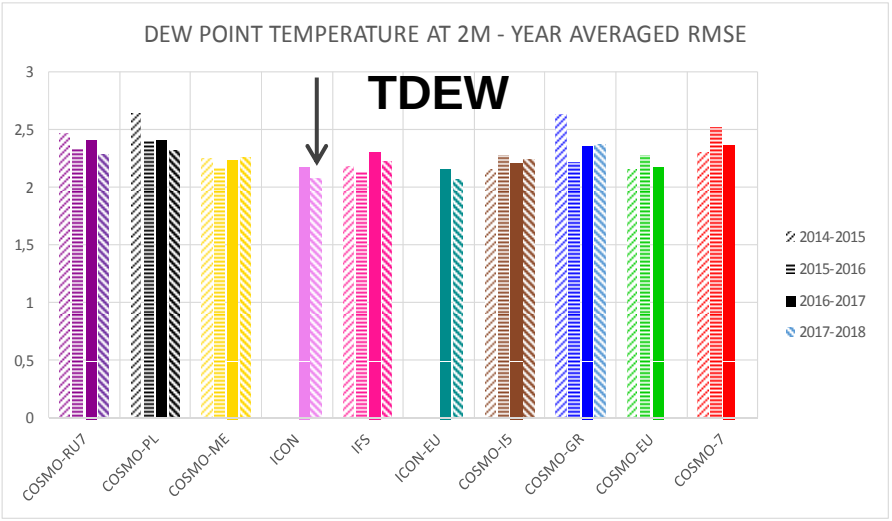

**> 20**

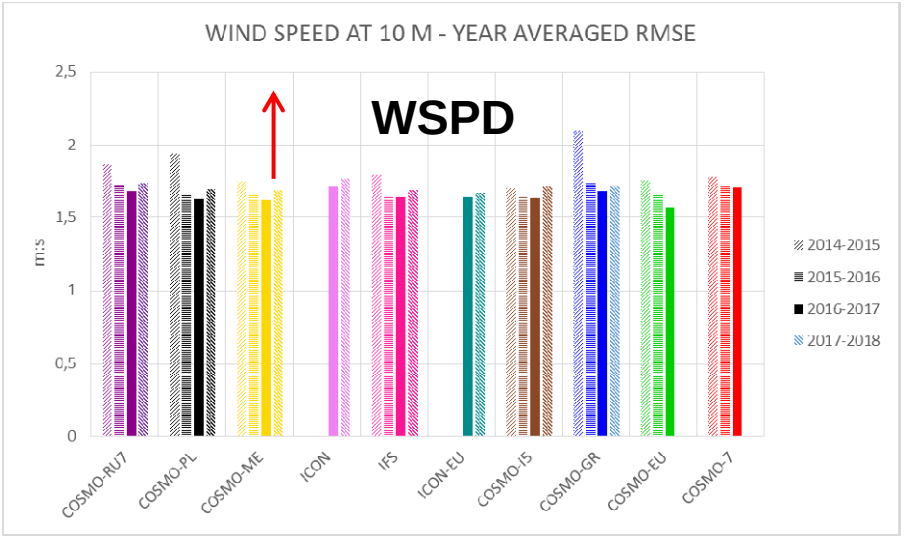
**There is Underestimation of occurrences for Higher Wind Gust threshold, ETS scores worse (closer to 0)**

# Annual RMSE TRENDS



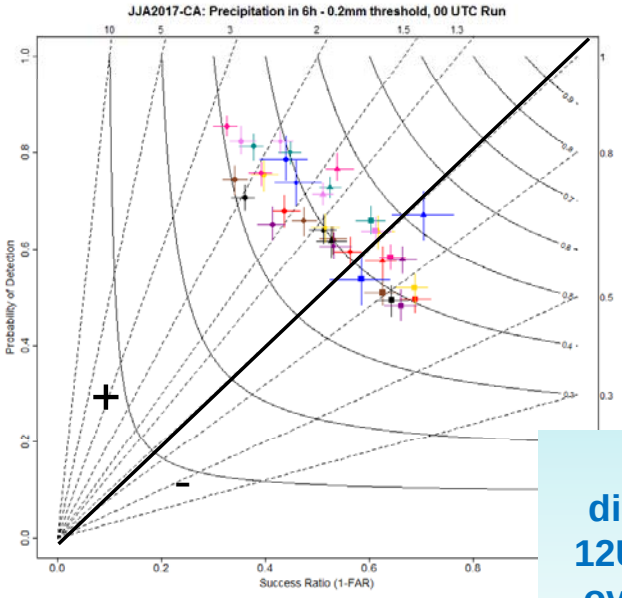



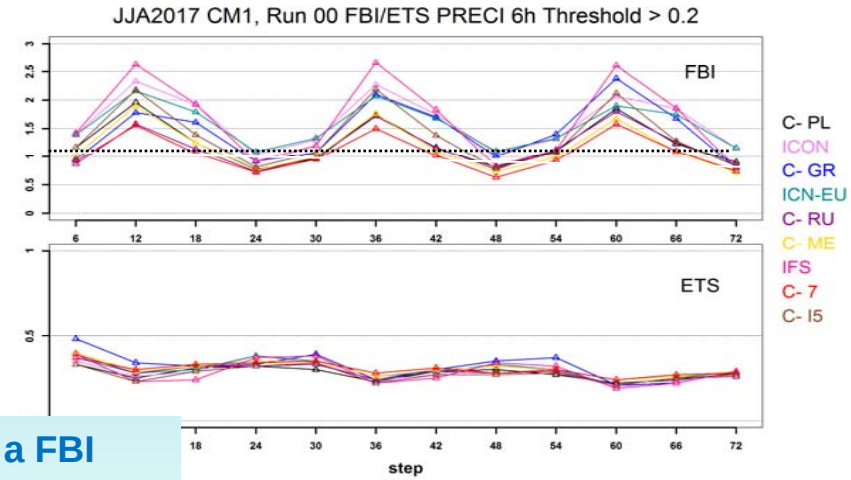




JJA ↑

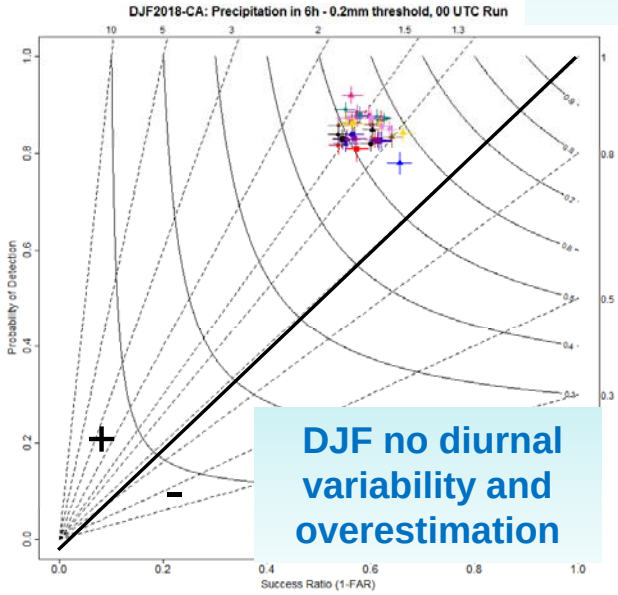
DJF ↓



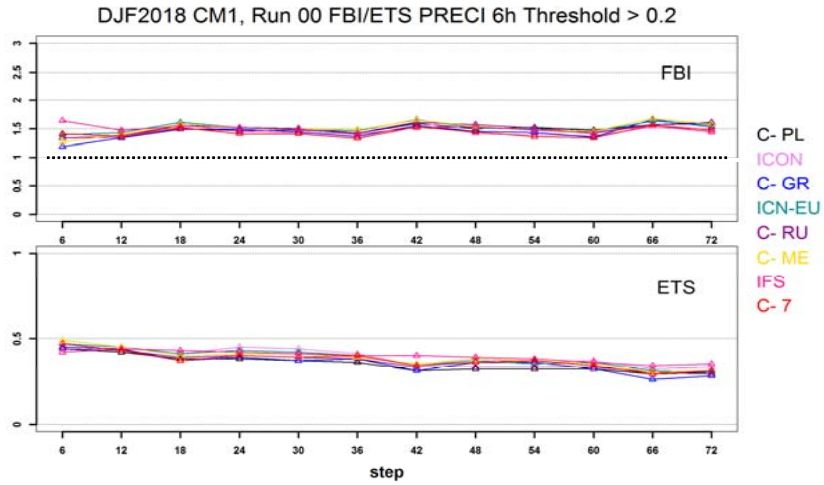

**In JJA there is a FBI diurnal cycle Max values 12UTC- Light precipitation overestimation mainly by IFS, ICON**

# Precipitation 6h > 0.2mm

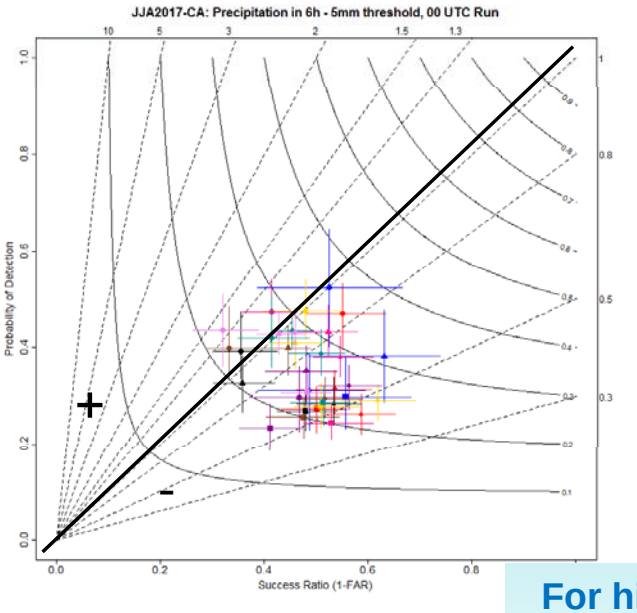



- COSMO-RU7
- COSMO-PL
- COSMO-ME
- ICON
- IFS
- ICON-EU
- COSMO-I7
- COSMO-GR
- COSMO-EU
- COSMO-7

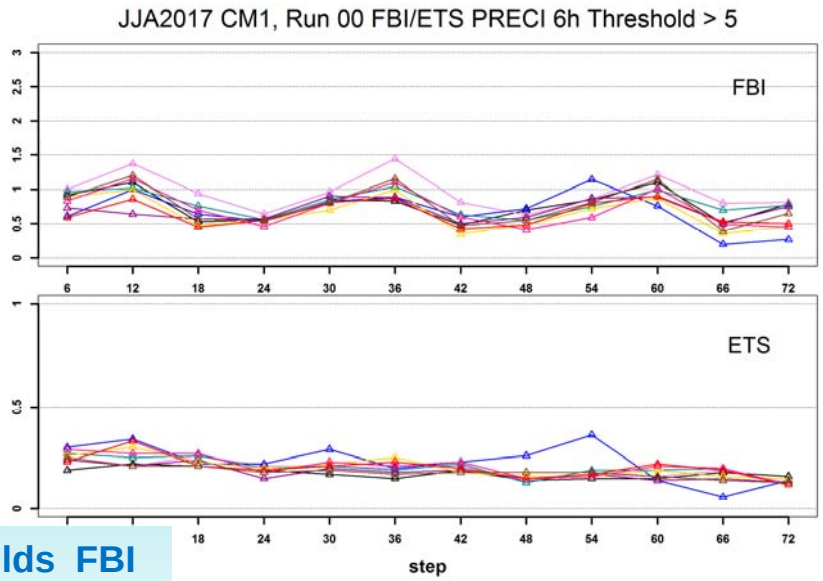
**DJF no diurnal variability and overestimation**



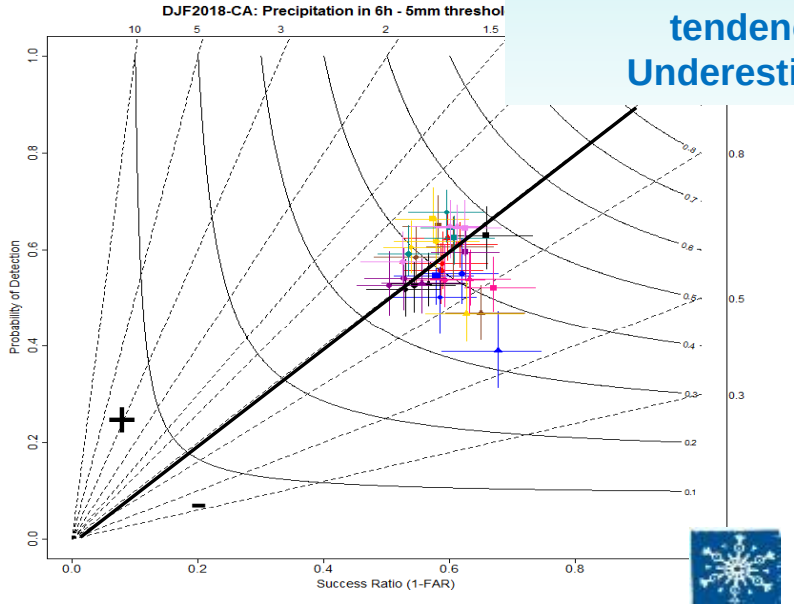

# Precipitation 6h > 5mm

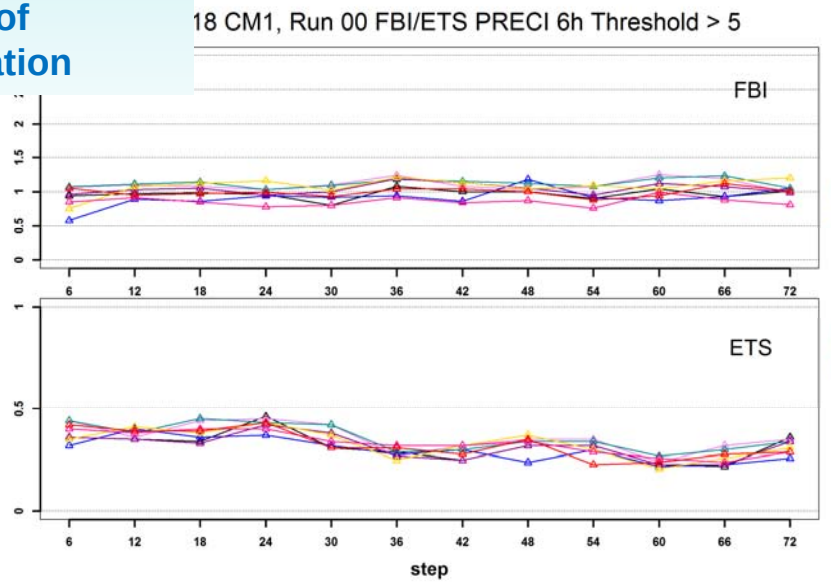



COSMO-RU7
COSMO-PL
COSMO-ME
ICON
IFS
ICON-EU
COSMO-I7
COSMO-GR
COSMO-EU
COSMO-7



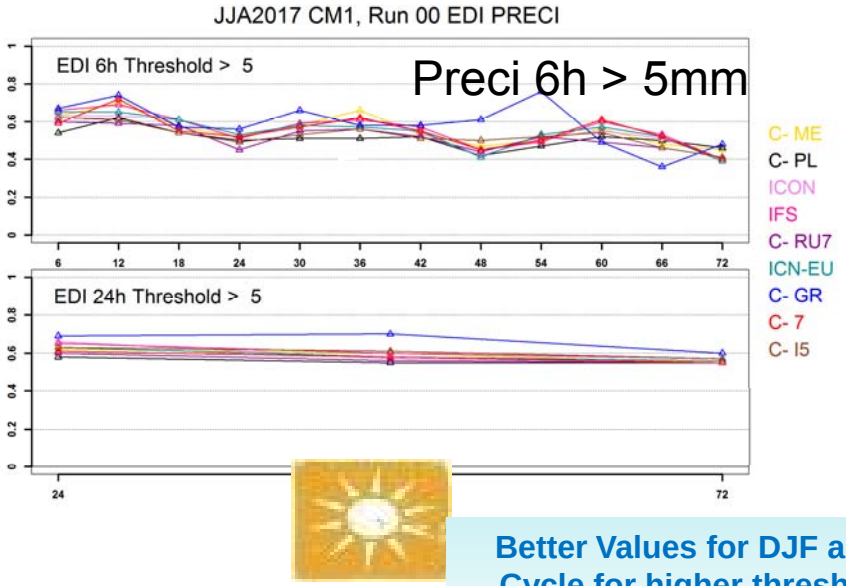

**For higher thresholds FBI weaker diurnal cycle- tendency of Underestimation**

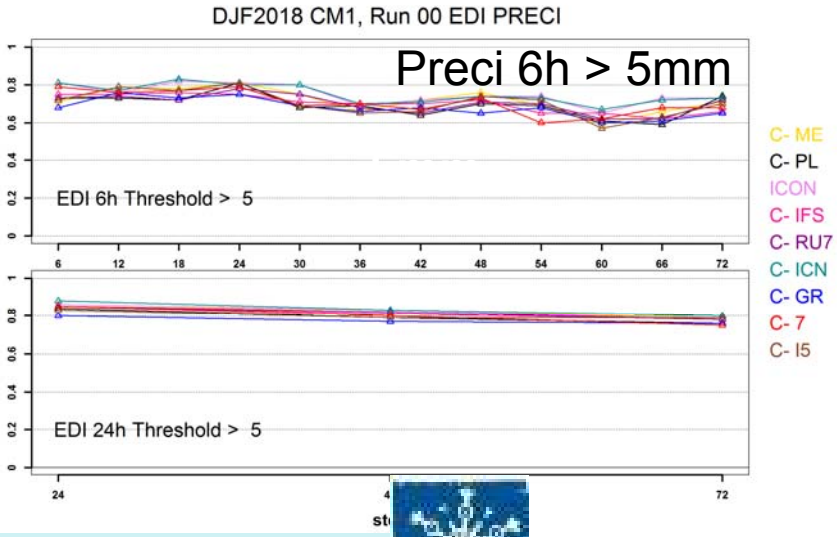


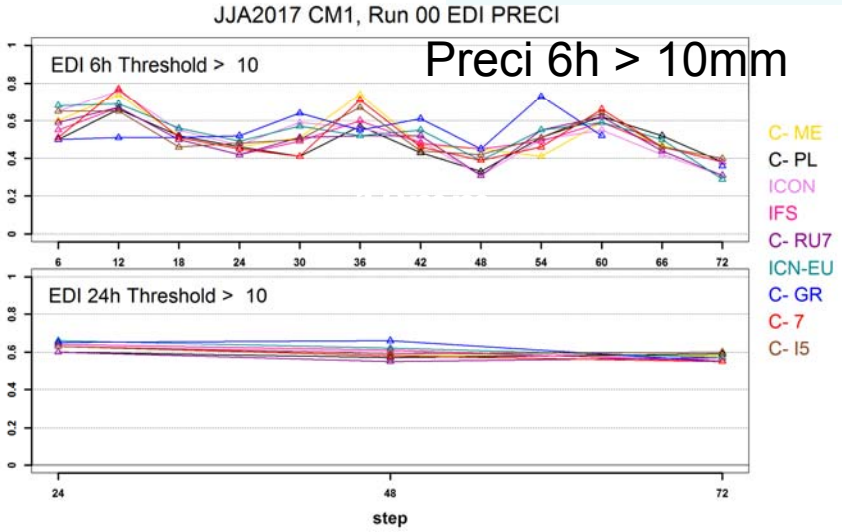

C- PL
ICON
C- GR
ICN-EU
C- RU
C- ME
IFS
C- 7

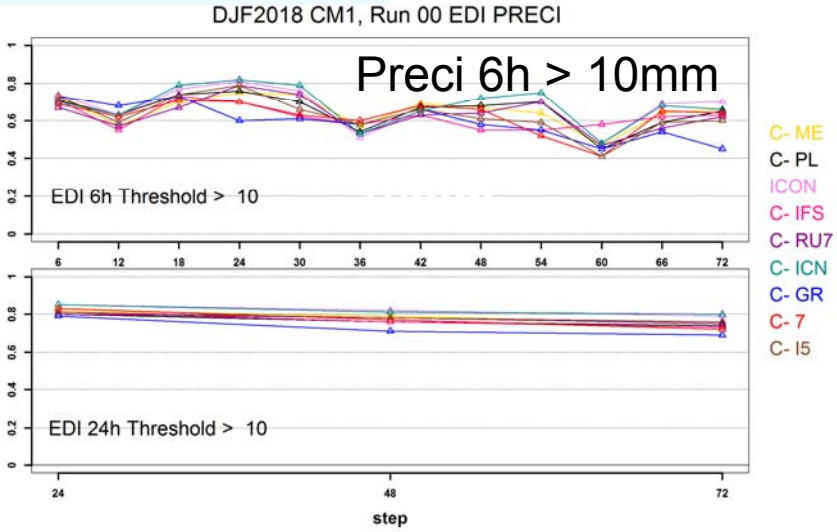
# Extremal Dependency Index 00 UTC Run



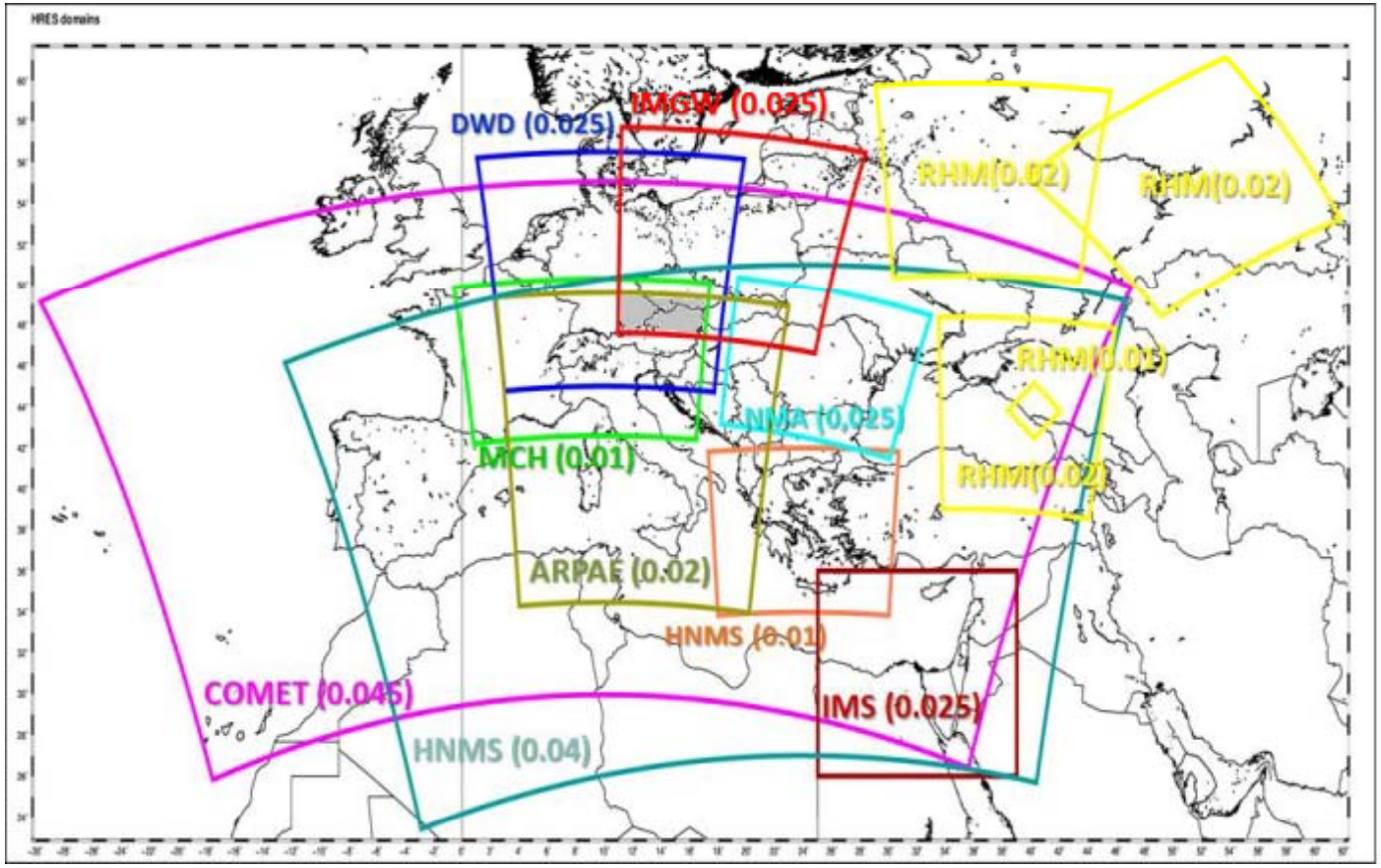




**Better Values for DJF and lower thresholds. Diurnal Cycle for higher thresholds. Index values decrease with time lead.**

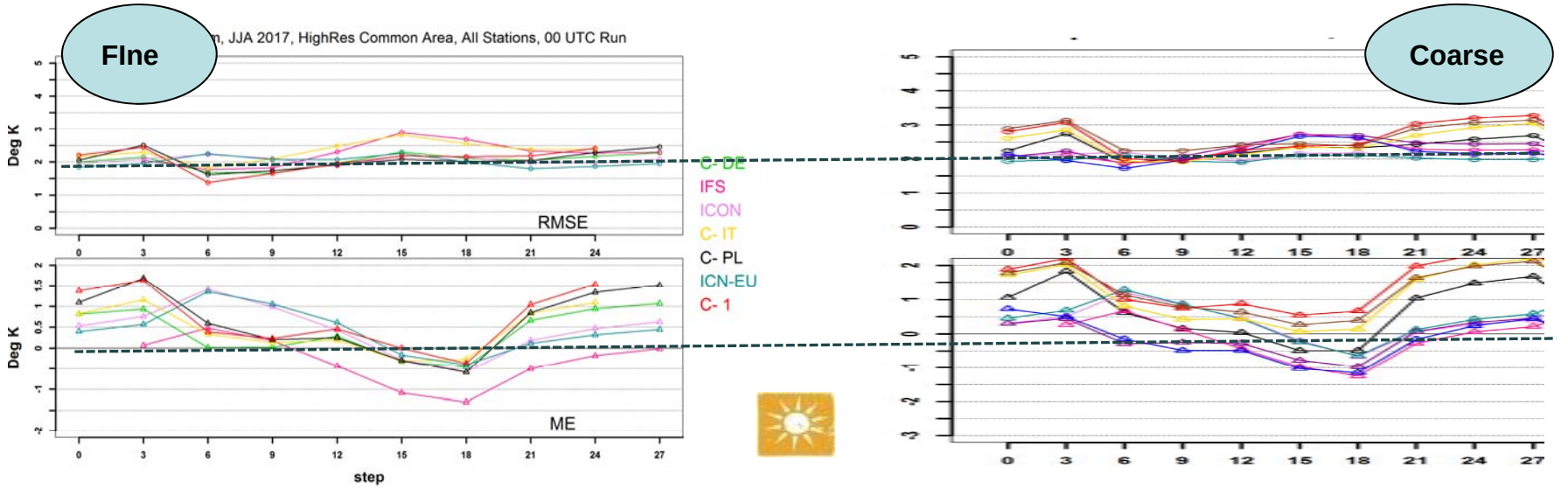


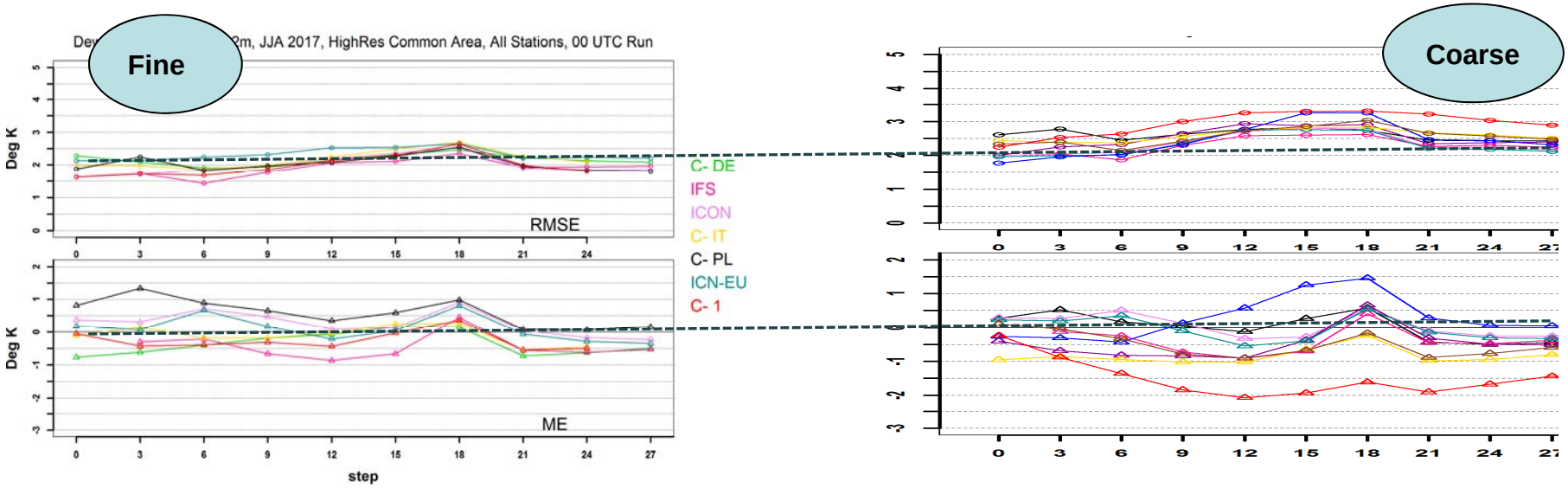
# Standard Verification on Common Area 2- High Resolution, COSMO-DE2, COSMO-1, COSMO-IT, COSMO-PL2, and ICON-EU, ICON and IFS for 00 UTC Run

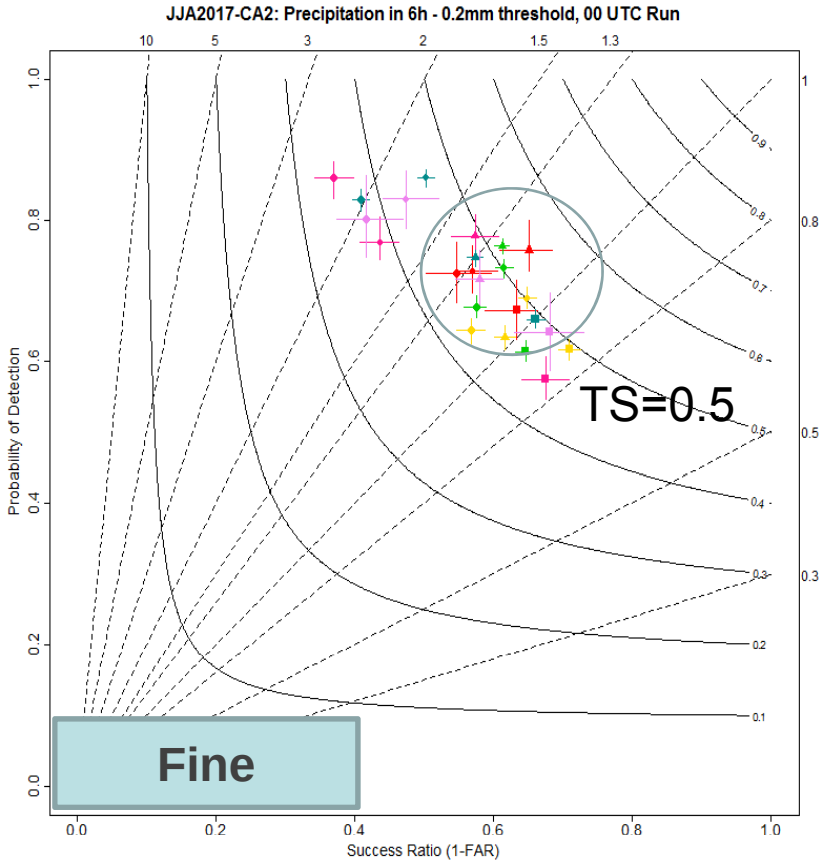


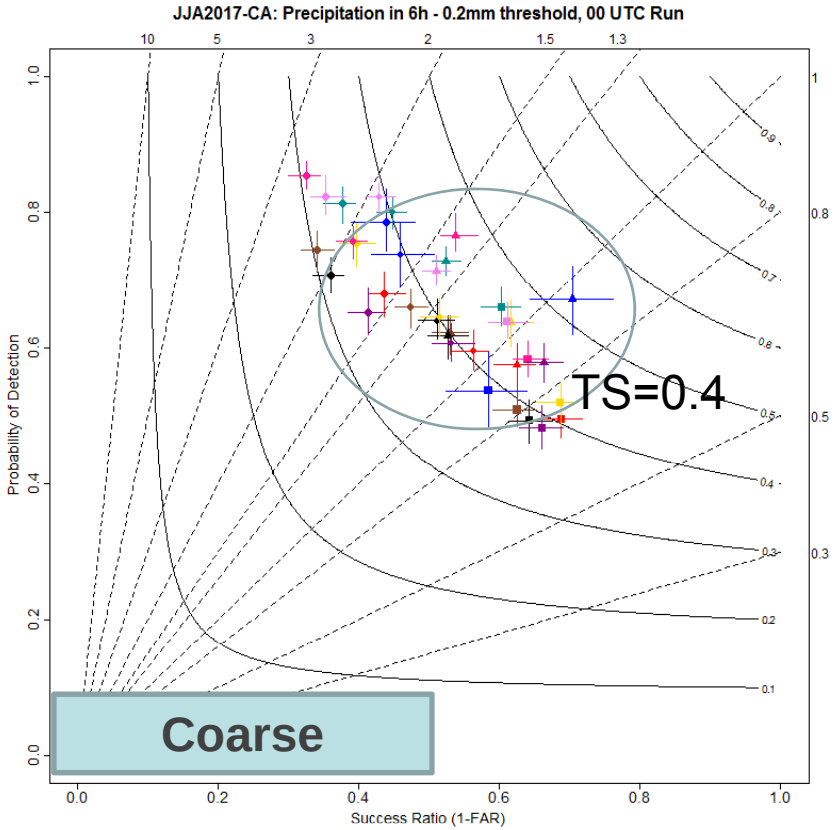
# 2m Temp




# 2m Dew Point T





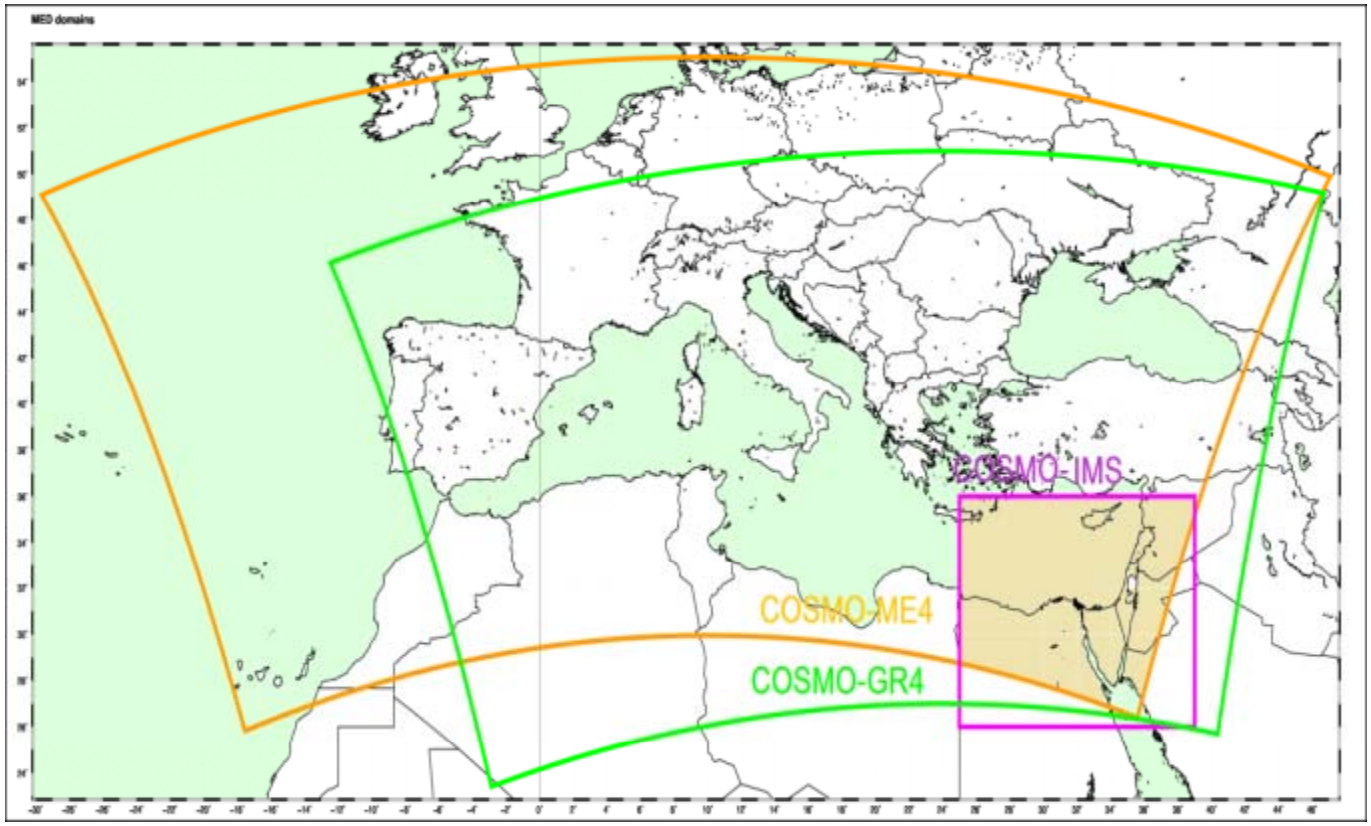

COSMO-PL
COSMO-IT
ICON
IFS
COSMO-DE
COSMO-1

COSMO-RU7
COSMO-PL
COSMO-ME
ICON
IFS
ICON-EU
COSMO-I7
COSMO-GR
COSMO-EU
COSMO-7

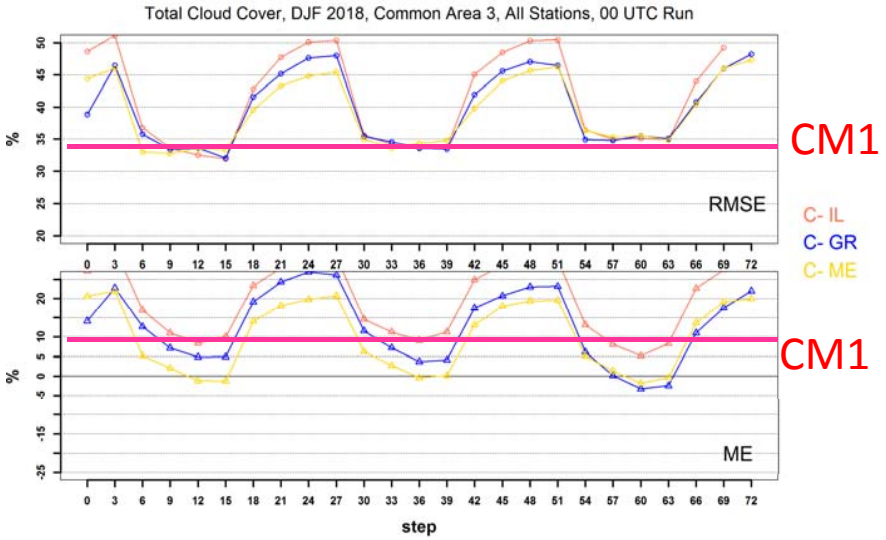
# Preci 6h > 0.2mm

**Better results for HR models, Higher ETS and FBI closer to 1**

# Standard Verification on Common Area 3- Dry Area Climate COSMO-IL(2.8km), COSMO-GR (4km) , COSMO-ME(5km) for 00 UTC Run

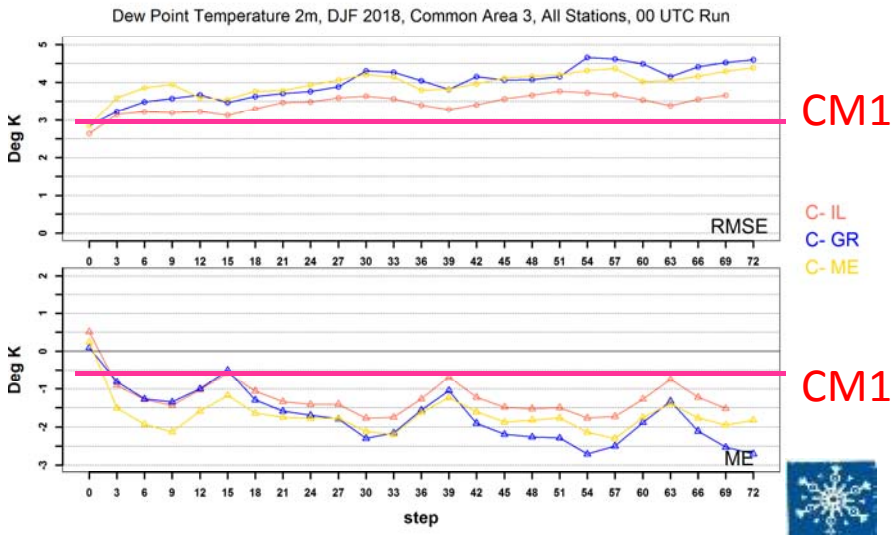


## TCC

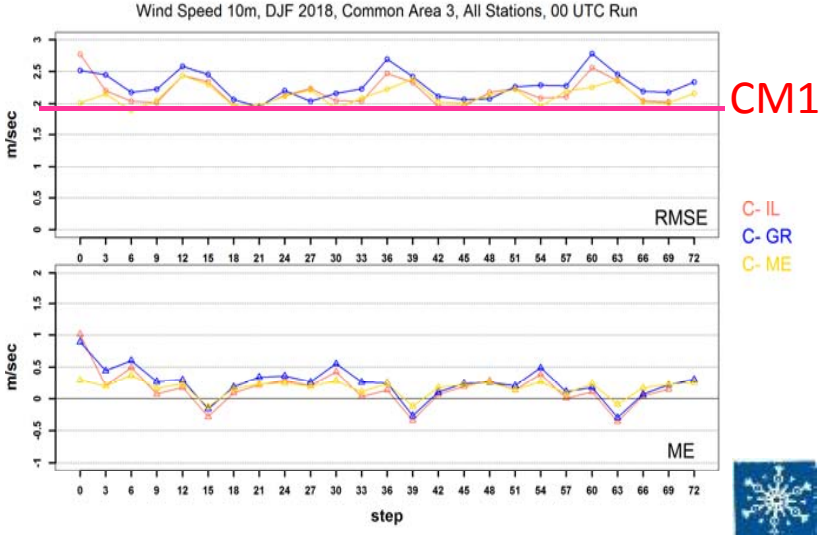



**TCC, Wspeed and Tdew scores generally worse than CM1**

## TDEW

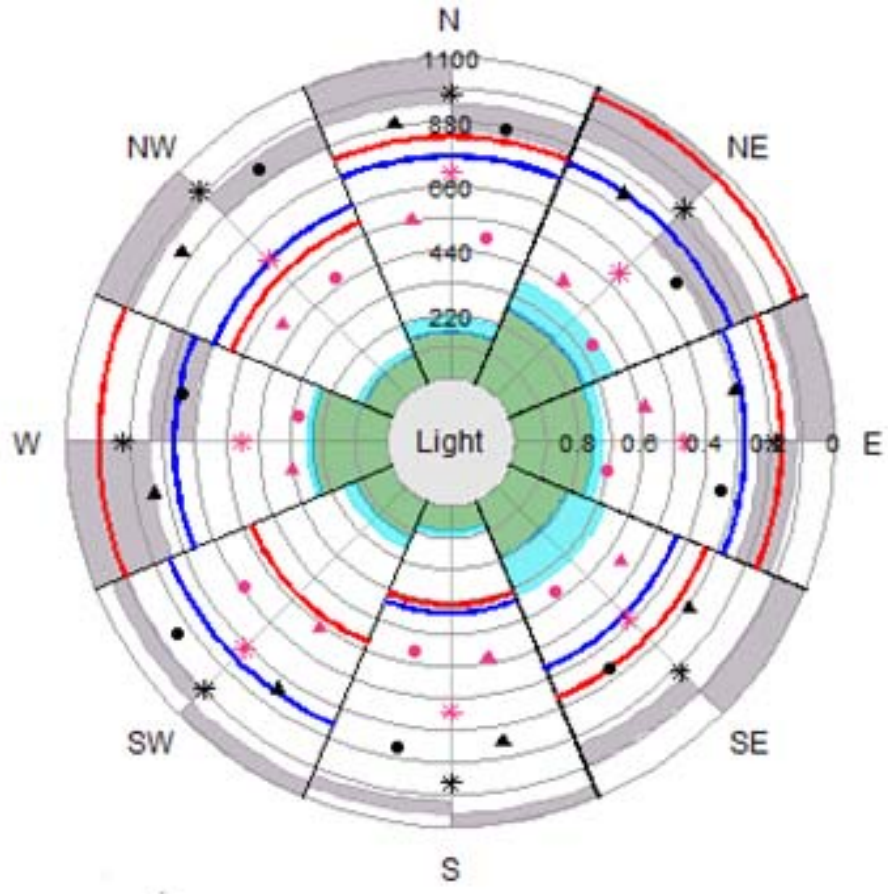



## WS



# The “performance-rose”

- A novel diagram in which scores and type of errors of wind forecast are summarized according to directions was introduced by Maria Stefania Tesini was applied for Common Area this year.



# How it works

For each station, obs and fcs wind data are categorized in octants for wind direction and in classes for wind speed.

**Light:** ws<10 knots

**Light-Moderate:**

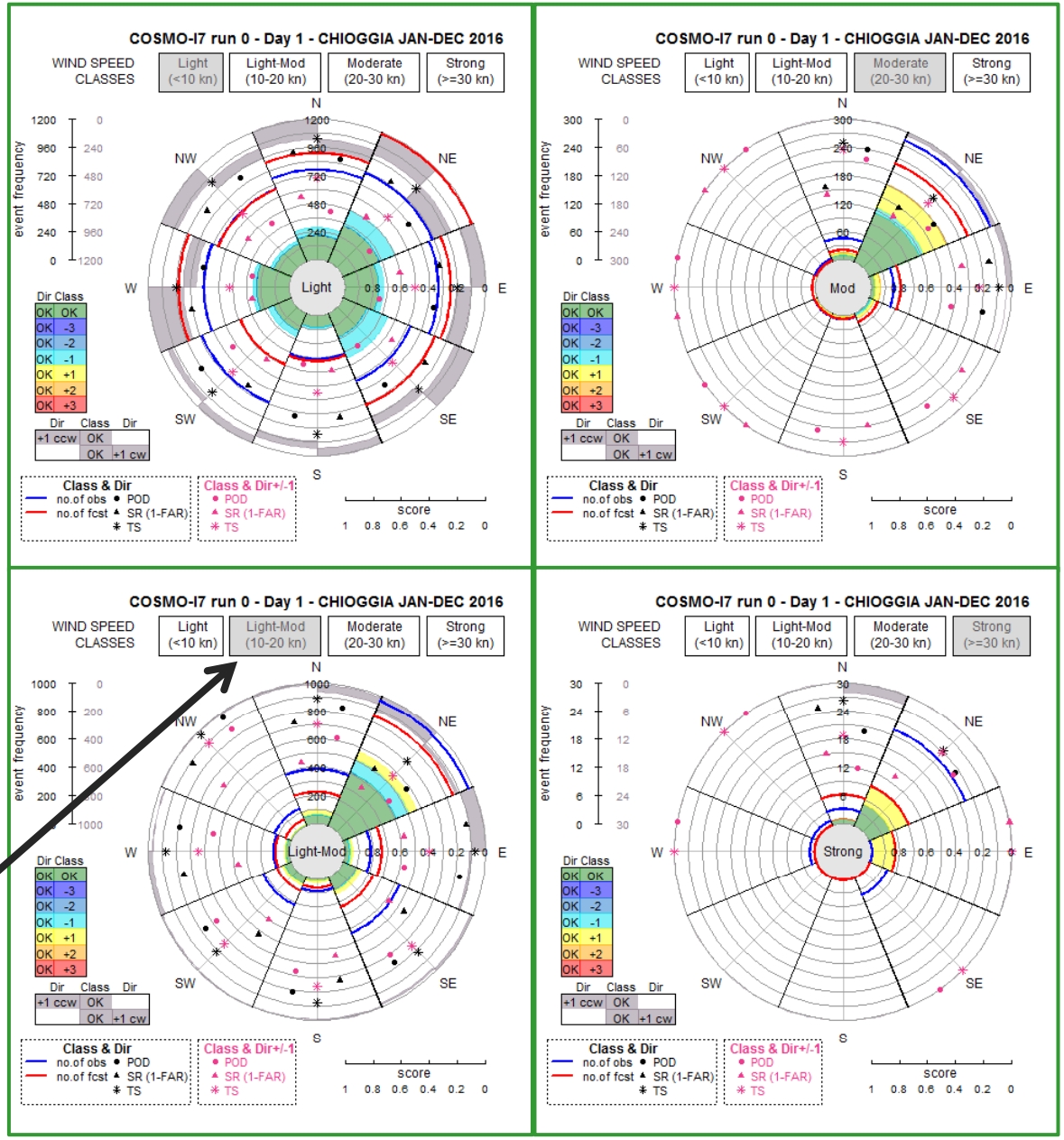
10≤ ws < 20 Knots

**Moderate:**

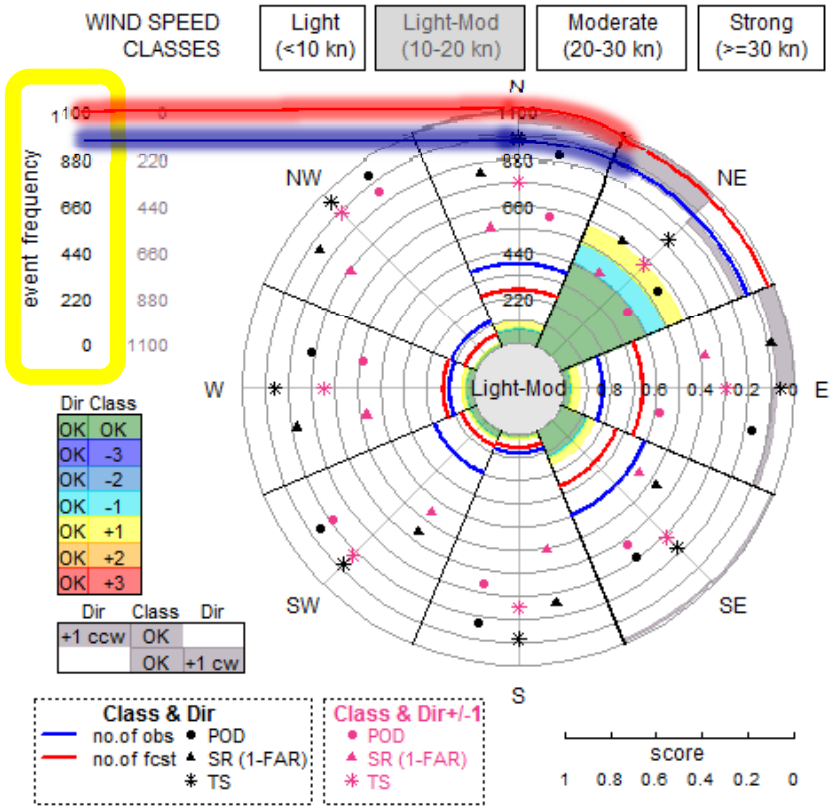
20≤ ws < 30 Knots

**Strong:** ≥30 Knots

**For each class a separate plot is done**

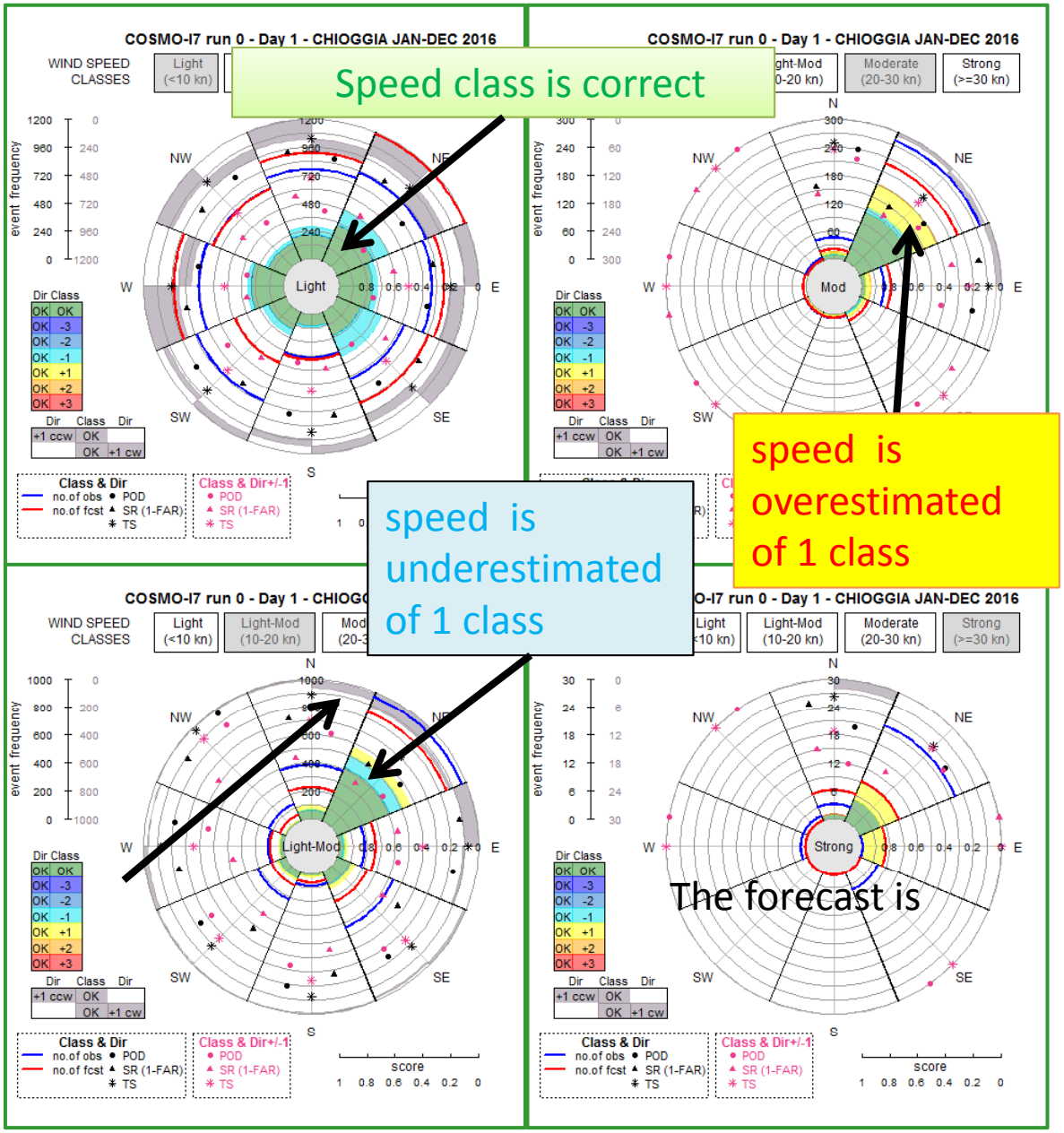


# How it works

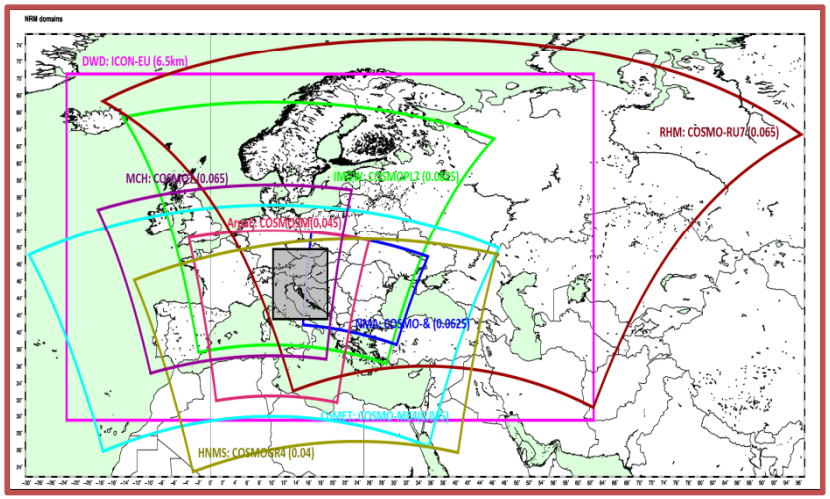


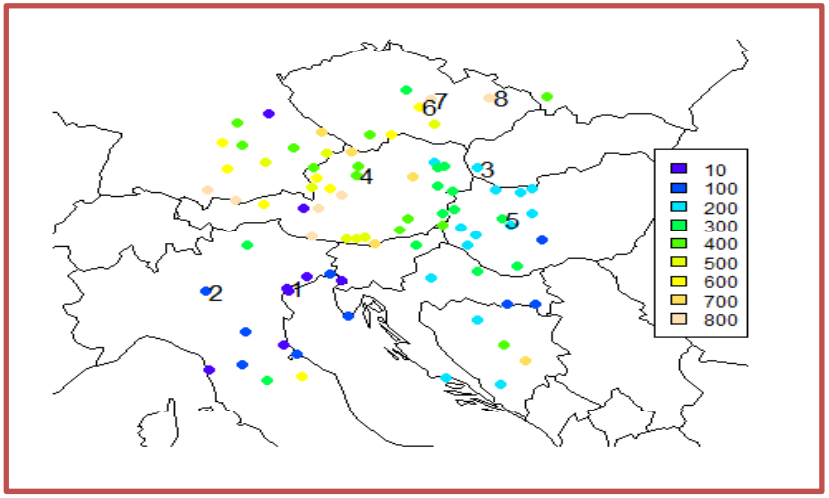

- Blue line is the **observed** frequency of the specific speed-class in each direction
- Red line is the **forecast** frequency of the specific speed-class in each direction
- The number of events can be read on the radial scale (frequency axis), increasing outward from the center
- The Frequency Bias can be easily deduced by **RELATIVE POSITION** of blue and red line
  - **Red outer** → overestimation
  - **Blue outer** → underestimation

Colored sectors represent how model predicts the reference speed class in each direction, <u>being the direction correct</u>



# Selected Stations for Common Area




Common Area 1

COSMO-5M (Arpae)

COSMO-PL (IMGW)

COSMO-GR4 (HNMS)

ICON-EU (DWD)

COSMO-DE /D2 (DWD)

**Selected stations:**

| | code | name | height |
|---|---|---|---|
| 1 | 16105 | VENEZIA TESSERA | 6 |
| 2 | 16088 | BRESCIA/GHEDI | 97 |
| 3 | 11816 | BRATISLAVA-LETISKO | 134 |
| 4 | 11012 | KREMSMUENSTER | 390 |
| 5 | 12830 | VESZPREM/SZENTKIRALYSZABADJA | 281 |
| 6 | 11659 | PRIBYSLAV | 536 |
| 7 | 11683 | SVRATOUCH | 740 |

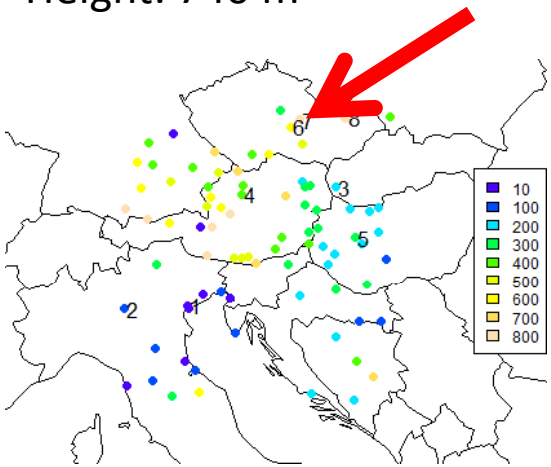
**According to this implementation of the PerfRose, some questions can be interesting for "COMMON PLOTS" users**

**Do COSMO models overestimate mountain 10m wind speed ?**

**Are there significant differences between forecasts of Day 1 and 2 ? (Useful for warnings)**

**Are there differences between a model and its driving model ?**

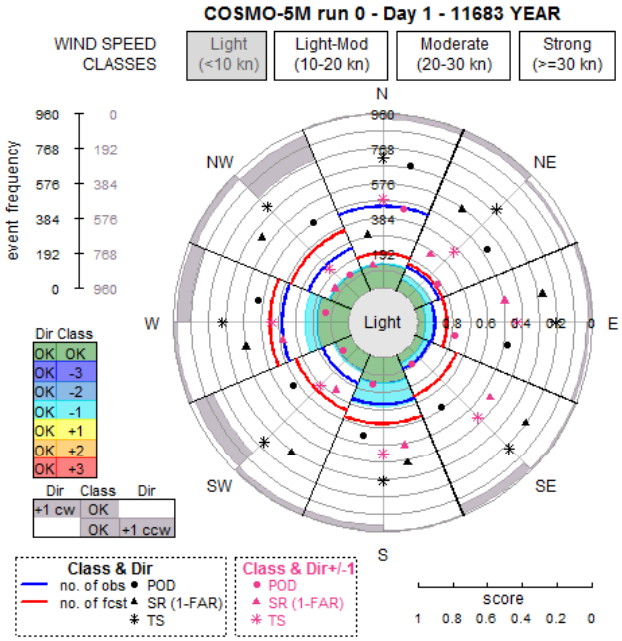
Station: 11683
SVRATOUCH
Height: 740 m

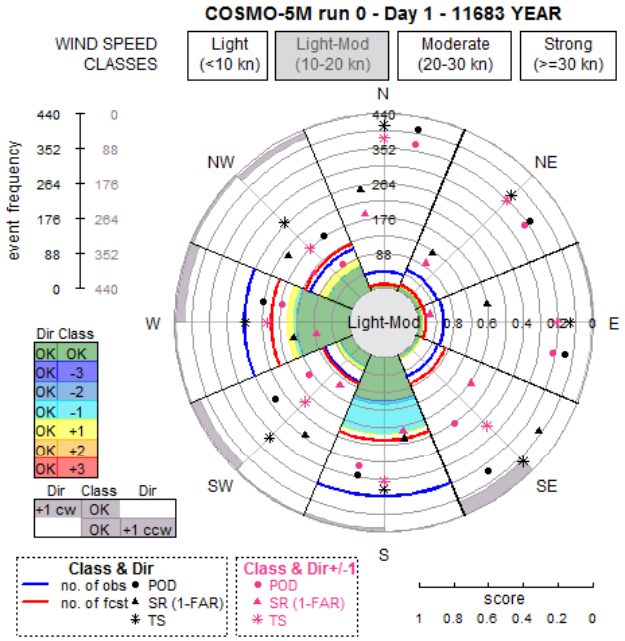


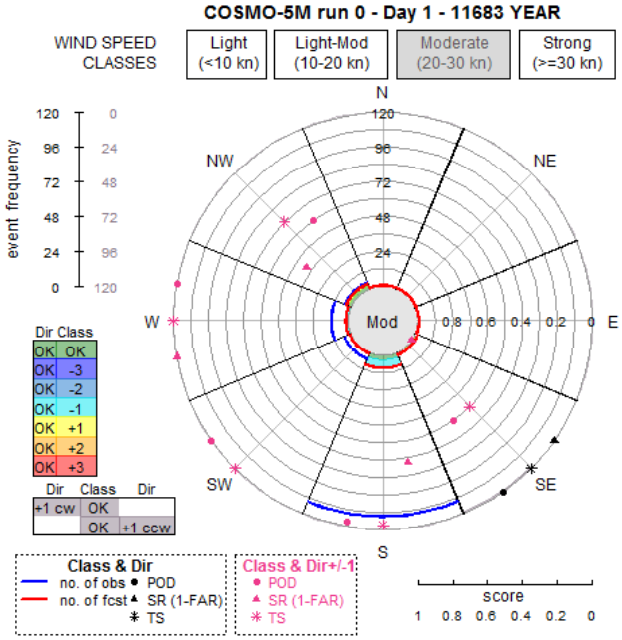

Period:
1 Year of data from
June 2017 to May 2018

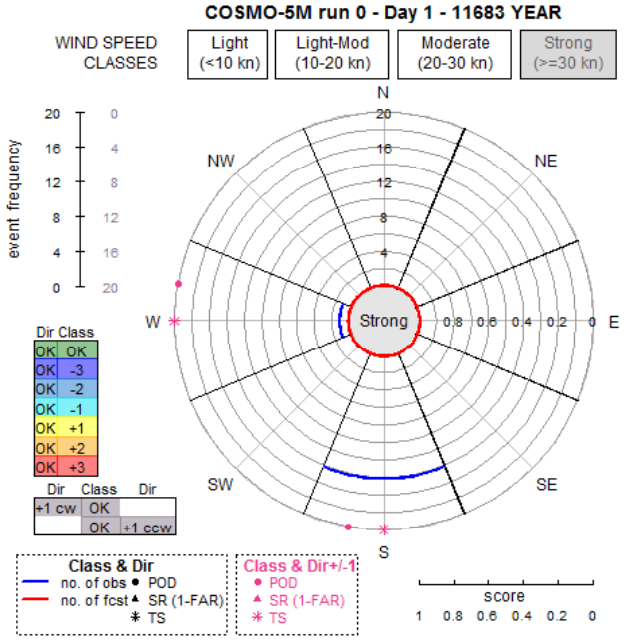
**Understimation of 10m winds in mountain**

Model: COSMO-5M
Run:00
Forecast Day: 1
Step: 3 hr



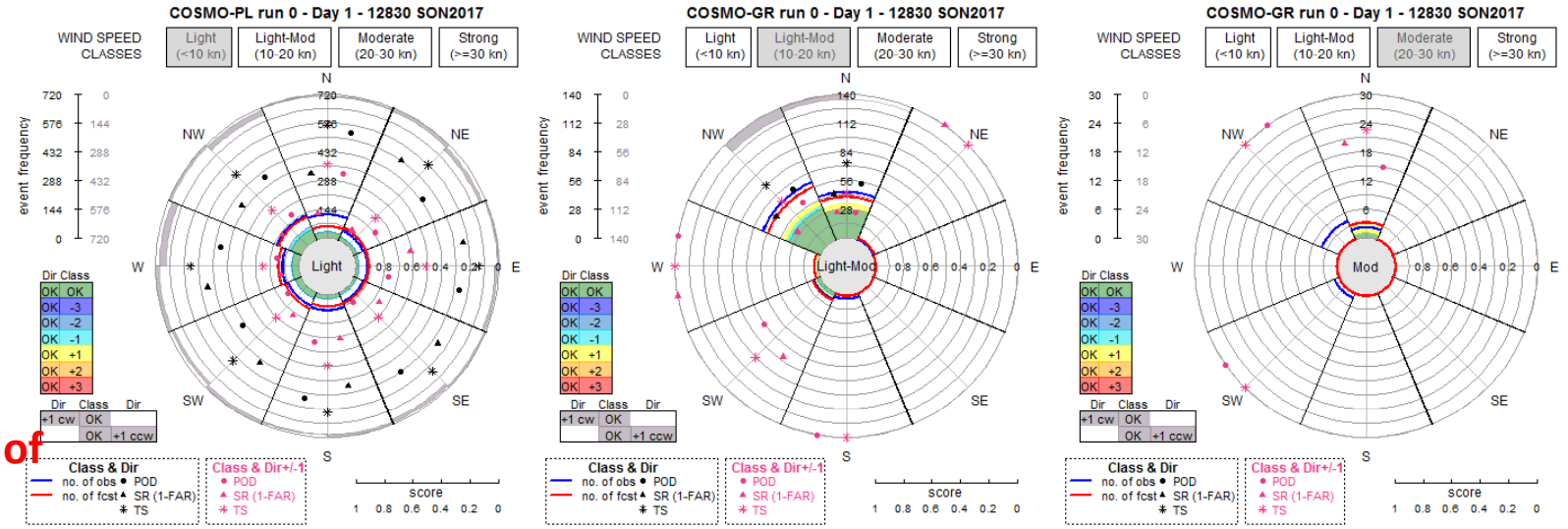



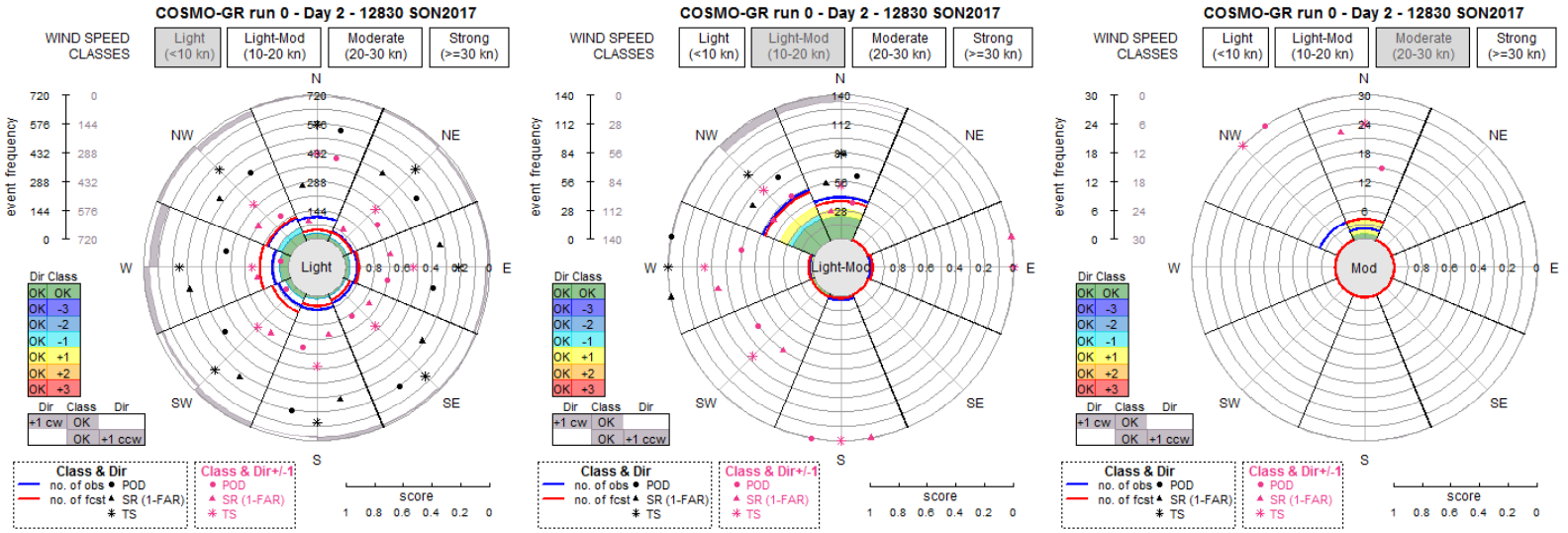
# SON2017 – 12830: COSMO-GR

**DAY 1**

**Differences of day 1 and 2**




**DAY 2**



# Summary of results I

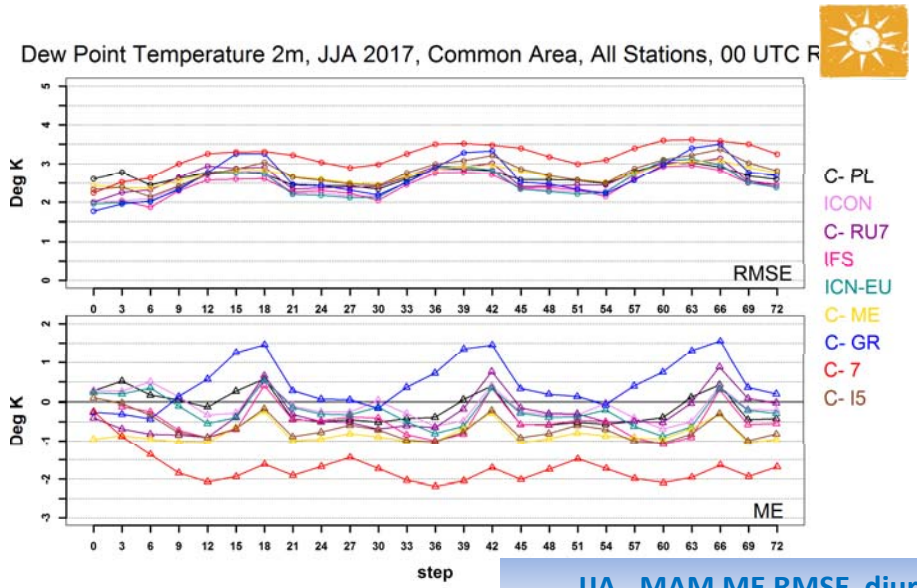
- **Total cloud cover:**
  - ***Positive BIAS especially at night. IFS, ICON, ICON-EU similar behavior with weaker variability and small negative values. Dichotomic TCC scores showed that the larger scale models overpredicted cases with low cloudiness.***
- **Temperature 2m:**
  - ***Clear diurnal cycle of BIAS with higher values during night .***
- **Dew point temperature 2m:**
  - ***Weaker variability in SON and DJF.***
- **Mean surface level pressure:**
  - ***All models (also IFS, ICON, ICON-EU) show a maximum of RMSE during summer at late afternoon. RMSE increases with time lead.***
- **Wind speed 10m:**
  - ***Lower BIAS amplitudes for IFS, ICON, ICON-EU.***
- **Wind Gust 10m:**
  - ***Scores are worse for higher thresholds.***

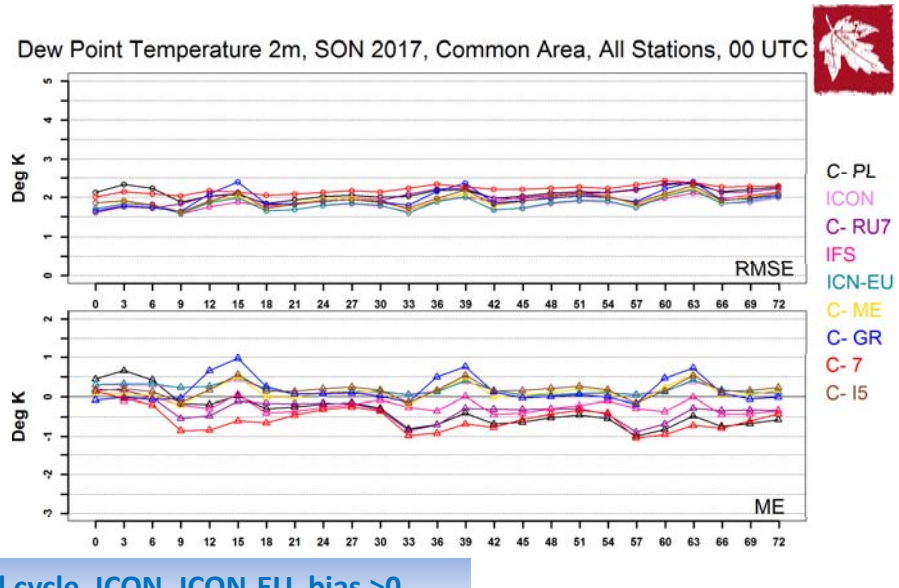
# Summary of results II

- **Precipitation:**
  - *Summer: Overestimation for occurences of low precipitation amounts during day especially for 06 - 12 UTC,– Underestimation for 18 – 24 UTC. (FBI decreases for higher precipitation amounts) .*
  - *Winter: Overestimation for occurences of low precipitation during the whole day. For higher precipitation amounts frequency bias is slightly greater than 1 with worse quality compared to low precipitation amounts*
  - *Overestimation for ICON, IFS for low precipitation amounts.*
- **Tendency of RMSE :**
  - *On annual basis most RMSE scores apart from Wind Speed imroved. JJA scores were slightly worse than last year (too warm and dry summer), while DJF scores improved.*
- **Test for Common Area II and III :**
  - *Scores for HR over CM2 were slightly better than CM1, but for dry Area CM3, some scores such as TCC, TDEW, WS were worse.*
- **Wind Rose Approach :**
  - *Errors of wind and direction plotted in a novel wind rose scheme can be applied and can be very useful tool for error classification.*

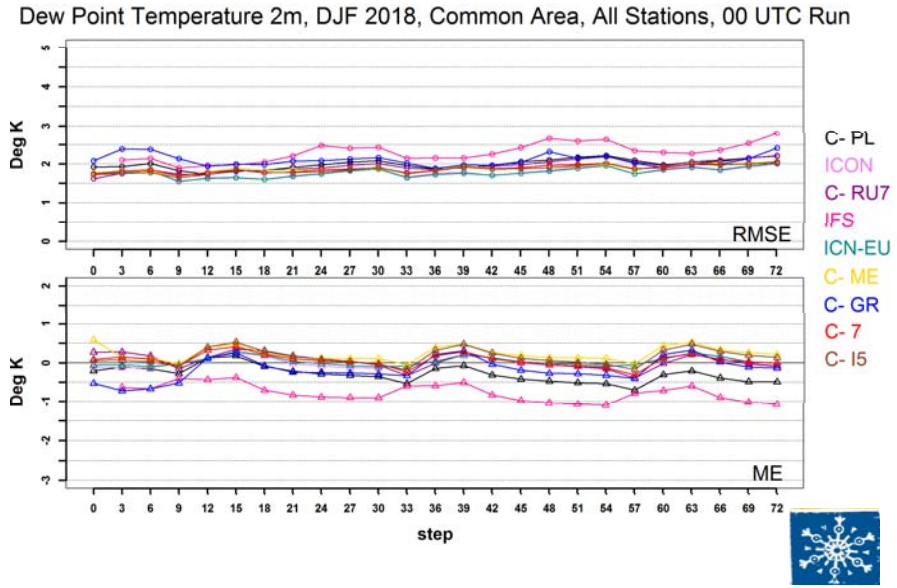

**Thank you for your attention**

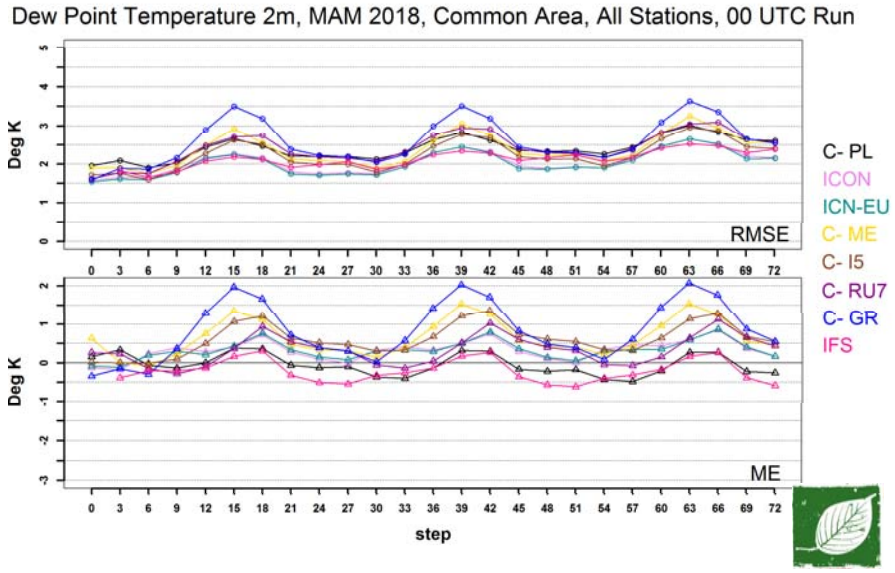
# T DEW 2m



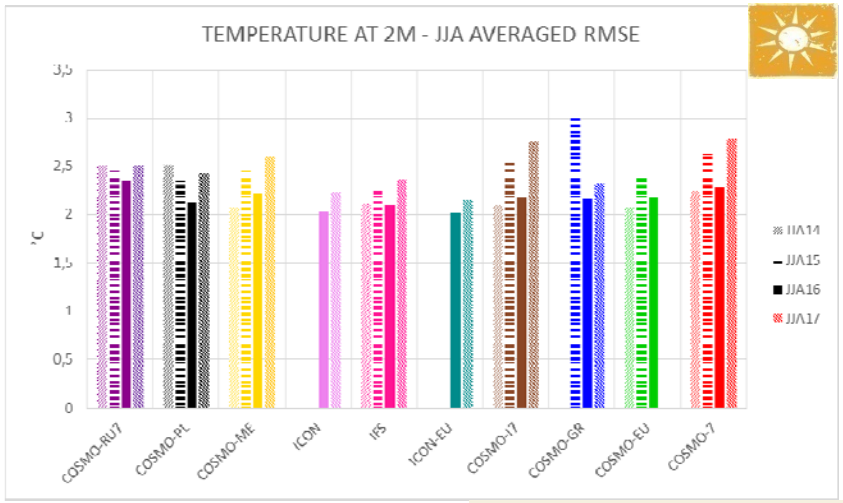



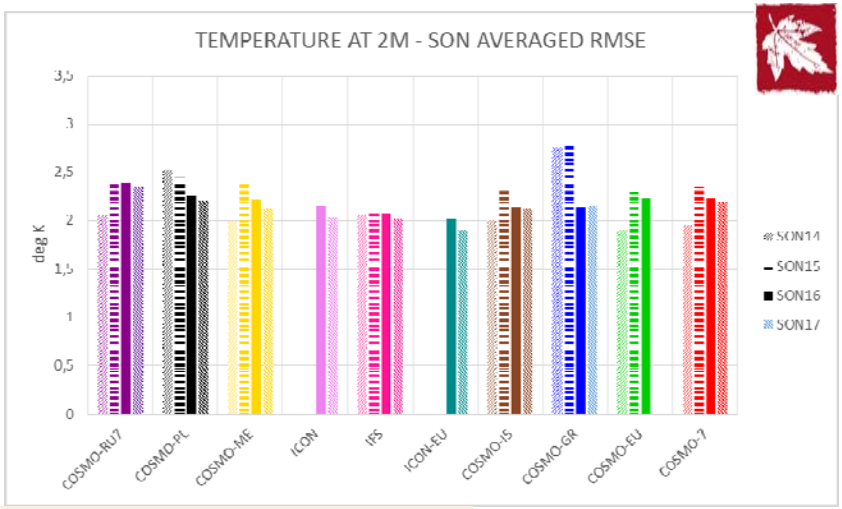

**JJA , MAM ME RMSE diurnal cycle. ICON, ICON-EU bias >0**



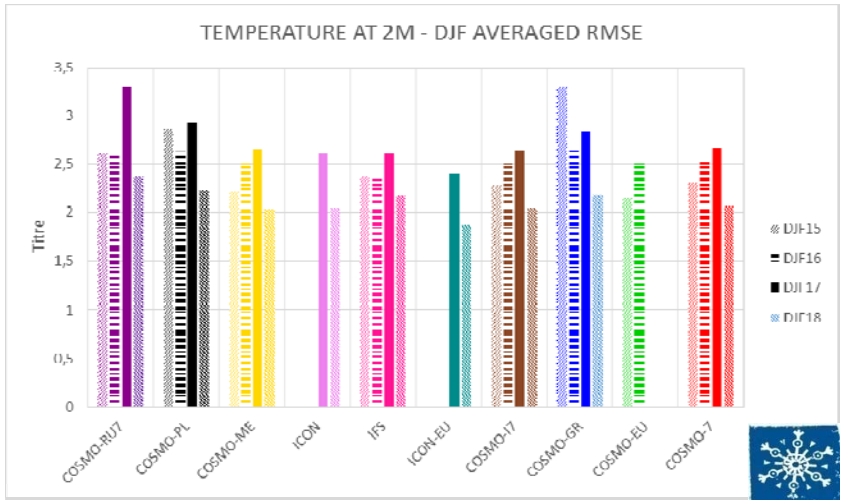



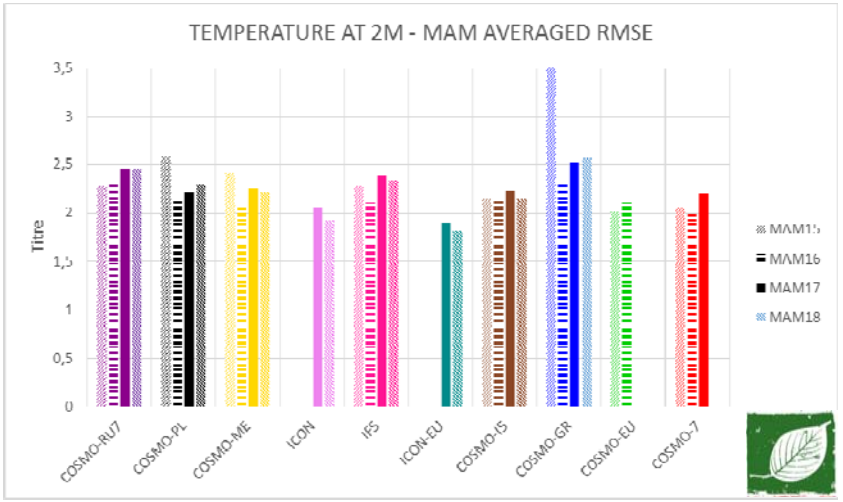
# Temp Seasonal TRENDS



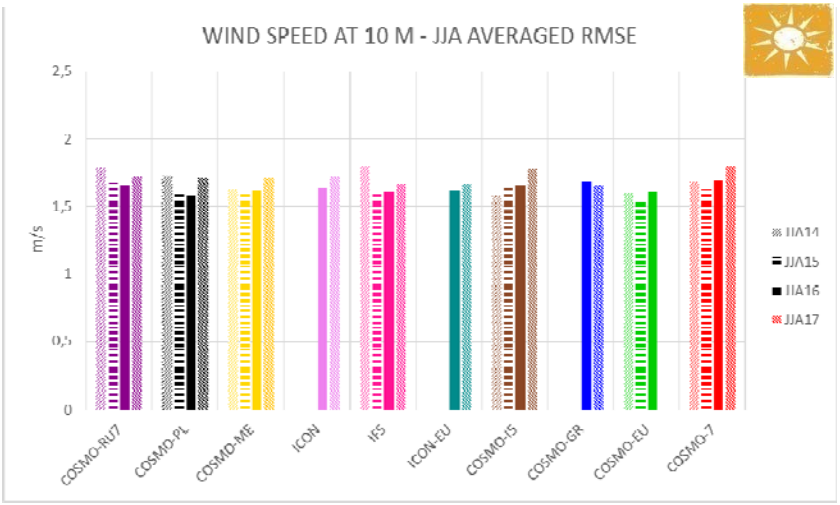



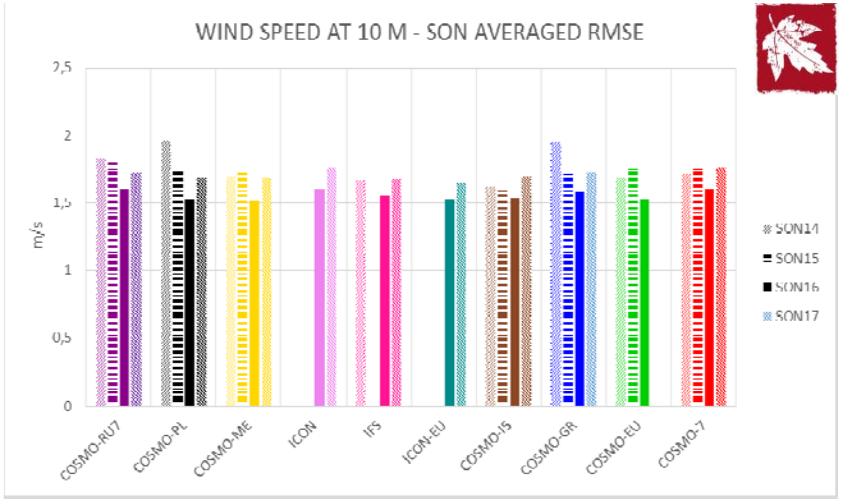

**Increase in JJA**



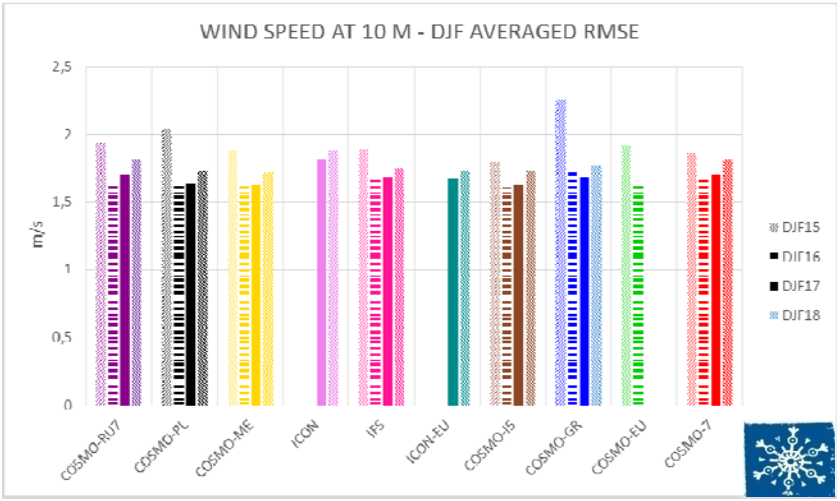



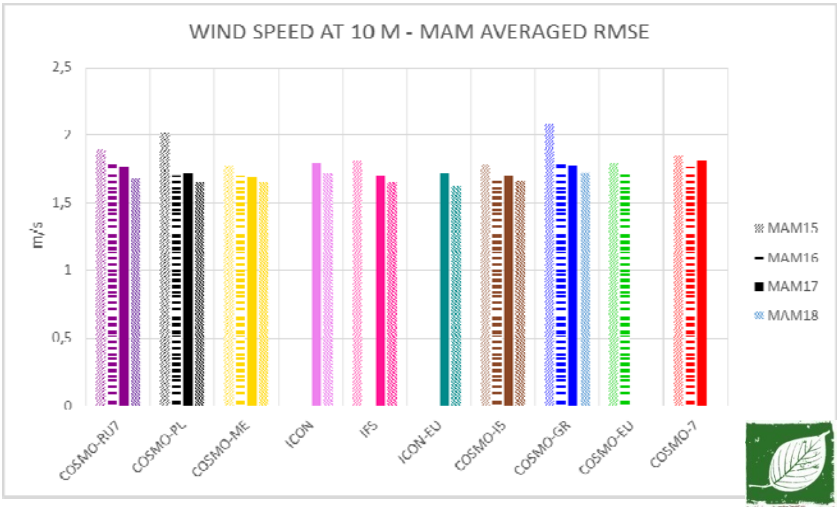
# Wind Speed Seasonal TRENDS




**Decrease in MAM**



# Conclusions I

- **We investigated the errors of COSMO models over common datasets from JJA 2017 to MAM 2018.**
- **The results showed that the general models behaviour did not significantly change from the previous years.**
- **The summertime JJA17 most RMSE scores increased compared to last year, possibly due to the extremely warm and dry summer.**
- **On the other side, winter DJF17-18 RMSE scores were generally better.**
- **On Annual basis, T, T2m, Tdew, TCC RMSE scores improved**
- **Cloud cover Global Models ME is negative and lower than COSMO and the dichotomic analysis showed that this mostly due to the overestimation of lower thresholds cases. The CSI and ETS scores are worse for 25-75 threshold.**

# Conclusions II

- **The precipitation FBI scores show diurnal variability in JJA (as in previous years) with drizzle overestimation of global models especially IFS**
- **Extremal Forecast Index was better for DJF and for lower precipitation thresholds**
- **HR models scores for CM2 are slightly better than CM1 especially for JJA**
- **CM3 Dry Area scores are worse than CM1 especially for TCC, Tdew and Wind Speed (mainly DJF which is the season that differs from CM1). Can this be a conclusion (together with the JJA CM1 results) that in drier regimes the models do not perform so well ?**