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

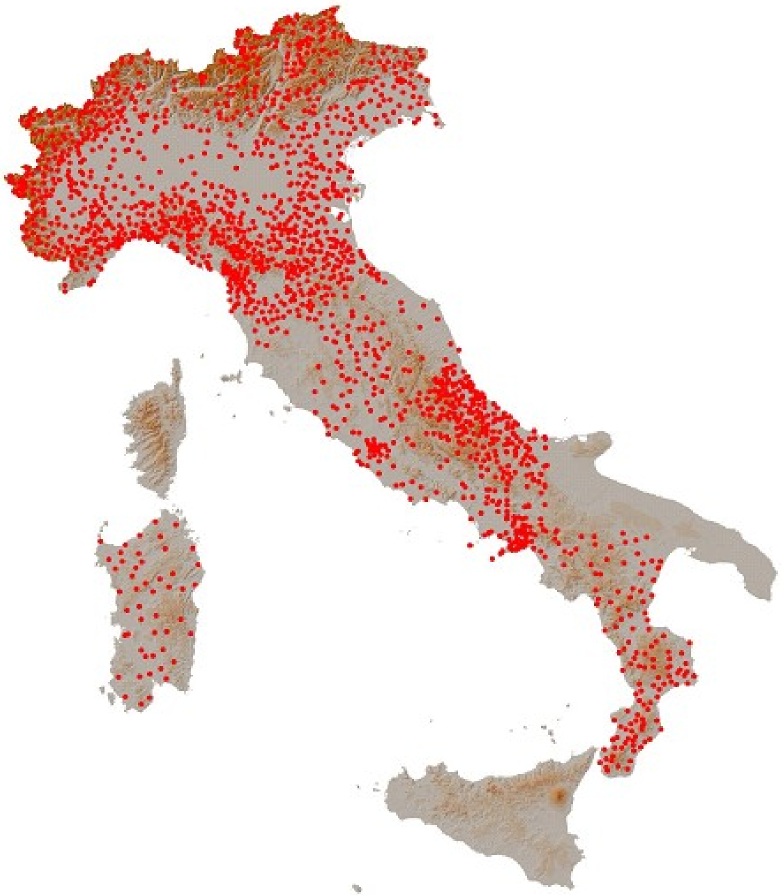
Elena Oberto
Naima Vela
Massimo Milelli

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.

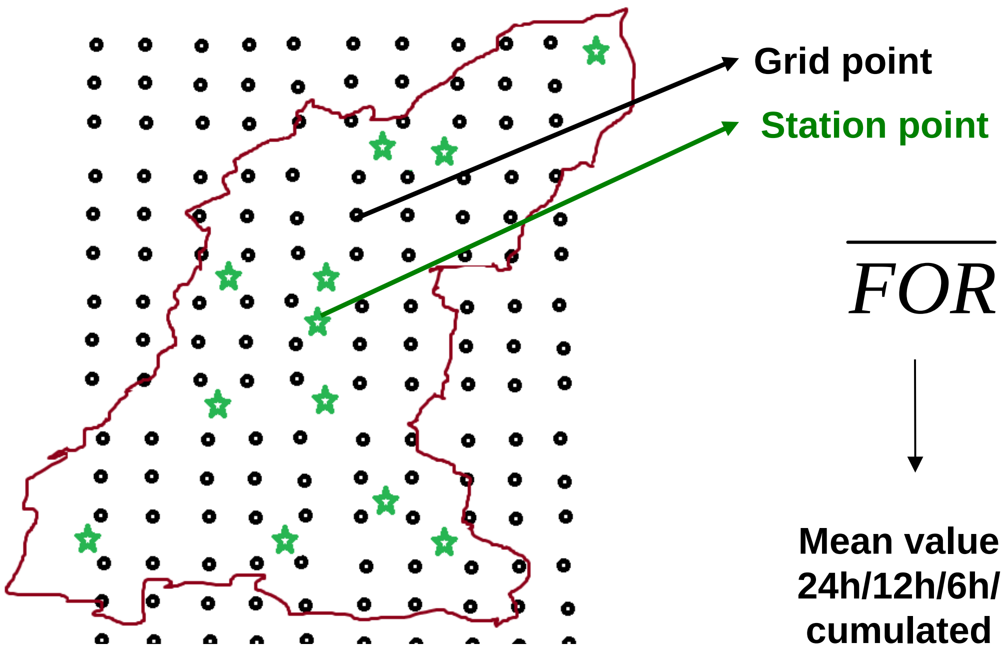
Cosmo General Meeting 2015 – Wroclaw (Poland)

The methods

Precipitation- high resolution network → problems with the data: dataset not stable in time, data not always reliable..



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



$$\overline{FOR} = \frac{1}{N}\sum_{i=1}^{N}(for)_i$$

Mean value 24h/12h/6h/ cumulated precipitation forecasted

Number of grid points inside alert area

i-th grid point precipitation value

$$FOR_{MAX} = MAX(for)_i$$

$$\overline{OBS} = \frac{1}{K}\sum_{i=1}^{K}(obs)_i$$

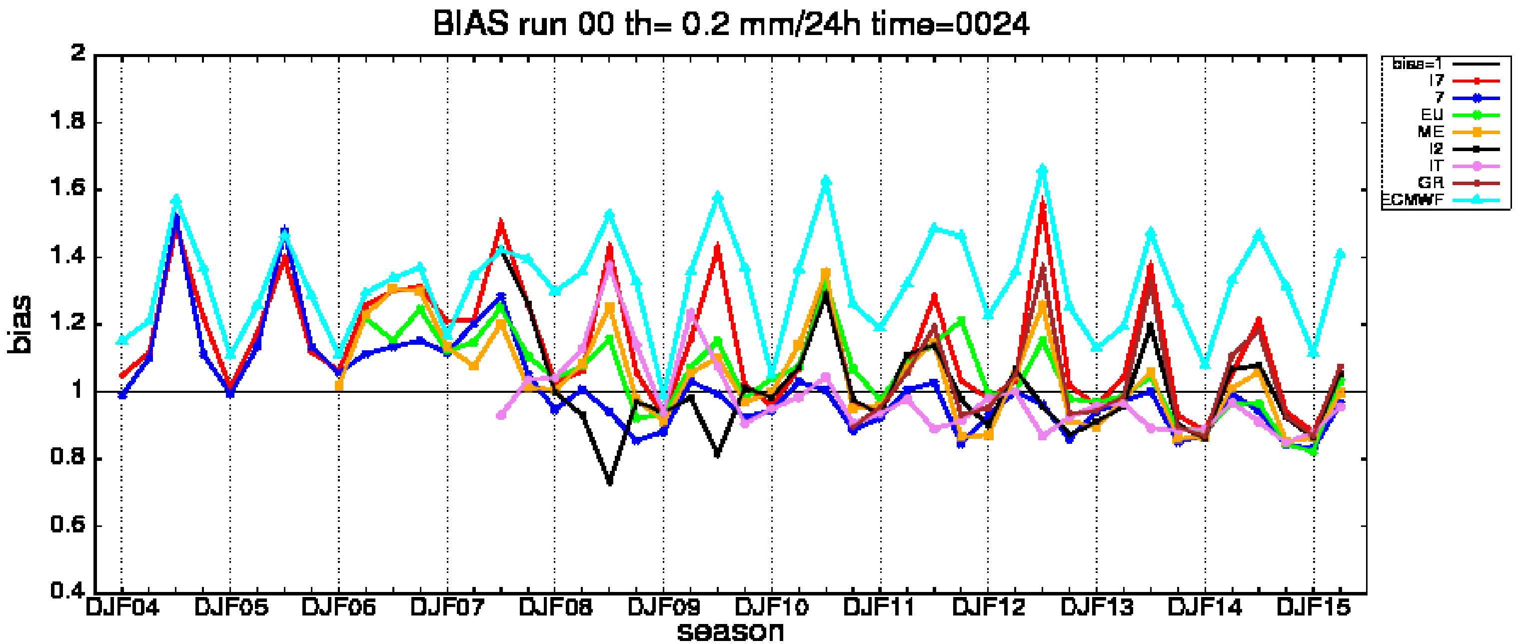
Mean value 24h/12h/6h/ cumulated precipitation observed

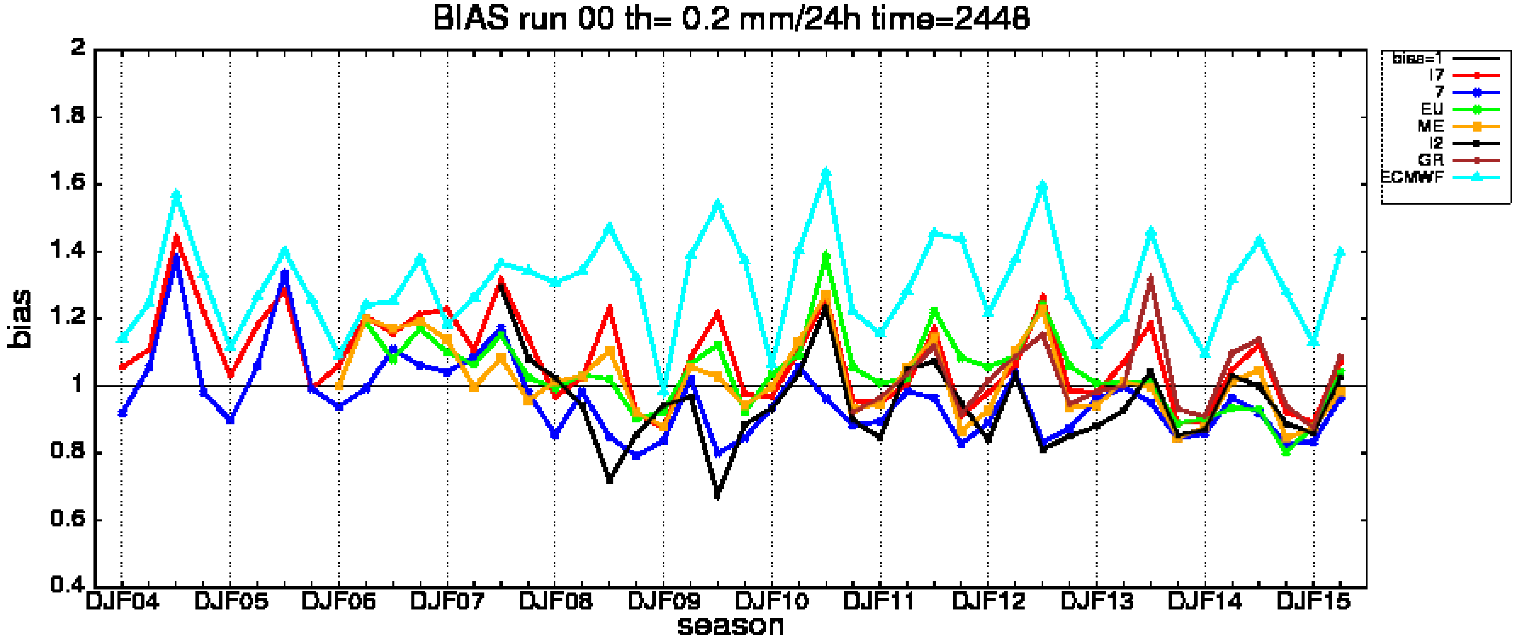
Number of station points inside alert area

i-th station precipitation value

$$OBS_{MAX} = MAX(obs)_i$$

LONG TREND PRECIPITATION with high resolution stations

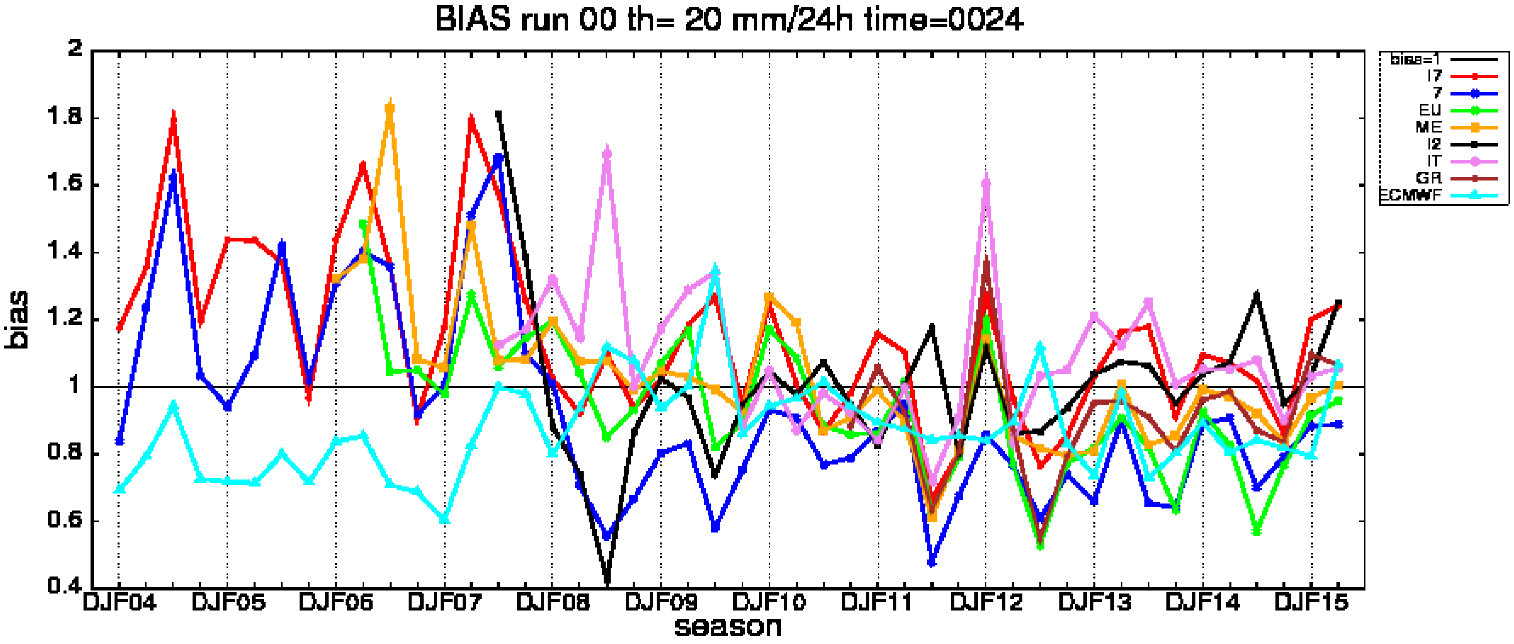




LOW THRESHOLDS

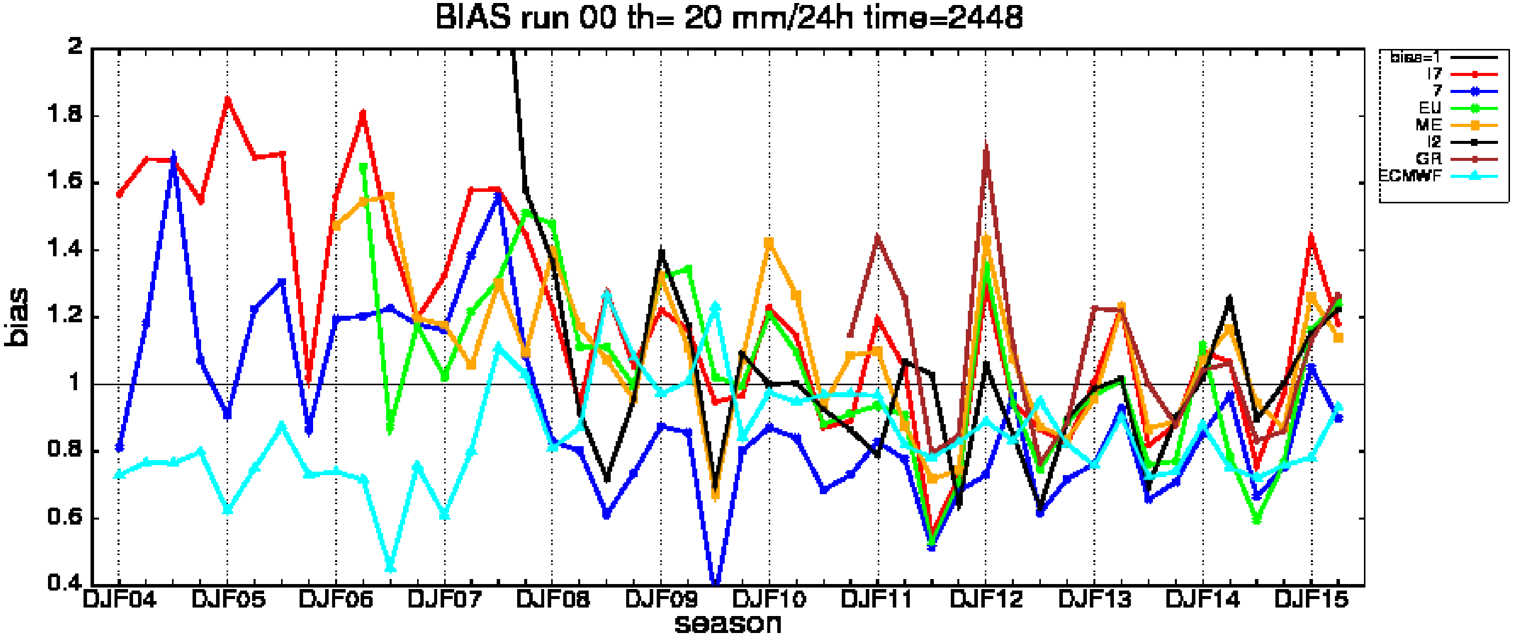
- 0.2 thres→ yes/no prec discriminant
- Seasonal cycle of error:summer over, winter under
- Ecmwf systematic overestimation
- During last years the spread has been reduced a little

LONG TREND PRECIPITATION with high resolution stations



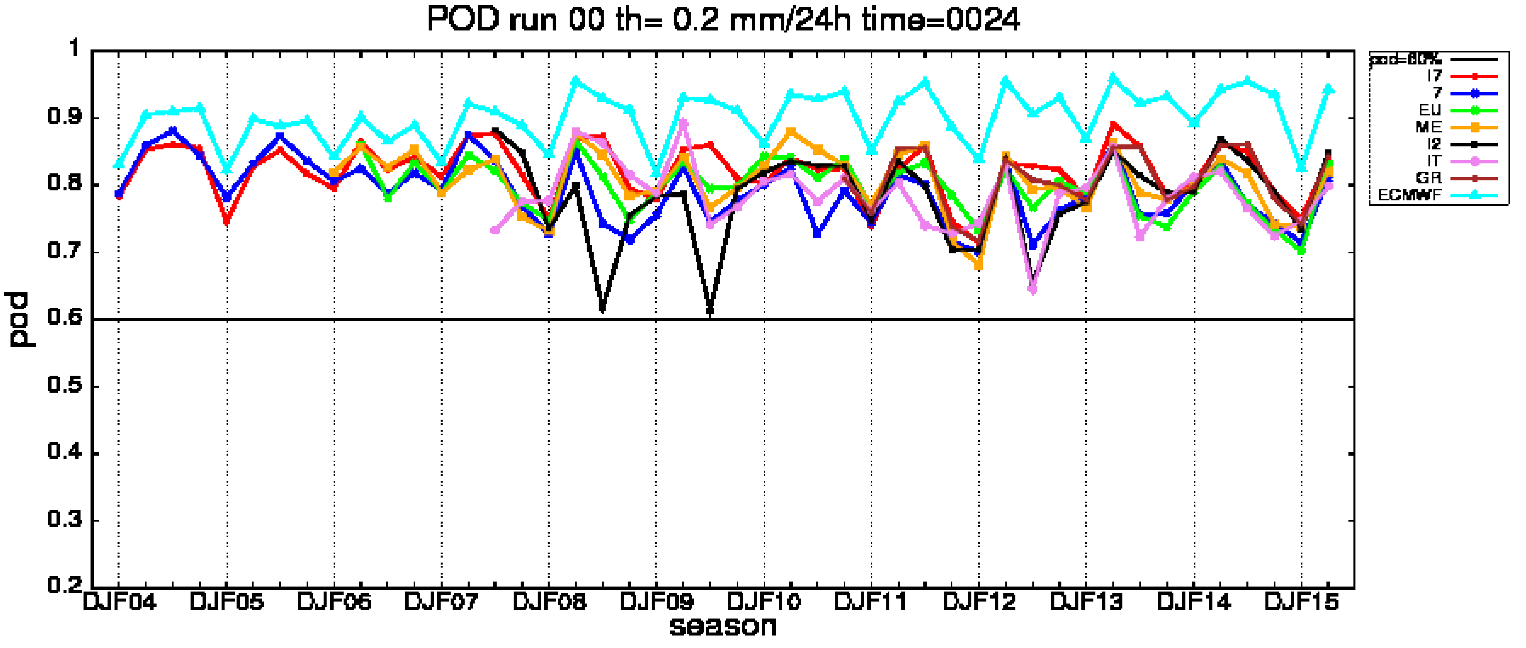
HIGH THRESHOLDS

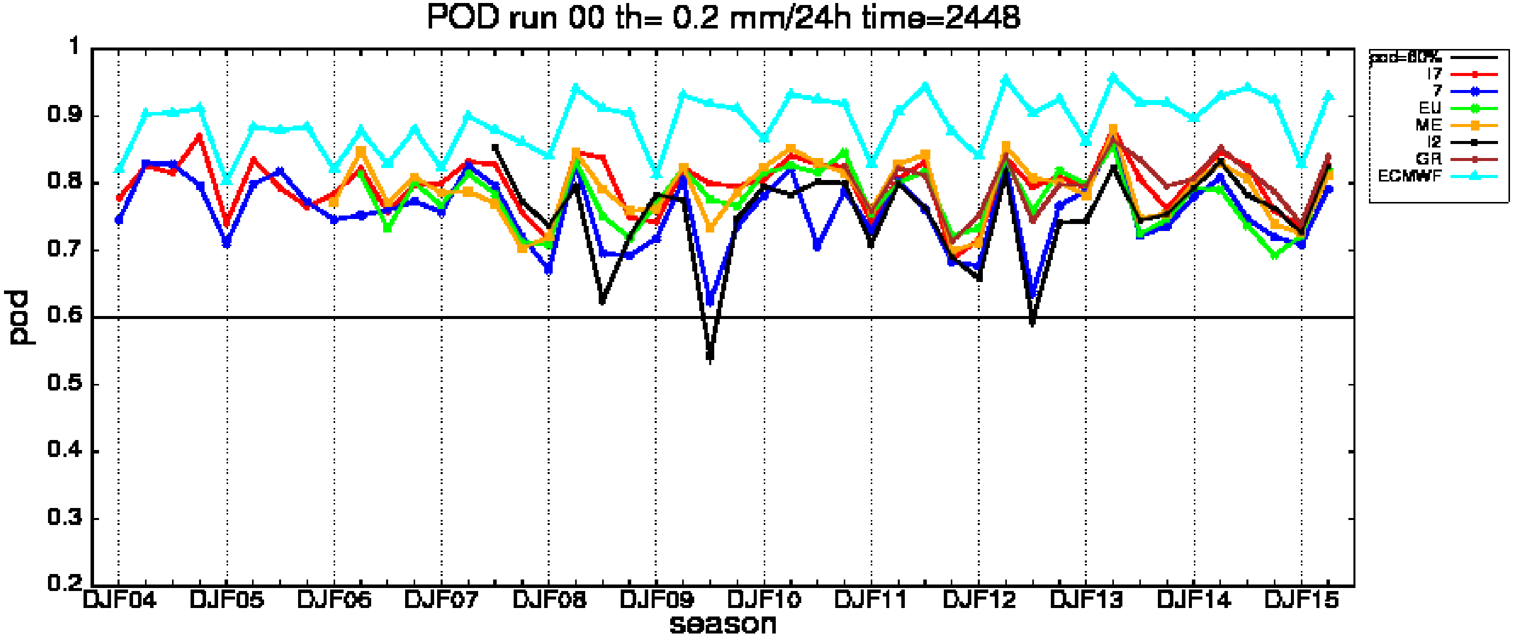
- Underestimation: 7, EU, ECMWF,GR
- Cycle error for I7 and ME



- The second day has got cycle error more pronounced
- The summer is the more difficult season

LONG TREND PRECIPITATION with high resolution stations



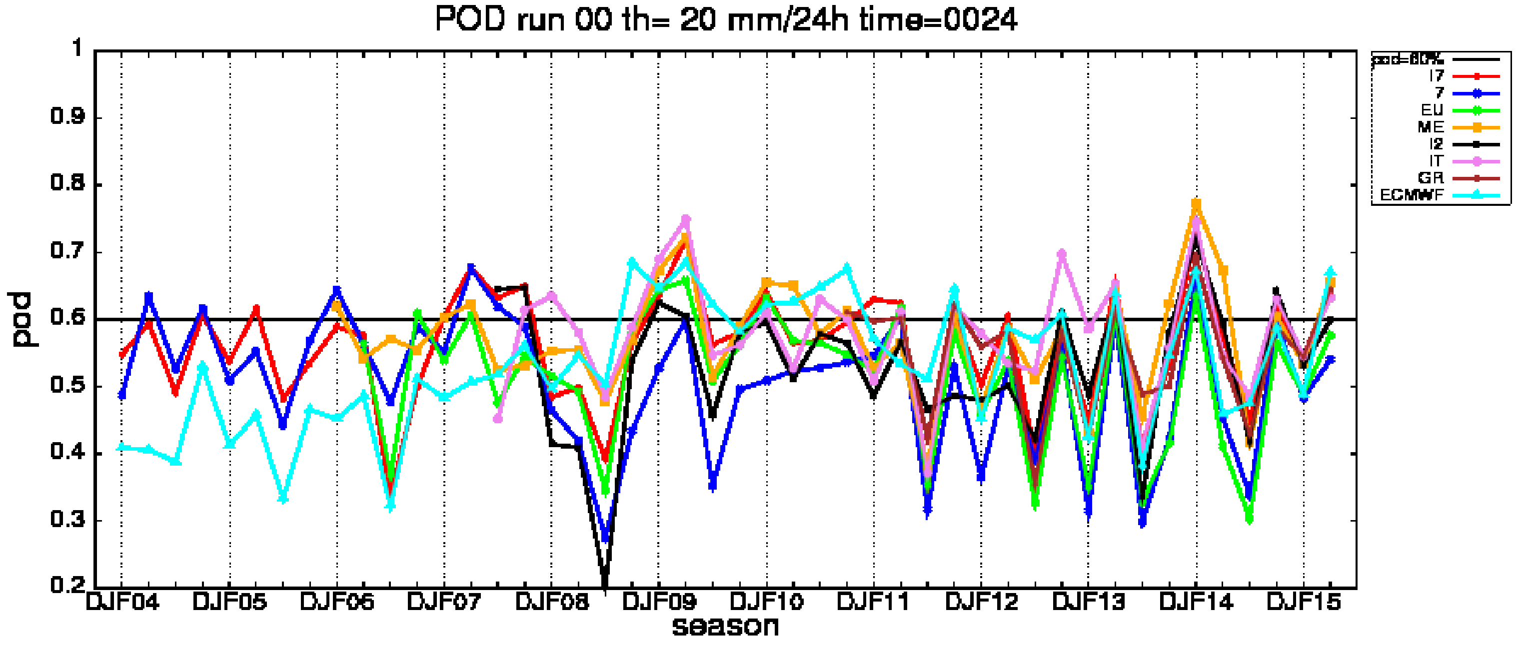


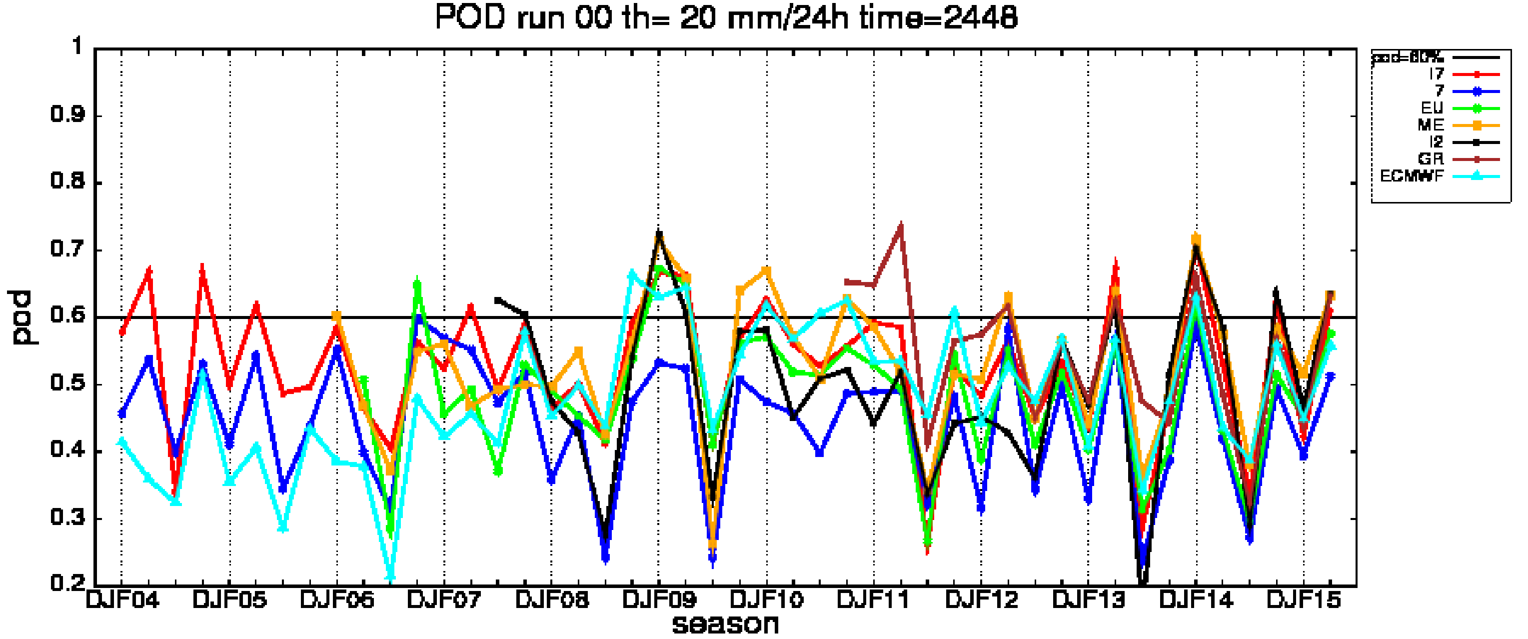
LOW THRESHOLDS

- Cycle error
- Summer is difficult

LONG TREND PRECIPITATION with high resolution stations

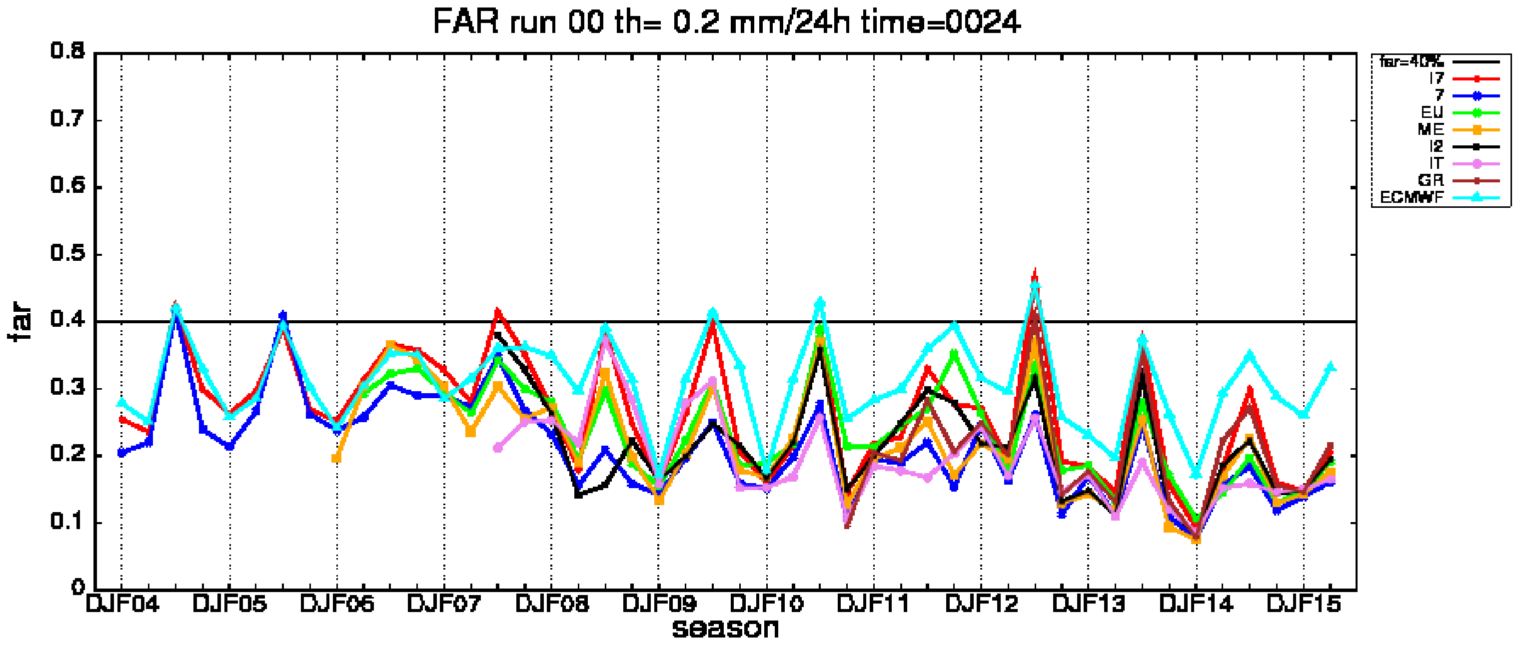
HIGH THRESHOLDS





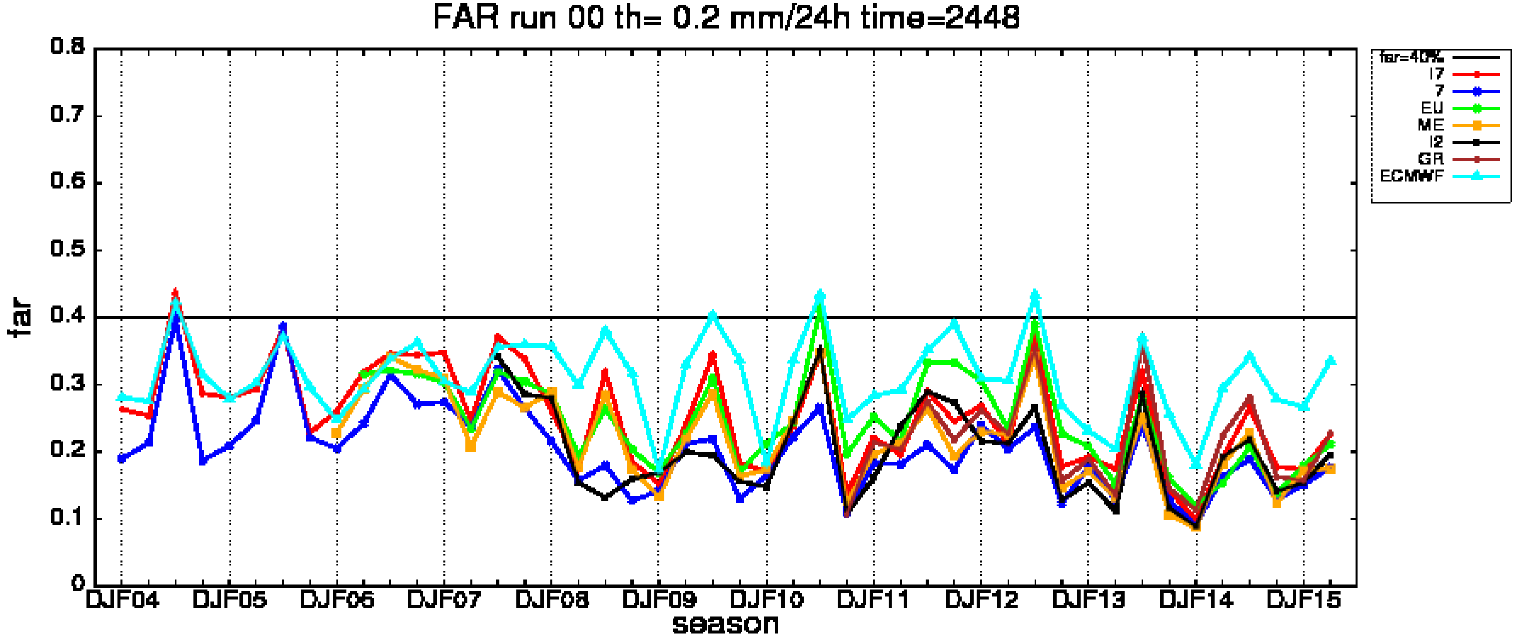
- •Cycle error with big spread
- •DJF14 warm, moist and rainy/snowy → strong dependency from pluviometric regime
- •Low POD in summer and winter

LONG TREND PRECIPITATION with high resolution stations

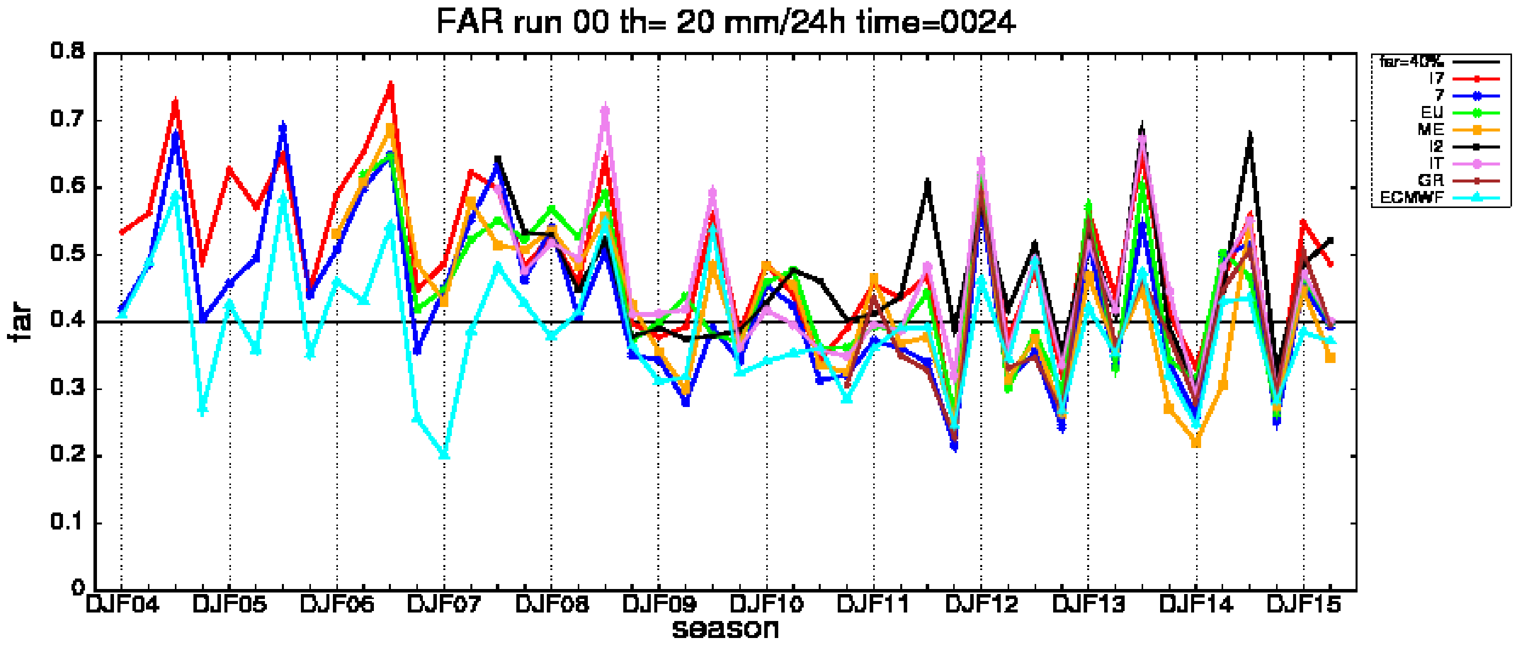


LOW THRESHOLDS

- Cycle error
- Summer is difficult

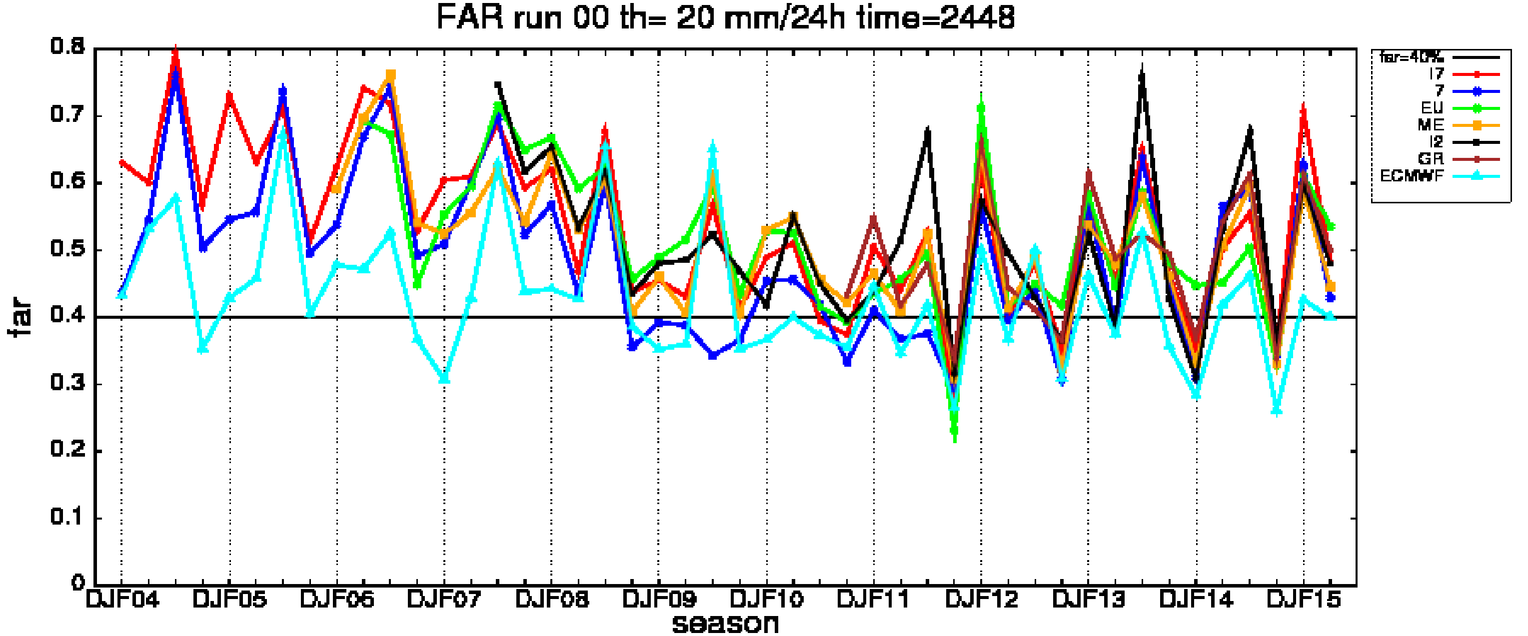


LONG TREND PRECIPITATION with high resolution stations

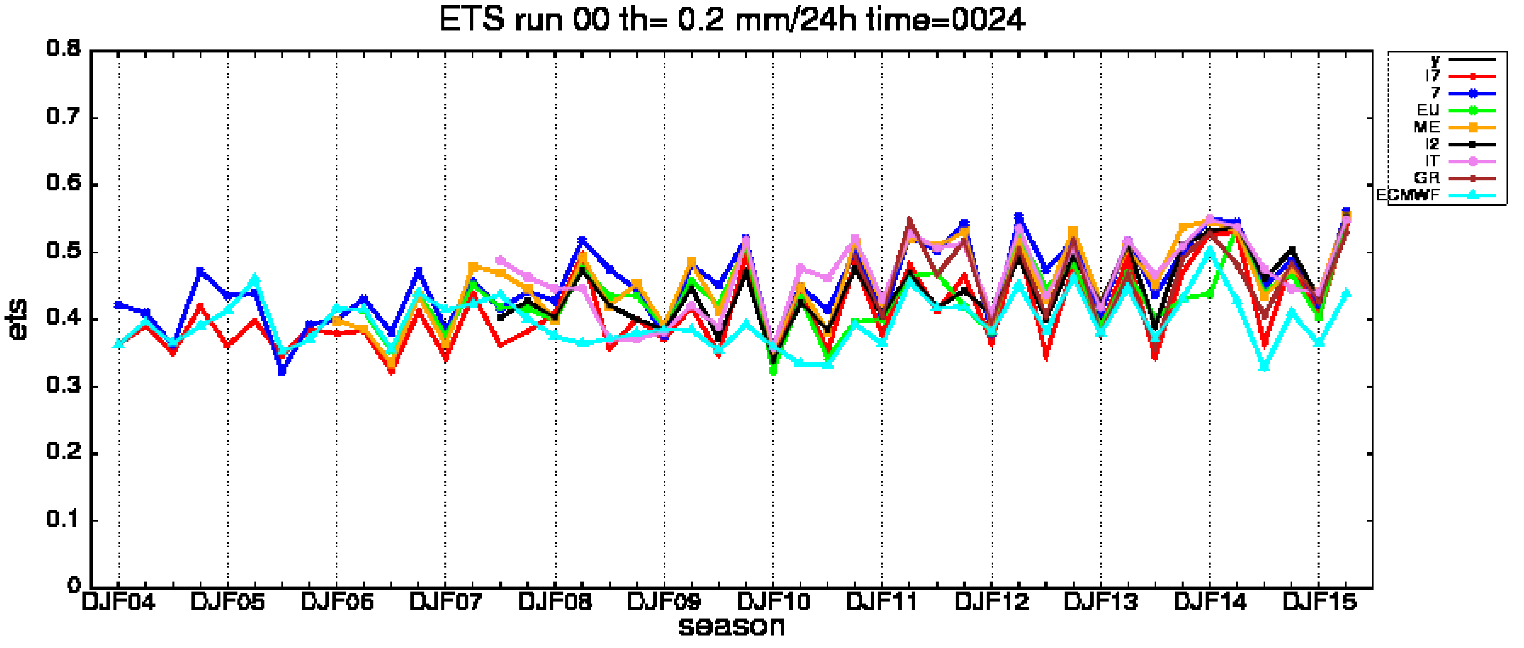


HIGH THRESHOLDS

- The results confirm POD

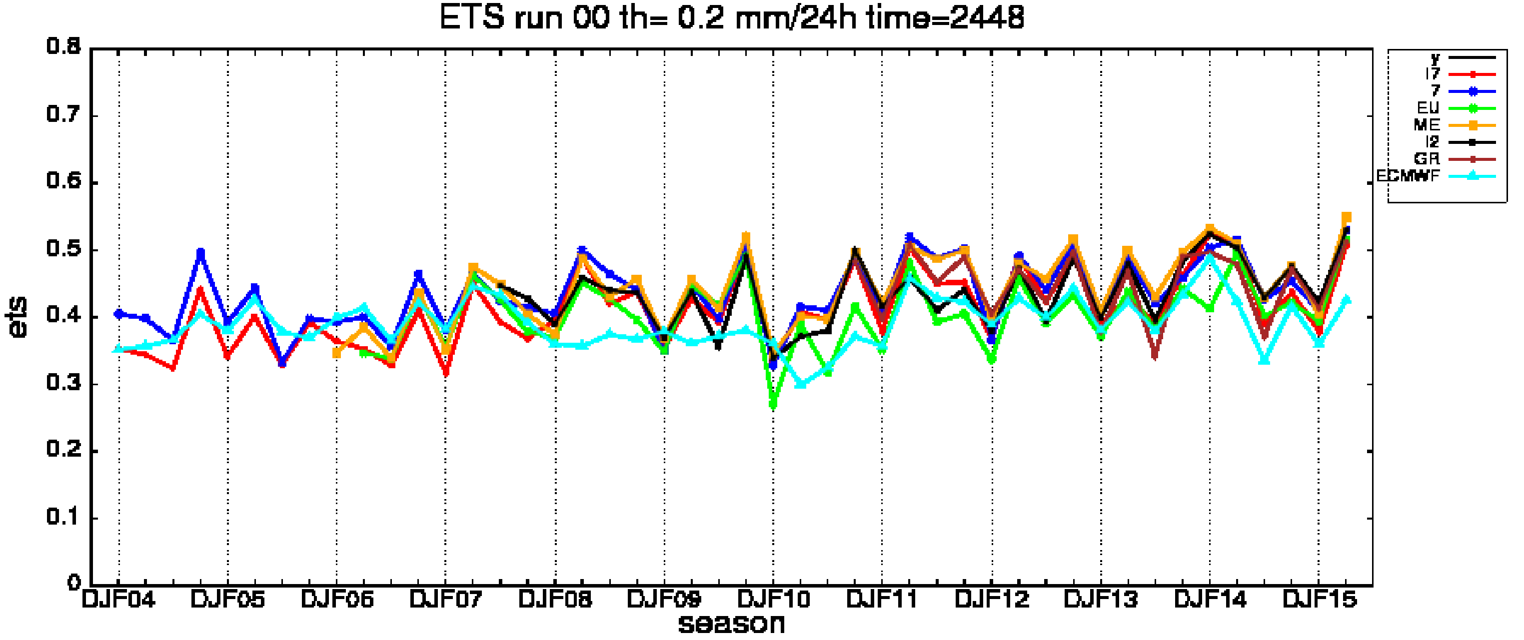


LONG TREND PRECIPITATION with high resolution stations

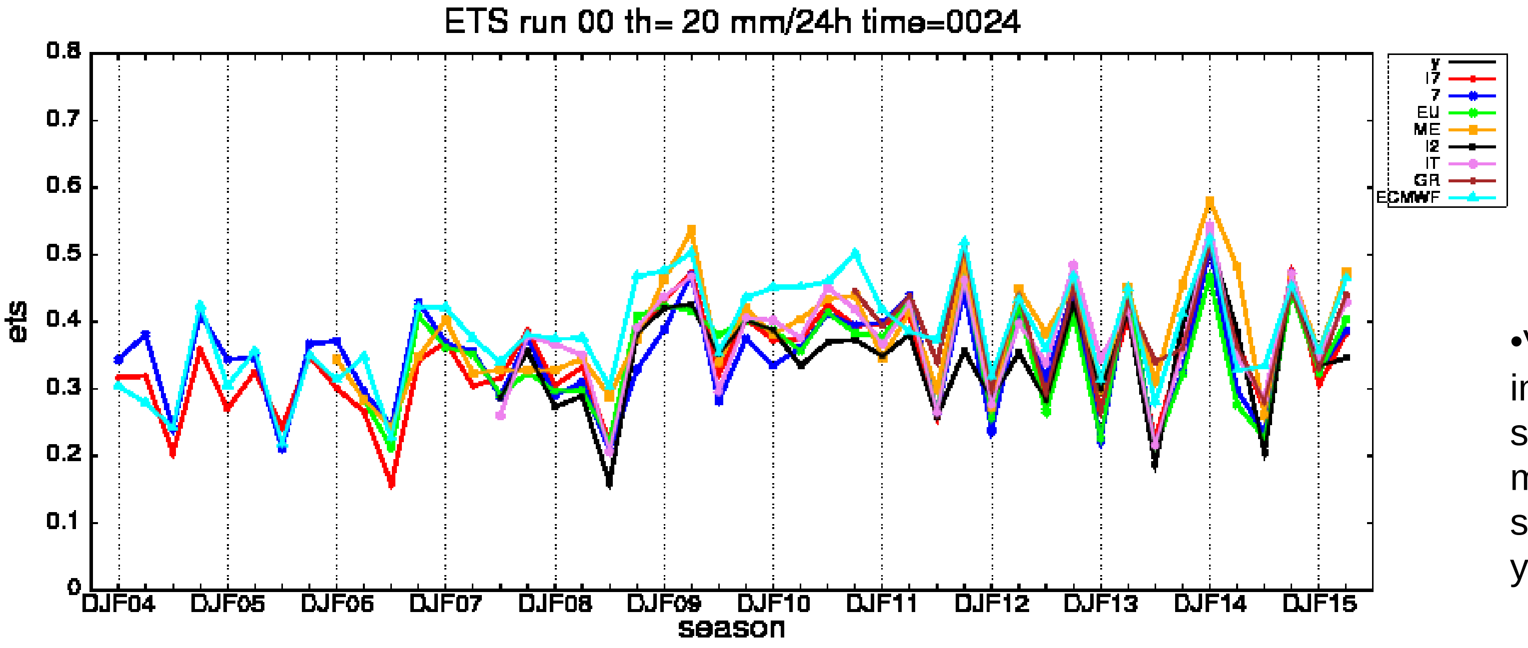


LOW THRESHOLDS

- Very slightly improvement in time

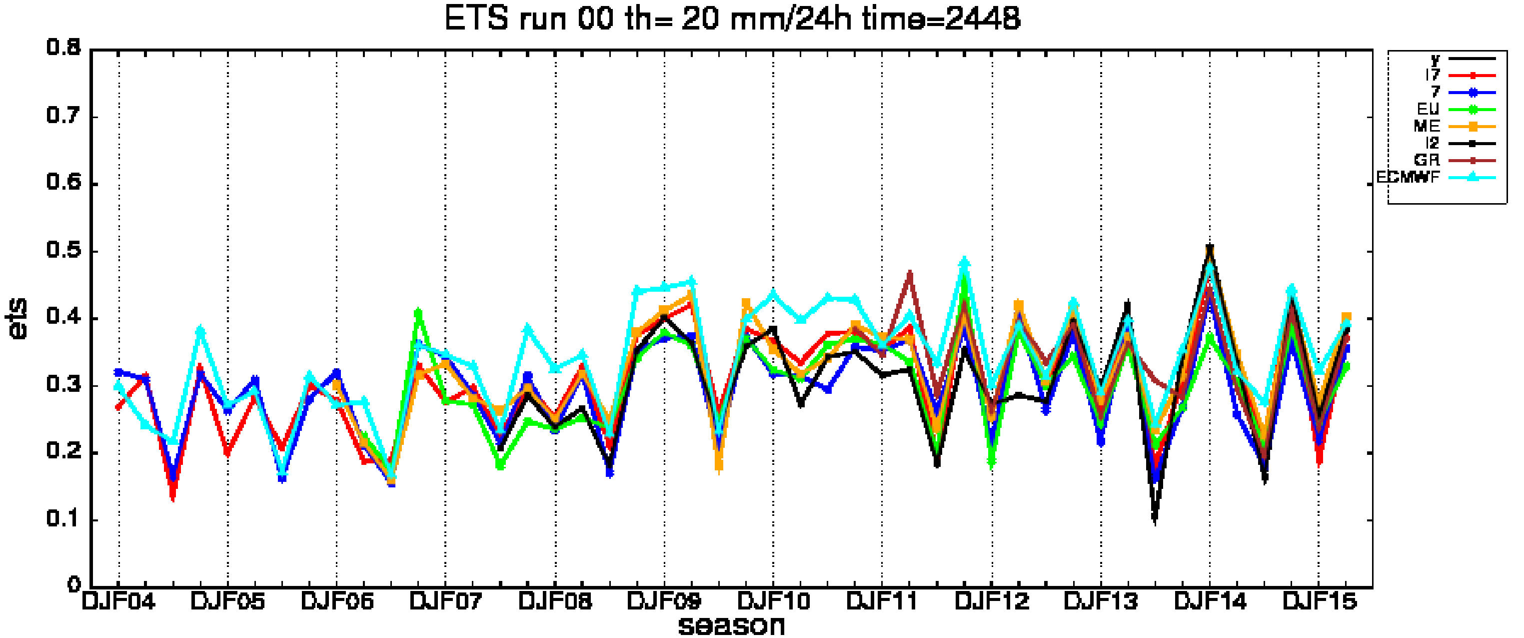


LONG TREND PRECIPITATION with high resolution stations

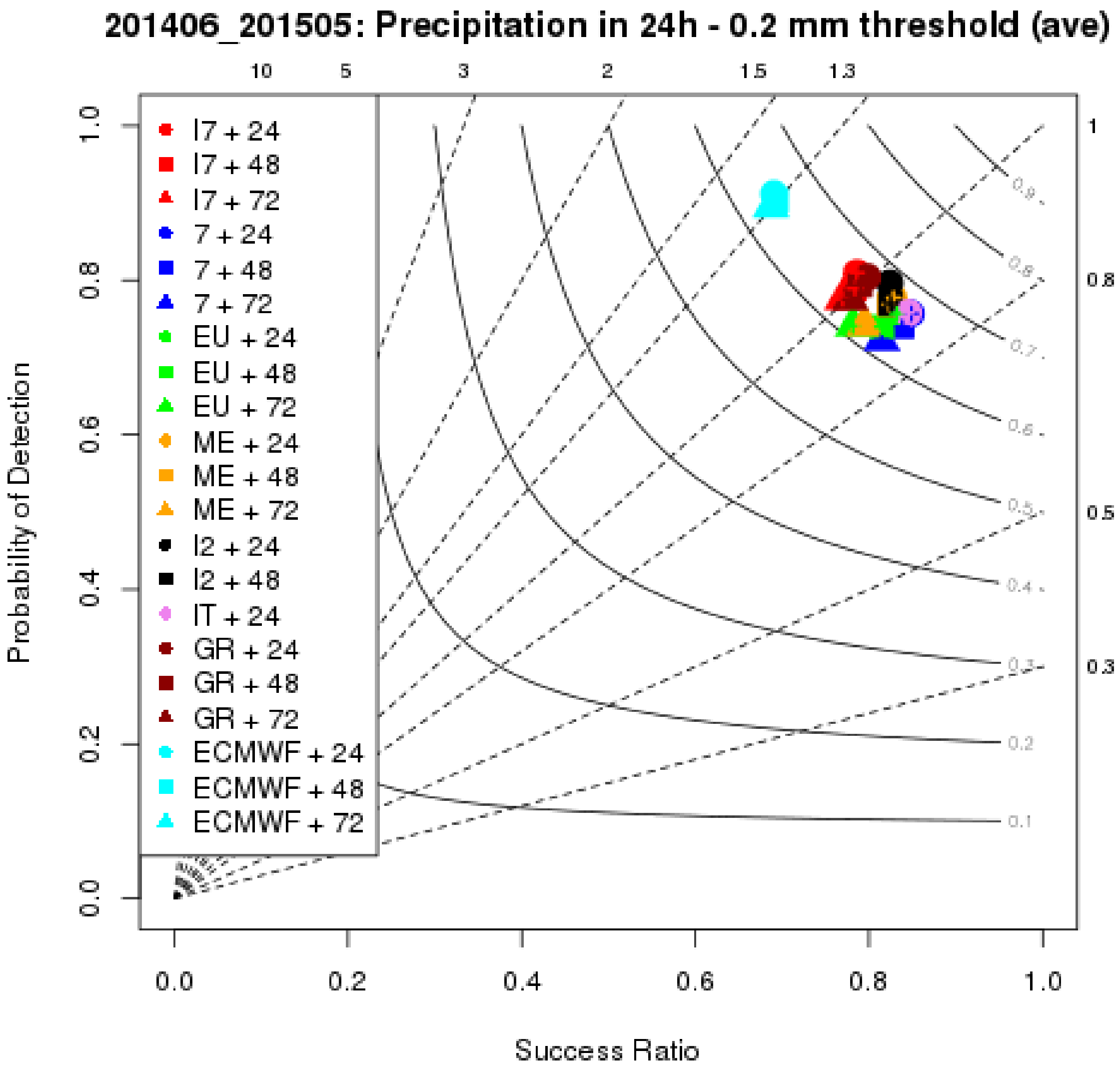


HIGH THRESHOLDS

- Very slightly improvement/quite stable in time but more pronounced spread during last years

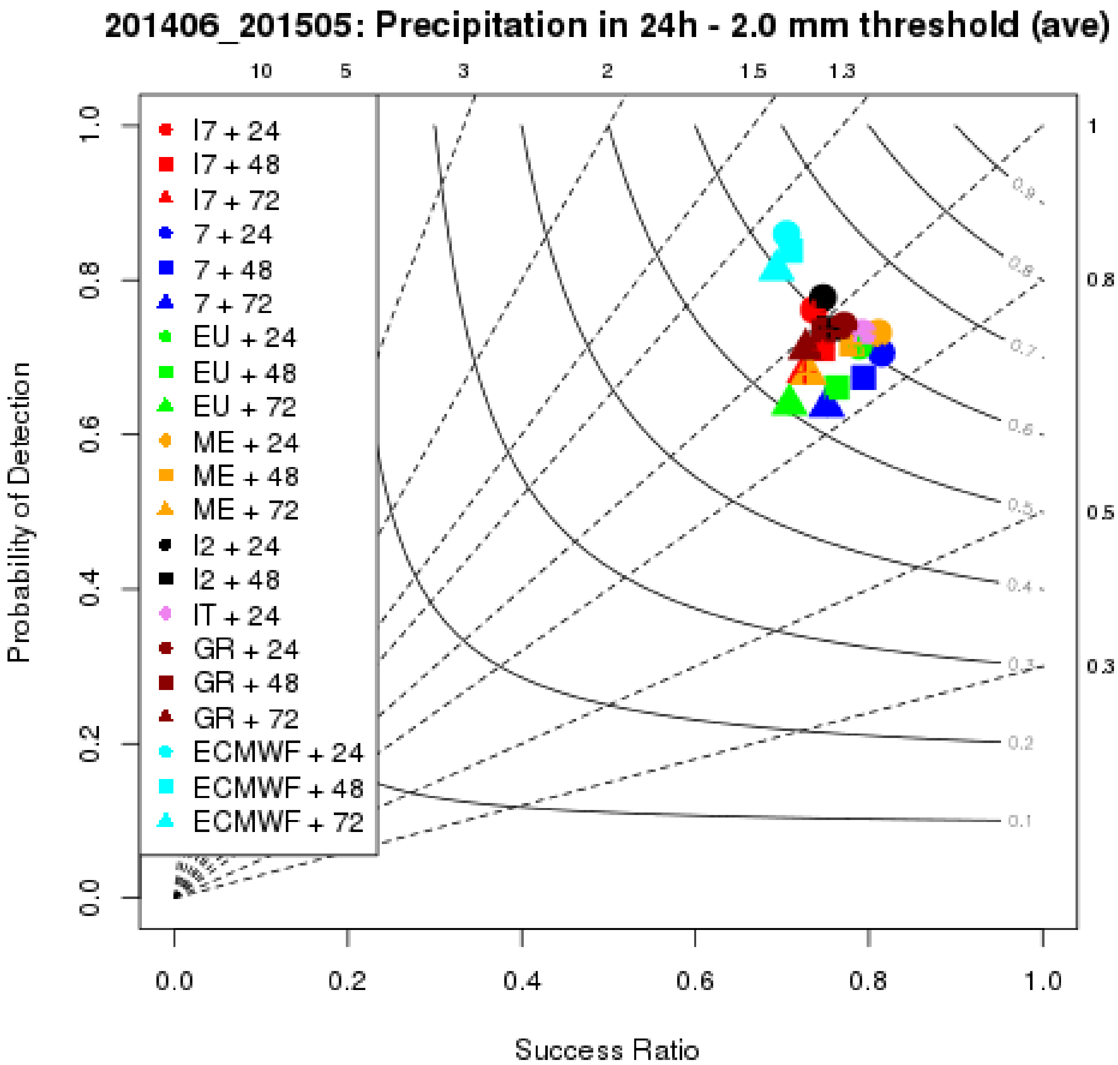


201406-201505: Average over area > 0.2 mm/24h



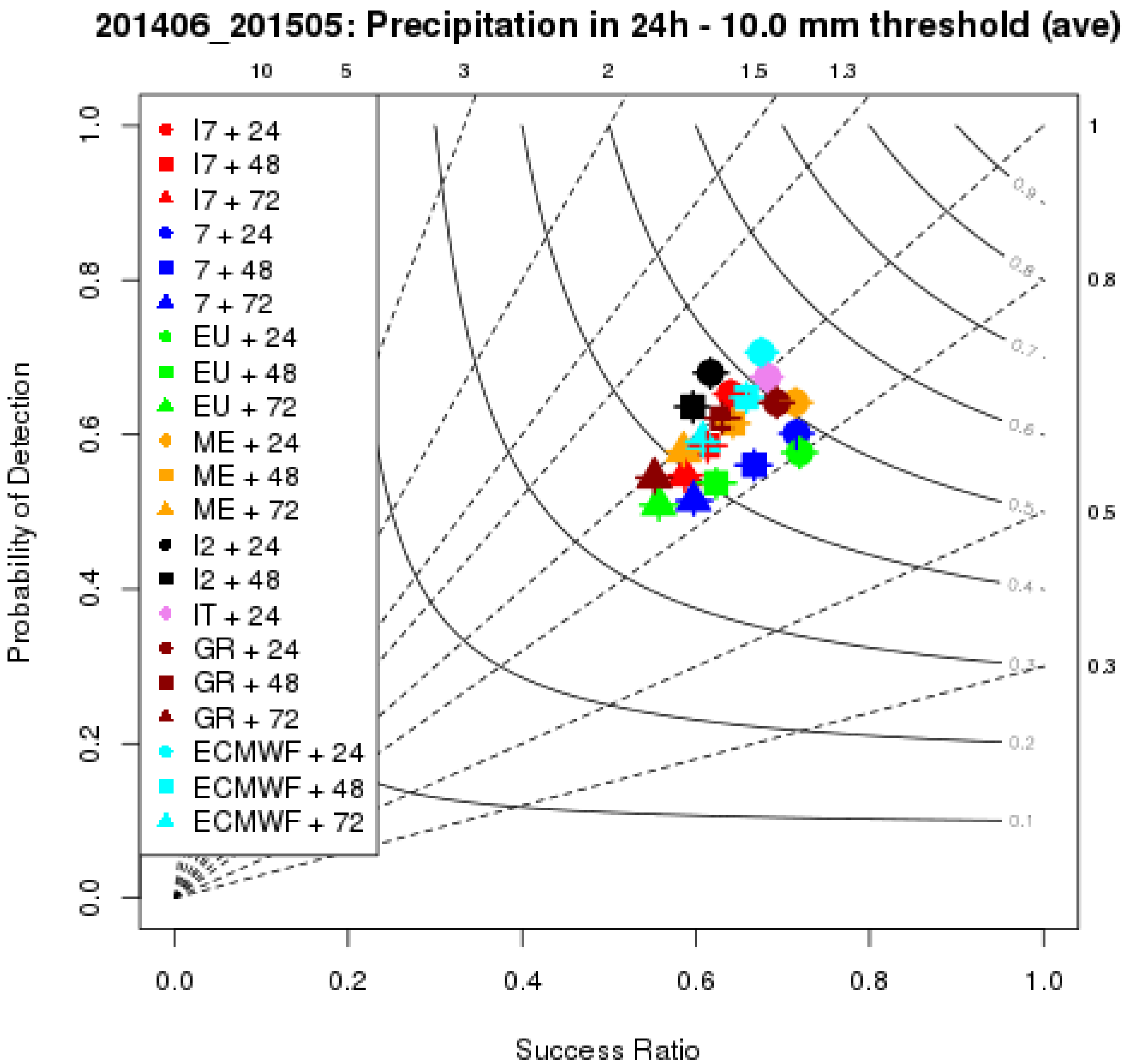
•Good/light underestimation

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



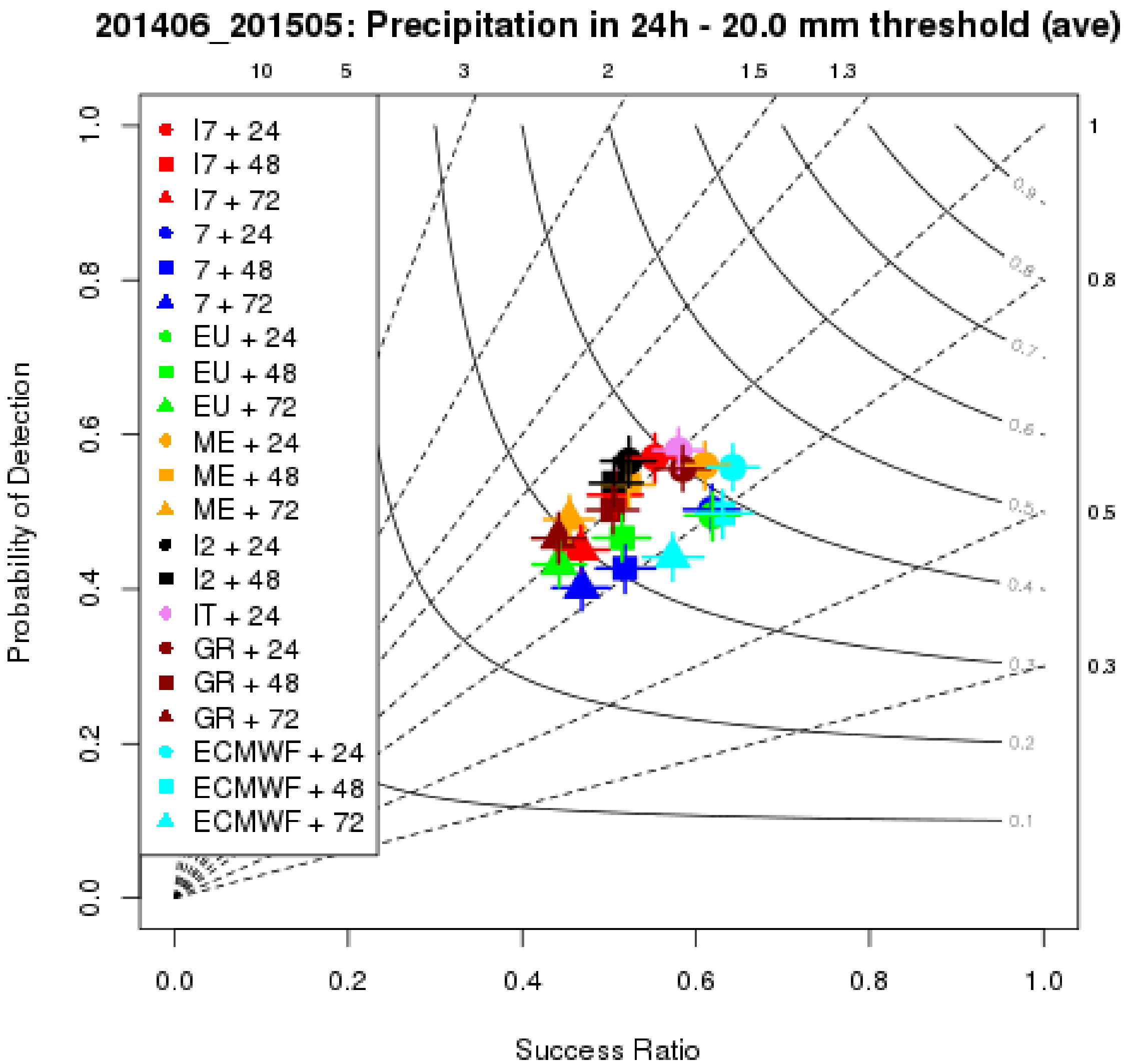
- Good/light underestimation

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



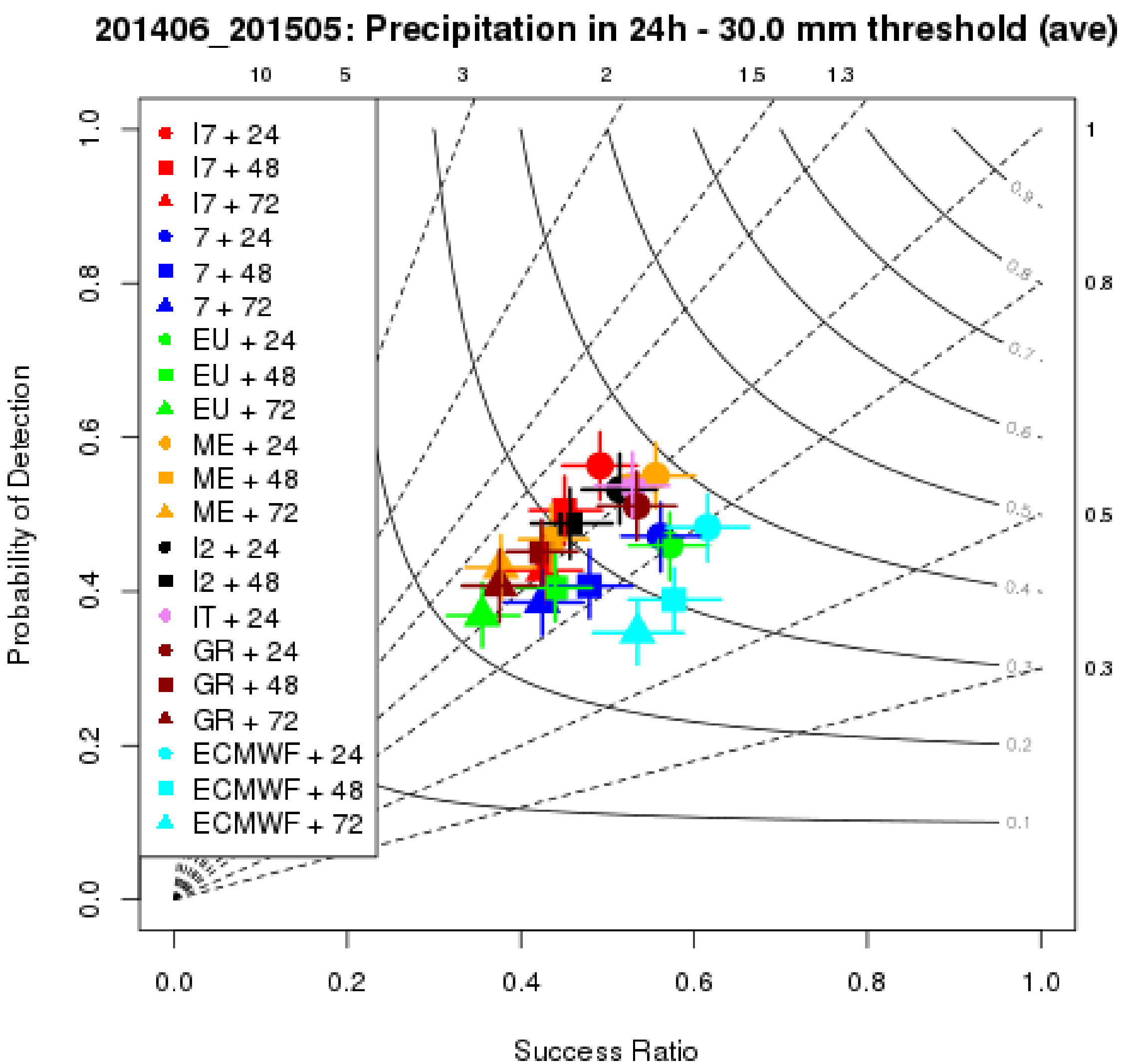
- Spread
- Good for I7,GR,ME
- Light underestimation for 7,EU

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



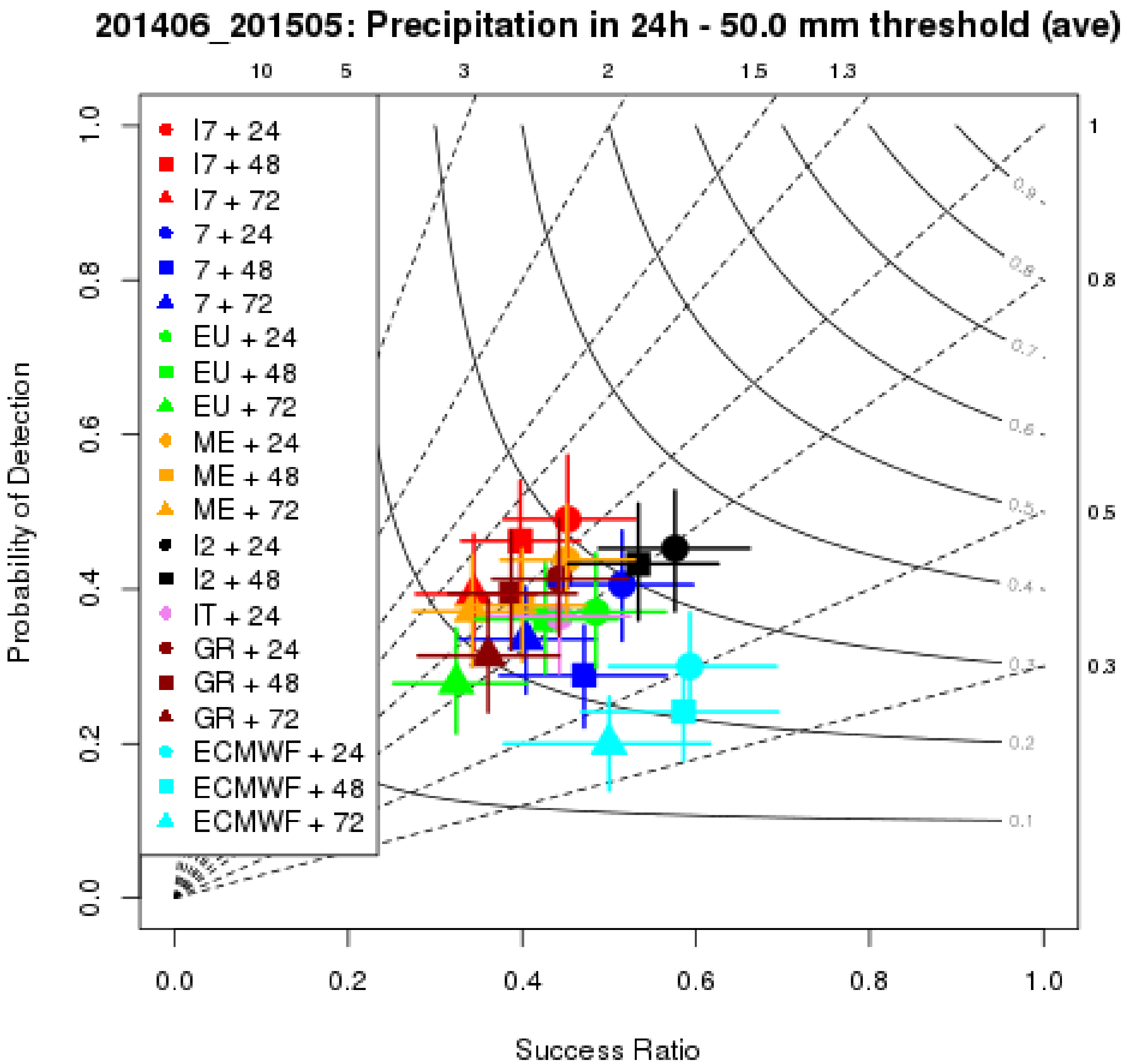
- Spread
- Good for I7,GR,ME,I2,IT
- Light underestimation for 7,EU,ECM

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



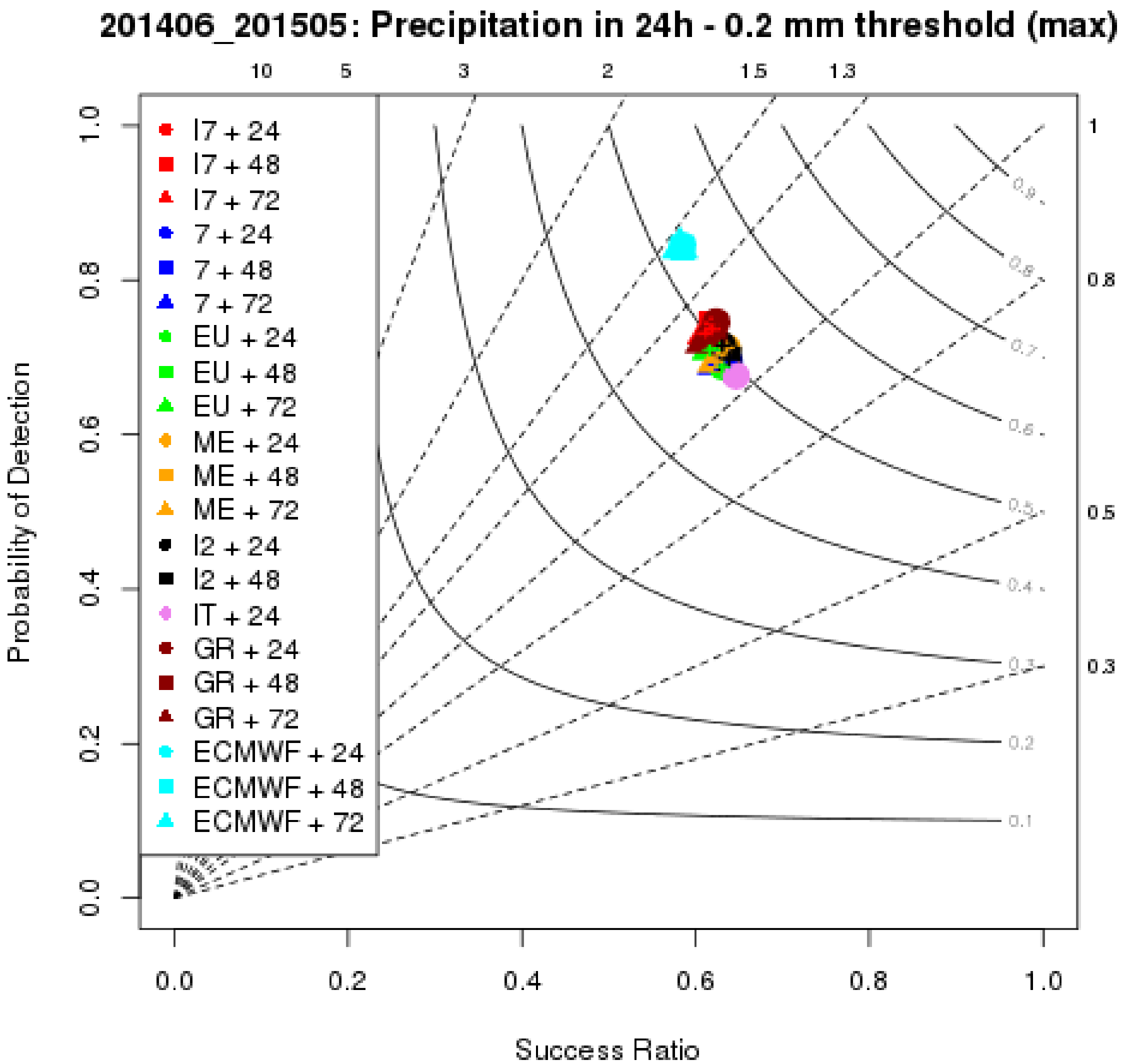
- Spread
- General skill decreasing
- Good for I7,GR,ME,I2,IT
- Light underestimation for 7,EU

201406-201505: Average over area > 50 mm/24h



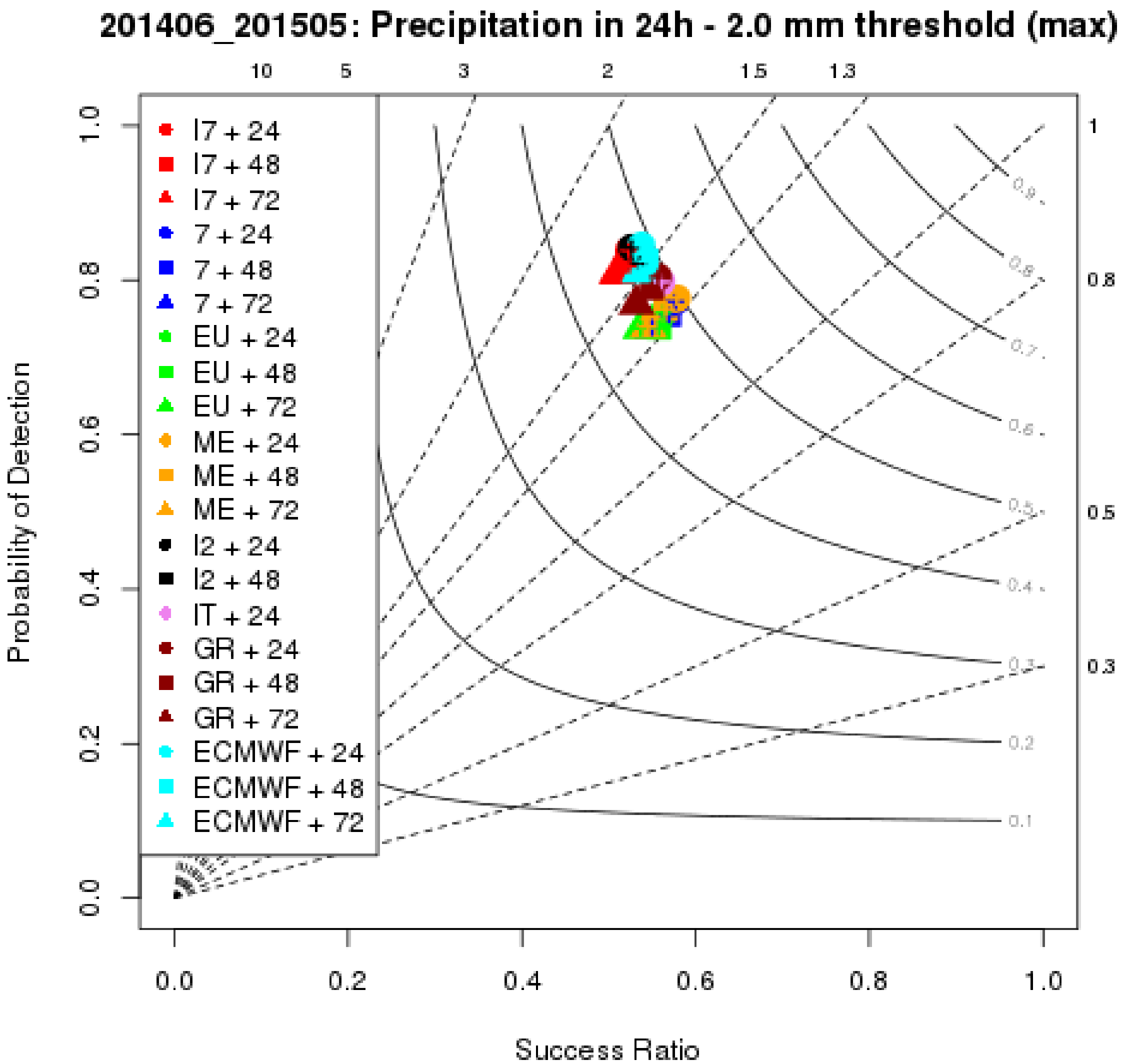
- Spread
- Good for I7,GR,ME,I2
- Light underestimation for 7,EU,IT

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



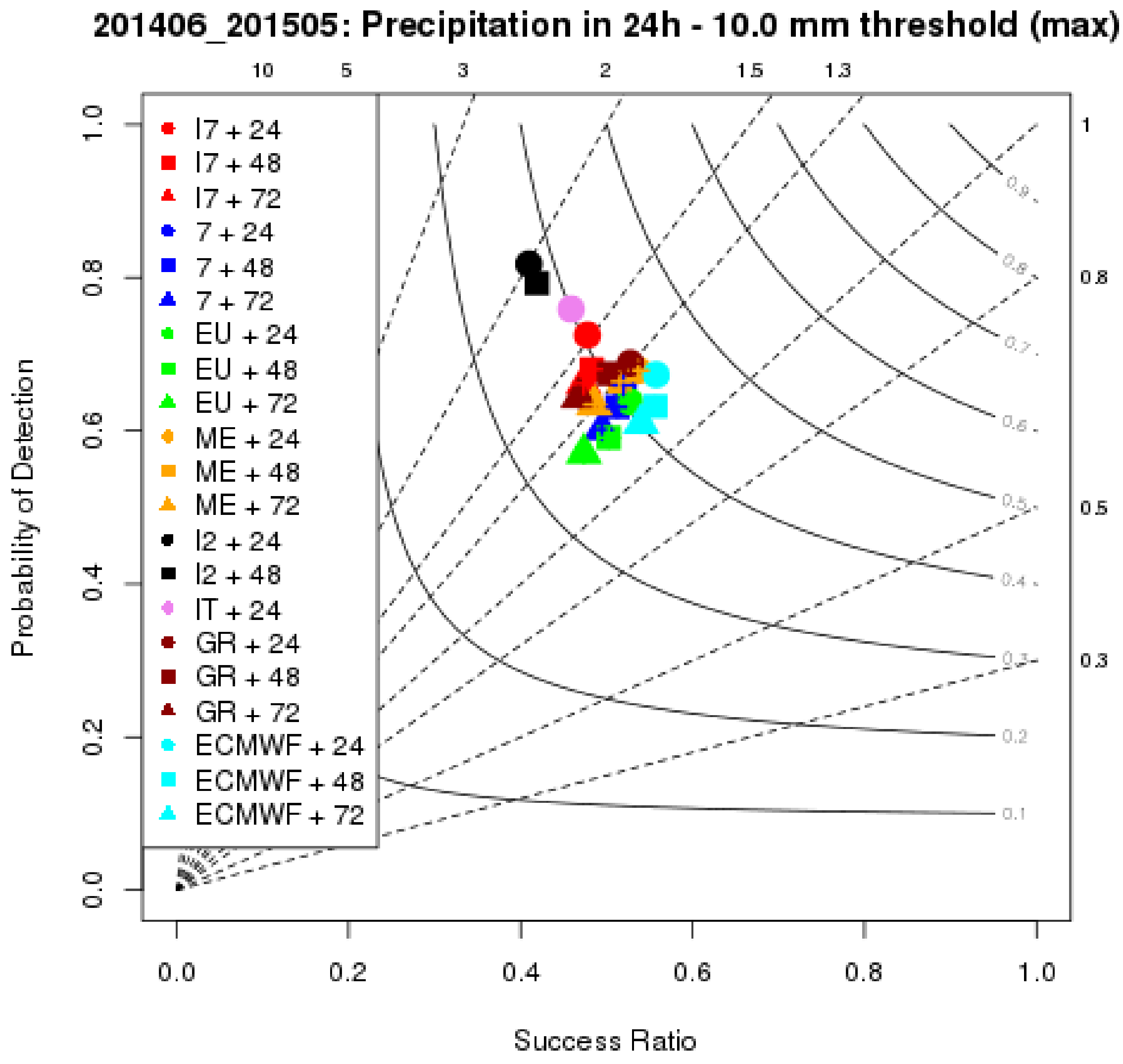
- Grouped dots
- Light overestimation

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



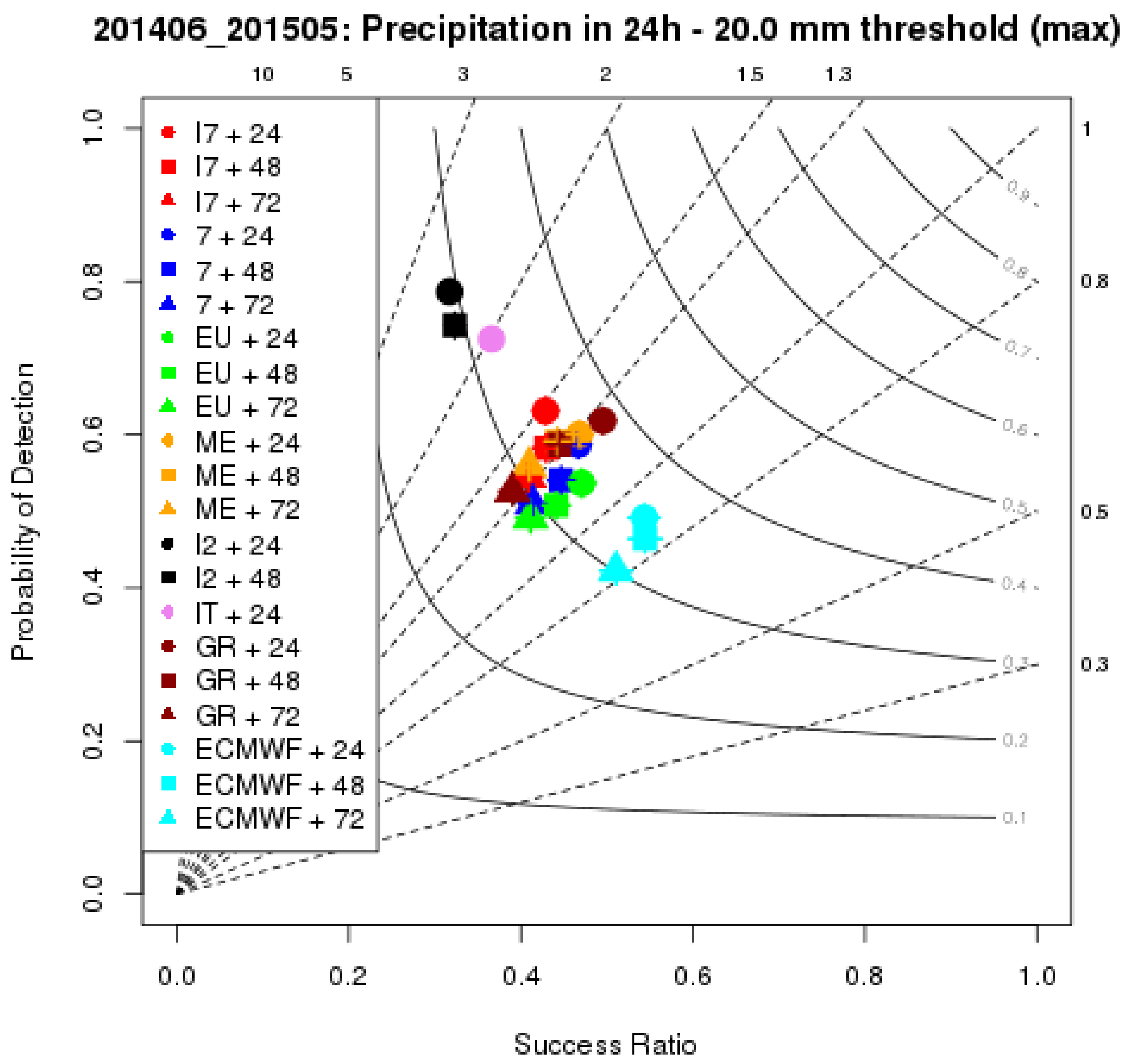
- Grouped dots
- general overestimation

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



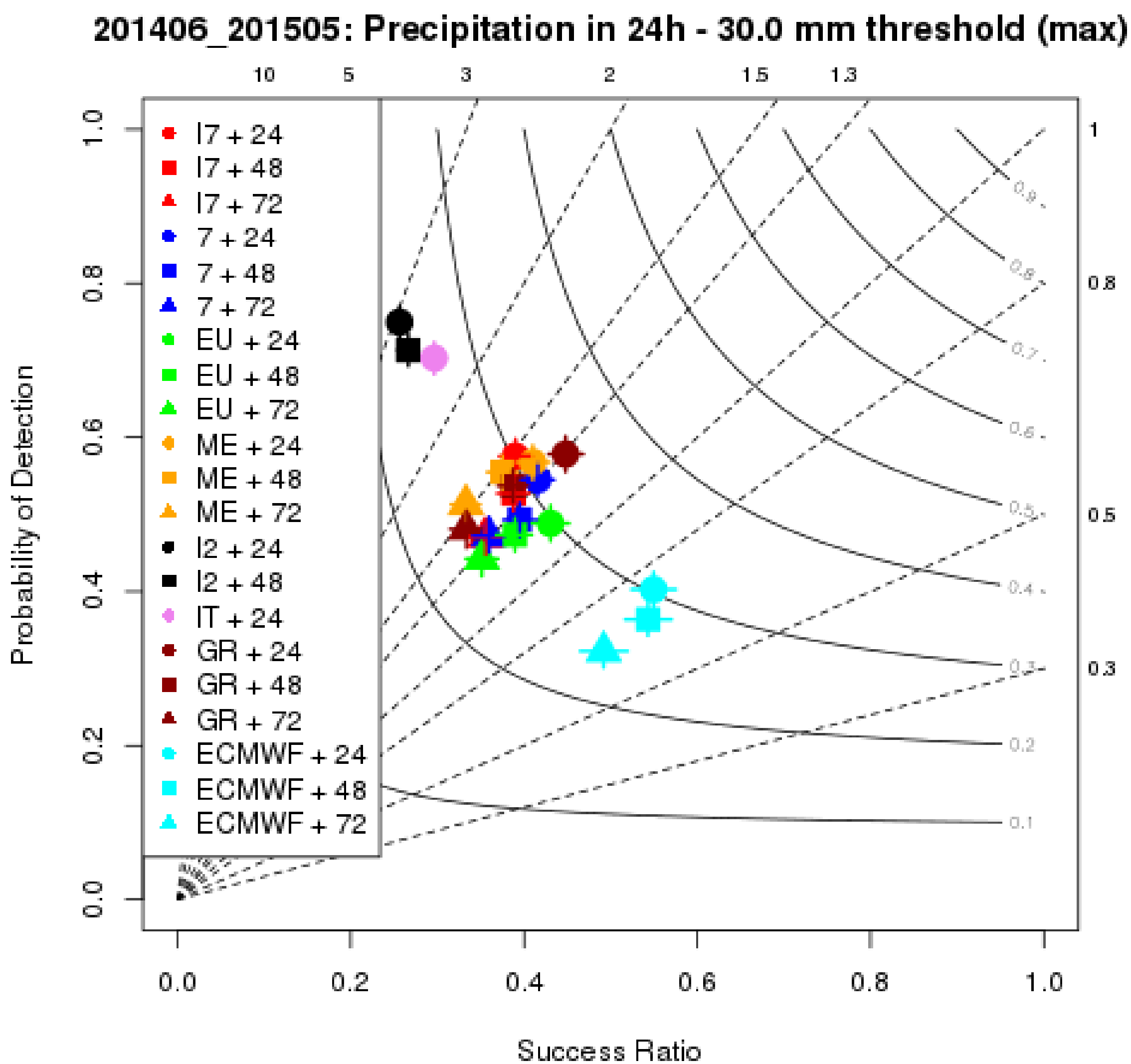
- •spread dots
- •general overestimation

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



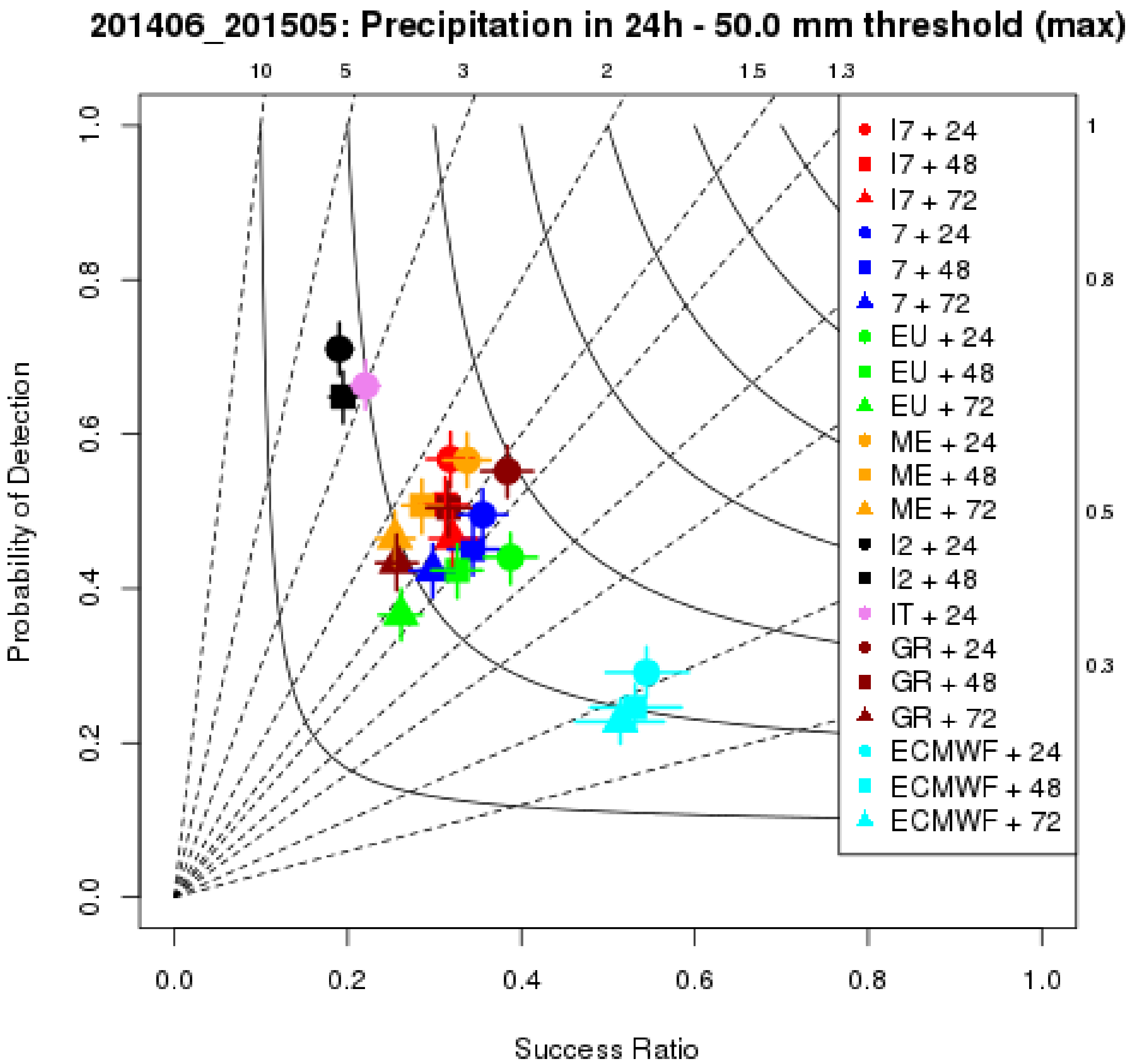
- spread dots
- 3 groups

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

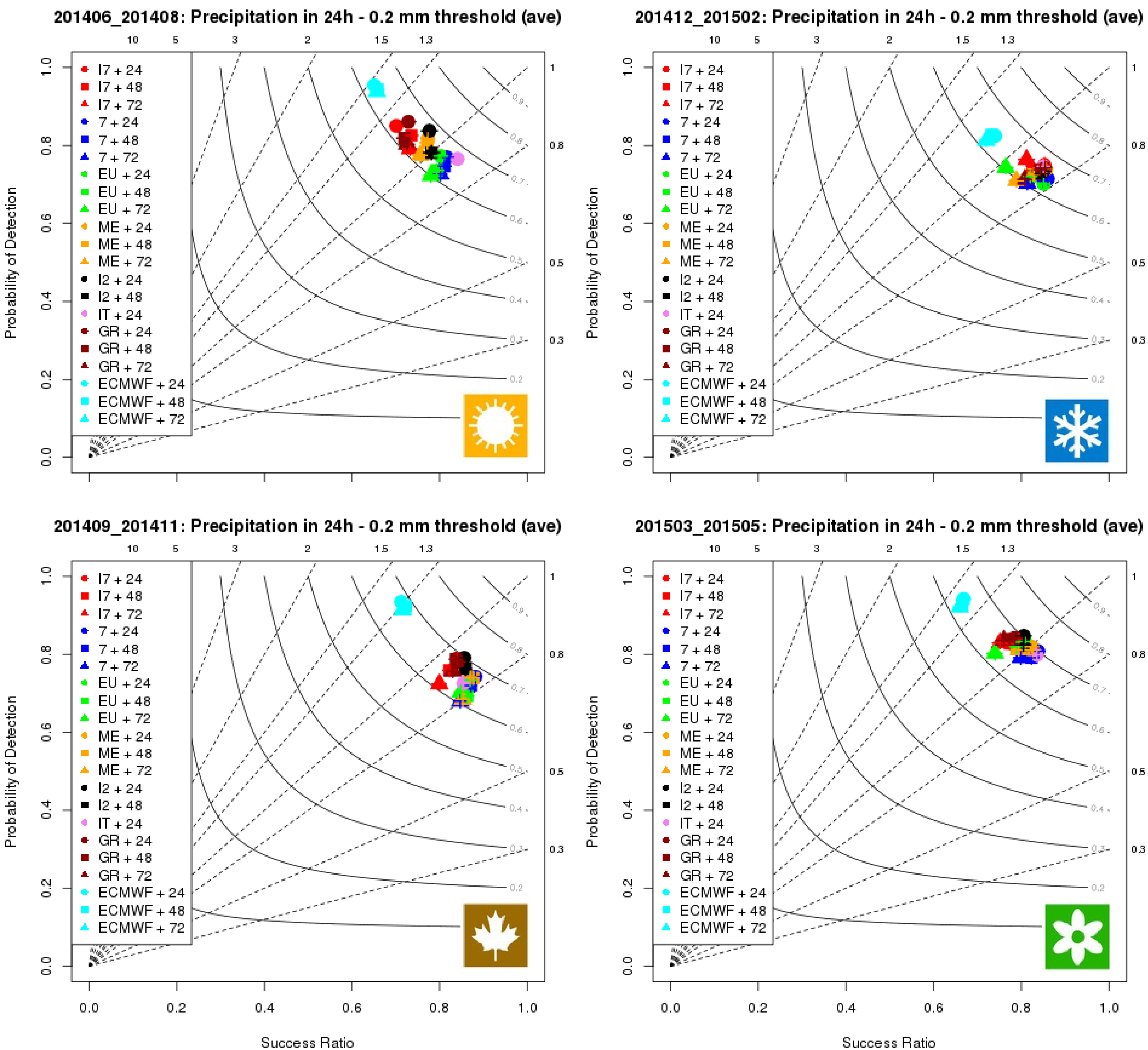


- spread dots
- 3 groups

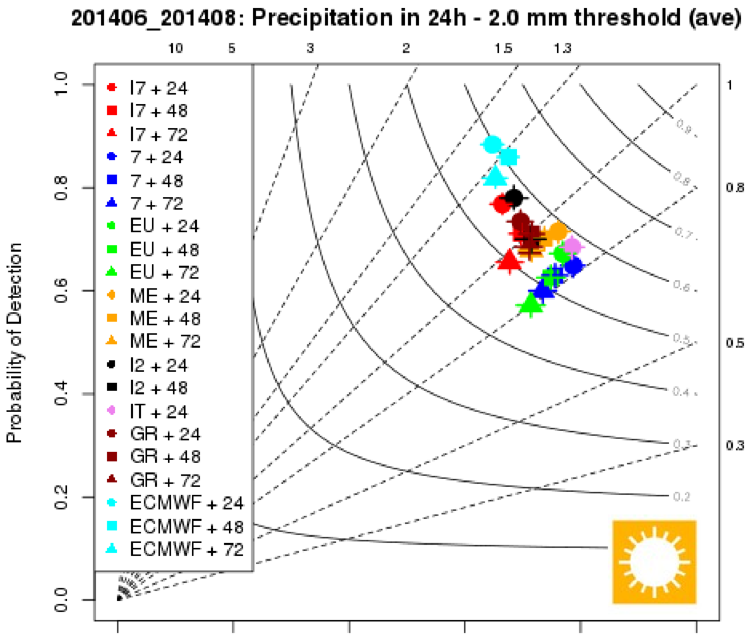
201406-201505: Maximum over area > 50 mm/24h

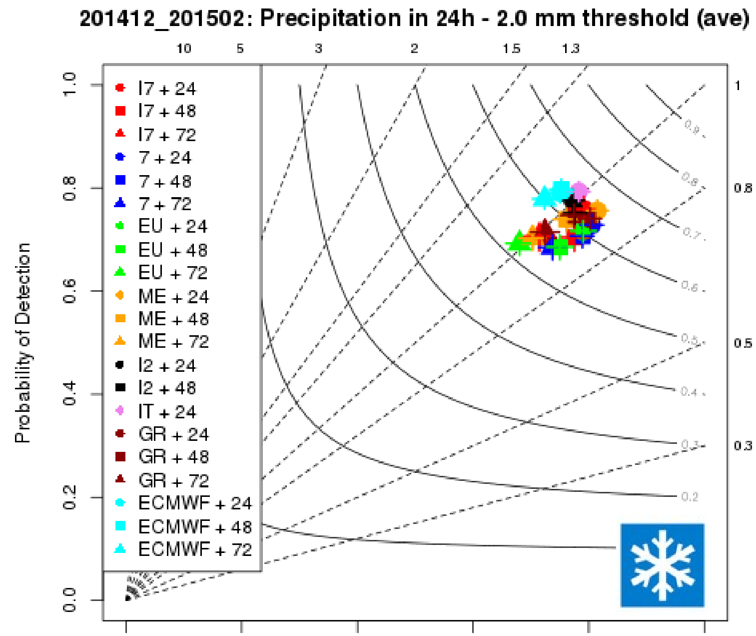


- •spread dots
- •3 groups

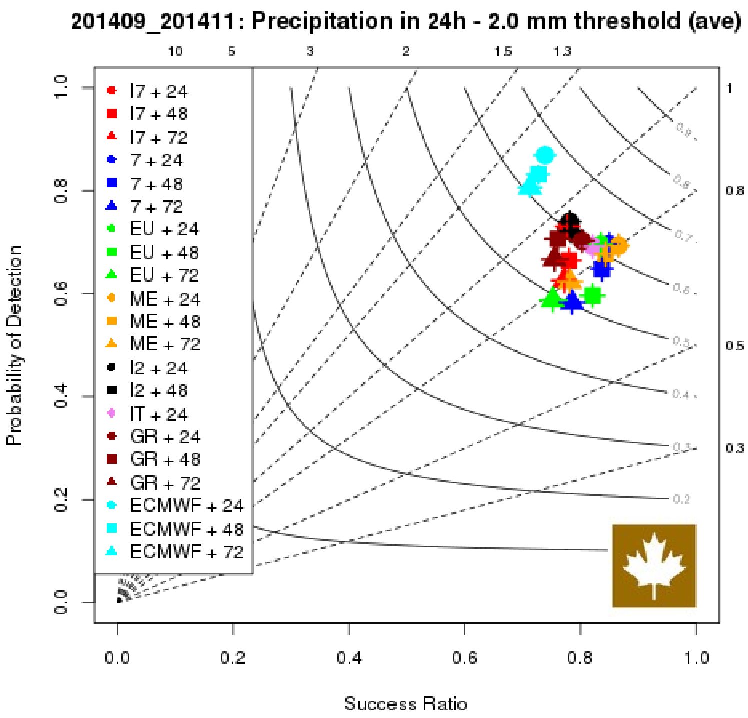


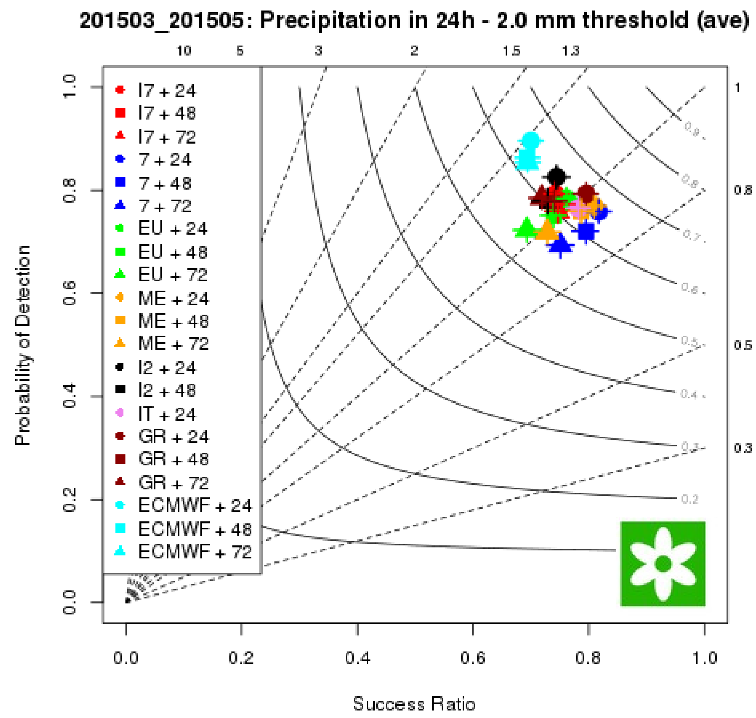
Average over area > 0.2 mm/24h



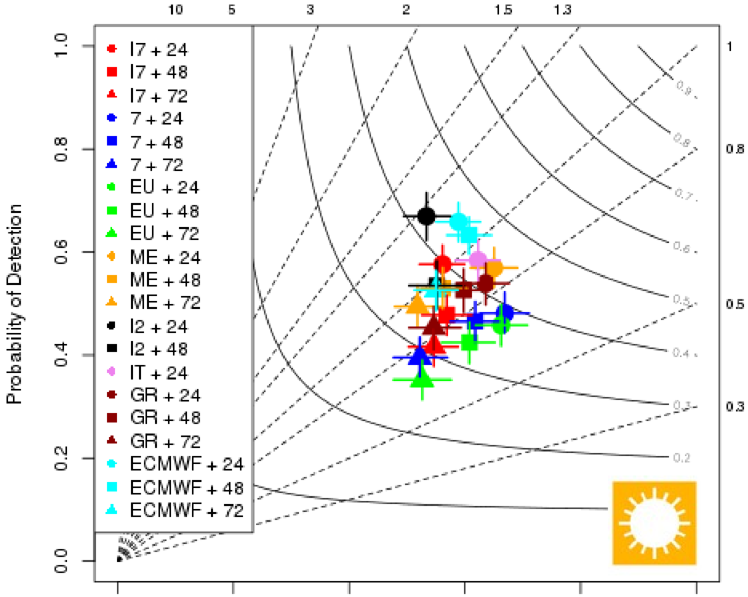


Average over area > 2 mm/24h

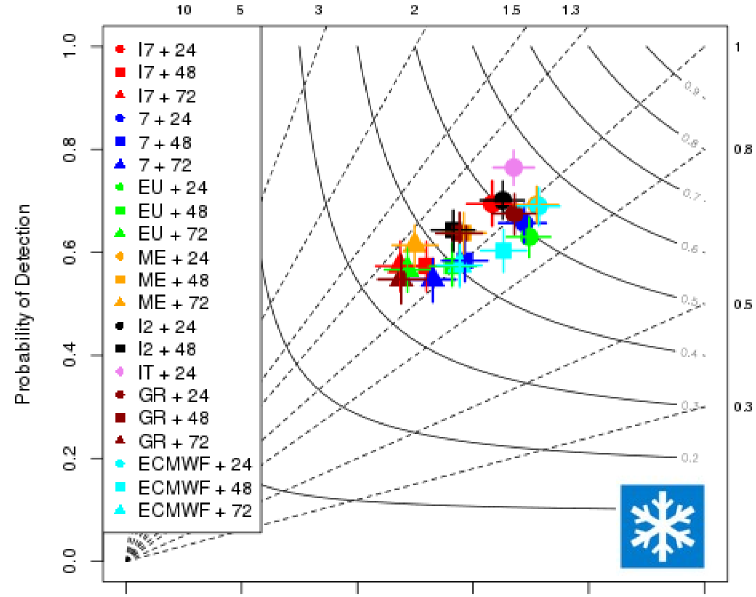




201406_201408: Precipitation in 24h - 10.0 mm threshold (ave)

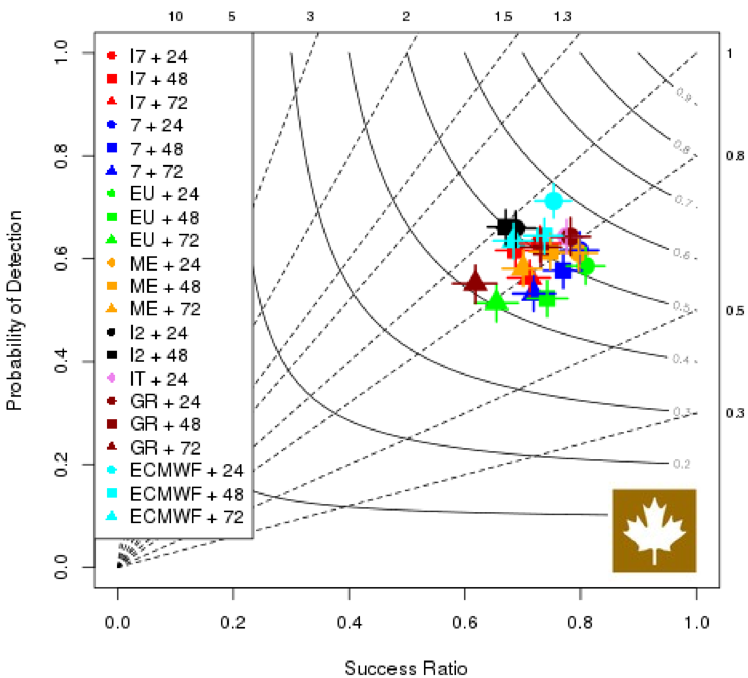


201412_201502: Precipitation in 24h - 10.0 mm threshold (ave)

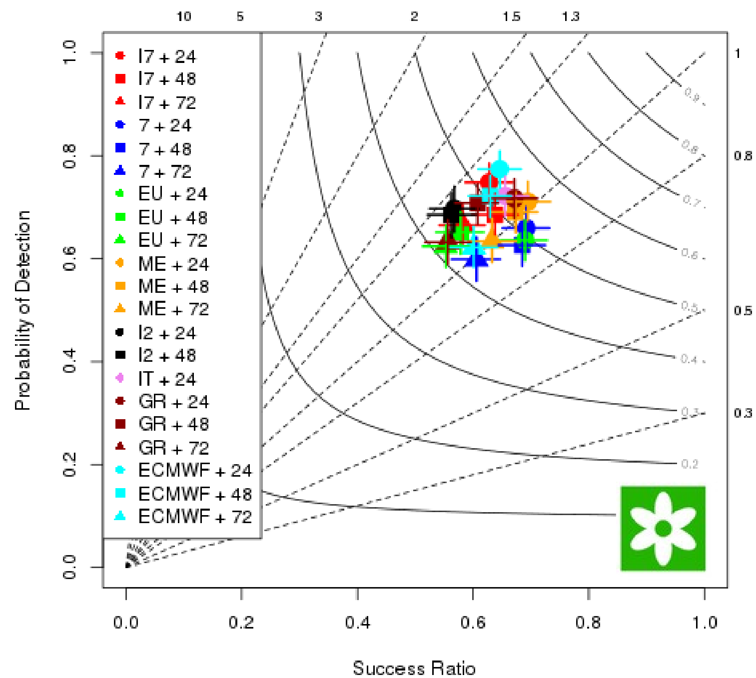


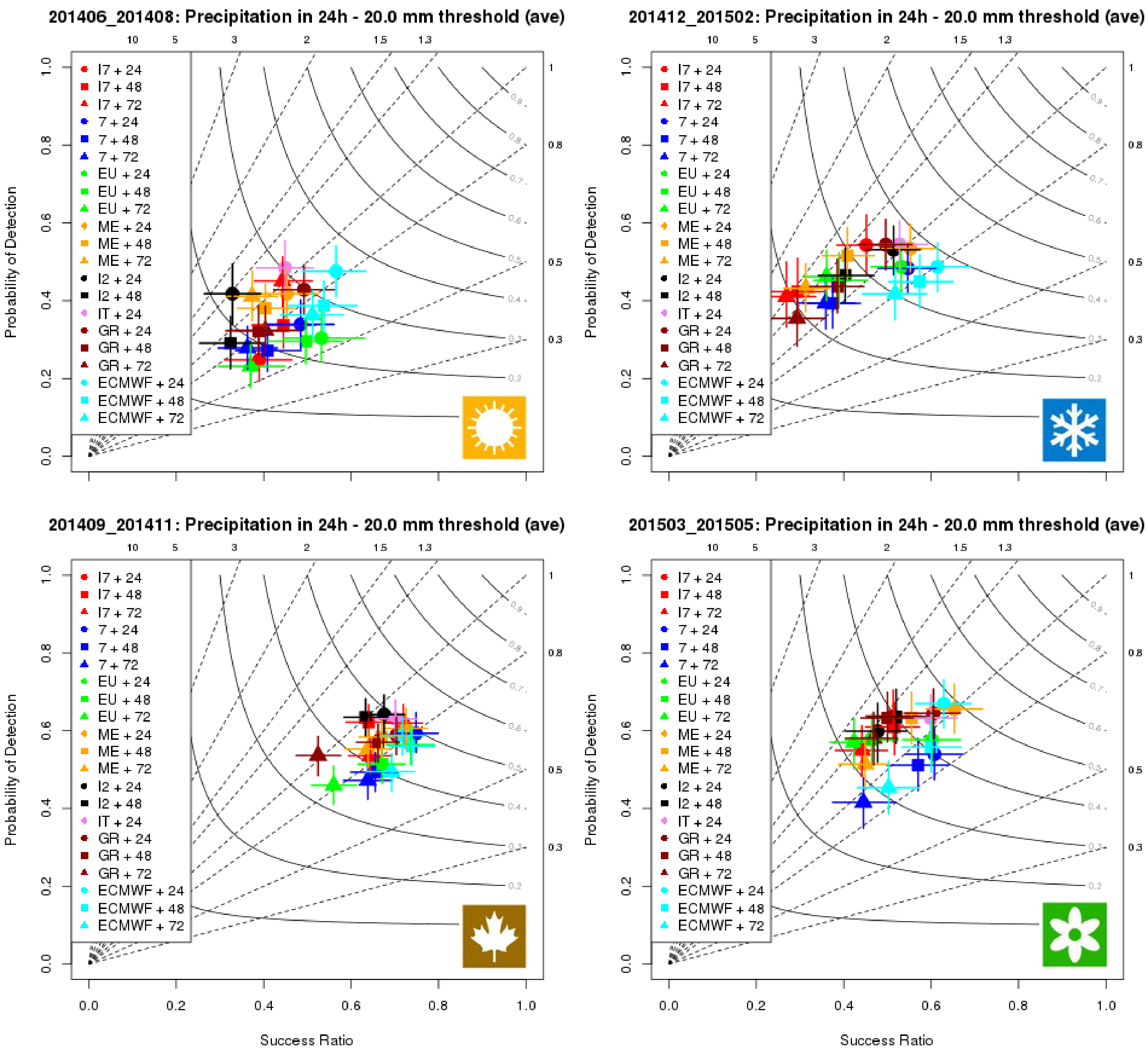
Average over area > 10 mm/24h

201409_201411: Precipitation in 24h - 10.0 mm threshold (ave)



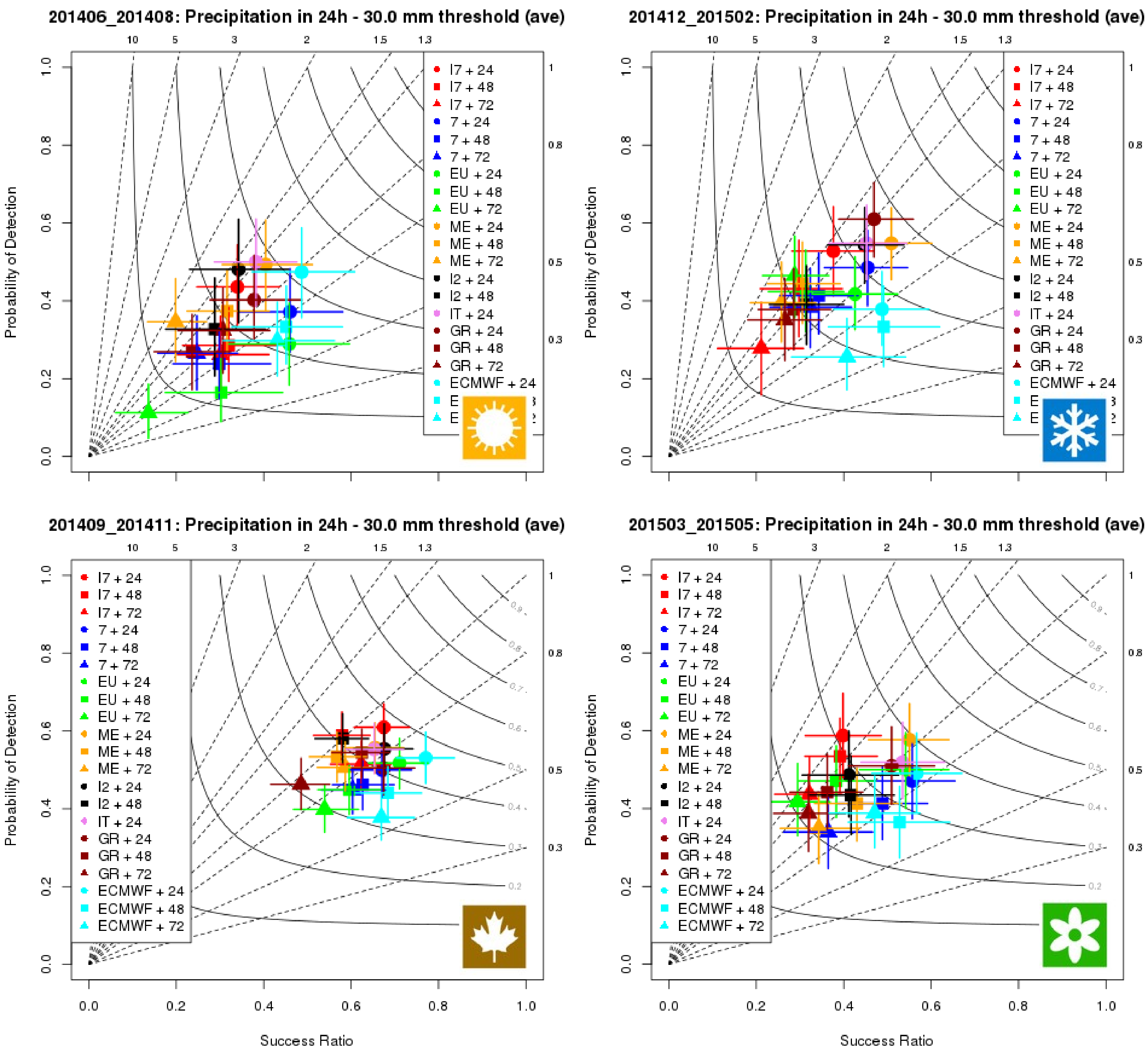
201503_201505: Precipitation in 24h - 10.0 mm threshold (ave)





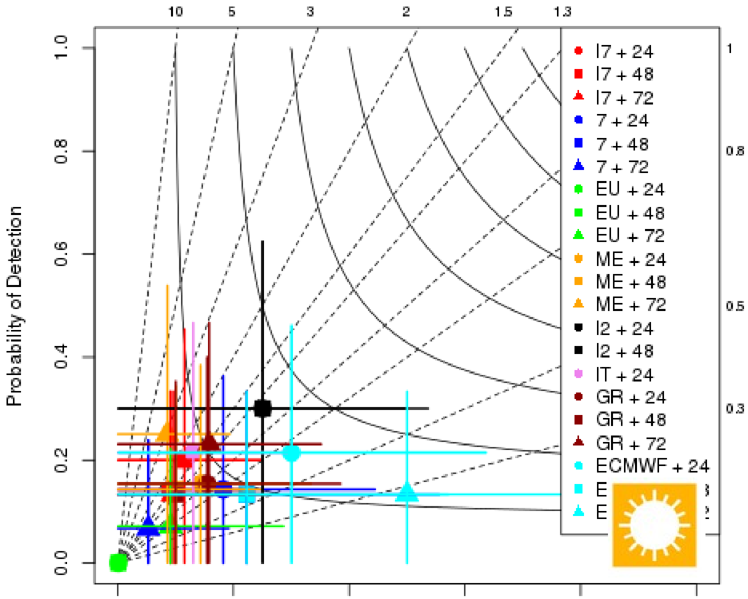
Average over area > 20 mm/24h

Average over area > 30 mm/24h

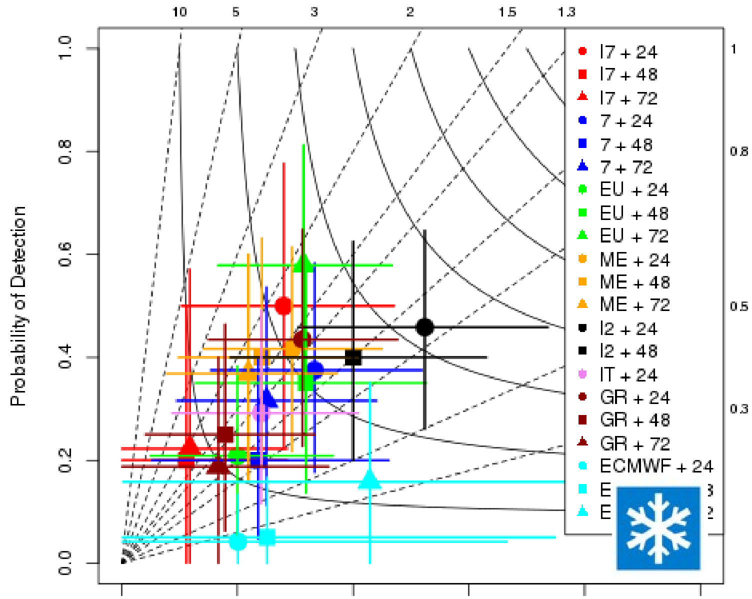


Average over area > 50 mm/24h

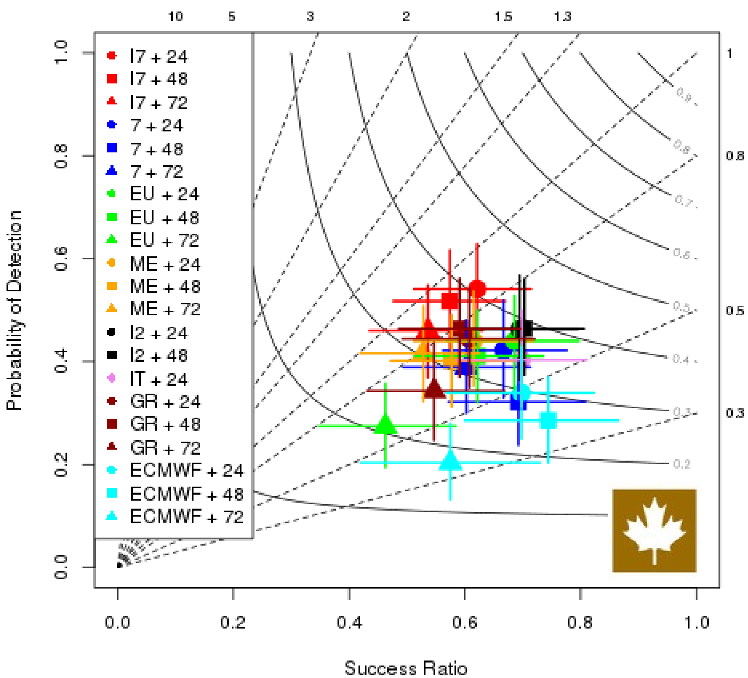
201406_201408: Precipitation in 24h - 50.0 mm threshold (ave)



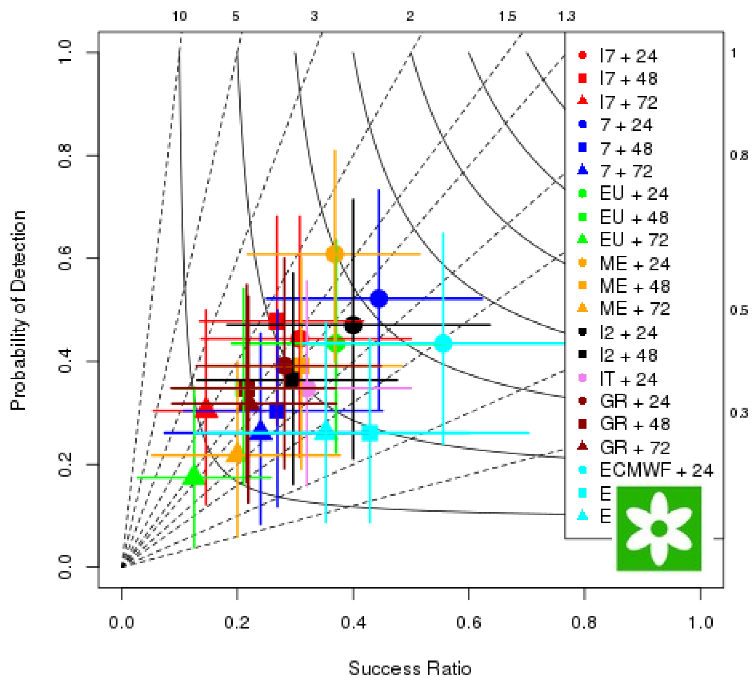
201412_201502: Precipitation in 24h - 50.0 mm threshold (ave)



201409_201411: Precipitation in 24h - 50.0 mm threshold (ave)

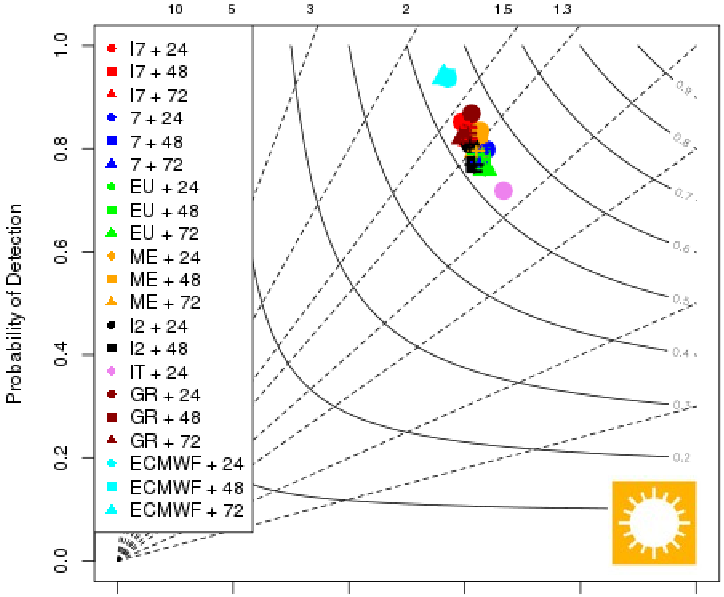


201503_201505: Precipitation in 24h - 50.0 mm threshold (ave)

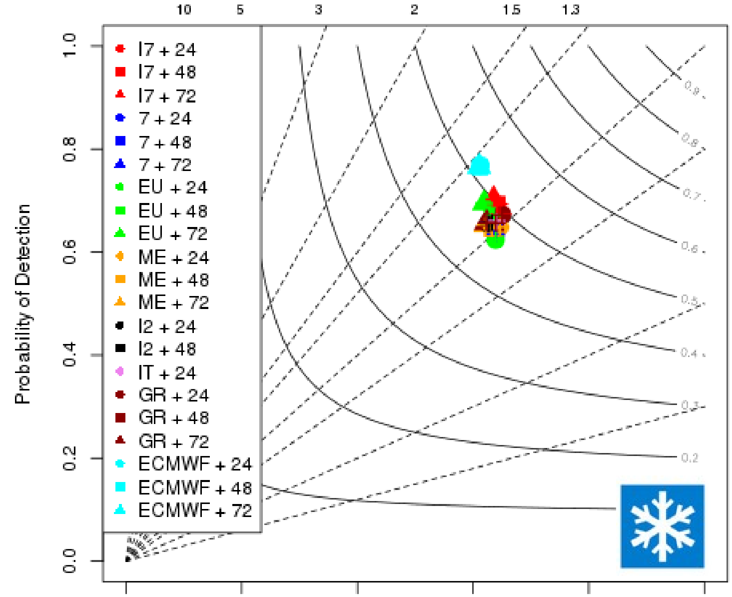


Maximum > 0.2 mm/24h

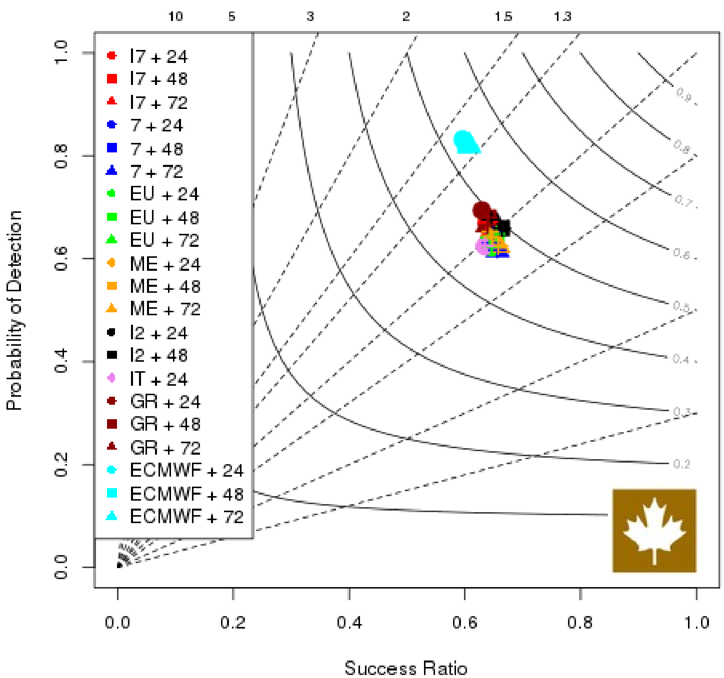
201406_201408: Precipitation in 24h - 0.2 mm threshold (max)



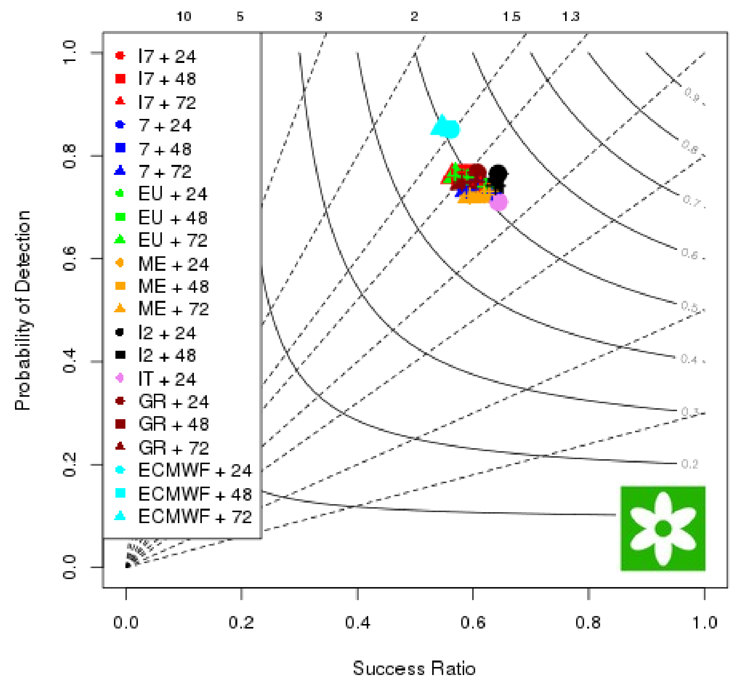
201412_201502: Precipitation in 24h - 0.2 mm threshold (max)

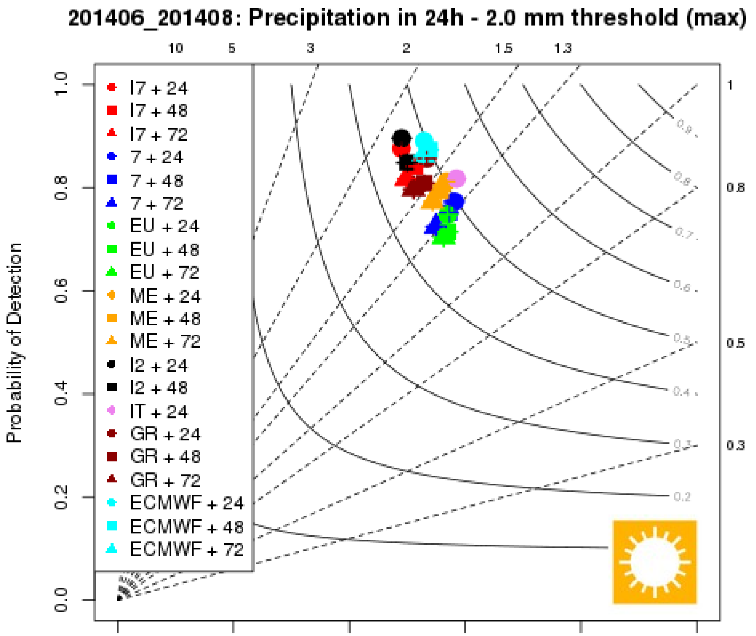


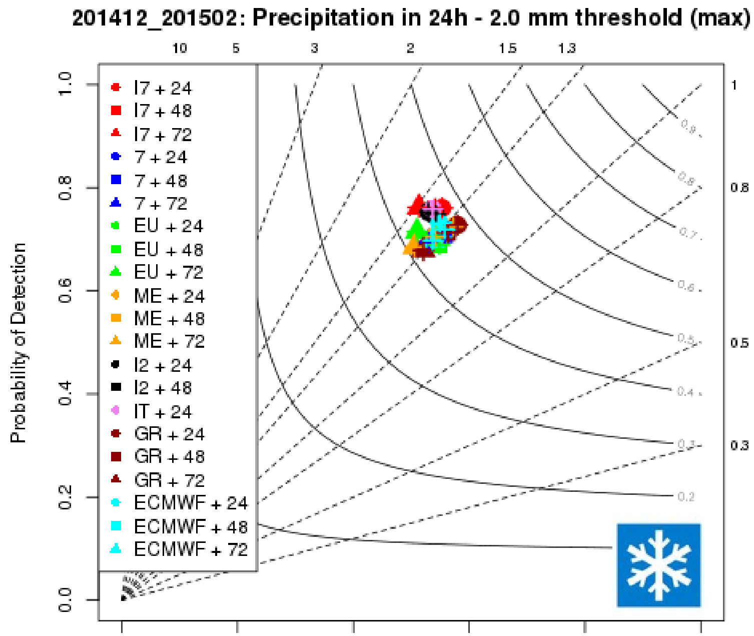
201409_201411: Precipitation in 24h - 0.2 mm threshold (max)



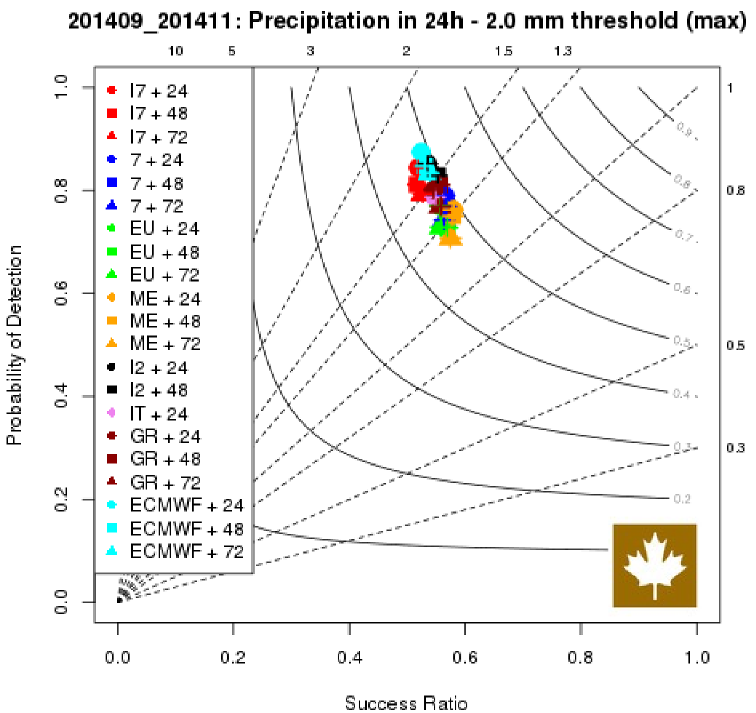
201503_201505: Precipitation in 24h - 0.2 mm threshold (max)

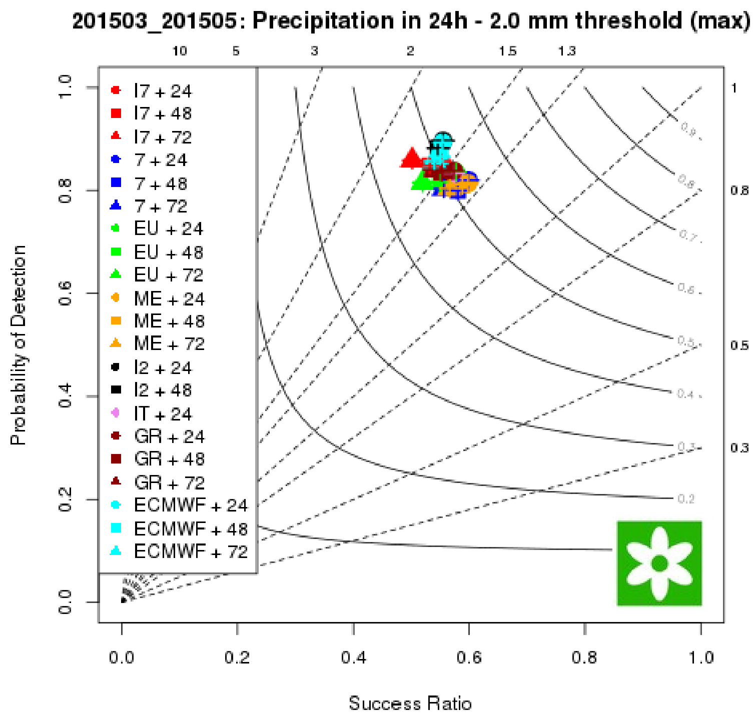






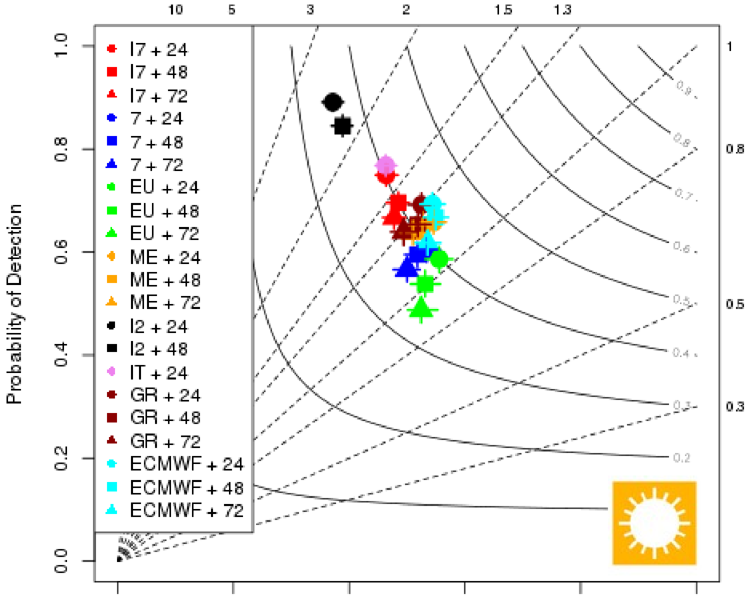
Maximum > 2 mm/24h



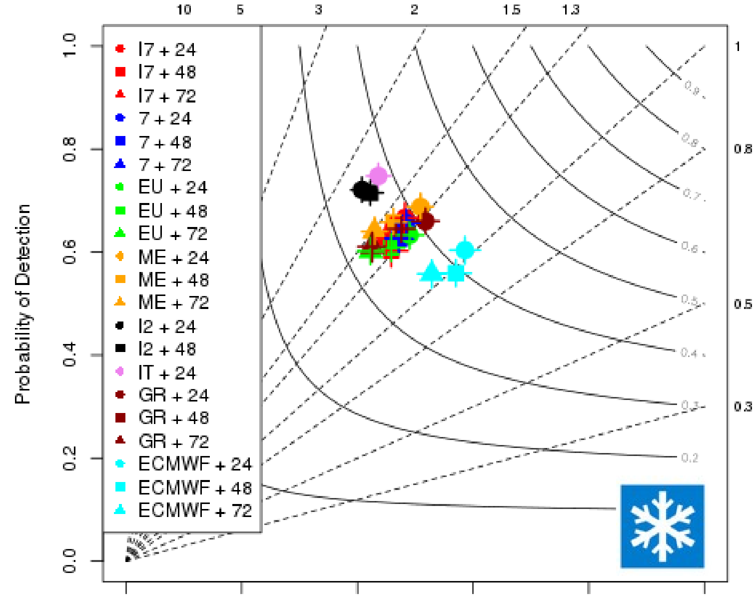


Maximum > 10 mm/24h

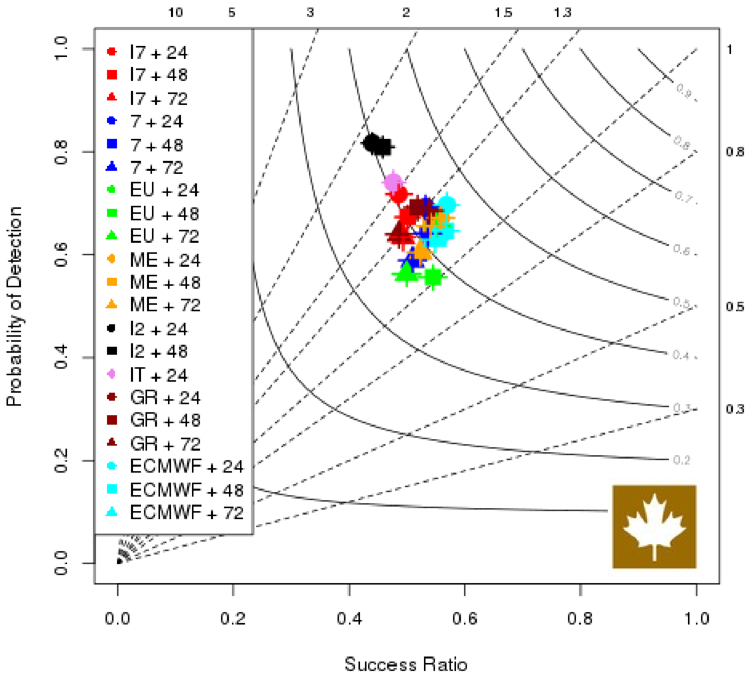
201406_201408: Precipitation in 24h - 10.0 mm threshold (max)



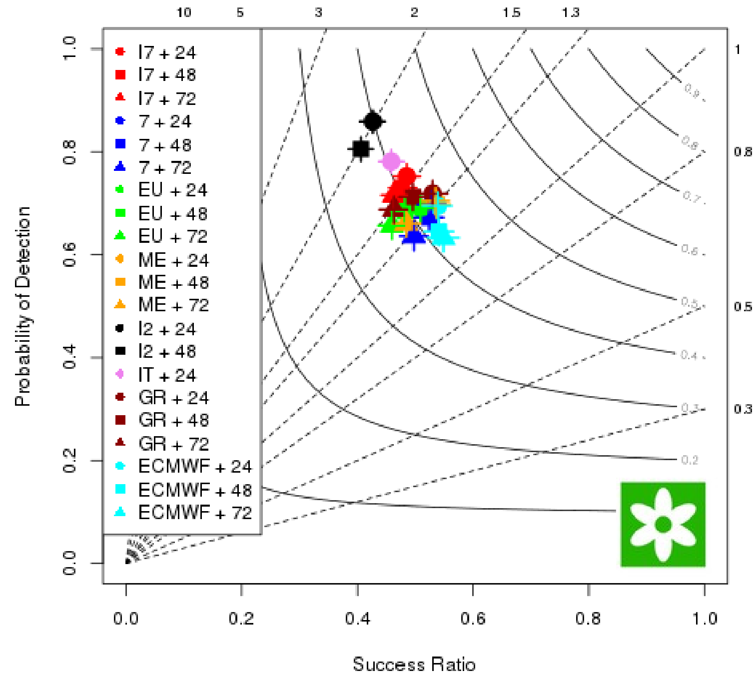
201412_201502: Precipitation in 24h - 10.0 mm threshold (max)

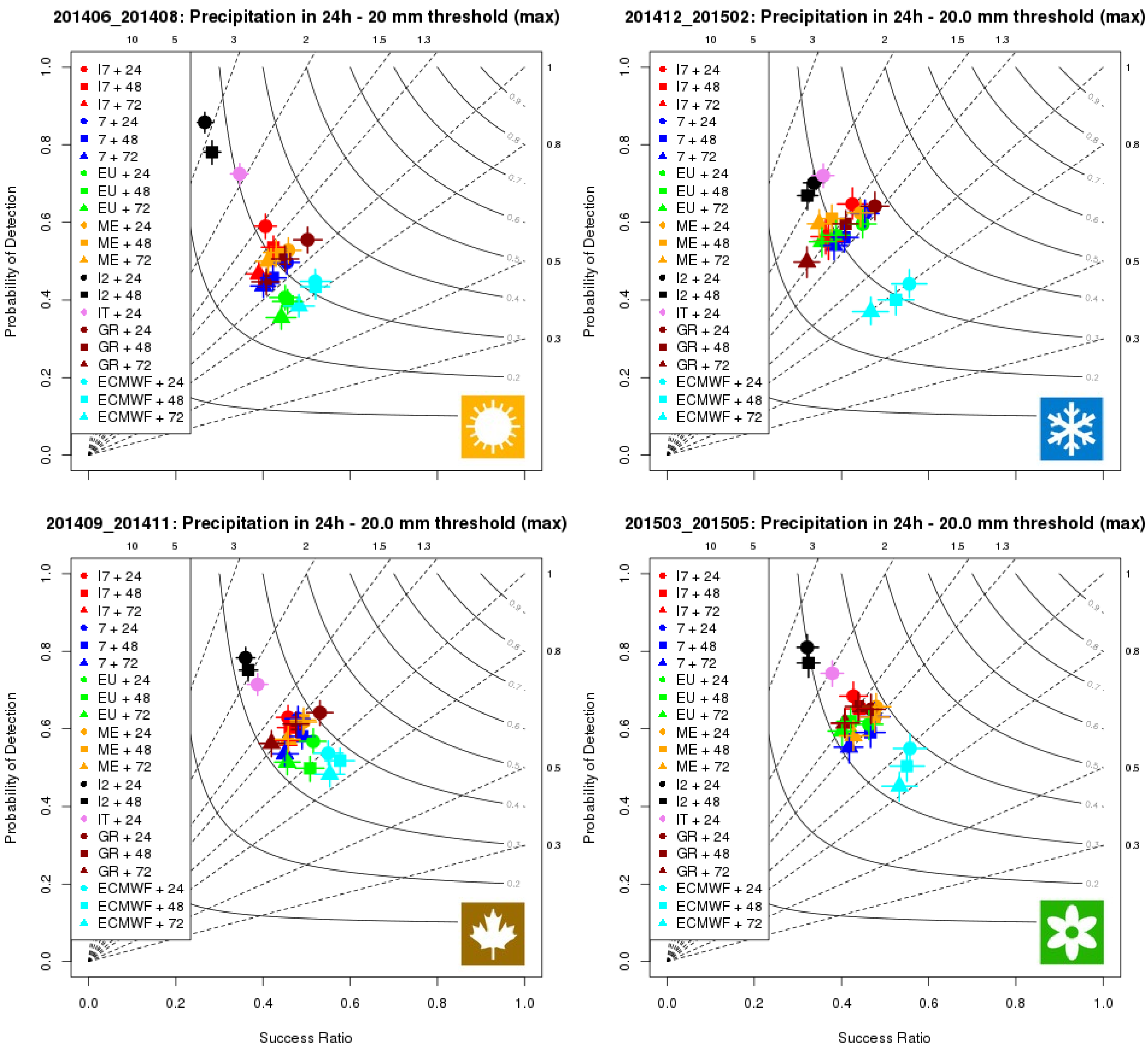


201409_201411: Precipitation in 24h - 10.0 mm threshold (max)

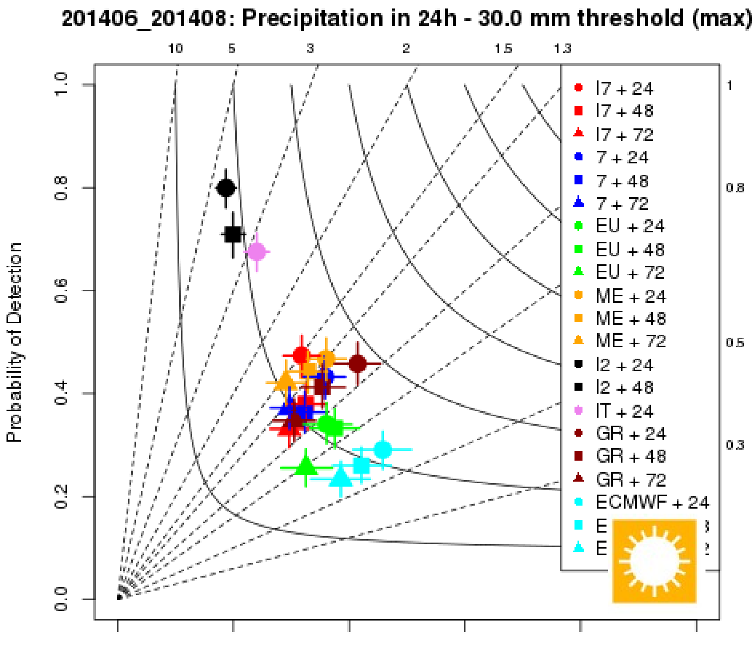


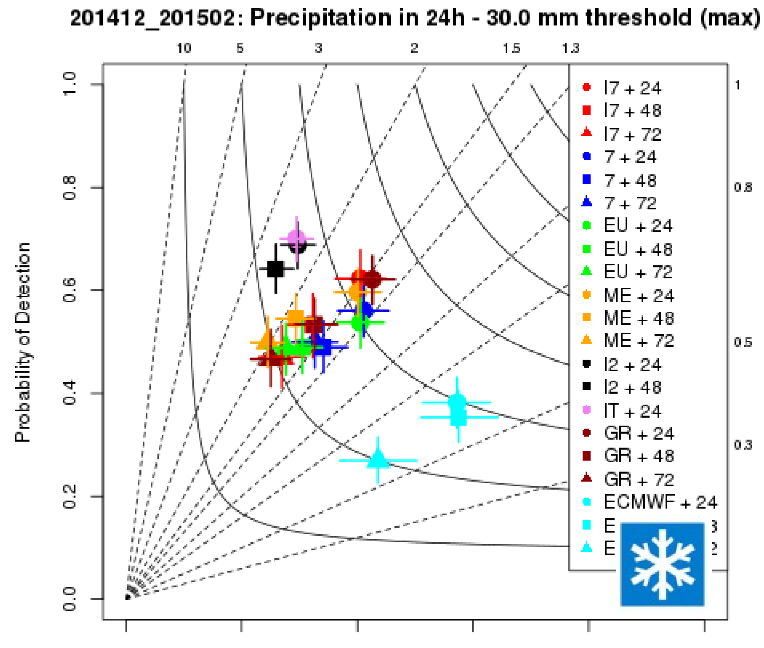
201503_201505: Precipitation in 24h - 10.0 mm threshold (max)



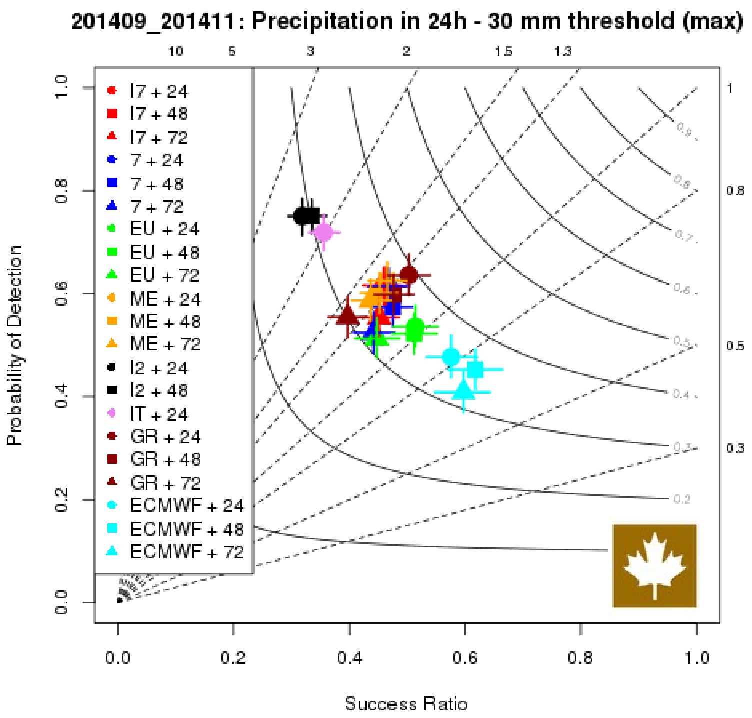


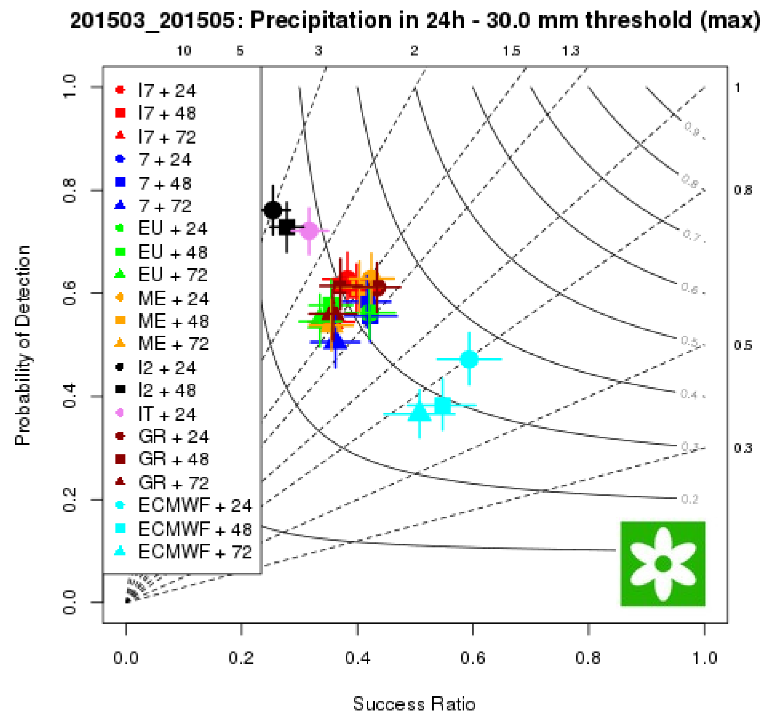
Maximum > 20 mm/24h





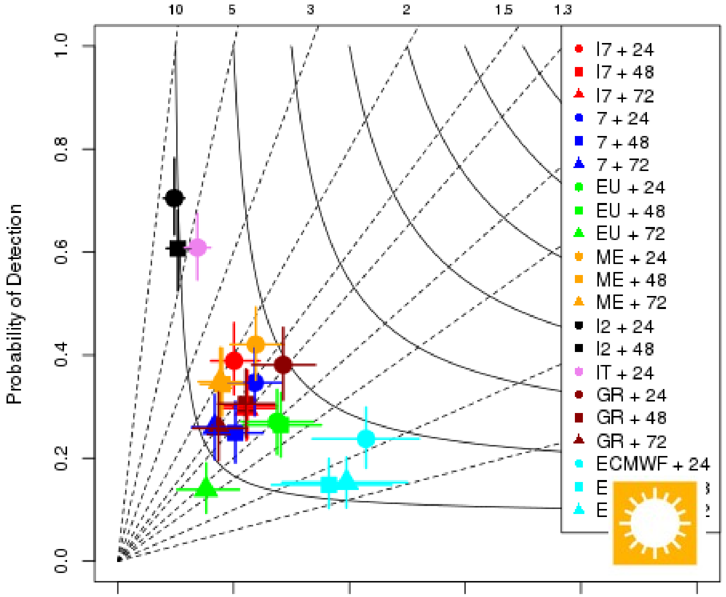
Maximum > 30 mm/24h



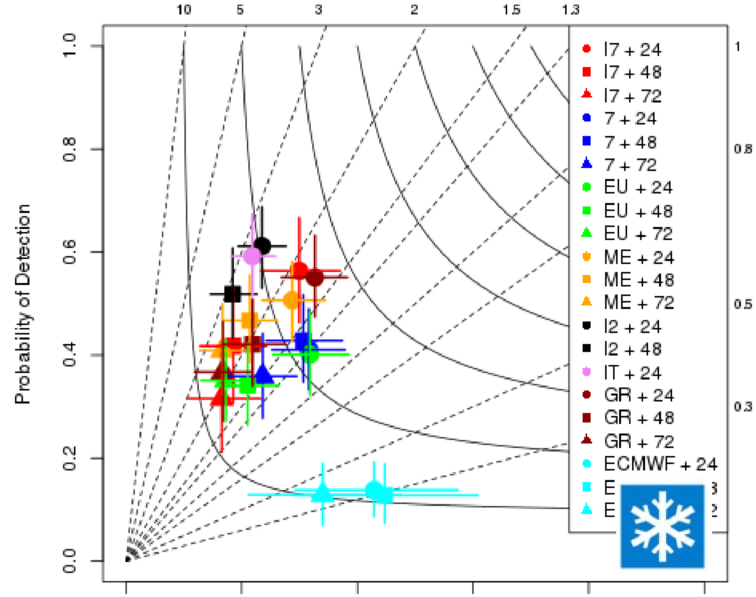


Maximum > 50 mm/24h

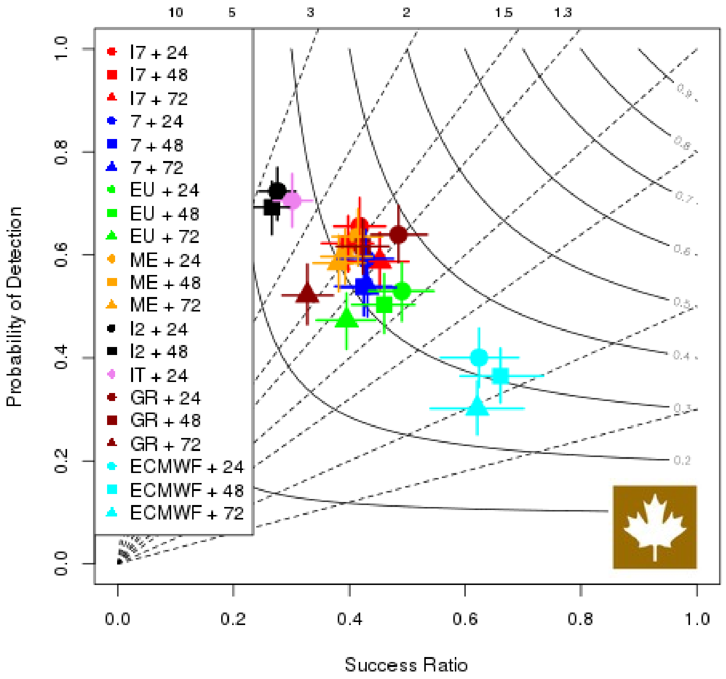
201406_201408: Precipitation in 24h - 50.0 mm threshold (max)



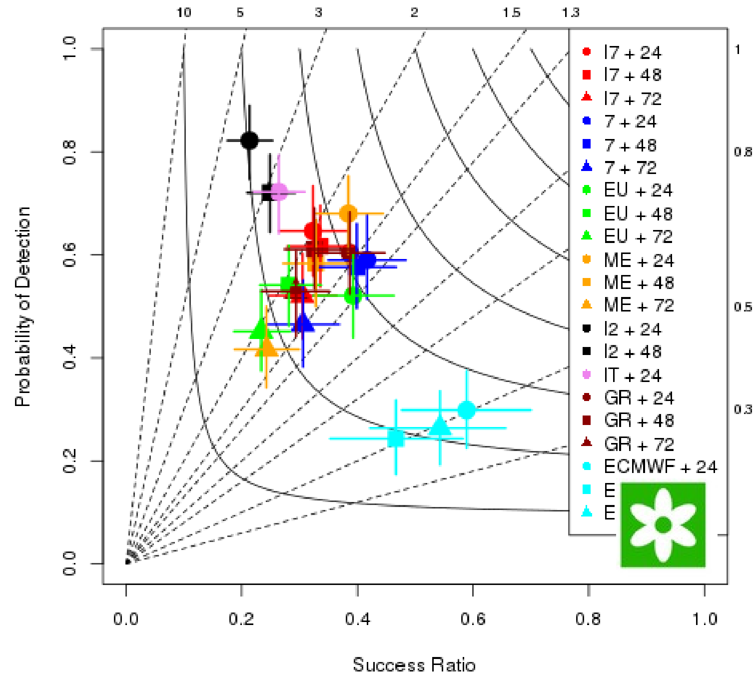
201412_201502: Precipitation in 24h - 50.0 mm threshold (max)



201409_201411: Precipitation in 24h - 50.0 mm threshold (max)

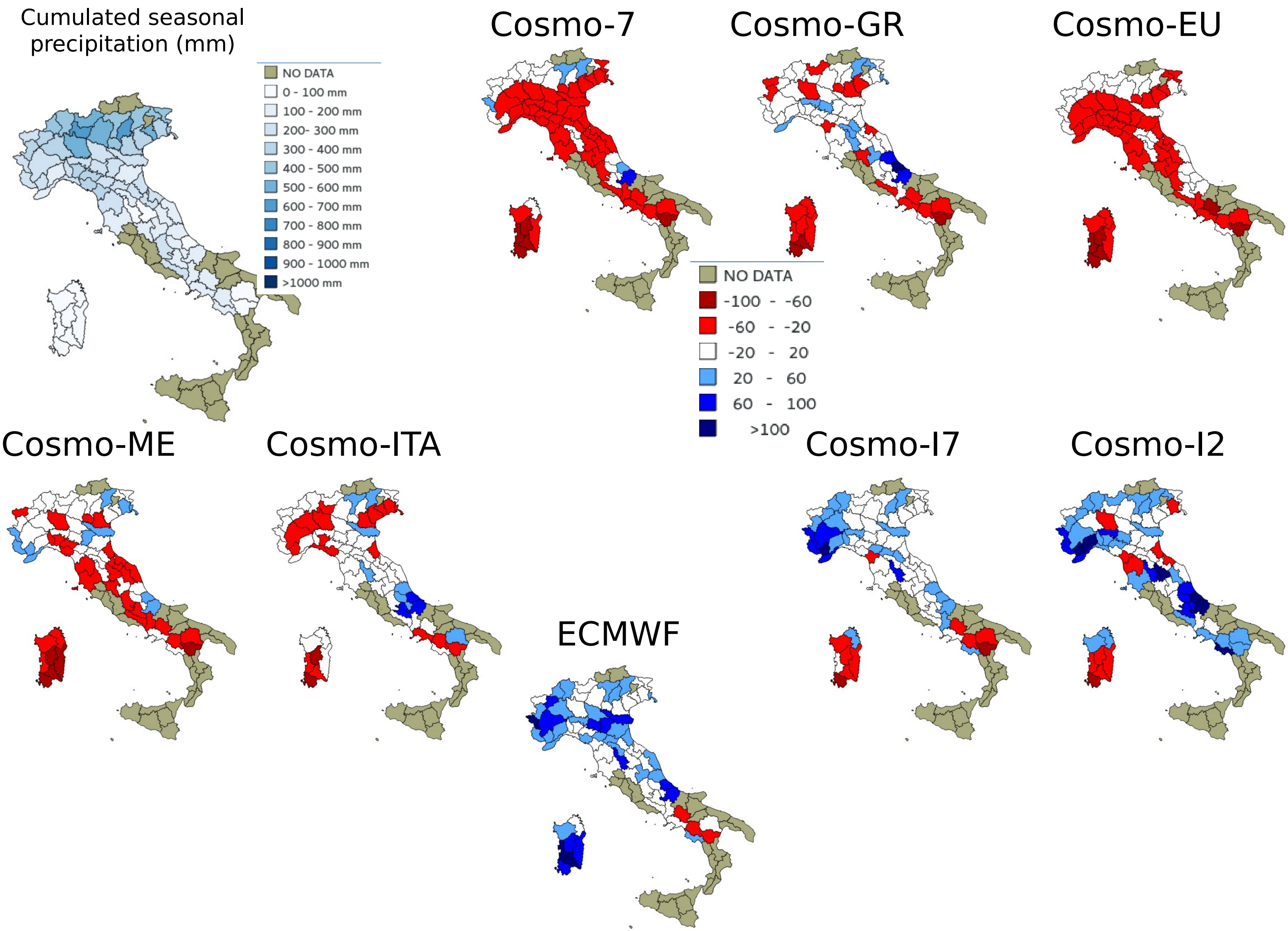


201503_201505: Precipitation in 24h - 50.0 mm threshold (max)



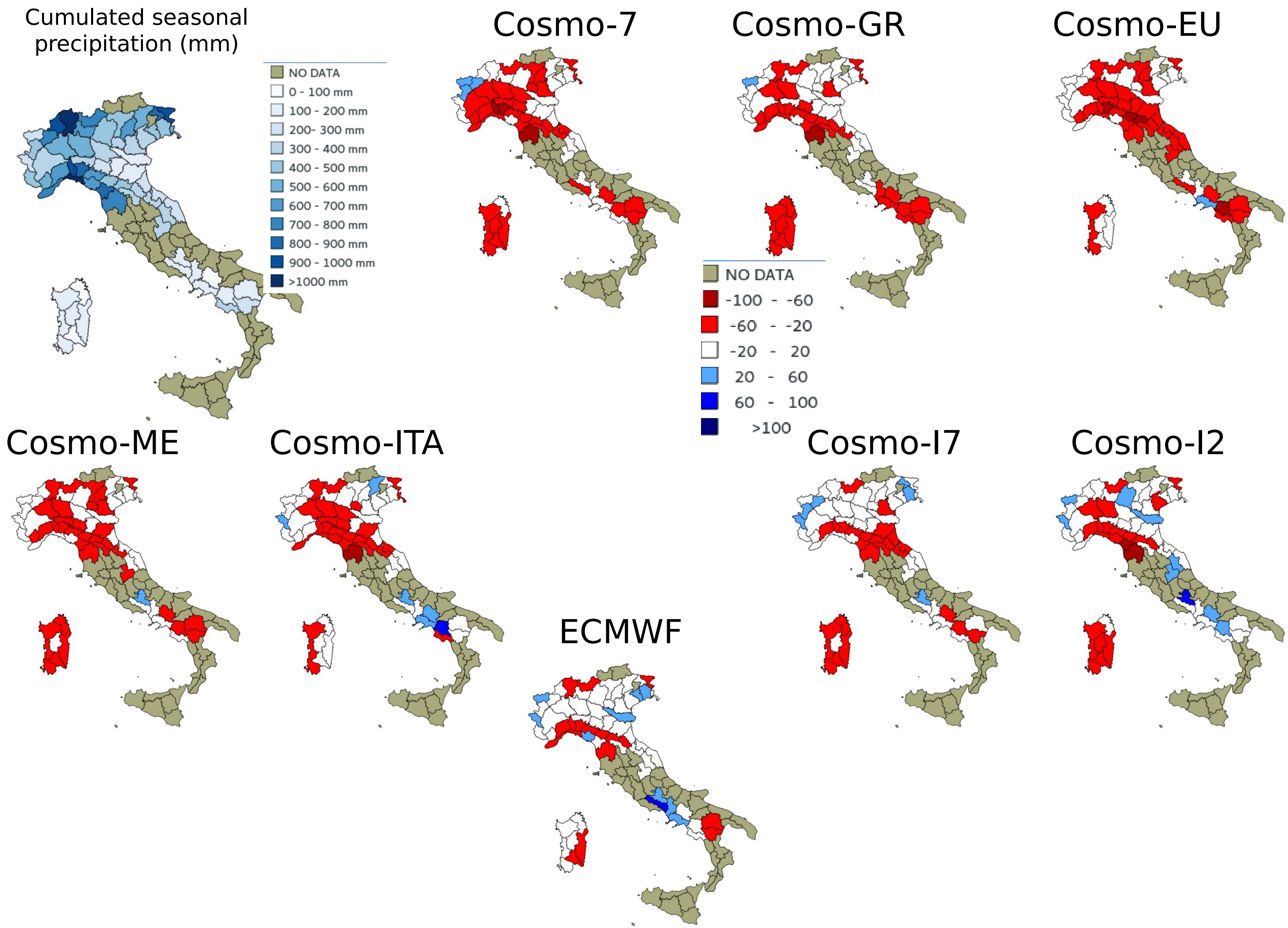
RELATIVE ERROR jja 2014

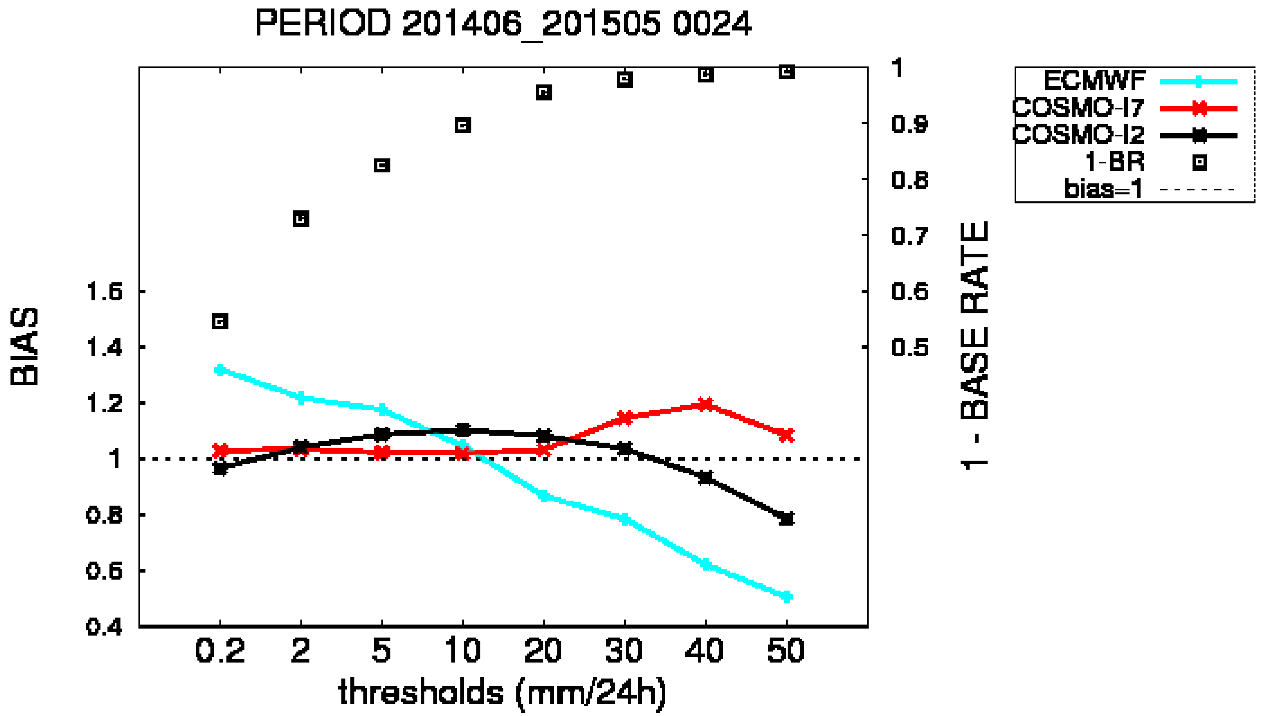
Rel Err= (for-obs)/obs %

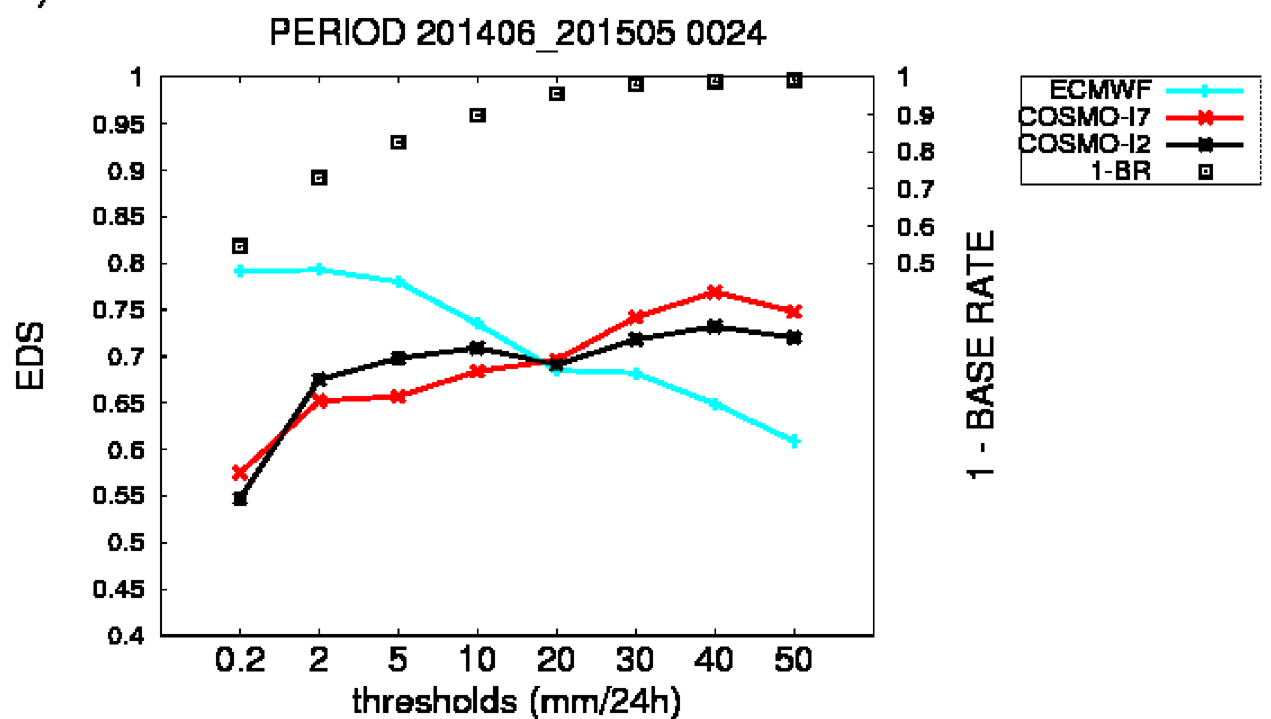


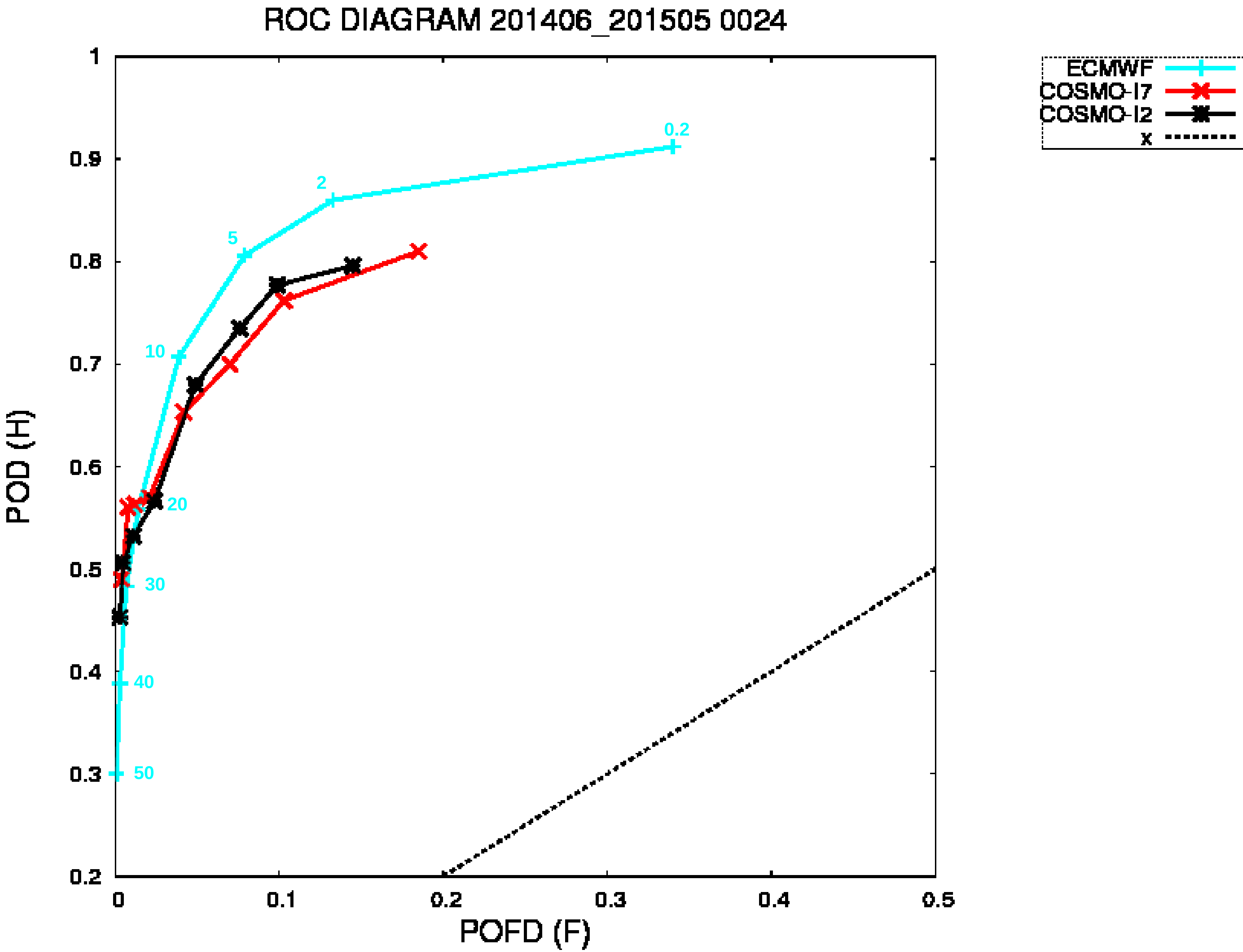
RELATIVE ERROR son 2014

Rel Err= (for-obs)/obs %




PERIOD 201406_201505 0024
BIAS
1 - BASE RATE
thresholds (mm/24h)
0.2
2
5
10
20
30
40
50
ECMWF
COSMO-I7
COSMO-I2
1-BR
bias=1


PERIOD 201406_201505 0024
EDS
1 - BASE RATE
thresholds (mm/24h)
0.2
2
5
10
20
30
40
50
ECMWF
COSMO-I7
COSMO-I2
1-BR



ROC DIAGRAM 201406_201505 0024
ECMWF
COSMO-I7
COSMO-I2
x
POD (H)
POFD (F)
0.2
2
5
10
20
30
40
50
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0
0.1
0.2
0.3
0.4
0.5

extreme dependency score → investigate the performance of an NWP model for rare events

Stephenson et al. Introduce the extreme dependency score (EDS) as a good alternative to standard scores for verification of rare events.

	Event observed yes	Event observed no	Total
Forecast yes	A	b	*a* + *b*
Forecast no	c	d	*c* + *d*
Total	*a* +*c*	*b*+*d*	*n*= *a* + *b* + *c* + *d*

frequency bias index	*FBI = (a + b)/(a + c)*	[0,∞] best 1	The frequency bias index indicates whether the forecasting system under or over-forecasts the number of events.
hit rate (POD)	*H = a/(a + c)*	[0,1] best 1	The hit rate represents the probability that the event is forecast when it occurs
false alarm rate (POFD)	*F = b/(b + d)*	[0,1] best 0	The false alarm rate represents the probability of forecasting the event when it did not occur. % not events obs. Not correctly forecasted. fraction of the observed "no" events were incorrectly forecast as "yes“.
true skill score	*T SS = H −F*	[-1,1] best 1	The true skill score gives information on how the forecasting system distinguishes between occurrences and not occurrences.
base rate	*BR = (a + c)/n*	[0,1]	The base rate represents the probability that the event occurs. By definition, 1-BR plotted versus increasing thresholds represents the probability that precipitation amount does not exceed a certain threshold.
extreme dependency score	*EDS= 2[ln((a+c)/n)/ln(a/n)]-1*	[-1,1] best 1	What is the association between forecast and observed rare events? Converges to 2η-1 as event frequency approaches 0, where η is a parameter describing how fast the hit rate converges to zero for rarer events. EDS is independent of bias, so should be presented together with the frequency bias

- To get clear information about how the forecasting system detects the extreme events, it would be fair if the EDS is compared for events having the same base rate. One has to investigate if better value of the EDS are related to an improvement in the quality of the forecasting system or if they are due to the event variability over the years.
- The equation defining the EDS uses the left hand side of a contingency table and the total number of cases (sample size). This results in an increased freedom for false alarms and correct negatives, which can freely vary with the only restriction that their sum has to be constant. Therefore, it is paramount to use the EDS in combination with other scores that include the right hand side of the contingency table, as the F and/or the FBI to show that improvements are not due to an increase of false alarms. (Ghelli&Primo,2009)

The affect of the base rate on the extreme dependency score (Ghelli&Primo,2009)

The Extreme Dependency Score (EDS) has been introduced as an alternative measure to verify the performance of numerical weather prediction models for rare events, taking advantage of the non-vanishing property of the score when the event probability tends to zero.
This score varies from 1 (best value) to −1 (worst value).

The EDS is written as a function of BR:

$EDS = [\ln(BR) - \ln(HR)]/[\ln(BR) + \ln(HR)]$

Equation presents the EDS as a function of the base rate and the hit rate.
when HR = 1, the EDS = 1 and when BR = 1, the EDS = −1.
On the other hand, when the base rate is equal to one, the event happens all the time and so the EDS is not an appropriate score since it is focused on verification of extreme events (low probability of occurrence). Therefore, if different data samples need to be compared, it is imperative to have similar base rate.

- Thus, even if there are no misses and the EDS value is maximum, the forecasting system might have a high number of false alarms. Therefore, an EDS = 1 does not imply a skilful system. If values of the EDS for different periods need to be compared, then the base rate must be constant in time to avoid changes in the EDS to be just a reflection of changes in the BR.
- If the base rate is constant, an increase of the EDS implies a better probability of detection (hit rate), i.e. a more skilful system. If only the hit rate is constant, then an increase of the EDS is only due to a higher event probability. If neither the base rate nor the hit rate is constant, then the improvement of the EDS could be due to any of the previous reasons.

The extreme dependency score: a non-vanishing measure for forecasts of rare events (Stephenson et al.)

EDS takes the value of 1 for perfect forecasts and 0 for random forecasts, and is greater than zero for forecasts that have hit rates that converge slower than those of random forecasts

EDS has demonstrated here that there is dependency between the forecasts and the observations for more rare events, which is masked by the traditional skill scores that converge to zero as the base rate vanishes. EDS does not explicitly depend on the bias in the system for vanishing base rate and so is less prone to improvement by hedging the forecasts. EDS has the disadvantage that it is based only on the numbers of hits and misses, and so ignores information about false alarms and correct rejections. Therefore, EDS is non-informative about forecast bias, and a forecasting system with a good EDS could be very biased. Therefore, one should present EDS together with the frequency bias as a function of threshold in order to provide a complete summary of forecast performance.