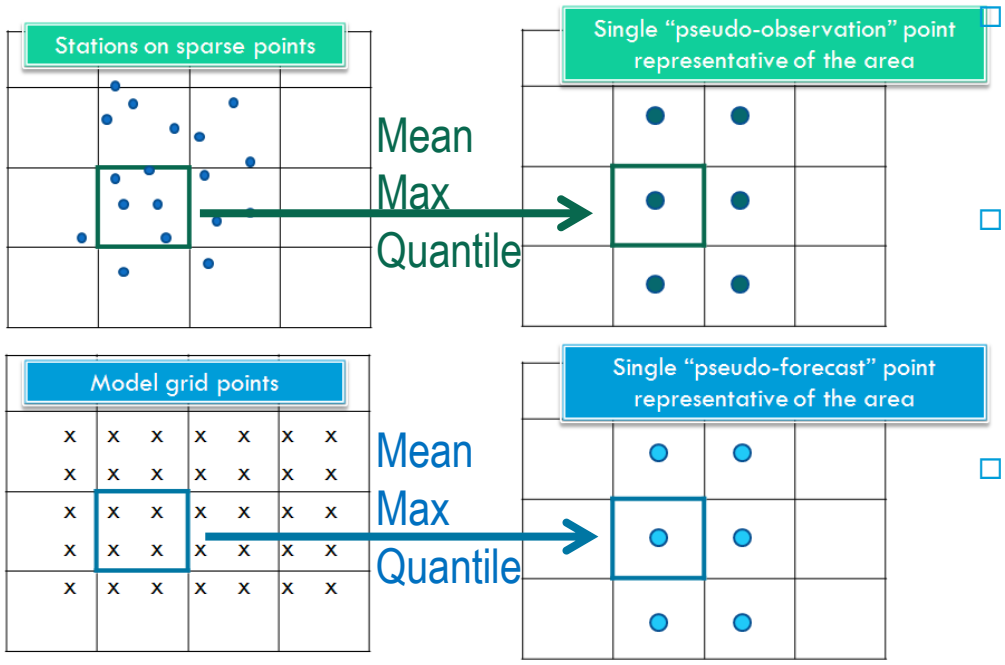
OPERATIONAL USE OF DIST FOR PRECIPITATION VERIFICATION

Maria Stefania Tesini
Chiara Marsigli (ARPAE Emilia-Romagna)
Elena Oberto (ARPA-Piemonte)

General Meeting
Jerusalem, 11-14 September 2017


Once
upon
a
time...

The "distributional method (DIST)"



- The verification domain is subdivided into a number of boxes, each of them containing a certain number of observed and forecast values.
- For each box, several parameters of the distribution of both the observed and forecast values falling in the box can be computed (mean, median, percentiles, maximum).
- Verification is then performed using a categorical approach, by comparing for each box one or more parameters of the forecast distribution against the corresponding parameters of the observed distribution, using a set of indices.

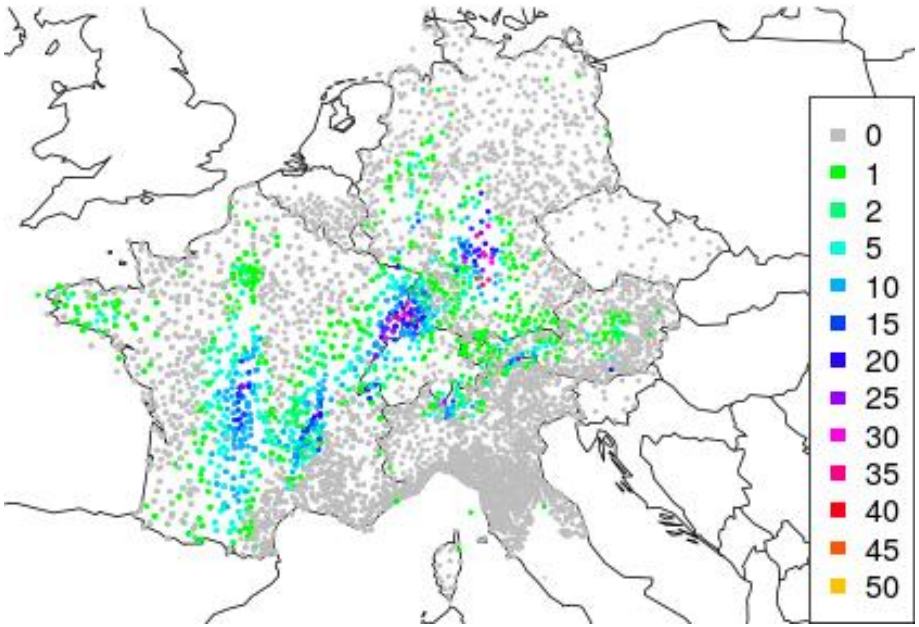
Marsigli, C., Montani, A. and Paccangnella, T. (2008), A spatial verification method applied to the evaluation of high-resolution ensemble forecasts. Met. Apps, 15: 125–143. doi: 10.1002/met.65

Why DIST is an handy Spatial method?

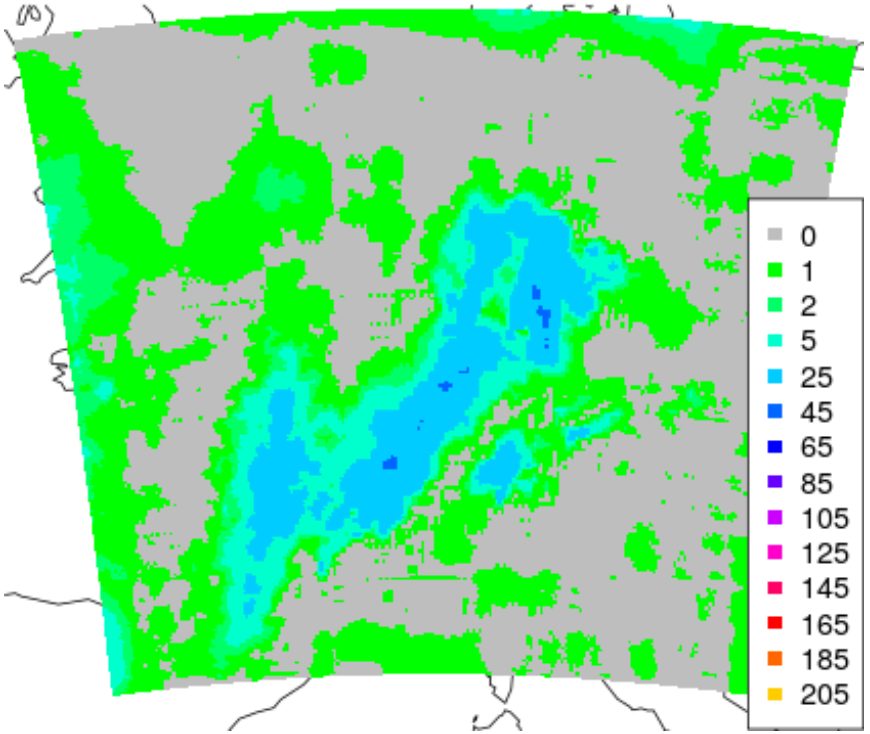
- One of the main problem for the application of several spatial verification methods is the difficulty to have gridded observation data
- DIST does not need gridded data and can be applied either to sparse points or gridded data
- MesoVict project gives us the possibility to compare the application of DIST for the verification of COSMO-1 using the VERA analysis or the JDC observations

MesoVict case 1:observational dataset

JDC – non GTS stations – 3h

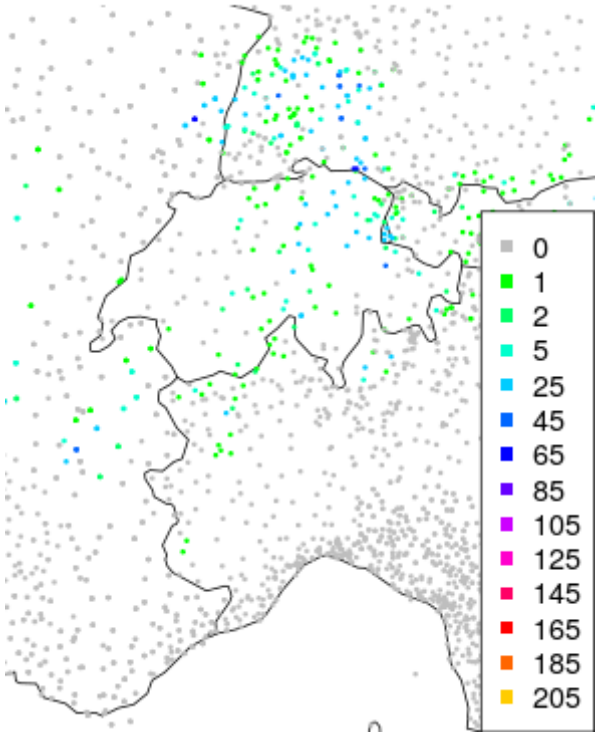


VERA analysis

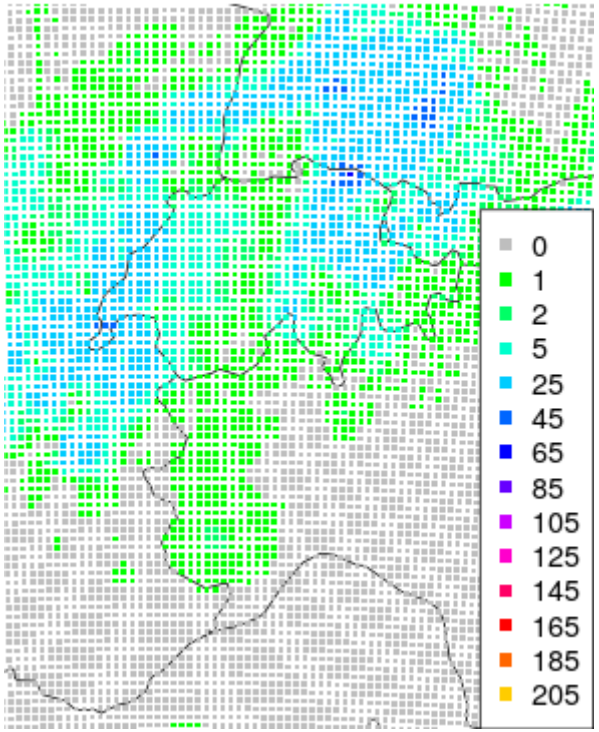


MesoVict case 1:observational dataset (ZOOM)

JDC – non GTS stations – 3h



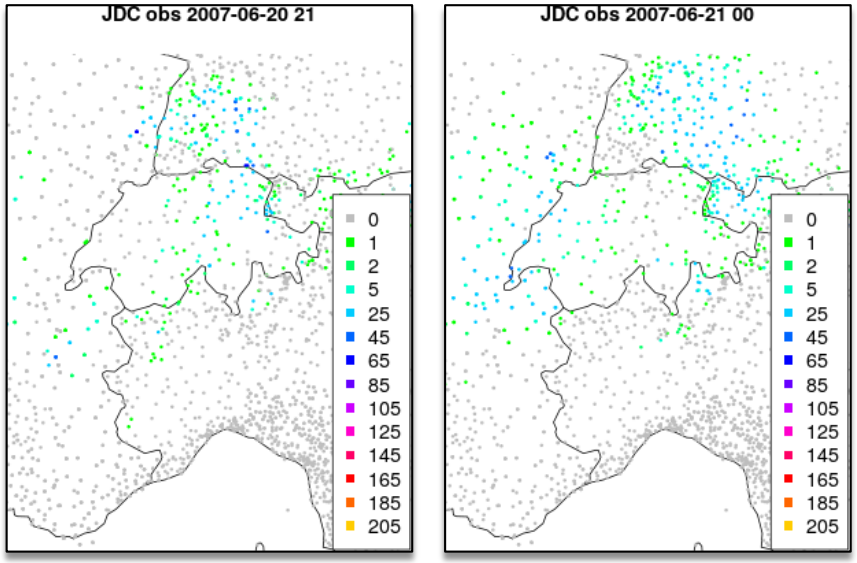
VERA analysis



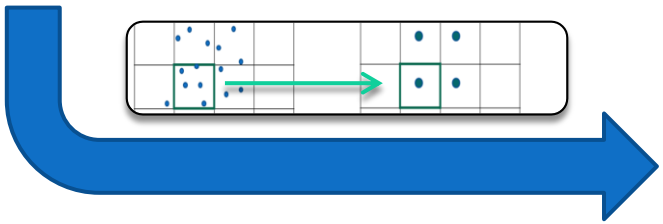
MesoVict case 1 – boxes definition

- We define different box size:
  - 8x8 Km^2 → 1 VERA grid-point
  - 24x24 Km^2 → 9 VERA grid-points
  - 40x40 Km^2 → 25 VERA grid-points
  - 80x80 Km^2 → 100 VERA grid-points

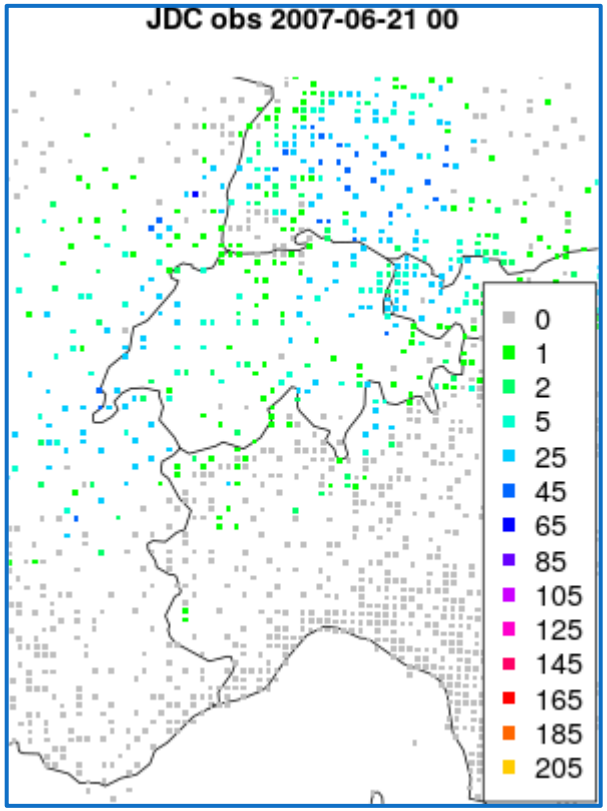
MesoVict case 1- box 24x24 Km^2



Sparse points JDC obs – 3 hr acc

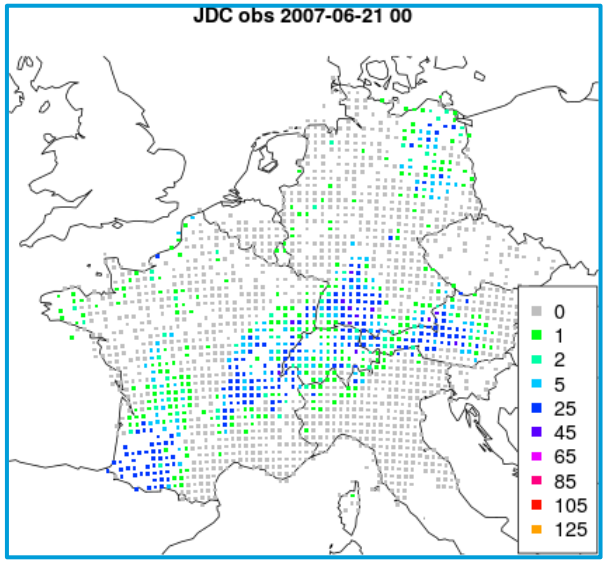


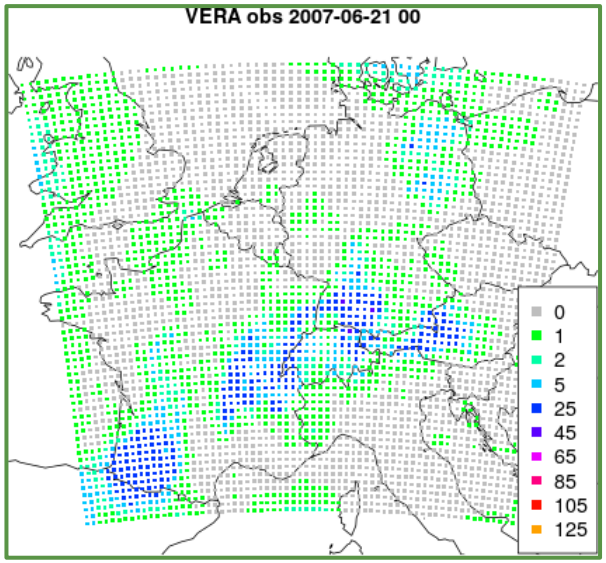
(Zoom)



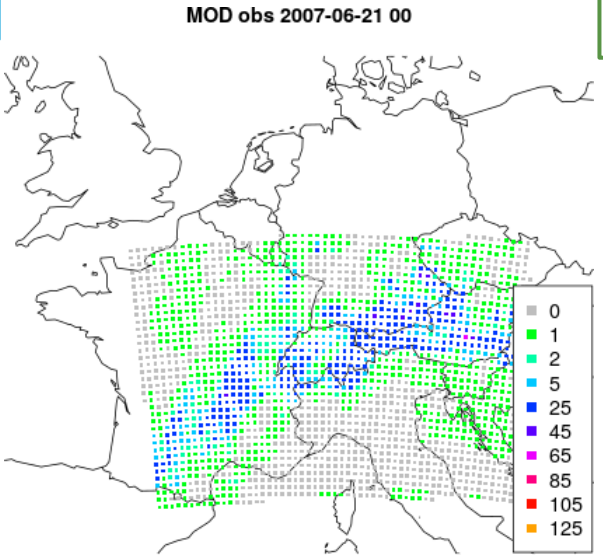
Boxed JDC obs – 6 hr acc

MesoVict case 1- box 24x24 Km^2

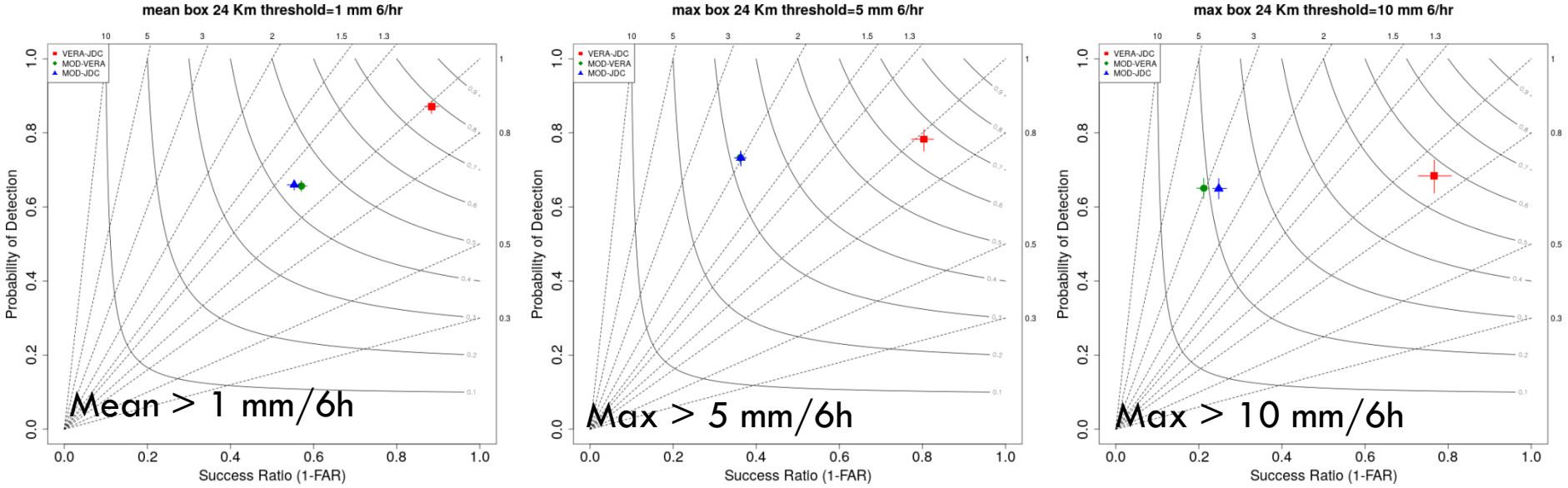




Cross verification between datasets



MesoVict case 1- box 24x24 Km^2

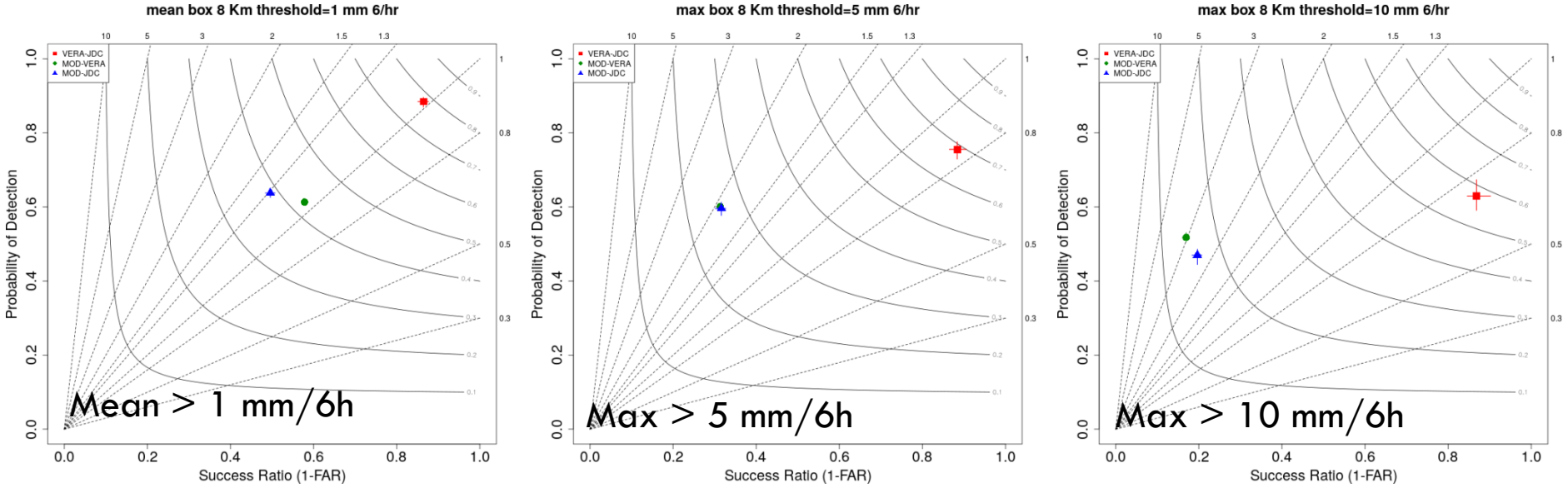


The differences in the scores relative to verification of **Model against VERA** and **Model against JDC** are not significant when **VERA against JDC** perform relatively well.

This means that the average of grid-points of the analysis that fall into each box (9 in this case) is very similar to the ones of the JDC stations.

In case of the maximum, when the threshold is low the differences are small, while when the threshold increases since the analysis tends to be smooth (for definition!). In this case the verification of Models compare to JDC perform a little better (less false alarms) , giving credit to the ability of the model (COSMO1 in the test case) in reproducing high precipitation values.

MesoVict case 1- box 8x8 Km^2

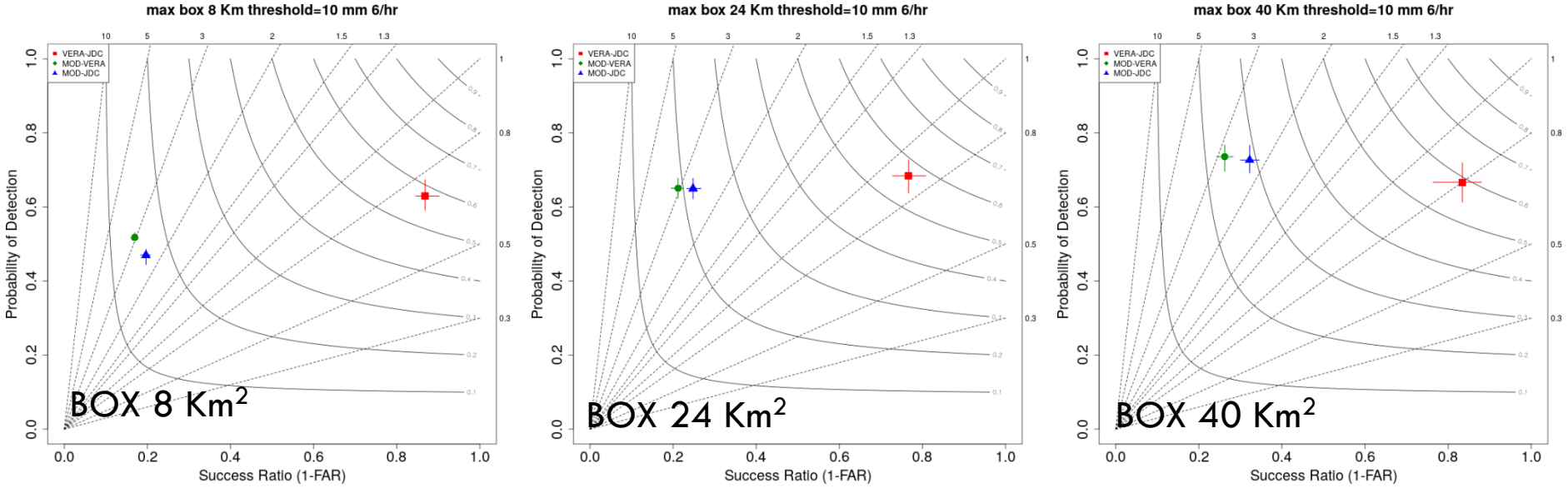


Considering the negative BIAS for "max" of VERA against the JDC sparse point observation, we can say that most of the information about high values of precipitation are missed during the analysis process (even in the VERA cases when thousand of observation were available!)

If we consider smaller boxes (1 VERA grid-point) the difference in the scores are bigger, at least for the mean >1 mm/6h and max>10 mm 6/h .

Nevertheless the information that comes from the performance diagram is the same either for Mod-JDC and Mod-Vera: respect to the bigger boxes, the scores are a slightly worst.

MesoVict case 1- positioning errors



Choosing the representation point for each box (mean, max, some percentile) of forecasts and observations is not only a methodology of upscaling/interpolating field or sparse point data of precipitation but it provides information about the distribution of the precipitation field over areas of various size.

For example, considering the maximum of precipitation over an area we can see that the increase of the box size improves the scores (*in a better way using JDC data*).

This mean that increasing the box size, is possible to take into account the positioning errors of the precipitation filed.

However the dimension of the boxes should be chosen according to the needs of the users.

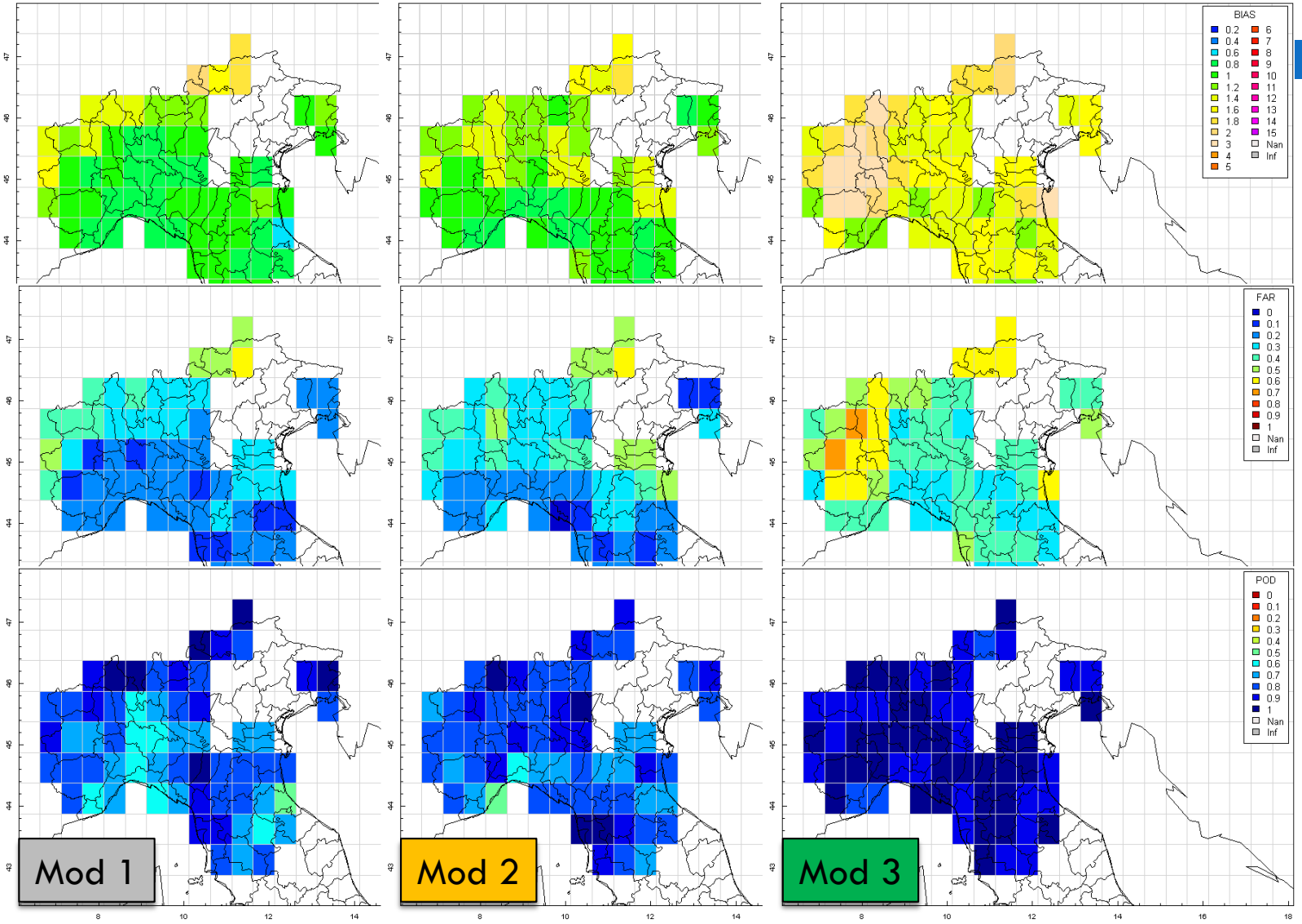
Example of scores:
MEAN VALUE > 1 mm/24h

Size of boxes:
50 Km x50 Km

BIAS

FAR

POD



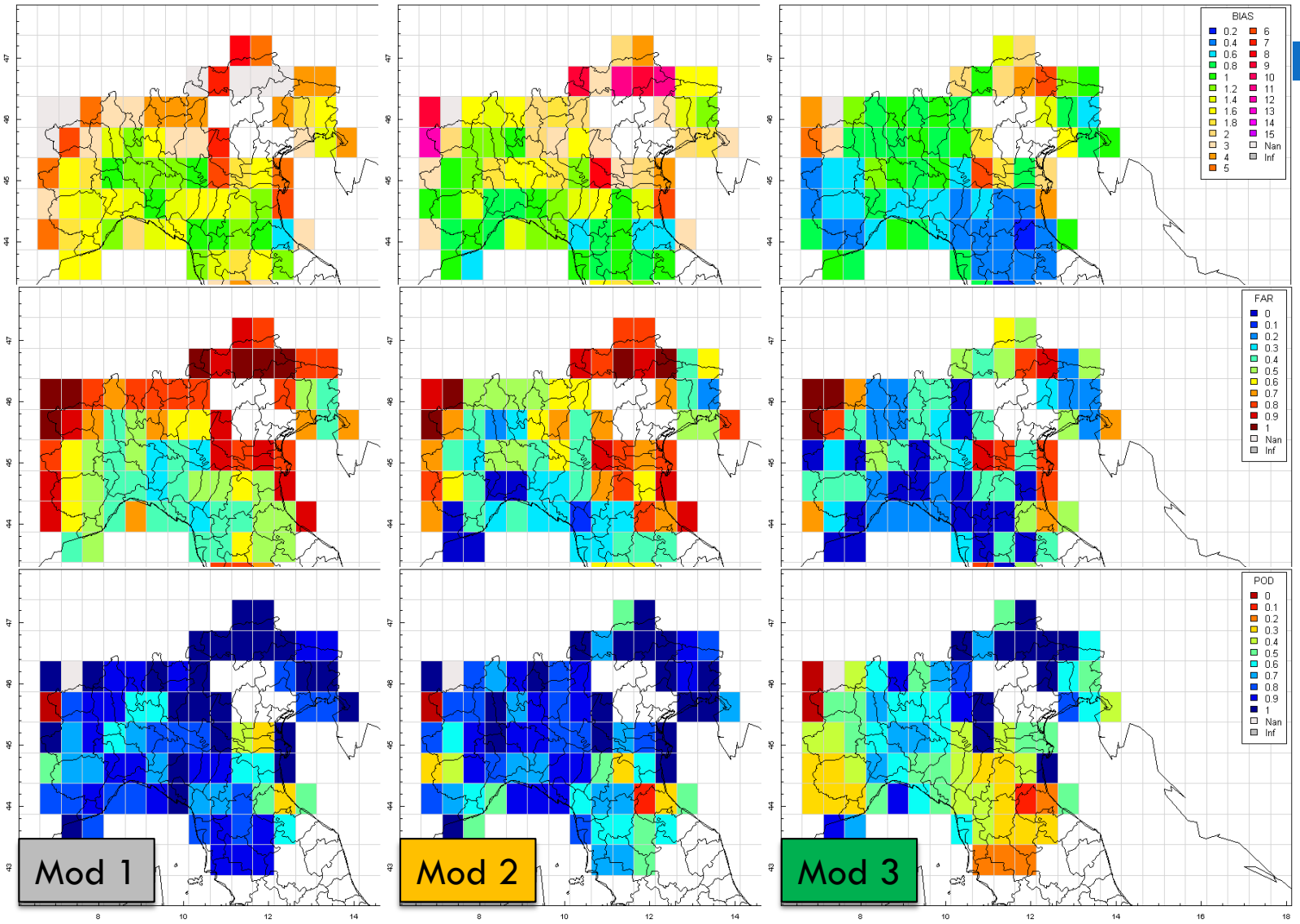
Example of scores:
MAX VALUE > 20 mm/24h

Size of boxes:
50 Km x50 Km

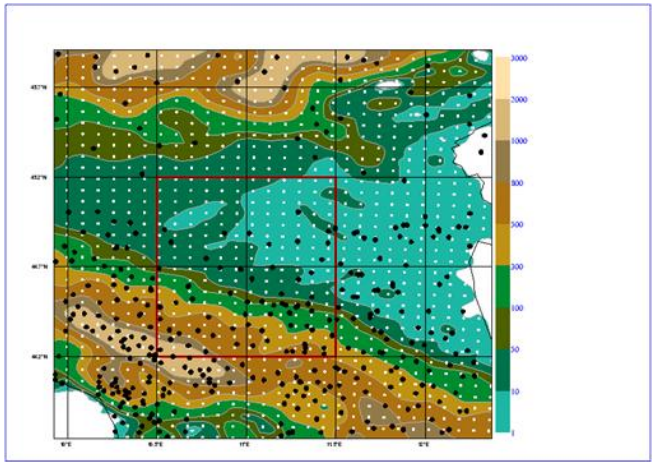
BIAS

FAR

POD

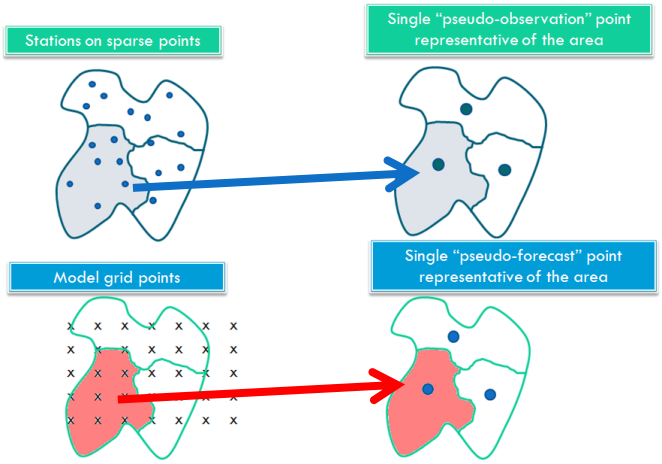


The location of the boxes



- The orographic features of the box are also important: if a ridge of a mountain divide the box this can give misleading results combining upwind and downwind situation
- The choice of regular squared boxes is simple from a computation point of view but is difficult to apply in case of complex terrain (or at least you need more care in interpreting the results)

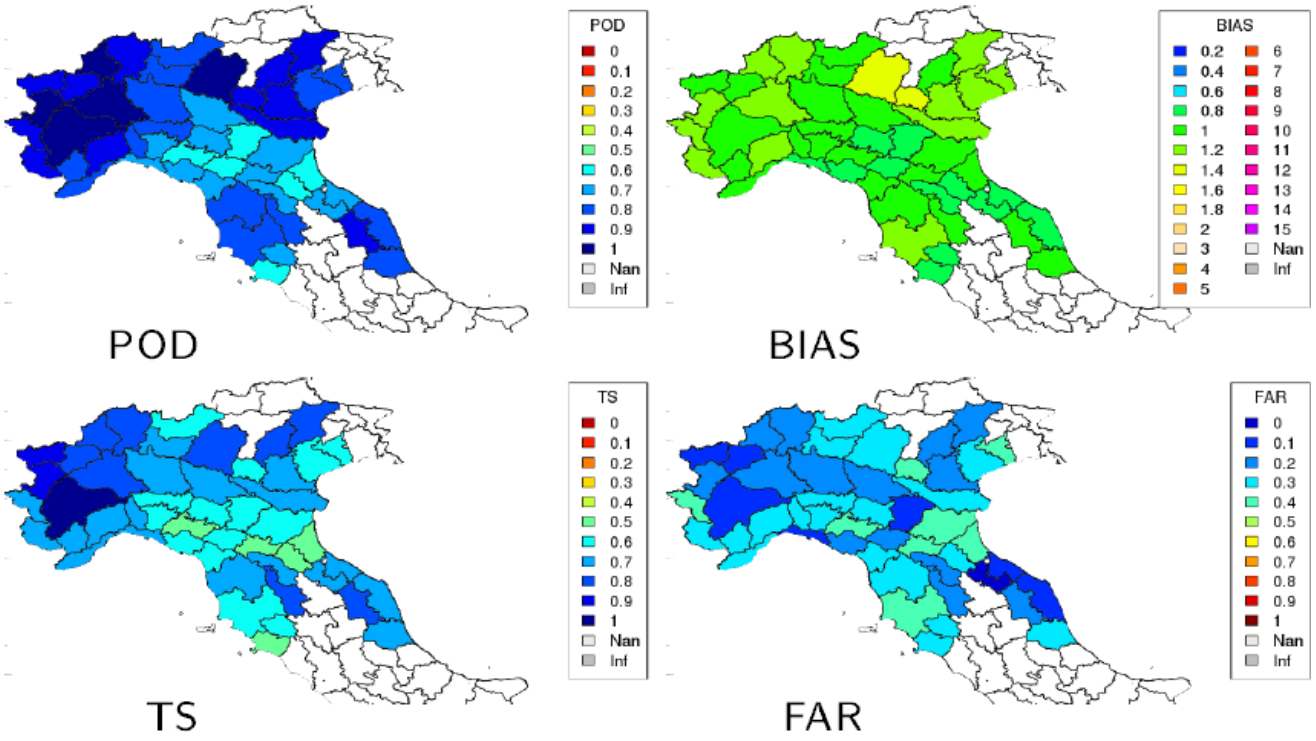
- Since one of the main goal of QPF verification is to give information about the usability of NWP models to forecaster and hydrologist (mainly for civil protection purpose), we decided to move from boxes to catchment area
- The methodology is the same but evaluating mean or maximum value in the area has now a more direct connection with the "real world"



Example of operational verification at

COSMO-I2 – MEDIA > 1mm/24h

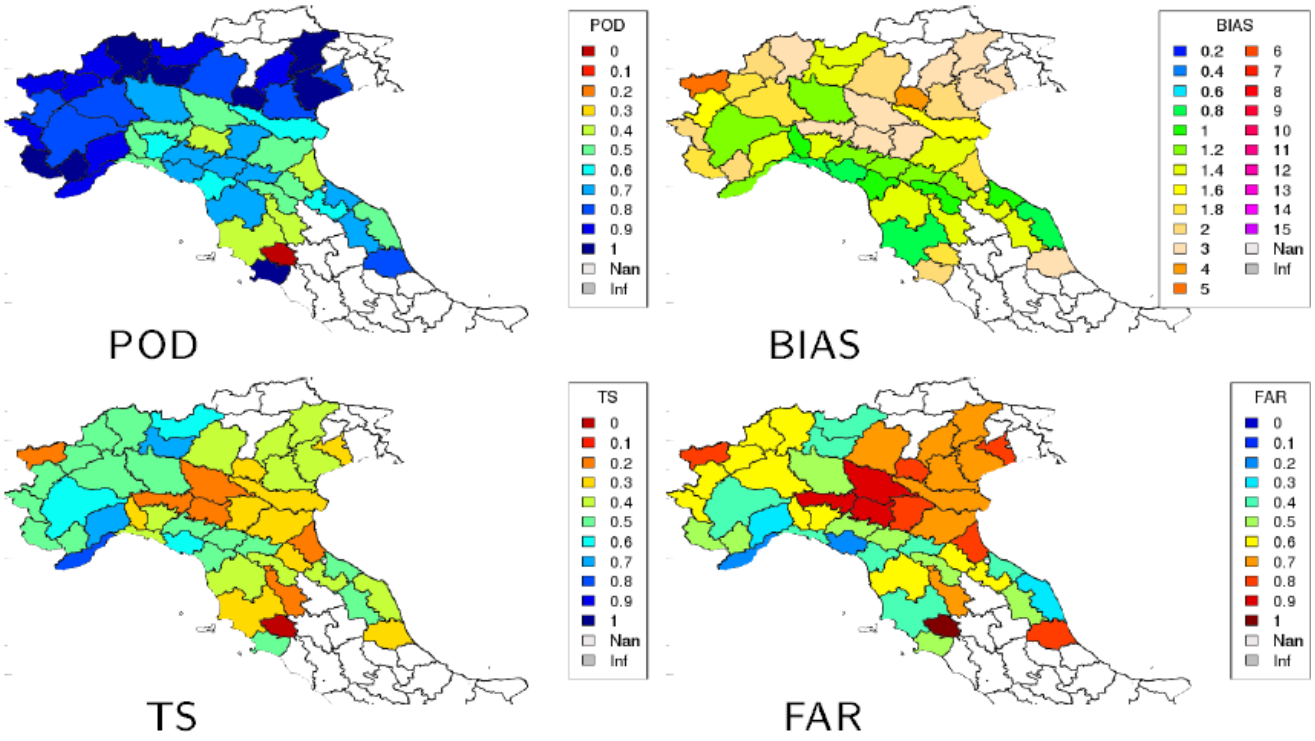
run 00 UTC - cumulata in 24 ore a +48h



Example of operational verification at

COSMO-I2 – MAX > 20mm/24h

run 00 UTC - cumulata in 24 ore a +48h

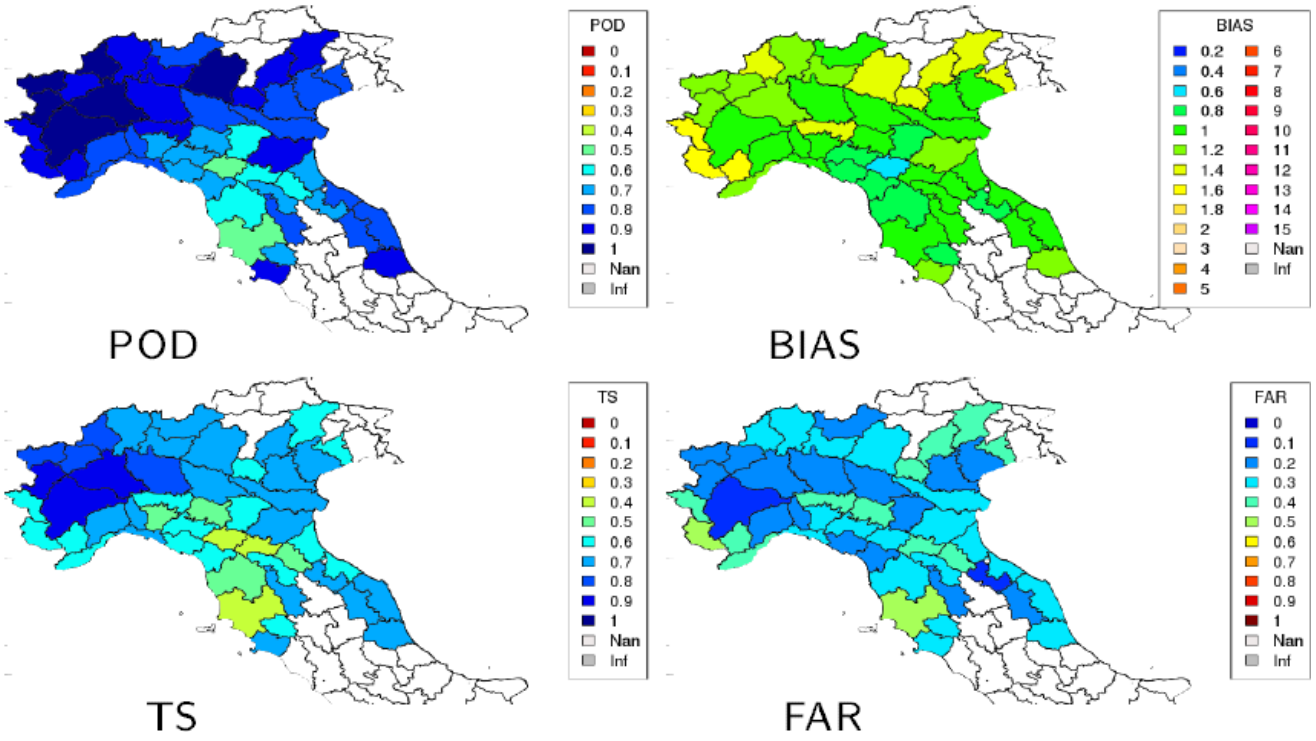


MAM2017

Example of operational verification at

COSMO-I7 – MEDIA > 1mm/24h

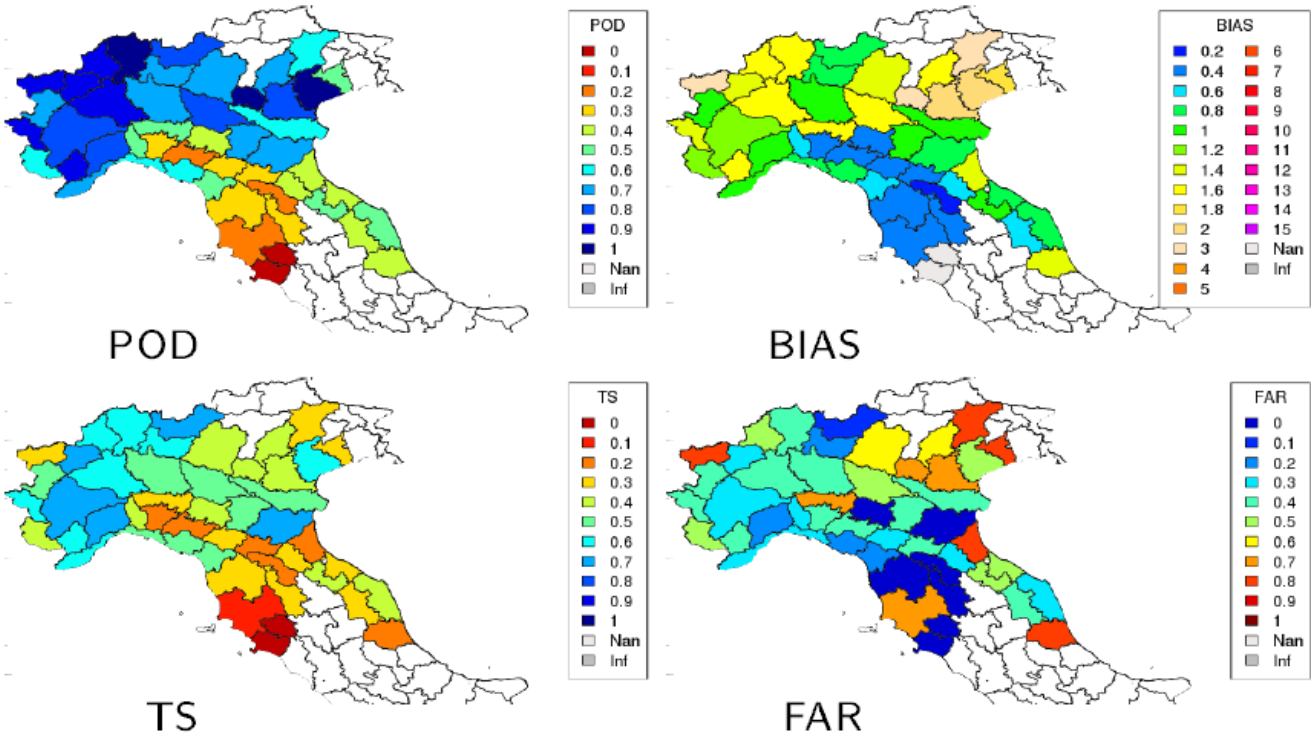
run 00 UTC - cumulata in 24 ore a +48h



Example of operational verification at

COSMO-I7 – MAX > 20mm/24h

run 00 UTC - cumulata in 24 ore a +48h



MAM2017

Example of operational verification at

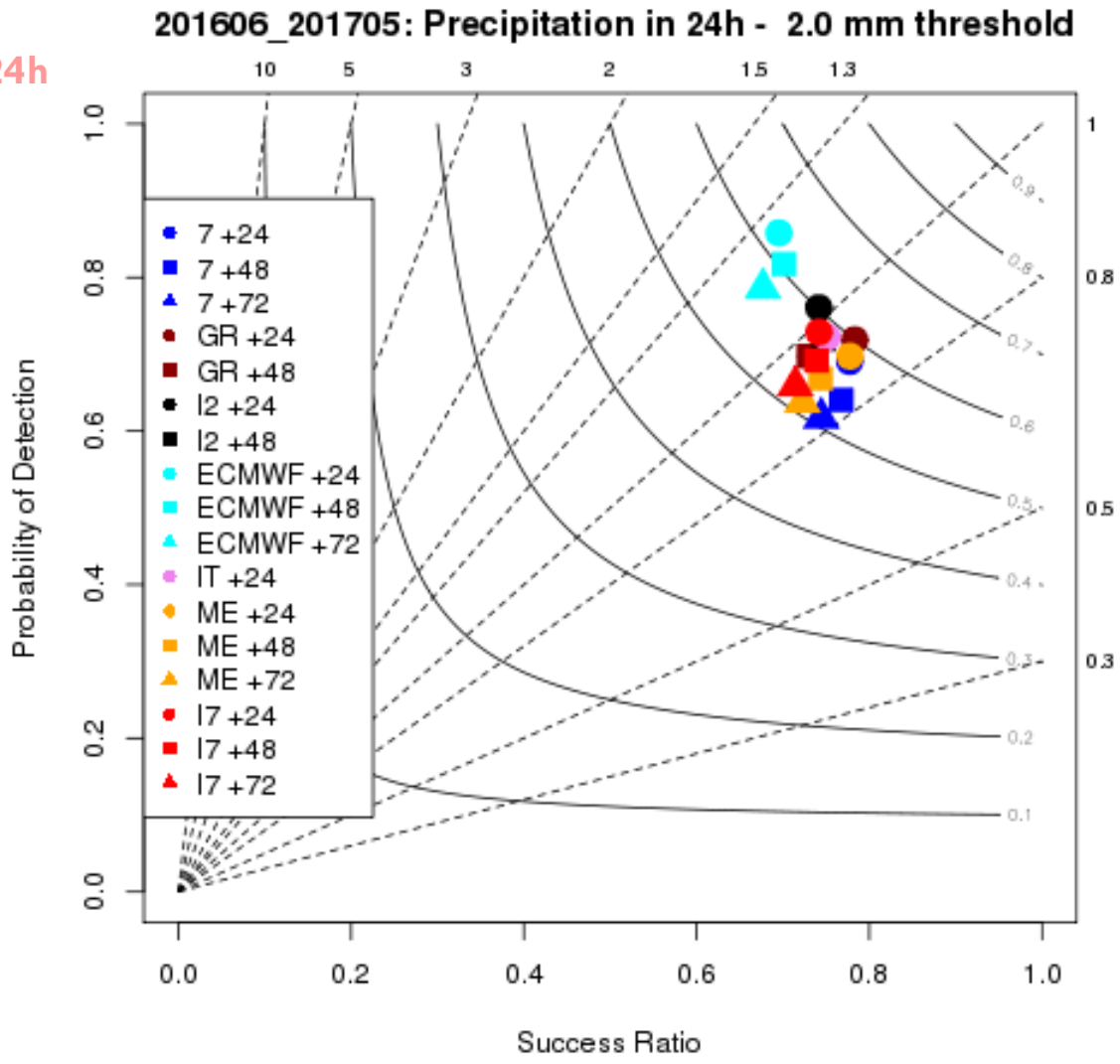
Model's availability:

- Cosmo-EU until 20161130
- Cosmo-ME until 20170508
- Ifs (res 0.125) until 20170430

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

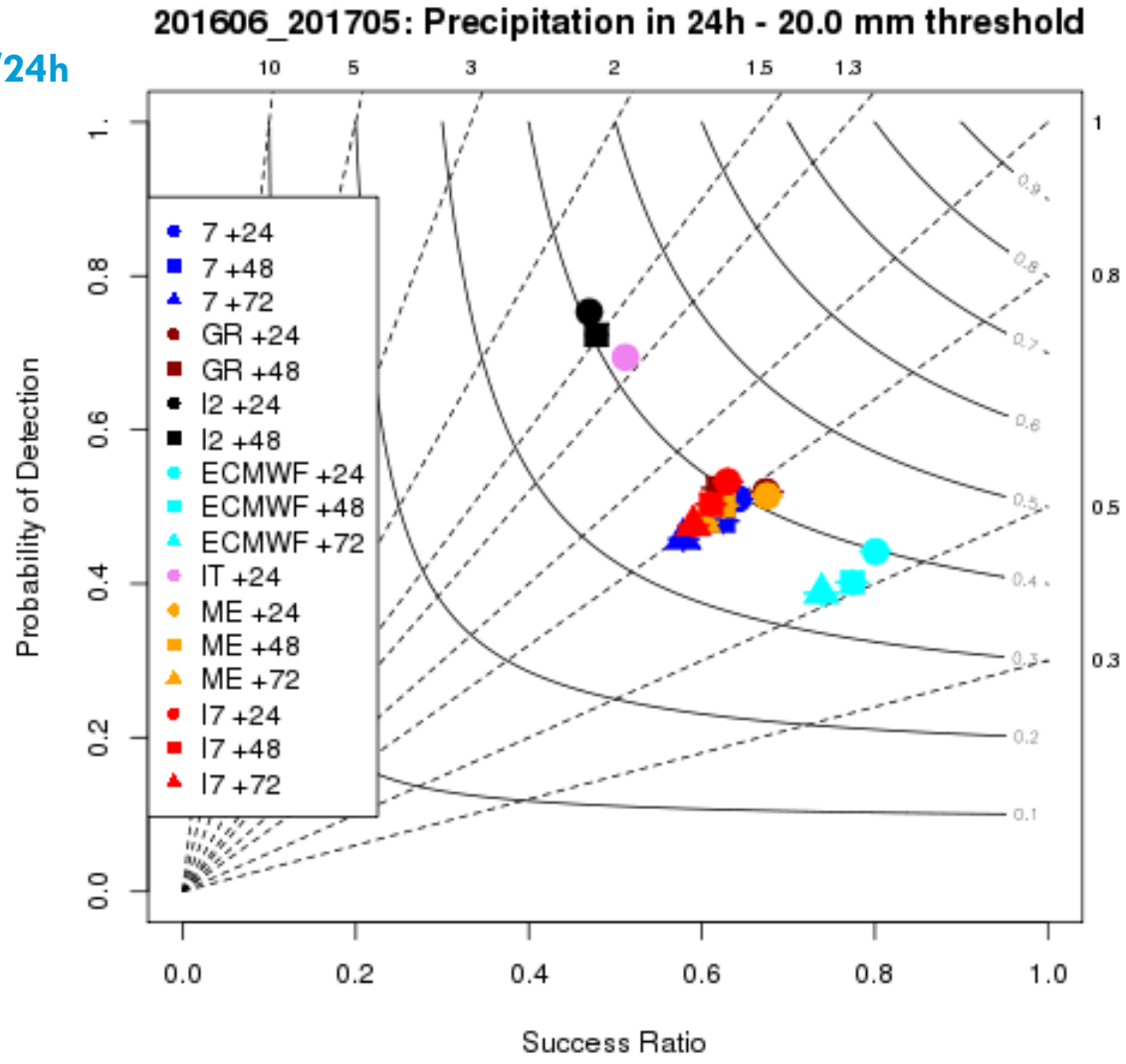
Example of operational verification at

Mean> 2 mm/24h



Example of operational verification at

Max> 20 mm/24h

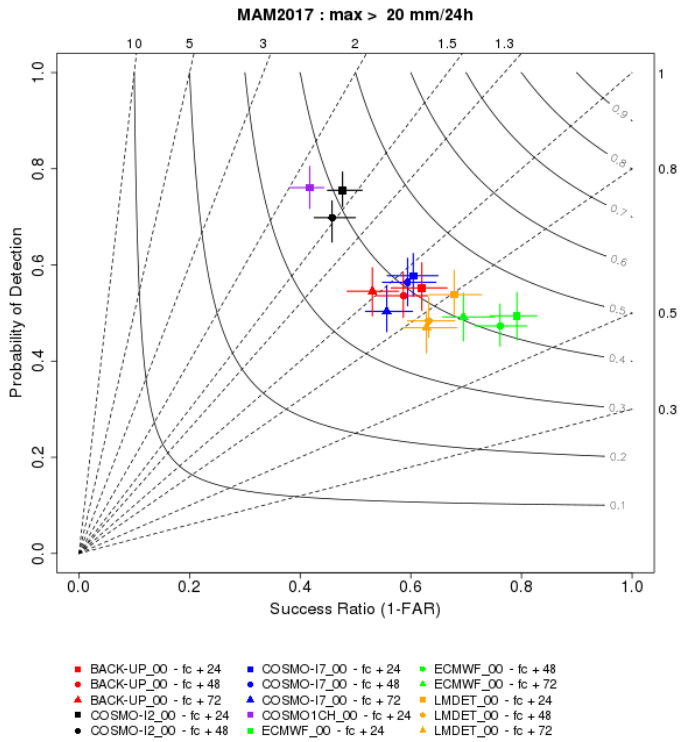


Advanced use of DIST

- Since now we have described the performance of models as a function of distinct distribution parameters (mean or maximum) and thresholds
- But it is also possible to use a combination of the distribution parameters to describe the type of precipitation, such as median and maximum

How to use a combination of the distribution parameters to describe the type of precipitation?

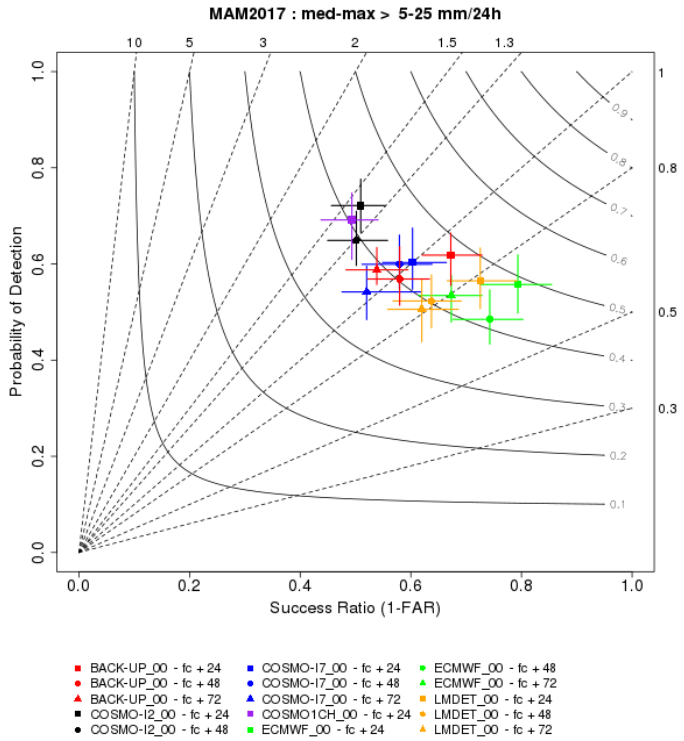
Max > 20 mm/24h



- At least one value in the area is greater than 20 mm/24h:
  - Local precipitation?
  - Widespread precipitation?
- *TS is almost the same for all the model, but higher resolution models (COSMO1 and COSMO12) have better POD but they overestimates the number of the events with a large number of False Alarm*
- *Low resolution model (ECMWF) underestimates the number of events with many misses*

How to use a combination of the distribution parameters to describe the type of precipitation?

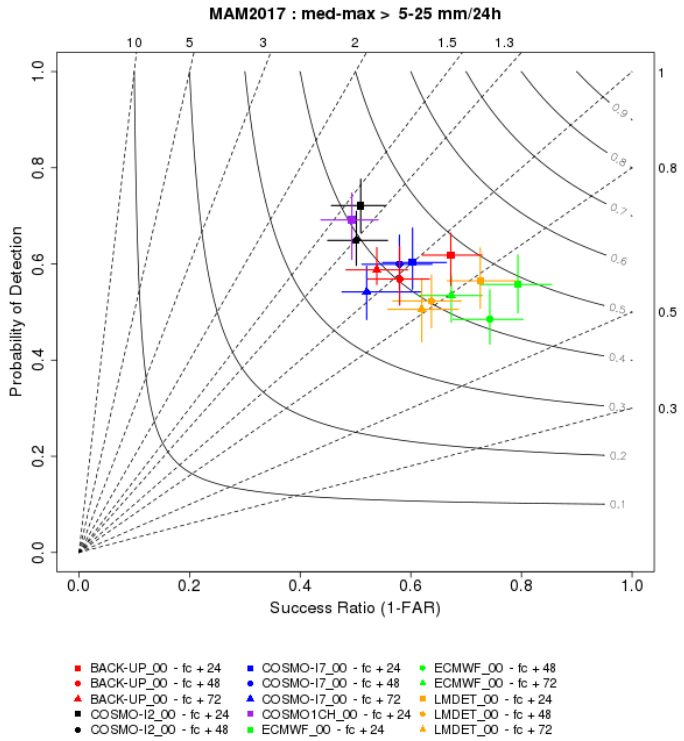
Median > 5 mm/24h
& Max >25 mm/24h



- 50% of the points in the area have more than 5 mm/24h and at least one value is greater than 25 mm/24h
  - The chance that there is an isolated high rain event is reduced
- *The FAR of COSMO-I2 and COSMO1 is reduced: in case of a widespread precipitation models perform better*
- *This is an important feedback for the forecaster as it allows them to be more confident on the predicted precipitation values*

How to use a combination of the distribution parameters to describe the type of precipitation?

Median > 5 mm/24h
& Max >25 mm/24h



- 50% of the points in the area have more than 5 mm/24h and at least one value is greater than 25 mm/24h
  - The chance that there is an isolated high rain event is reduced
- *The FAR of COSMO-I2 and COSMO1 is reduced: in case of a widespread precipitation models perform better*
- *This is an important feedback for the forecaster as it allows them to be more confident on the predicted precipitation values*

Conclusion

- The application of DIST to MesoVict case 1showed that:
  - the use of sparse point observations (JDC) gives results comparable to that obtained with gridded observations (VERA)
  - using sparse point observations for the verification provides best results for maxima of precipitation.
    This aspect is very useful for the verification of very high resolution models.
- Since the methodology is the same, DIST can be used both with gridded or sparse points data according to the availability of them.

Since we don't need an analysis of precipitation to perform the verification!!