

Focus on Spatial Verification Filtering techniques

Flora Gofa



ΕΘΝΙΚΗ
ΜΕΤΕΩΡΟΛΟΓΙΚΗ
ΥΠΗΡΕΣΙΑ

HELLENIC NATIONAL METEOROLOGICAL SERVICE

Approach

- Attempt to introduce alternative methods for verification of spatial precipitation forecasts and study their relative benefits
- Techniques that allow to diagnose the skill of a system as a function of the **spatial scale of the forecast error and intensity** of the precipitation events
- Aim is to compare the performance of precipitation forecasts for two different resolutions of COSMO model over the core MesoVICT case
- Precipitation events on different spatial scales are caused by different physical processes and the evaluation of forecast skill on various scales becomes important



Preparation/Adaptation of MesoVICT datasets as input for VAST

MesoVICT datasets

Forecast Data:

1. **Model Data interpolated on the VERA grid
(resolution 8 km)**

ASCII format on a Cartesian grid (non regular)

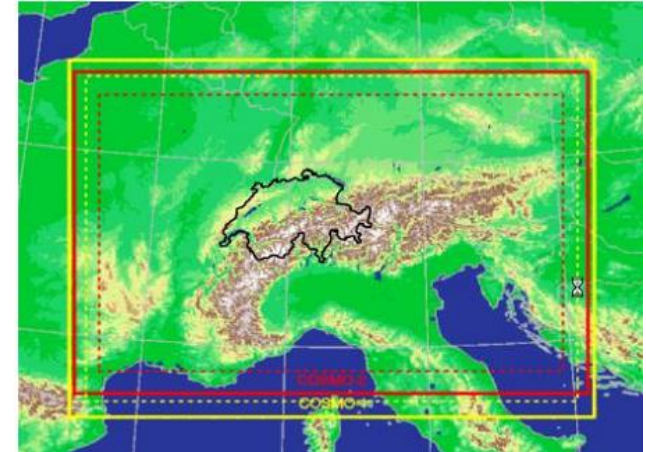
Models available:

COSMO-2 (old runs – version of model):

COSMO-1 (new runs)

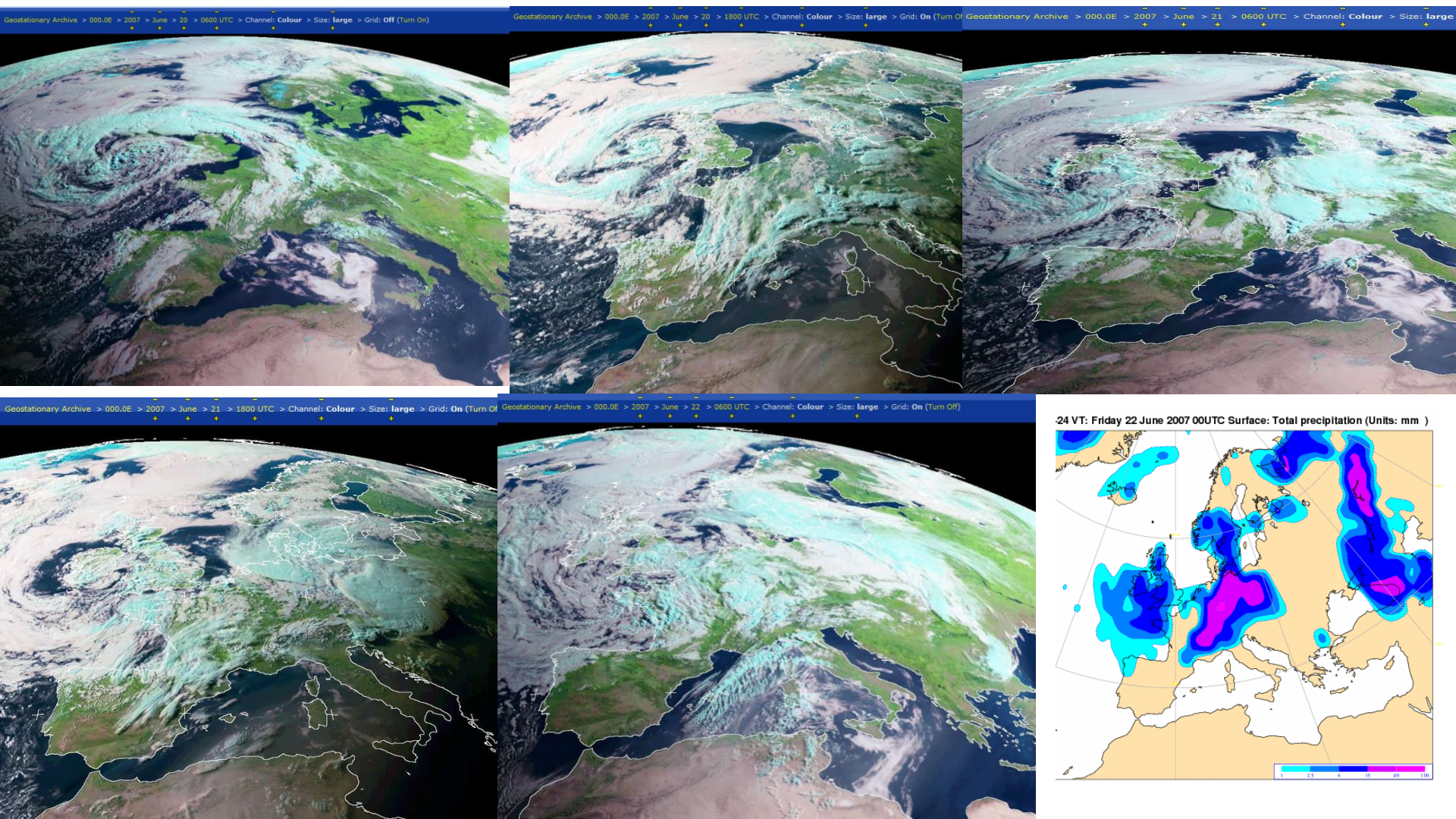
Data:20, 21,22.06.07:00-24UTC. Precipitation, 1h accumulation

- **VAST can process only regular lat-lon files so these data will have to be interpolated for a second time (!) on a regular grid, to be usable by software**
- Observation data used: VERA analysis in ~8km resolution



Case 1 (core case): 20-22 June 2007

Ahead of a trough, located over the British Isles, warm moist air is advected towards the Alpine Region. This leads to strong convective events in the evening of 20 June, in the area north of the main mountain range. On the next day (21st) a cold front is reaching the Alps from the west and moves to the east rather quickly. Ahead of the front again convective events are observed. With the passage of the front strong westerly winds occurred.

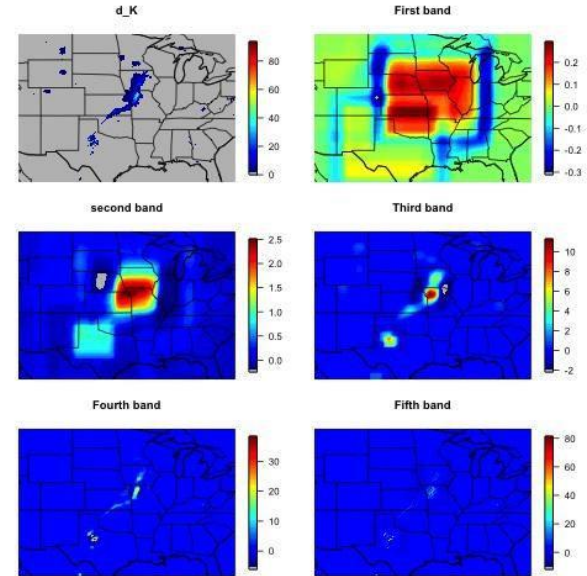
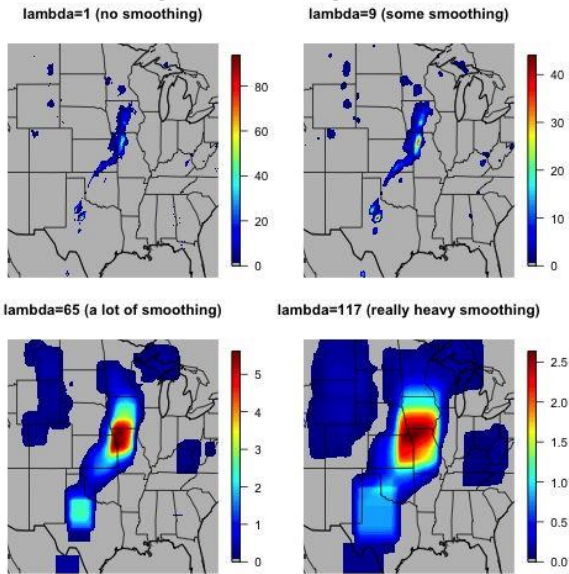


smoothing / neighborhood



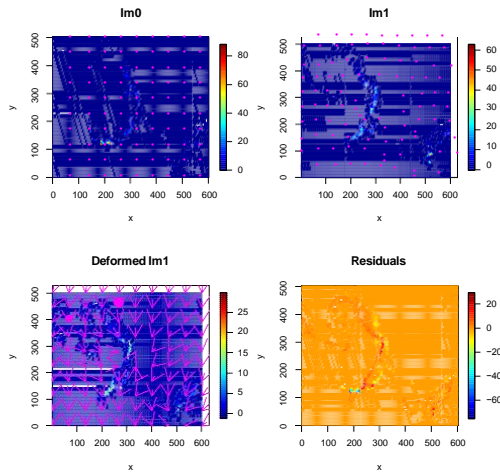
NCAR

Filter

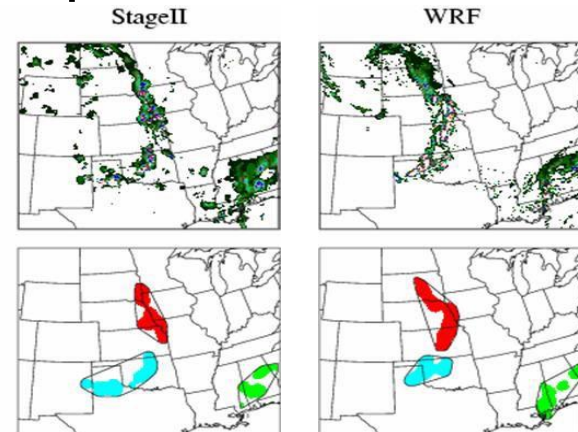


Scale Separation

Entire field



specific features



Spatial Displacement

Intensity-scale verification

B. Casati, G. Ross and D.B. Stephenson (2004) “A New intensity-scale verification approach for the verification of spatial precipitation forecasts”, Meteorol Appl, vol 11, 141-154 pp

Evaluate the forecast skill as a function of the precipitation **intensity** and the **spatial scale** of the error

NOTE: **scale** = single band spatial filter → features of different scales → feedback on different physical processes and model parameterizations

In the **neighborhood based (fuzzy) verification**, the **scale** is the neighborhood size (**low band pass filter**): as the scale increases the exact positioning requirements are more and more relaxed

Steps in IS verification

Binary error decomposition: Thresholding is used to convert the forecast and analysis into binary images for each of the rainfall rates.

Binary error is the difference of this $Z = I_y - I_x$

Binary error image is then expressed as the sum of components on different spatial scales by performing a 2-dimensional discrete Haar wavelet decomposition $Z = \sum_{l=1}^L Z_l$

Most substantial binary error image of the mother wavelet components are calculated for various spatial scales ($l = 1, \dots, L = 7$). **The spatial scales refer to the spatial scale of the error and not that of the precipitation features or their displacement as it happens in the neighborhood methods**

- The MSE of the binary error image is calculated from $MSE = \sum_{i=1}^L MSE_l$ $MSE_l = \overline{Z_l^2}$ while for each threshold the skill score can be calculated from

$SS = \frac{MSE - MSE_{random}}{MSE_{best} - MSE_{random}}$, where MSE_{random} is associated with a random forecast calculated from the bias and the base rate at each threshold

- Intensity scale verification technique is a spatial generalization of traditional binary verification ([HSS](#))
- Application using SpatialVx: wavelS routine

Two-dimensional discrete Haar wavelet filter (Casati et al, 2004)

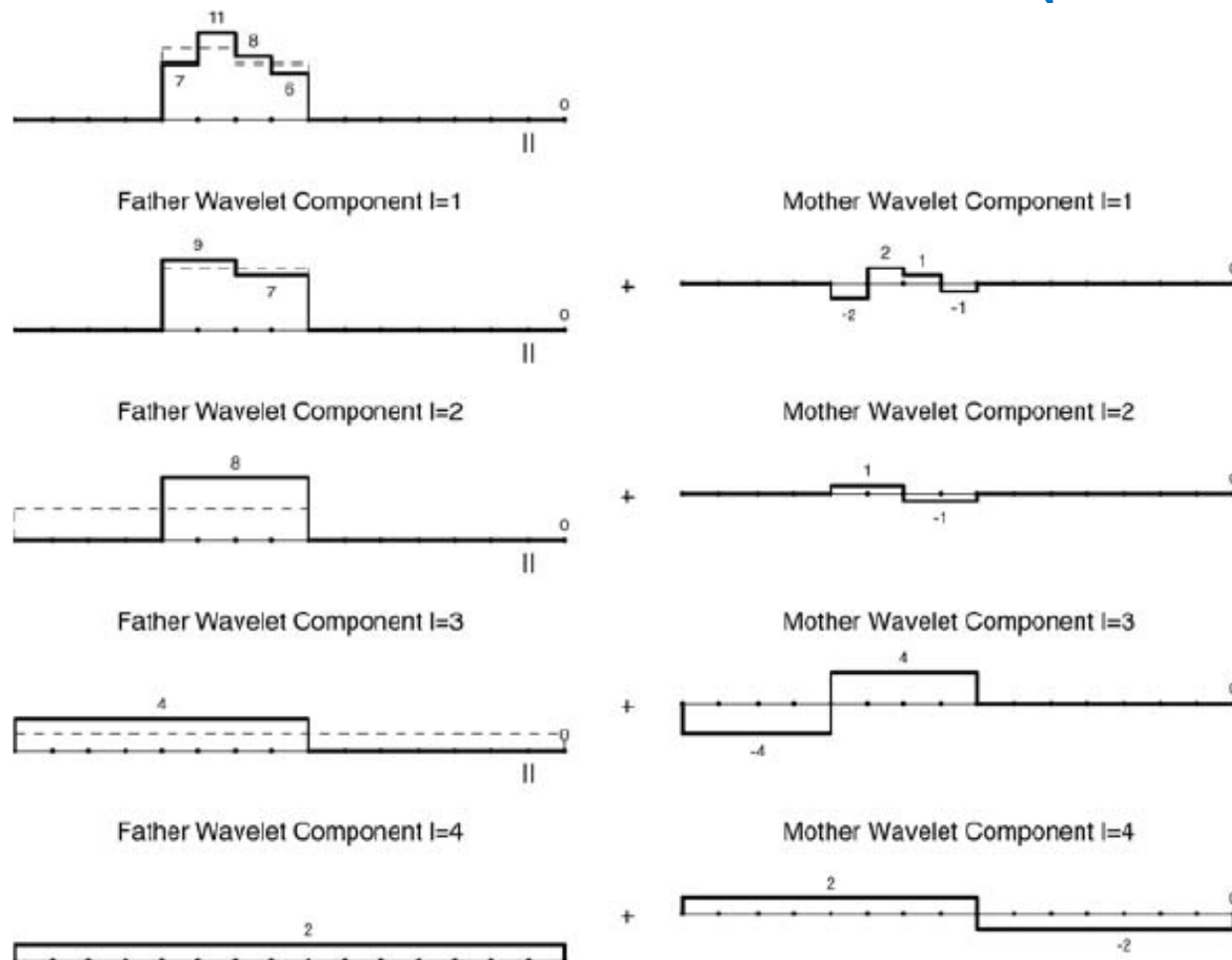
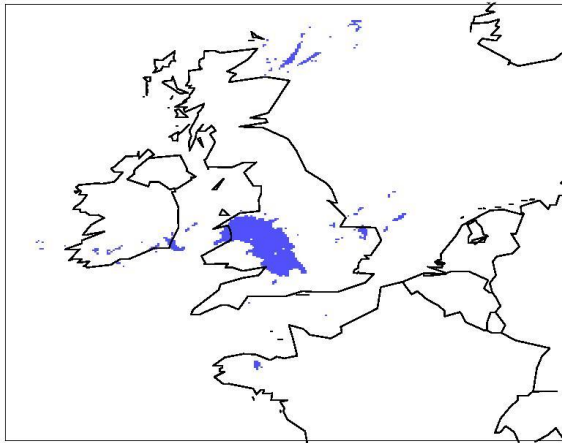


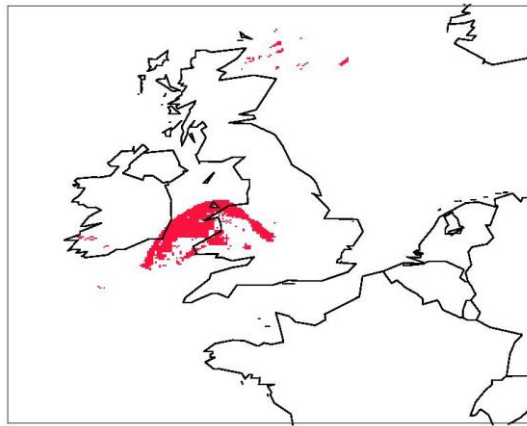
Figure 11. Example of the one-dimensional discrete Haar wavelet filter applied to an example function (top left panel). At the first step the function is decomposed into the sum of a coarser mean function (the first father wavelet component) and a variation-about-the-mean function (the first mother wavelet component). At each step the Haar wavelet filter decomposes the father wavelet component obtained from the previous step into the sum of a coarser mean function (the l^{th} father wavelet component) and a variation-about-the-mean function (the l^{th} mother wavelet component). The l^{th} father wavelet component is obtained from the initial function by a spatial averaging over 2^l pixels. The process stops when the largest father wavelet component (mean over the whole domain) is found.

Links with categorical verification

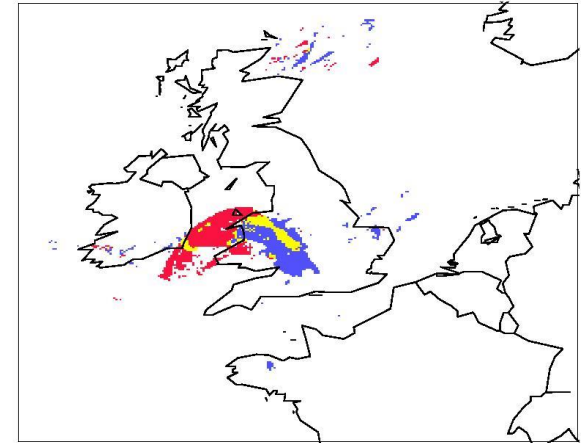
Binary Analysis



Binary Forecast



Overlapping



$$MSE_u = \frac{b+c}{n}$$

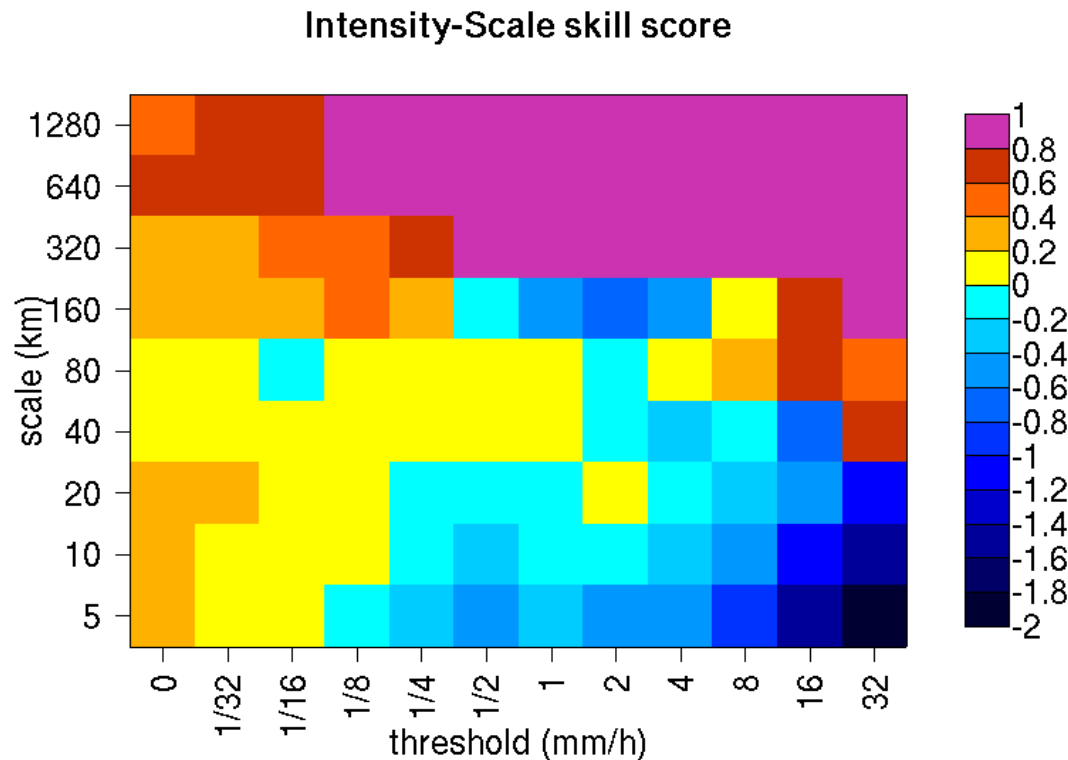
$$SS_u = HSS$$

	$X > u$	$X < u$	
$Y > u$	Hits a	False Alarms b	a+b
$Y < u$	Misses c	Correct Rejections d	c+d
	a+c	b+d	a+b+c+d=n

Intensity-scale skill score (SS)

For each threshold and scale component: skill score associated to the MSE of binary images (= HSS). Skill versus random chance, equally partitioned across the scales. The IS skill score is capable of isolating specific scale-dependent errors.

Usually, small scales exhibit negative skill, whereas large scales exhibit positive skill

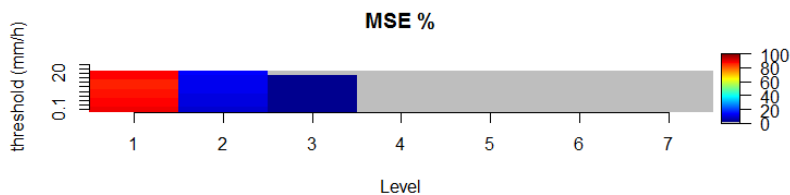
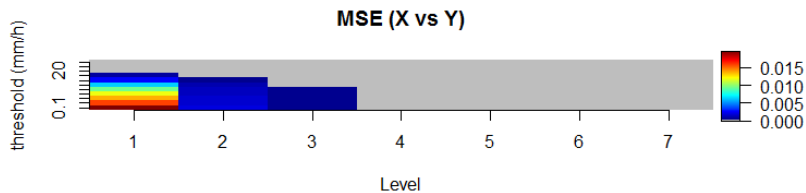
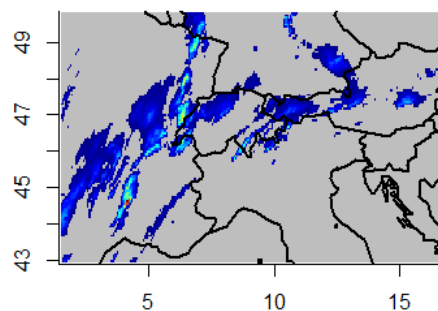
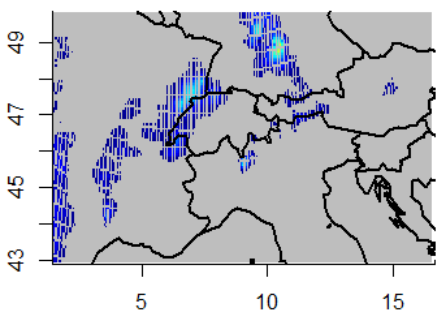


Cosmo1

New
Rainfall (mm/h)

obs

fcs

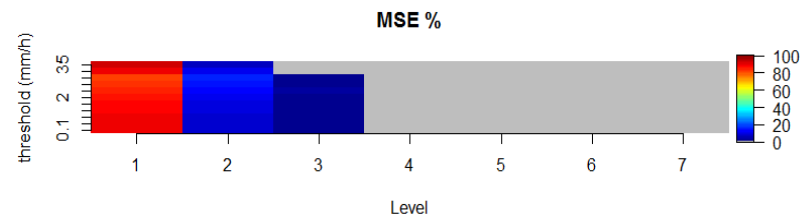
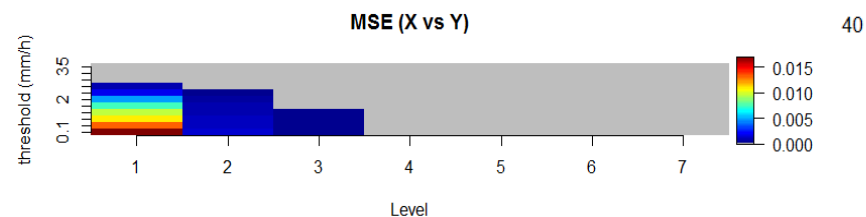
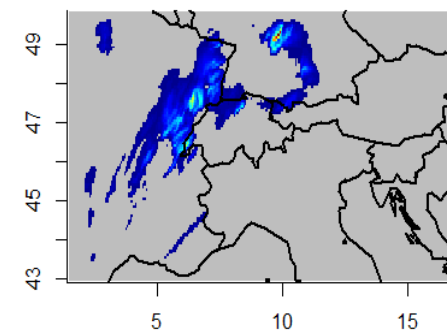
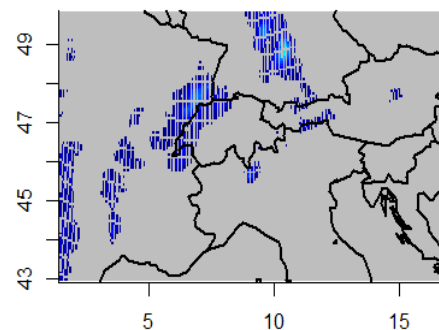


Cosmo2

New
Rainfall (mm/h)

obs

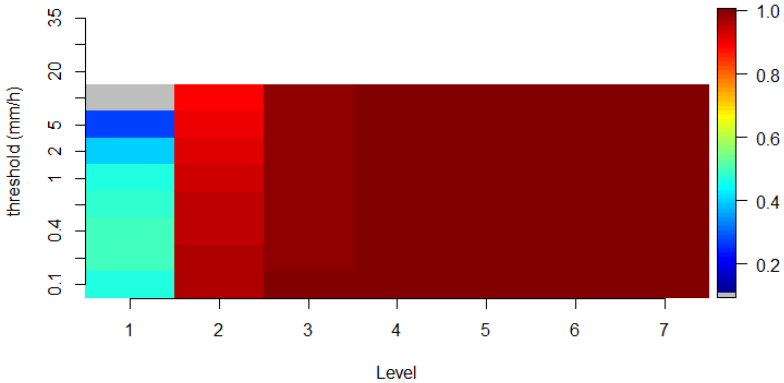
fcs



20070621-01: Skill

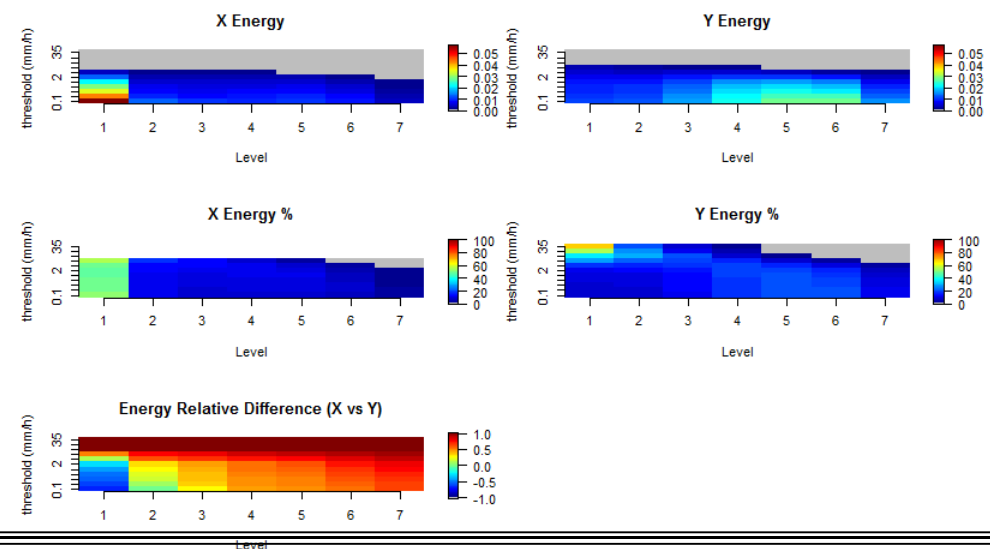
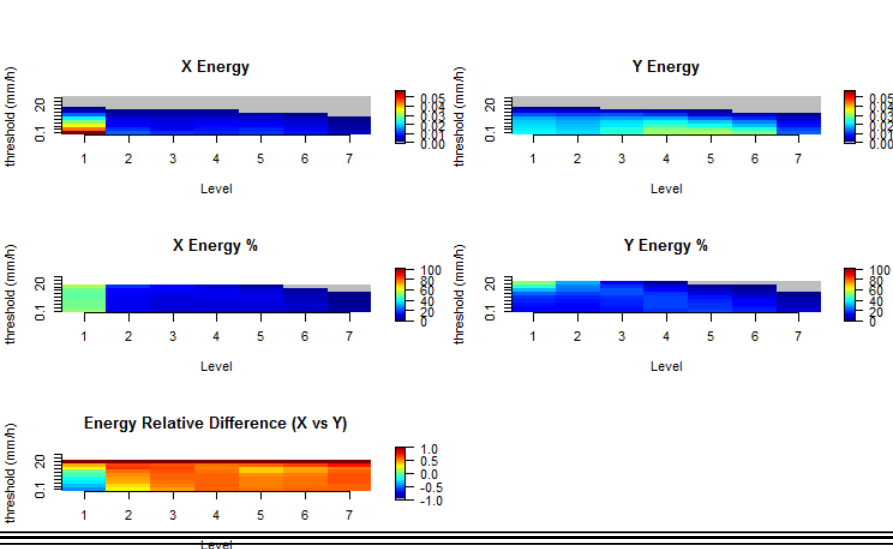
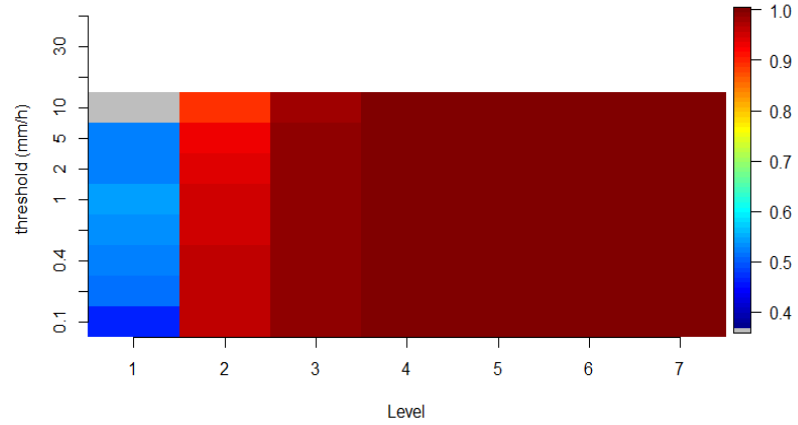
Cosmo1

IS Skill Score (X vs Y)



Cosmo2

IS Skill Score (X vs Y)

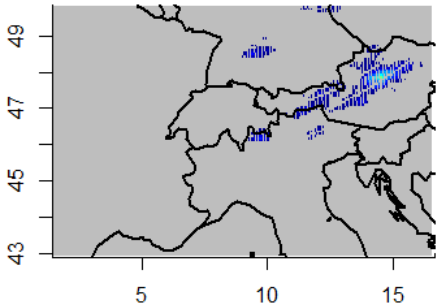


20070621-15:map

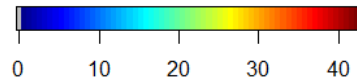
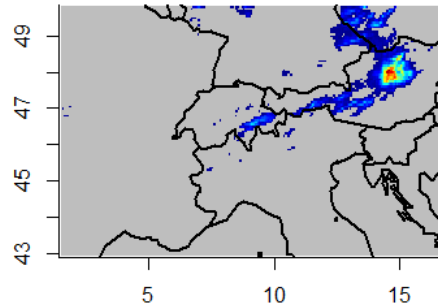
New
Rainfall (mm/h)

Cosmo1

obs



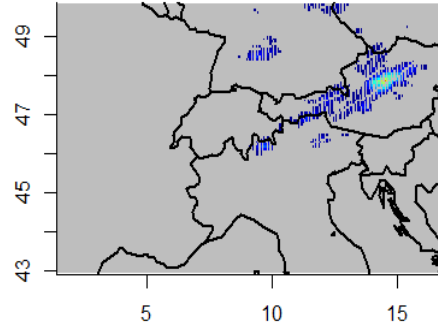
fcs



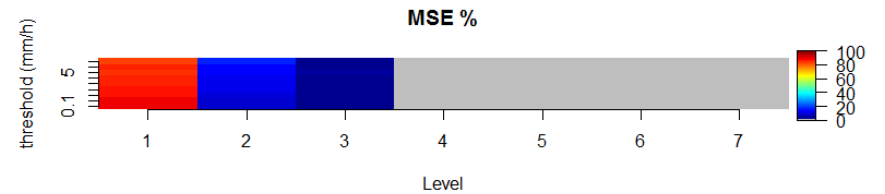
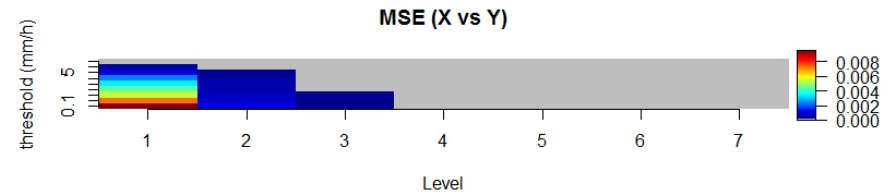
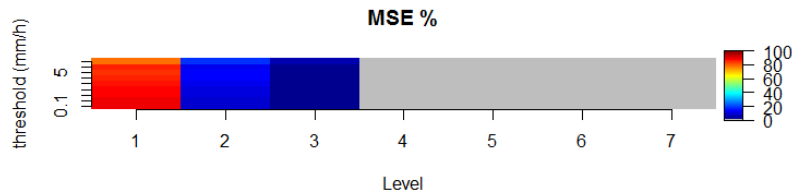
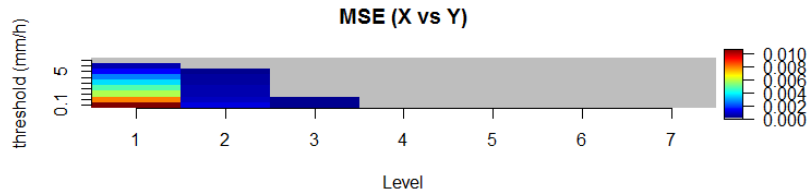
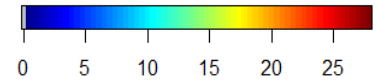
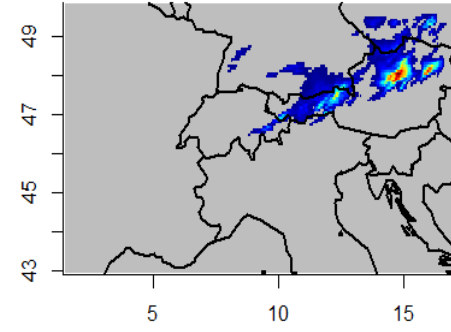
New
Rainfall (mm/h)

Cosmo2

obs



fcs

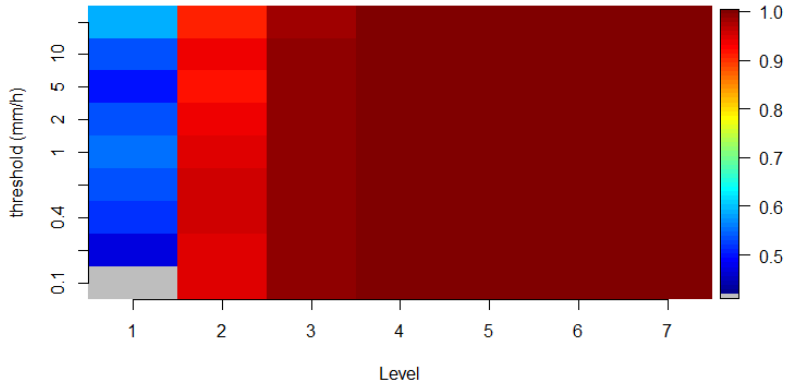


20070621-15: Skill

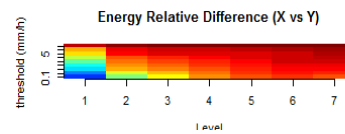
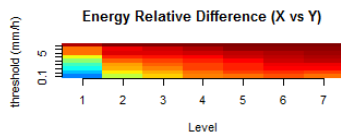
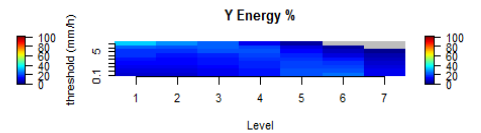
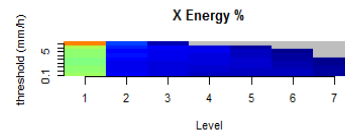
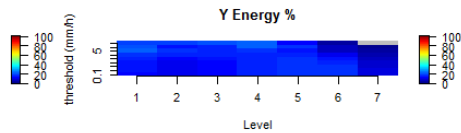
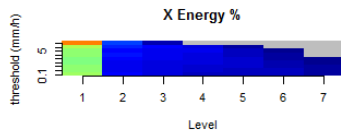
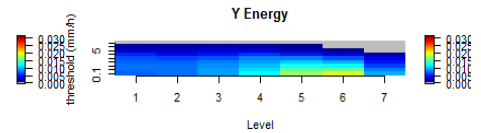
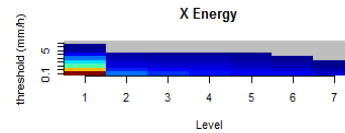
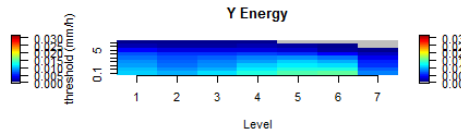
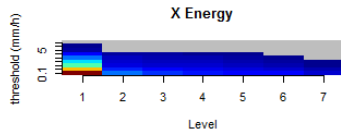
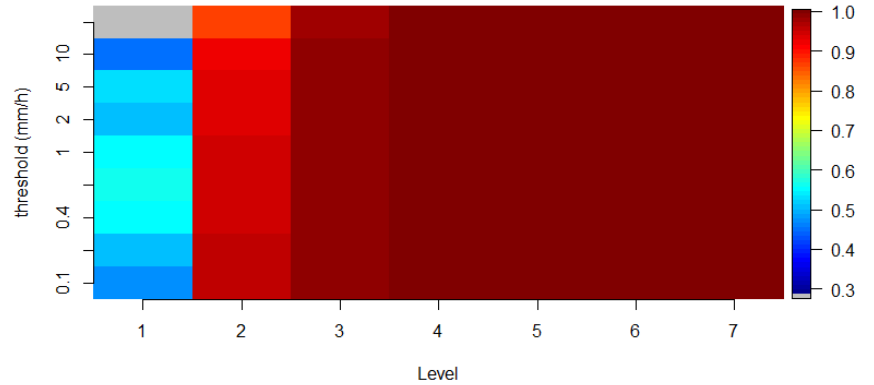
Cosmo1

Cosmo2

IS Skill Score (X vs Y)



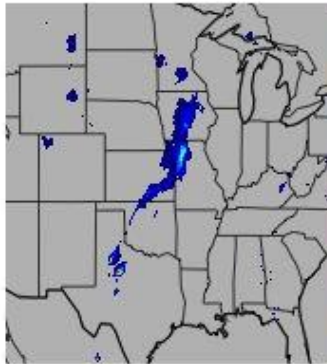
IS Skill Score (X vs Y)



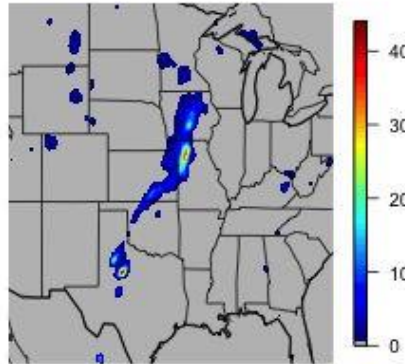
Recalling the ICP

smoothing filter / Neighborhood Methods

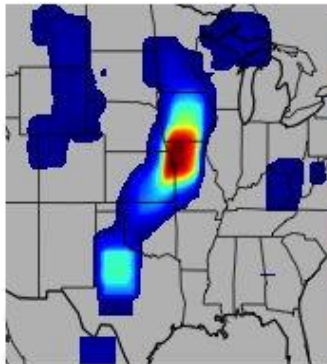
lambda=1 (no smoothing)



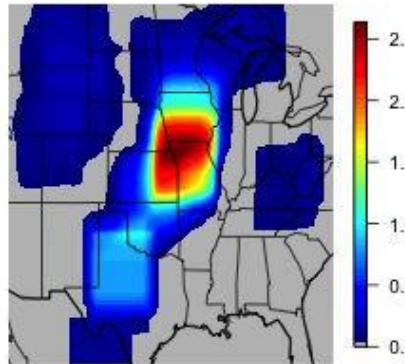
lambda=9 (some smoothing)



lambda=65 (a lot of smoothing)



lambda=117 (really heavy smoothing)



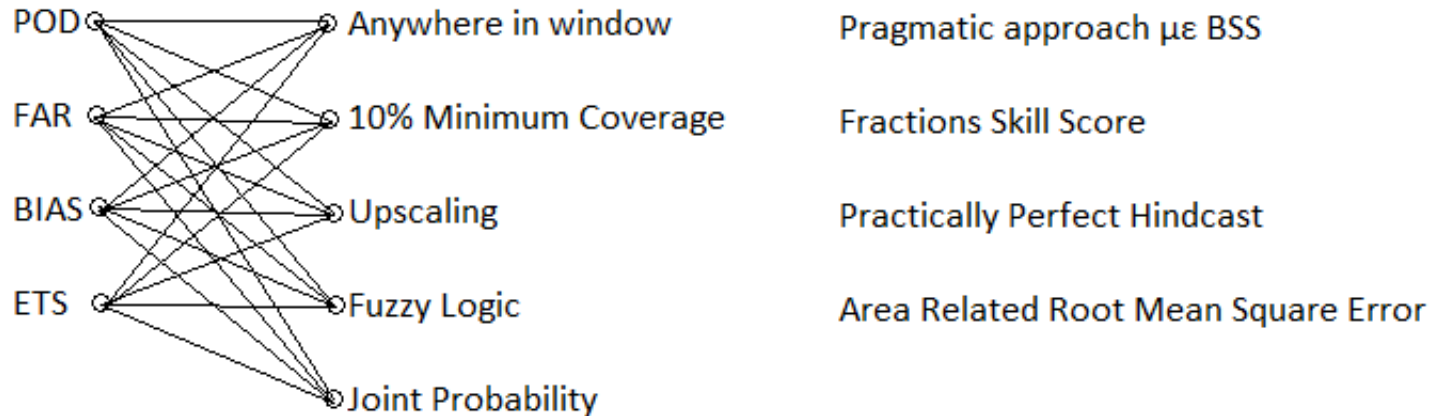
Apply filter to:

- raw field
- binary “thresholded” field (event field)

Apply filter to:

- both fields
- forecast only
- observations only

• Neighborhood methods/scores :VAST



• Categorical verification

BIAS

FAR

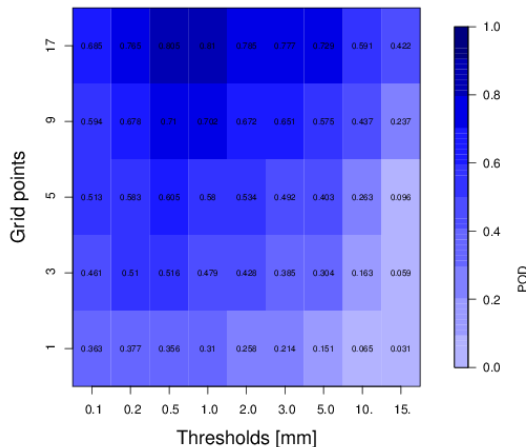
POD

COSMO-1

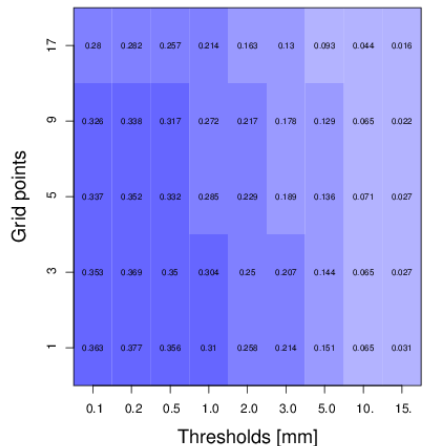
POD

Neighborhood method	Decision Model
Upscaling	In a good forecast it is predicted the same average value of precipitation with the predicted
Fraction Skill Score	A forecast is useful when the frequency of forecasted events is similar with the frequency of the observed events
Minimum coverage	A forecast is useful if the event is forecasted in a minimum fraction of the area of interest
Anywhere in the window	A forecast is useful if it is more true than untrue, characterizing the events as probabilities
Pragmatic approach	A useful forecast has higher probability to discriminate the events and the non events.(in reference with the climatological value derived from the observations)
Area related RMSE	A useful forecast has the same distribution of intensity as the observations

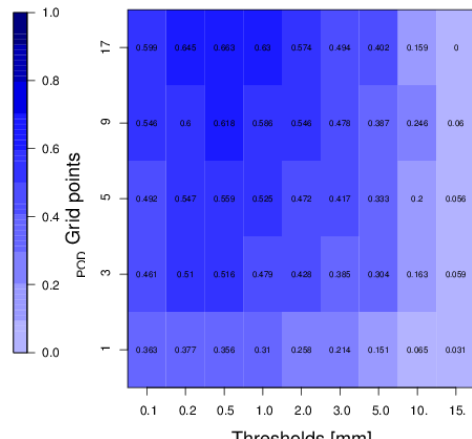
Anywhere in window Cosmo1 – POD – 200706_20–23



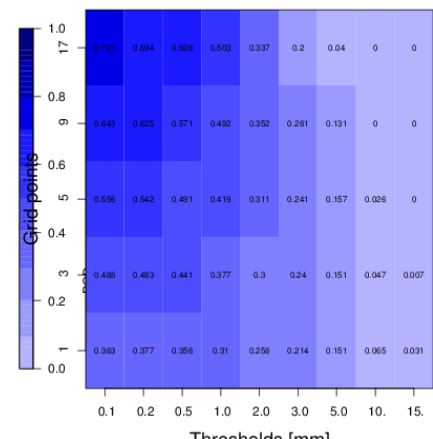
Joint probability Cosmo1 – POD – 200706_20–23



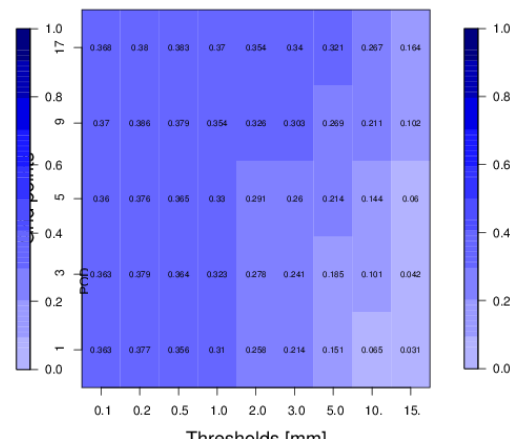
10% coverage Cosmo1 – POD – 200706_20–23



Upscaling Cosmo1 – POD – 200706_20–23



Fuzzy logic Cosmo1 – POD – 200706_20–23

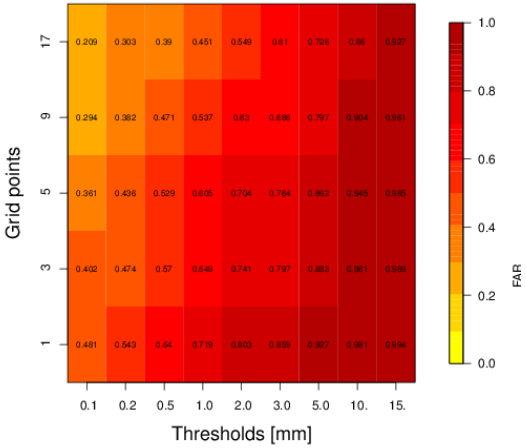


COSMO-1

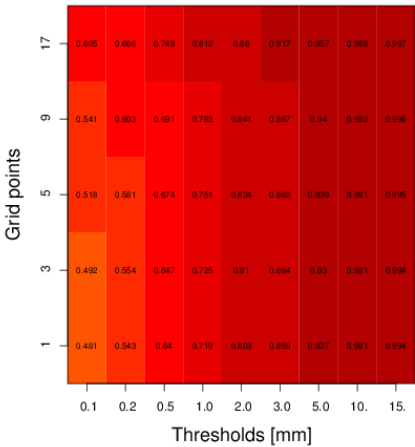
- FAR**

Neighborhood method	Decision Model
Upscaling	In a good forecast it is predicted the same average value of precipitation with the predicted
Fraction Skill Score	A forecast is useful when the frequency of forecasted events is similar with the frequency of the observed events
Minimum coverage	A forecast is useful if the event is forecasted in a minimum fraction of the area of interest
Anywhere in the window	A forecast is useful if it is more true than untrue, characterizing the events as probabilities
Pragmatic approach	A useful forecast has higher probability to discriminate the events and the non events.(in reference with the climatological value derived from the observations)
Area related RMSE	A useful forecast has the same distribution of intensity as the observations

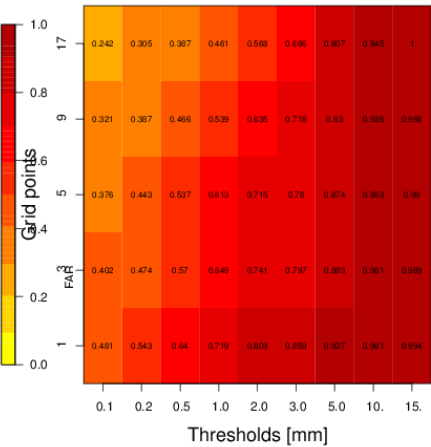
Anywhere in window Cosmo1 – FAR – 200706_20–23



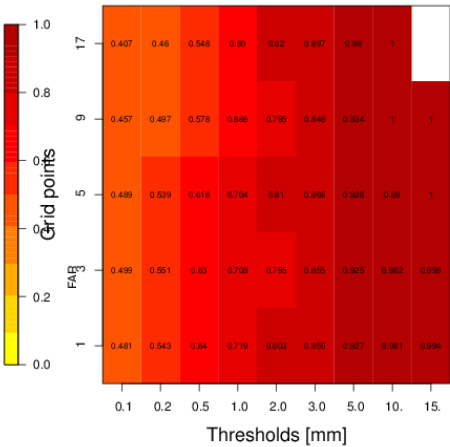
Joint probability Cosmo1 – FAR – 200706_20–23



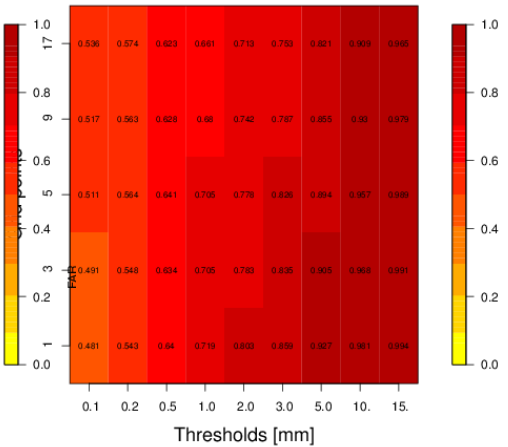
10% coverage Cosmo1 – FAR – 200706_20–23



Upscaling Cosmo1 – FAR – 200706_20–23

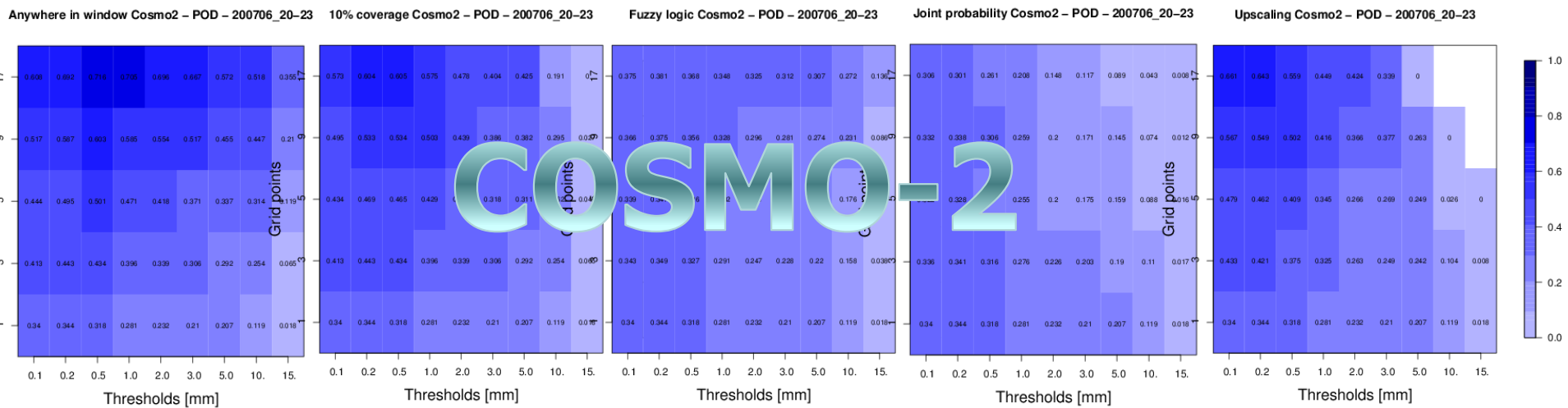
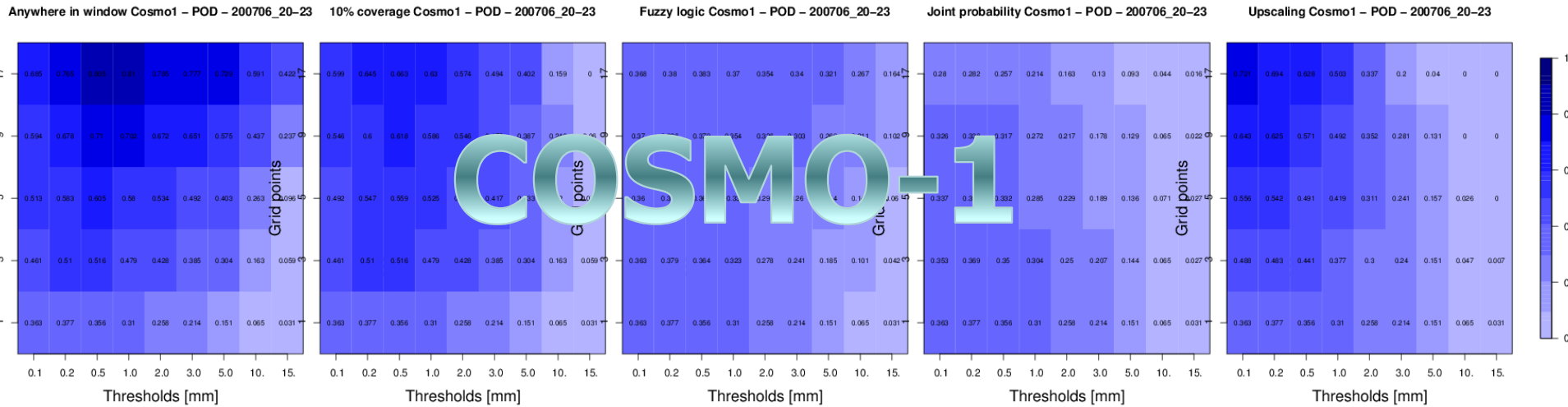


Fuzzy logic Cosmo1 – FAR – 200706_20–23



Neighborhood method scores COSMO-1 vs COSMO-2

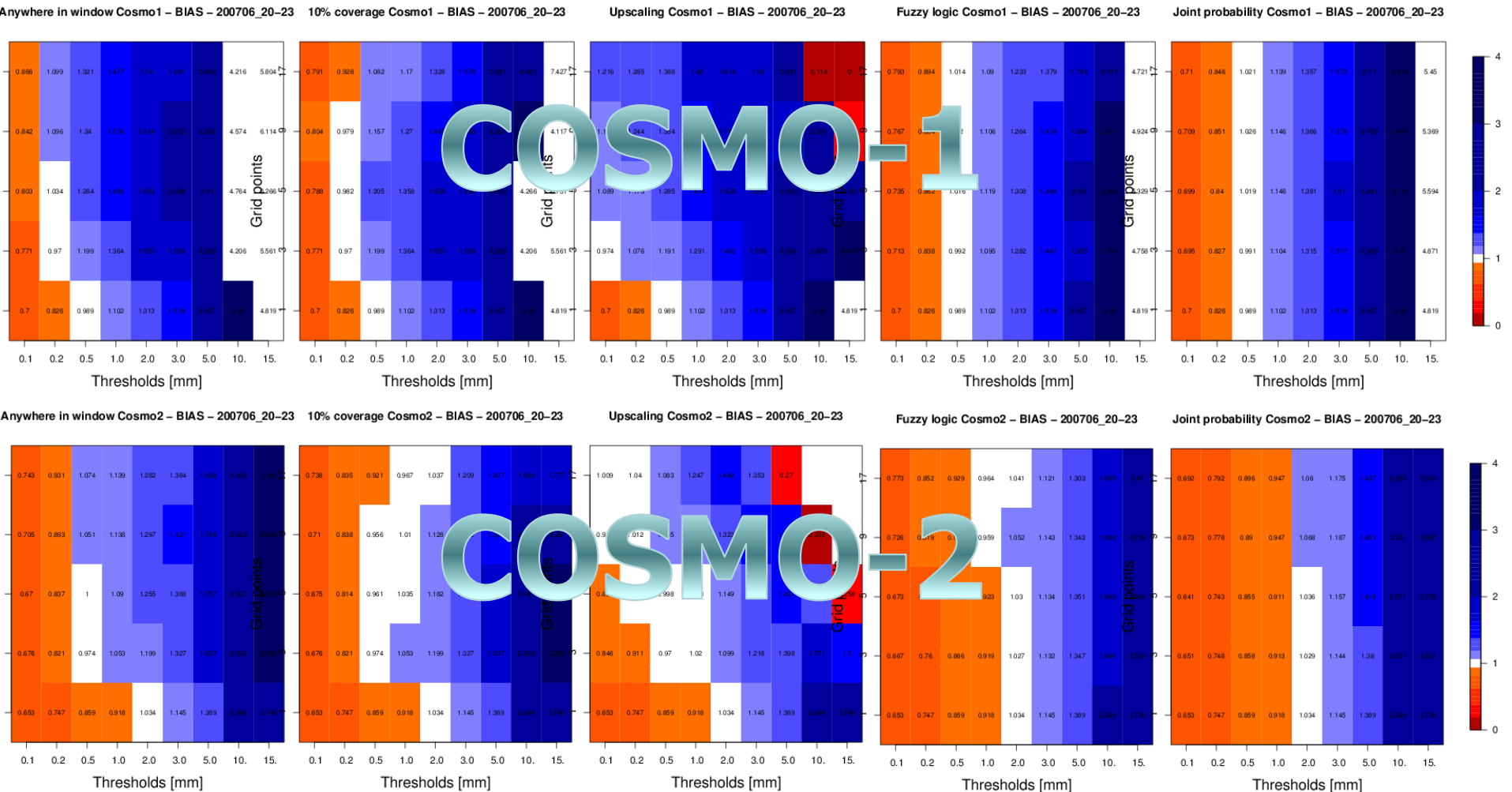
- POD**



Neighborhood method scores

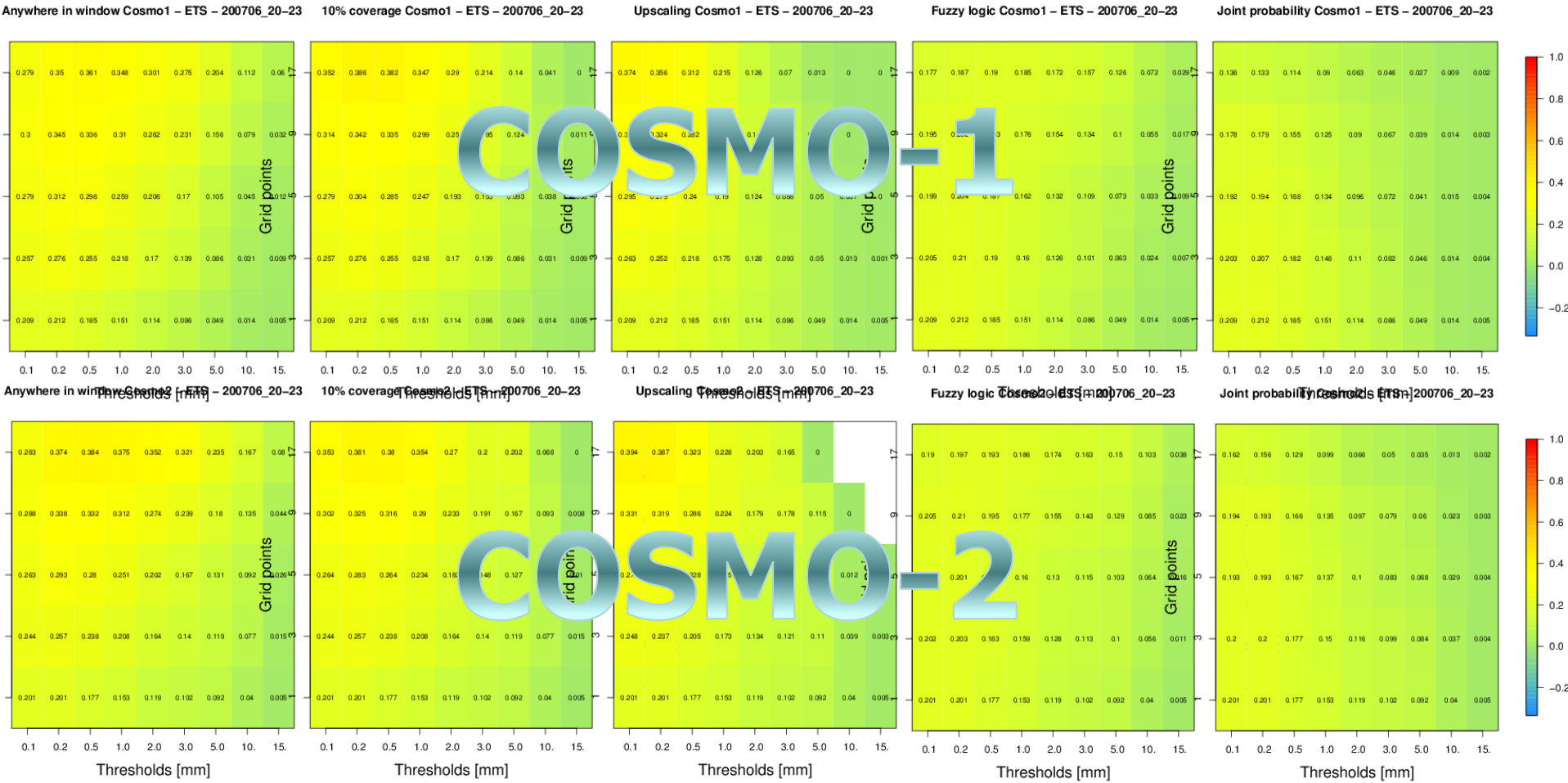
COSMO-1 vs. COSMO-2

- BIAS**



Neighborhood method scores COSMO-1 vs COSMO-2

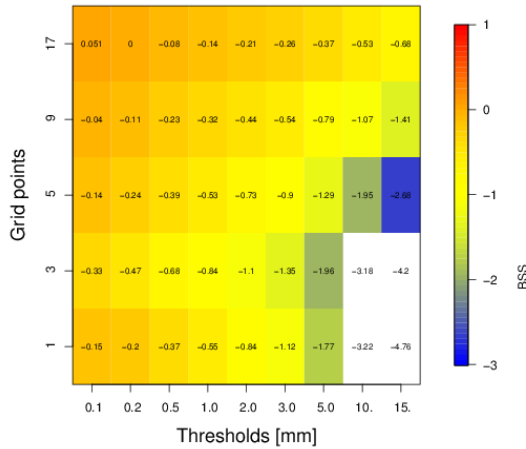
- ETS



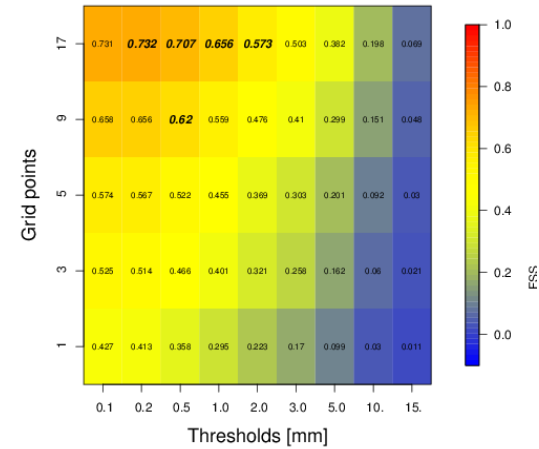
Neighborhood method scores

COSMO-1

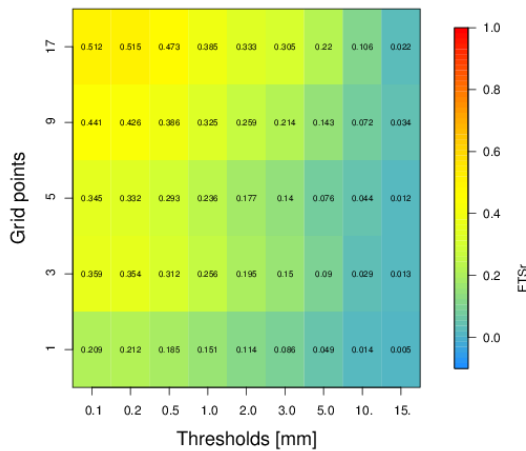
Pragmatic approach Cosmo1 – BSS – 200706_20-23



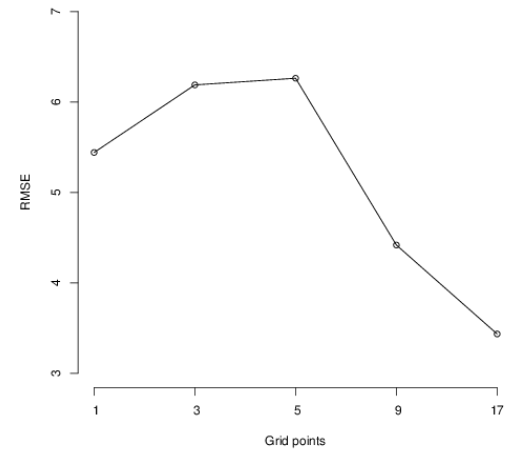
Fractions skill score Cosmo1 – FSS – 200706_20-23



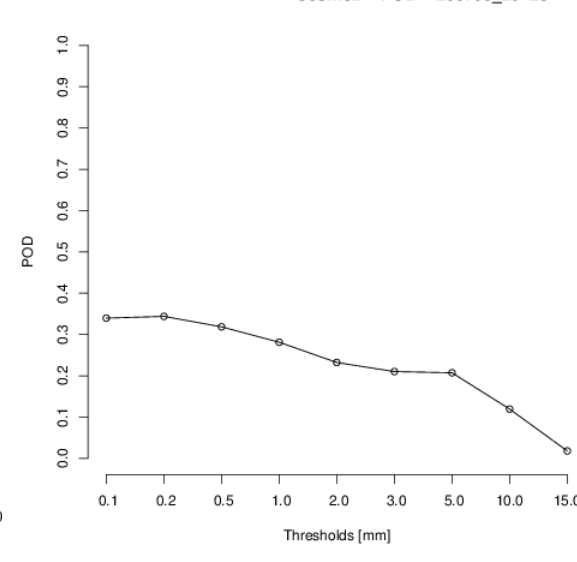
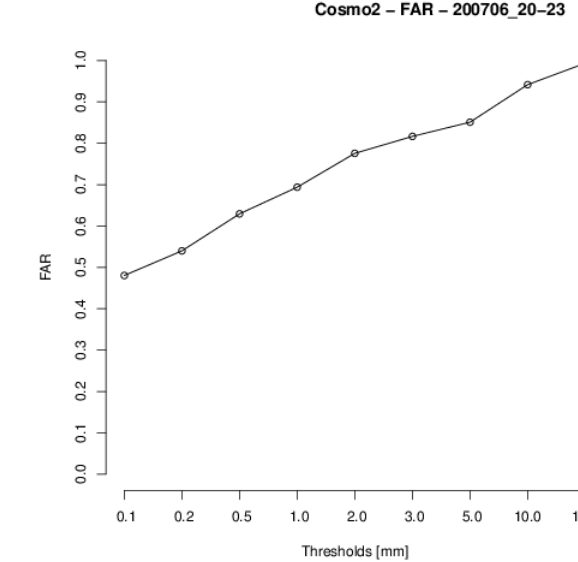
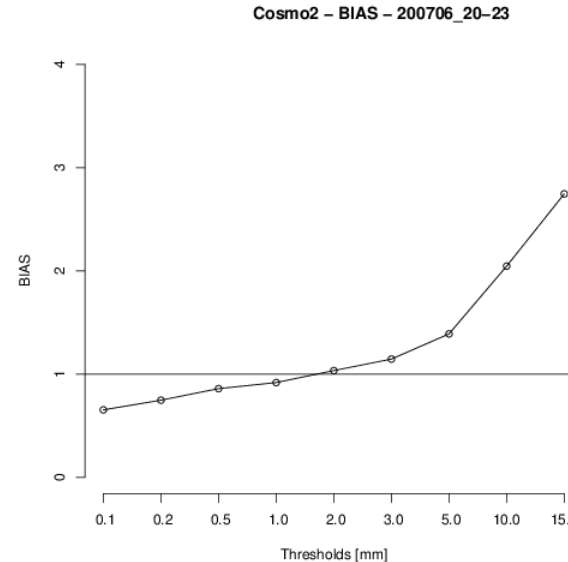
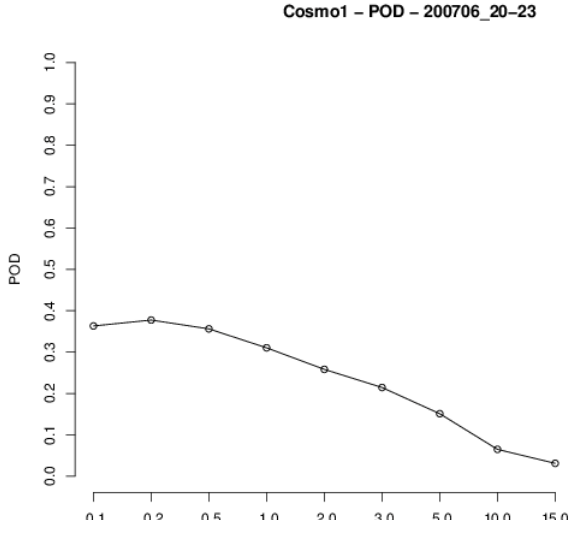
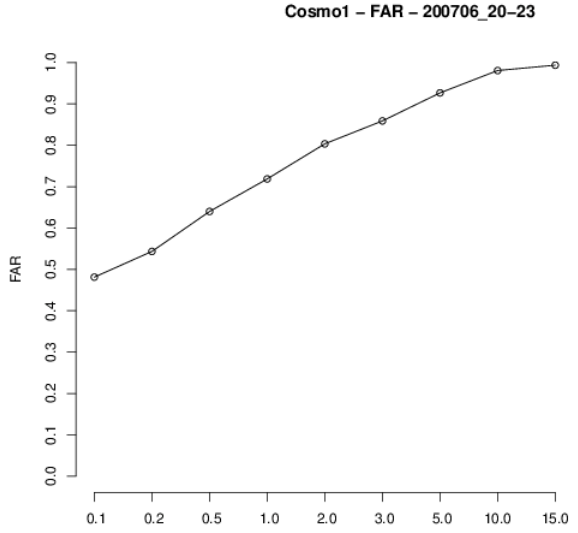
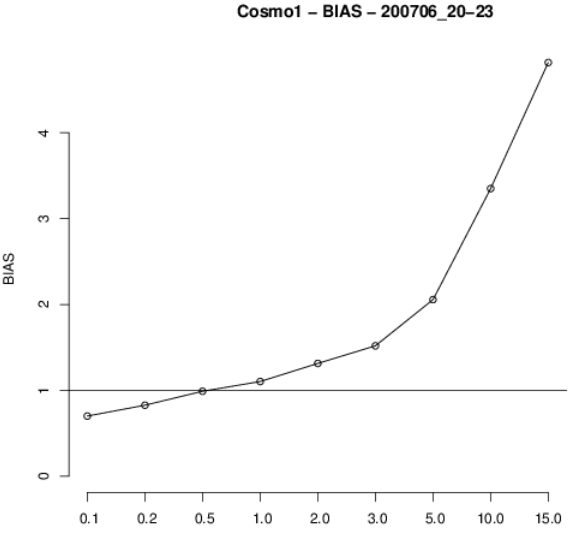
Prac. perf. hindcast Cosmo1 – ETSratio – 200706_20-23



Area related RMSE Cosmo1 – RMSE – 200706_20-23



Categorical scores: COSMO-1 vs. COSMO-2



Few thoughts/comments for filtering method applications

- Application of IS spatial method was attempted over a few instances related to the MesoVICT core case of 21-23/06/2007
- Main issue was the shift of the frontal system which was well described by the small (even negative) skill score in the pertinent scales
- Both models have negative skills at the smallest scale but skill improves when considering larger spatial scales.
- ISS decreases as the precipitation threshold is increased and this is due to the poor ability of the model to go beyond just the yes/no rain discrimination
- IS method constraints are related to the request to have precipitation analysis available for each grid point of the forecast field, and to the fact that Haar wavelet decomposition is designed for a square domain defined by a grid $2^l \times 2^l$



ΕΘΝΙΚΗ
ΜΕΤΕΩΡΟΛΟΓΙΚΗ
ΥΠΗΡΕΣΙΑ

HELLENIC NATIONAL METEOROLOGICAL SERVICE

Few thoughts/comments for filtering method applications

- IS kind of analysis and graphical representation of scores is not suitable for the operational verification as it is not concentrated in the average behavior of the model over areas but on single forecast. Neighborhood methods however with the right choice of decision model and aggregation on several timesteps/runs can provide a more operationally “useful” type of information
- While all methods measure intensity bias, no single method addresses all types of errors and so it is necessary to either prioritize which types of errors are most important to the user and choose the appropriate verification approach, or preferably apply more than one type of verification method.



ΕΘΝΙΚΗ
ΜΕΤΕΩΡΟΛΟΓΙΚΗ
ΥΠΗΡΕΣΙΑ
HELLENIC NATIONAL METEOROLOGICAL SERVICE

Ευχαριστώ πολύ



ΕΘΝΙΚΗ
ΜΕΤΕΩΡΟΛΟΓΙΚΗ
ΥΠΗΡΕΣΙΑ
HELLENIC NATIONAL METEOROLOGICAL SERVICE