First tests of QPF verification using SAL at AEMET

Jose A. Garcia-Moya. AEMET. Spain.
On behalf of Carlos Santos, Anna Ghelli (ECMWF) & Imanol Guerrero
ECMWF & AEMET collaboration

Hirlam-Aladin All Staff Meeting, Krakow.
April 13-16, 2010

Introduction

- Classical problem of how to show that higher resolution models are doing better than lower resolution ones.
- Even worse for mesoscale models which have much more grid points than observations.
- Then:
  - Focus on surface parameters. Mainly precipitation.
  - Looking for new scores giving more importance to what we see as more useful information but from an objective point of view.
- Object oriented methods may be the solution.
- We start using SAL to see what we can get from it (collaboration with ECMWF).

Limitations of classical verification methods

- Spatial scale models and observations
  - Interpolation methods
  - Correlation
  - E.g. patterns, structures
  - E.g. double penalty can give better scores to a coarser grid model
  - "New" methods: Up-scaling, Fuzzy, Feature/Object-oriented (e.g. SAL)

- **S**tructure**A**mplitude**L**ocation measure
  - Wernli et al, source code (fortran) provided by Marcus Paulat (DWD)
  - Collaboration ECMWF-AEMET

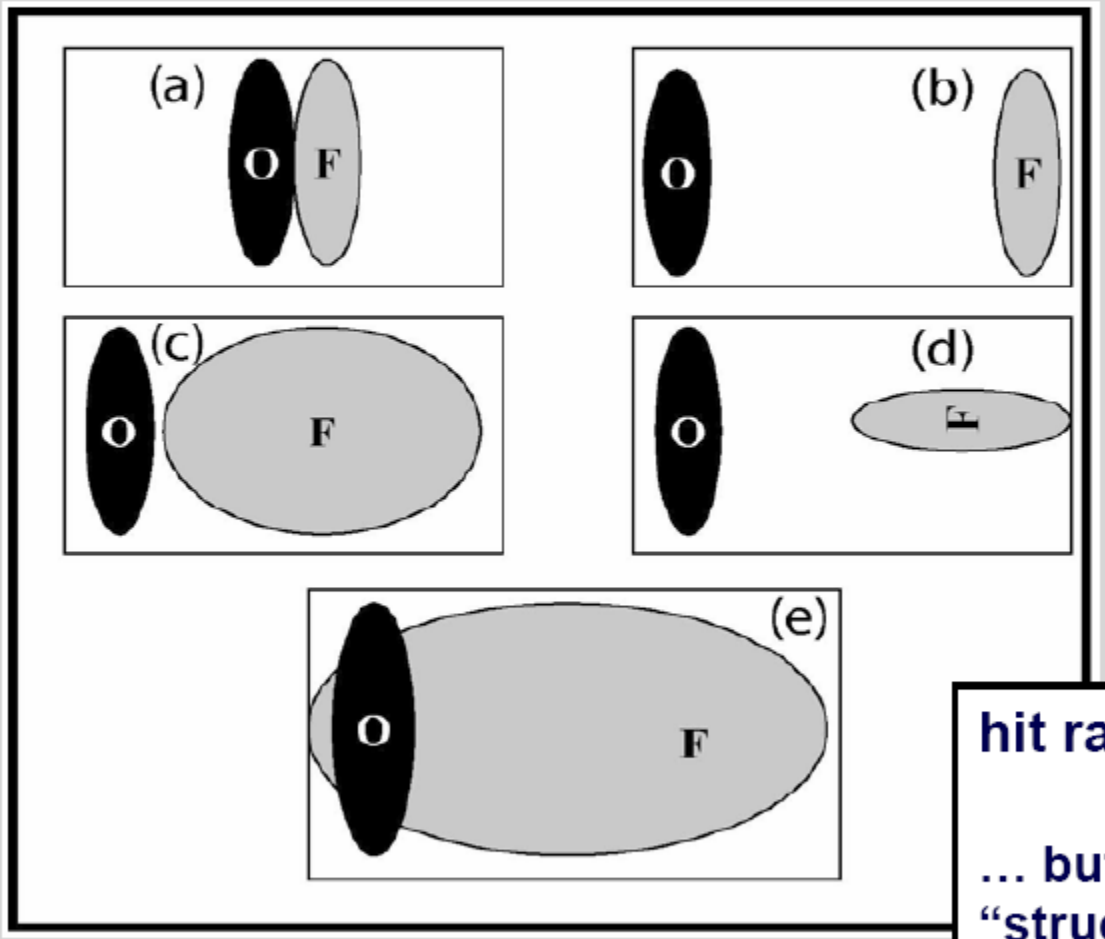
- Observational Uncertainty/Error:
  - Specific "new" methods: Saetra&Hersbach, Candille&Talagrand "Observational Probability"

- Sampling Uncertainty/Error:
  - Can lead to false conclusions
  - Specific methods: Confidence Intervals, Bootstrap

- Severe and extreme weather
  - Severe ≠ Adverse
  - A forecasting system can be useful on detecting signals even without good scores
  - Distributions-oriented verification, extreme events scores e.g. EDS

SAL

Problematic aspects of grid point based error scores



Davis et al. 2006 (MWR)

hit rate = 0

… but (a) is "better" than (b)-(d)

hit rate > 0

… but forecast has not the right "structure"

SAL

- Classical problem of double penalty
- Feature-oriented → ~ subjective verification
- E.g: SAL measure
  - S (Structure)
  - A (Amplitude)
  - L (Location)
- Perfect forecast: S = A = L = 0
- **S requires patterns/objects definition, currently simple algorithms, need improvement**

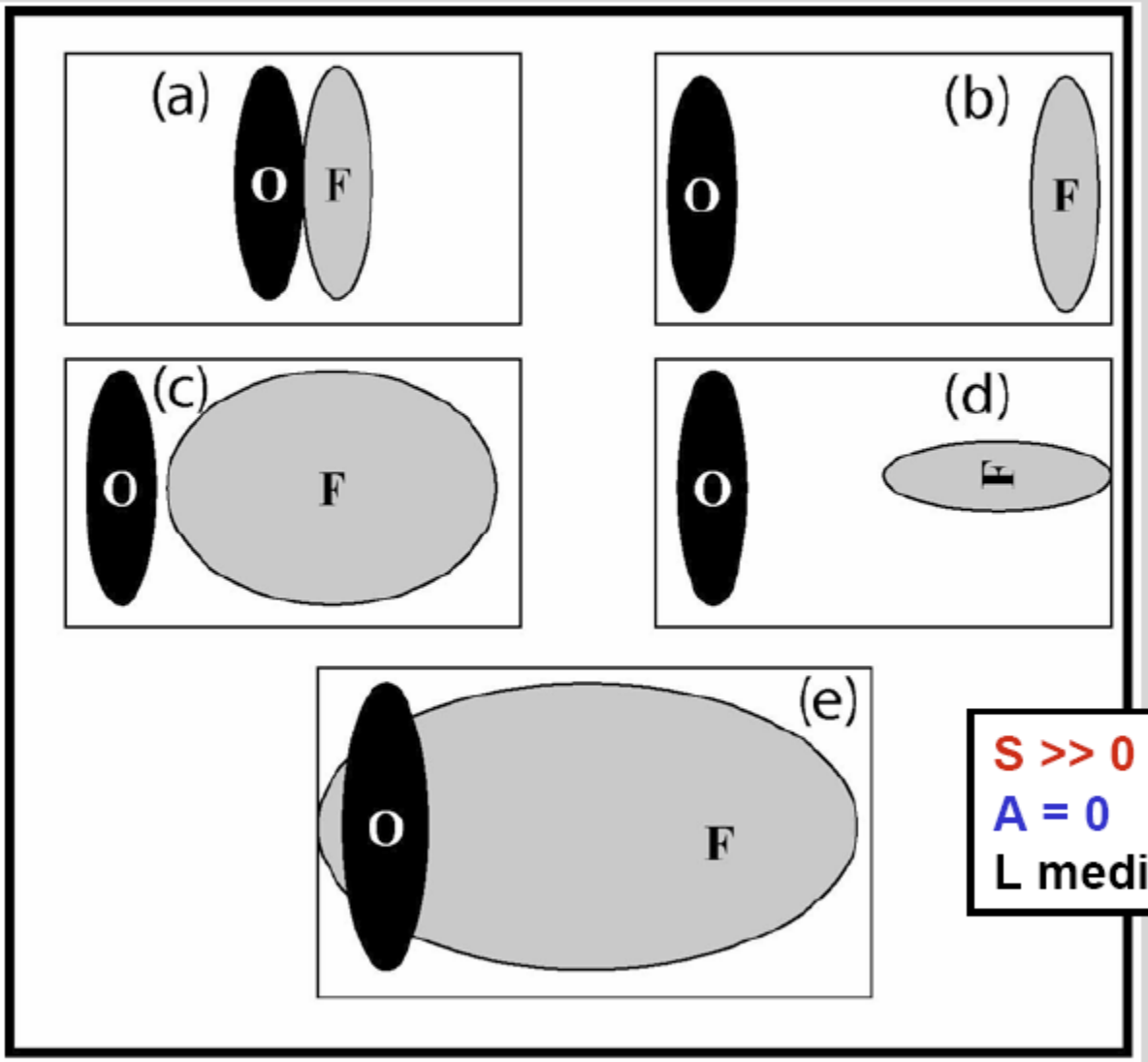
SAL

S:	Structure	-2	…	0	…	+2
		objects too small or too peaked		Perfect		objects too large or too flat
A:	Amplitude	-2	…	0	…	+2
		averaged QPF under-estimated		Perfect		averaged QPF over-estimated
L:	Location			0	…	+2
				Perfect		wrong location of Total Center of Mass (TCM) and / or of objects relative to TCM

SAL

S = 0
A = 0
L small

S > 0
A = 0
L medium



S = 0
A = 0
L large

S = 0
A = 0
L large

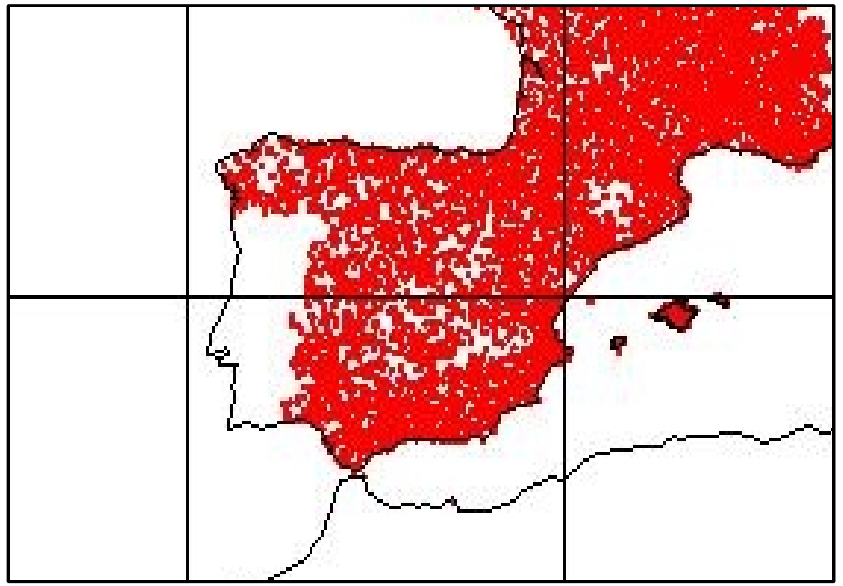
S >> 0
A = 0
L medium

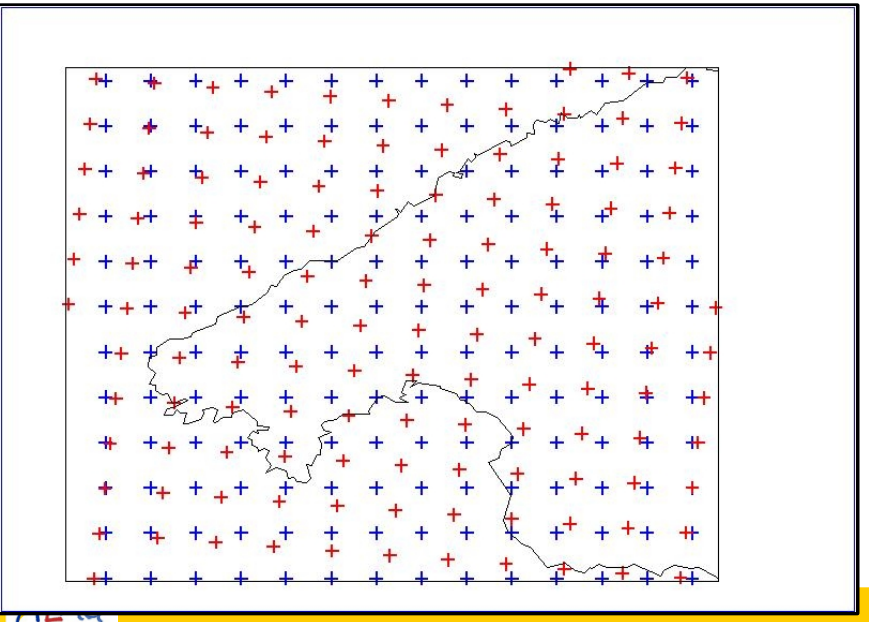
Davis et al. 2006

Work at AEMET

- SAL code adaptation (provided by Marcus Paulat, DWD)
- Up-scaling code implementation of two algorithms
  - Cell (with problems of missing data)
  - Structure functions r^{α}
- Research about models QPF SAL performance on one season
  - SON 2008
  - Iberian Peninsular
  - Up-scaling 3000 stations
- Research impact of:
  - Pcp threshold R* = f Rmax, over Spain f = 1/5
  - Model resolution: Hirlam_0.05, Hirlam_0.16, ECMWF T799
  - Model interpolation (original rotated to regular)

Model grid and Up-scaling grid

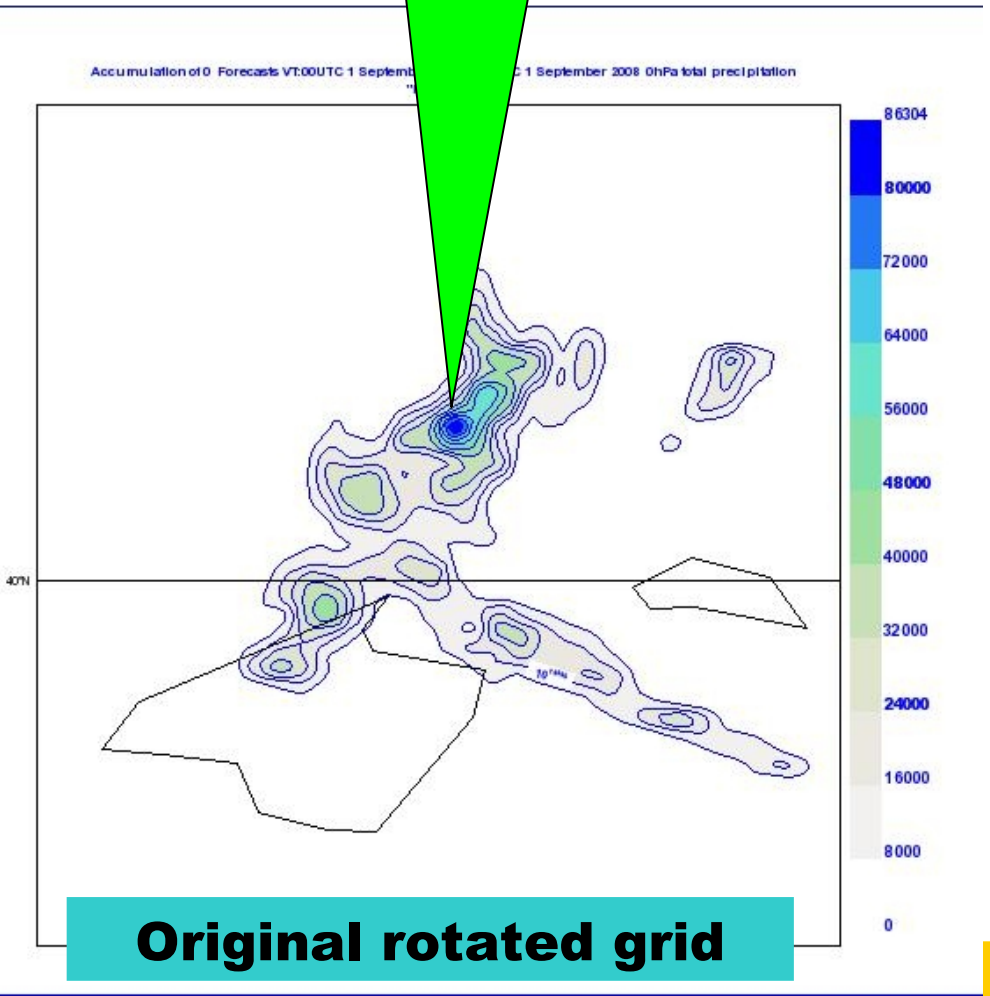




- What is the **truth**?
- Problems: model interpolation
  - Interpolation rotated → regular can smooth max model pcp
- Problems: verification on obs points
  - Spatial scale model-obs
  - Independent realizations?
- Up-scaling Europe/Spain HR data: a first simple approach:
  - #obs < 5 → grid box rejected
  - #obs >= 5 → grid box OK → avg (?)
  - Regular latlon interpolated model grid
  - Original rotated latlon model grid

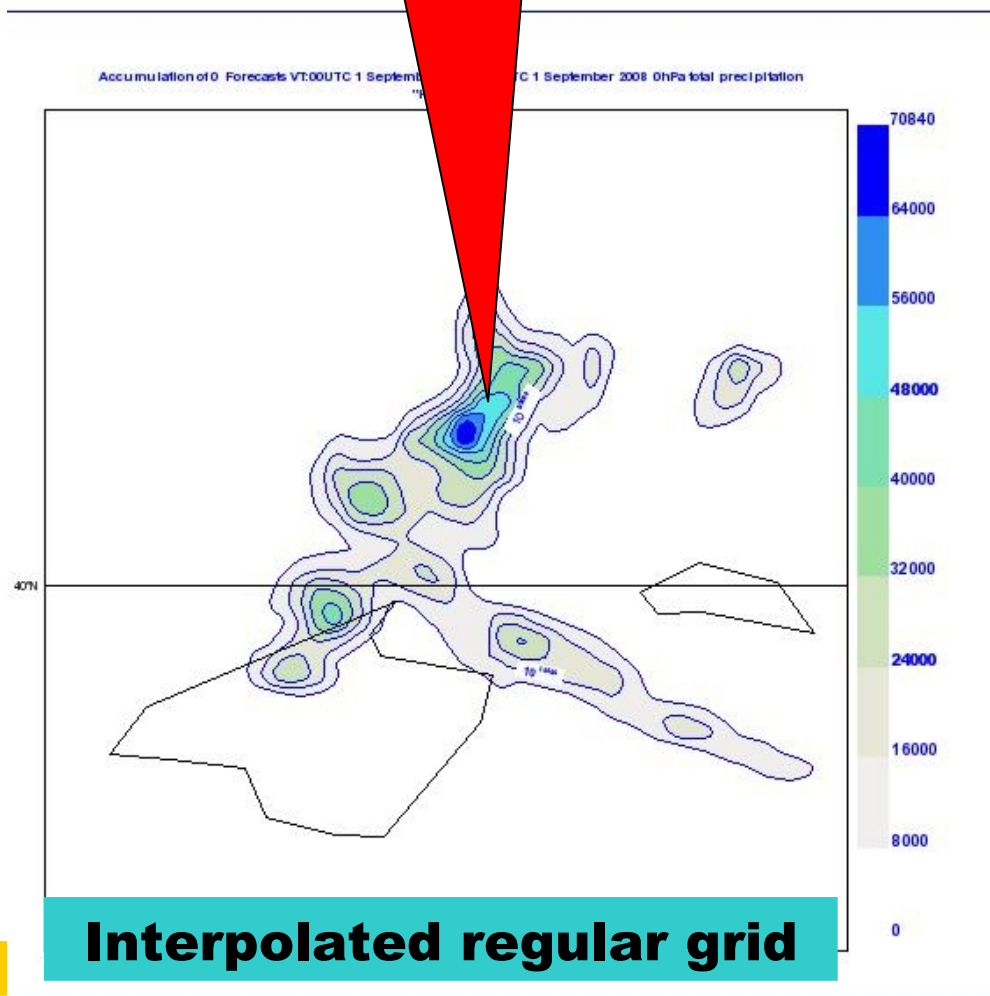
Smoother local max pcp

max pcp ~ 86mm

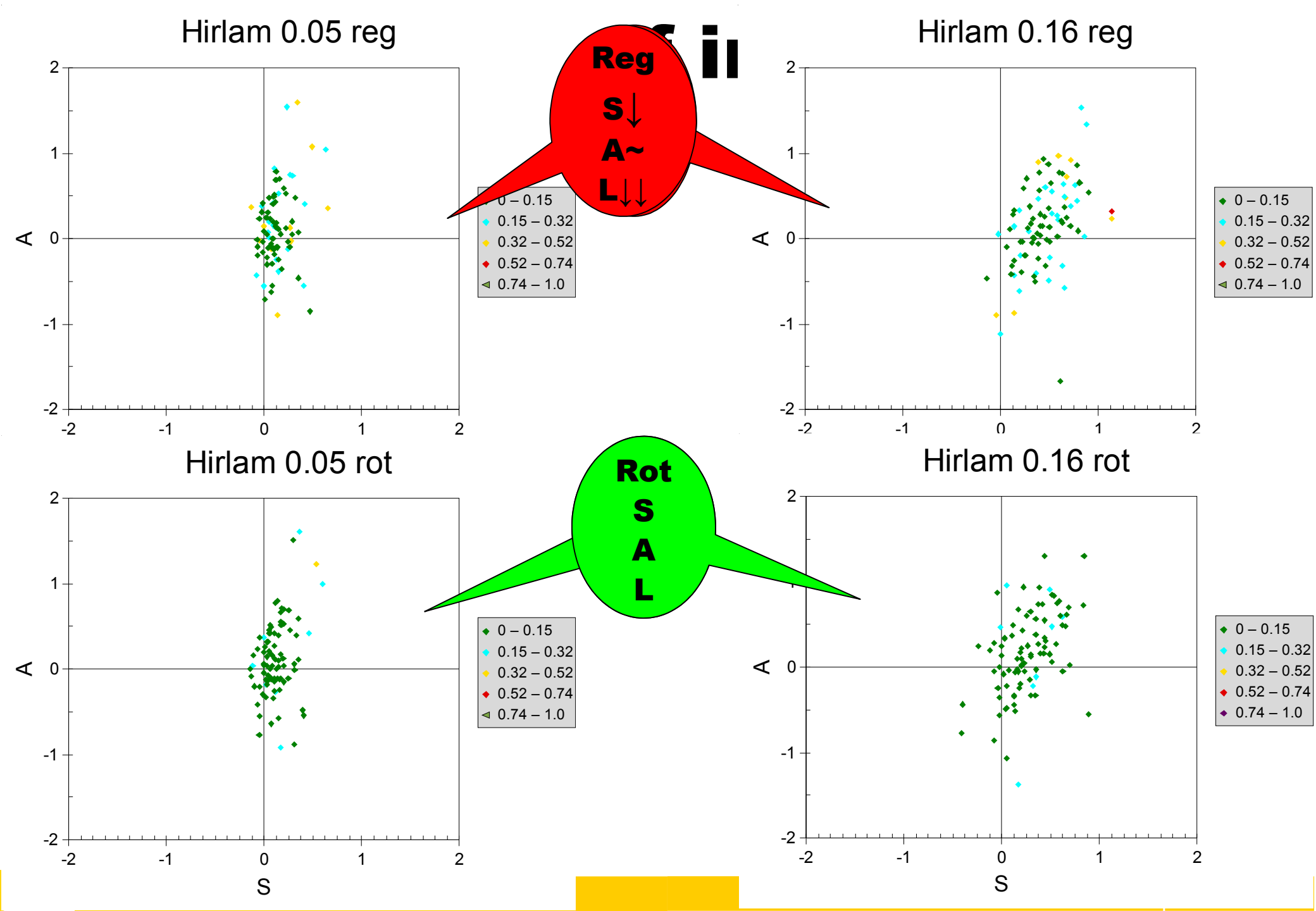


Original rotated grid

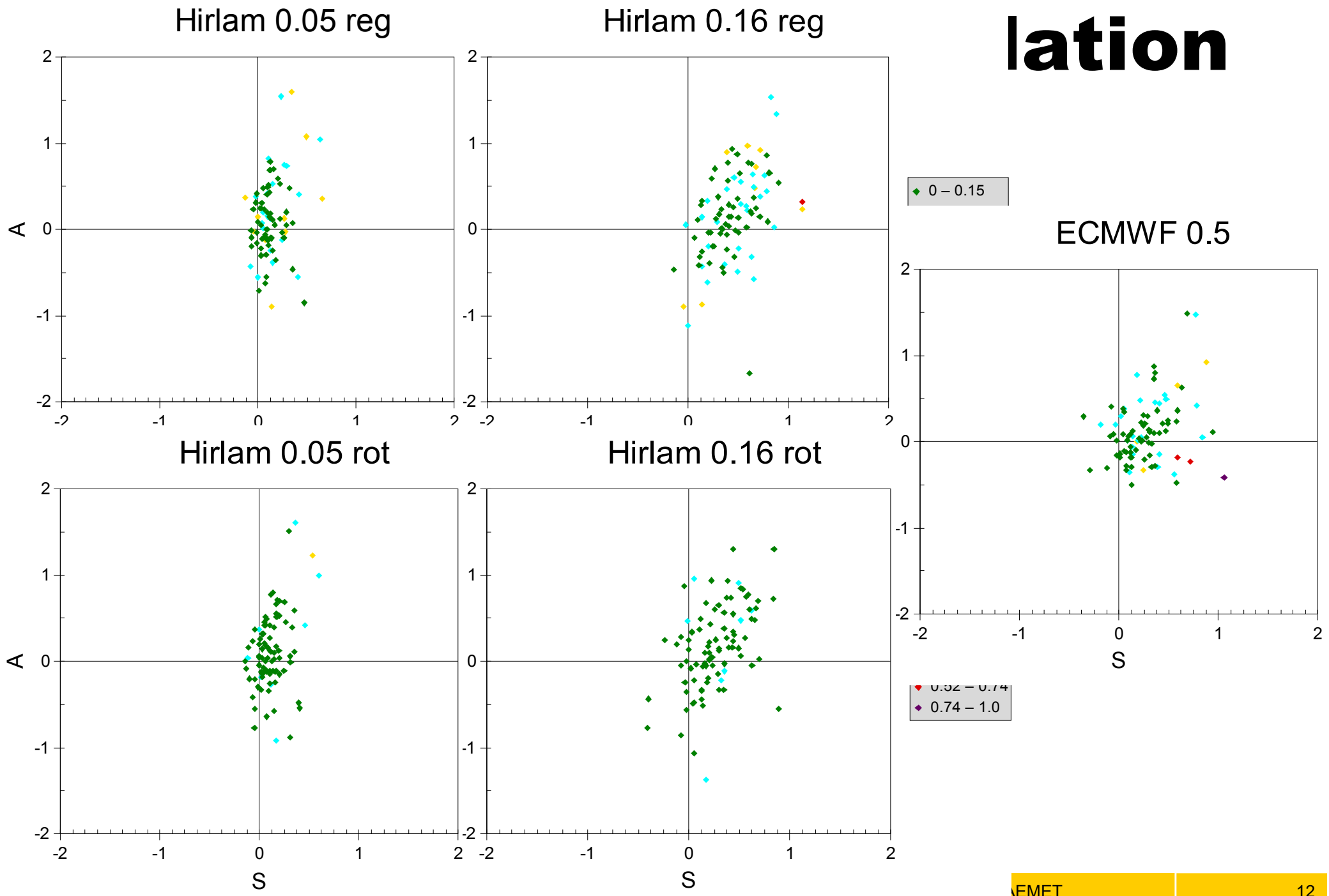
max pcp ~ 70mm
Smoother shapes



Interpolated regular grid

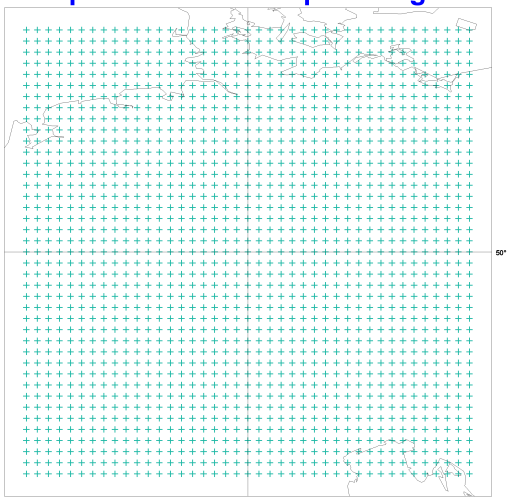

Hirlam 0.05 reg
Hirlam 0.16 reg
Hirlam 0.05 rot
Hirlam 0.16 rot
Reg
S↓
A~
L↓↓
Rot
S
A
L
A
S
0 – 0.15
0.15 – 0.32
0.32 – 0.52
0.52 – 0.74
0.74 – 1.0

lation


Hirlam 0.05 reg
Hirlam 0.16 reg
Hirlam 0.05 rot
Hirlam 0.16 rot
ECMWF 0.5
A
S
0 – 0.15
0.74 – 1.0

Work at ECMWF

- Collaboration framework ECMWF-AEMET
  - Anna Ghelli, Carlos Santos
  - Up-scaling & SAL code installed on linux cluster
- Research about models QPF SAL performance on one year
  - 2008
  - Central Europe (55N/5E/45N/15E)
  - Up-scaling 3000 stations
- Research impact of:
  - Pcp threshold R* = f Rmax, f = 1/15, stratification on 1.0mm pivot
  - Model resolution: T799, T399 (cf)
  - Forecast step: D+2, D+5

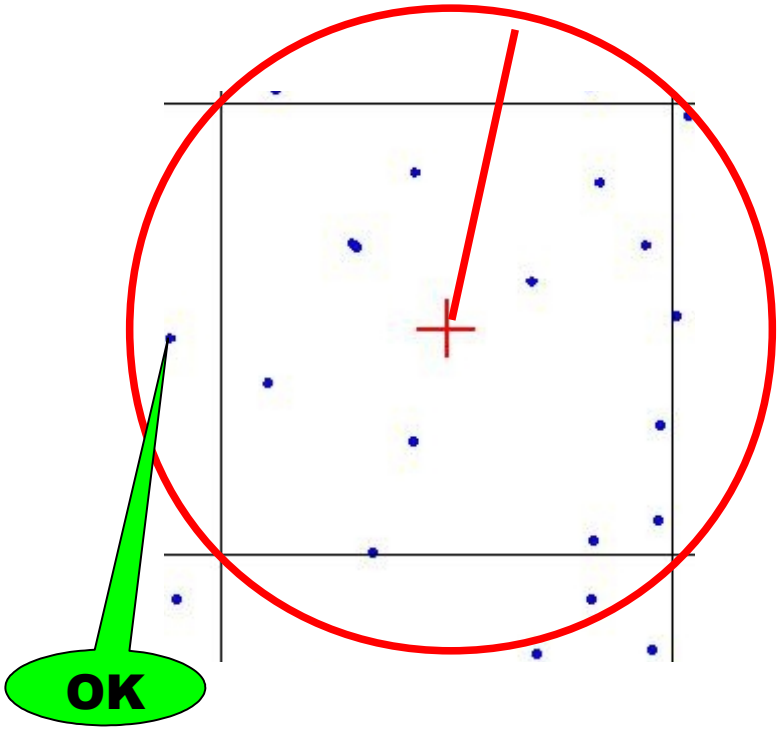


Up-scaling

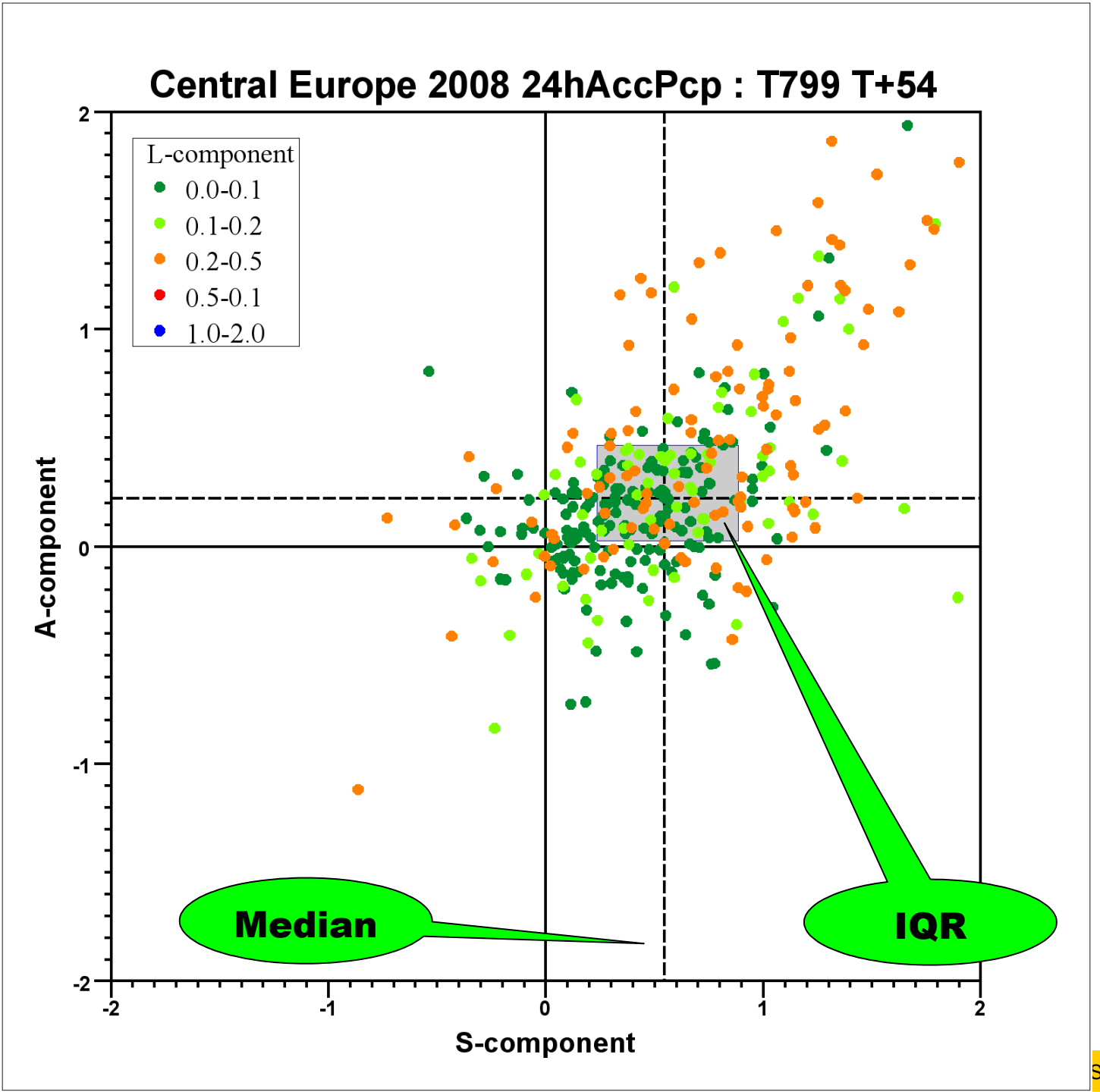
What is the **truth**?

Up-scaling Europe HR obs available at ECMWF: a first simple approach:

- For each grid point consider d
- obs r < d → ob considered
- $R = \sum r^{-\alpha} R_i / \sum r^{-\alpha}$ with e.g. α=2
- Overcome missing data at most resolutions



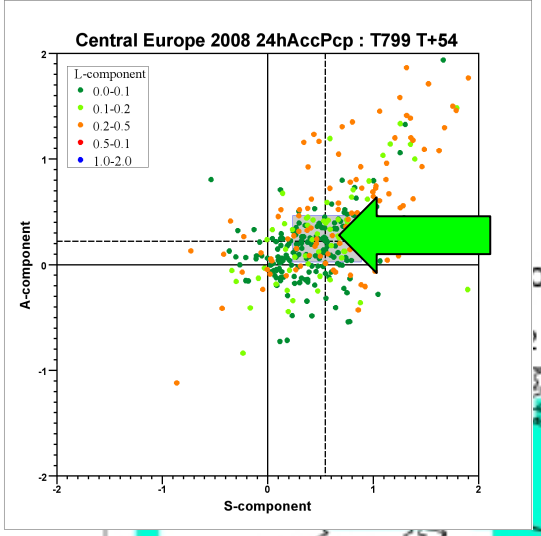
- In this work
  - Each model is compared with its own "natural up-scaling"
  - T799 with up-scaling 0.25
  - T399 with up-scaling 0.50

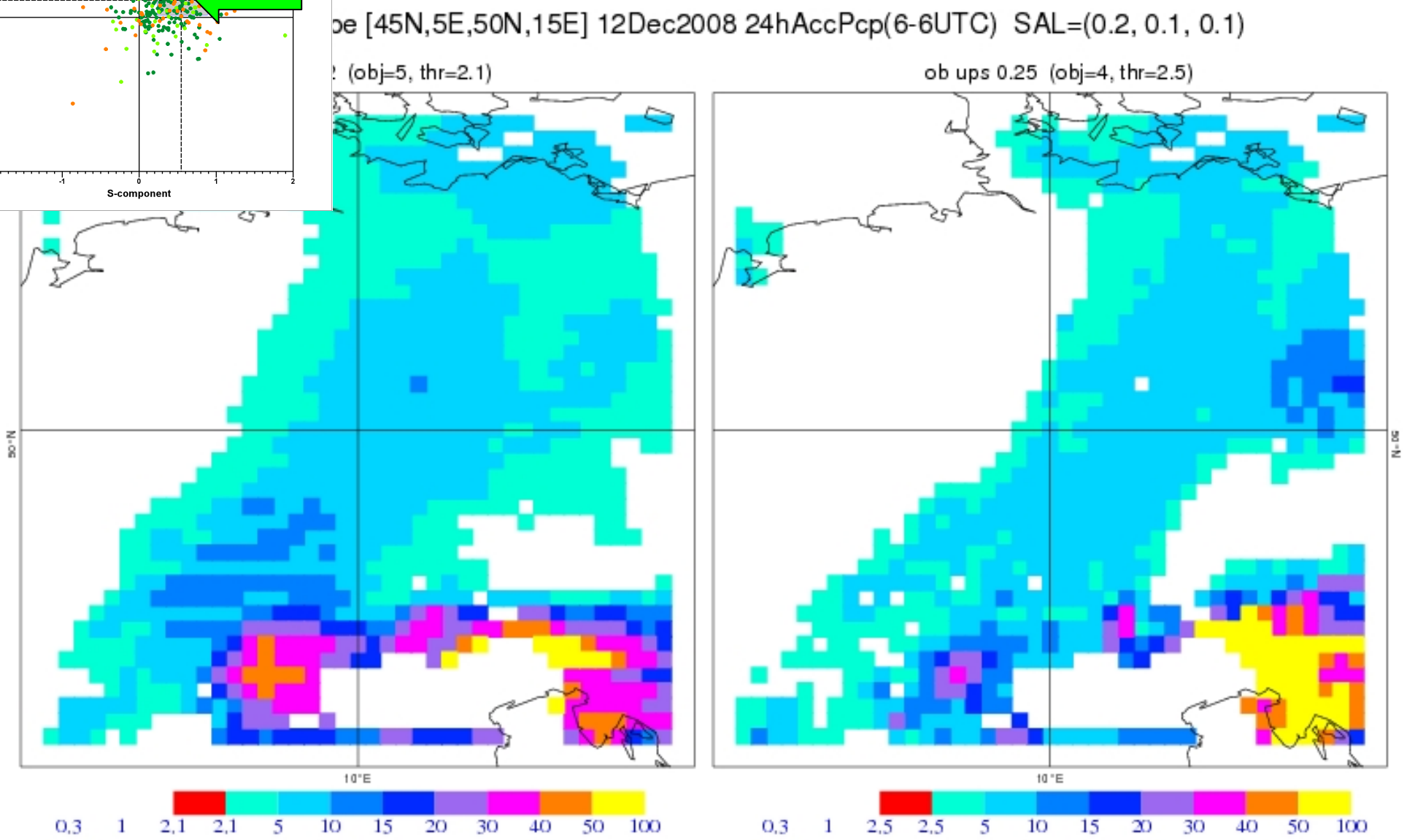


The SAL plot:

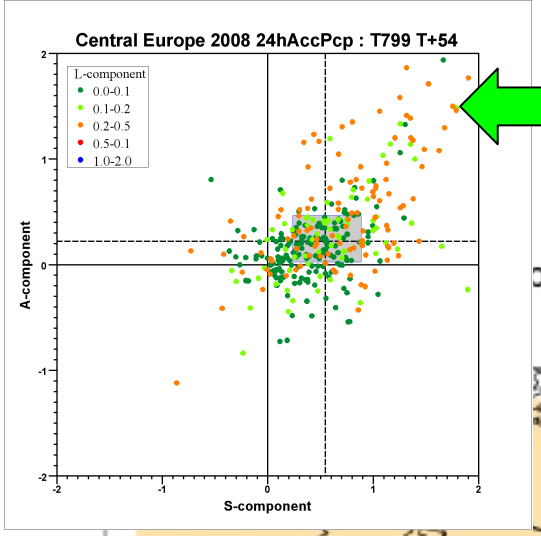
T799 D+2

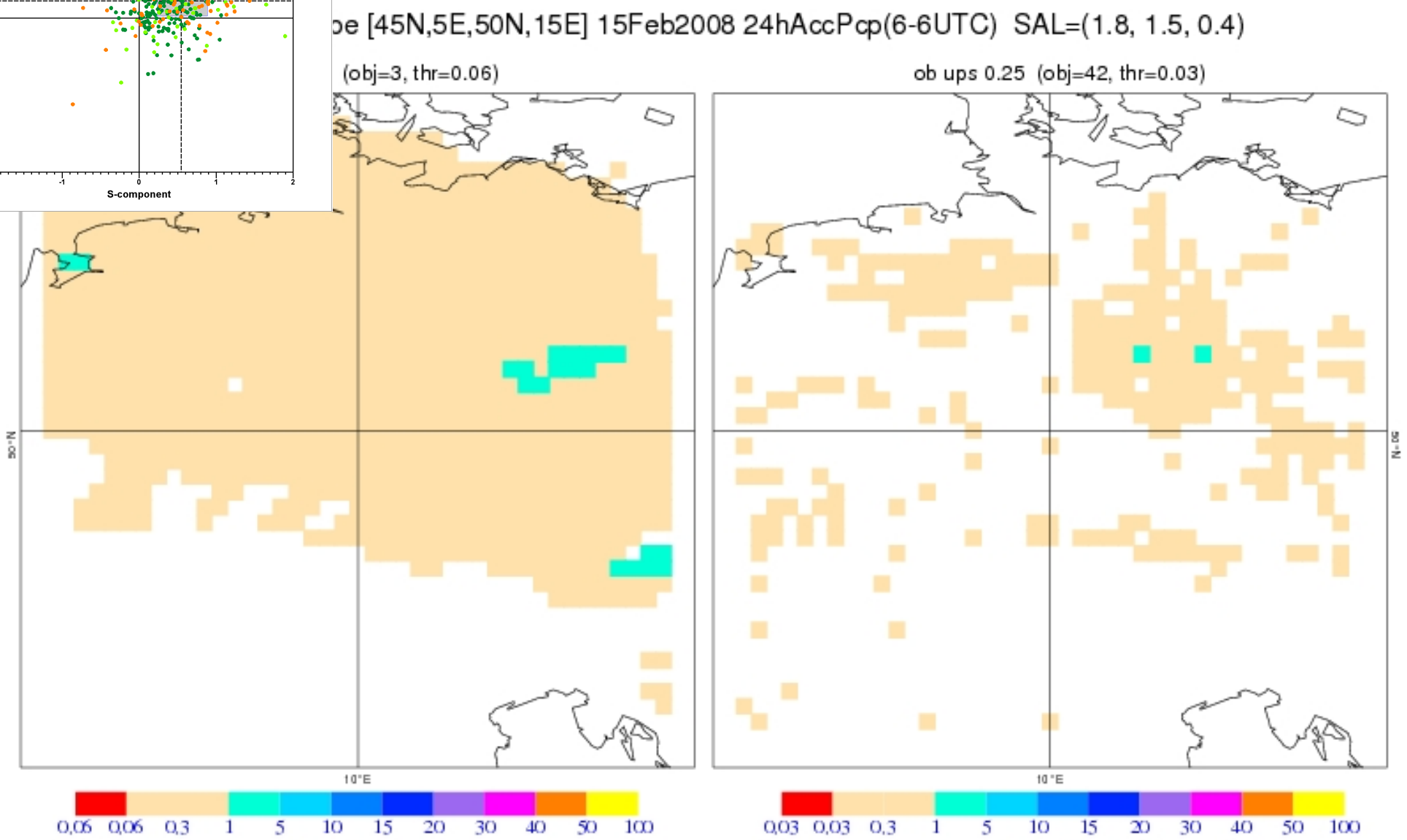
Inside IQR


Central Europe 2008 24hAccPcp : T799 T+54
L-component
0.0-0.1
0.1-0.2
0.2-0.5
0.5-0.1
1.0-2.0
A-component
S-component

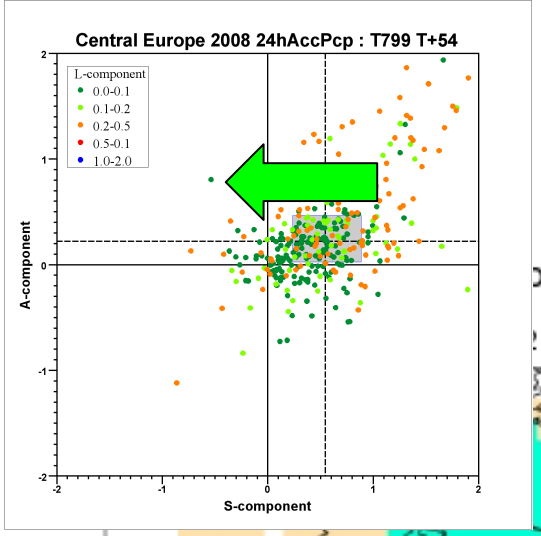

be [45N,5E,50N,15E] 12Dec2008 24hAccPcp(6-6UTC) SAL=(0.2, 0.1, 0.1)
(obj=5, thr=2.1)
ob ups 0.25 (obj=4, thr=2.5)
10°E
10°E
0.3 1 2.1 2.1 5 10 15 20 30 40 50 100
0.3 1 2.5 2.5 5 10 15 20 30 40 50 100

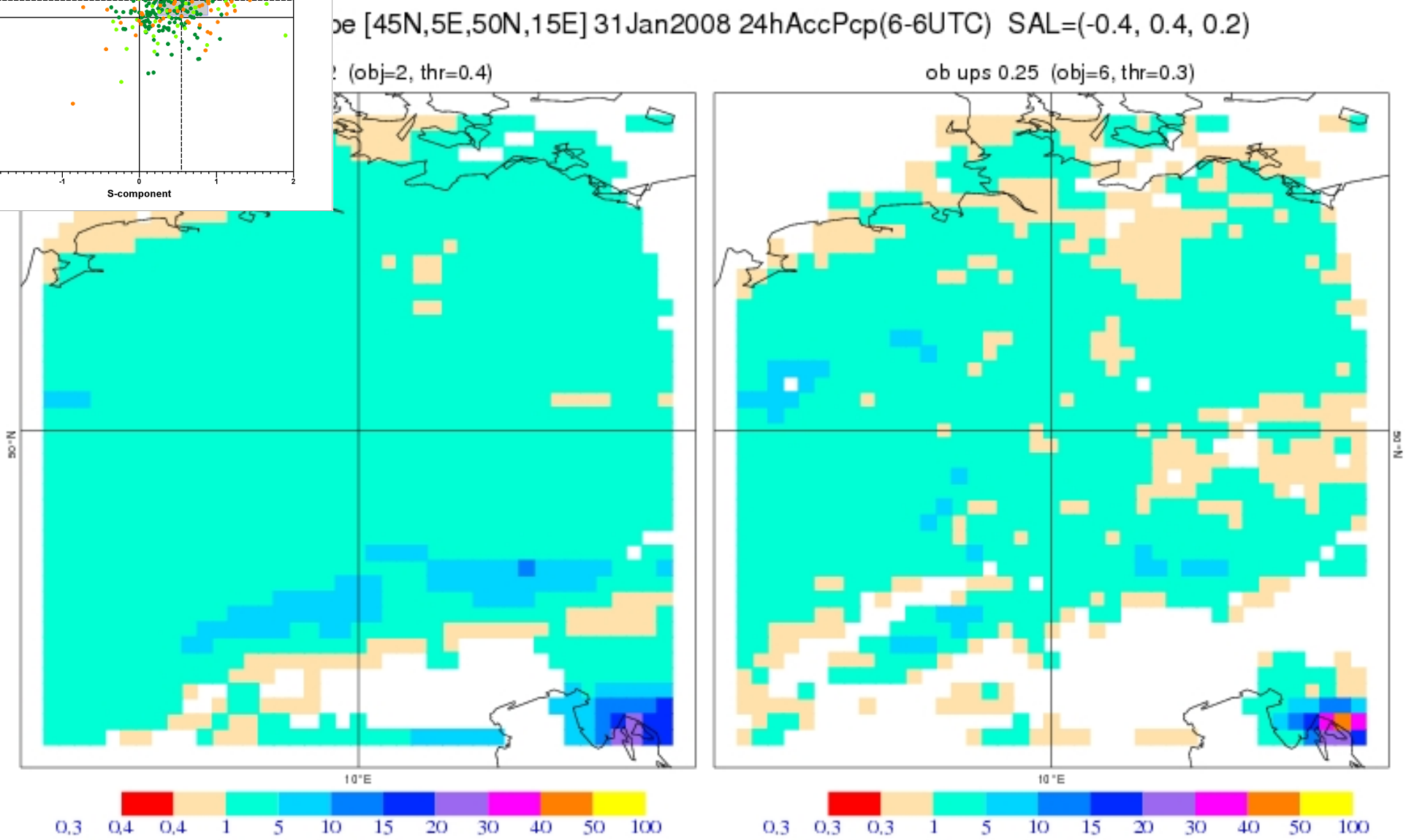
Tail I


Central Europe 2008 24hAccPcp : T799 T+54
L-component
0.0-0.1
0.1-0.2
0.2-0.5
0.5-0.1
1.0-2.0
A-component
S-component

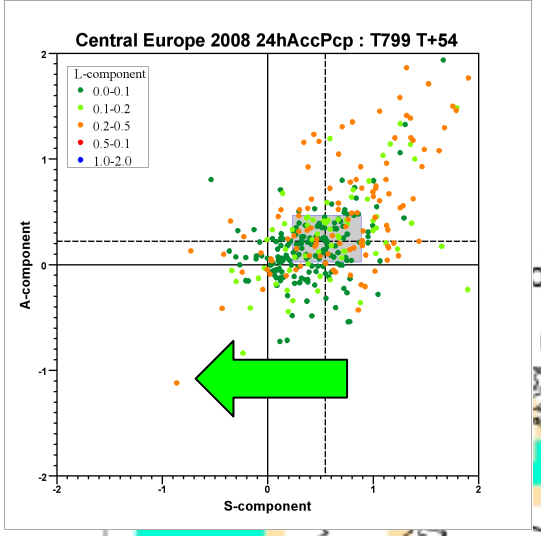

be [45N,5E,50N,15E] 15Feb2008 24hAccPcp(6-6UTC) SAL=(1.8, 1.5, 0.4)
(obj=3, thr=0.06)
ob ups 0.25 (obj=42, thr=0.03)
10°E
10°E
0,06 0,06 0,3 1 5 10 15 20 30 40 50 100
0,03 0,03 0,3 1 5 10 15 20 30 40 50 100

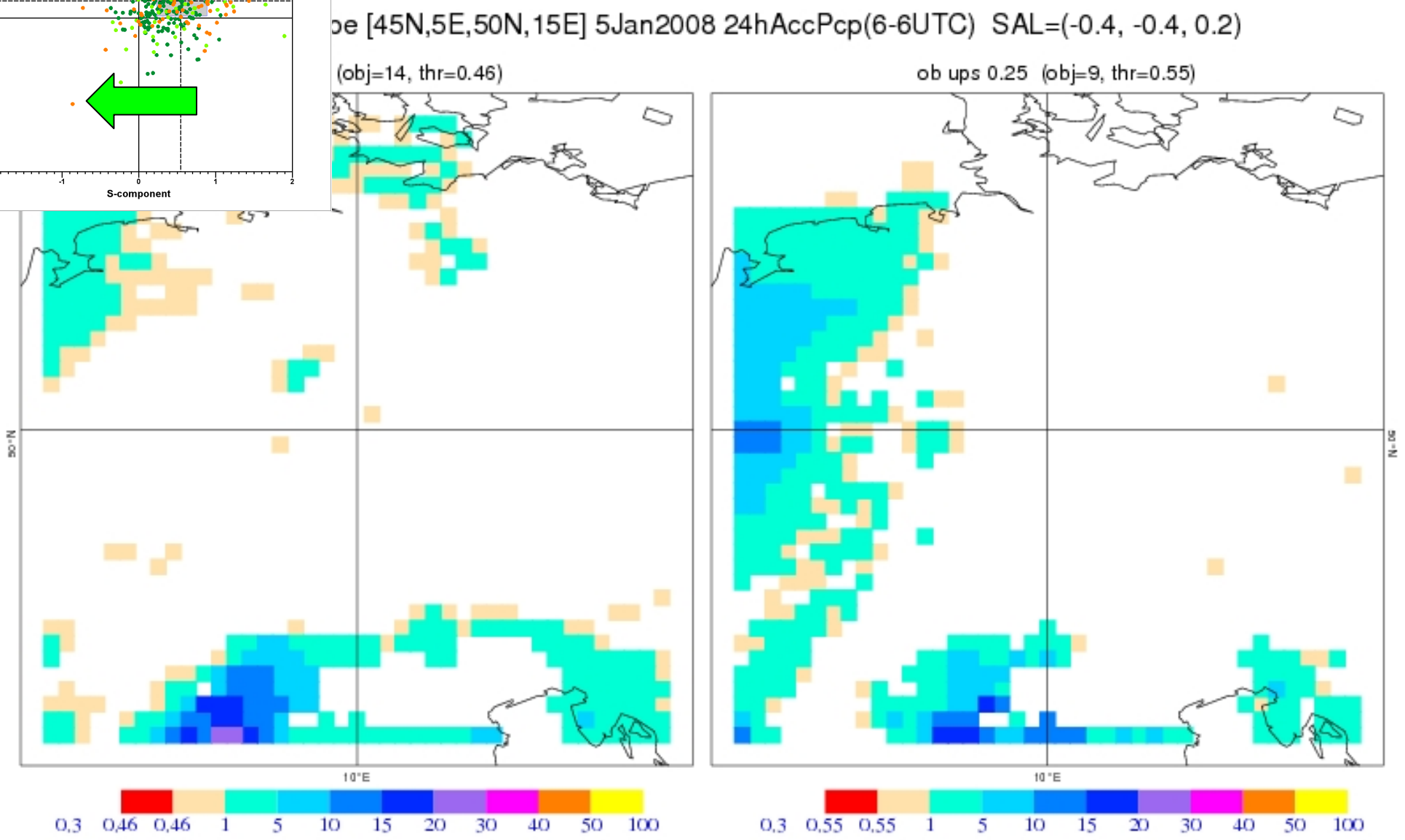
Tail II


Central Europe 2008 24hAccPcp : T799 T+54
L-component
0.0-0.1
0.1-0.2
0.2-0.5
0.5-0.1
1.0-2.0
A-component
S-component

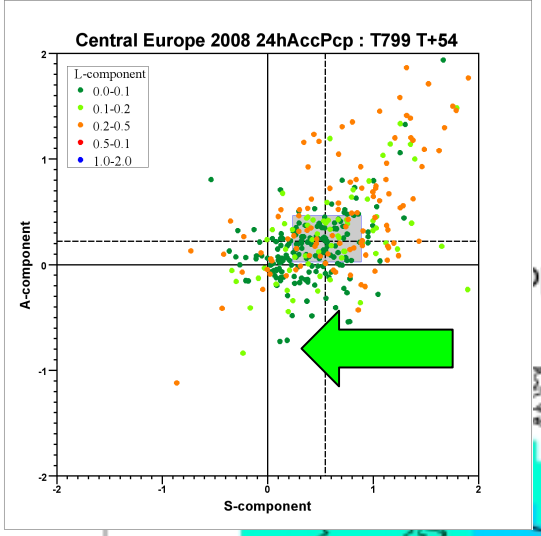

be [45N,5E,50N,15E] 31Jan2008 24hAccPcp(6-6UTC) SAL=(-0.4, 0.4, 0.2)
(obj=2, thr=0.4)
ob ups 0.25 (obj=6, thr=0.3)
10°E
10°E
0.3 0.4 0.4 1 5 10 15 20 30 40 50 100
0.3 0.3 0.3 1 5 10 15 20 30 40 50 100

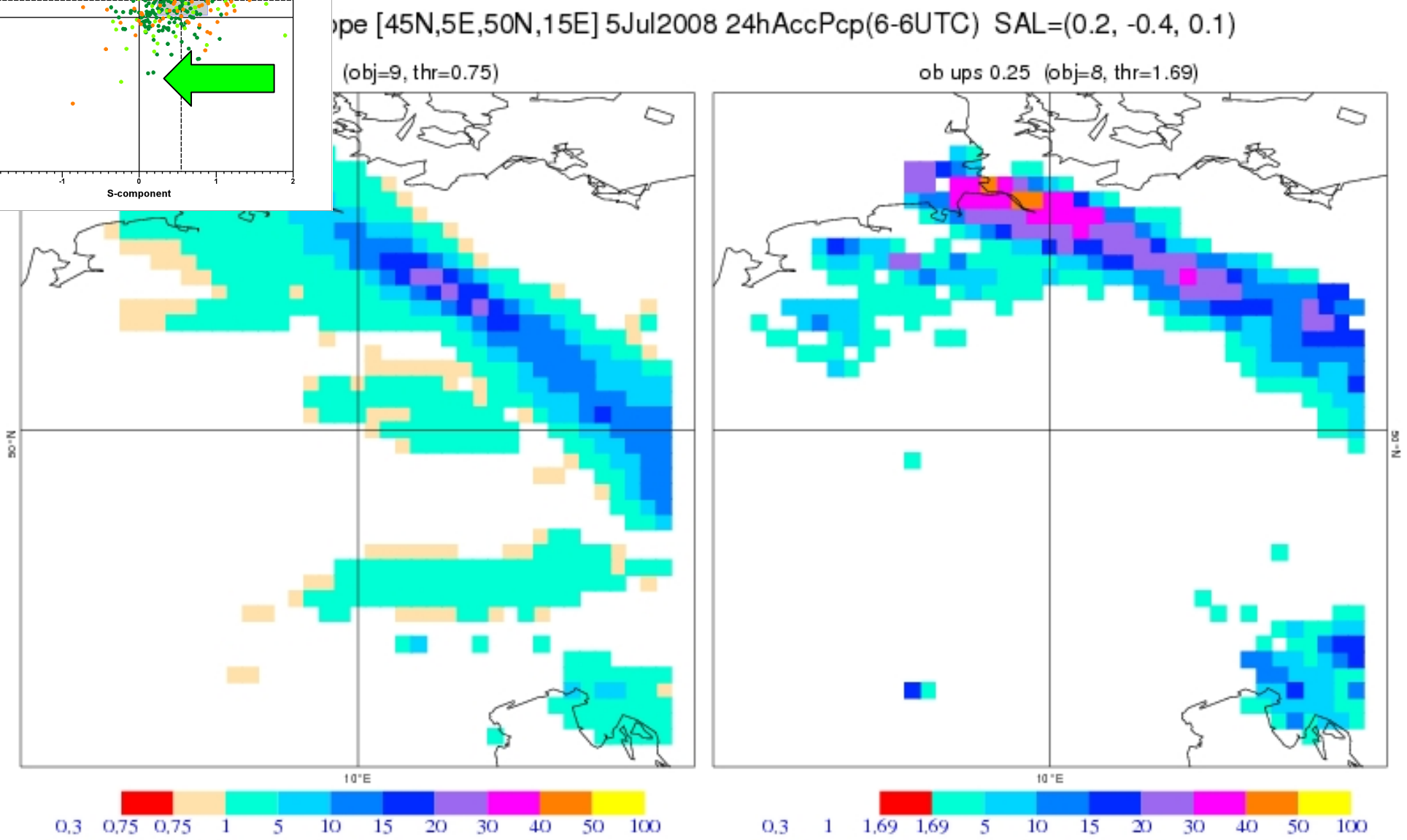
Tail III


Central Europe 2008 24hAccPcp : T799 T+54
L-component
0.0-0.1
0.1-0.2
0.2-0.5
0.5-0.1
1.0-2.0
A-component
S-component

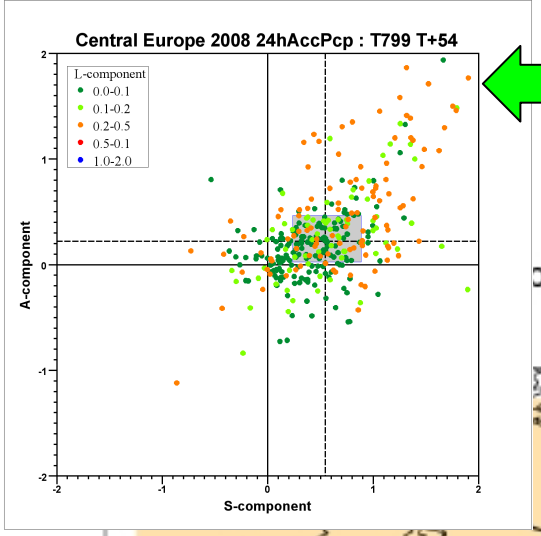

be [45N,5E,50N,15E] 5Jan2008 24hAccPcp(6-6UTC) SAL=(-0.4, -0.4, 0.2)
(obj=14, thr=0.46)
ob ups 0.25 (obj=9, thr=0.55)
10°E
10°E
0.3 0.46 0.46 1 5 10 15 20 30 40 50 100
0.3 0.55 0.55 1 5 10 15 20 30 40 50 100

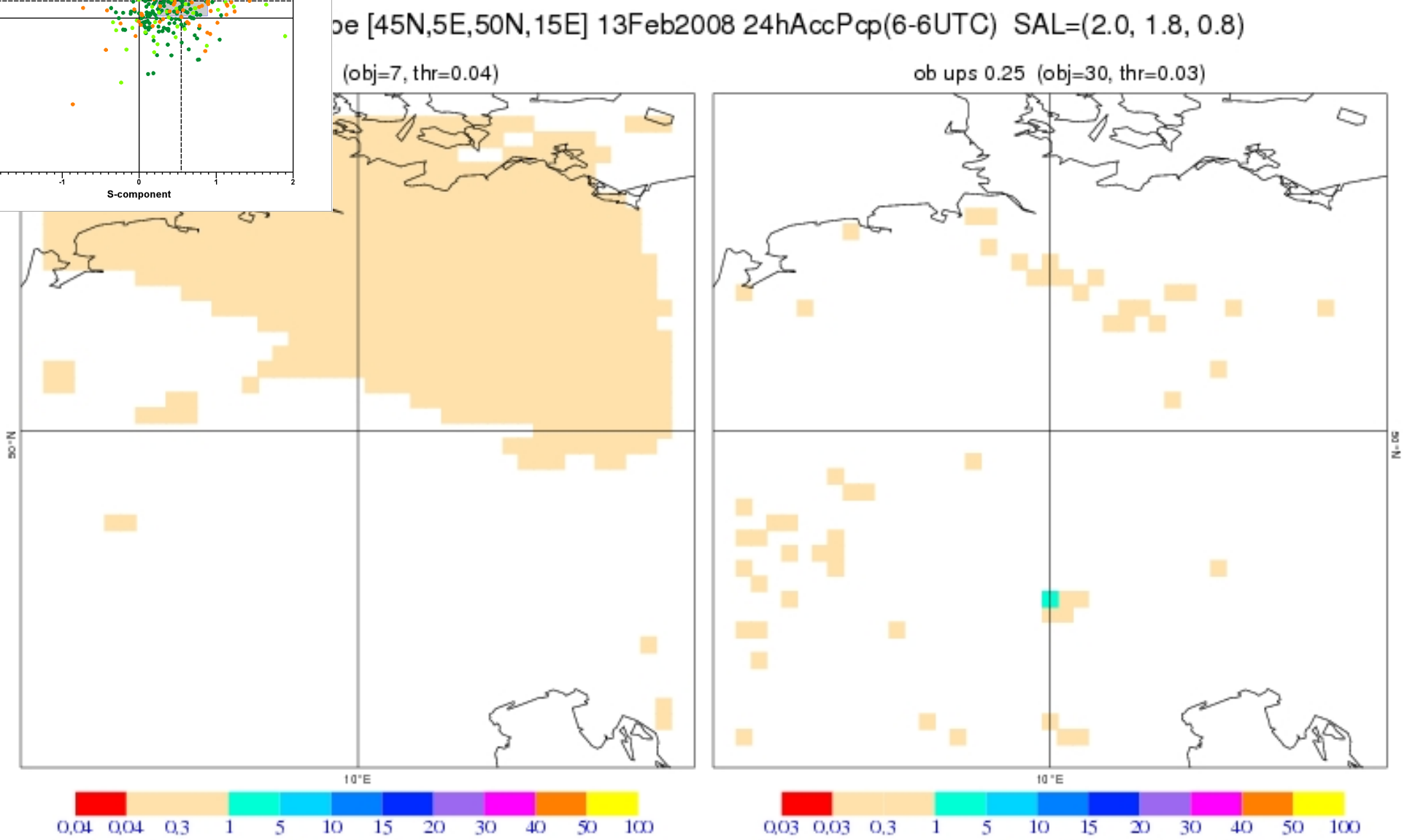
Tail IV


Central Europe 2008 24hAccPcp : T799 T+54
L-component
0.0-0.1
0.1-0.2
0.2-0.5
0.5-0.1
1.0-2.0
A-component
S-component

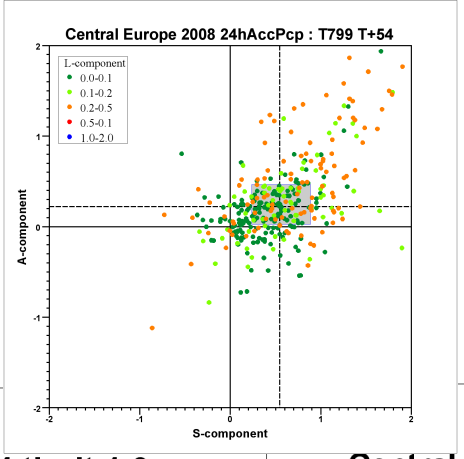

pe [45N,5E,50N,15E] 5Jul2008 24hAccPcp(6-6UTC) SAL=(0.2, -0.4, 0.1)
(obj=9, thr=0.75)
ob ups 0.25 (obj=8, thr=1.69)
10°E
10°E
0,3 0,75 0,75 1 5 10 15 20 30 40 50 100
0,3 1 1,69 1,69 5 10 15 20 30 40 50 100

Tail L

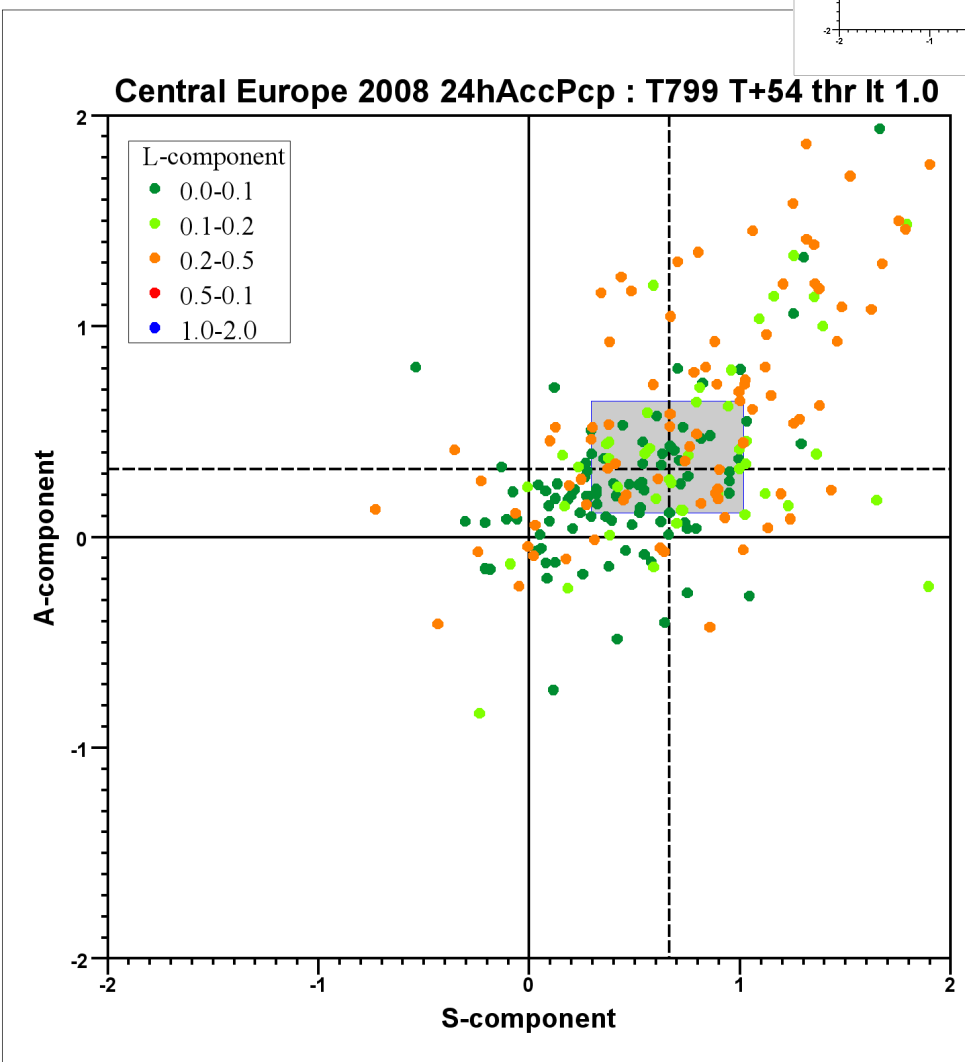

Central Europe 2008 24hAccPcp : T799 T+54
L-component
0.0-0.1
0.1-0.2
0.2-0.5
0.5-0.1
1.0-2.0
A-component
S-component

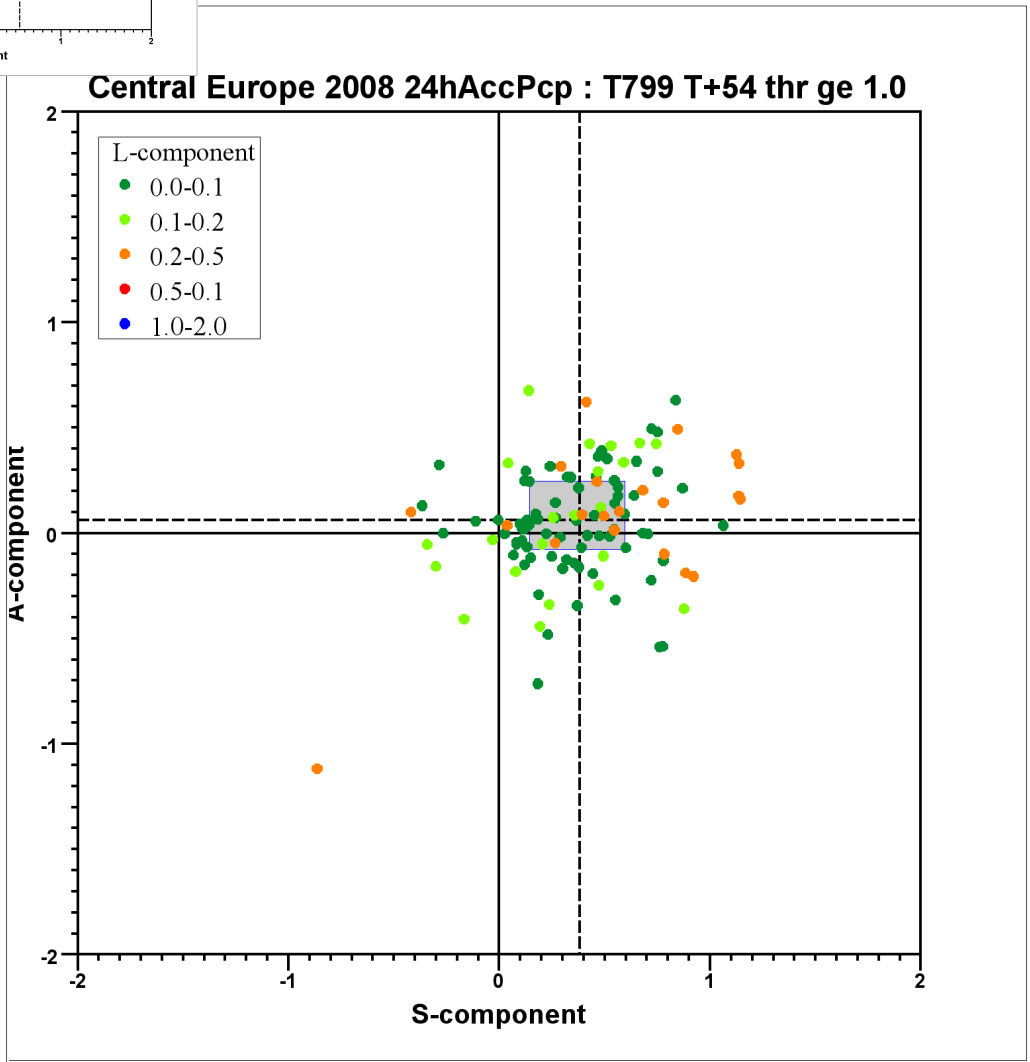

be [45N,5E,50N,15E] 13Feb2008 24hAccPcp(6-6UTC) SAL=(2.0, 1.8, 0.8)
(obj=7, thr=0.04)
ob ups 0.25 (obj=30, thr=0.03)
10°E
10°E
0,04 0,04 0,3 1 5 10 15 20 30 40 50 100
0,03 0,03 0,3 1 5 10 15 20 30 40 50 100

Pcp thr <1.0mm

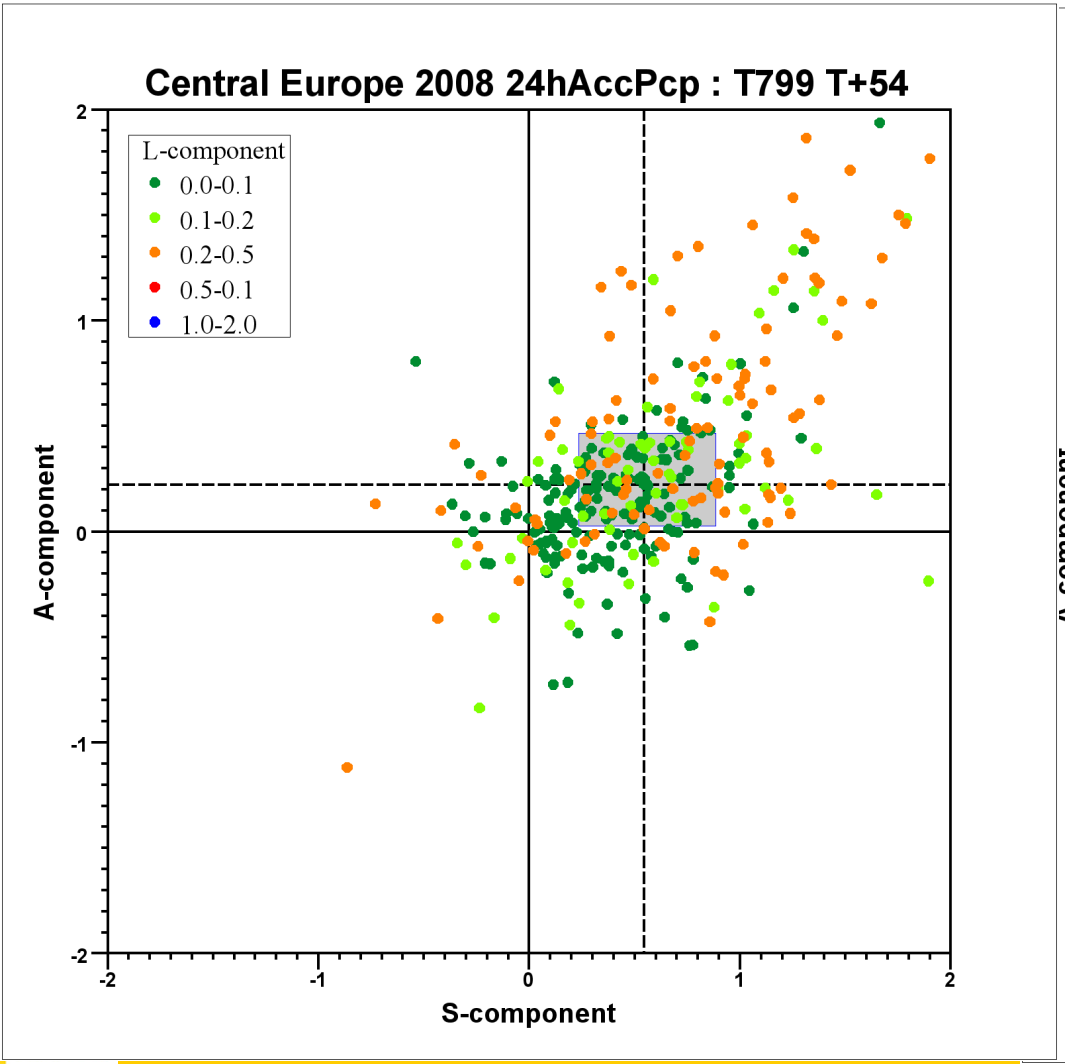


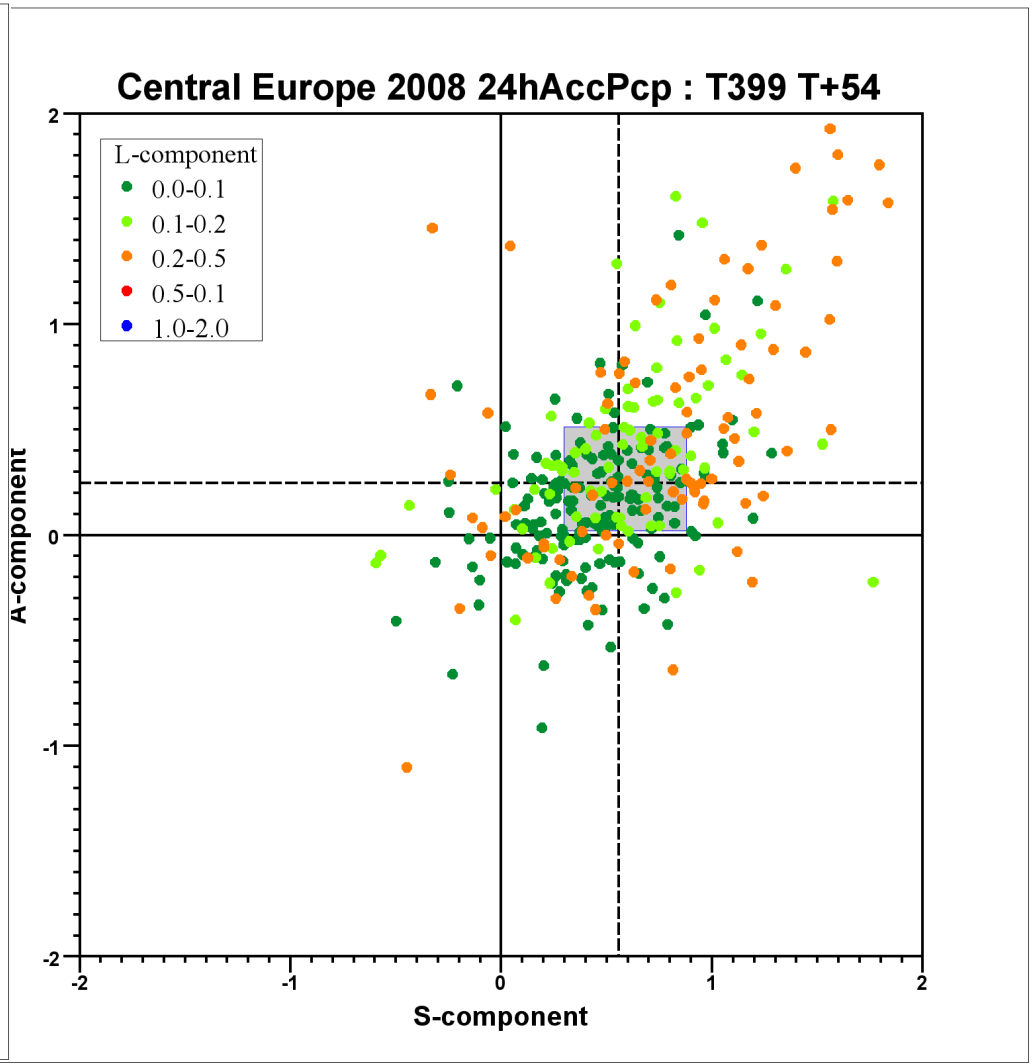
>=1.0mm



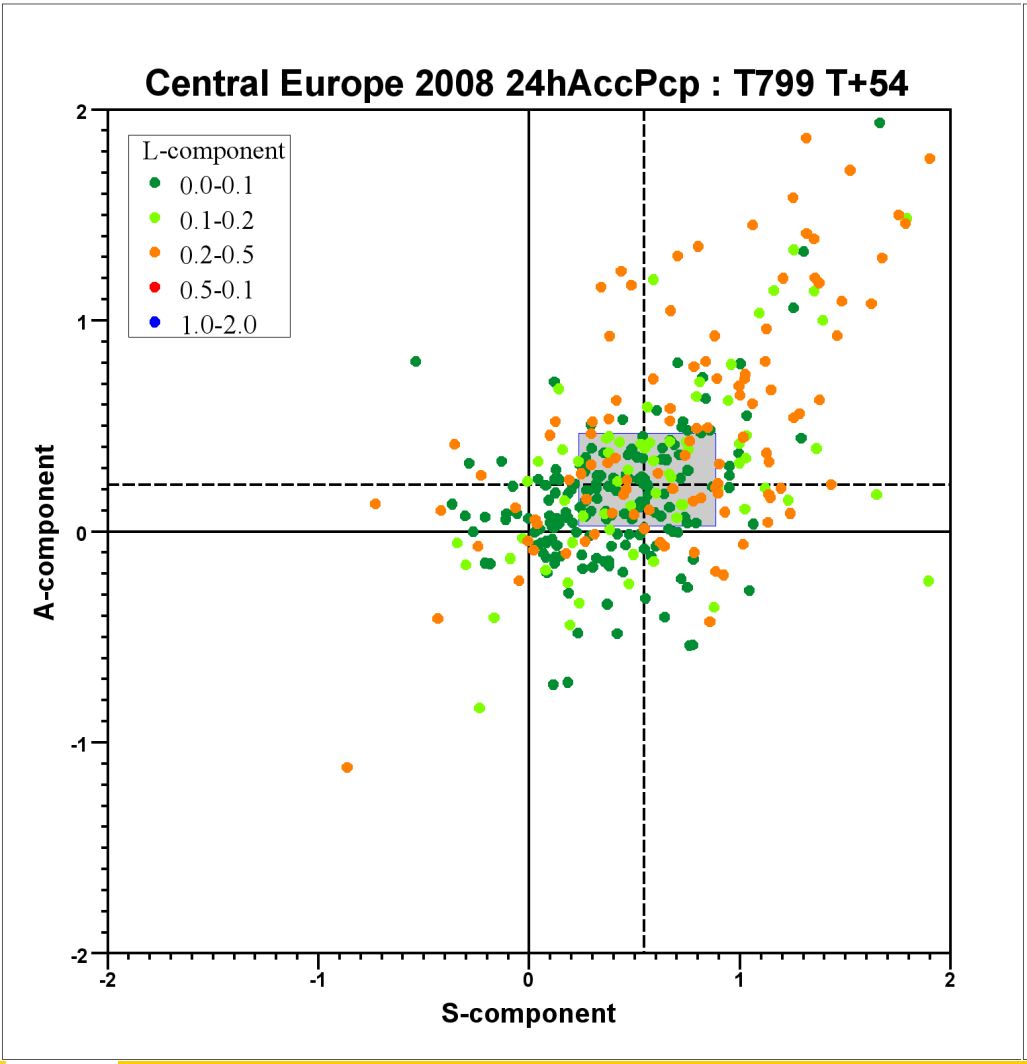


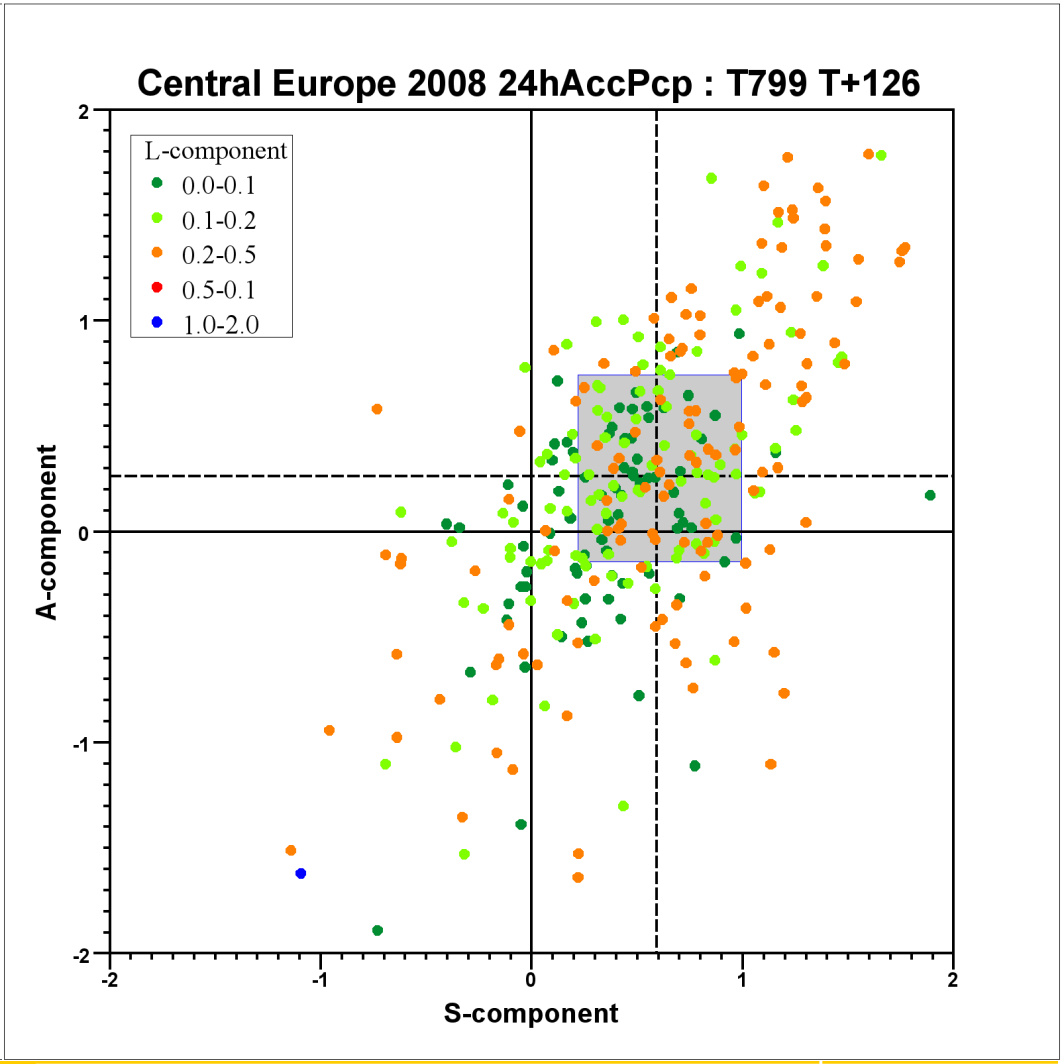
T799 & T399


Central Europe 2008 24hAccPcp : T799 T+54
L-component
0.0-0.1
0.1-0.2
0.2-0.5
0.5-0.1
1.0-2.0
A-component
S-component
2
1
0
-1
-2
-2
-1
0
1
2


Central Europe 2008 24hAccPcp : T399 T+54
L-component
0.0-0.1
0.1-0.2
0.2-0.5
0.5-0.1
1.0-2.0
A-component
S-component
2
1
0
-1
-2
-2
-1
0
1
2

T799 D+2 & D+5


Central Europe 2008 24hAccPcp : T799 T+54
L-component
0.0-0.1
0.1-0.2
0.2-0.5
0.5-0.1
1.0-2.0
A-component
S-component
2
1
0
-1
-2
-2
-1
0
1
2


Central Europe 2008 24hAccPcp : T799 T+126
L-component
0.0-0.1
0.1-0.2
0.2-0.5
0.5-0.1
1.0-2.0
A-component
S-component
2
1
0
-1
-2
-2
-1
0
1
2

Conclusions

- Problems of classical verification methods, e.g. double penalty can give better scores to a coarser grid model, so new methods must be explored
  - **S**tructure**A**mplitude**L**ocation measure (ϵ object-oriented methods) gives quantitative and detailed information about different aspects of model QPF performance
- First tests at ECMWF:
  - Collaboration with AEMET: SAL (original provided by Marcus Paulat, DWD) and up-scaling
  - Research about models 24h QPF SAL performance on 2008 over Central Europe
- Results look promising:
  - T799 D+2 Overall behaviour: overestimation of structure size (S), overestimation of pcp (A), location to improve (L)
  - Pcp threshold: Above 1mm much better performance not only on A, but also S and L
  - Model resolution: T799 and T399 perform similarly (each one at its own resolution)
  - Forecast step: D+5 performs worse: more L outliers, S and A keep the bias and open IQR
  - Still looking for other patterns: seasonal, flow-dependent, number of objects, etc.
- On-going work
  - Explore other clustering algorithms
  - Research on factor f for R* = f Rmax (regional sensitivity, introduce variability...)

Thanks to

- **Marcus Paulat (DWD), Matthias Zimmer (Univ. Mainz), Heini Wernli (Univ. Mainz)**
- Pertti Nurmi (FMI)
- Martin Goëber (DWD)
- AEMET & ECMWF computing support staff