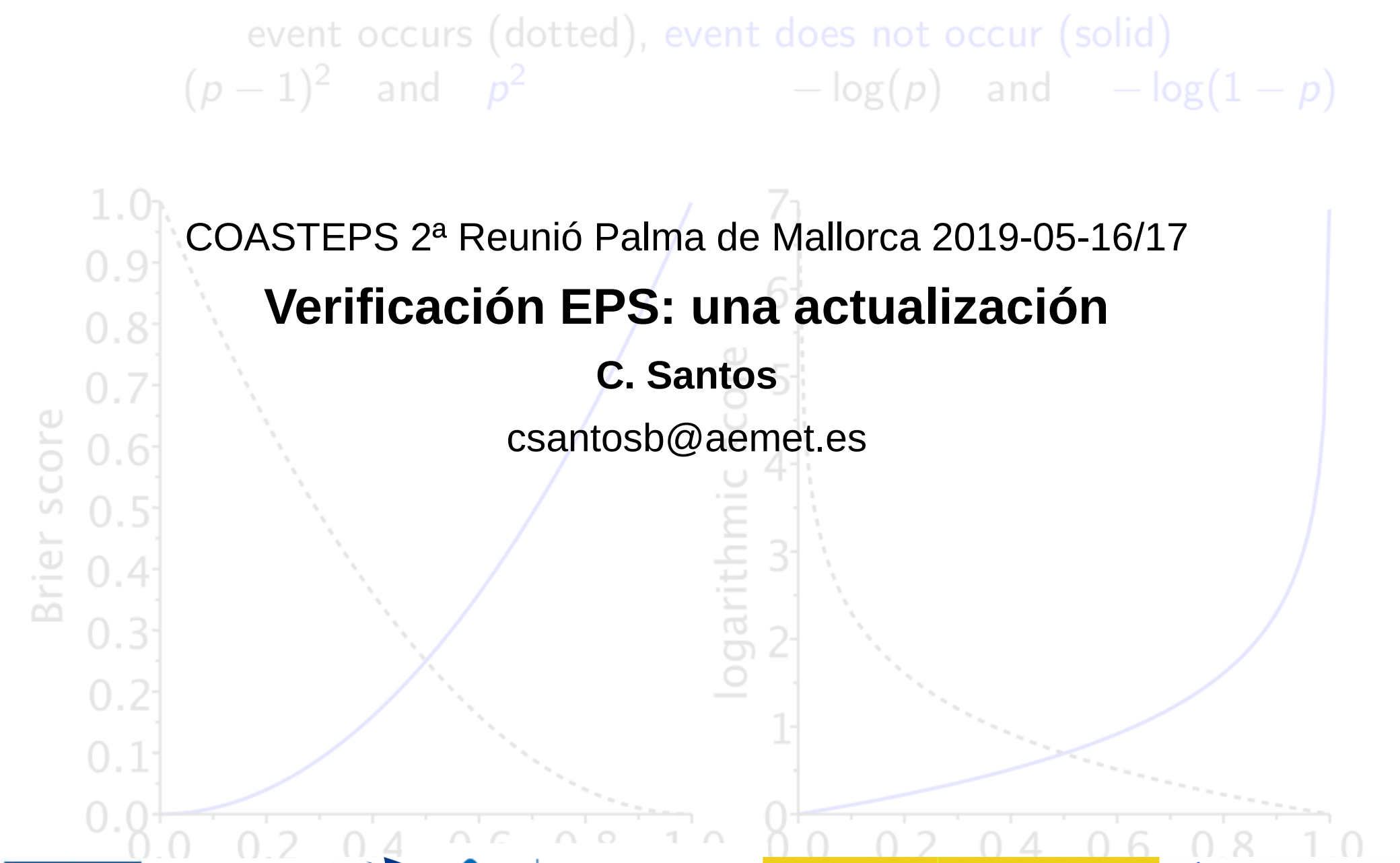


COASTEPS 2ª Reunió Palma de Mallorca 2019-05-16/17

Verificación EPS: una actualización

C. Santos

csantosb@aemet.es

EUROPEAN UNION
EUROPEAN REGIONAL DEVELOPMENT FUND
"A way to make Europe"
AGENCIA ESTATAL DE INVESTIGACIÓN
UIB
Universitat de les Illes Balears
GOBIERNO DE ESPAÑA
MINISTERIO PARA LA TRANSICIÓN ECOLÓGICA
AEMet
Agencia Estatal de Meteorología

Contenido

Repaso de métodos espaciales de verificación

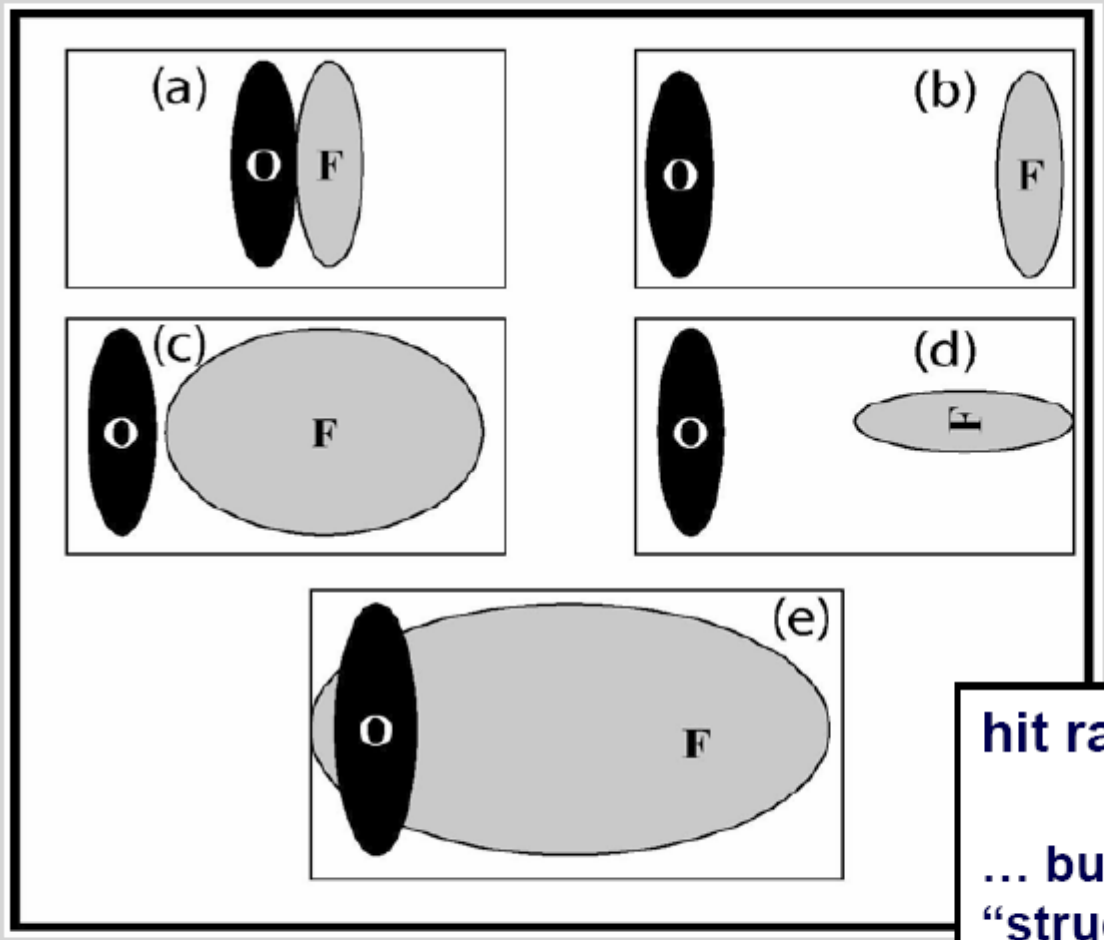
- ICP project
- SAL
- MODE

Ignorance score

Sensible probabilities

Double penalty

Problematic aspects of grid point based error scores



Davis et al. 2006 (MWR)

hit rate = 0

… but (a) is "better" than (b)-(d)

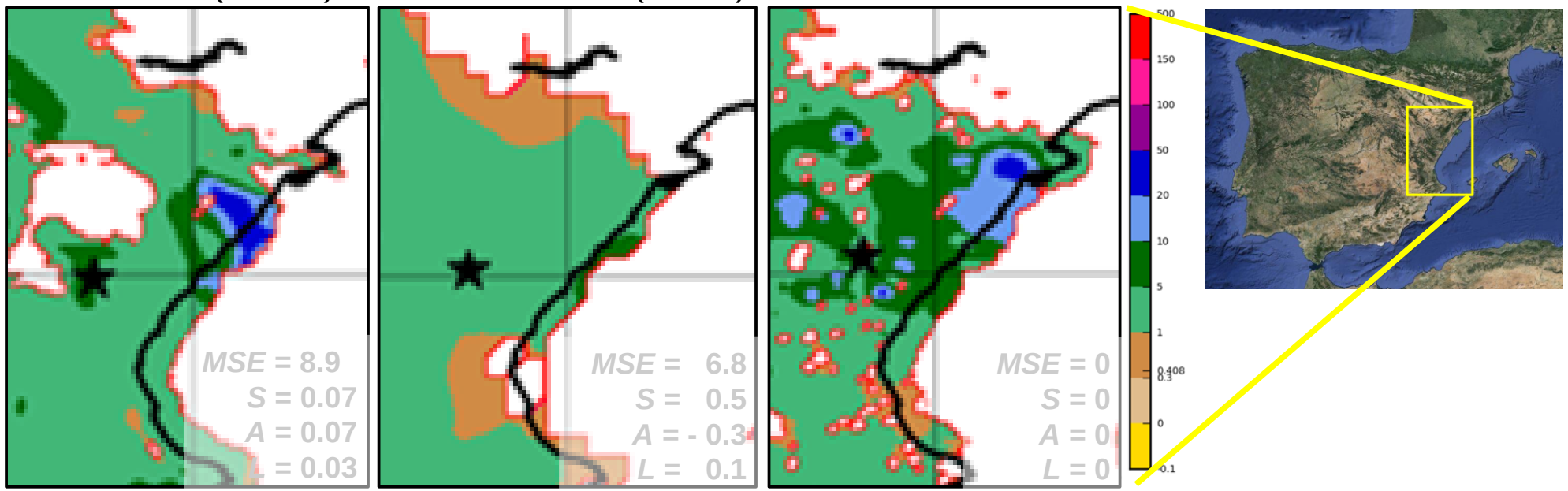
hit rate > 0

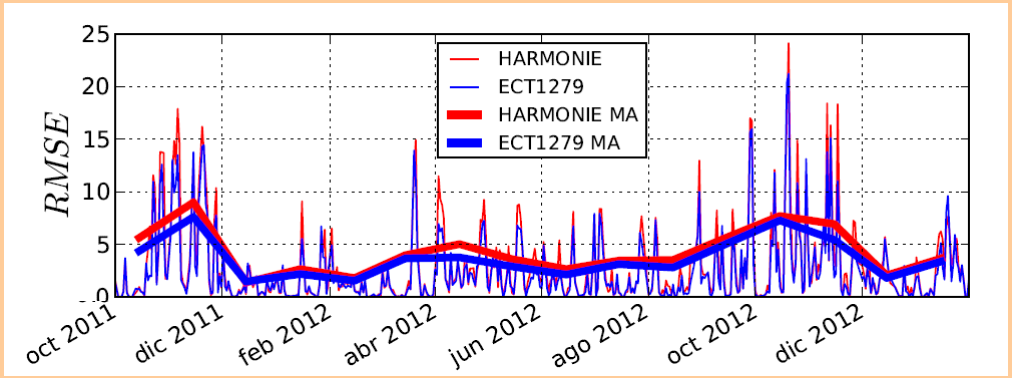
… but forecast has not the right "structure"

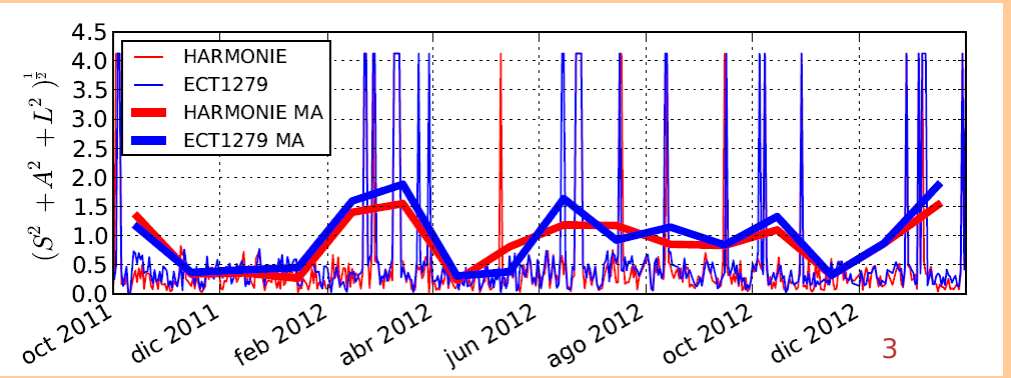
Double penalty problem

Which rain forecast would you rather use?

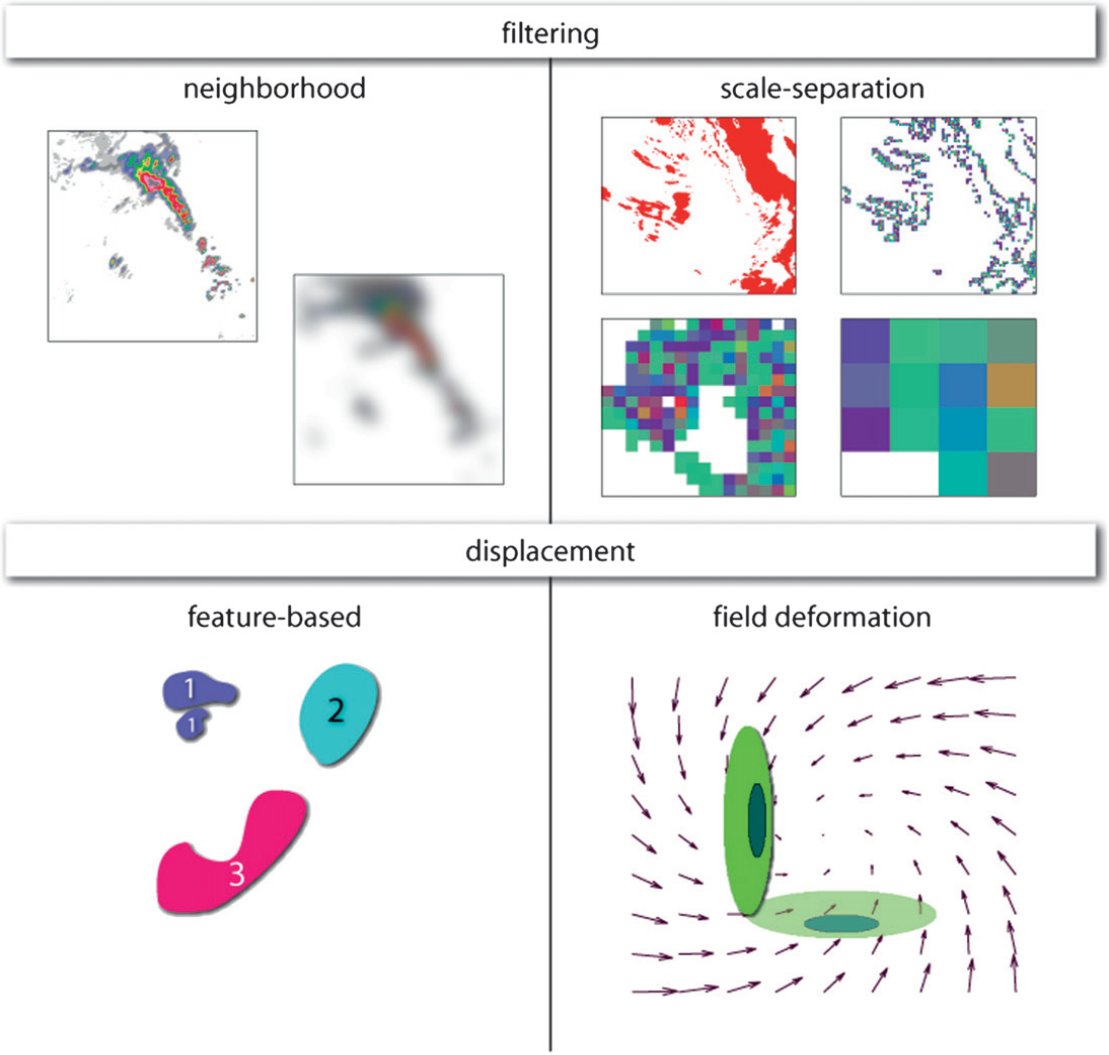
HARMONIE (2.5 km) ECMWF T1279 (16 km) OBSERVATIONS







Intercomparison Of Spatial Fc Verif Methods Project



Gilleland et al, 2009

How do the methods inform about performance at different scales?

How do the methods provide information on location errors?

Do the methods provide information on intensity errors and distributions?

Do the methods provide information on structure errors?

Do the approaches have the ability to provide information about hits, misses, false alarms, and correct negatives?

Do the methods do anything that is counterintuitive?

Do the methods have selectable parameters and how sensitive are the results to parameter choice?

Can the results be easily aggregated across multiple cases?

Can the methods be used to identify timing errors?

Can confidence intervals or hypothesis tests be readily computed for the method?

SAL (Wernli et al 2008)

(1) Upscale observations to model grid

(2) Search of objects in forecast and observation fields separately:

Thresholding: subjetive task (visual inspection)

$R_{ij} \geq R^* = f\, R_{95}$, f=1/15 empirical

Clustering: select grid points belonging to object (8 ngb)

(3) Object properties

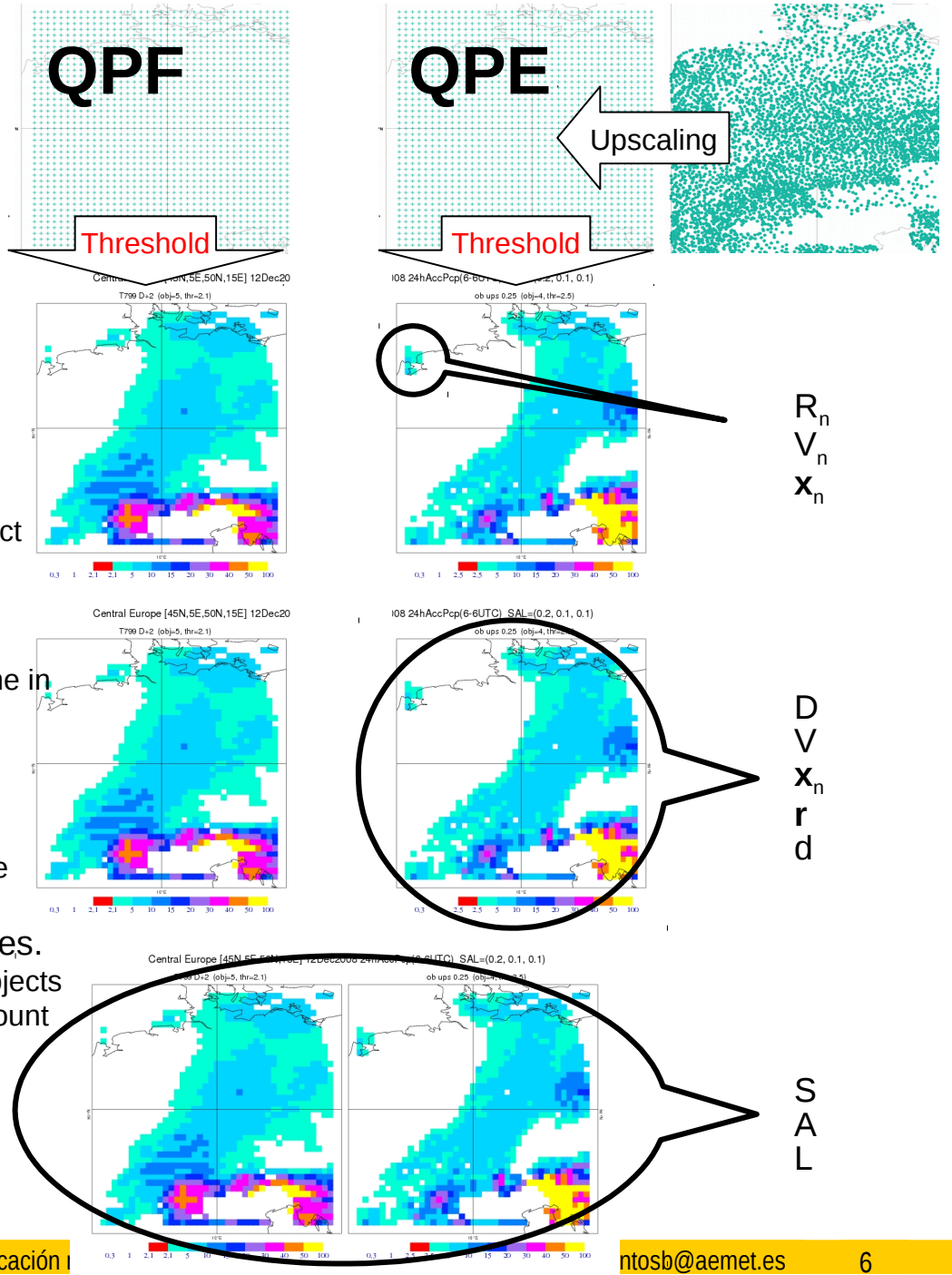
$R_n = \sum R(n)_{ij}$	total pcp of the object
$\mathbf{x}_n = \sum R(n)_{ij}\mathbf{x}(n)n_{ij} / \sum R(n)_{ij}$	object centre of mass
$V_n = R_n / R(n)_{(i,j),\,max}$	scaled volume or total pcp of the object normalized by its max value

(4) Field properties

$D = (1/IJ)\sum R_{ij} \sim E[R]$	domain-average pcp value in the field
$V = \sum R_n V_n / \sum R_n \sim E[R^2]$	weighted mean of objects scaled volume in the field
$\mathbf{x} = \sum R_{ij}\mathbf{x}_{ij} / \sum R_{ij}$	field centre of mass
$\mathbf{r} = \sum R_n\lvert\mathbf{x}-\mathbf{x}_n\rvert / \sum R_n$	weighted mean of distance between objects MCs and field MC
d	max distance between two points inside domain

(5) SAL measures normalized differences of field properties.

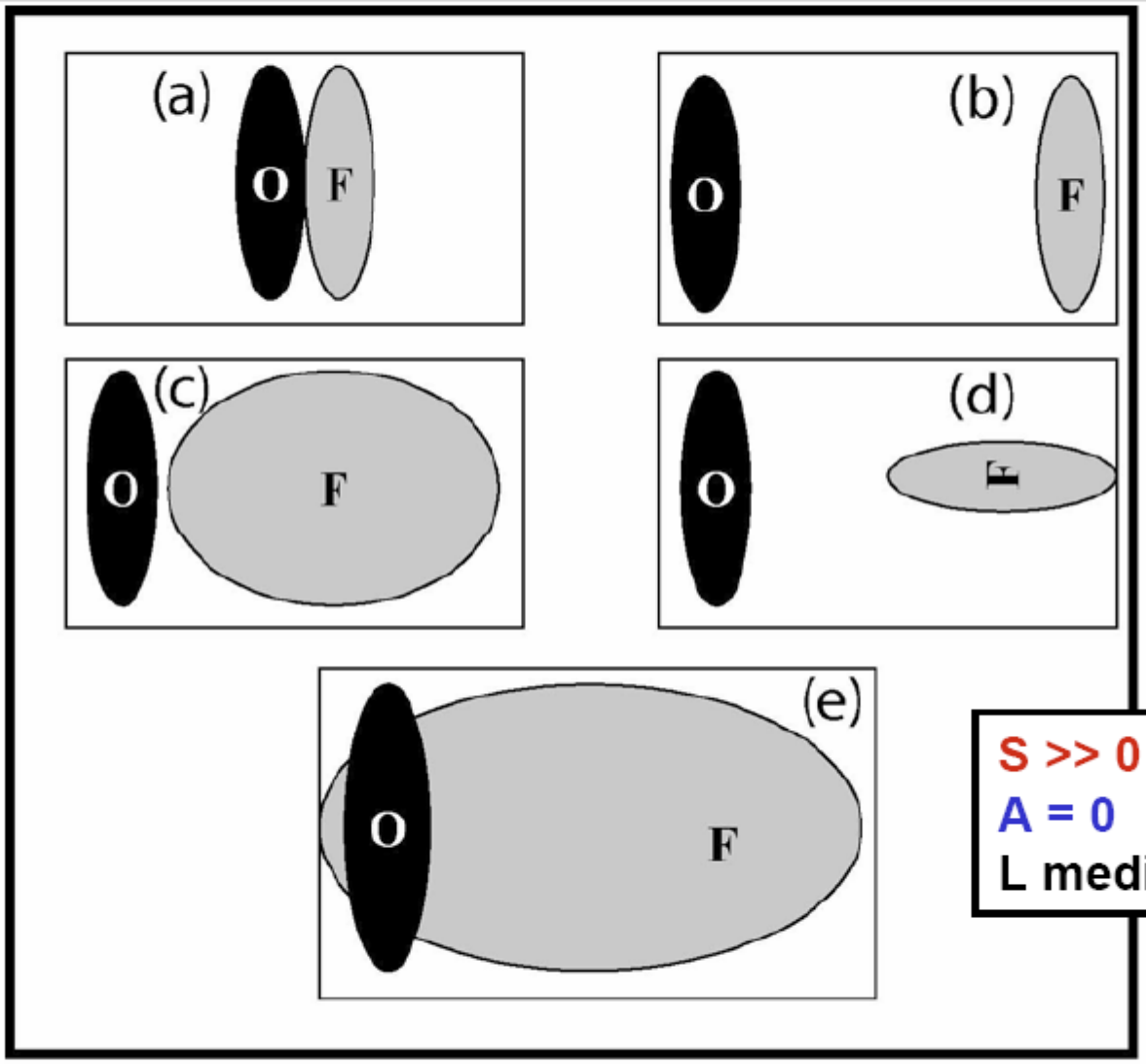
$S = 2\,(Vf-Vo) / (Vf+Vo)$	**Structure** → size and shape of pcp objects
$A = 2\,(Df-Do) / (Df+Do)$	**Amplitude** → accuracy of the total amount of pcp
$L_1 = \lvert\mathbf{x}f-\mathbf{x}o\rvert / d$	**Location** → accuracy of pcp
$L_2 = 2\,(r_f-r_o) / d$	distribution and relative
$L = L_1 + L_2$	positions of objects in the field.



SAL

S = 0
A = 0
L small

S > 0
A = 0
L medium



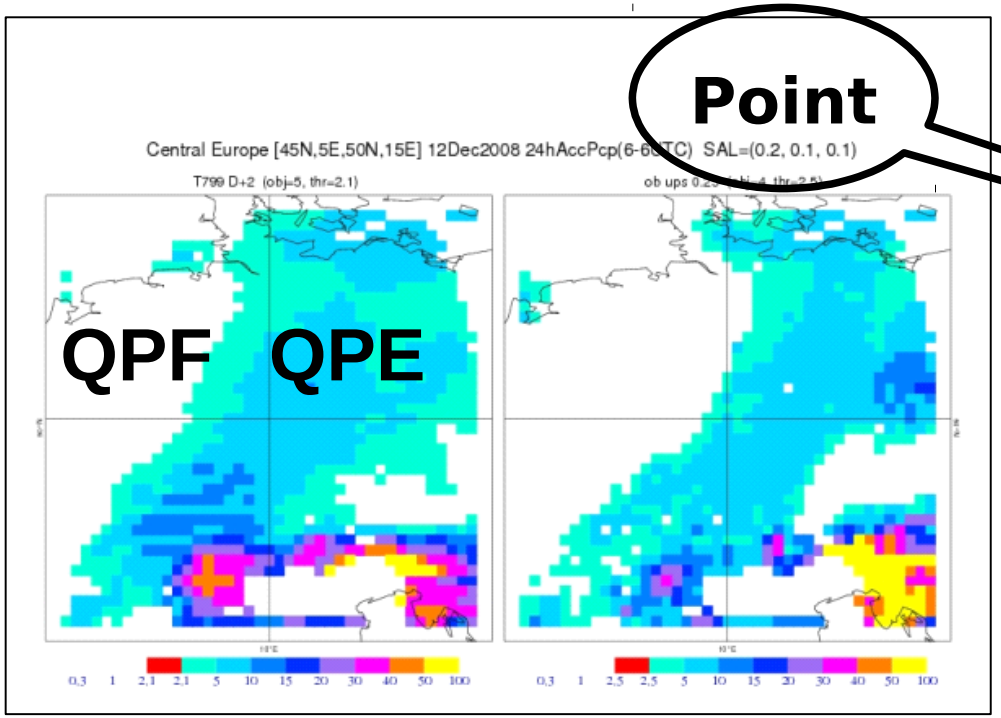
S = 0
A = 0
L large

S = 0
A = 0
L large

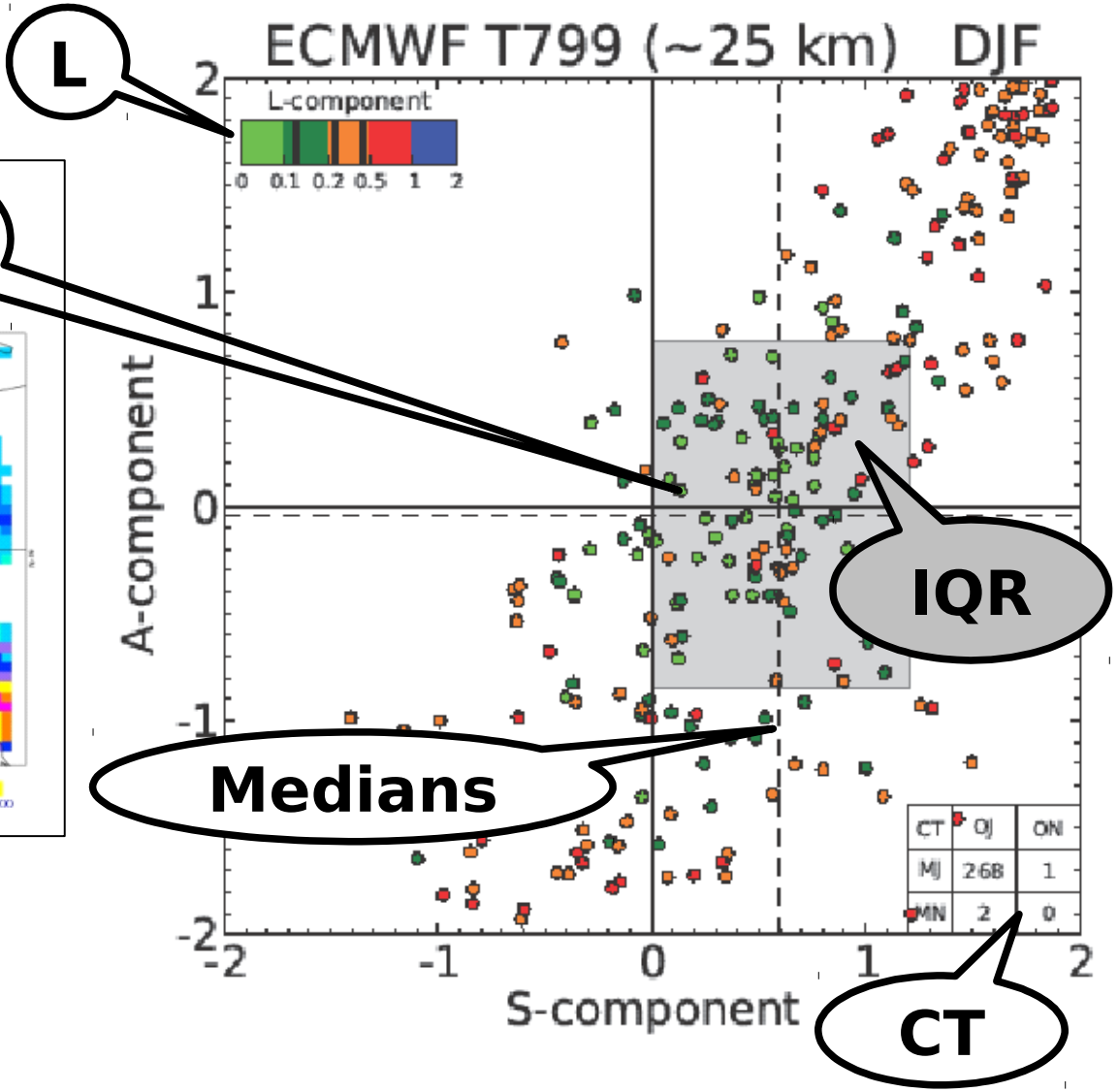
S >> 0
A = 0
L medium

Davis et al. 2006

SAL plot



Point

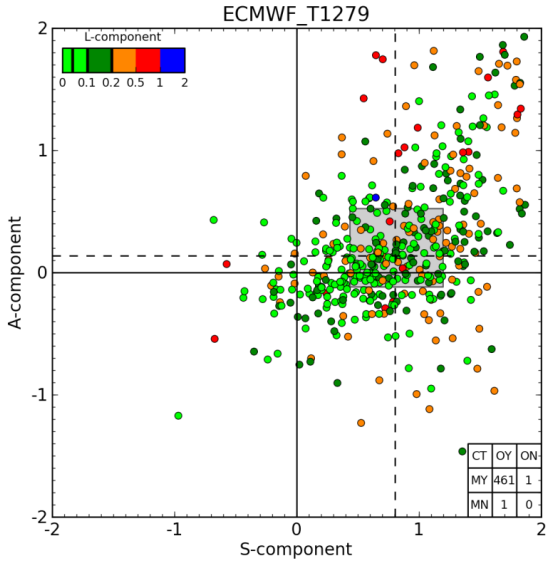


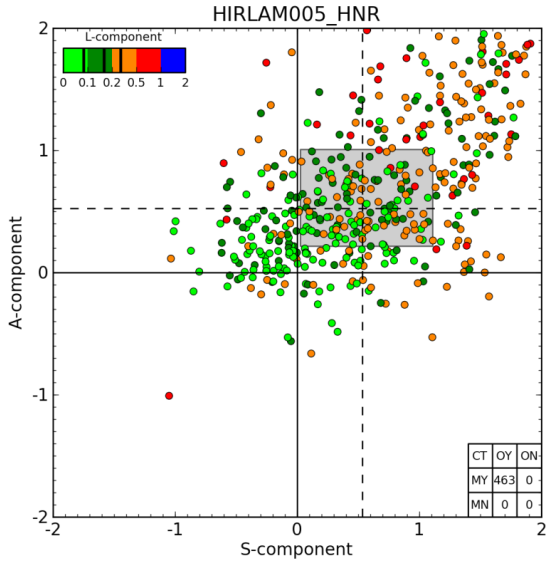
S:	**Structure**	**-2** … objects too small or too peaked	**0** Perfect	… **+2** objects too large or too flat
A:	**Amplitude**	**-2** … averaged QPF under-estimated	**0** Perfect	… **+2** averaged QPF over-estimated
L:	**Location**		**0** Perfect	… **+2** wrong location of Total Center of Mass (TCM)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

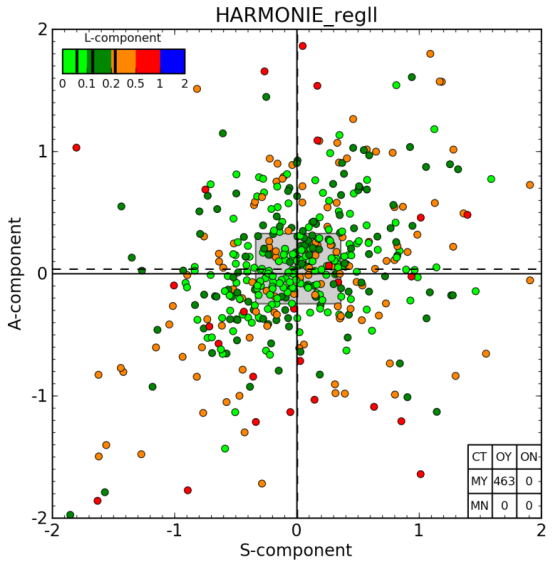
Quantitative, explicit and accumulative information about different aspects of QPF performance

SAL

Objective and fair comparison for models of different resolution, without penalty for mesoscale ones; Structure (X) Amplitude (Y) Location (color)

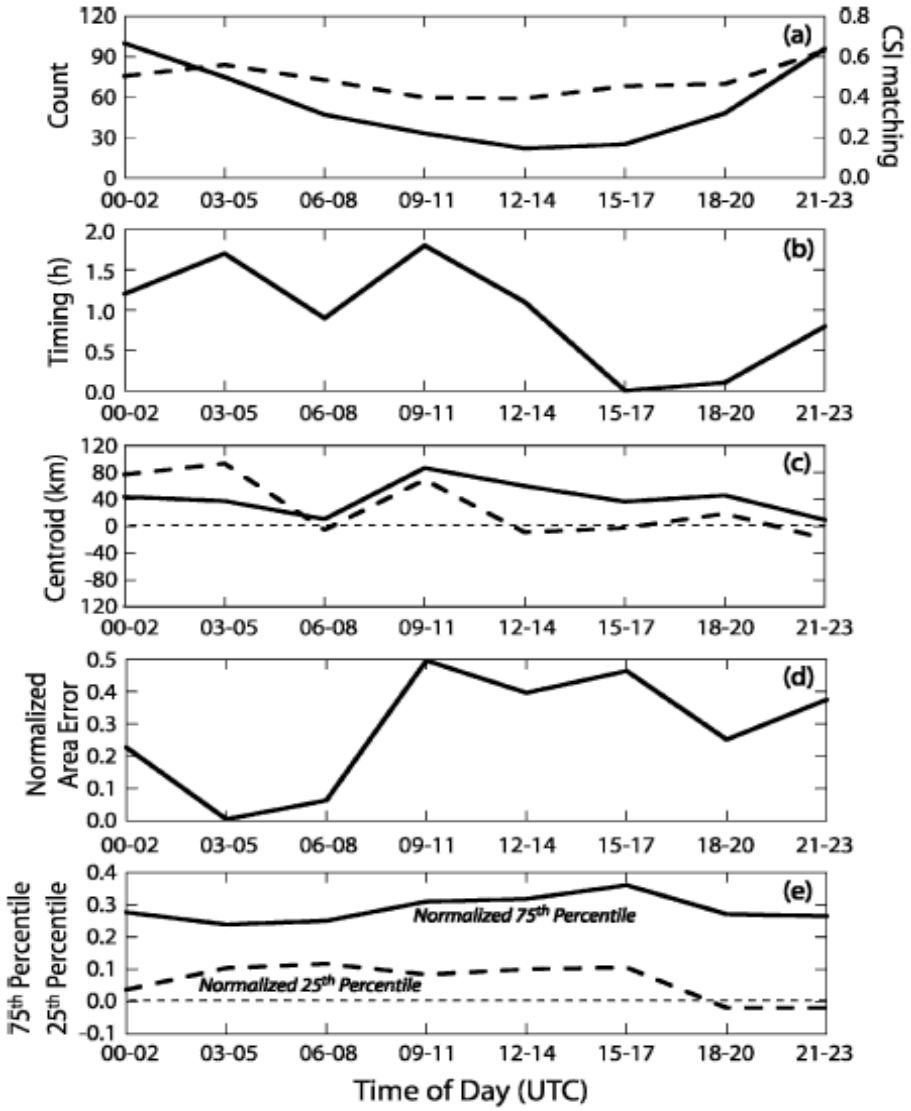






ECMWF T1279 (17 km) **HIRLAM 0.05 (5 km)** **HARMONIE 2.5 km**

MODE (Davis et al 2008)



MODE proporcionado por el paquete de verificación de NCAR:
Model Evaluation Tools (MET).

Modificaciones posteriores del MODE (interés total):
Davis et al. 2009

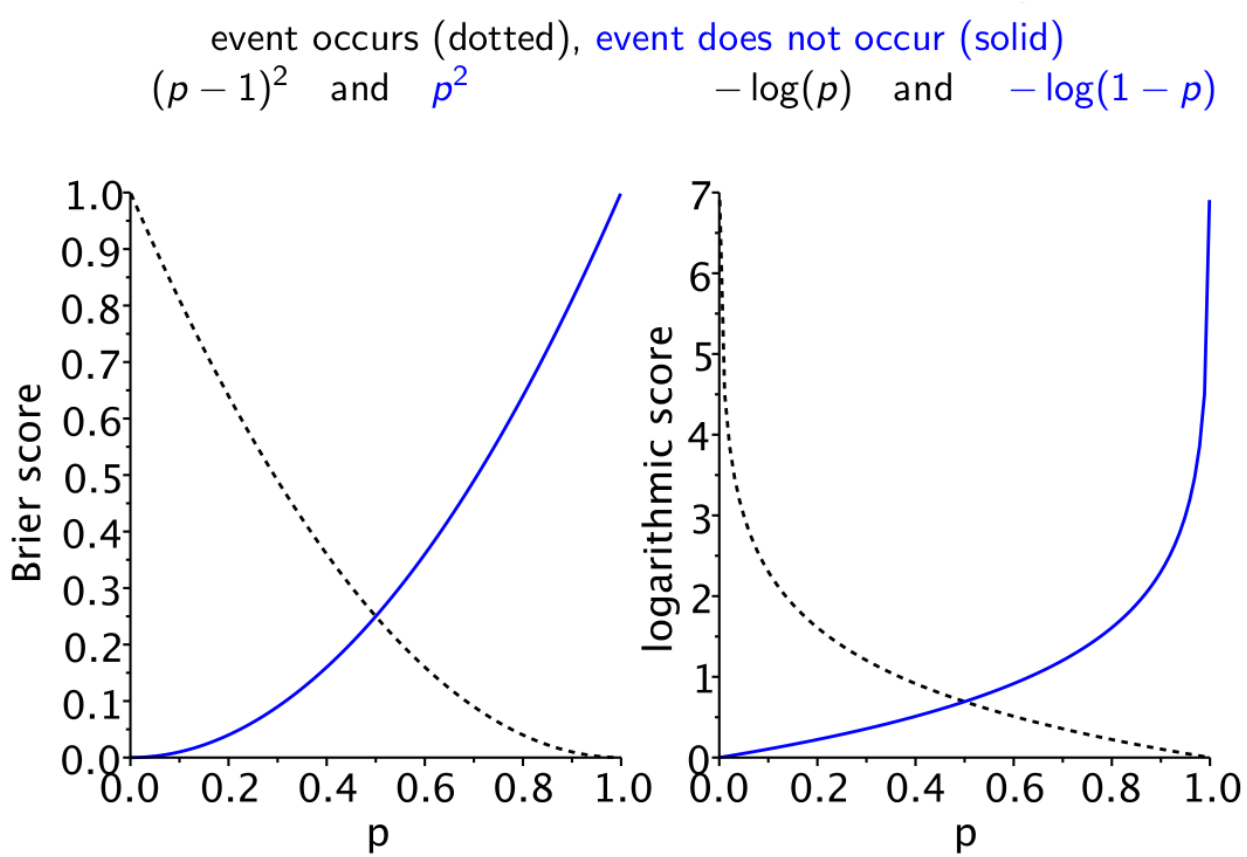
MODE aplicado a ensembles:
Gallus, 2010

Ignorance score (Good 1952, Roulston and Smith 2000)

$$LS = -\frac{1}{N}\sum_{k=1}^{N} o_k \log p_k + (1-o_k)\log(1-p_k)$$

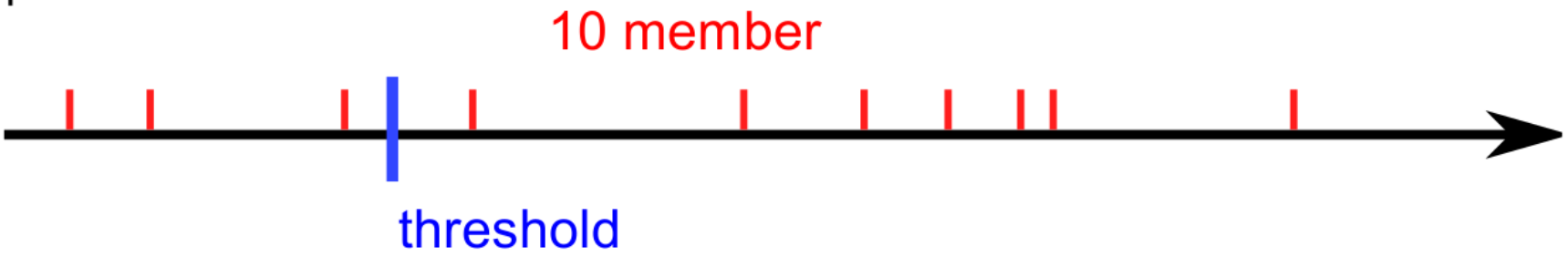
$0 < LS < \infty$

El valor ifinito se da cuando p=0 y el evento ocurre, o viceversa.
Moraleja: no predecir nunca p=0 o p=1
Incluso cuando la verdadera probabilidad no sea 0 (1) habrá casos en que ningún (todos) miembros(s) prevea el evento.



Sensible probabilities (Leutbecher 2019)

- Never forecast $p = 0$ or $p = 1$ unless you are really certain!
- If the true probability is not equal to zero (or one), there will still be cases when no member (or all members) predict(s) the event.
  Sampling uncertainty!
- Wilks proposed to estimate cumulative probabilities using Tukey's plotting positions



- When n members of an M-member ensemble have a value less than the threshold θ, the probability to not exceed θ is set to

$$p^{(T)}(n) = \frac{n + 2/3}{M + 4/3}$$

- Consider for example $M = 10$:

n	0	1	2	3	4	5	6	7	8	9	10
p	0.06	0.15	0.24	0.32	0.41	0.50	0.59	0.68	0.76	0.85	0.94

Gracias