EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

Assessing performance of the European Flood Alert System EFAS

J.Bartholmes
& EFAS team

EC, Joint Research Centre, IES-LMU, Ispra (IT)

Land Management Unit

EFAS skill assessment 2005-2007

The period analyzed spans 25 consecutive months from 1st Jan. 2005 to 31st Jan. 2007 (only the 12:00 forecasts)

The quantitative approach assesses performance of forecast with regard to all four fields of the contingency table. For all 5x5 km2 pixels > 4000 km^2 ups on the grid for the whole of Europe :

- Hits, false alarms, misses and positive rejects
- Leadtime
- Persistence criteria

The EFAS forecasts are compared against a proxy, which is the simulated discharge obtained with observed meteorological input data

Questions :

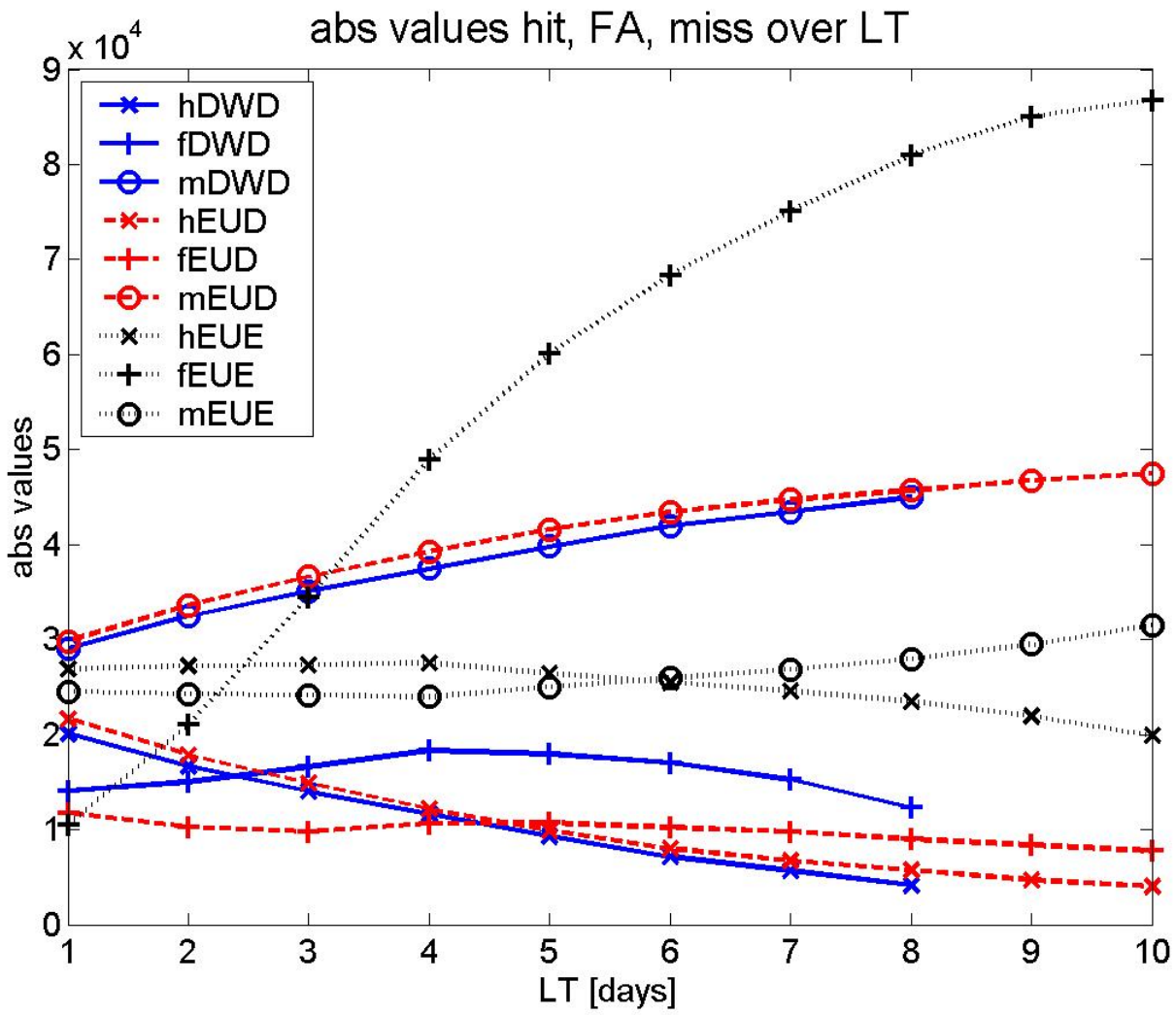
- What is the relation between lead time and skill ?
- What is the improvement when using the persistence criteria ?
- What is the *real* probability of a forecast ?
- Which skill scores should be used ?
- Are there regional differences in skill ?

Persistence criteria
or the good old “wait and see…”

- Only if the threshold exceedance in a river stretch is forecasted continuously on 2 consecutive dates it is considered as persistent.
- The definition of persistence regarding EFAS EPS is linked to 2 thresholds: the EFAS alert threshold (HAL or SAL) and a threshold concerning the number of EPS members that are persistently above the respective alert threshold

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

5 EPS persistent



Joint Research Centre

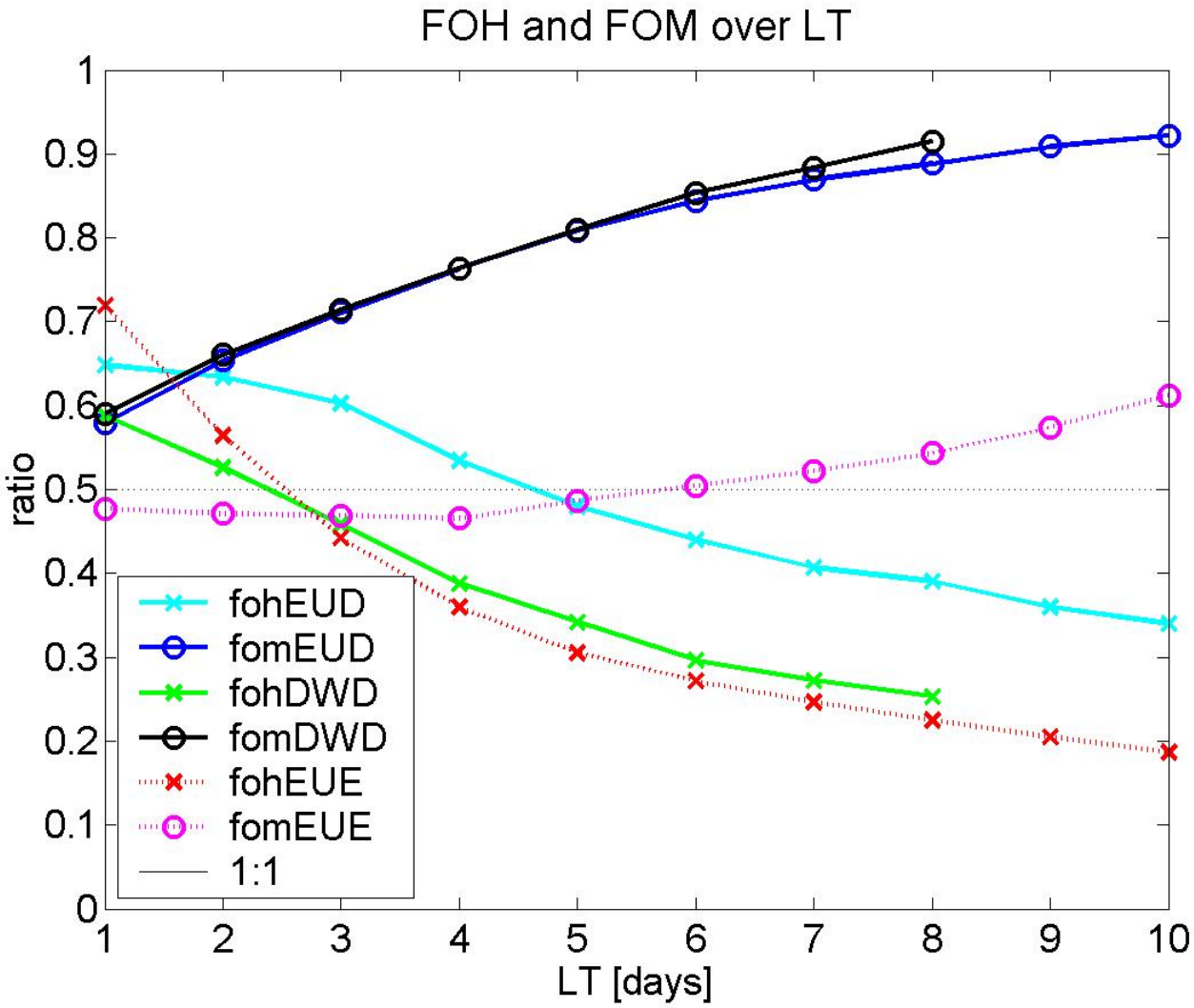
Land Management Unit

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

5 EPS persistent

FOH=hits/(hits+false)

FOM=misses/(hits+misses)



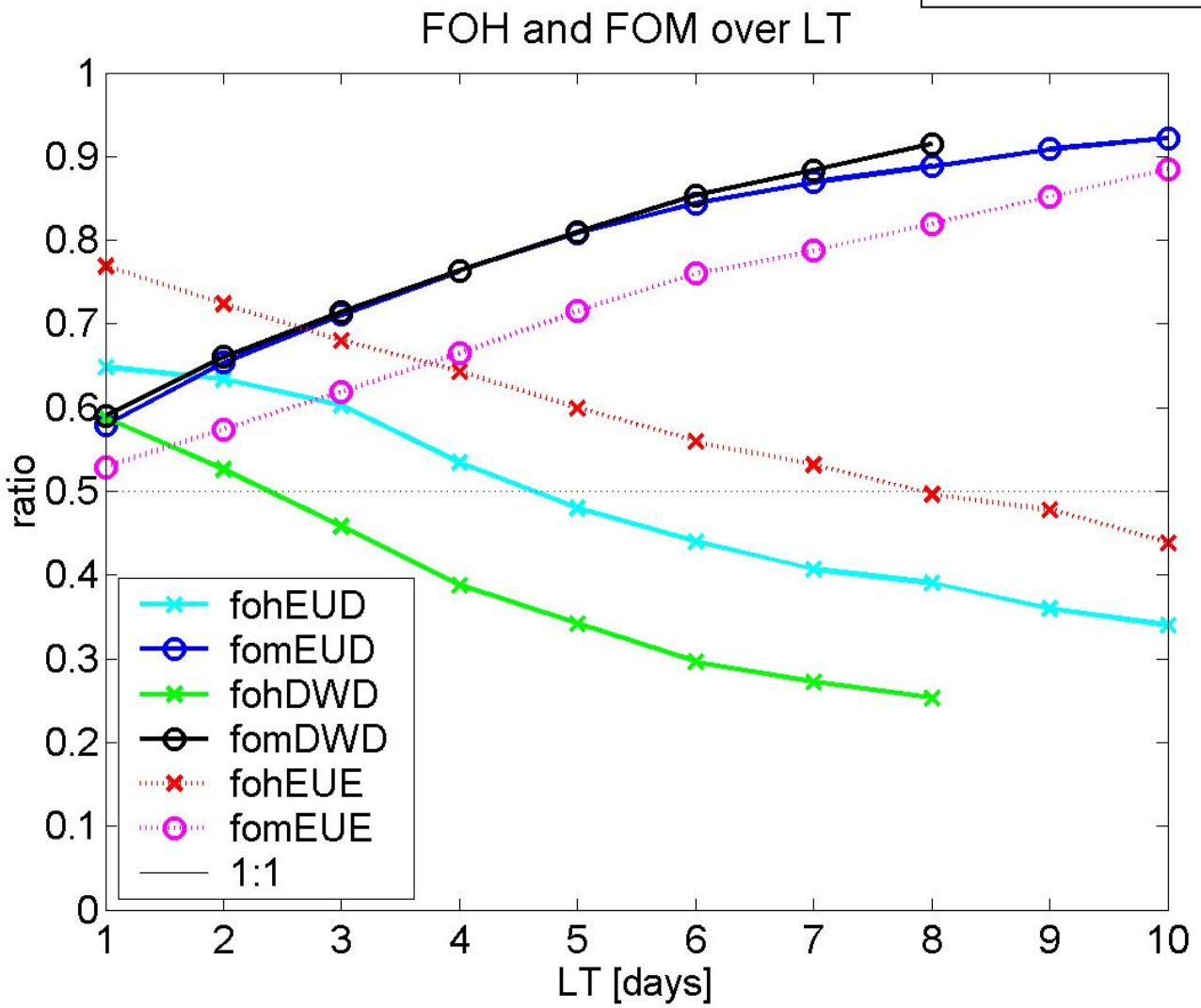
Joint Research Centre

Land Management Unit

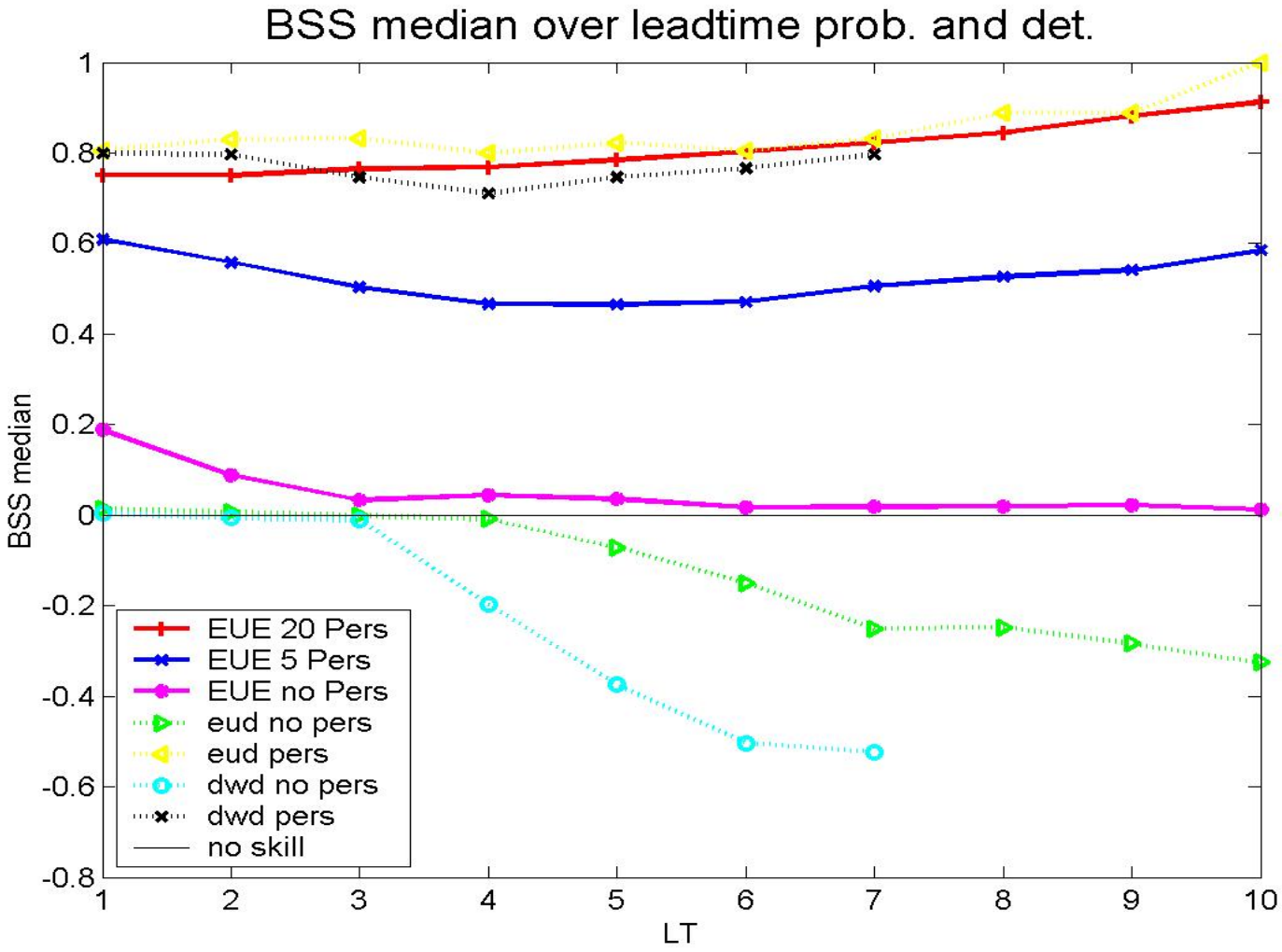
25 EPS persistent

FOH=hits/(hits+false)

FOM=misses/(hits+misses)



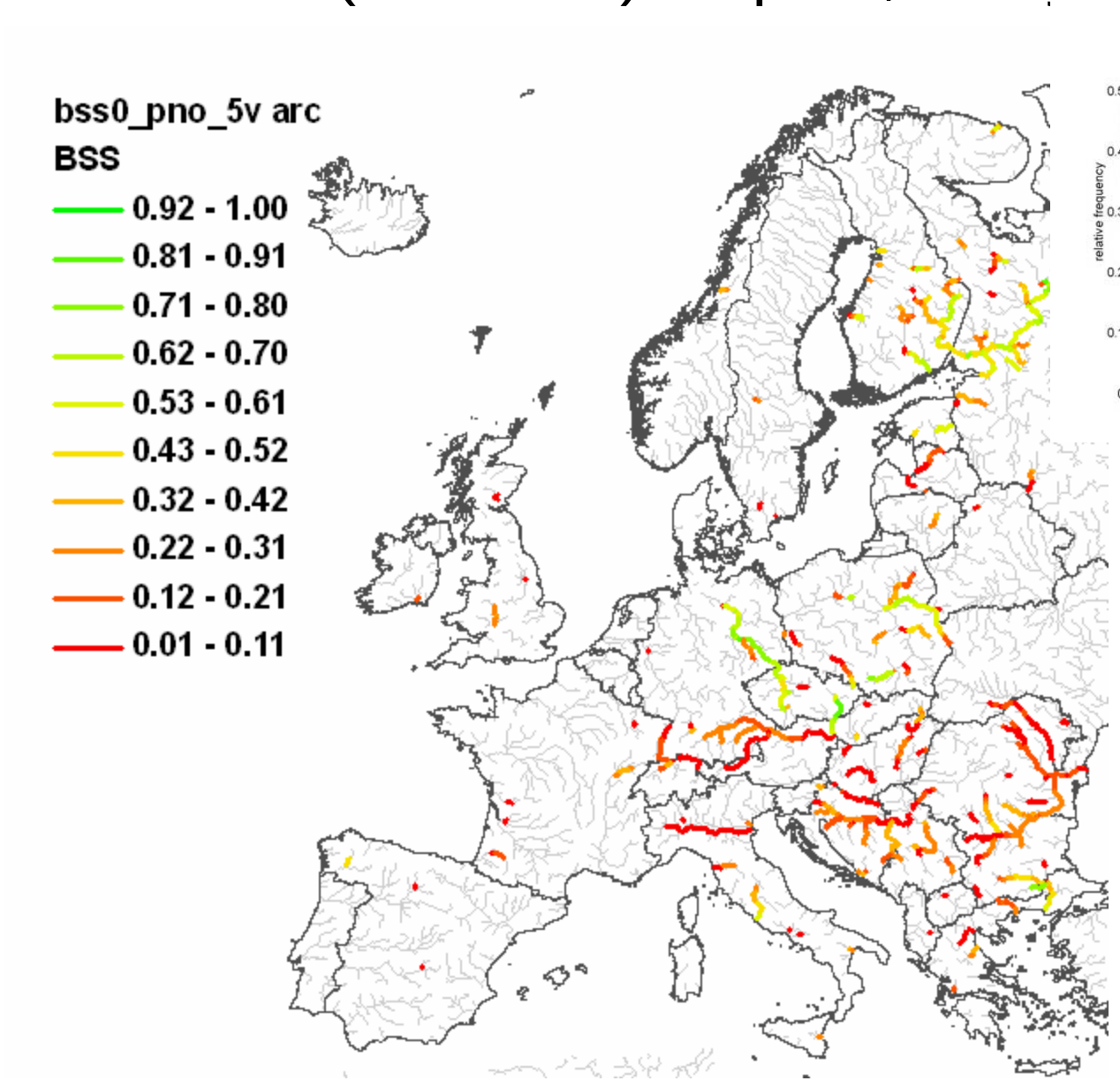
$$BSS = 1 - \frac{BS_f}{BS_{c\,\lim}}$$ range[0,1], best: 1 With: $$BS = \frac{1}{N}\sum_{1}^{N}(p-o)^2$$

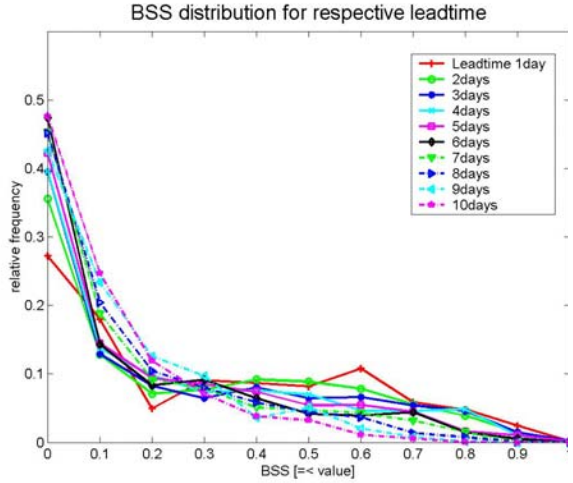


EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

BSS (clim 0.00) no pers, 6 days leadtime

Joint Research Centre



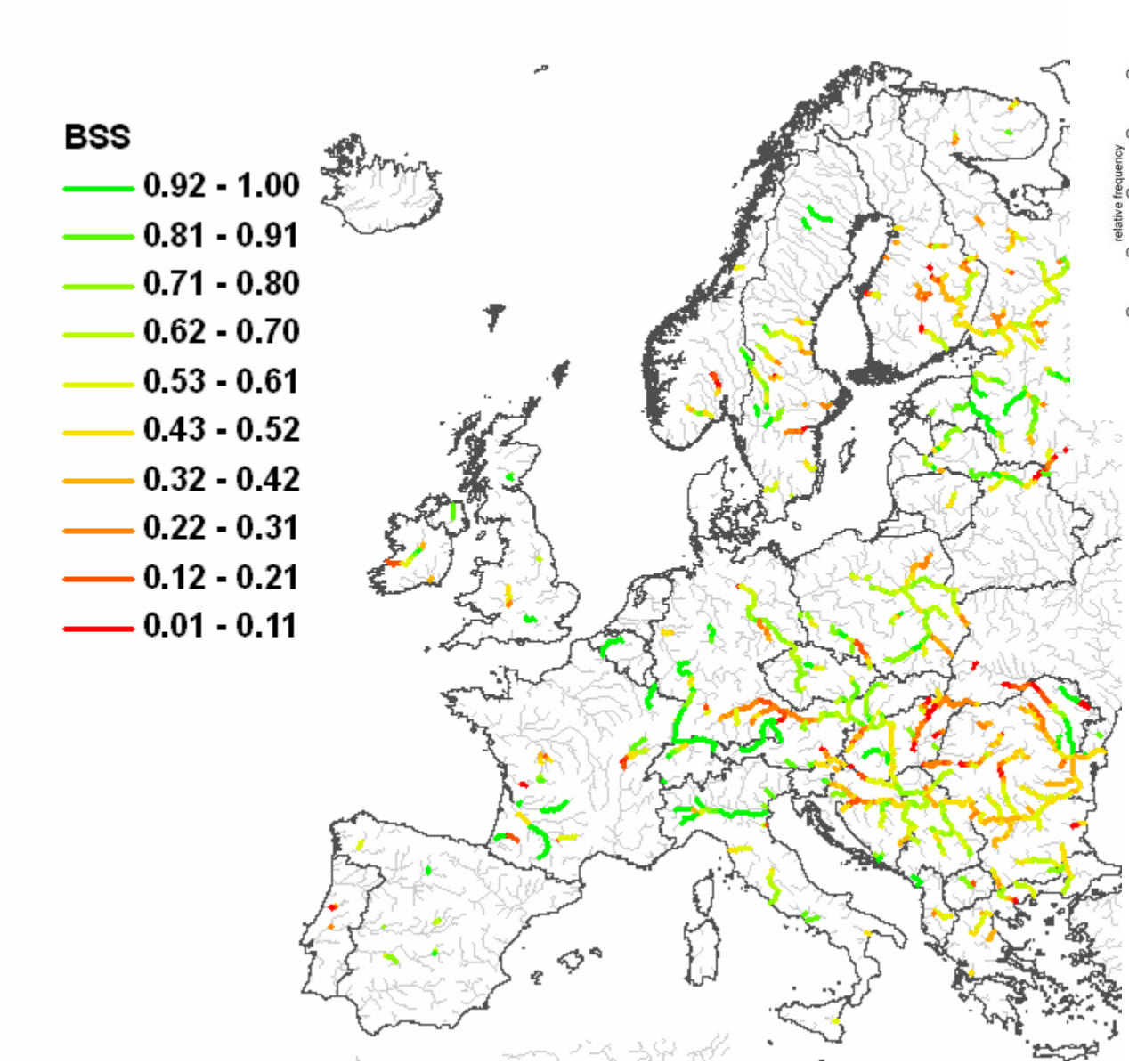


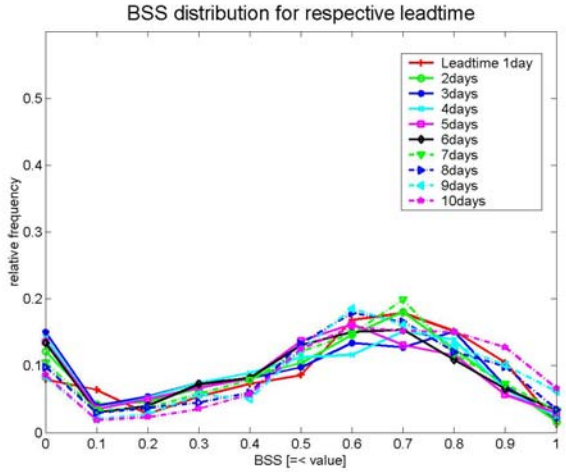
Land Management Unit

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

BSS (clim 0.00) 5 pers, 6 days leadtime

Joint Research Centre



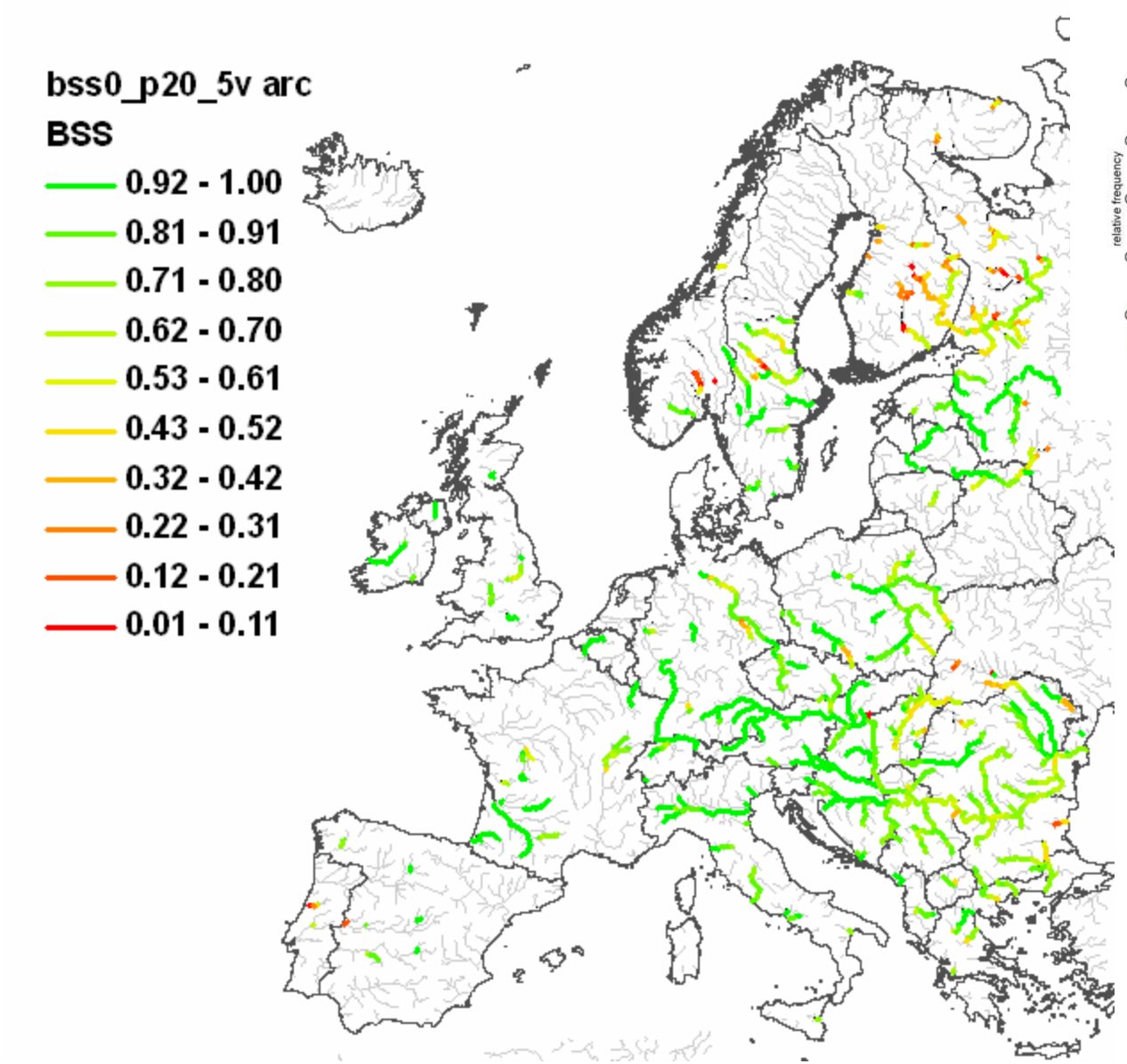


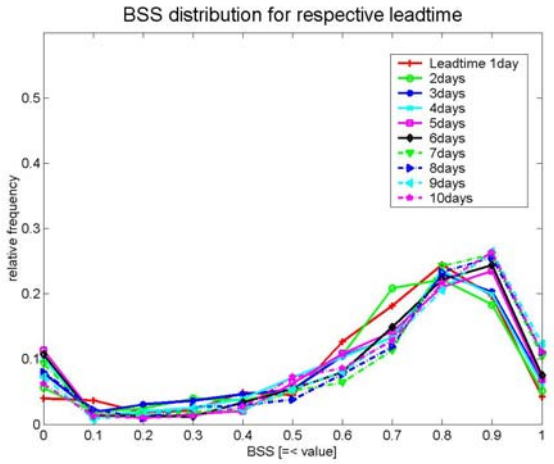
Land Management Unit

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

BSS (clim 0.00) 20 pers, 6 days leadtime

Joint Research Centre



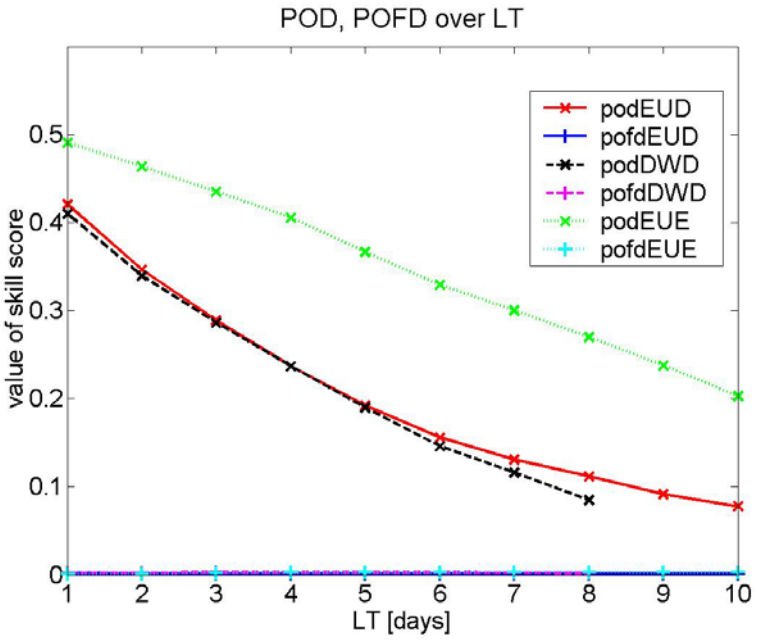


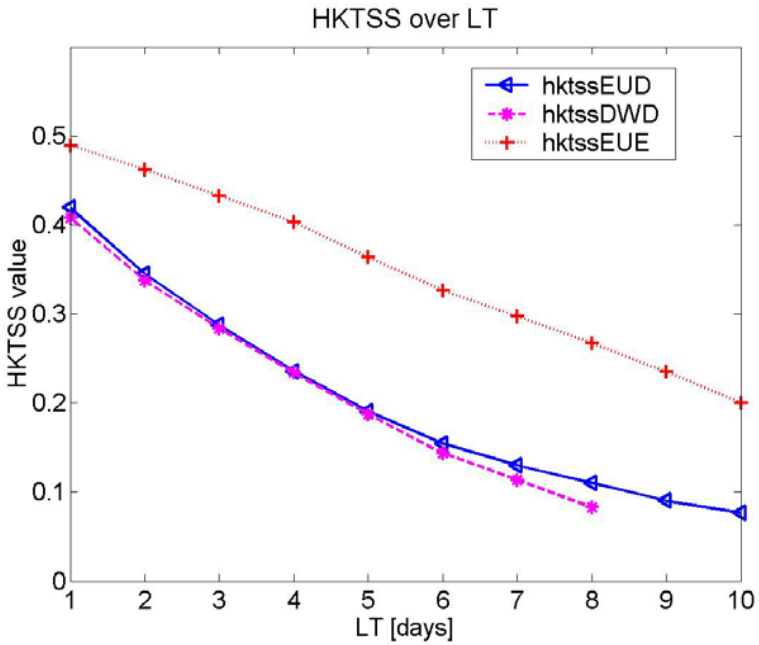
Land Management Unit

- The HKTSS and the OR are described as equitable (Gandin and Murphy,1992), least influenced climatology and representative for different climatic regimes (McBride and Ebert, 2000)
- Stephenson (2000) advocated a combination of those two with the bias score.

POD=hits/(hits+misses) POFD=false/(false+posReject)

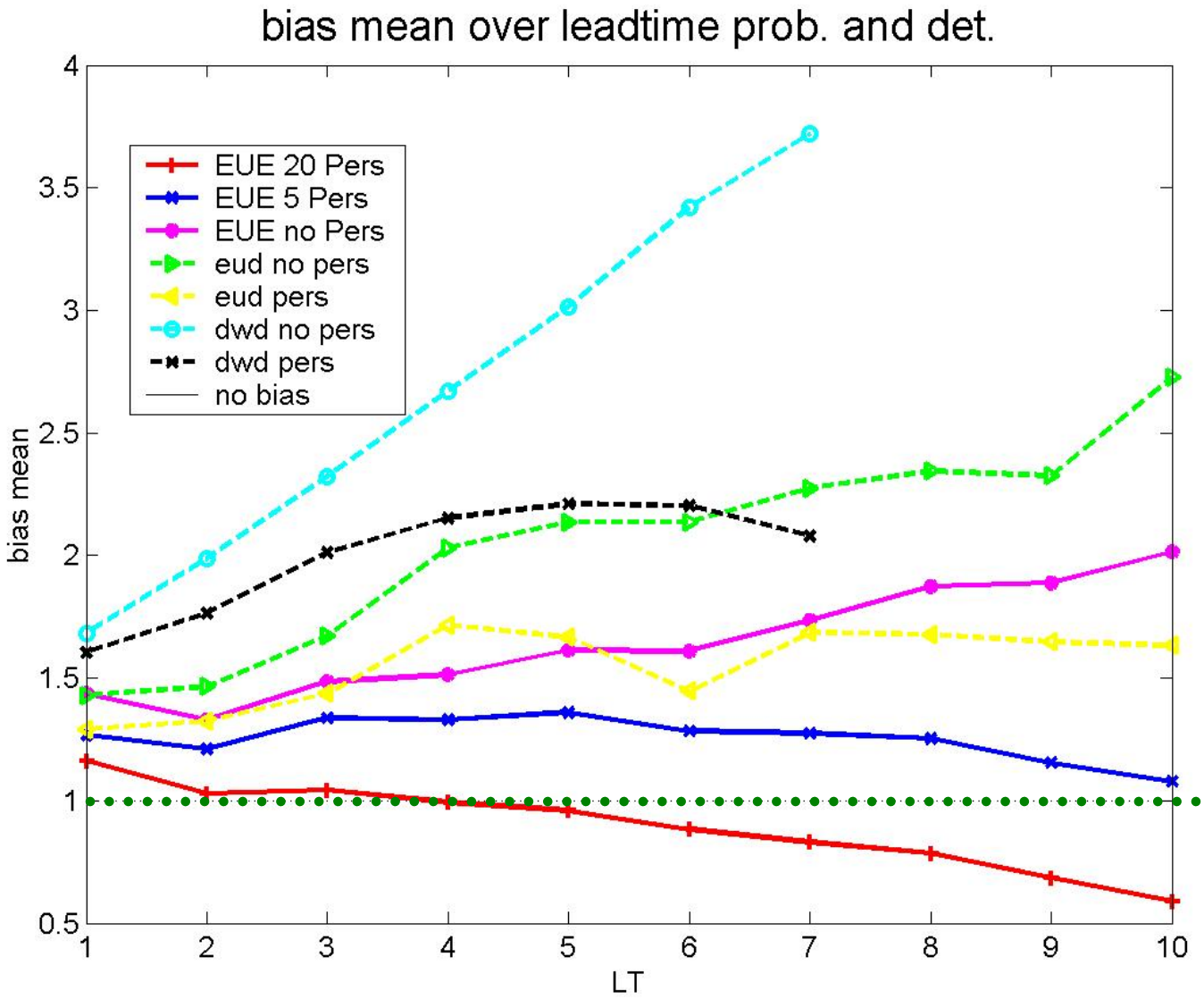
HKTSS=POD-POFD





15 EPS persistent

Bias= (hits+false) / (hits + misses)

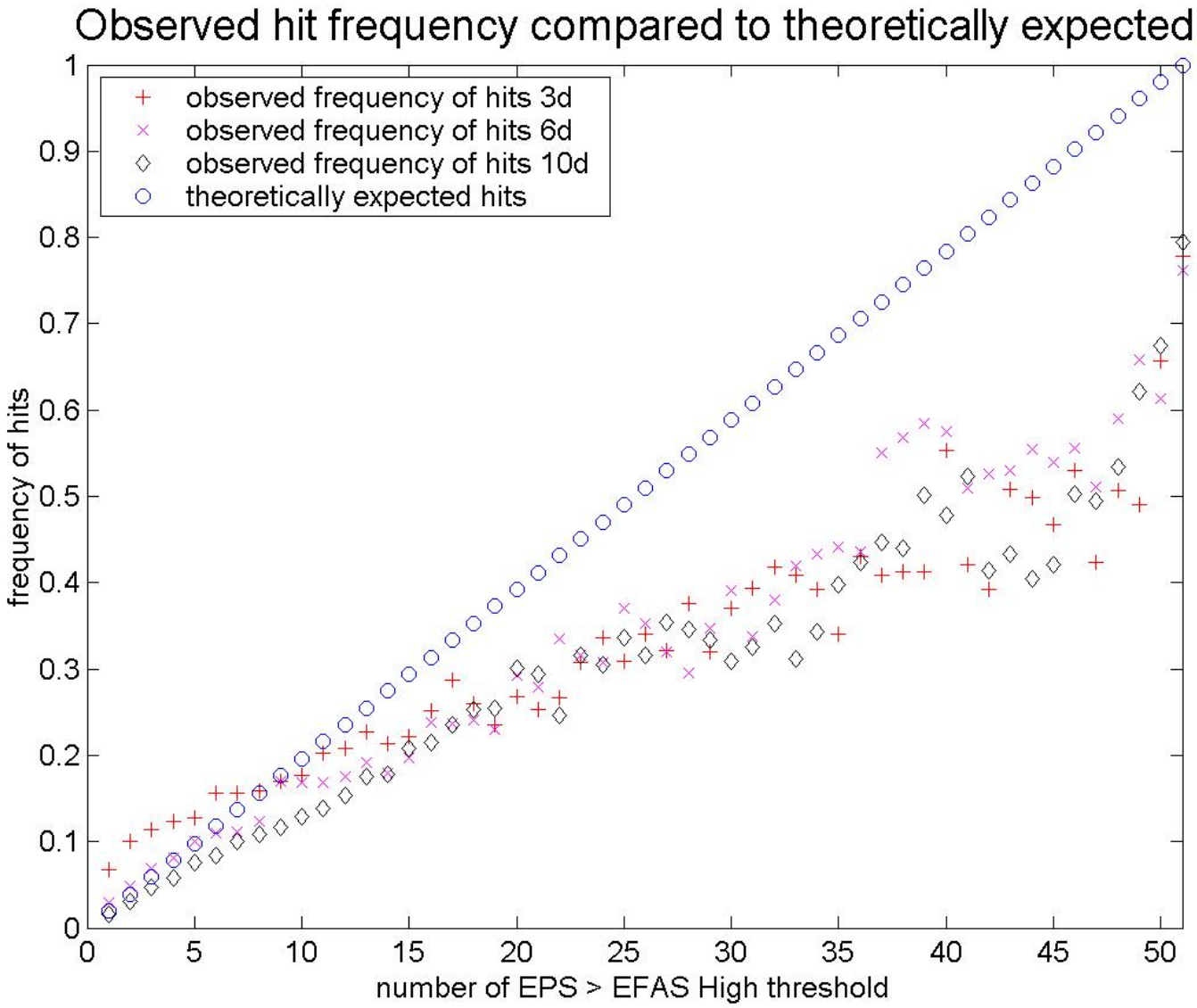


No bias

b4

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Reliability diagram, no pers



Joint Research Centre

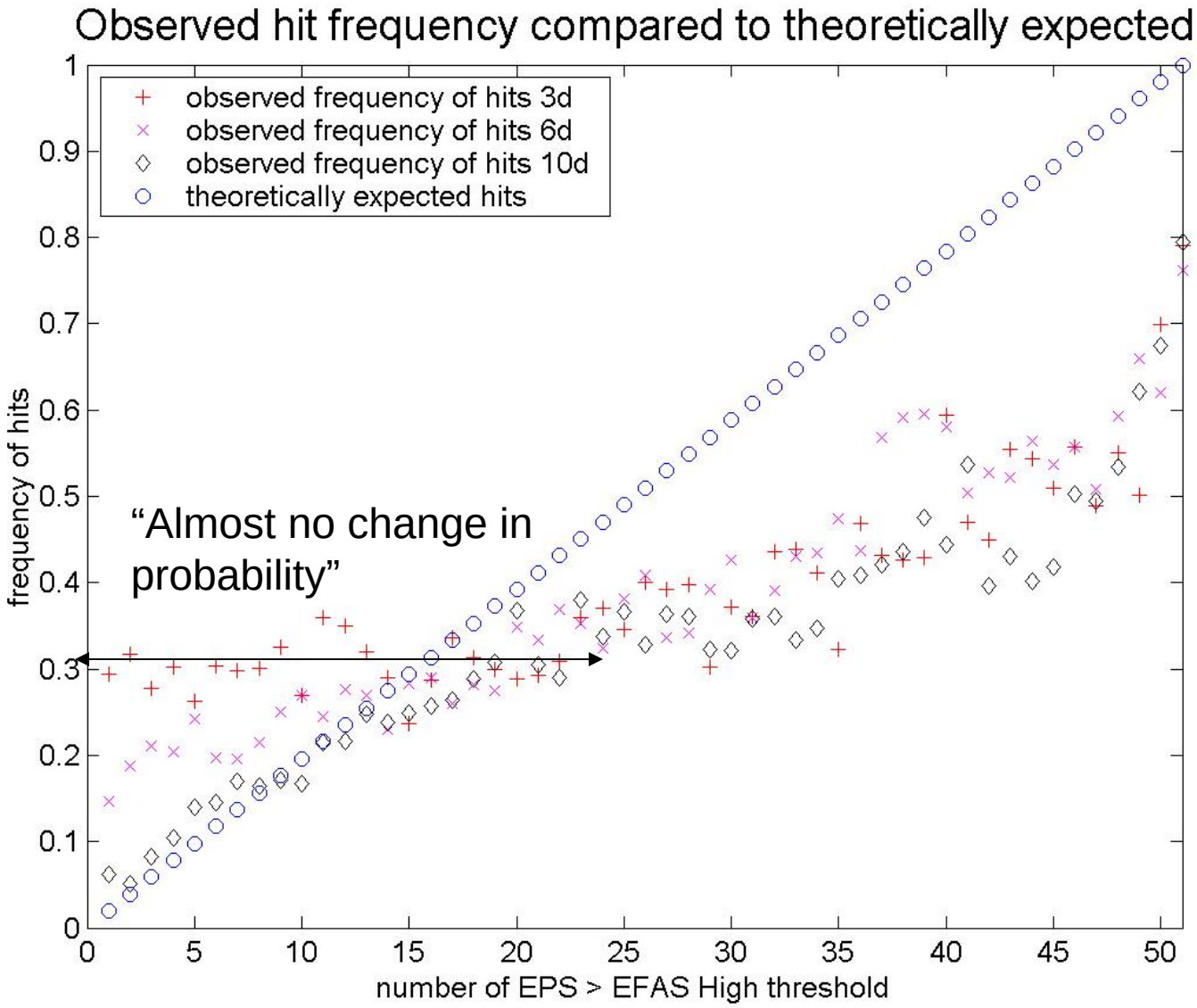
Land Management Unit

Slide 14

b4 no filter for number of EPS over the year etc...
barthje, 5/8/2007

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Reliability diagram last forecast >= 20EPS



Joint Research Centre

Land Management Unit

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Conclusions

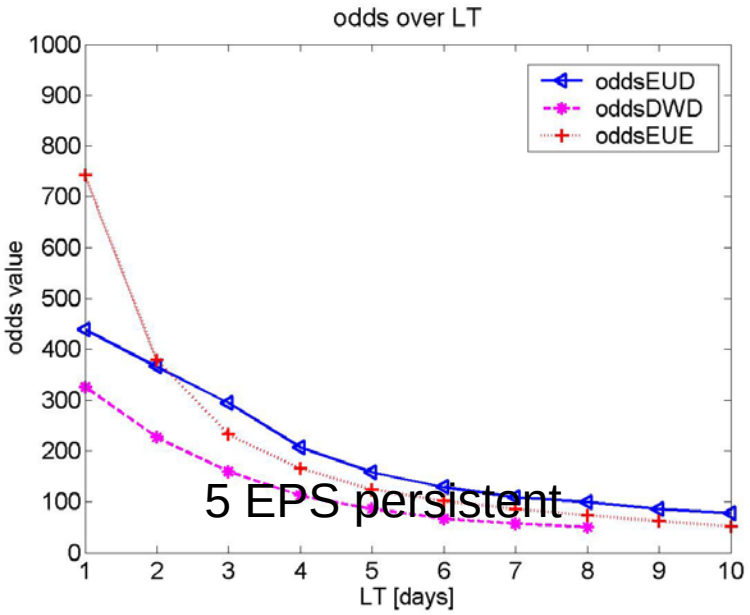
- Positive influence of persistence criterion
- Counter-intuitive BSS behaviour
- Bias to over-forecast
- Probability of previous forecast has high influence on current one.
- No absolute skill measure. Skill depends on trade-offs the end-users have to define.
- For rare events some skill scores are dominated by certain fields of the contingency table
- Better simple score that are easier to interpret

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

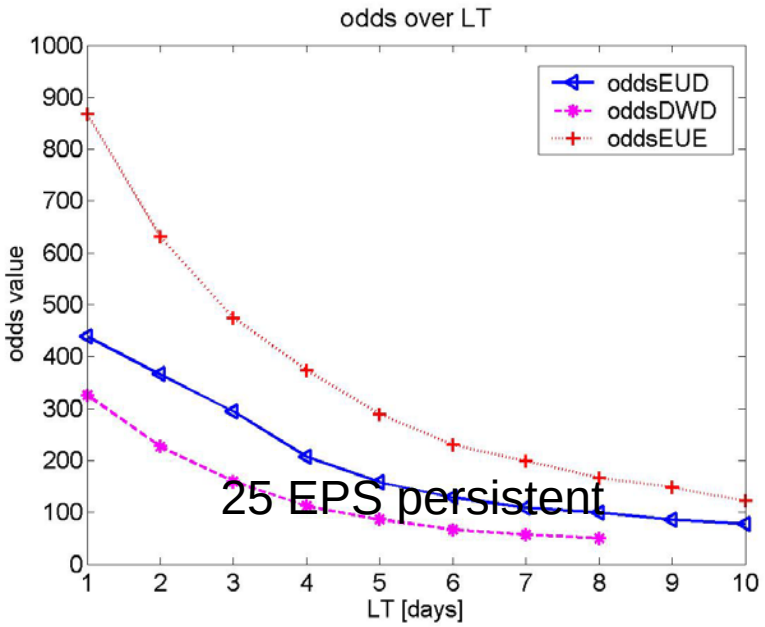
Joint Research Centre

Land
Management Unit

Odds = (hits*posReject) / (false*misses)



5 EPS persistent



25 EPS persistent

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

EFAS Threshold approach

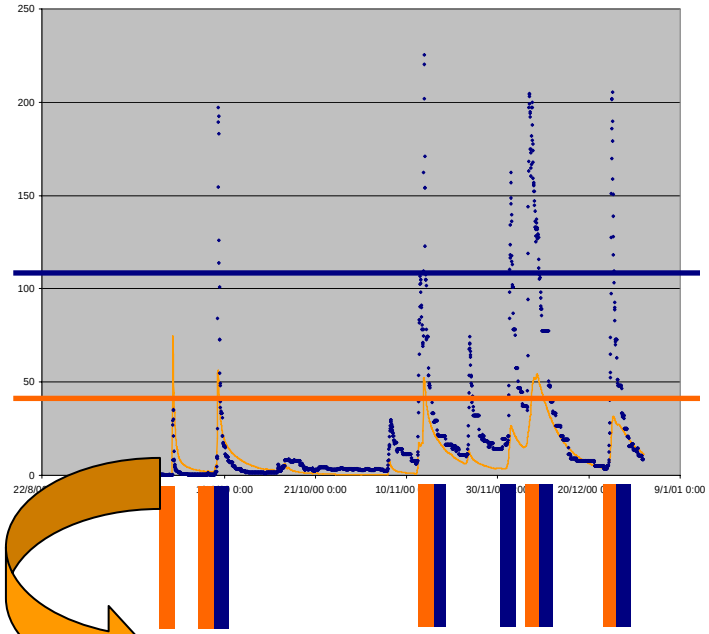
Alarm levels are EFAS / LISFLOOD specific and linked to internally calculated discharges

Joint Research Centre

f(t) (Q_{sim})
(natural)

~~f(t)=F(t)~~

F(t) (Q_{obs})
(regulated)



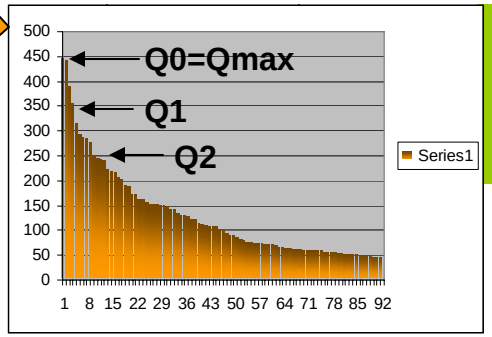
f*(t)=1 : f(t) > $f_{0(f(t))}$

f*(t)=0

f'(t)~F'(t)

F*(t) : F(t) > $F_{0(F(t))}$

F*(t)=0



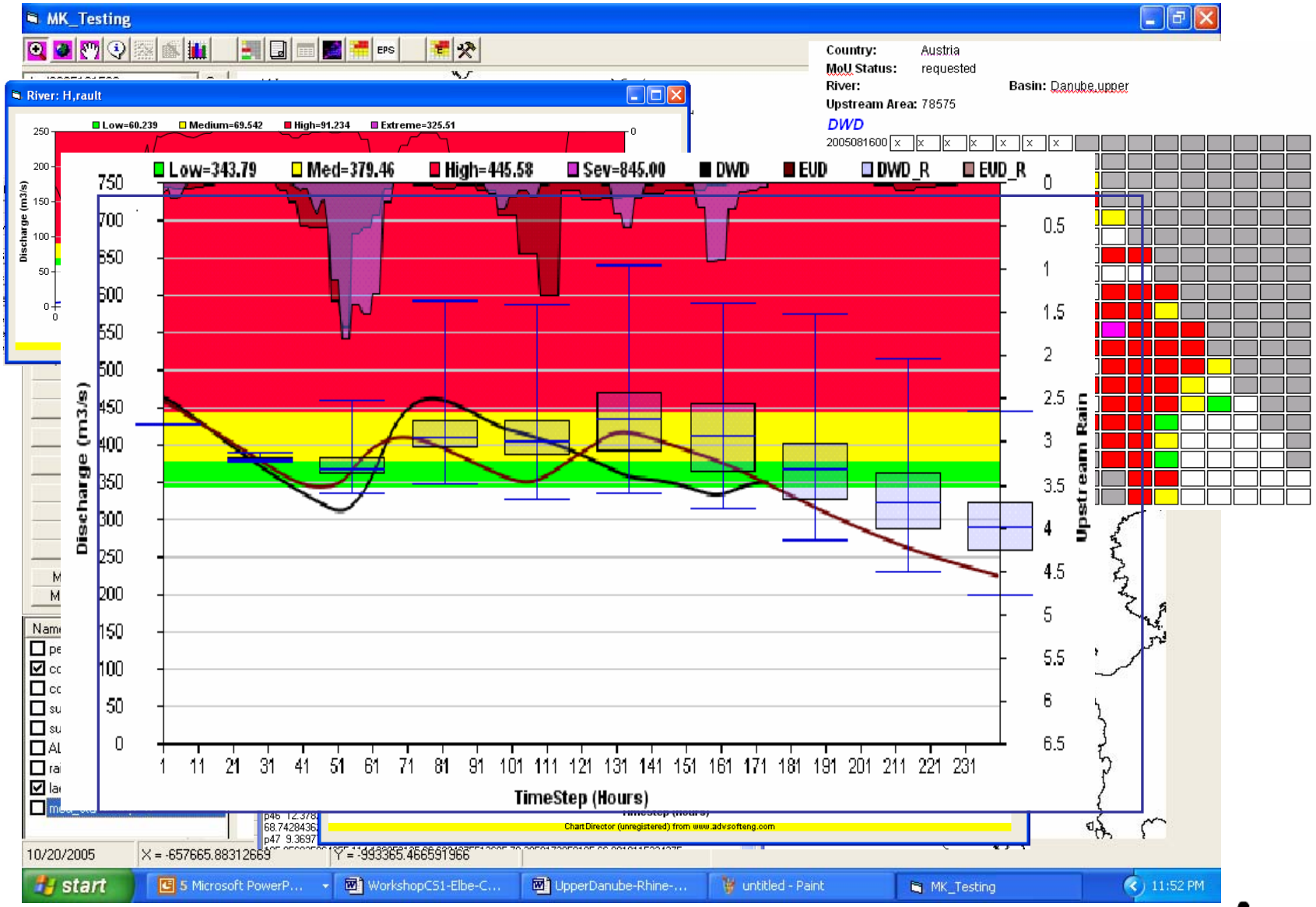
f vs F: 5 missed

f* vs F*: 4 ok, 1 false

Land Management Unit

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre



Land Management Unit

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

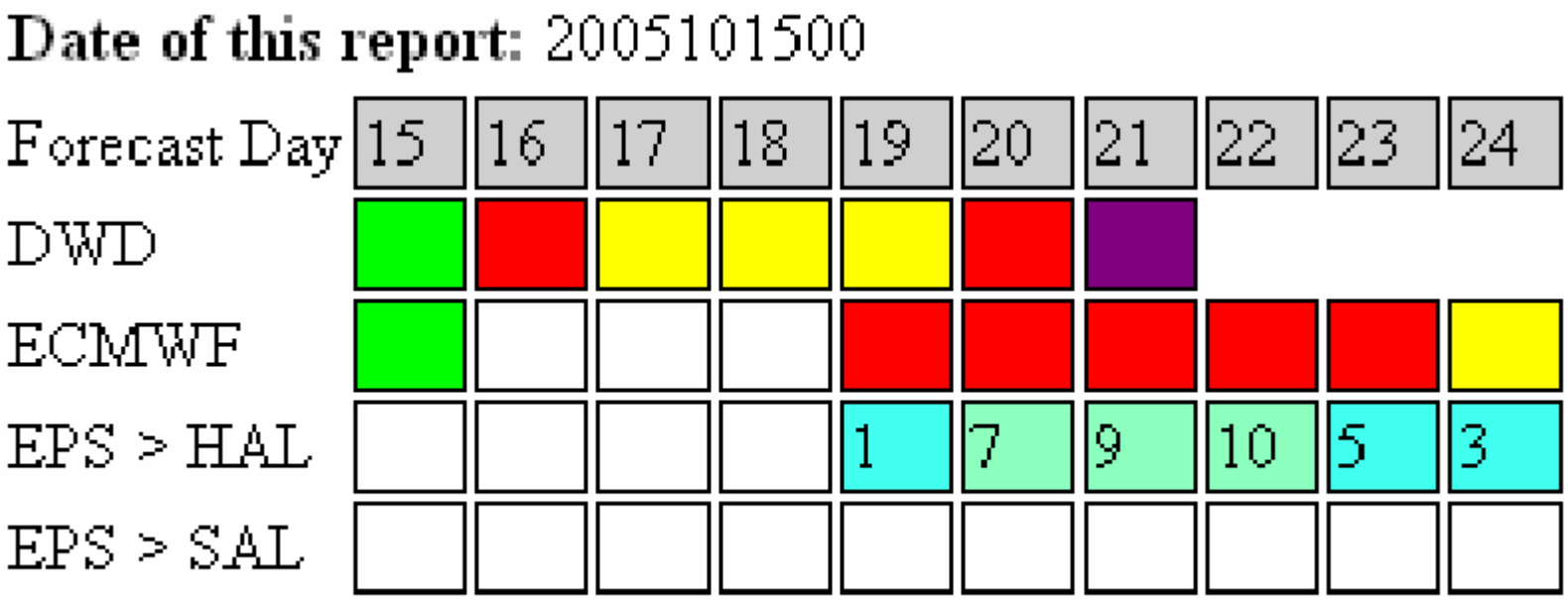
Joint Research Centre

EFAS: User-friendly simplification of complex forecast information

Forecast Day	8	9	10	11	12	13	14	15	16
EPS > HAL		1	8	22	24	22	17	14	5
EPS > SAL									

Land Management Unit

Communication to the end-user



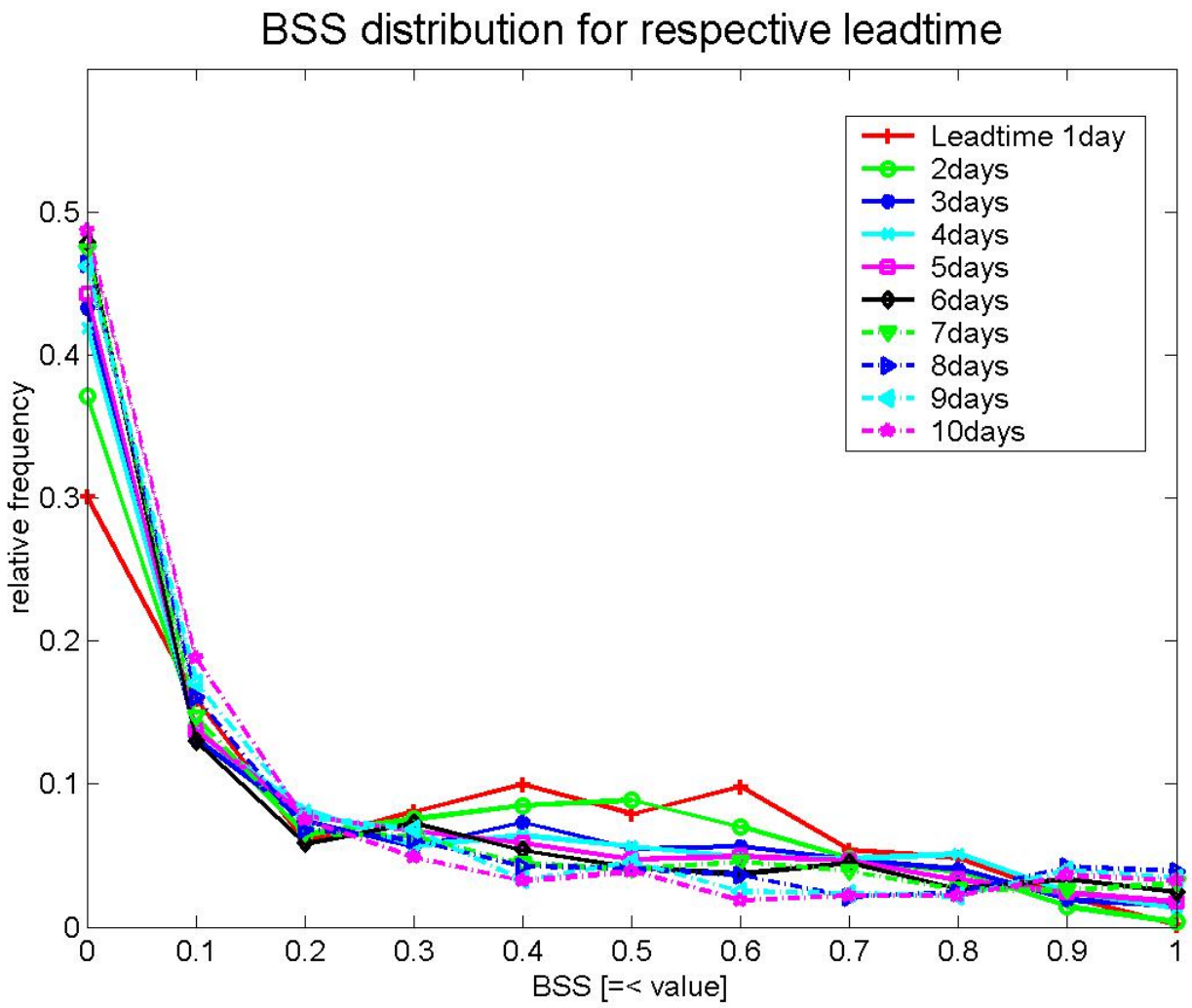
- Others scores like the critical success index CSI (also thread score TS and Gilbert skill score GSS) (Gilbert, 1884, Schaefer, 1990) or Heidke (1926) skill score strongly depend on the frequency of certain (precipitation) events (Ebert and McBride, 1997). Likewise, Wilson (2001) stated that the ETS (Schaefer, 1990) is a reasonable score but that it is not independent from the observed event distribution.

- Influence of persistency on overall skill (BSS). The Higher the persistency threshold the higher average BSS
- Influence of used climatology for BS_{clim} Hamill et al. (2005) have shown that the interpretation of BSS is very sensitive to the choice of reference climatology. Difference between P_{clim}=0.00 for rare events (Legg and Mylne, 2004) and P_{clim}=0.01 (EFAS HAL 1% quantile) exists but is fairly small.

b6

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

BSS clim 1% no pers



Relative frequency of BSS values. Different leadtimes. BSS=0=>no skill, BSS=1=> perfect

Only pixels where something happened (>5 times a forecast or >1 observed) to avoid Finley's Tornado effect

b5

Values (BSS>0.95) still had to be filtered as they are an artefact of f.e. 2 years no event but 6 times forecast 1 EPS (=BSS 0.96)

Joint Research Centre

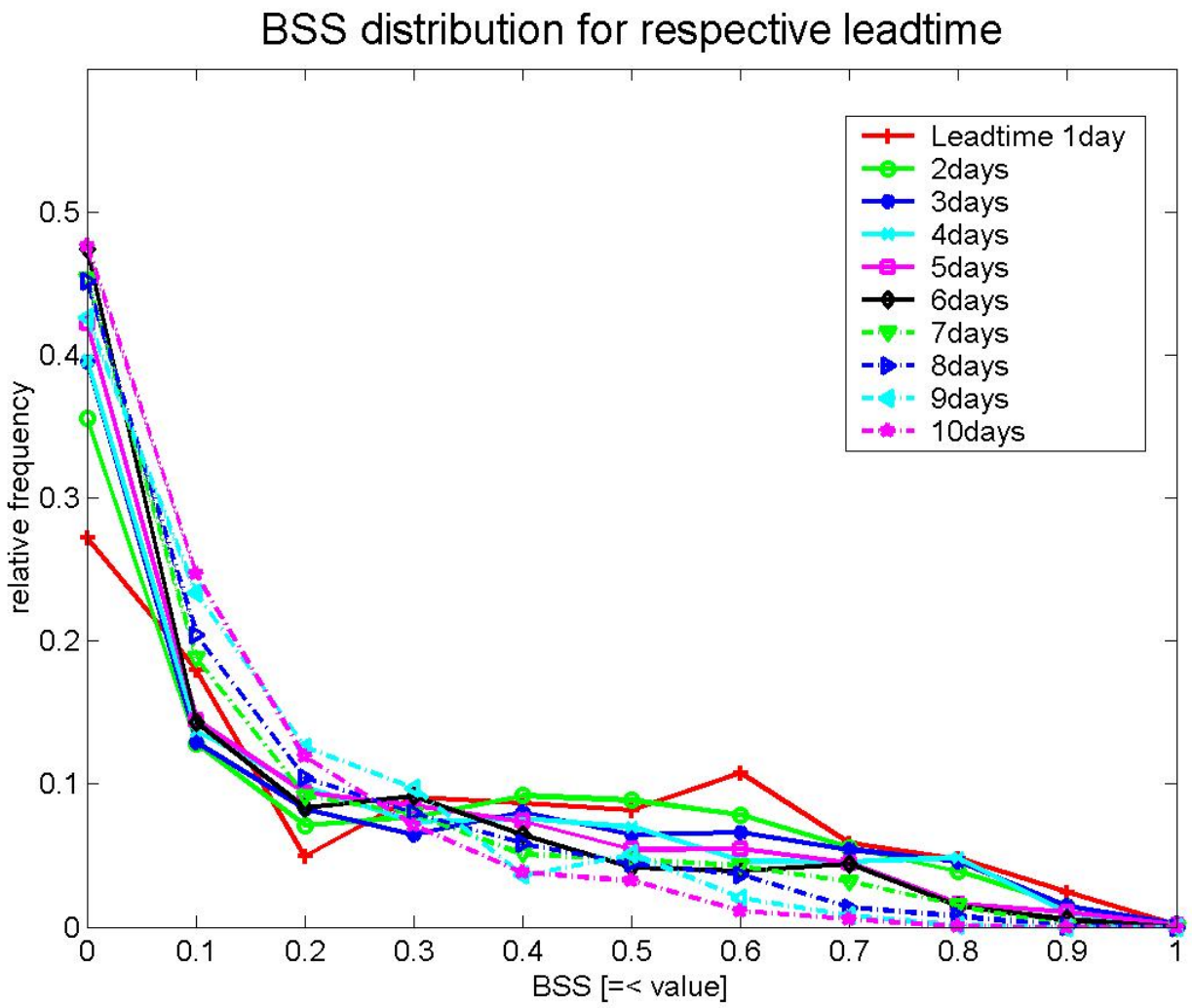
Land Management Unit

Slide 25

b5 threshold of 5: after some analysis of the data to avoid too strong bias from prositive rejection (finley)
barthje, 5/8/2007

b6 majority of curves lies in the area of low-no skill
barthje, 5/8/2007

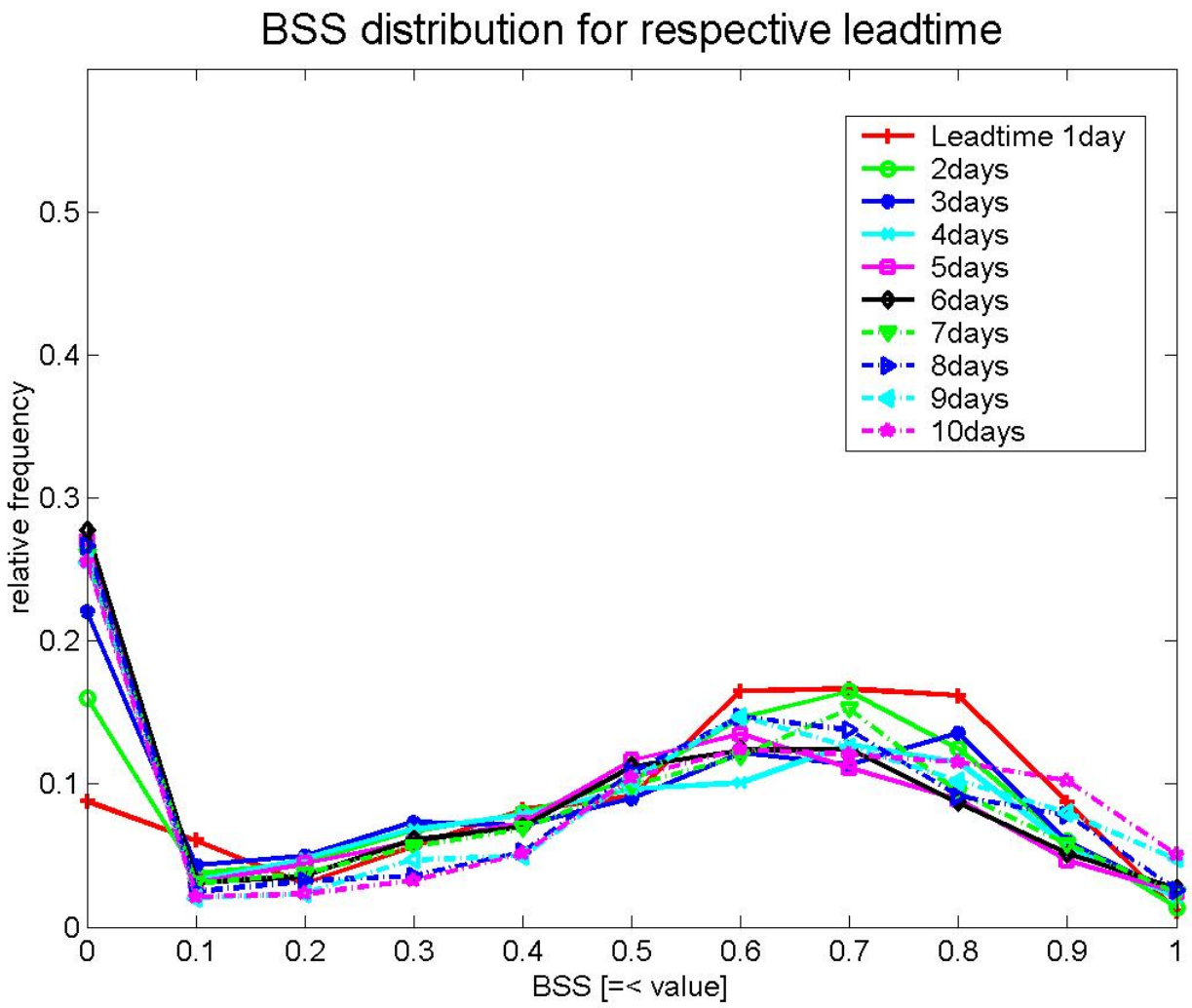
BSS clim 0% no pers



b7

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

BSS clim 1% pers 5EPS



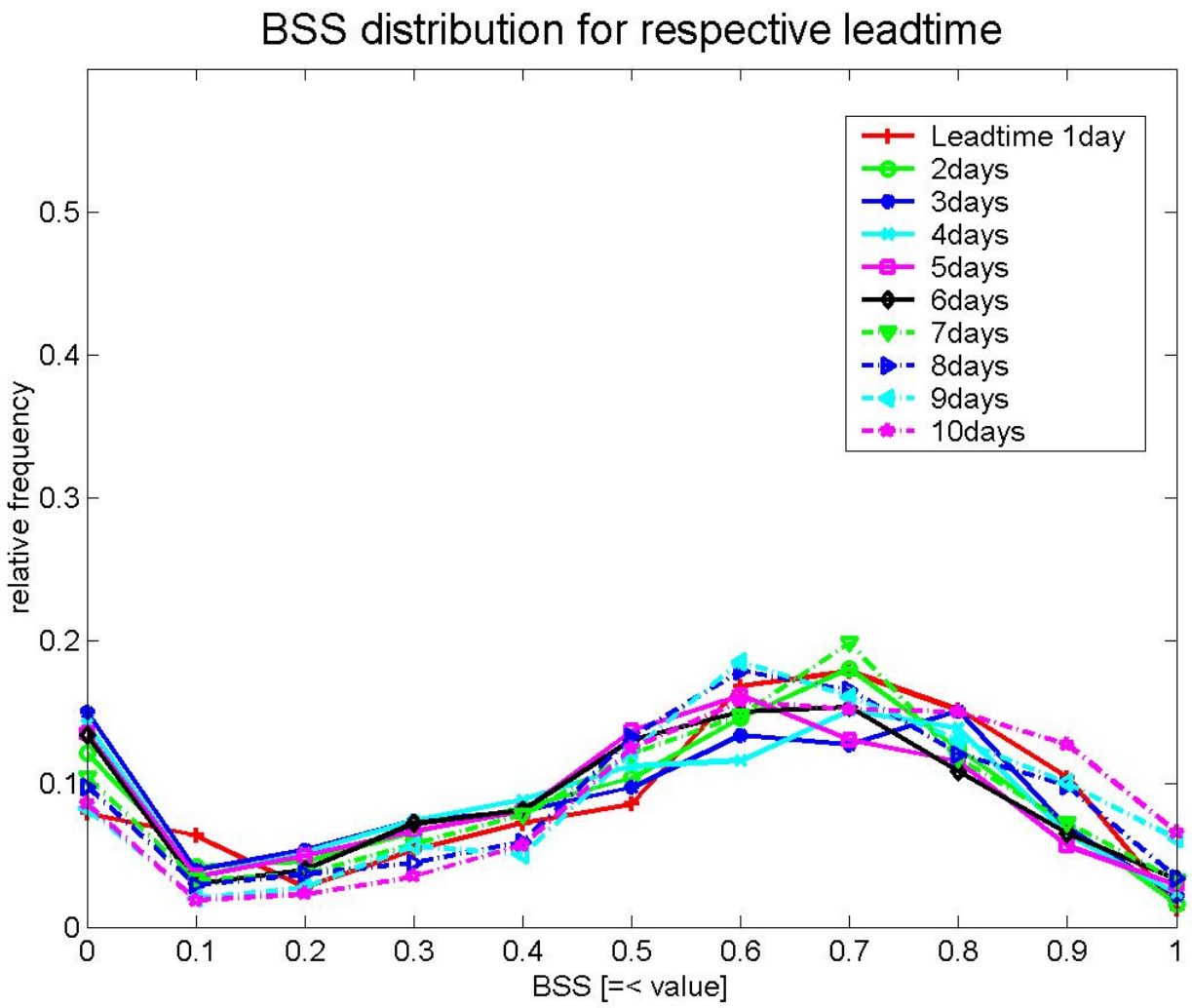
Joint Research Centre

Land Management Unit

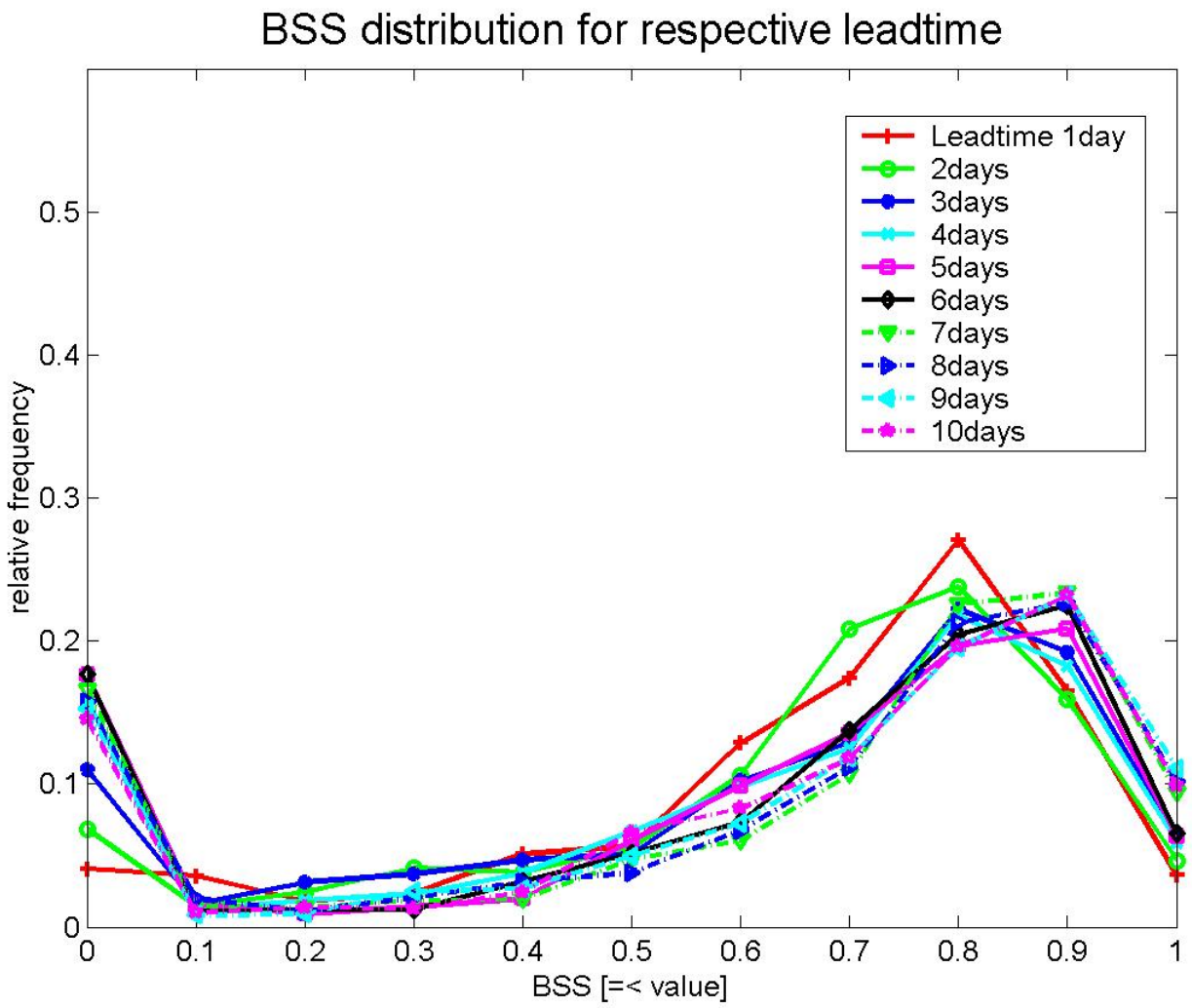
b7 persistence over 2 forecasts, (48 hours)

barthje, 5/8/2007

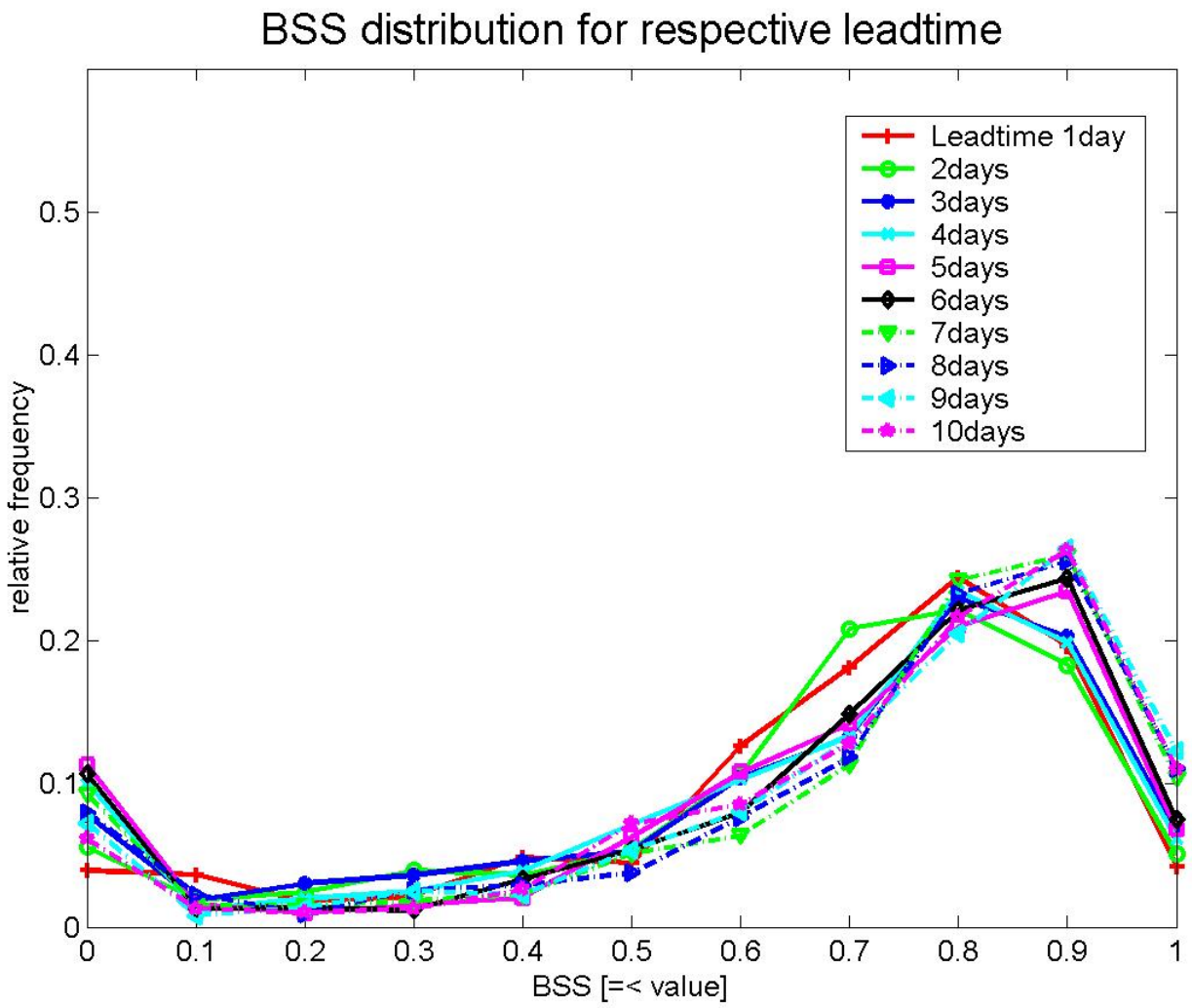
BSS clim 0% pers 5EPS



BSS clim 1% pers 20EPS



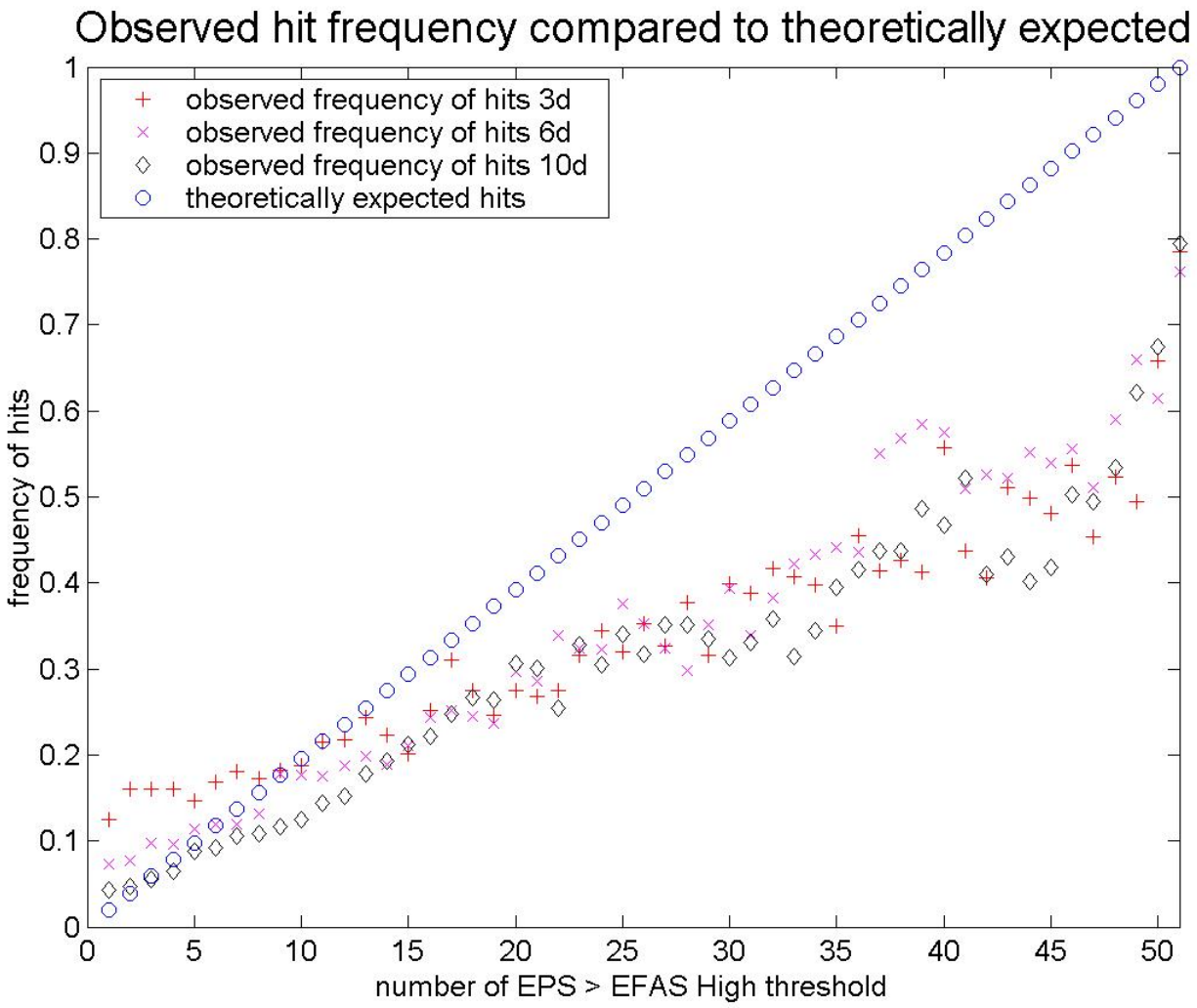
BSS clim 0% pers 20EPS



EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Reliability diagram

last forecast >= 5EPS



Joint Research Centre

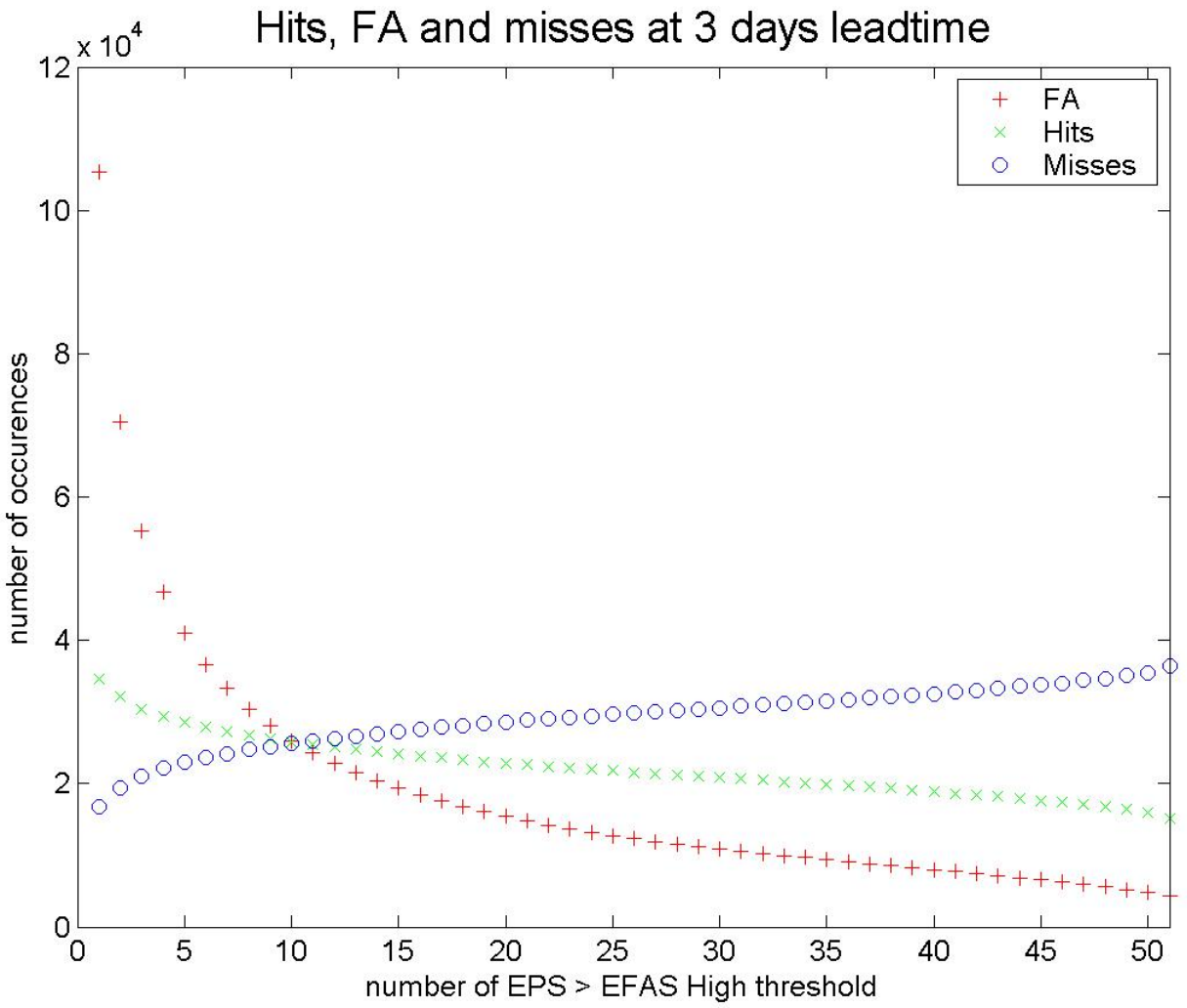
Land Management Unit

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Hits, misses and FA absolute numbers

No pers, 3days

Intersection around 10 EPS: hits=misses=FA



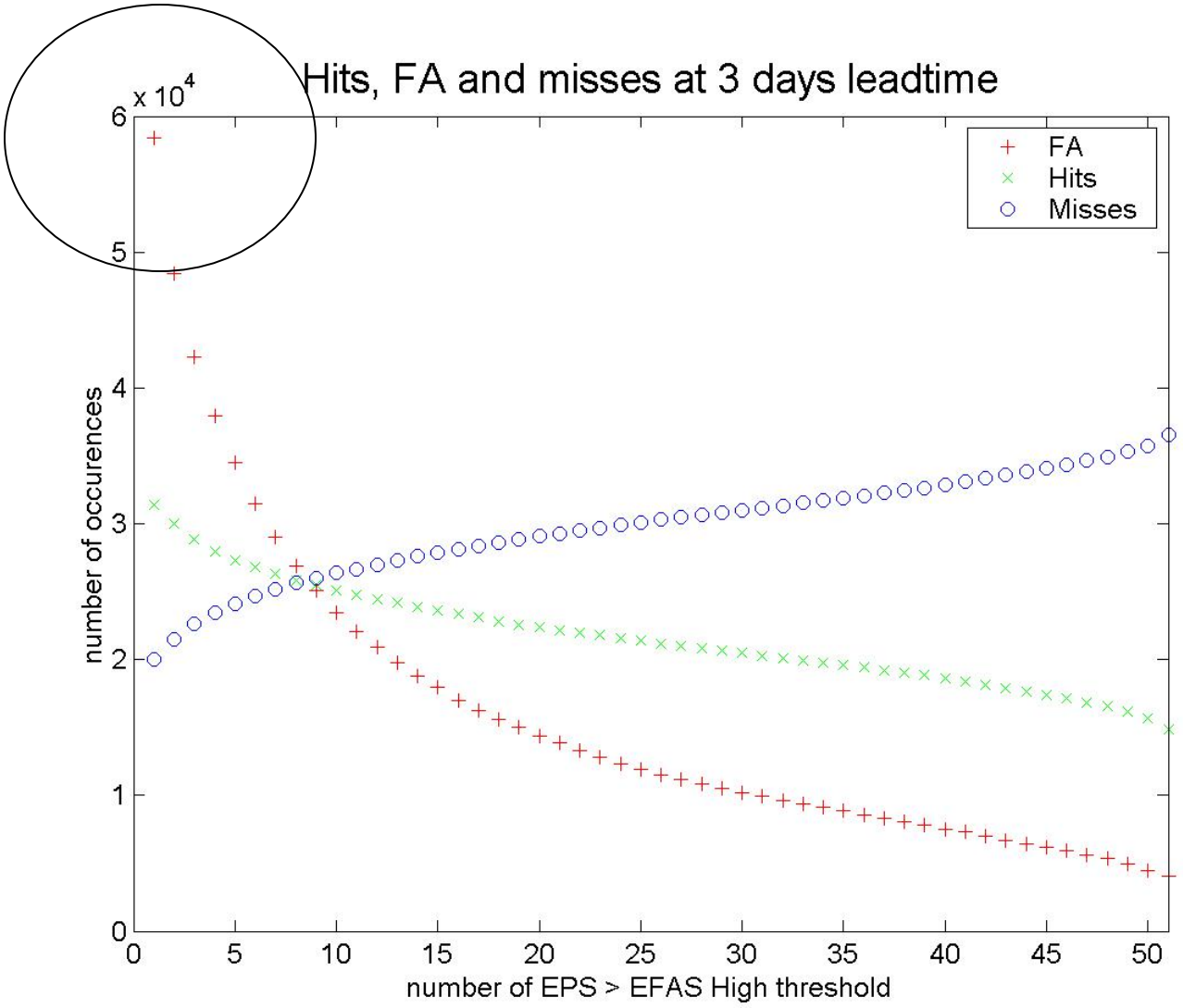
Joint Research Centre

Land Management Unit

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Hits, misses and FA

5EPS pers, 3days



Intersection still around 10 EPS: hits=misses=FA but drastically reduced FA numbers for few EPS

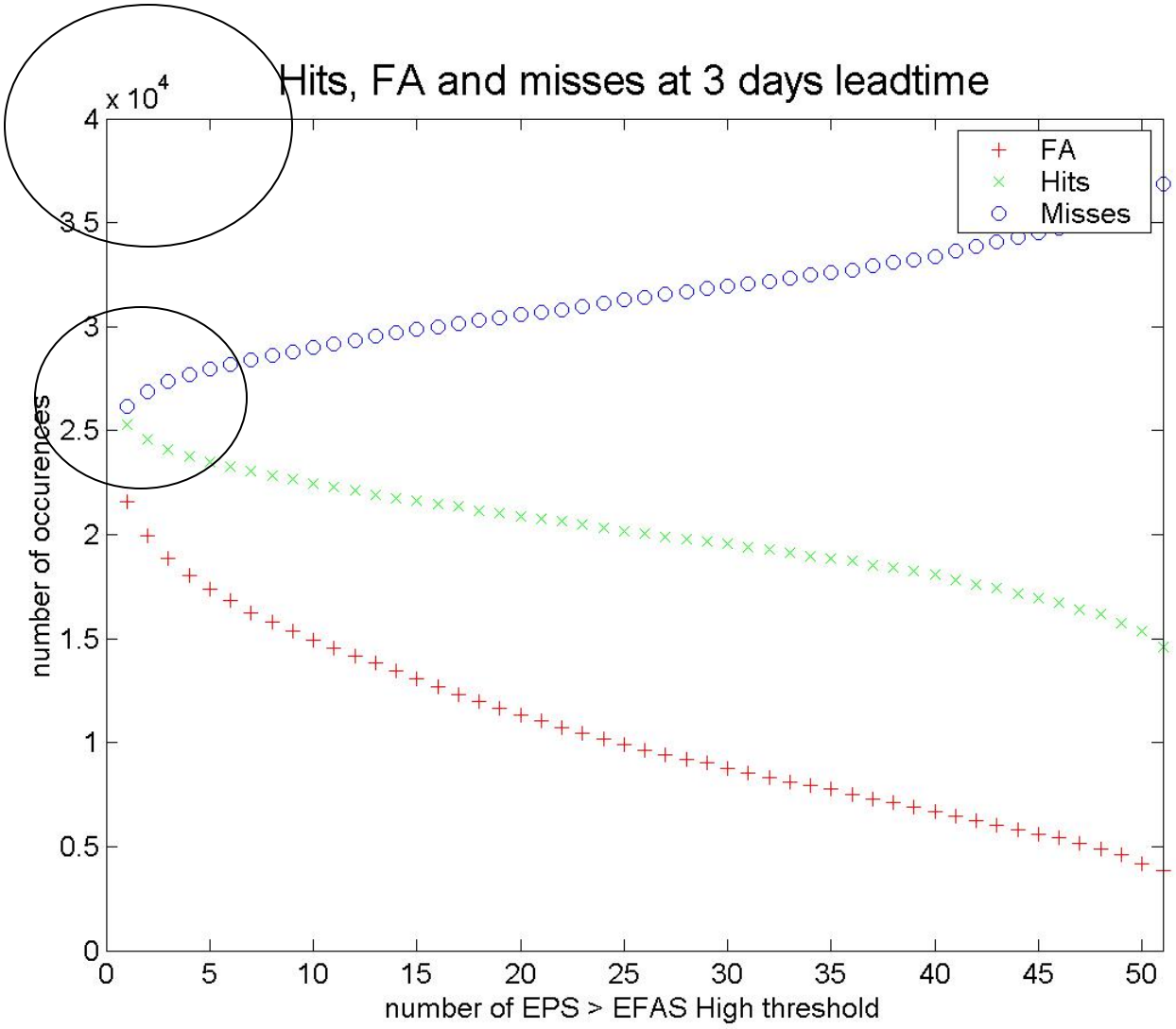
Joint Research Centre

Land Management Unit

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Hits, misses and FA

20EPS pers, 3days



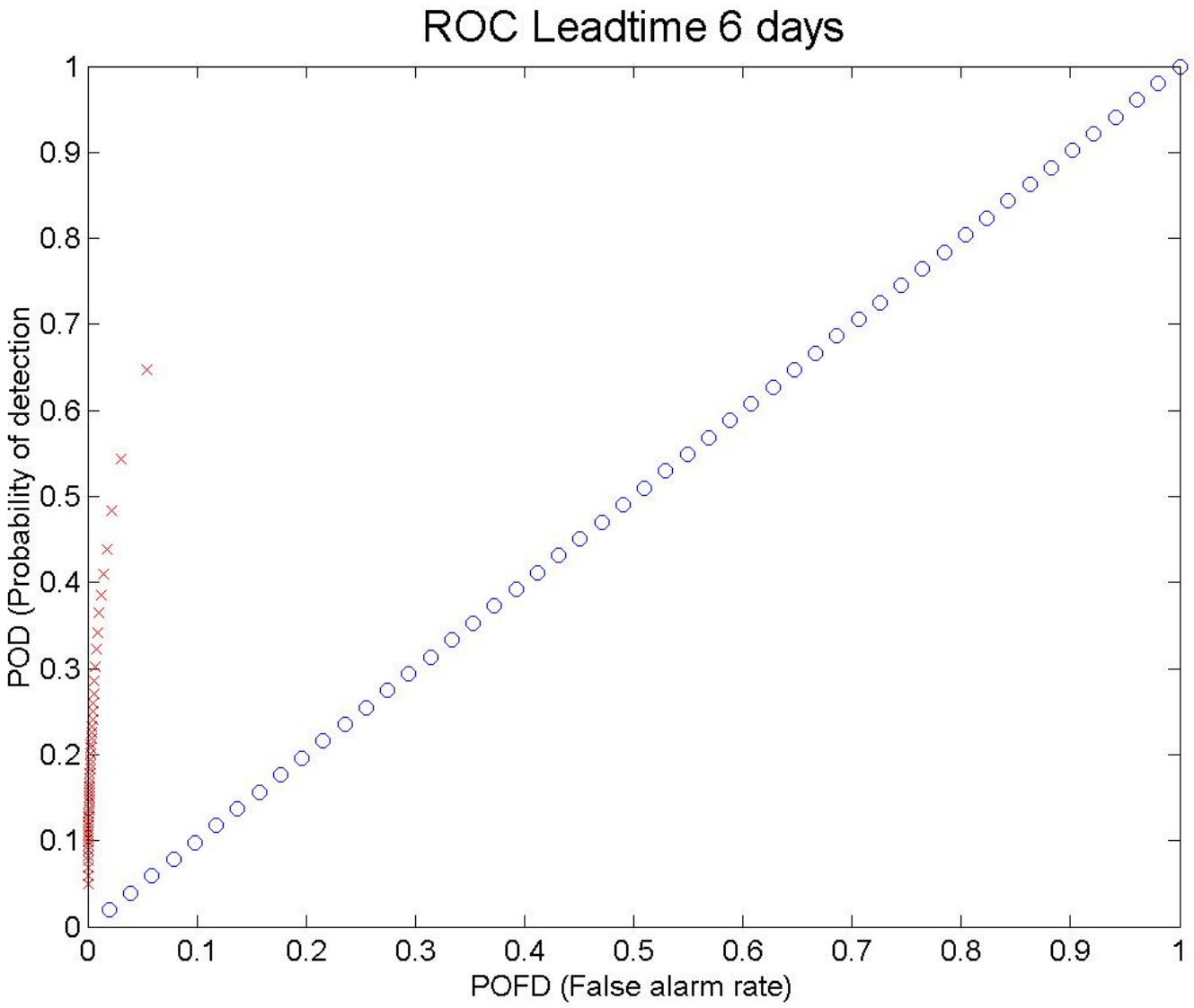
At pers20 leadtime 3days always less FA then hit BUT always more miss than hit, too !

Joint Research Centre

Land Management Unit

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Roc diagram, no pers 6days



Joint Research Centre

Land Management Unit