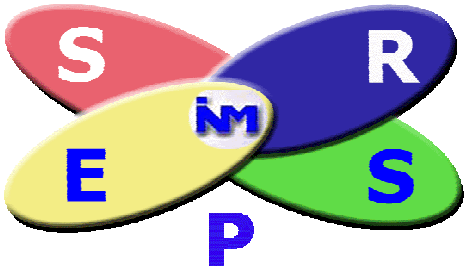


Multi-model Short-Range Ensemble Forecasting at Spanish Met Institute (INM)

J. A. García-Moya

NWP – Spanish Met Service INM

NCEP
September 2006

Acknowledgments

Special thanks to Dr. Eugenia Kalnay for her special support at the beginning of this project

Outline

- Introduction
- EPS for short-range forecast
- SREPS system at INM
- Verification exercise
  - Verification against observations
    - Deterministic scores
    - Probabilistic scores
  - Verification against analysis
    - Deterministic scores
    - Probabilistic scores
- Conclusions

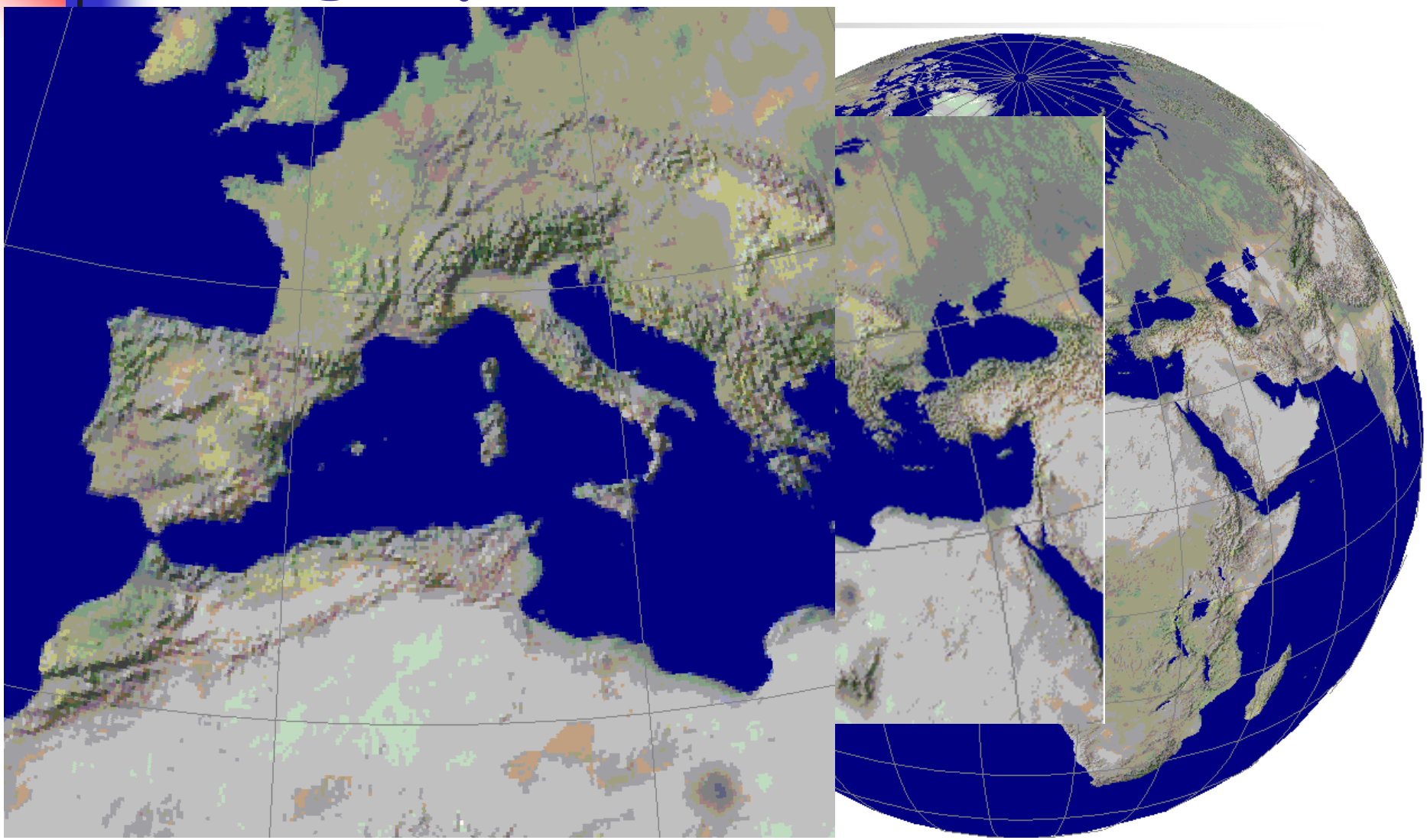
Introduction

- INM is the Spanish Meteorological Institute.
- Headquarter is located in Madrid.
- About 1500 people working for INM.
- It is mostly an operational institute, research is about 10% of resources.

Meteorological Framework

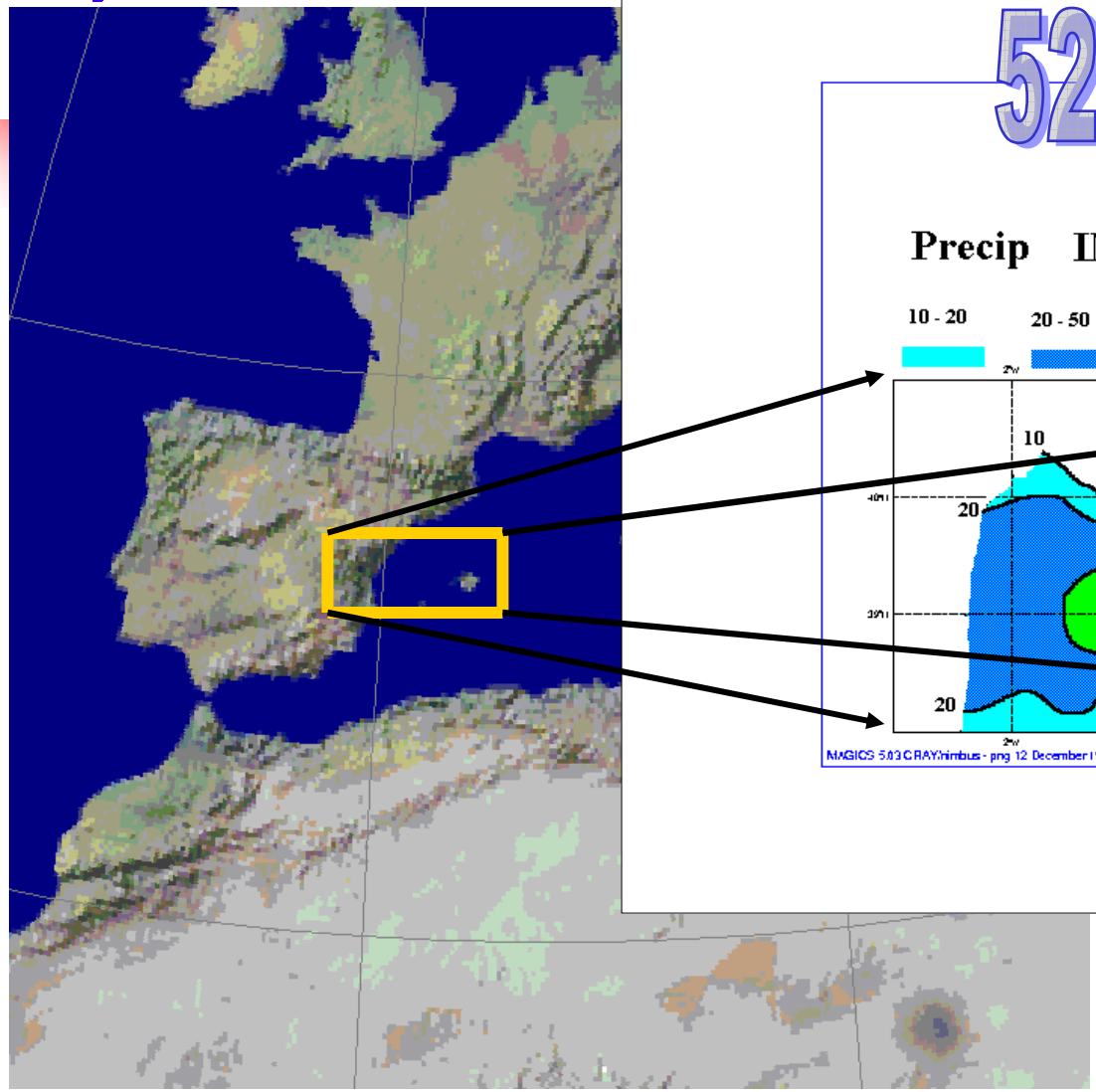
- INM has one CFO (Central Forecast Office) at the HQ, 10 RFO (Regional Forecast Offices) and one AFO (Forecast Office for the Army).
- Main Weather Forecast issues are related with Short-Range forecast of extreme events.
- Convective precipitation is the most dangerous weather event in Spain.

Geographical Framework

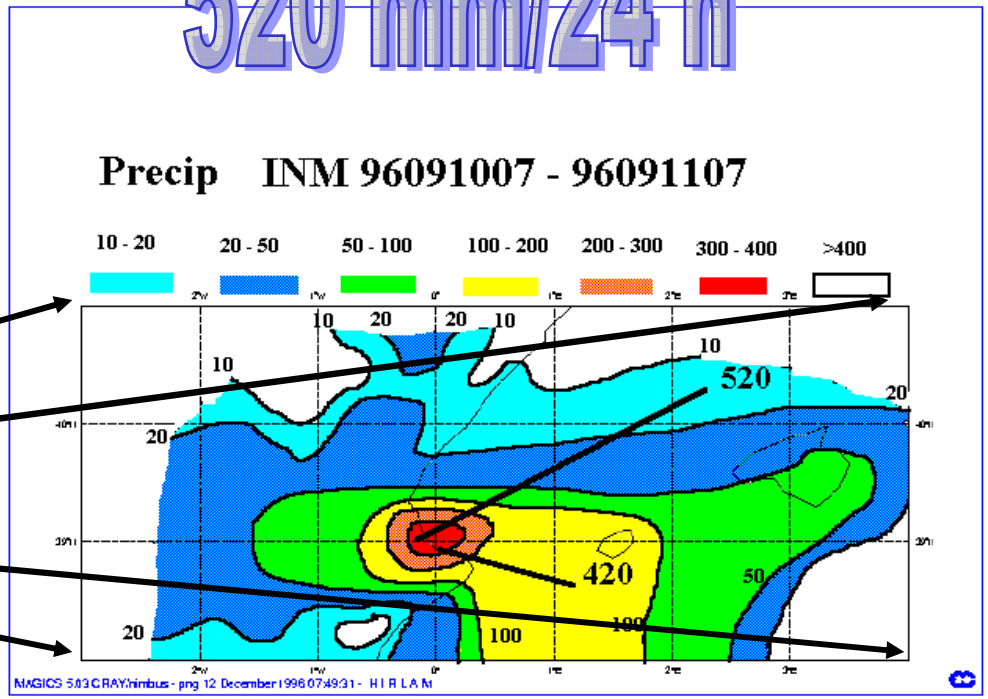


Geographical Framework

- Western Mediterranean is a close sea rounded by high mountains.
- In autumn sea is warmer than air.
- Several cases of more than 200 mm/few hours occurs every year.
- Some fast cyclogenesis like "tropical cyclones" also appears from time to time.

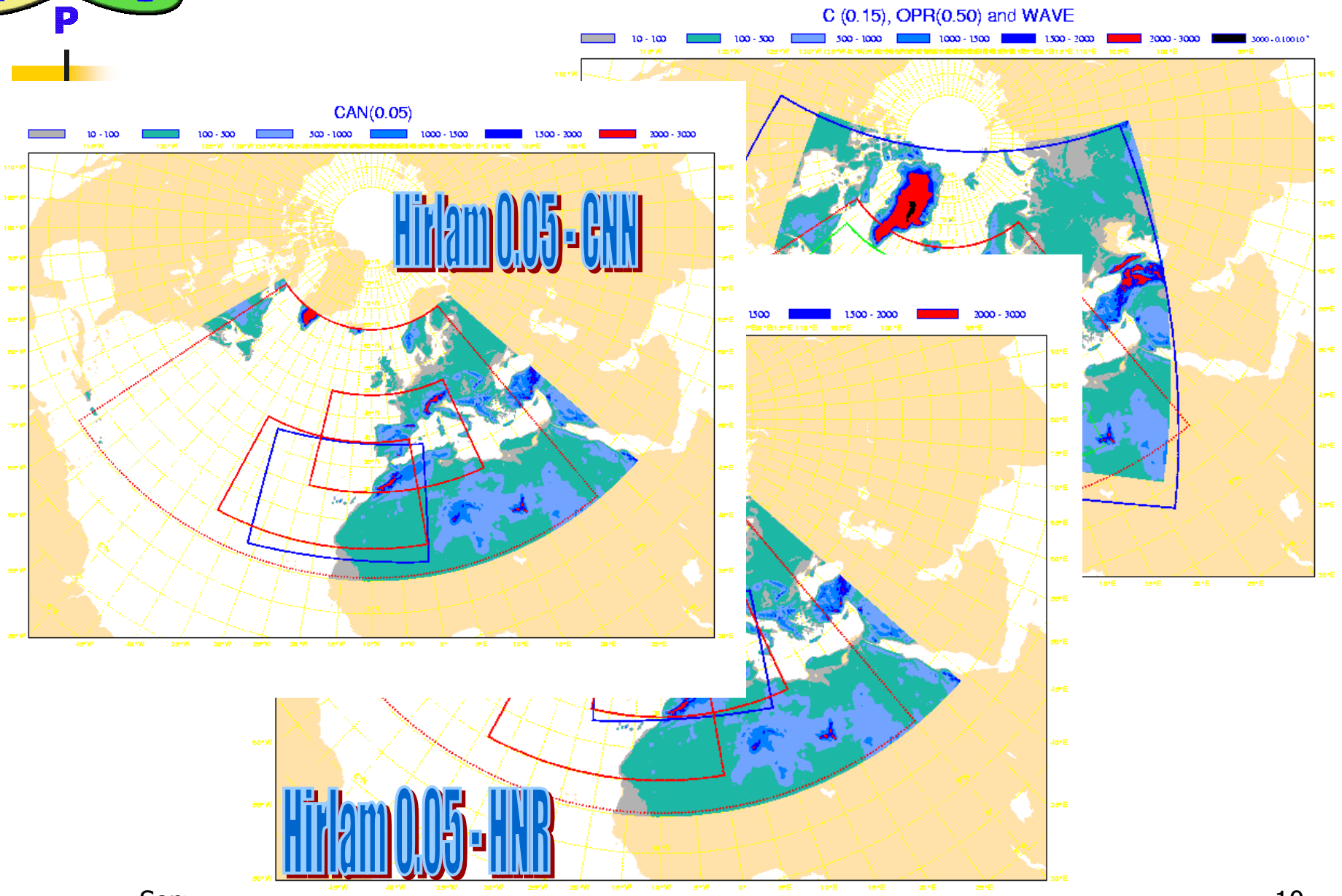


520 mm/24 h


Precip INM 96091007 - 96091107
10 - 20
20 - 50
50 - 100
100 - 200
200 - 300
300 - 400
>400
520
420
100
50
20
10

Numerical Modelling

- INM use numerical model Hirlam for operational NWP.
- **Hirlam** (http://hirlam.org) is an European Project (8 countries) to develop operational NWP models.
- Short range forecast up to 72 hours every 6 hours (00, 06, 12 and 18 UTC).
- Two different horizontal resolutions 0.16 and 0.05 deg. latxlon. 40 vertical levels.


C (0.15), OPR(0.50) and WAVE
10 - 100
100 - 300
300 - 1000
1000 - 1500
1500 - 2000
2000 - 3000
3000 - 0.100 10 *
CAN(0.05)
10 - 100
100 - 300
300 - 1000
1000 - 1500
1500 - 2000
3000 - 3000
Hirlam 0.05 - CNN
1500
1500 - 2000
3000 - 3000
Hirlam 0.05 - HNR

HP Computer Cray X1e

- 16 nodes, 8 MSP's each (~2.4 Tf peak perf.)
  - Deterministic Forecast
  - SREPS
  - Climatic runs

Ensemble for Short Range

- Surface parameters are the most important ones for weather forecast.
- Forecast of extreme events (convective precip, gales,…) is probabilistic.
- Short Range Ensemble prediction can help to forecast these events.
- Forecast risk (Palmer, ECMWF Seminar 2002) is the goal for both Medium- and, also, Short-Range Prediction.

Errors of the short-range forecast

- Due to model formulation.
- Due to simplifications in parameterisation schemes.
- Due to uncertainty in the initial state.
- Special for LAMs, due to errors in lateral boundary conditions.
- Due to uncertainties in soil fields (soil temperature and soil water content, …).

SREPS I

- Multi-model approach (Hou & Kalnay 2001).
- Stochastic physics (Buizza et al. 1999).
- Multi-boundaries:
  - From few global deterministic models.
  - From global model EPS (ECMWF).
  - SLAF technique (Hou & Kalnay 2001).

SREPS II

- Different assimilation techniques:
  - Optimal Interpolation.
  - Variational (3D or 4D).
- Perturbed analysis:
  - Singular vectors (ECMWF, Palmer et al. 1997).
  - Breeding (NCEP, Toth & Kalnay 1997).
  - Scaled Lagged Average Forecast (SLAF, Hou & Kalnay 2001).

Multi-model

- Hirlam (http://hirlam.org).
- HRM from DWD (German Weather Service).
- MM5 (http://box.mmm.ucar.edu/mm5/).
- UM from UKMO (Great Britain Weather Service).
- LM (Lokal Model) from COSMO consortium (http://www.cosmo-model.org).

Multi-Boundaries

From different global deterministic models:

- ECMWF
- UM from UKMO (Great Britain Weather Service)
- AVN from NCEP
- GME from DWD (German Weather Service)

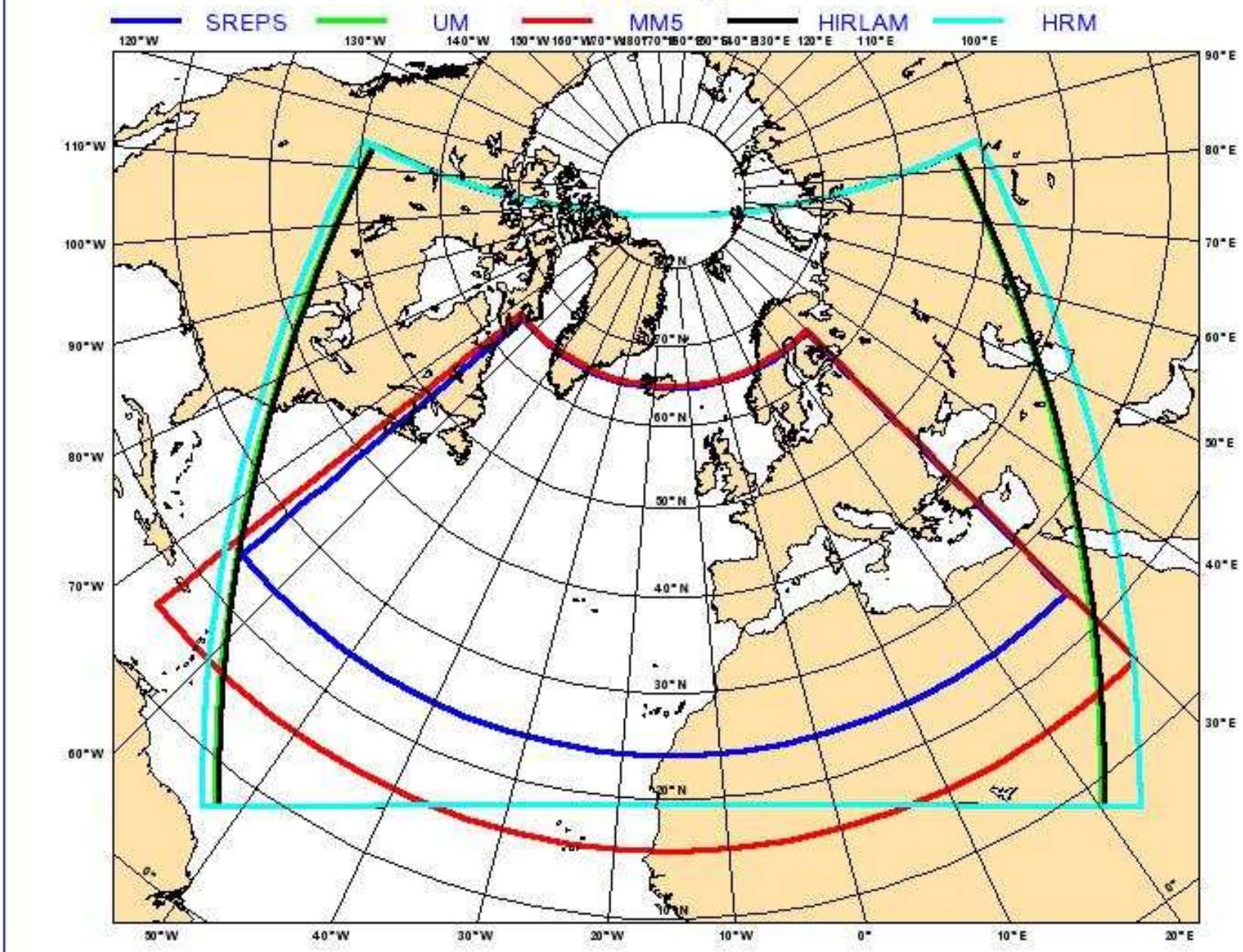
SREPS at INM

- Mummub: Multi-model Multi-boundaries
- 72 hours forecast four times a day (00, 06, 12 y 18 UTC).
- Characteristics:
  - 5 models.
  - 4 boundary conditions.
  - 4 latest ensembles (HH, HH-6, HH-12, HH-18).
- 20 member ensemble every 6 hours
- Time-lagged Super-Ensemble of 80 members every 6 hours.

Current Ensemble

- Mummub: Multi-model Multi-boundaries
- 72 hours forecast twice a day (00 & 12 UTC).
- Characteristics:
  - 5 models.
  - 4 boundary conditions.
- 20 member ensemble every 12 hours

Coverage

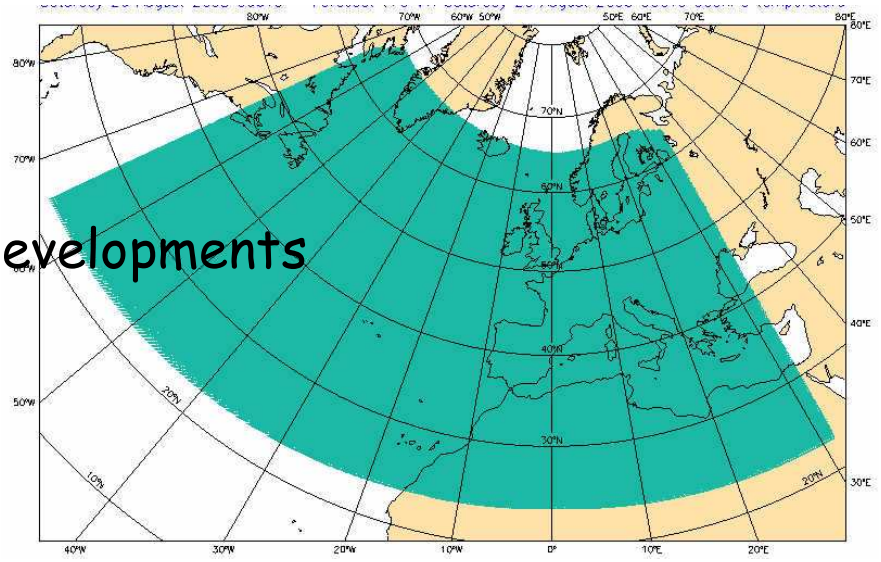

SREPS
UM
MM5
HIRLAM
HRM
120°W
130°W
140°W
110°W
100°W
90°W
80°W
70°W
60°W
90°E
80°E
70°E
60°E
50°E
40°E
30°E
50°N
40°N
30°N
20°N
10°N
50°W
40°W
30°W
20°W
10°W
0°
10°E
20°E

Road Map

2003-2004	Research to find best ensemble for the Short Range	
Jun 04 – Jun 05	Building Multimodel System	
Jun 05-Dec 05	Mummub n/16 members	Daily run non-operational
Mar 06	Mummub 16/16 members	Once a day
Jun 06	Mummub 20 members	Twice a day
July 06	Obs verfication	
September 06	**40 member lagged Super-ensemble**	**Twice a day**
October 06	BMA Calibration	
January 07	Broadcast products	Experimental

Post-processing

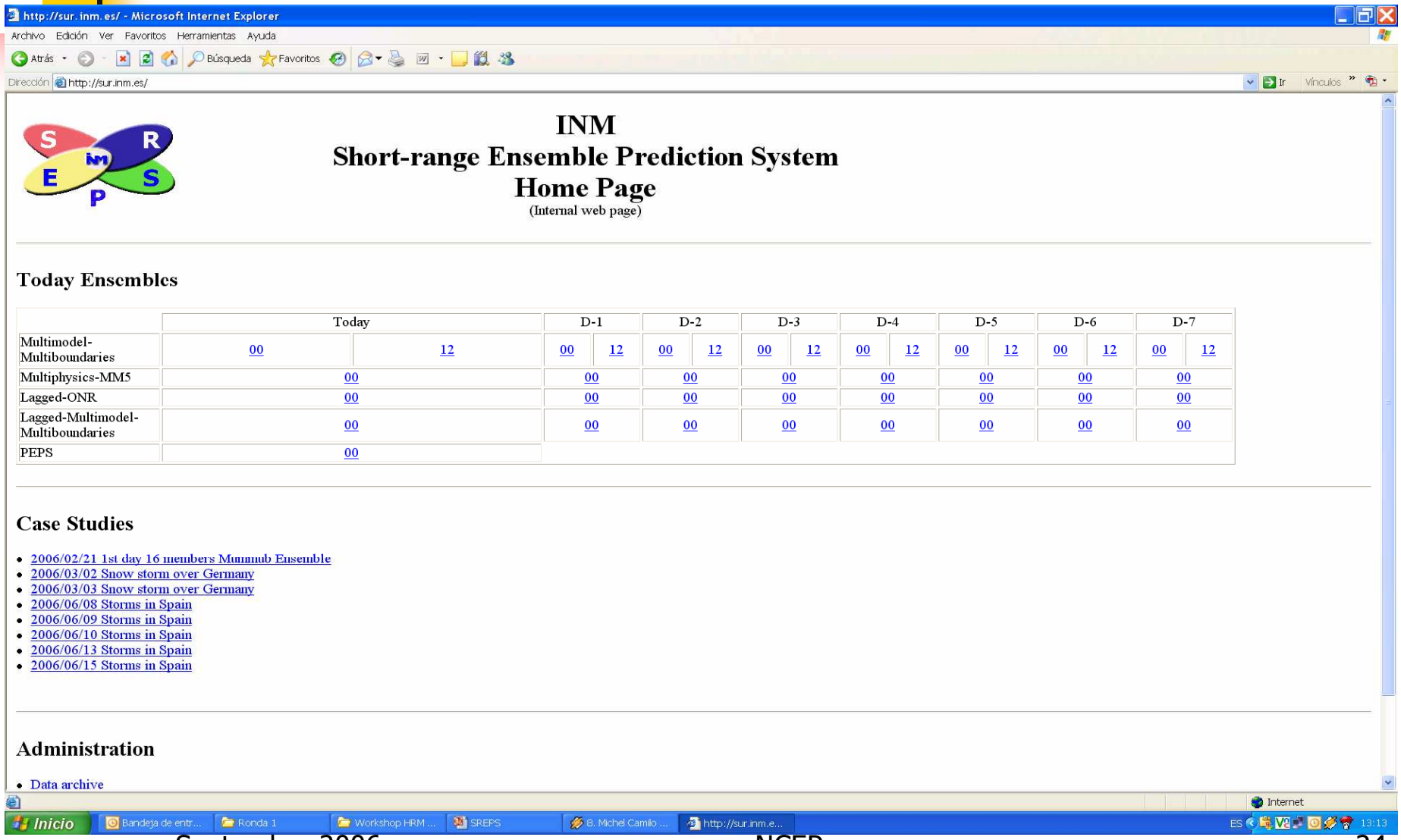
- Integration areas 0.25 latxlon, 40 levels
- Interpolation to a common area
  - ~ North Atlantic + Europe
  - Grid 380x184, **0.25°**
- Software
  - Enhanced PC + Linux
  - ECMWF Metview + Local developments
- Outputs
  - Deterministic
  - Ensemble probabilistic



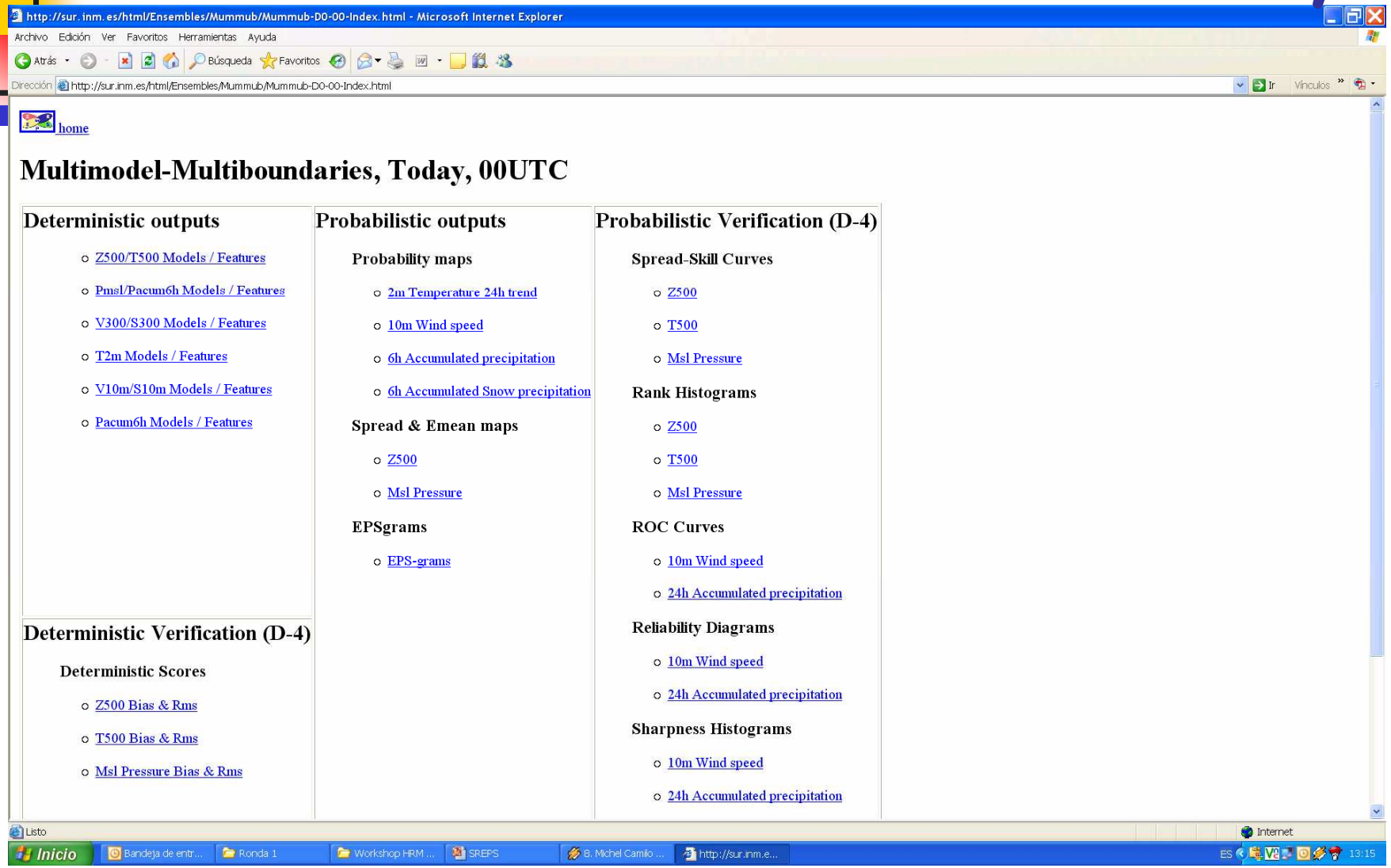
Monitoring in real time

- Intranet web server
- Deterministic outputs
  - Models X BCs tables
  - Maps for each couple (model,BCs)
- Ensemble probabilistic outputs
  - **Probability maps**: 6h accumulated precipitation, 10m wind speed, 24h 2m temperature trend
  - Ensemble mean & Spread **maps**
  - **EPSgrams** (work in progress)
- Verification: Deterministic & Probabilistic
  - Against ECMWF analysis
  - Against observations

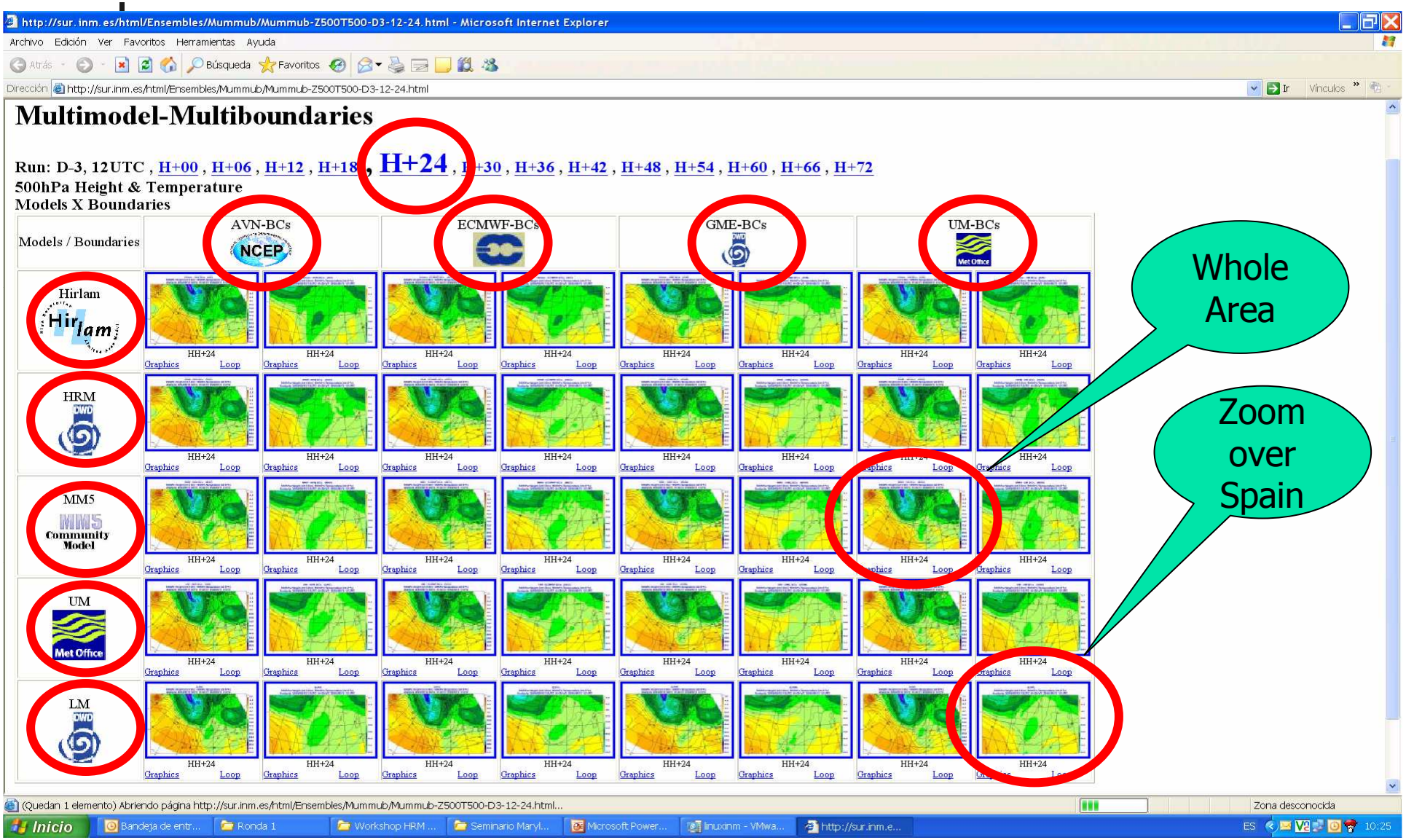
Intranet web server



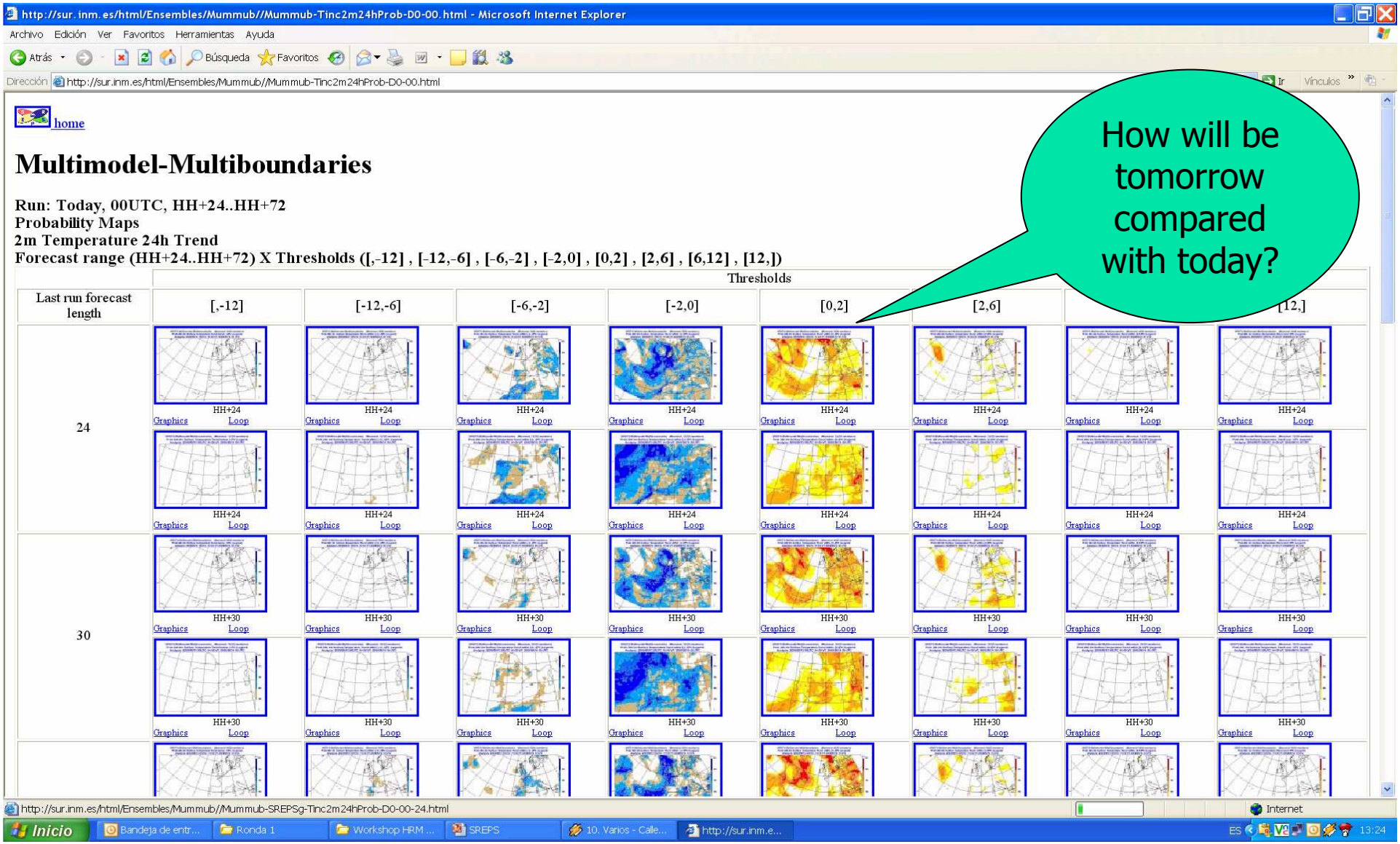
Monit 1: multi-boundaries today



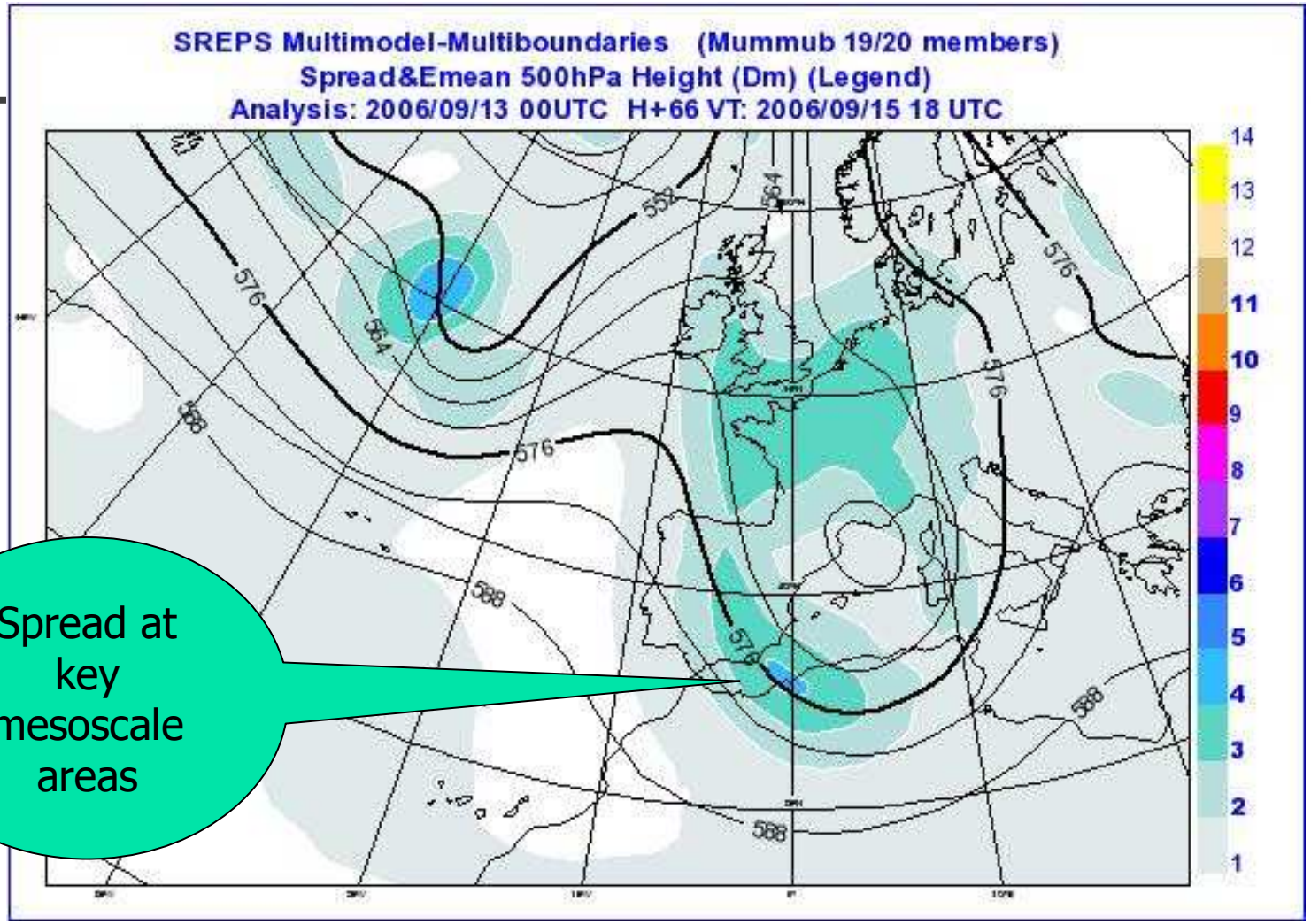
Monit 2: all models X bcs



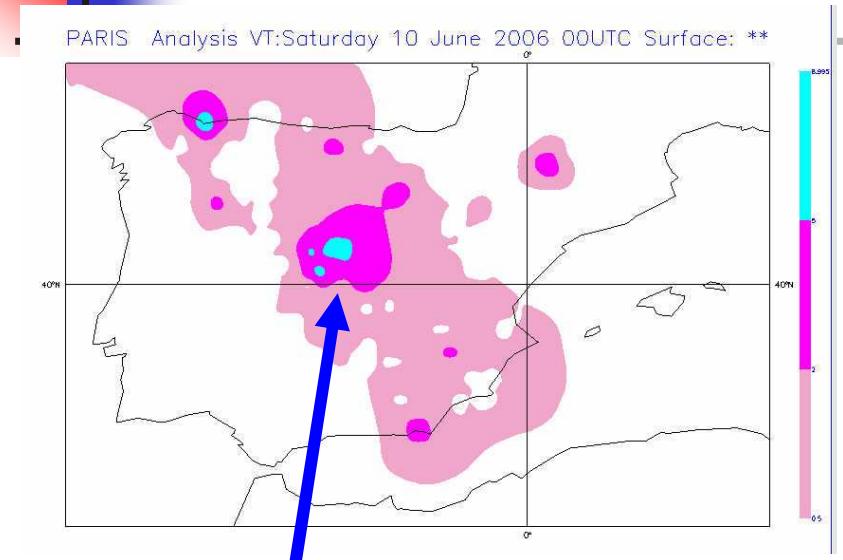
Monit 3: Prob 24h 2m T trend

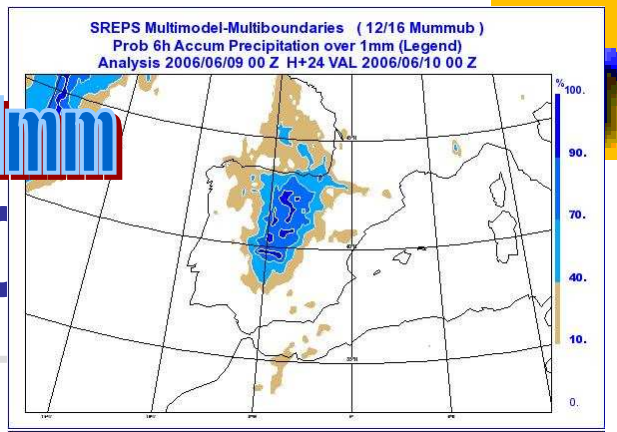


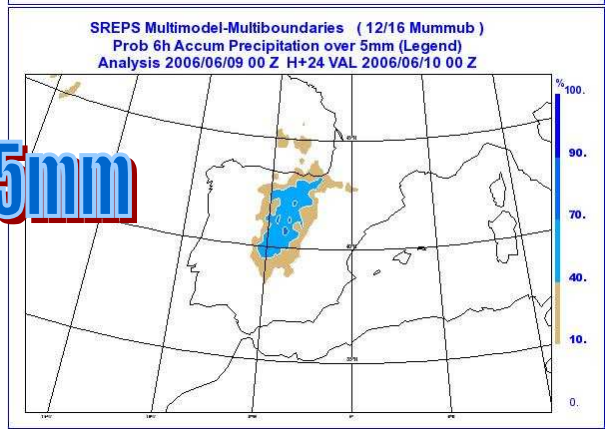
Monit 4: Spread – Ensemble mean maps

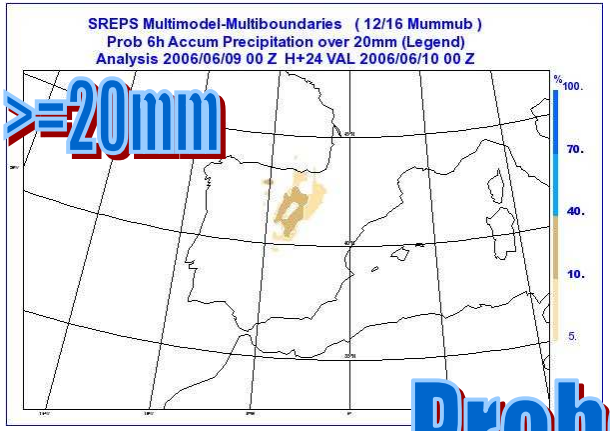


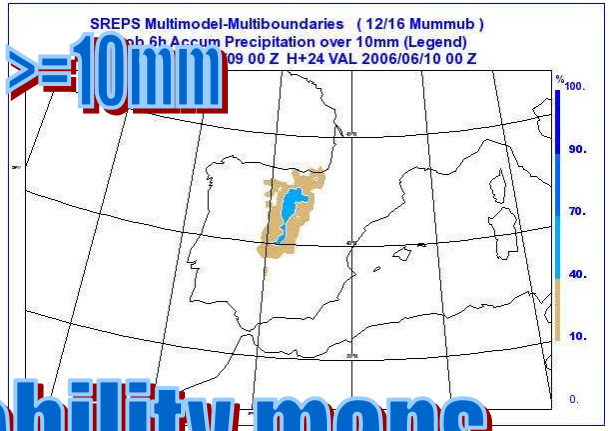
Case Study 06/10/2006 at C











- More than 15 mm/6 hours

Probability maps

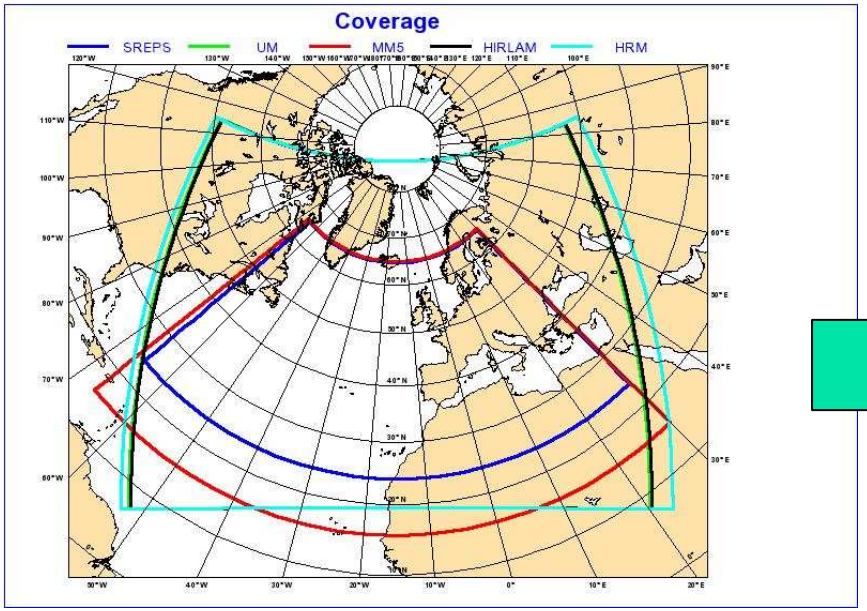
Verification

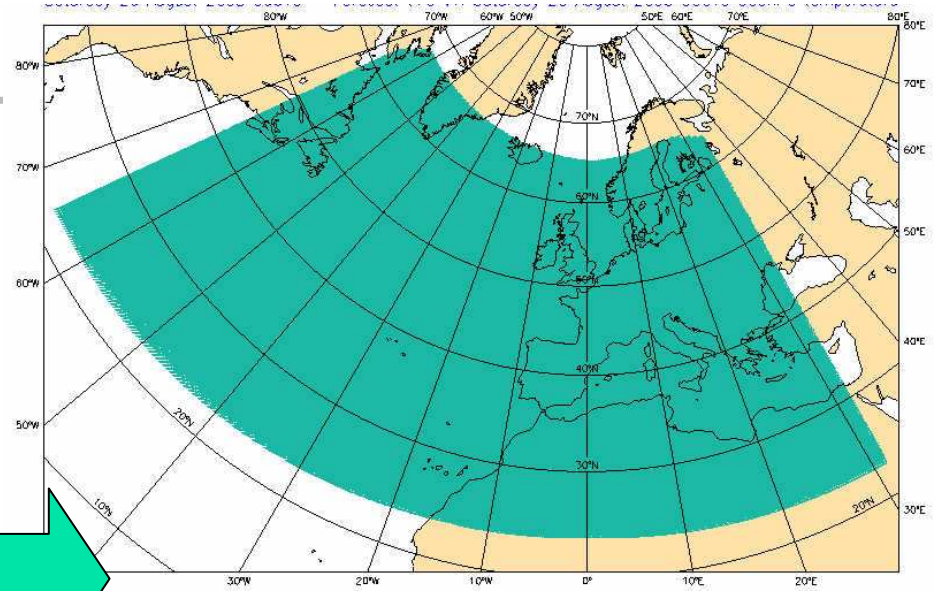
- We show here the 2006 first half (6months) verification results, against both references observations and ECMWF analysis, which look promising:
  - **Calibration:** with synoptic variables Z500, T500, Pmsl
  - **Response to binary events:** reliability and resolution of surface variables 10m surface wind, 6h and 24h accumulated precipitation

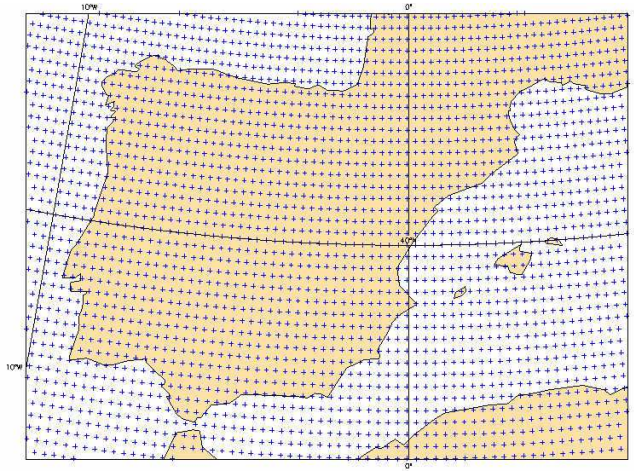
Verification exercise

- Interpolation to a common area
  - ~ North Atlantic + North Africa + Europe
  - Lat-lon Grid 380x184, 0.25º
- ~180 days (Jan1 to Jun30 2006).
- Two different references:
  - Analysis: ECMWF (6h and 24h det fc for Acc. Prec.)
  - Observations: TEMP & SYNOP
- Verification software
  - ~ ECMWF Metview + Local developments
- Deterministic scores
  - Synoptic variables: Bias & RMSE for each member & Ens Mean
- Probabilistic ensemble scores
  - Synoptic variables: Calibration
  - Surface variables: Response to binary events

Interpolation

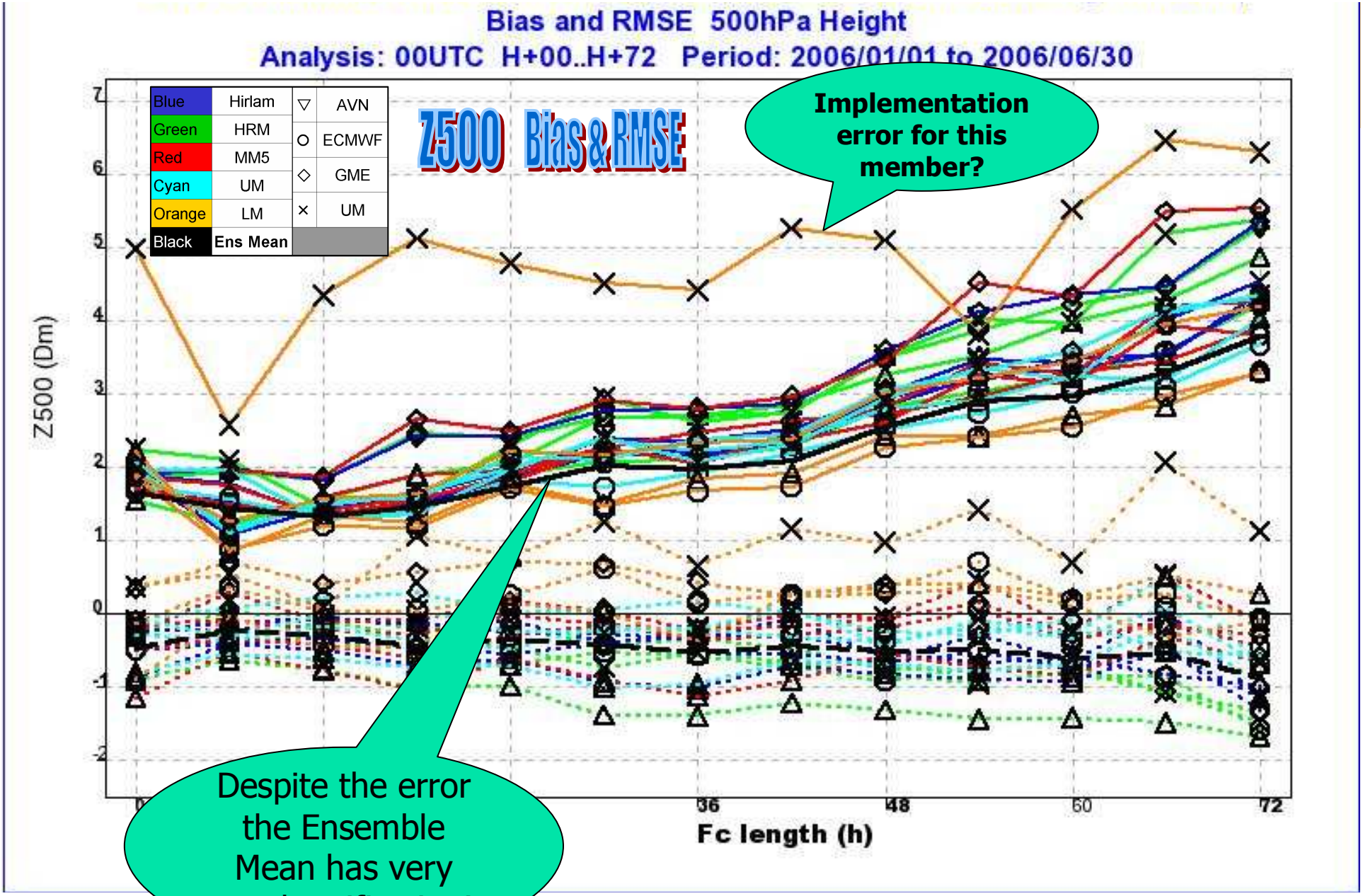


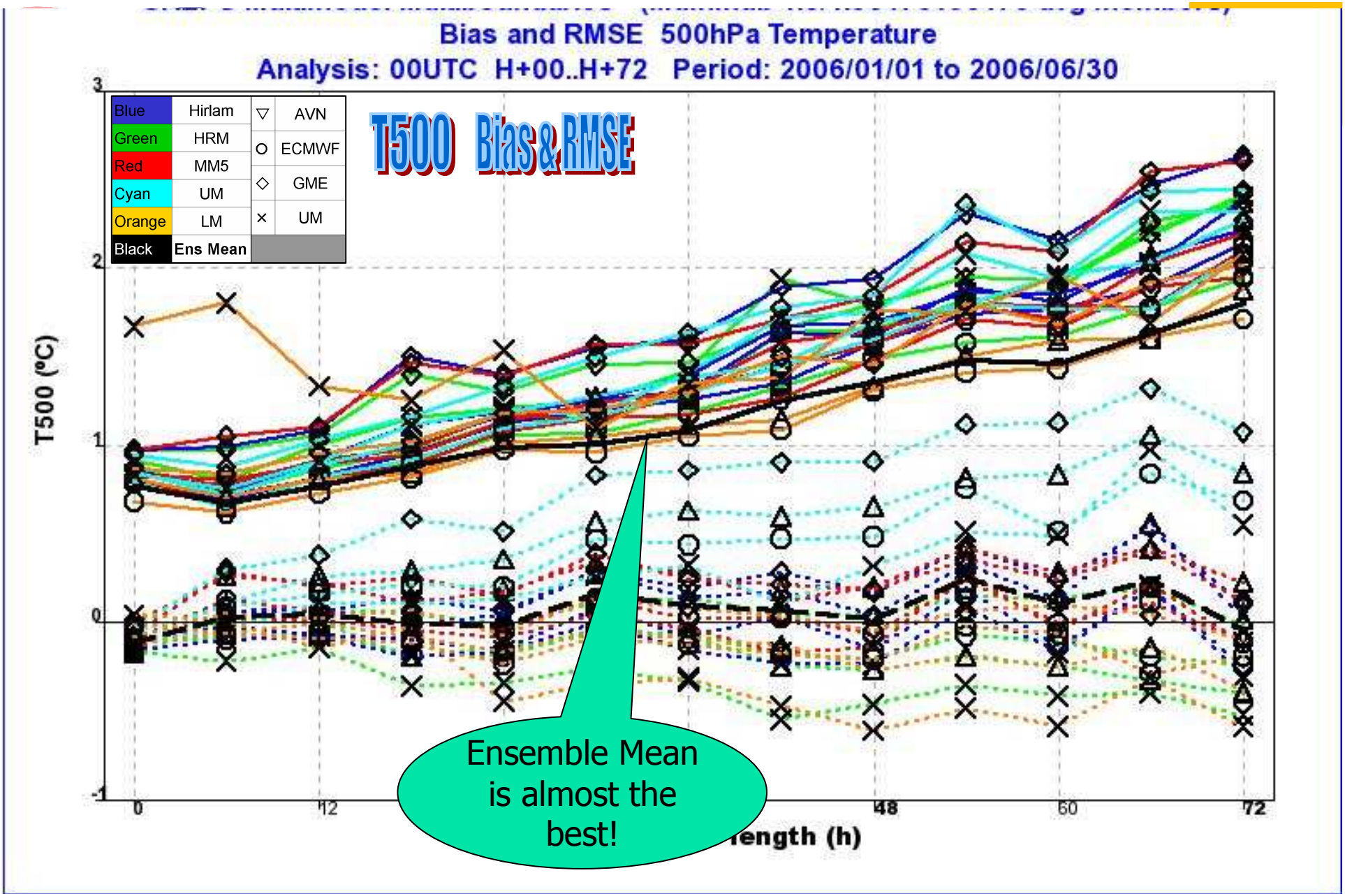


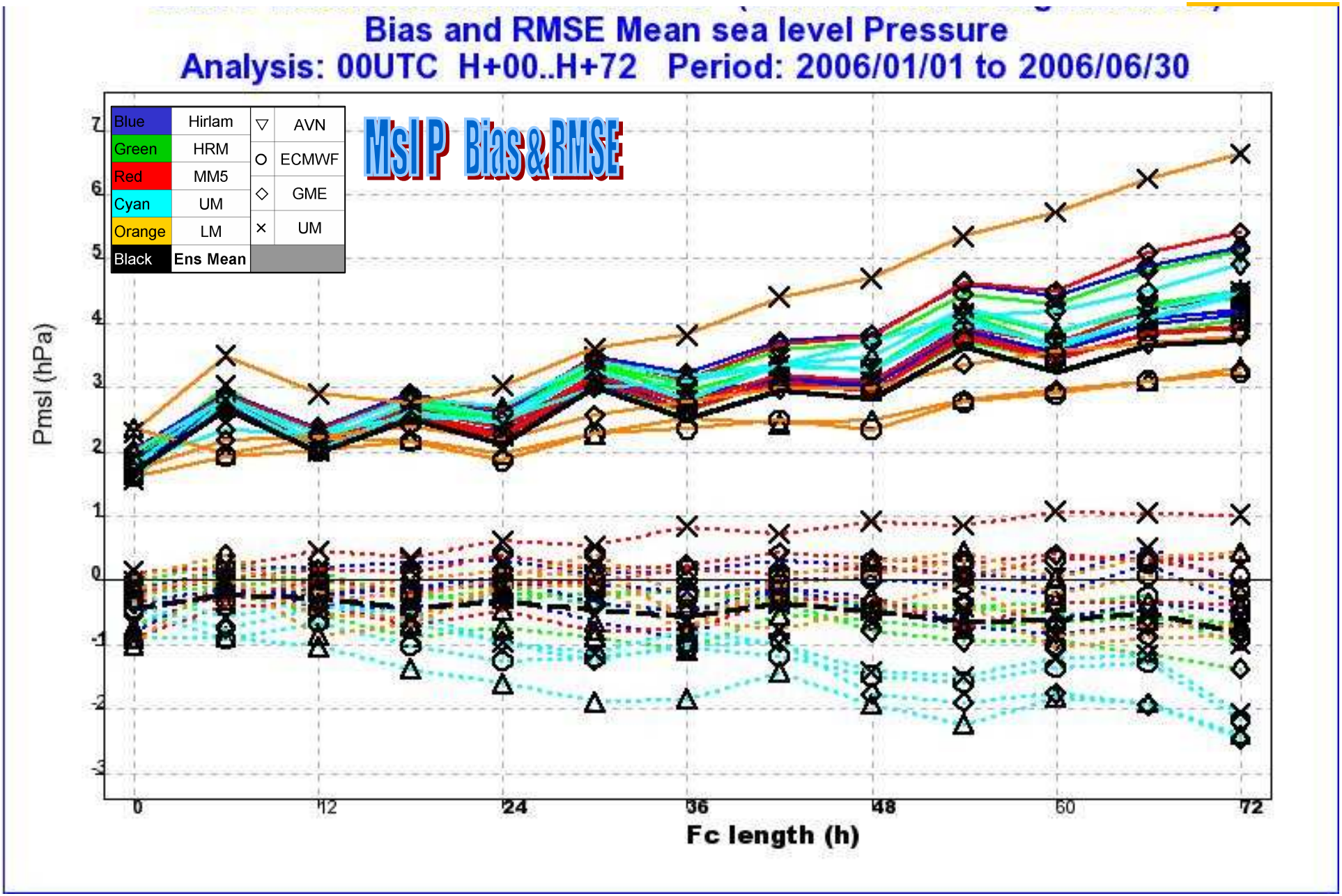


Obs verification - deterministic scores

- Bias & RMSE with forecast length
- For:
  - All members
  - And ensemble mean
- Z500, T500, Msl Pressure

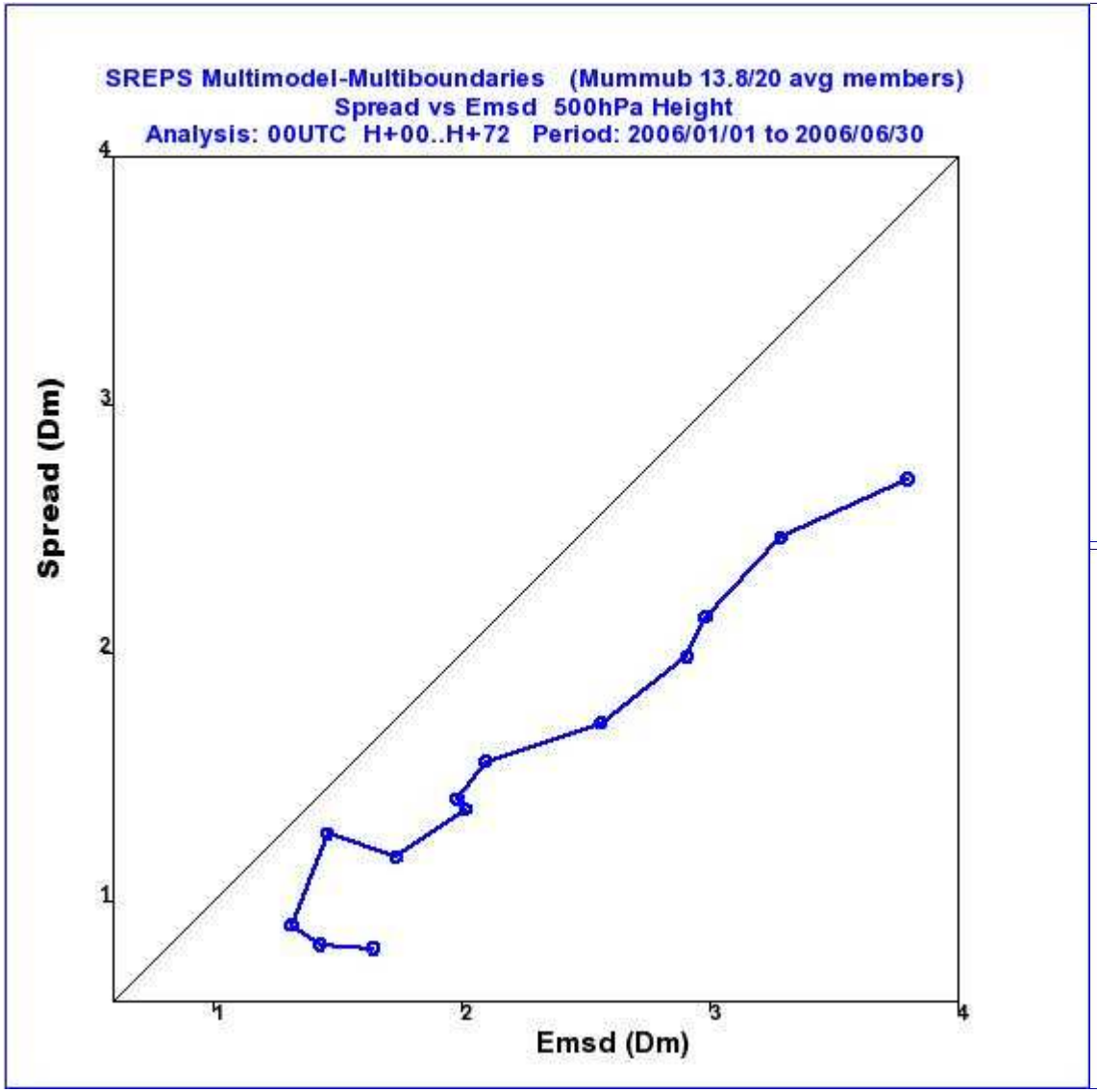

Bias and RMSE 500hPa Height
Analysis: 00UTC H+00..H+72 Period: 2006/01/01 to 2006/06/30
Blue Hirlam
Green HRM
Red MM5
Cyan UM
Orange LM
Black Ens Mean
AVN
ECMWF
GME
UM
Z500 Bias & RMSE
Implementation error for this member?
Z500 (Dm)
Fc length (h)
Despite the error the Ensemble Mean has very good verification!

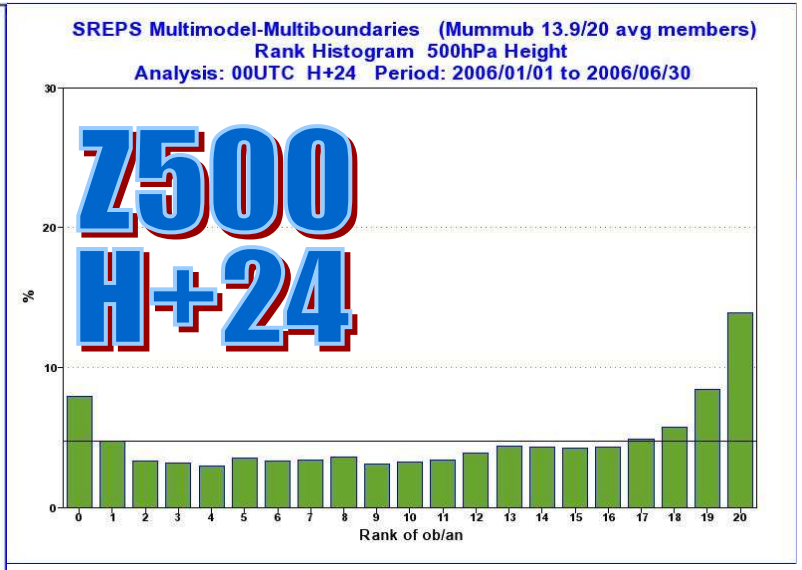

Bias and RMSE 500hPa Temperature
Analysis: 00UTC H+00..H+72 Period: 2006/01/01 to 2006/06/30
T500 Bias&RMSE
Blue Hirlam
Green HRM
Red MM5
Cyan UM
Orange LM
Black Ens Mean
AVN
ECMWF
GME
UM
T500 (ºC)
3
2
1
0
-1
0
12
48
60
72
Ensemble Mean is almost the best!

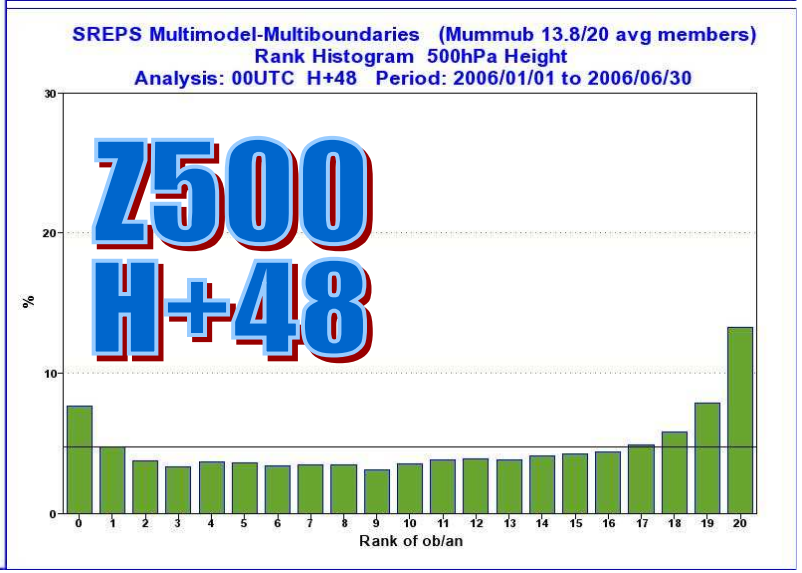

Bias and RMSE Mean sea level Pressure
Analysis: 00UTC H+00..H+72 Period: 2006/01/01 to 2006/06/30
MslP Bias&RMSE
Blue Hirlam
Green HRM
Red MM5
Cyan UM
Orange LM
Black Ens Mean
AVN
ECMWF
GME
UM
Pmsl (hPa)
Fc length (h)
0
12
24
36
48
60
72

Obs verification - Probabilistic scores

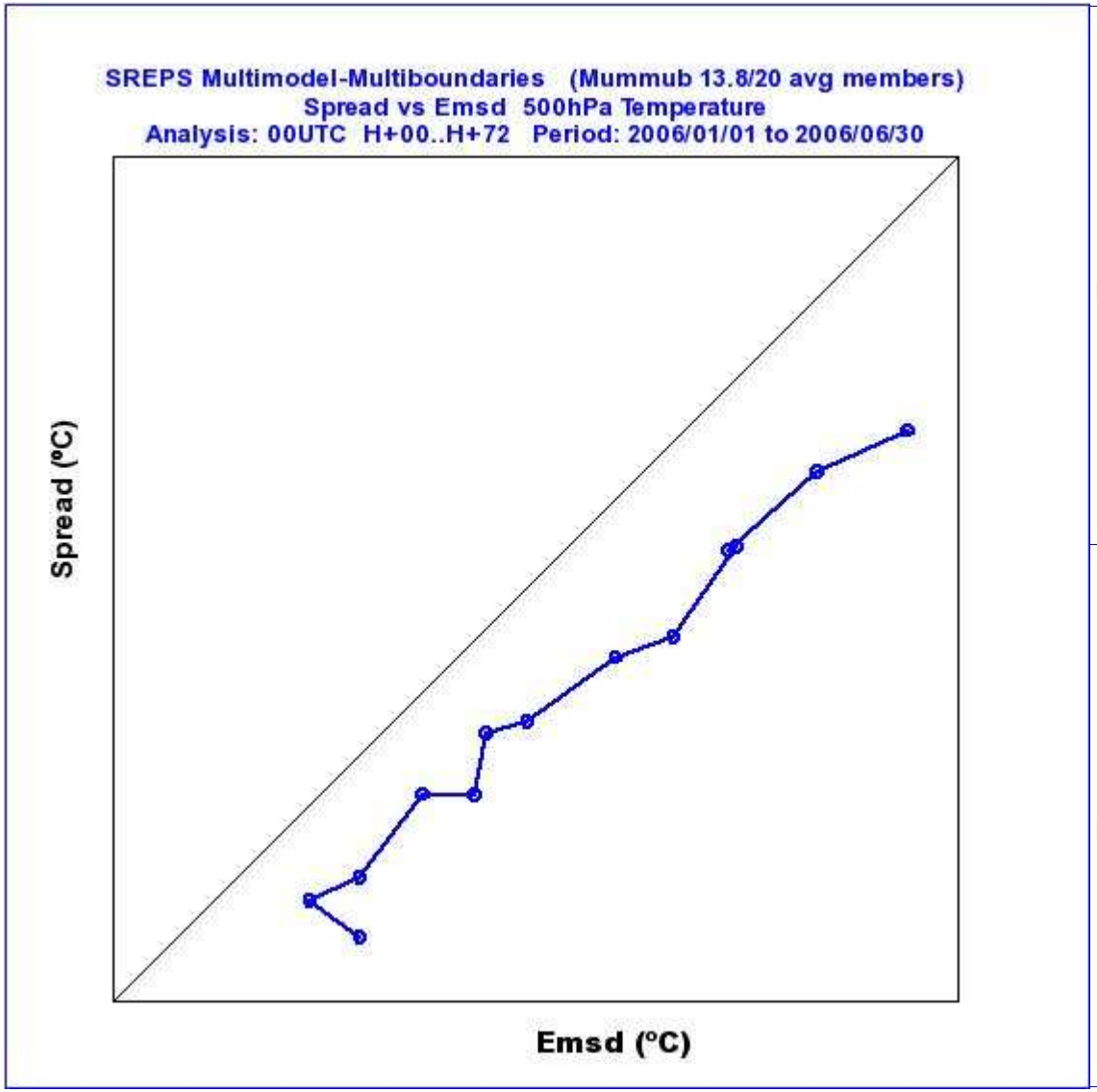
- Ensemble calibration:
  - Synoptic variables:
    - Z500, T500, Pmsl
  - Scores:
    - Rank histograms
    - Spread-skill
- Response to binary events:
  - Surface variables:
    - 10m surface wind (10,15,20m/s thresholds)
    - 6h accumulated precipitation (1,5,10,20mm thresholds)
    - 24h accumulated precipitation (1,5,10,20mm thresholds)
  - Scores:
    - Reliability, sharpness (H+24, H+48)
    - ROC, Relative Value (H+24, H+48)
    - BSS, ROCA with forecast length

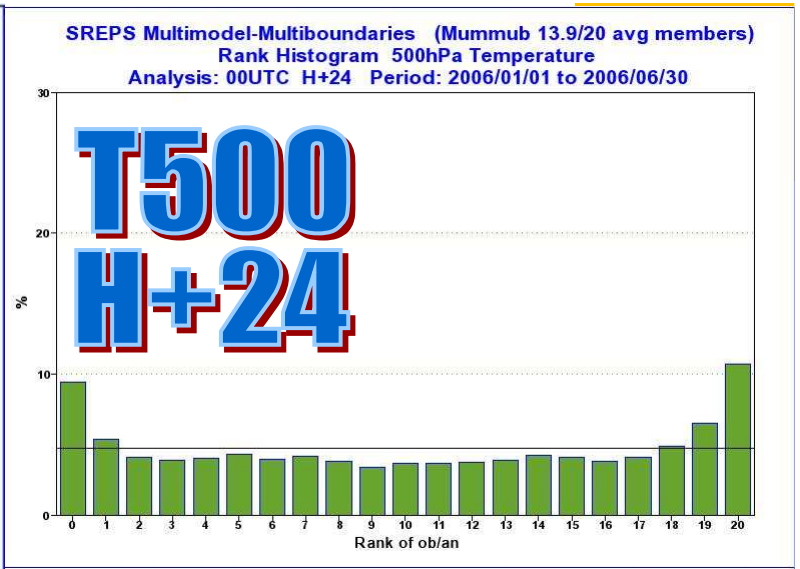


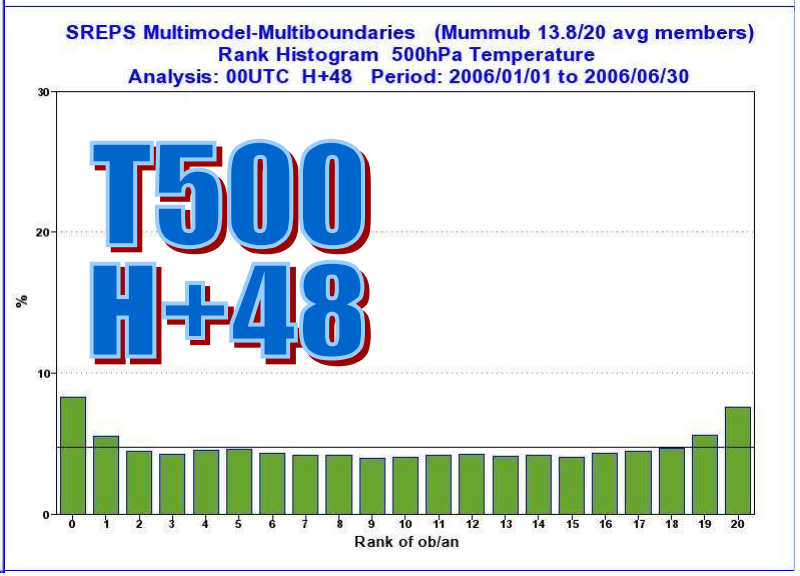




- Z500 under-dispersive
- A bit under-forecast







- T500 under-dispersive, but better
- Fairly flat rank histograms

Binary events

- Binary events X = {0,1} at every point
- Accumulated precipitation in 24 hours >= 5mm
- Useful to decompose the forecast in thresholds
- Performance computed using contingency tables (CT's)

Contingency tables

- It is the best way to characterize a binary event

fc(X)={1,0}

ob(X)={1,0}

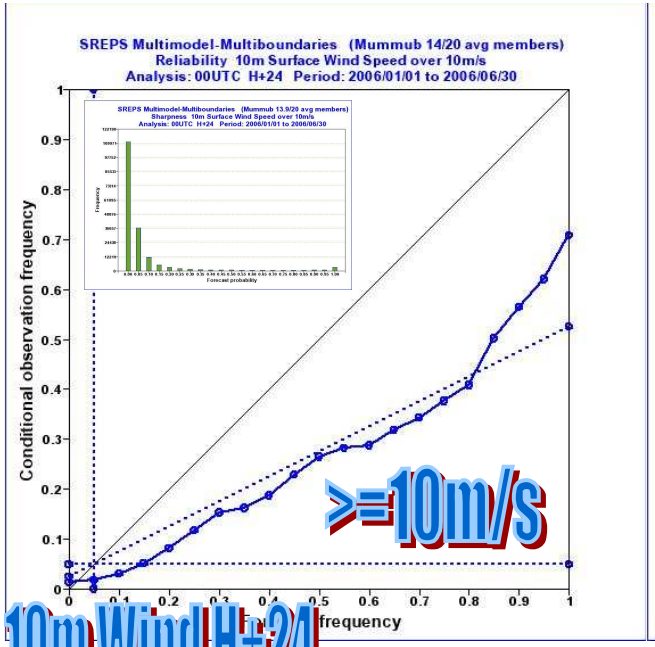
		ob		
		1	0	
fc	1	a	b	a+b
	0	c	d	a+d
		a+c	b+d	a+b+c+d = N

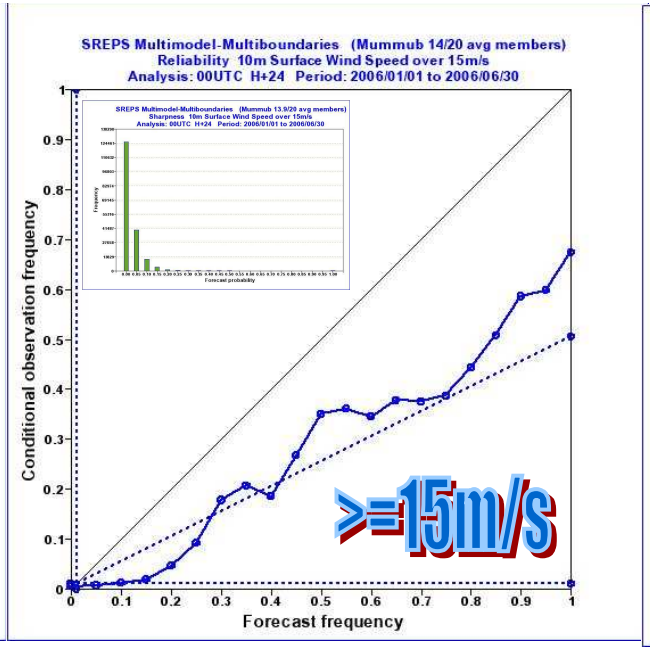
Contingency tables: scores

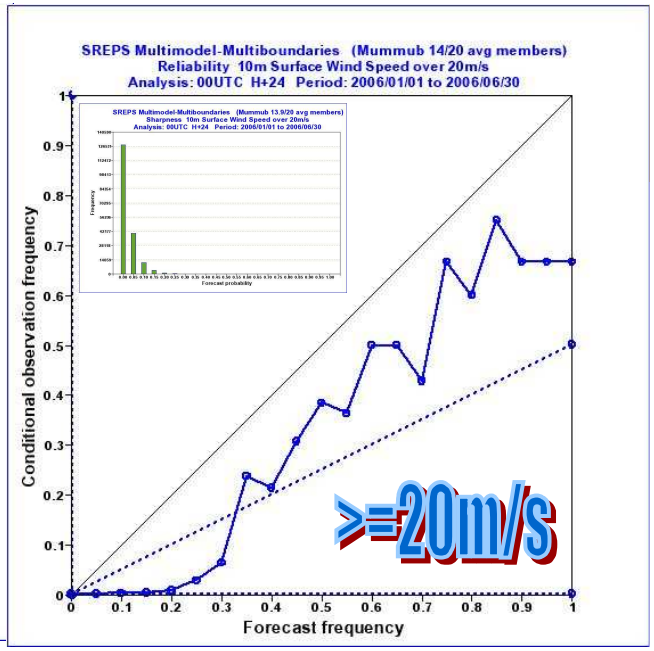
- Several scores can computed from CT's

		ob		
		1	0	
fc	1	a	b	a+b
	0	c	d	a+d
		a+c	b+d	a+b+c+d = N

Base Rate	s	(a + c) / n
HitRate	HIR	a / (a + c)
False Alarm Rate	FAR	b / (b + d)
False Alarm Ratio	FARatio	b / (a + b)
Proportion Correct	PC	(a + d) / n
...		



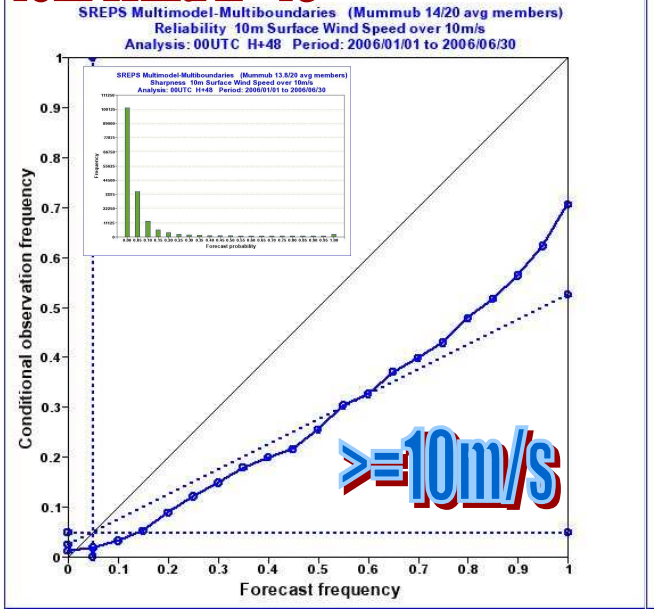


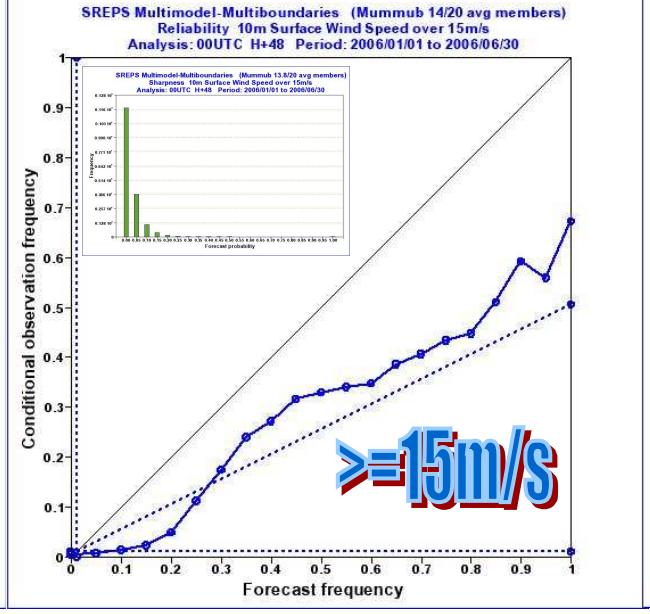


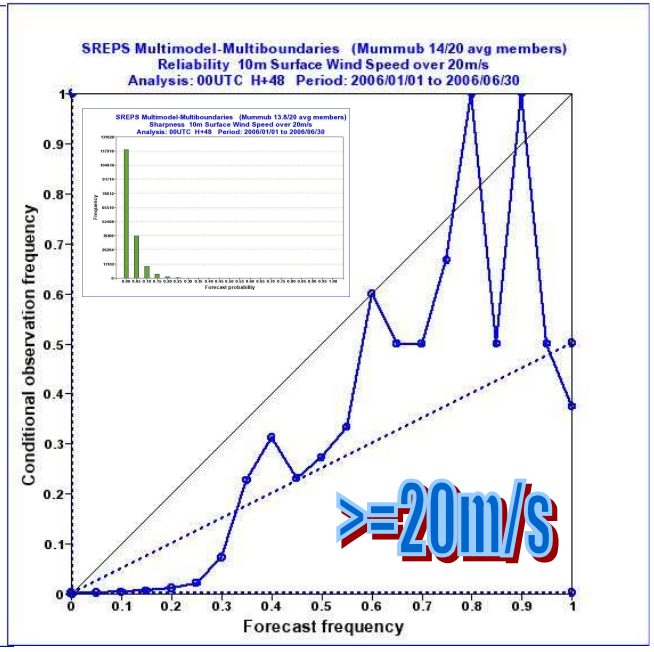
10m Wind H+24

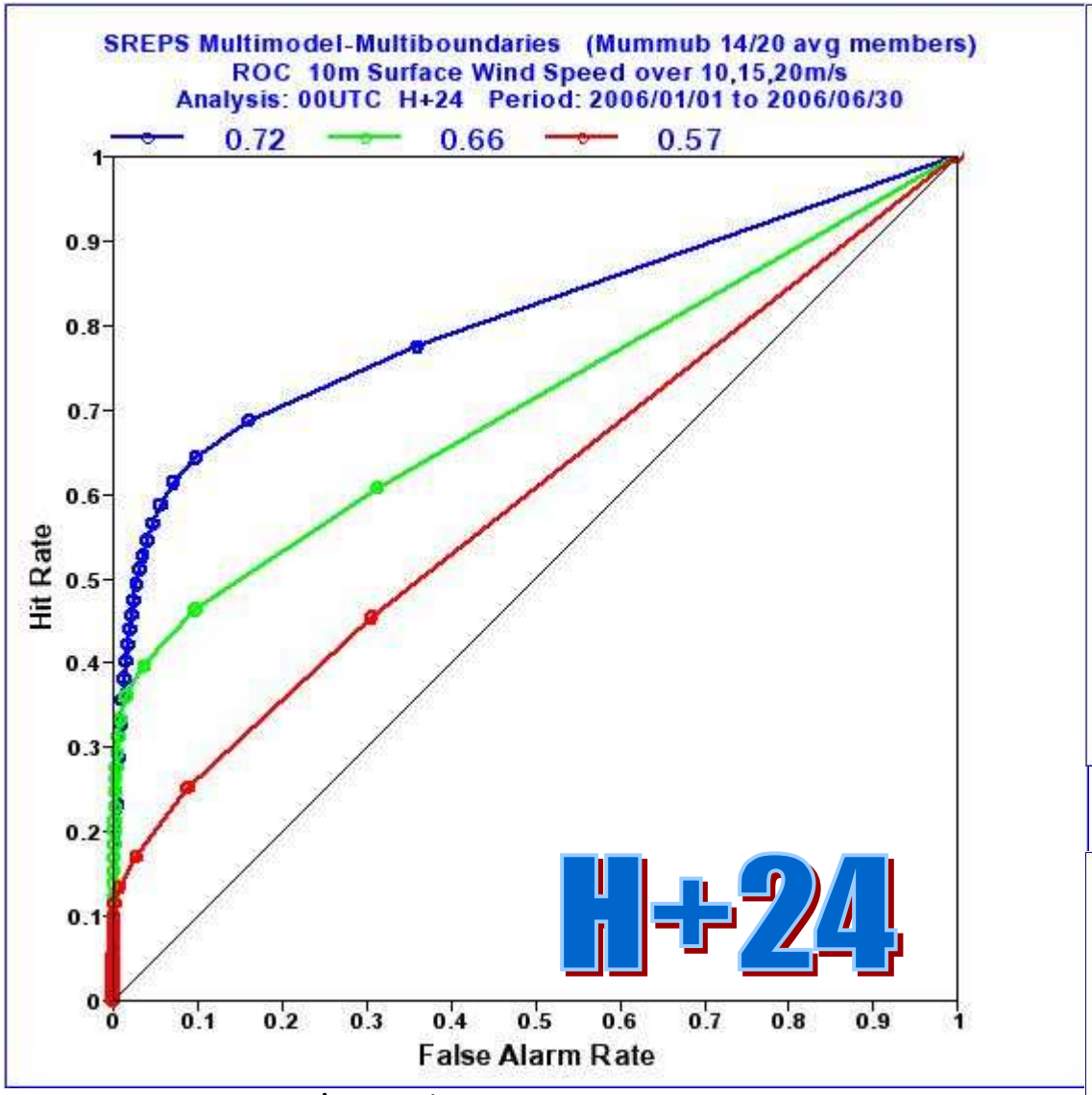
- Quite reliable
- Near/over no-skill line, 20 m/s few cases

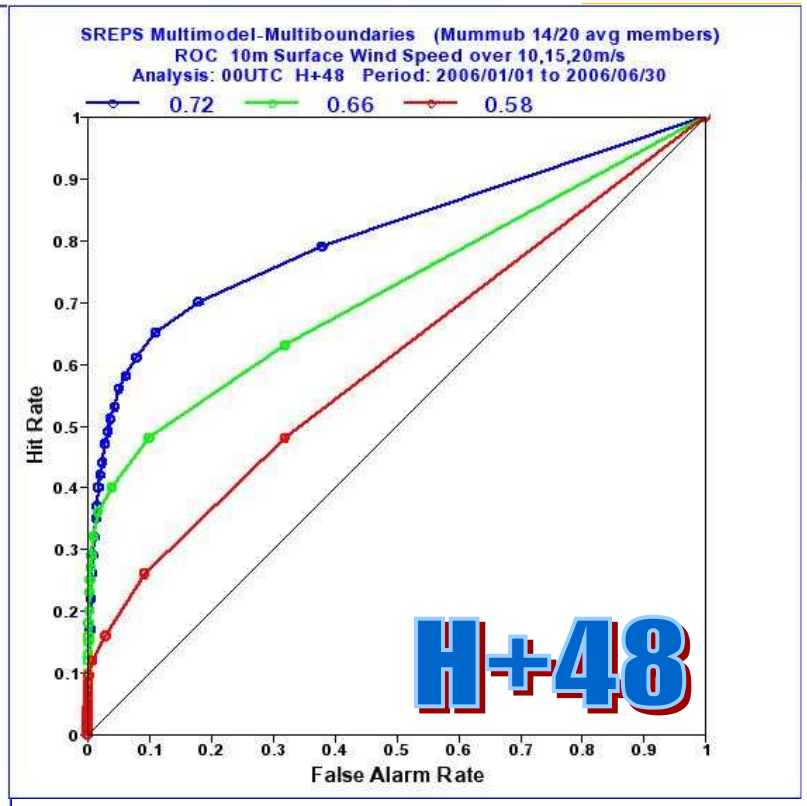
10m Wind H+48



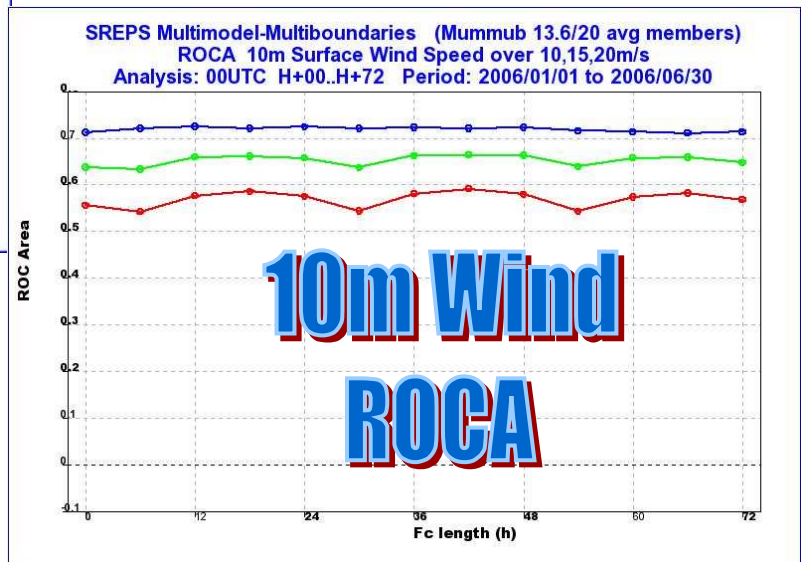






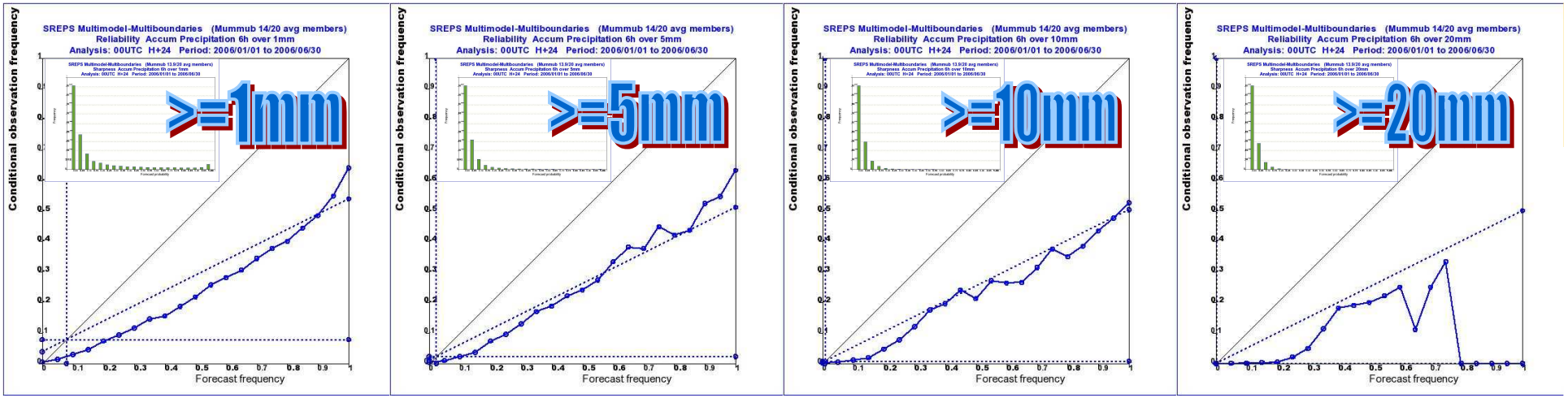


- Quite good ROC curves
- Decreasing performance with threshold
- Same performance with FC length?



6h accumulated precipitation

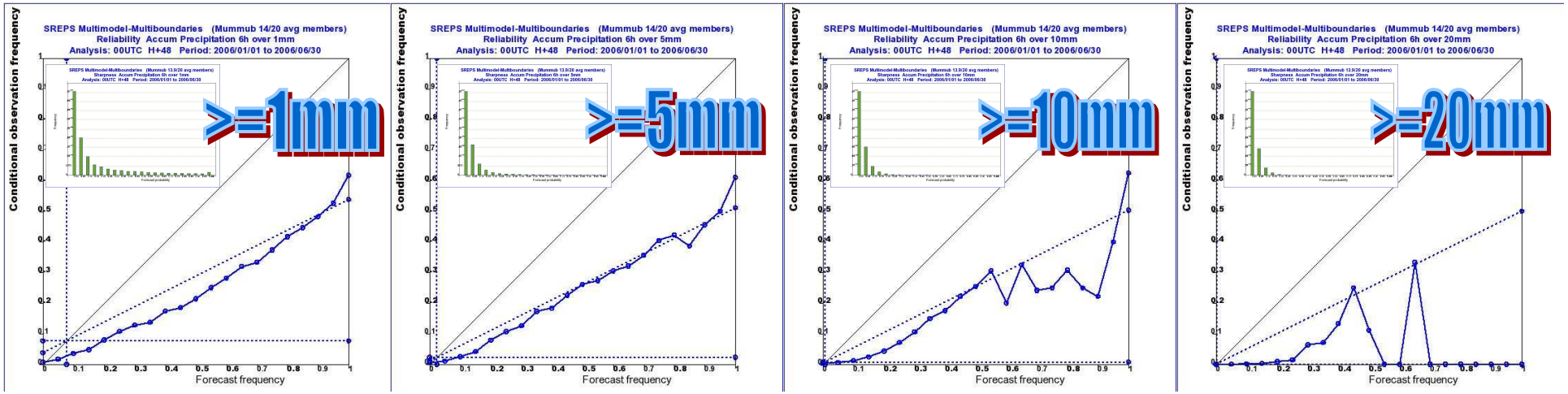
- Binary event thresholds
  - 1, 5, 10, 20mm / 6h
- Reliability & sharpness
- Resolution
  - ROC
  - ROCA
- Brier Skill Score
- Economic Value RV

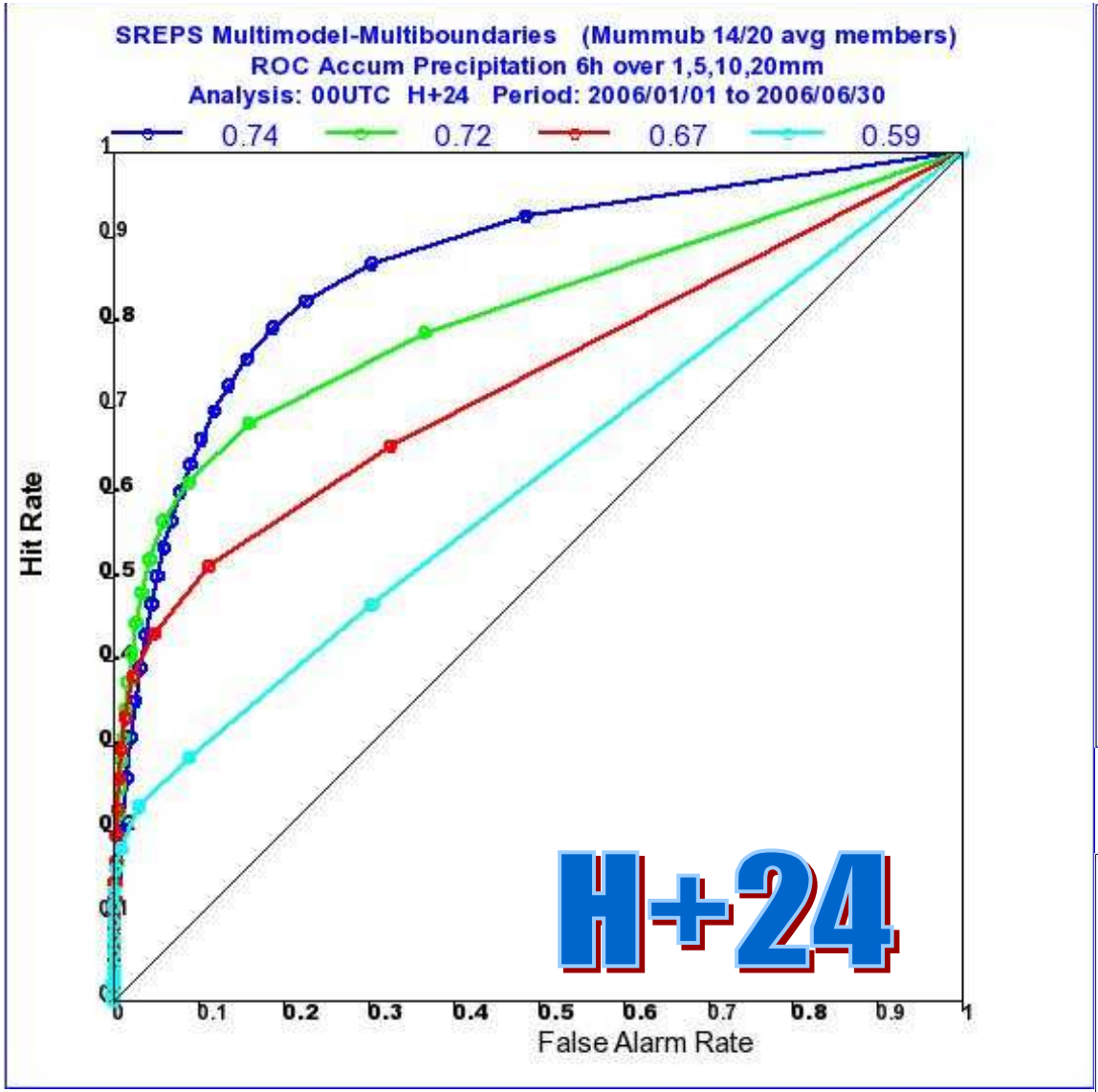


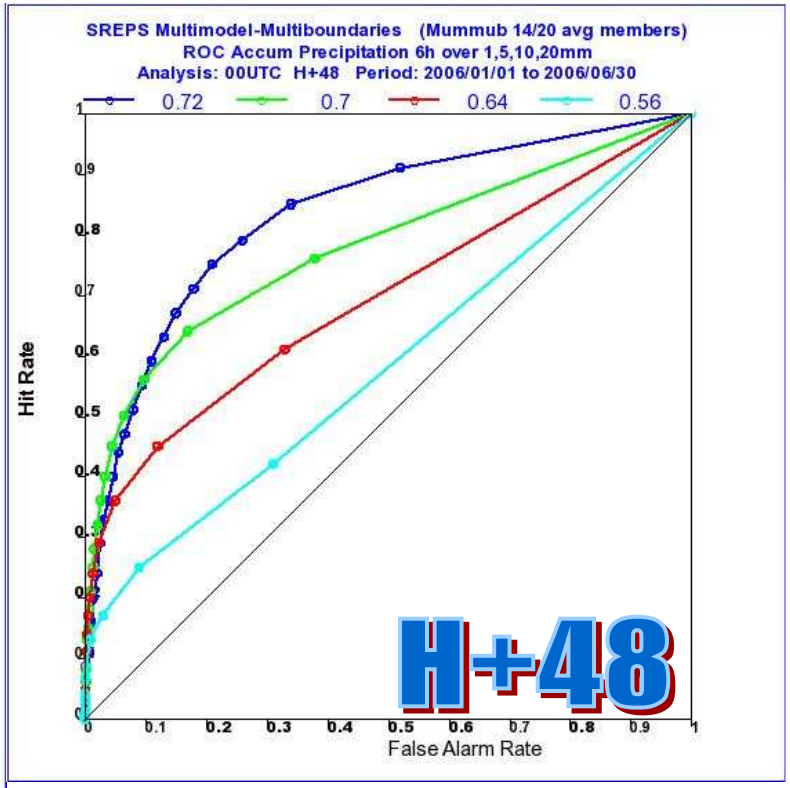
6hAccPrec H+24

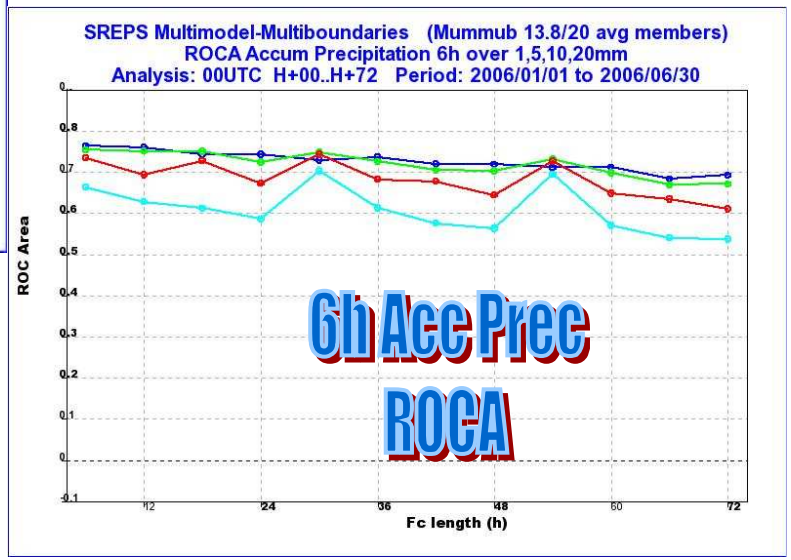
- Quite reliable, apart from 20mm
- 1,5,10mm near no-skill line
- 20mm under no-skill line, few cases

6hAccPrec H+48





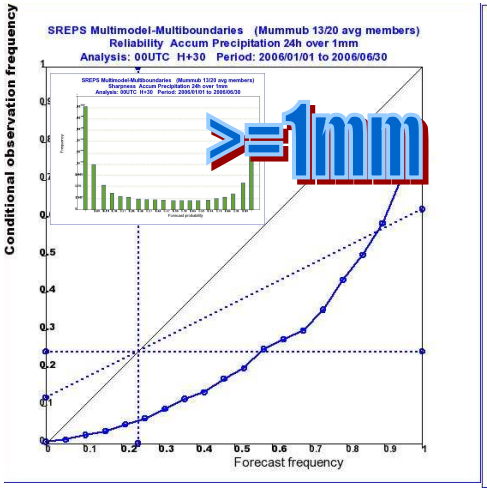


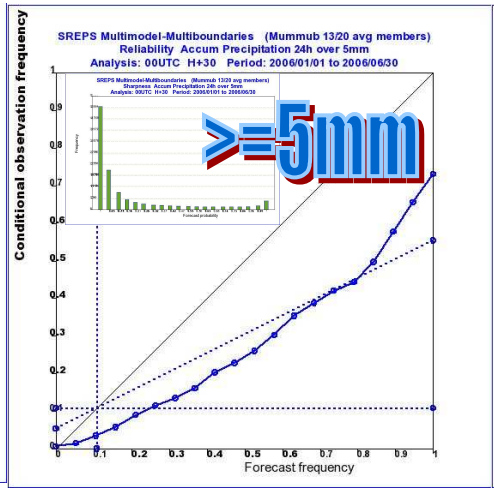


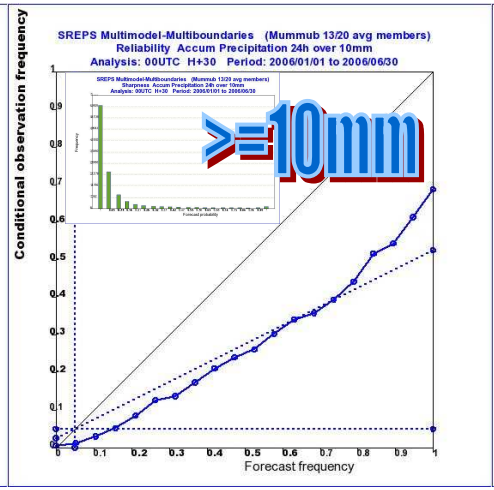
- Quite good ROC curves
- Worse with threshold
- Slightly worse with FC length

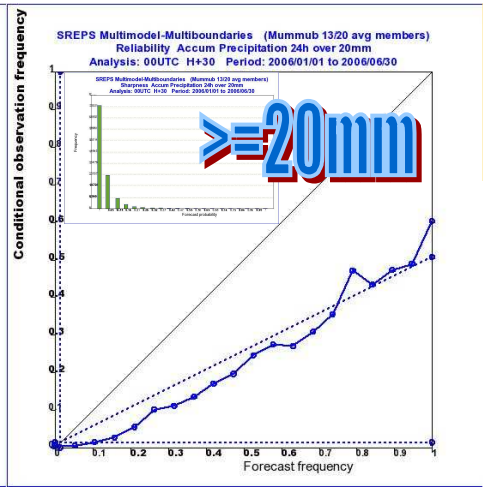
24h accumulated precipitation

- Binary event thresholds
  - 1, 5, 10, 20mm / 24h
- Reliability & sharpness
- Resolution
  - ROC
  - ROCA
- Brier Skill Score
- Economic Value RV





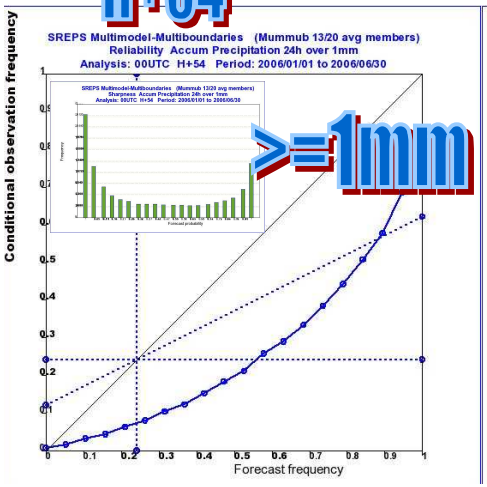


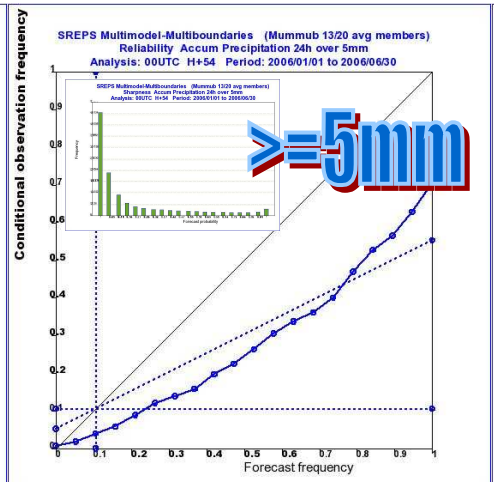


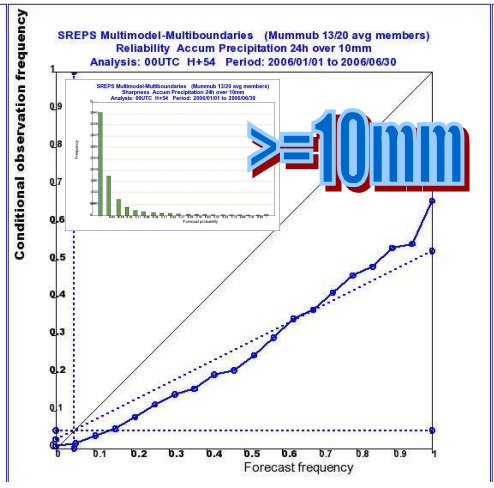
24h Acc Prec H+30

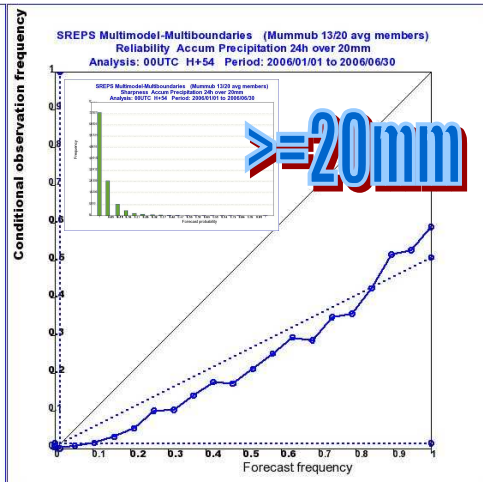
- 1mm Poorly reliable, very sharp
- 5,10,20mm quite reliable
- 1mm far under no-skill line
- 5,10,20mm near no-skill line

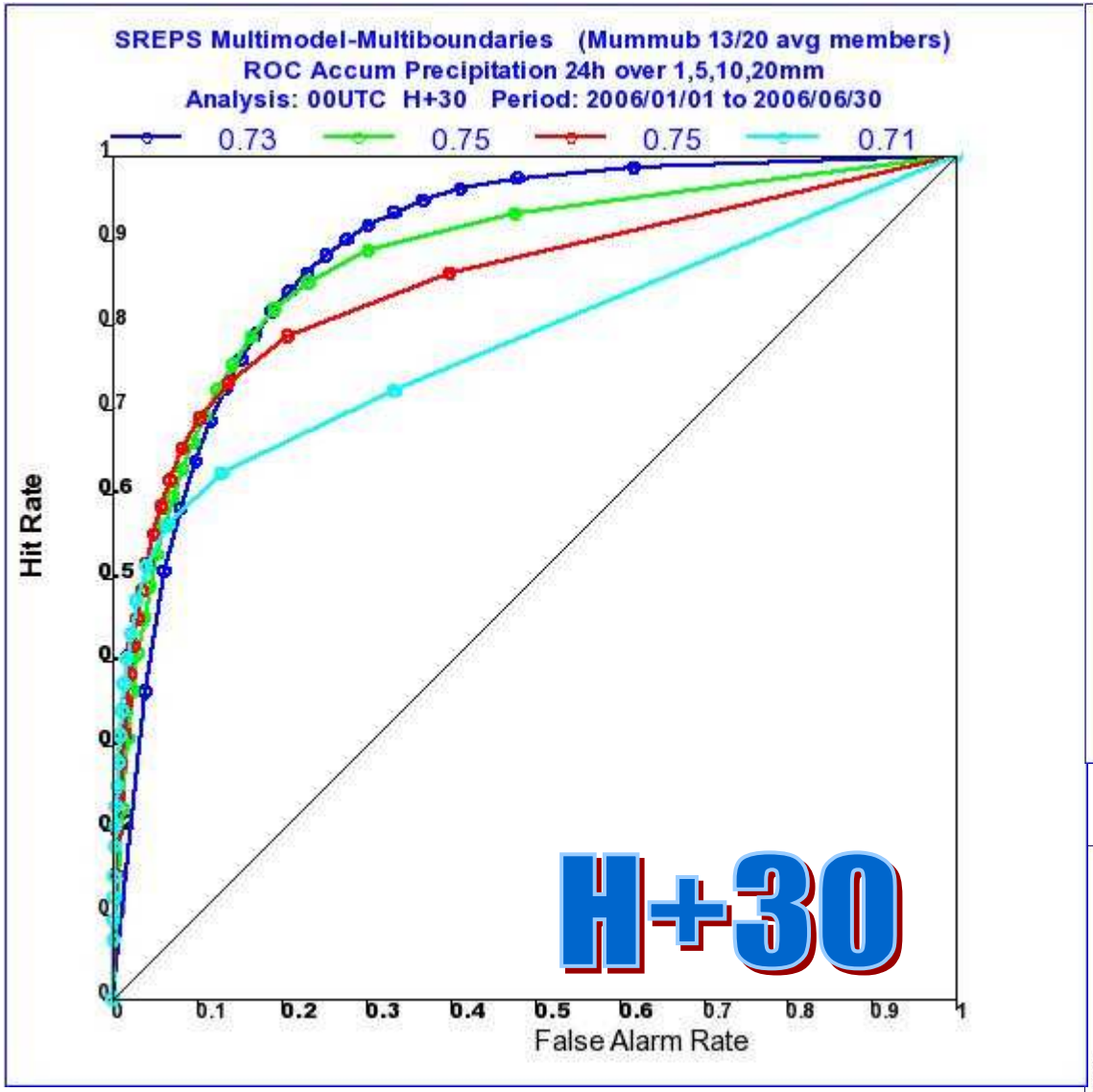
24h Acc Prec H+54

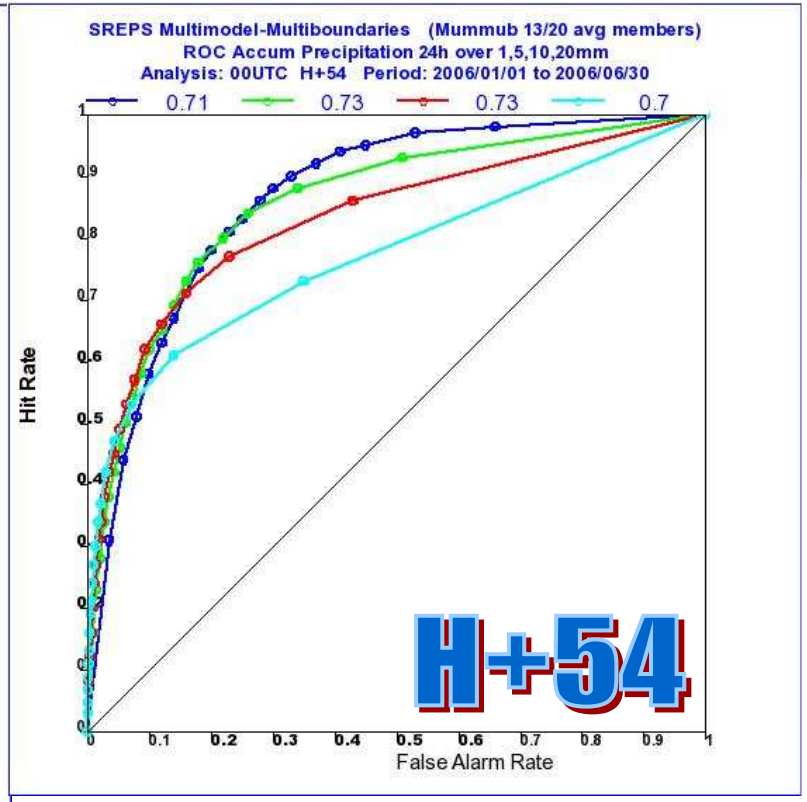


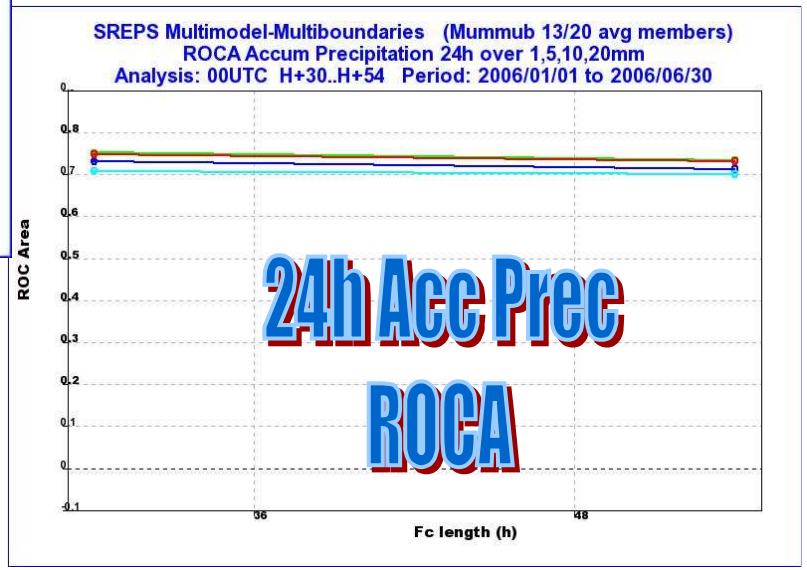






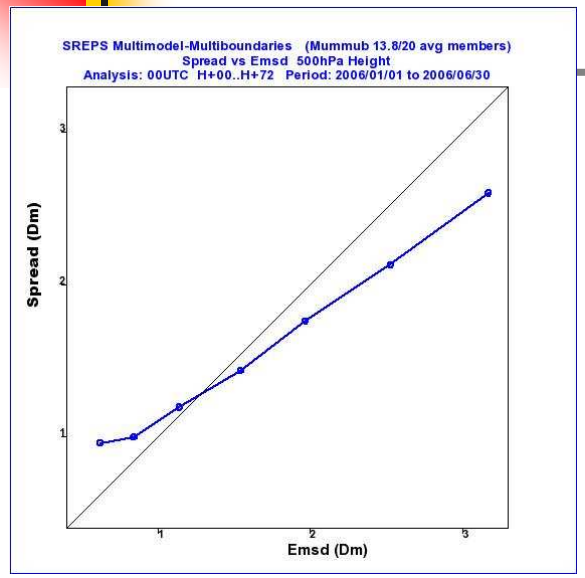


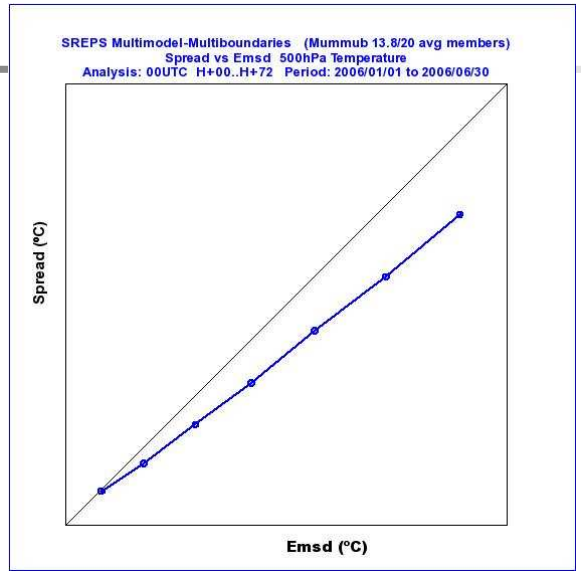


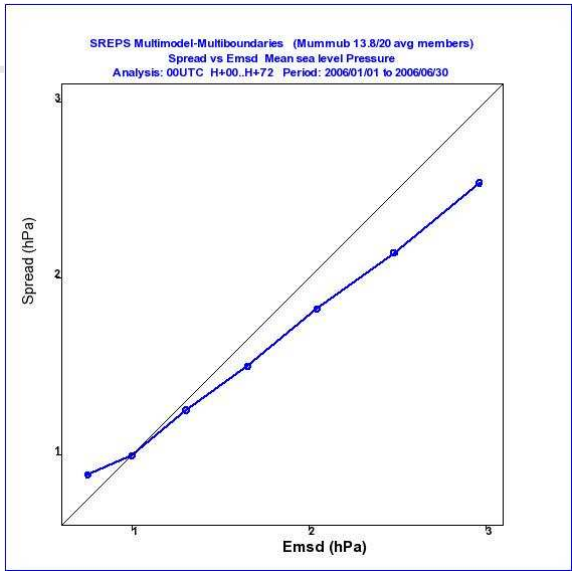


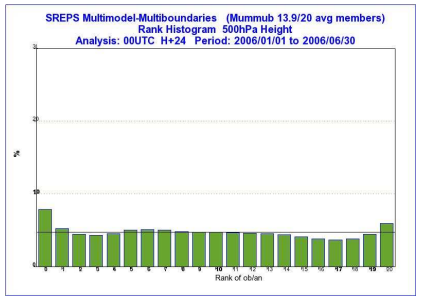
- Good ROC curves
- Small degradation with threshold
- Flat with FC length?

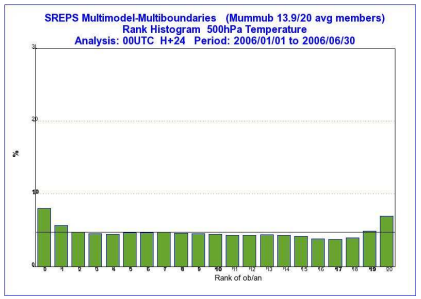
Against analysis: synoptic variables

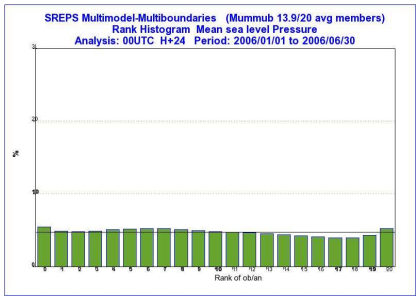






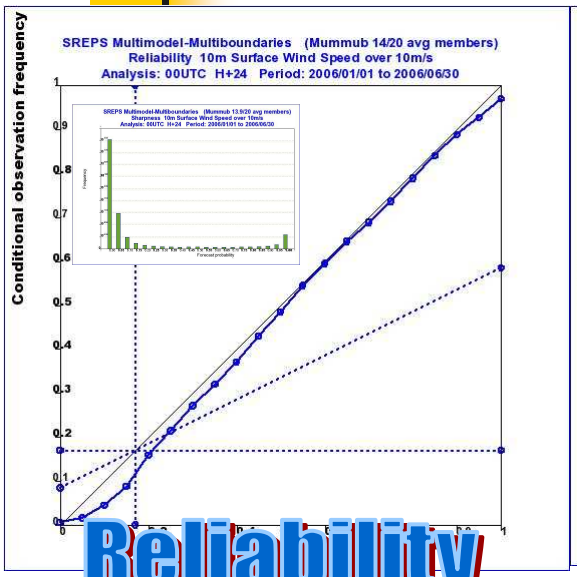


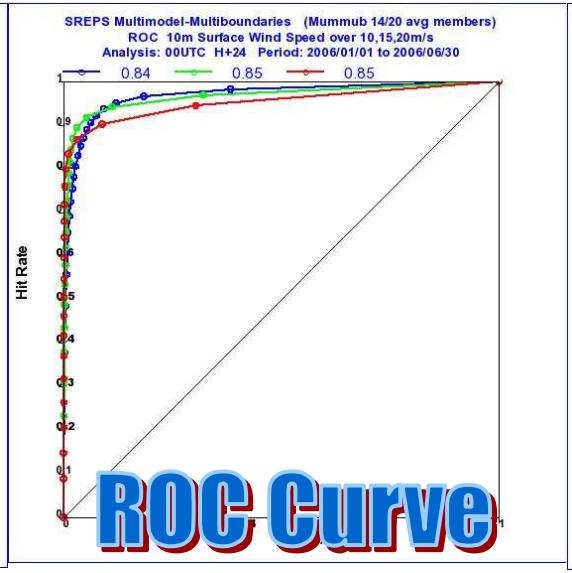


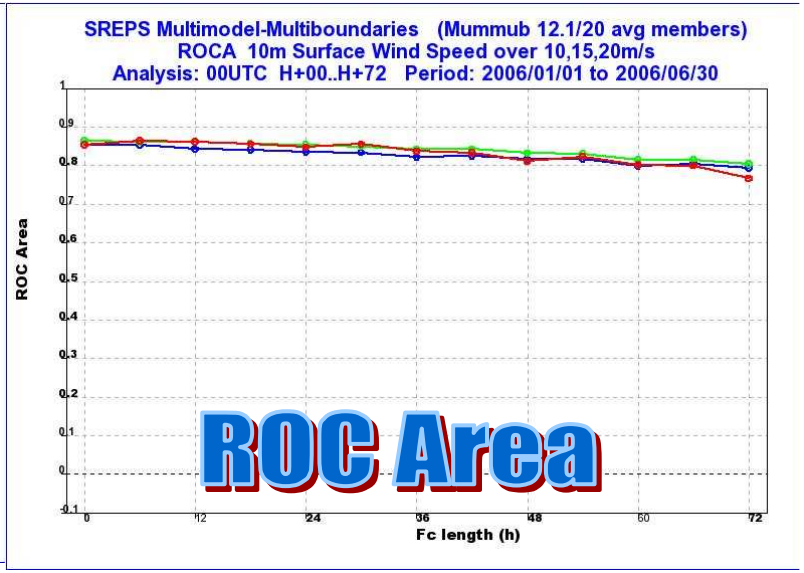


- Much better

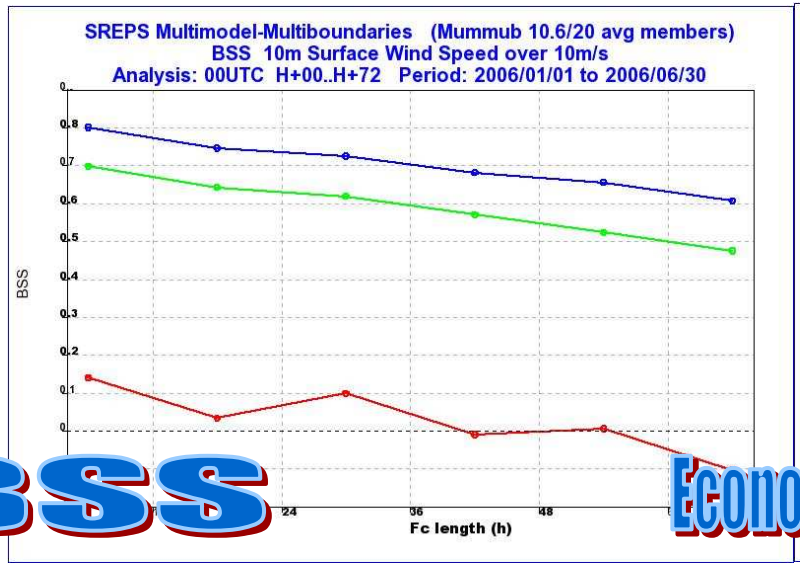
Against analysis: 10m wind (H+24) >= 10 m/s

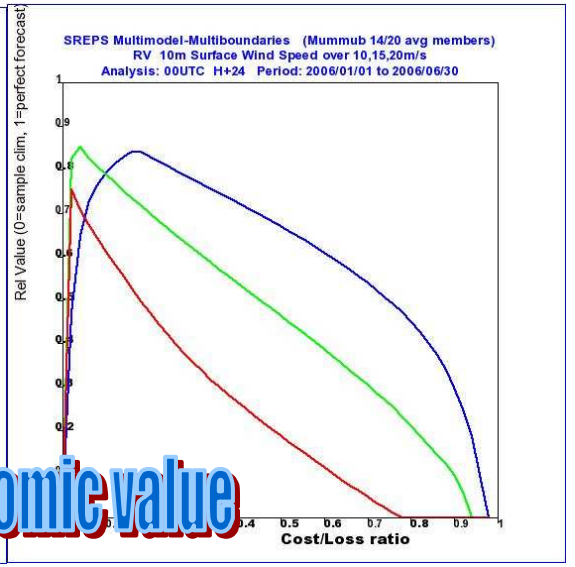




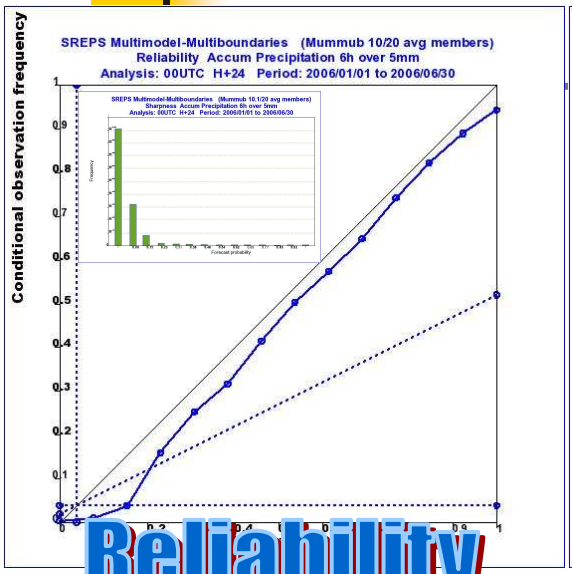


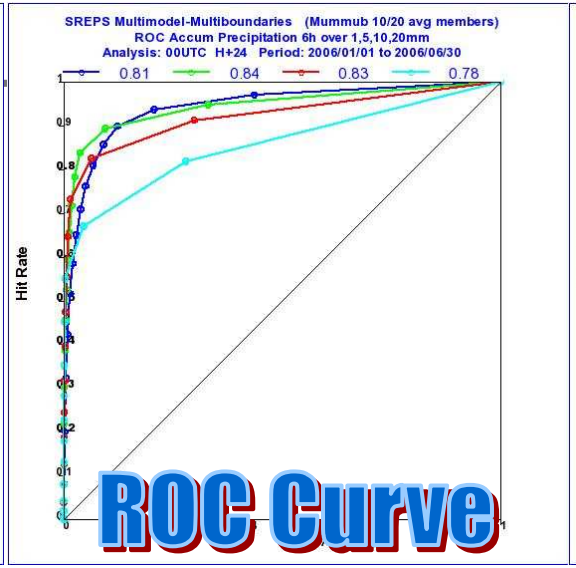
- Much better

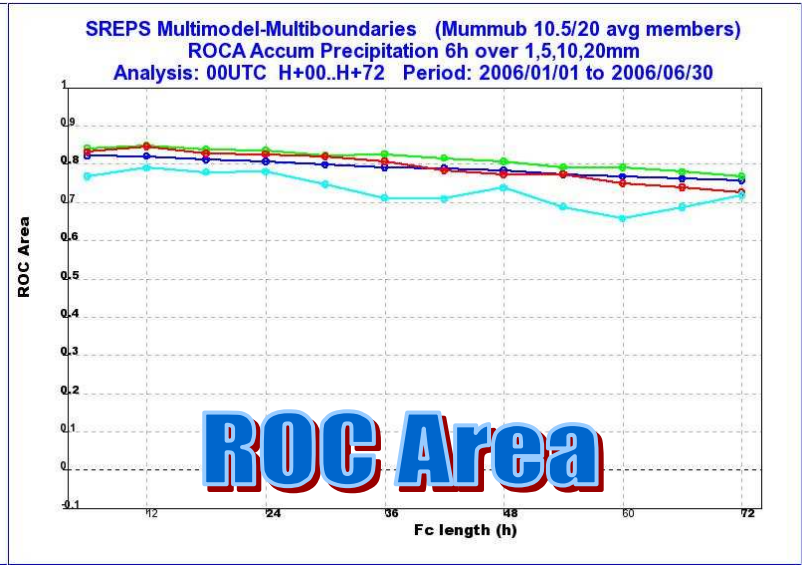




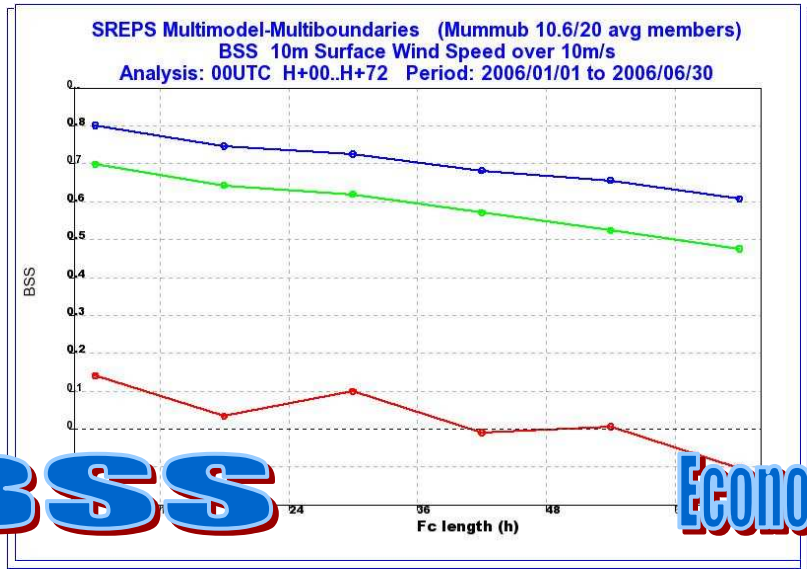
Against EMCWF 24 h forecast: 24h Acc Precip >= 5 mm

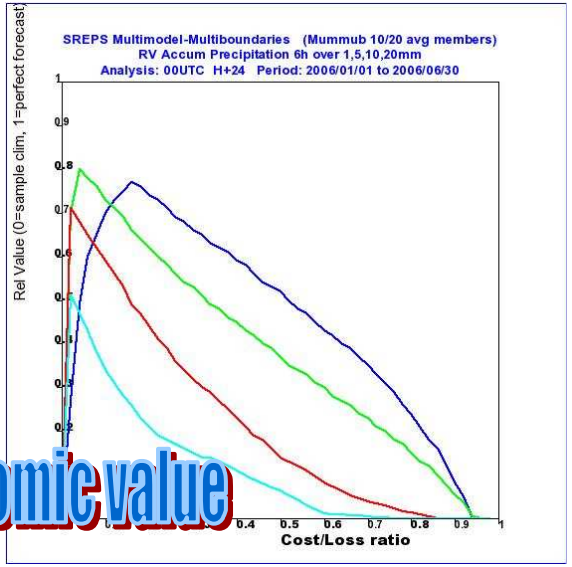






- Much better

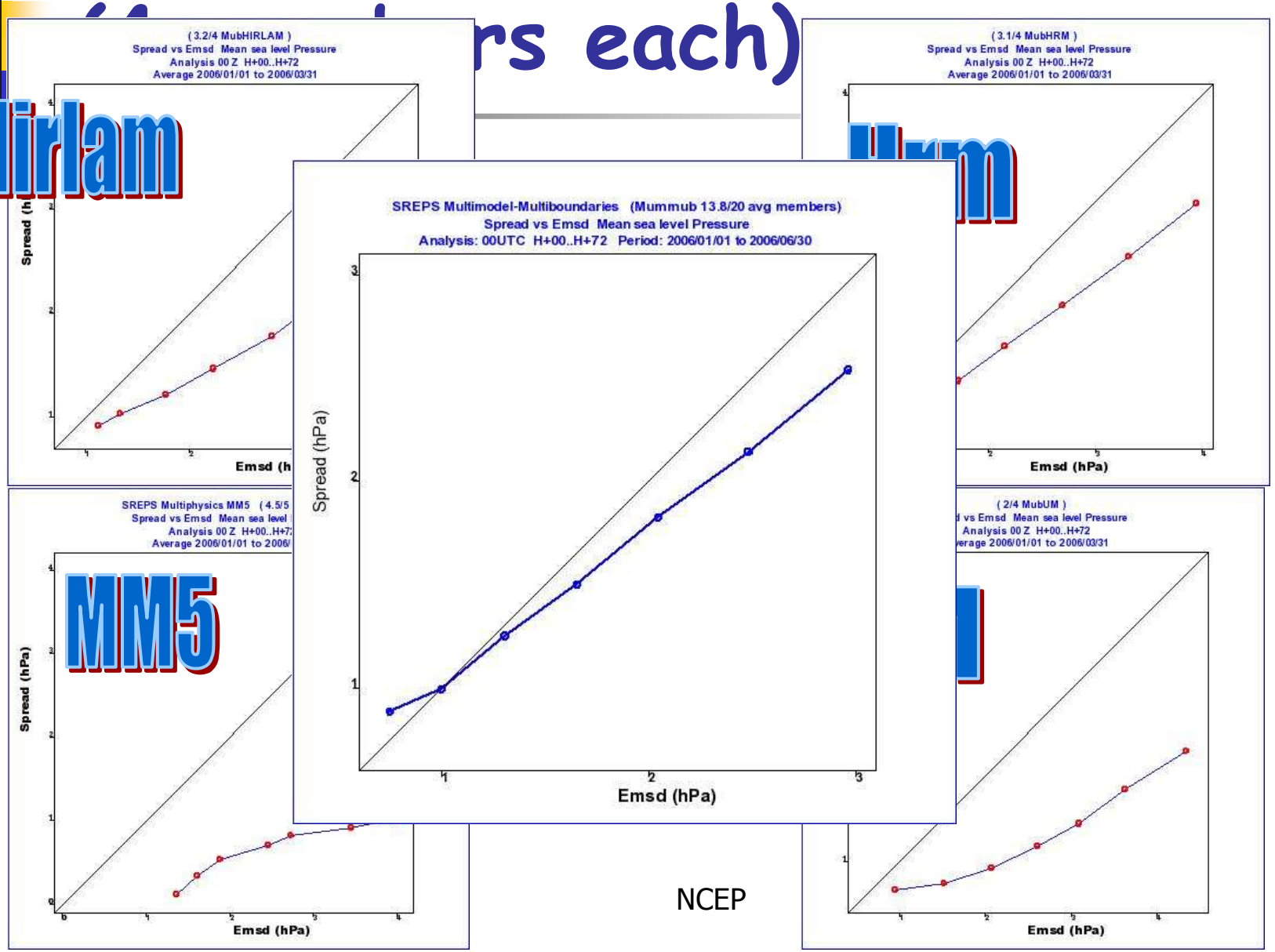




Why Multi-model?

- Better representation of model errors (SAMEX, Hou & Kalnay 2001, and DEMETER).
- Consistent set of perturbations of initial state and boundaries.
- Better results than any single model ensemble (SAMEX, DEMETER, Arribas et al., 2005).

Single model Ensembles (rs each)



Conclusions

- A Multi-model-Multi-boundaries Short Range Ensemble Prediction System (MMSREPS), has been develop at the INM-Spain
- We show here 6months verification results (2006 first half), against both observations and ECMWF analysis:
  - Calibration: with synoptic variables Z500, T500, Pmsl
  - Response to binary events: reliability and resolution of surface variables 10m surface wind, 6h and 24h accumulated precipitation

Conclusions (cont.)

- These first results look promising:
  - Verification against ECMWF analysis shows very good results
  - Verification against observations shows quite good results
    - Ensemble is under-dispersive
    - Good response to binary events
- The ensemble performance could be improved with some post-processing, today under development:
  - Bias correction
  - Calibration using Bayesian Model Averaging (BMA)

Coming Future

- **Bias correction**
- **Calibration**: Bayesian Model Averaging
- Verification against **climate observations** (high resolution precipitation network over Europe)
- Time-lagged 40 members twice a day
- More post-process software (clustering)

References

- (2001) Hou D., Kalnay E., & Droegemeier, K.K.: Objective Verification of the SAMEX'98 Ensemble Forecasts. M.W.R., 129, 73-91.
- (1999) Buizza R., Miller, M., & Palmer, T.N.: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. Q. J. R. Meteorol. Soc., 125, 2887-2908.
- (1997) Toth Z., & Kalnay E.: Ensemble Forecasting at NCEP and the breeding method. M.W.R., 125, 3297-3319.
- (1997) Palmer T.N., Barkmeier J., Buizza R., & Petroliagis T.: The ECMWF Ensmble Prediction System. Meteorol. Appl., 4, 301-304.
- (2004) Palmer, T.N., et al.: Development of a European multi-model ensemble system for seasonal to inter-annual prediction (DEMETER). B.A.M.S., 85, 853-872.
- (2005) Arribas A., Robertson K.B., & Mylne, K.R.: Test of Poor Man's Ensemble Prediction System. M.W.R., 133, 1825-1839.

Questions

j.garciamoya@inm.es