Ensemble forecasting probability for thresholds, lead times and models

Kadarsah Binsukandar Riadi (Indonesia)
Sei-Young Park (R. Korea)
Joanna Linkowska (Poland)

Question of Project

- Are ensembles better at forecasting probabilities with respect to higher precipitation thresholds rather than lower precipitation thresholds?

Introduction (Verification)

- **Brier score**: Accuracy

$$BS = \frac{1}{N}\sum_{k=1}^{K} n_k (p_k - \bar{o}_k)^2 - \frac{1}{N}\sum_{k=1}^{K} n_k (\bar{o}_k - \bar{o})^2 + \bar{o}(1-\bar{o})$$

reliability resolution uncertainty

- **Brier skill score** : Skill

  - proportion of improvement of accuracy over the accuracy of a standard forecast, climatology or persistence

- **Reliability Diagrams**

  - A graphical method for assessing reliability, resolution, and sharpness of a probability forecast

- **Discrimination**

  - The ability of the forecast system to clearly distinguish situations leading to the occurrence of an event of interest from those leading to the non-occurrence of the event.

- **ROC** (Relative Operating Characteristic) curve

  : False Alarm Rate vs. Hit Rate

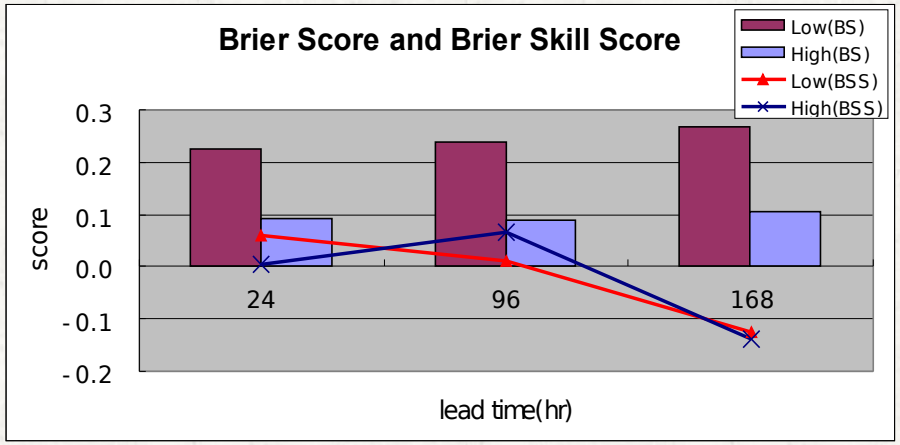
Introduction (Data)

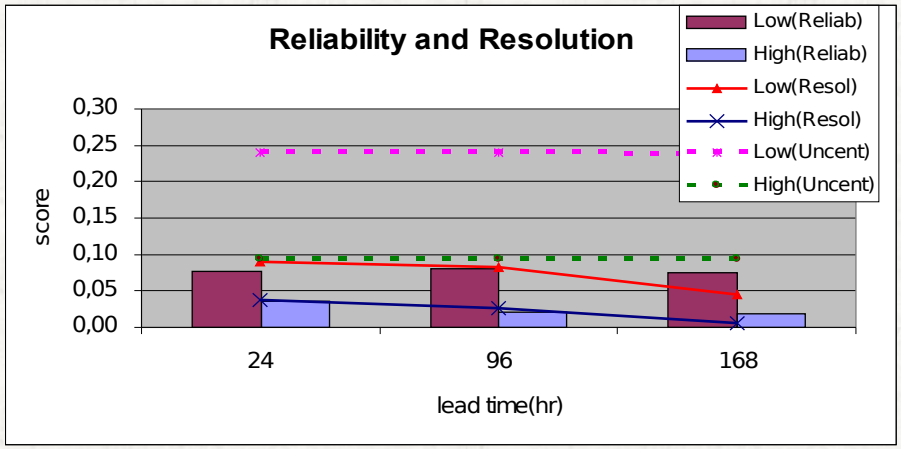
- Model : Ensemble forecasting systems
  - European Centre for Medium-Range Weather Forecasts (ECMWF) (#51)
  - Met Service of Canada (MSC) (#21)
- Thresholds
  - Low : 0.2mm/24hr
  - High : 5.0mm/24hr
- Lead times
  - 24, 96, 168hr
- Cases : Montreal,Quebec station (#315)

  - Cases with missing data for any of the models have been removed and has been matched with the ensemble forecasts at the nearest grid-point.

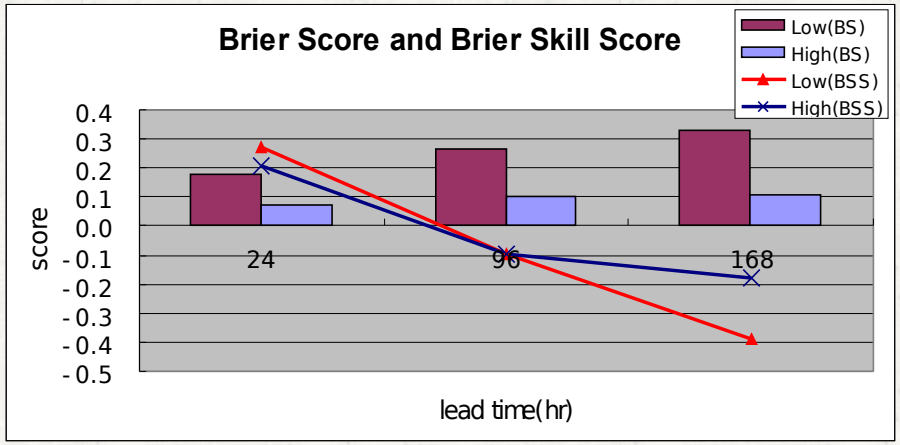
Brier Score and Skill Score

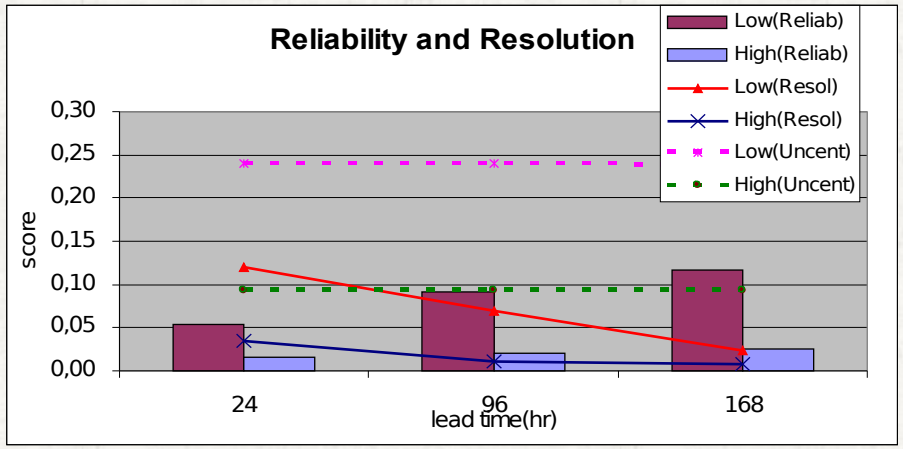
ECMWF





MSC





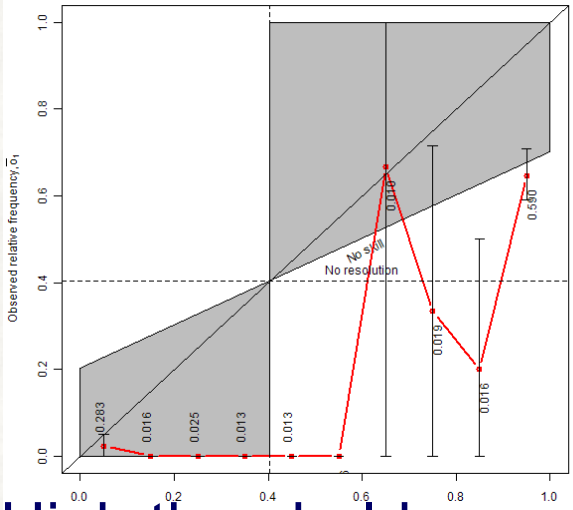
$$BS = \frac{1}{N}\sum_{k=1}^{K} n_k (p_k - \bar{o}_k)^2 - \frac{1}{N}\sum_{k=1}^{K} n_k (\bar{o}_k - \bar{o})^2 + \bar{o}(1-\bar{o})$$

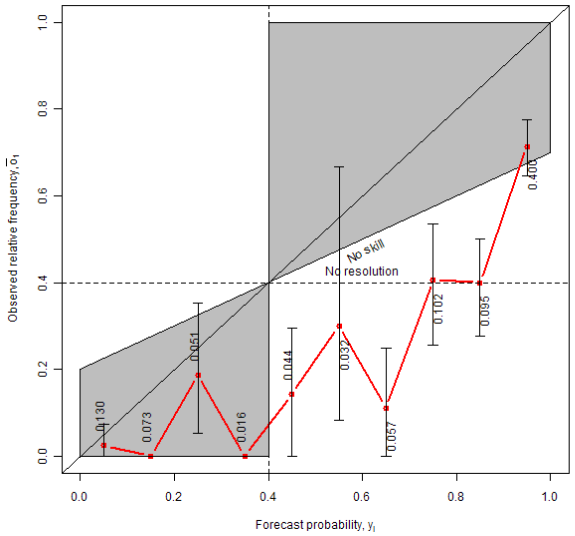
reliability resolution uncertainty

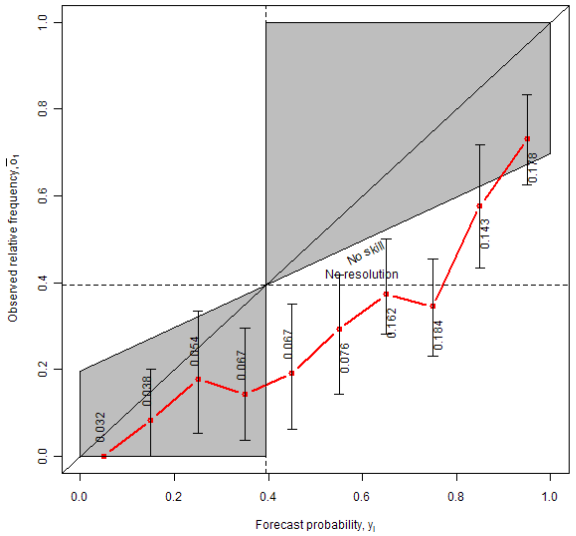
Reliability

ECMWF

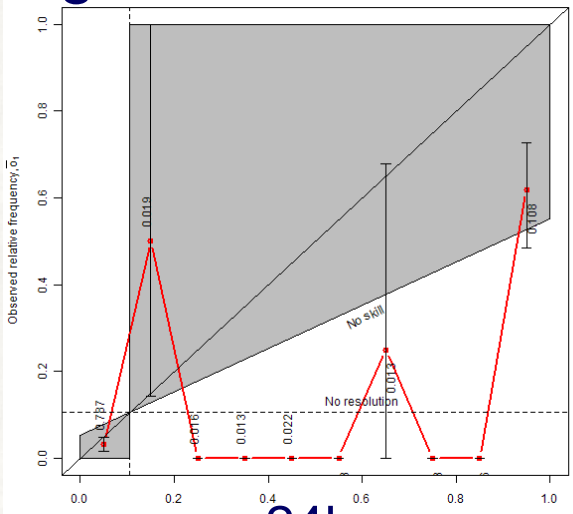
- Low threshold



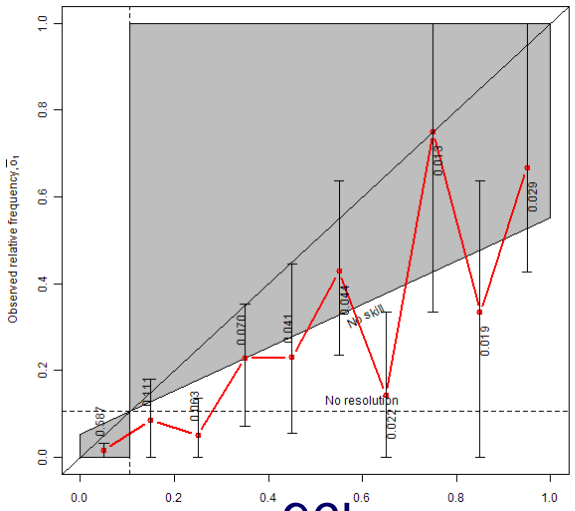




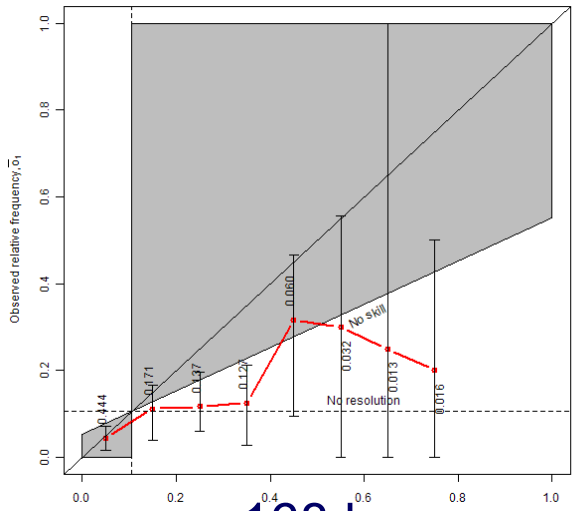
- High threshold



24hr



96hr

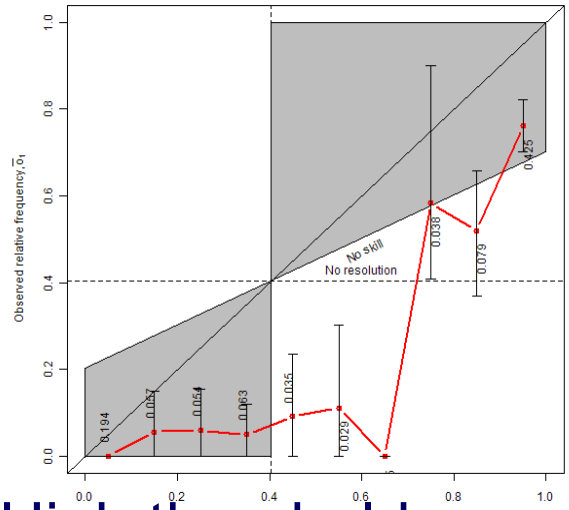


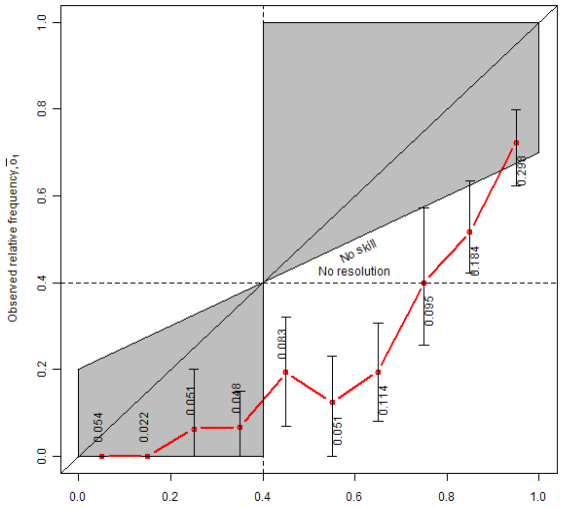
168 hr

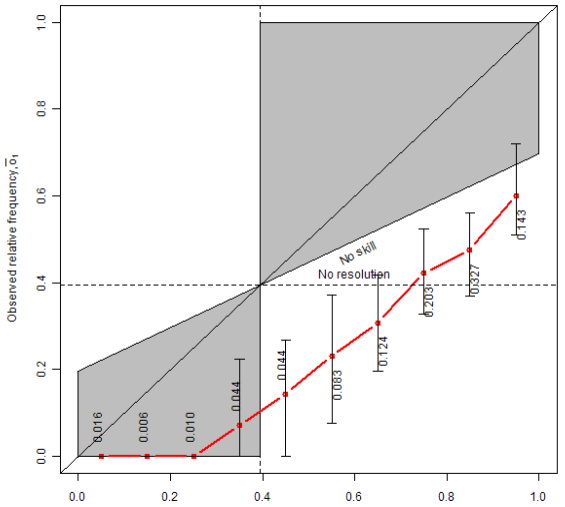
Reliability

- Low threshold

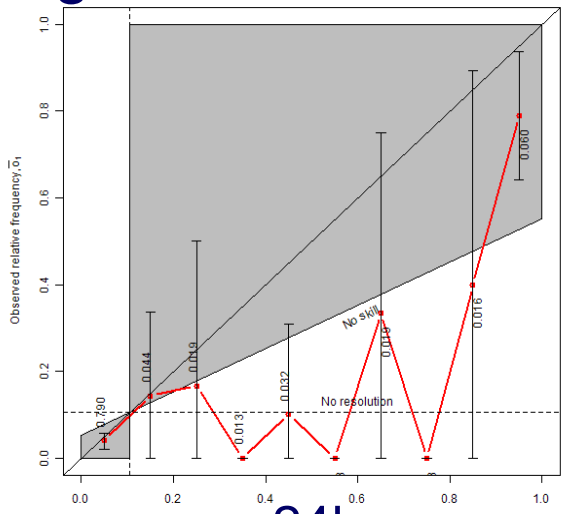
MSC



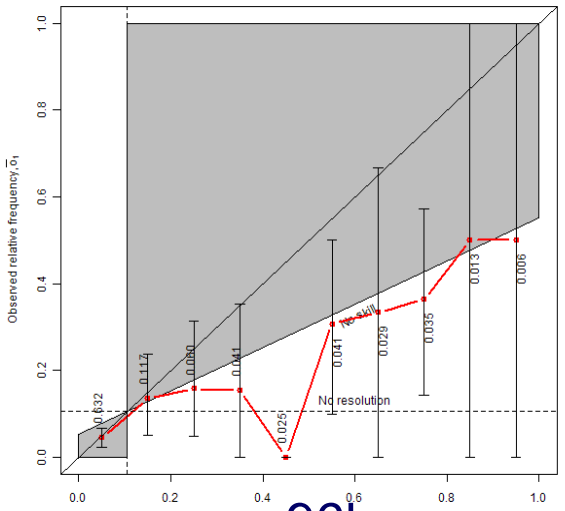




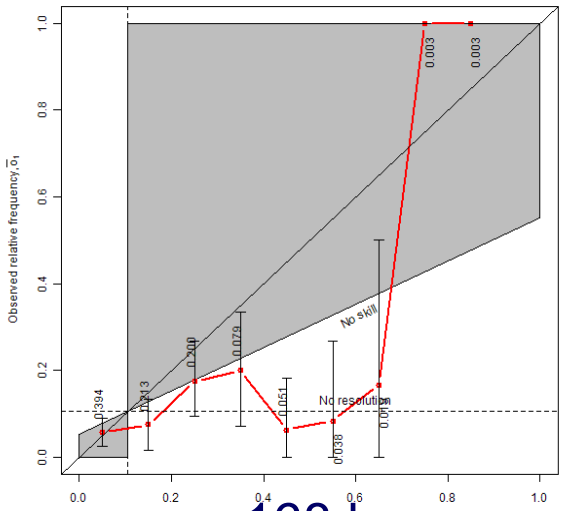
- High threshold



24hr



96hr

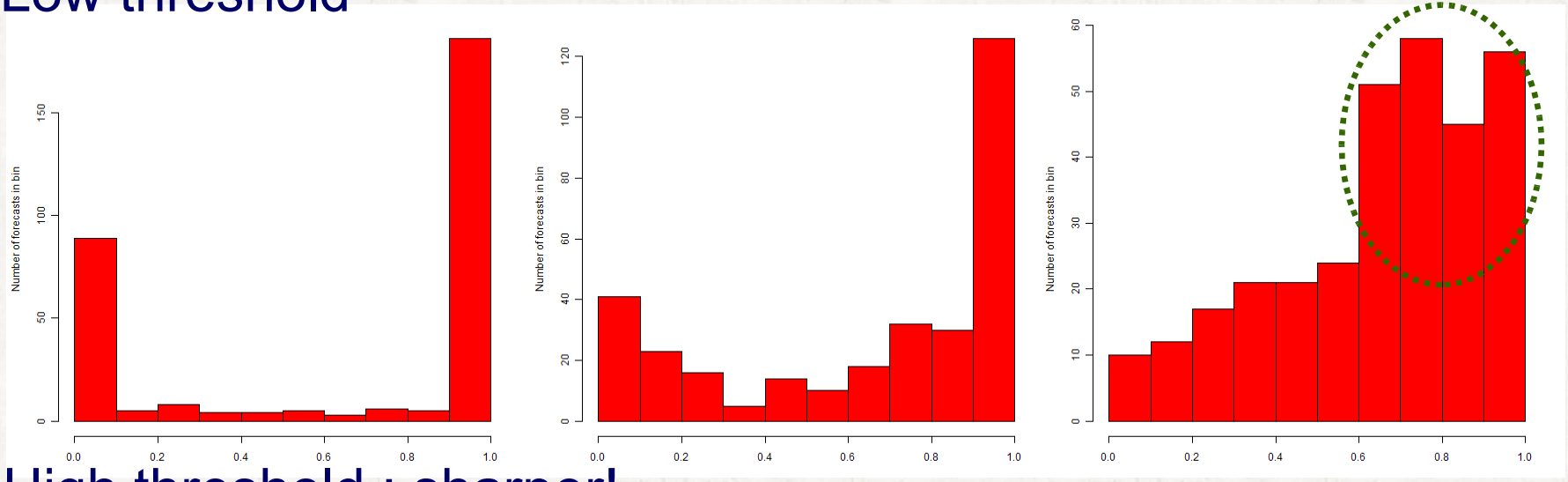


168 hr

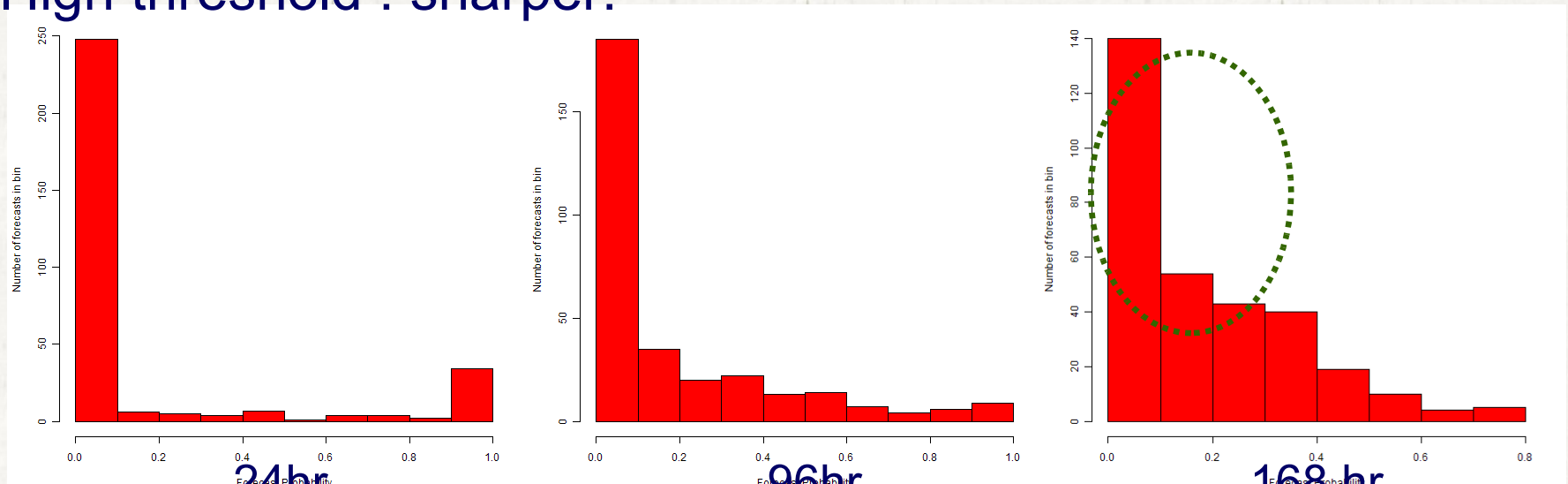
Sharpness

- Low threshold

ECMWF



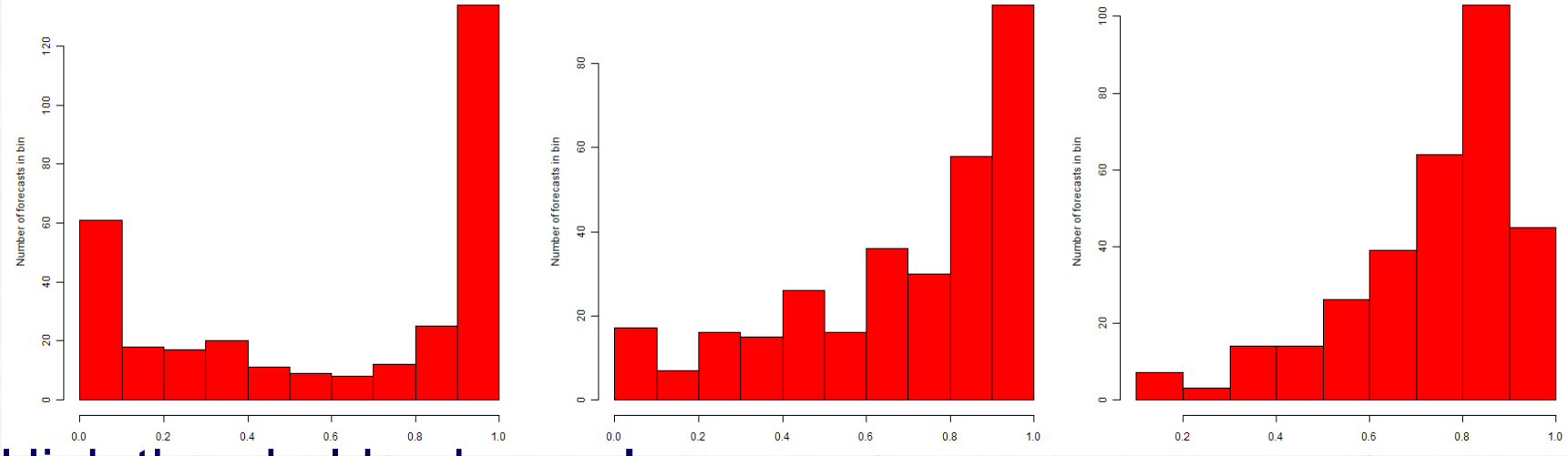
- High threshold : sharper!



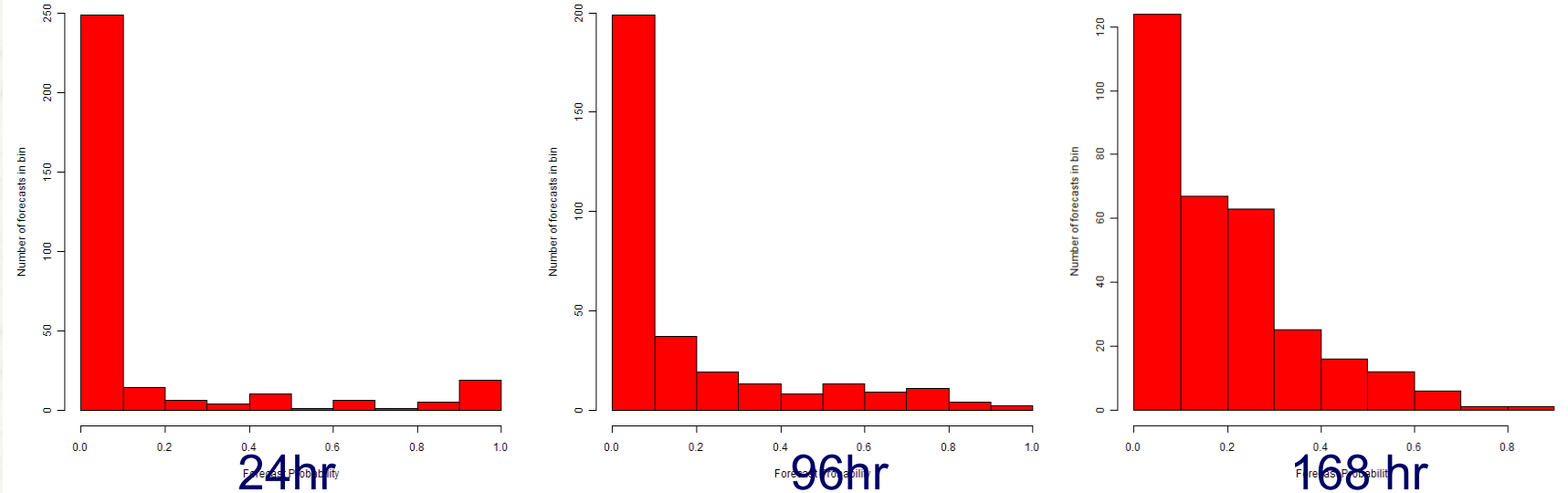
Sharpness

- Low threshold

MSC



- High threshold : sharper!



Discrimination & ROC

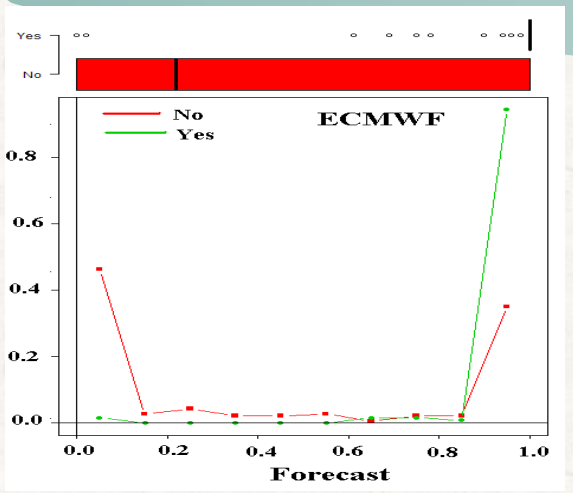
- *Discrimination* : The ability of the forecast system to clearly distinguish situations leading to the occurrence of an event of interest from those leading to the non-occurrence of the event.
  - Depends on:
    - Separation of means of conditional distributions
    - Variance within conditional distributions

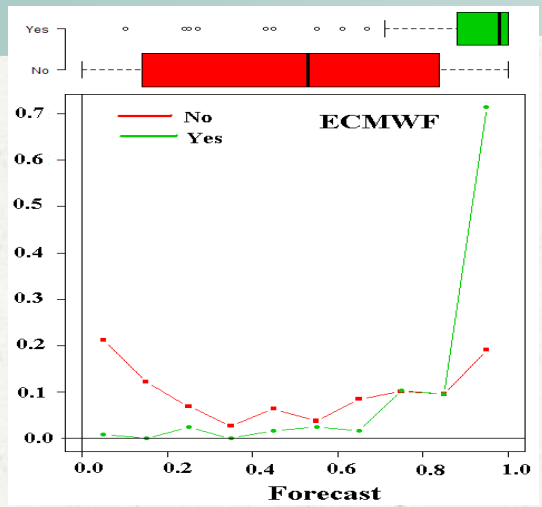
- *ROC* : The relative operating characteristic, ROC (Mason 1982), is being considered by the World Meteorological Organization as a recommended method of indicating the skill of probabilistic weather and climate forecasts. The ROC is a highly flexible system that can be used to assess the skill level of dichotomous, categorical, continuous, and probabilistic forecasts. The ROC curve is useful for identifying an optimal strategy for issuing warnings, by indicating the trade-off between false alarms and misses.

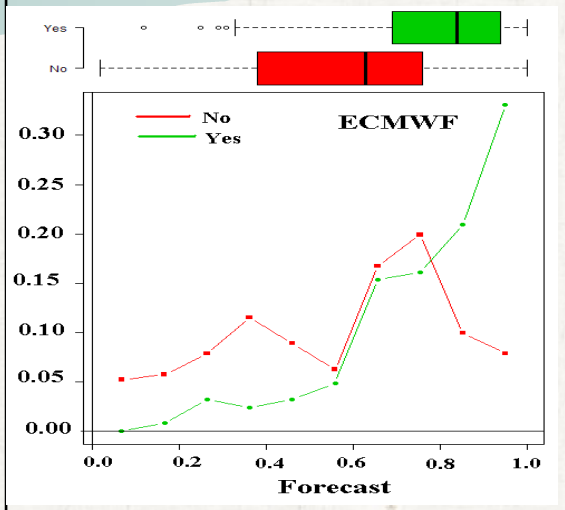
Discrimination

ECMWF

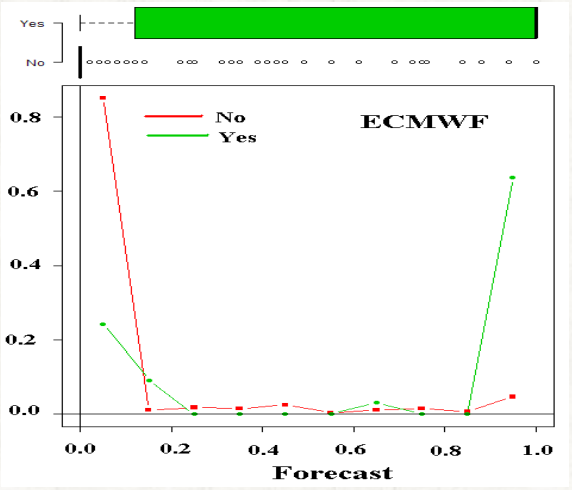
Low threshold

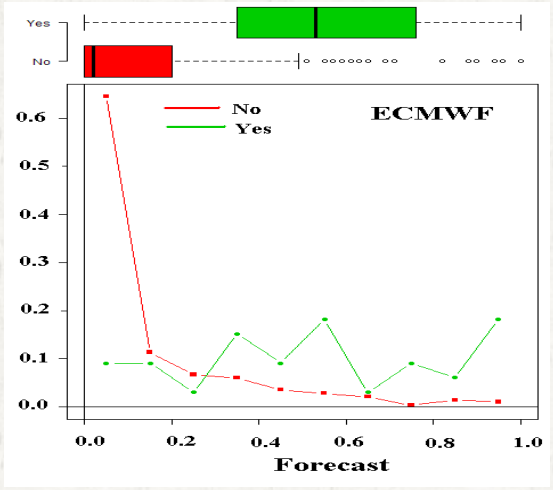


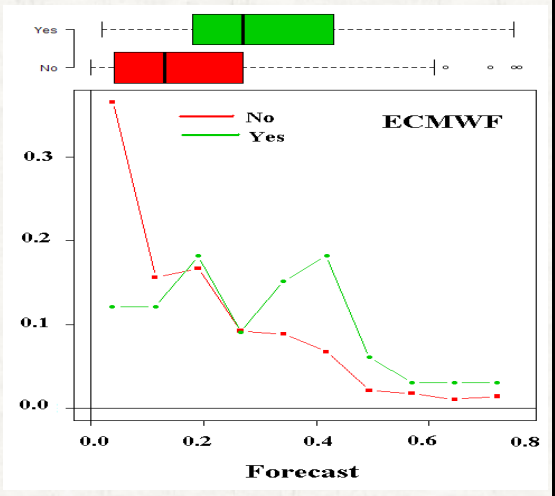




High threshold







24hr

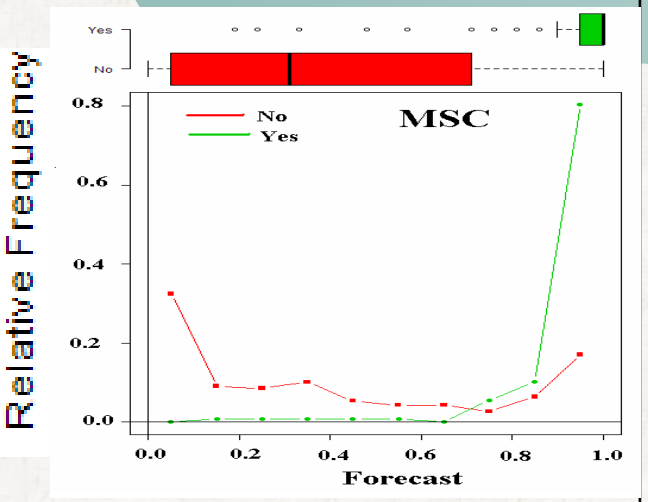
96hr

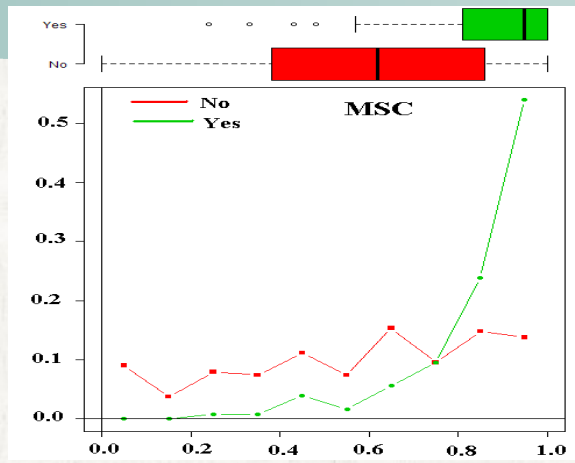
168hr

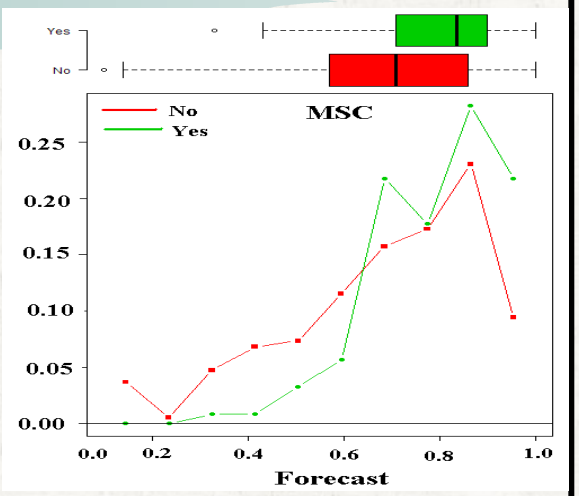
Discrimination

MSC

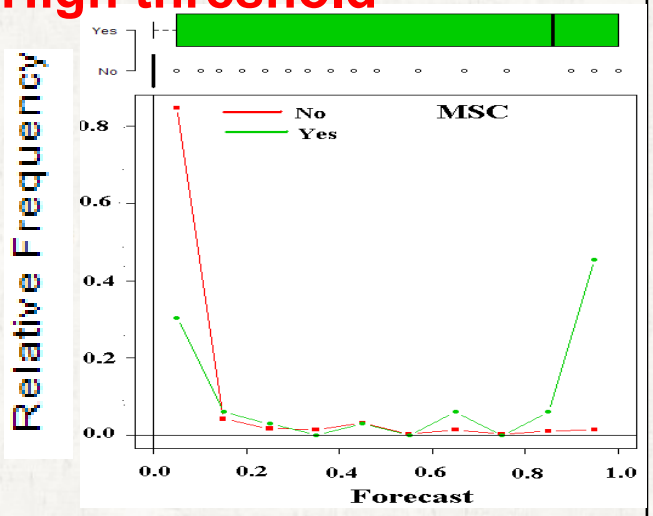
Low threshold

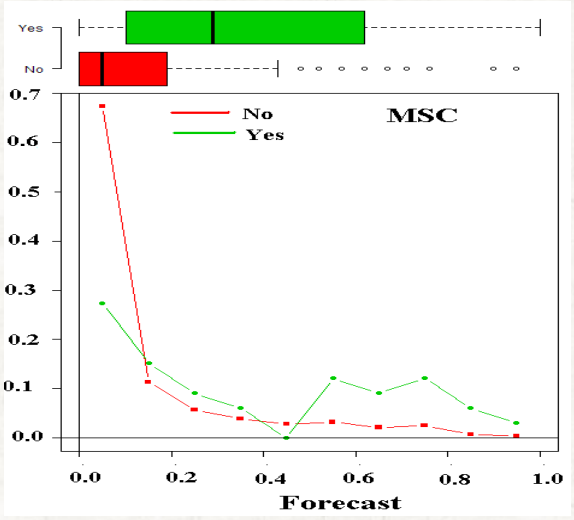


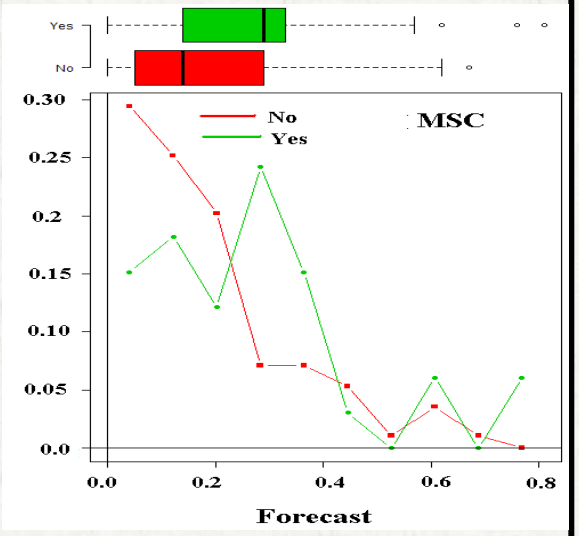




High threshold





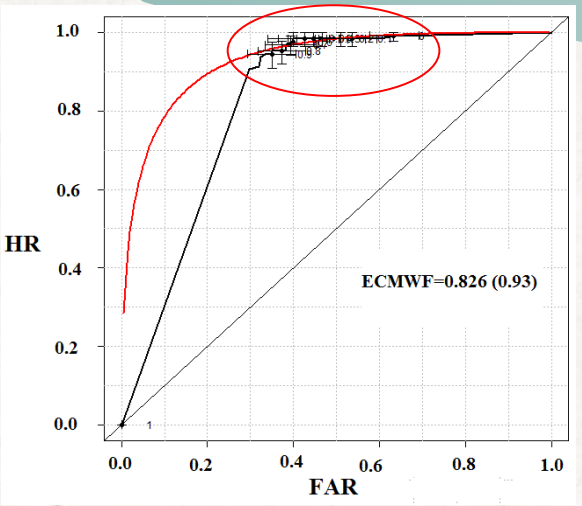


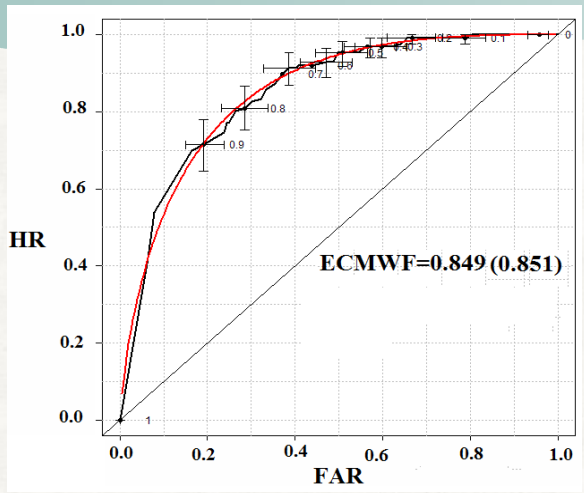
24hr **96hr** **168hr**

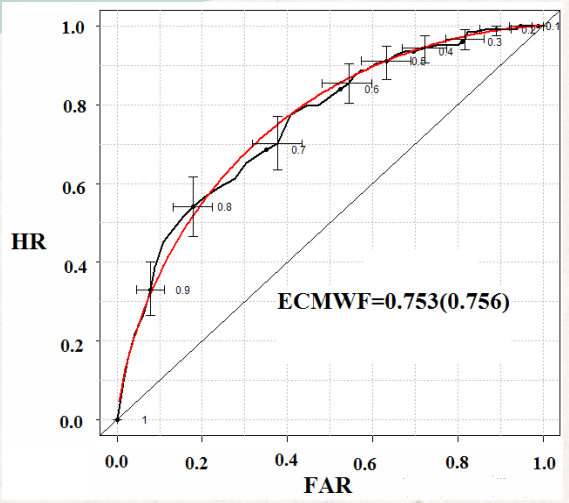
ROC Curve

ECMWF

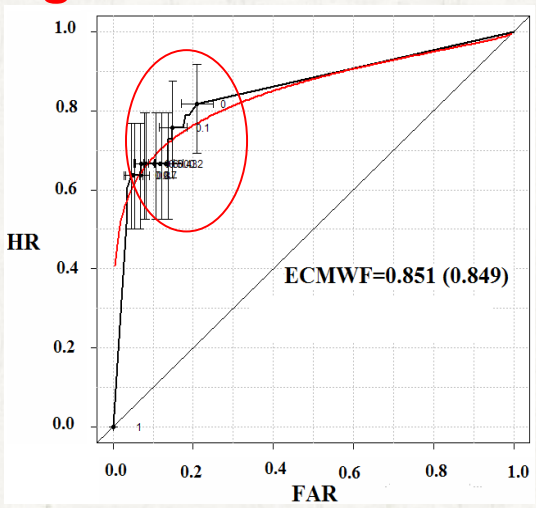
Low threshold

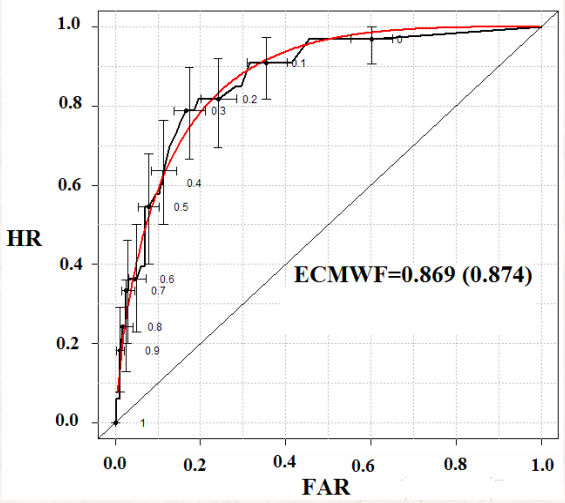


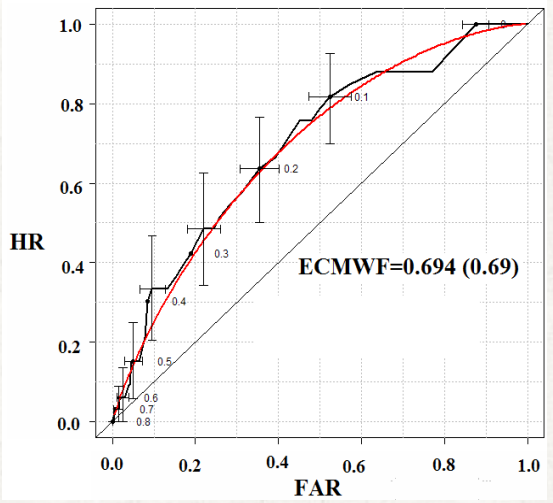




High threshold





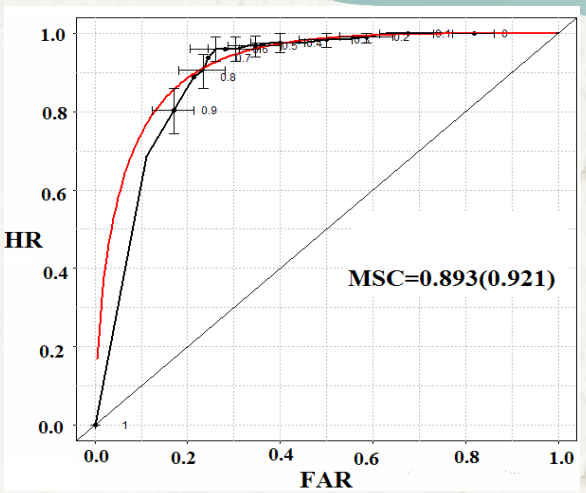


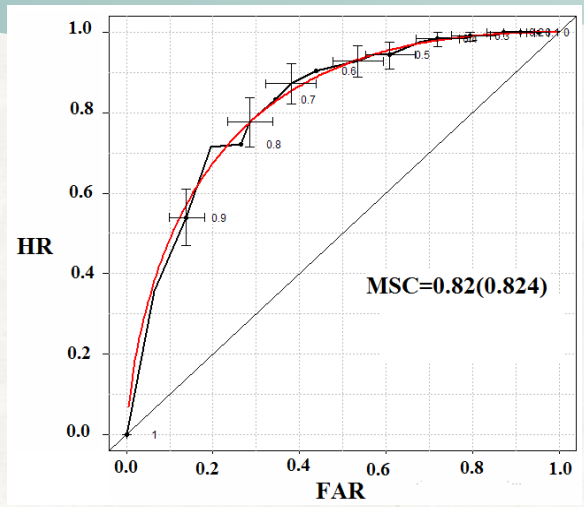
24hr 96hr 168hr

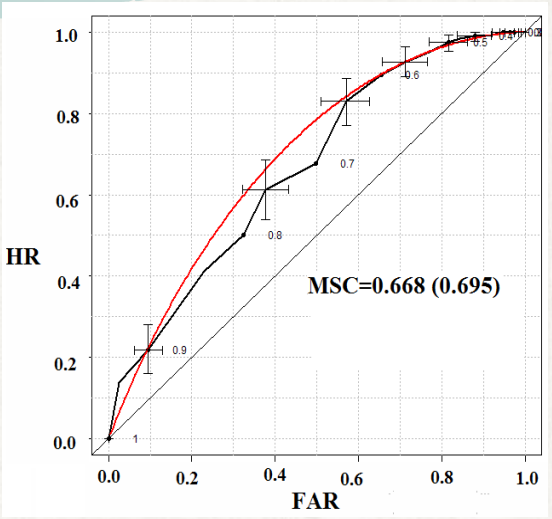
ROC Curve

MSC

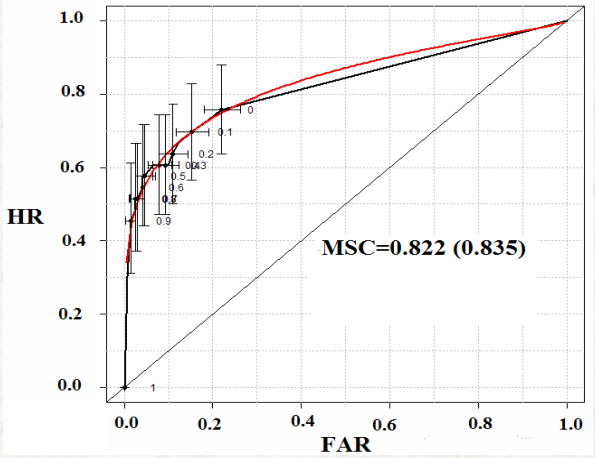
Low threshold

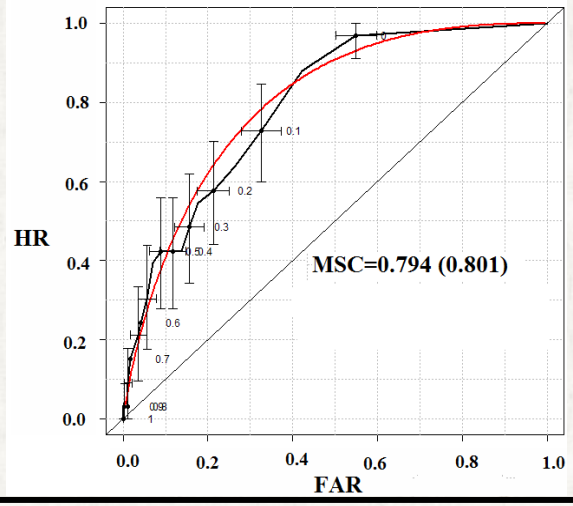


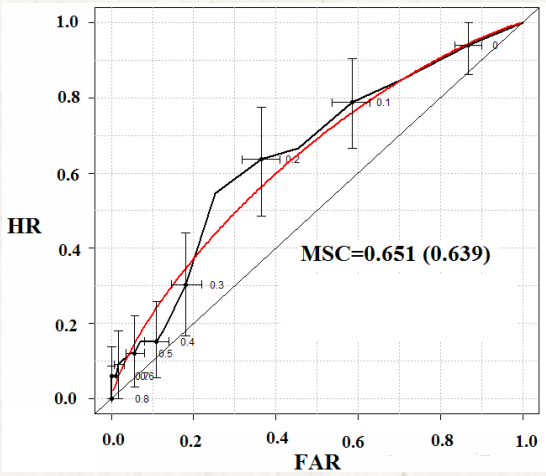




High threshold







24hr

96hr

168hr

ROC Curve

Threshold	ROC score					
	ECMWF			MSC		
Lead Time(hr)	24	96	168	24	96	168
Low	0,826	0,849	0.753	0,893	0,820	0,668
High	0.851	0,869	0,694	0.822	0,794	0,651

ECMWF:

- For 24hr-96hr, Hightreshold is better than lowthreshold
- At 168 hr, ROC score is decreased both high and low threshold

MSC:

- For 24hr-168hr, lowthreshold is better than highthreshold

Summary

- The reliability is going up as the lead time is longer. For high threshold, it looks better than low threshold. But because the sample size is so small, it's hard to compare their reliabilities.
- BS of Low is higher than High and the scores are going up as the lead time is longer for both models.

  - That's why the reliabilities are high and resolutions are low.
- BSS is also decreased as the lead time is longer.
- High threshold is sharper than low. And there are the tendencies that low threshold is shifted to high Forecast probability and high threshold is shifted to low one.
- Generally, the discriminations of ECMWF &MSC are increasing as the lead time (24hr-96hr)
- ROC score of
  - ECMWF is increased as the lead time (24hr-96hr)
  - MSC is decreased as the lead time.

Conclusion

- High thresold shows the better result than low.
- ECMWF has the best result at 96hr lead time, but MSC has at 24hr.
- But!! To get the better results, we need sufficient data.