Verification of Ensemble Probability of Precipitation Forecasts

Project No.1

Tirivanhu Muhwati (Zimbabwe)
Petra Roiha (Finland)
Flora Gofa (Greece)

4th International Verification Methods Tutorial
HELSINKI, FINLAND
JUNE 2009

DATASET AND MAIN AIM

- Dataset of precipitation accumulation observations for one location in Greece accompanied by the corresponding COSMO-LEPS 16 member ensemble forecast.

- Main Aim – to evaluate the probability of precipitation forecasts generated from the ensemble.

DATA SPECIFICATION

Forecasted 20 day values for a single station (16741 - Athens Airport) for October 2007.

- Forecast period of 72hrs.

- Weather parameter examined: 24hr precipitation.

- Rainfall amounts were very little as a result the use of a rain or no rain predictant.

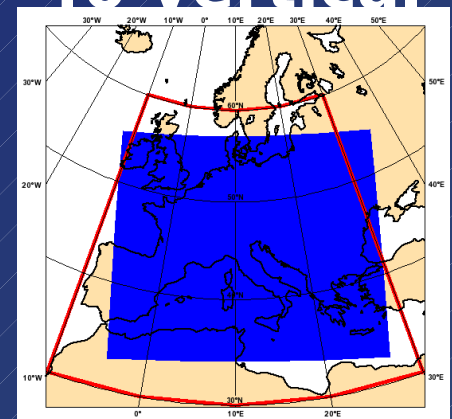# System setup: ic and bc: multi-model approach

IFS – ECMWF global

COSMO at 25 km on IFS

GME – DWD global

COSMO at 25 km on GME

UM – UKMO global

COSMO at 25 km on UM

AVN – NCEP global

COSMO at 25 km on NCEP

by INM Spain

P1: conv. scheme (Tke)

P2: conv. scheme (KF)

P3: turb. Parameter 1

P4: turb. parameter 2

16 COSMO runs 10 km hor. res. 40 vertical

Brier Skill Score

- BSS measures the improvement of the probabilistic forecast relative to the sample climatology.

$$BSS = \frac{resolution - reliability}{uncertainty}$$

- **Brier Score** : Scalar summary measure for the assessment of the probabilistic forecast performance, mean-squared error of the probability forecast
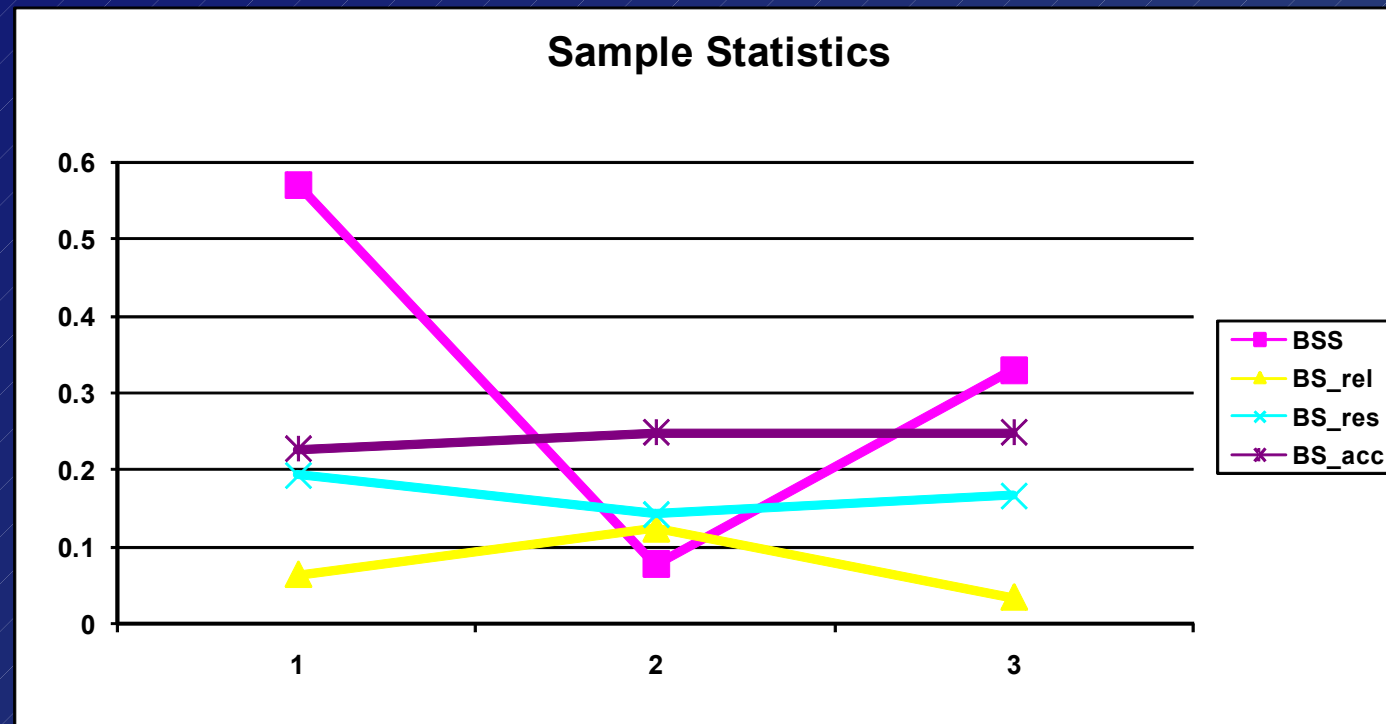
Statistical Properties

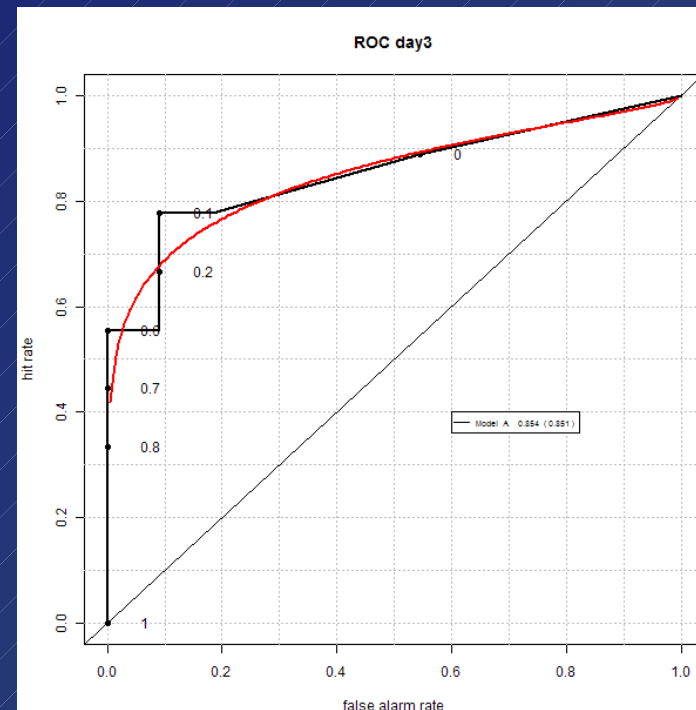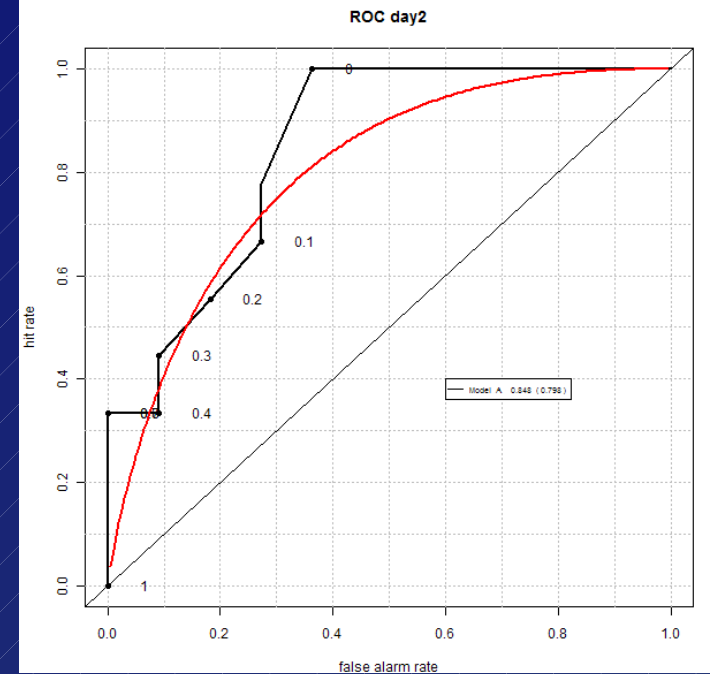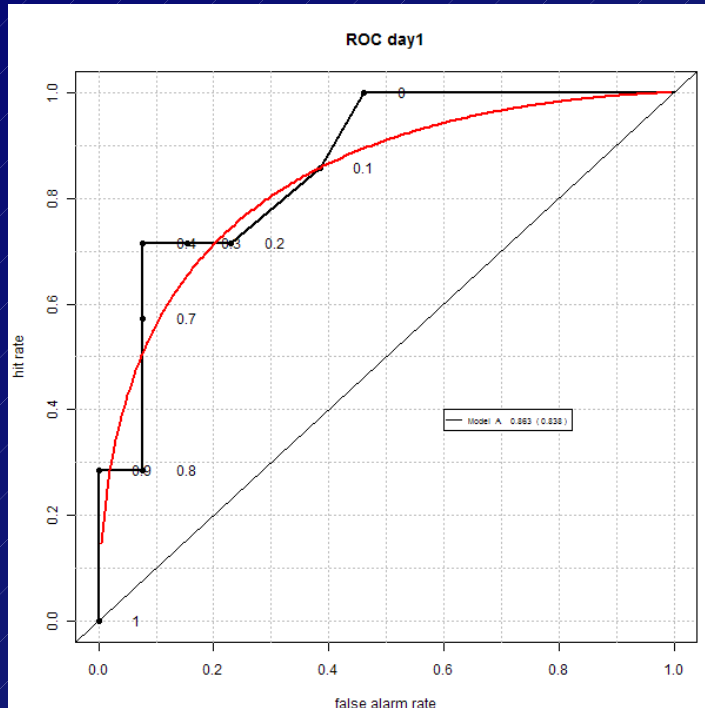|	DAY1	DAY2	DAY3
BS	0,098	0,22	0,168
BSS	0,57	0,077	0,33
BS_resol	0,194	0,143	0,166
BS_reliab	0,064	0,124	0,034
BS_acc	0,227	0,247	0,248

Perfect Brier Score = 0

Day 1 exhibits the best results

Perfect Brier Skill Score = 1

Day 1 exhibits the best results.

Sample Statistics

ROC Area Plots

ROC measures the ability of the forecast to discriminate between two alternative outcomes.
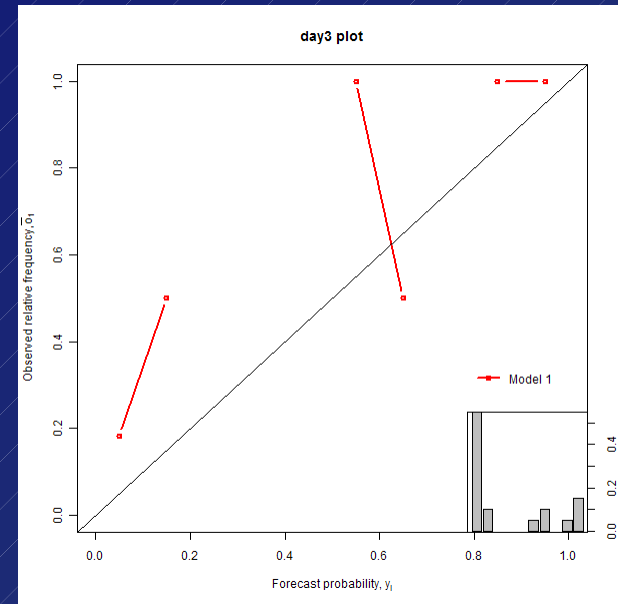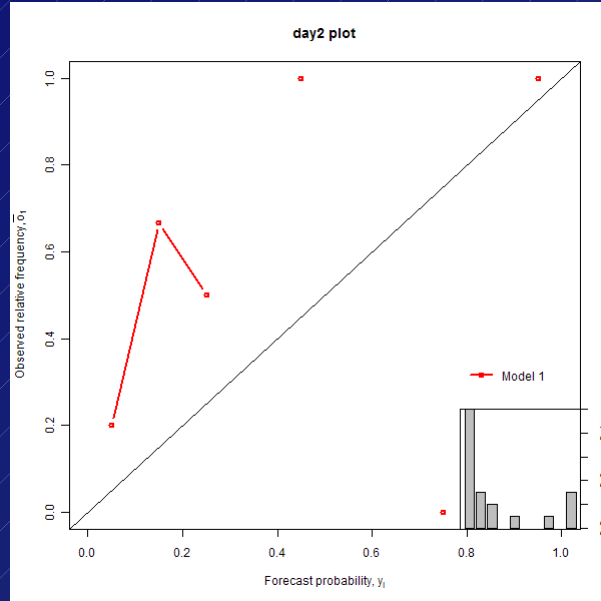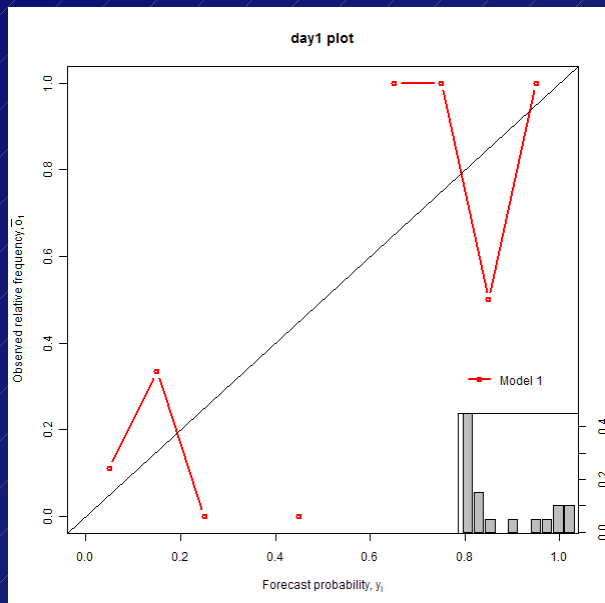
ROC area day1=0.838

day2=0.798

day3=0.851

The forecast has skill. Least skill for day 2.

Reliability Diagrams

- A graphical method for assessing reliability, resolution, and sharpness of a probability forecast
- Sometimes called "attributes" diagram.

Reliability Diagrams

- Incomplete reliability plots for forecast days 1, 2 and 3 suggesting too small a sample size

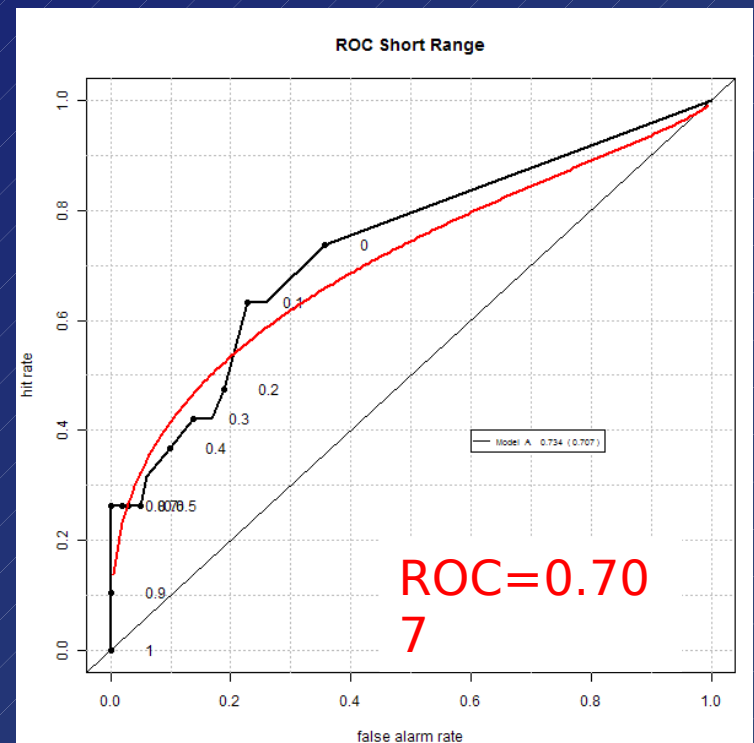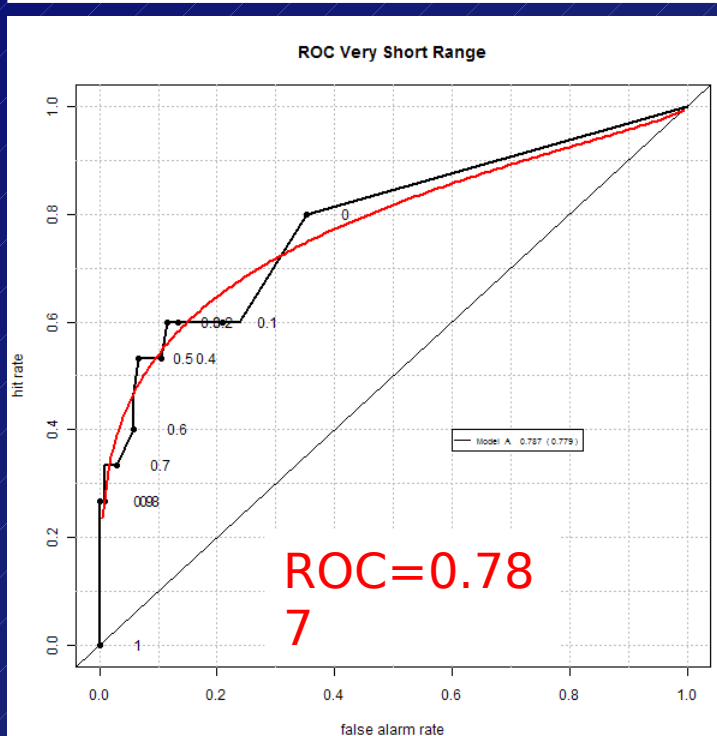# Need to aggregate the data to provide larger data set

- Very Short Range

  00h-36h

- Short Range

  42h-72h
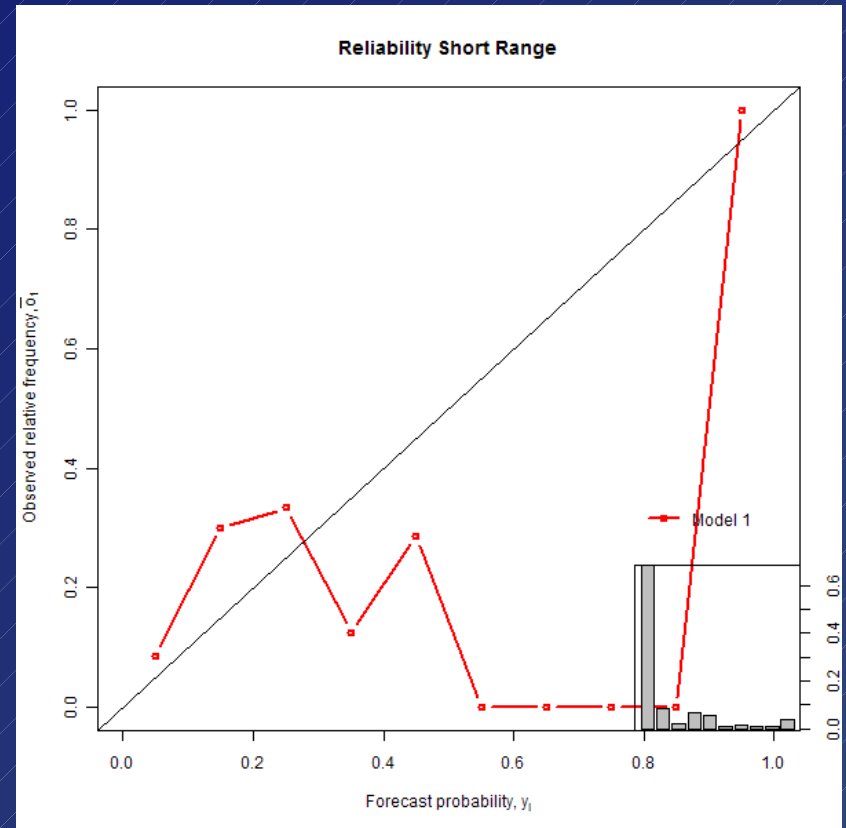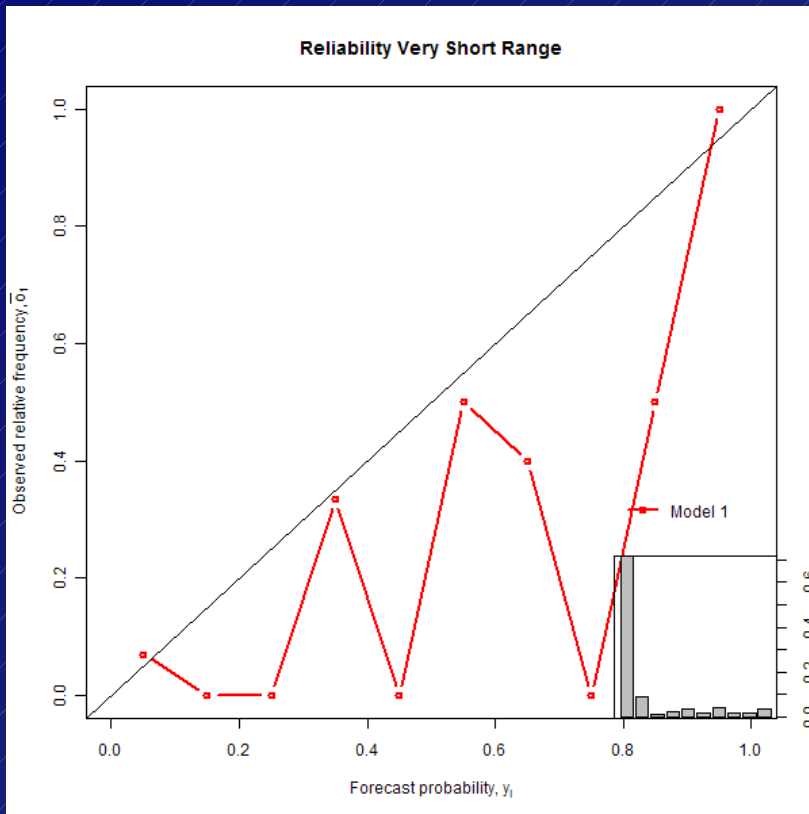
Very short range and short range – BS, BSS and ROC

	VS	S
BS	0,094	0,124
BSS	0,14	0,071
BS_resol	0,039	0,038
BS_reliab	0,024	0,0282
BS_acc	0,11	0,133

	DAY1	DAY2	DAY3
BS	0,098	0,22	0,168
BSS	0,57	0,077	0,33
BS_resol	0,194	0,143	0,166
BS_reliab	0,064	0,124	0,034
BS_acc	0,227	0,247	0,248

ROC=0.787

ROC=0.707

Very short range and short range - reliability

CONCLUSIONS

- Reliability diagram computation requires a fairly large dataset, because of the need to partition the sample into subsamples conditional on forecast probability.

- ROC scores better for the first unaggregated data sample with the BS showing otherwise.

- Although the scores showed some level of skill, it is not possible to come up with concrete conclusions on the quality the forecasts. Need for

Thank You