www.ec.gc.ca
Lawrence.wilson@ec.gc.ca

Verification of multi-model ensemble forecasts using the TIGGE dataset

Laurence J. Wilson
Environment Canada

Anna Ghelli
ECMWF

Outline

- What is TIGGE?
- Some early verification results using TIGGE data
- Precipitation verification project:
  - plan
  - early results
- Summary

What is TIGGE?

- A Subproject of THORPEX
- Goals:

(i) enhancing collaboration on ensemble prediction, between operational centres and universities;
(ii) facilitating research on ensemble prediction: development of methods to combine ensembles from different sources; to correct for systematic errors;
(iii) enabling evolution towards a prototype operational system, the “Global Interactive Forecast System (GIFS)”

- Effectively – a very large database of global ensembles from 10 centers, collected in 3 distribution centers, ECMWF, NCAR and CMA
- Basic TIGGE question: Can probability forecasts be improved by combining ensembles from several centers compared to single-center ensembles?
- VERIFICATION needed to answer this question.

Structure of TIGGE Database

- Up to 73 surface and upper air variables
- 6 hr frequency, one to 4 runs per day, projections as far as model runs, up to 16 days.
- Native model resolution, but archive centers have “standardizing” software
- Typically 1 by 1 degree
- ~300 GB per day.
- Data available to all, “near” real time.
- Metadata available, about all models, but not yet complete

Status of Verification of TIGGE ensembles

- Mostly model-oriented verification so far
  - Upper air data
  - Against analyses
  - Standard scoring and case studies
- Studies on the TIGGE website
  - Park et al, 2008
    - First study involving several months of data
    - RPSS main score used; climate from ERA 40
    - Found modest improvement with combined ensembles, greatest benefits in tropics
    - Including “poorer” ensembles didn’t help
  - Pappenberger et al.
    - Case study of flooding event in Romania
    - User-oriented, Q-Q plots, RPS and RMSE main scores used
    - Individual ensembles and combined ensemble compared.

Park et al study – impact of analysis used as truth in verification

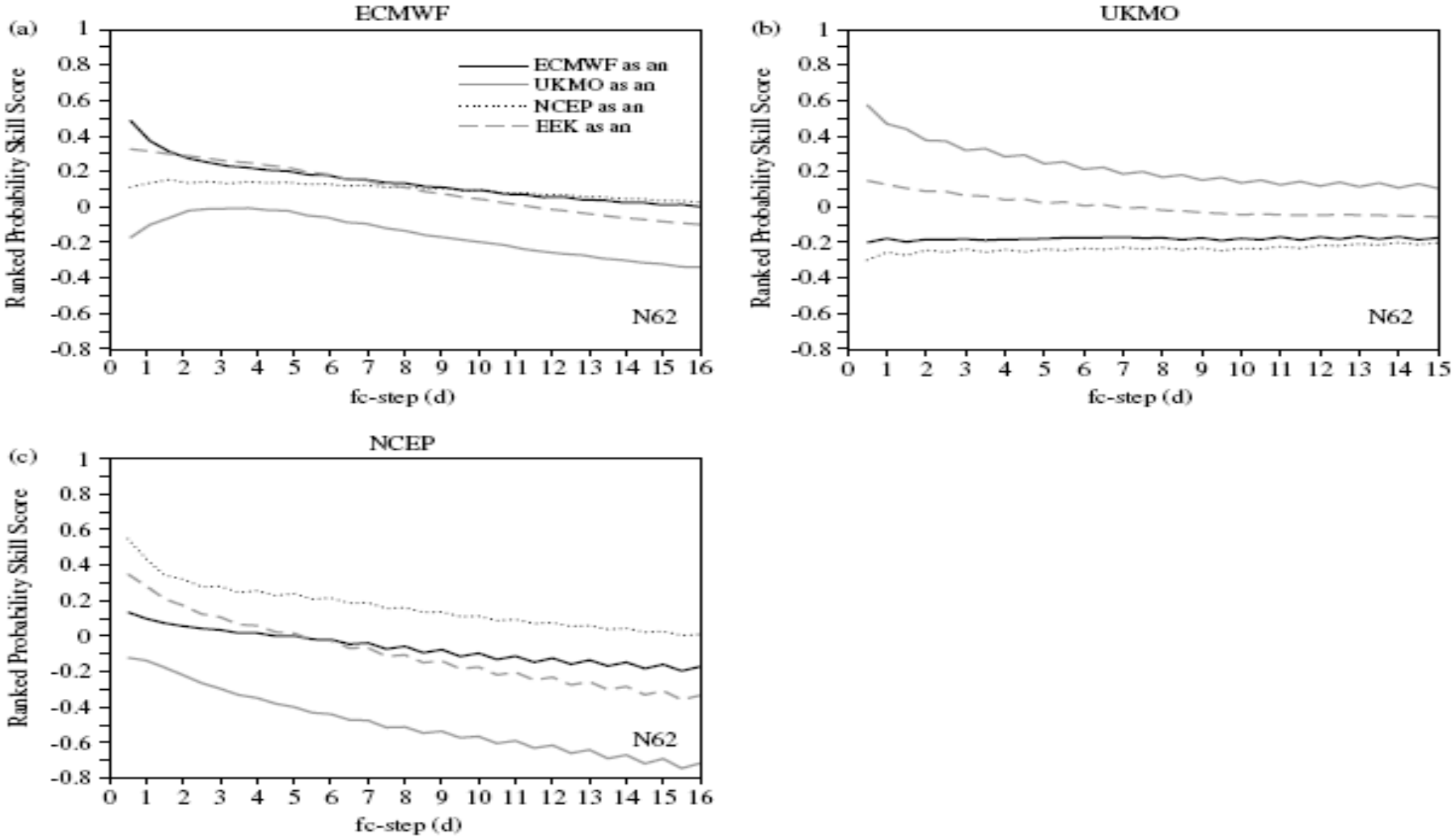


Figure 14. Sensitivity to verification analysis: AM07 (62 cases) average rank probability skill score of the probabilistic forecasts of T850 over the Tropics given by (a) the EC ensemble, (b) the UKMO ensemble and (c) the NCEP ensemble verified against the EC analysis (solid black line), the NCEP analysis (dotted black line), the UKMO analysis (solid grey line) and the mean of the three analyses (dashed grey line).

Status of TIGGE – related verification

- Current efforts – Is verification of surface variables happening?
  - Studies using spatial methods
    - Ebert – application of CRA technique to ensemble forecasts. So far, only ECMWF.
    - Application of Wilks minimum spanning tree – rank histogram for TC centers. (idea stage)
  - Precipitation verification project:

Precipitation verification project

- Goal: to verify global 24h precipitation forecasts from all the ensembles in the TIGGE archive
- One region at a time, using highest density observations
- Canada and Europe so far
- Methodology:
  - Cherubini et al upscaling, verify only where data available
  - Kernel density fitting following Peel and Wilson to look at extremes of distributions.

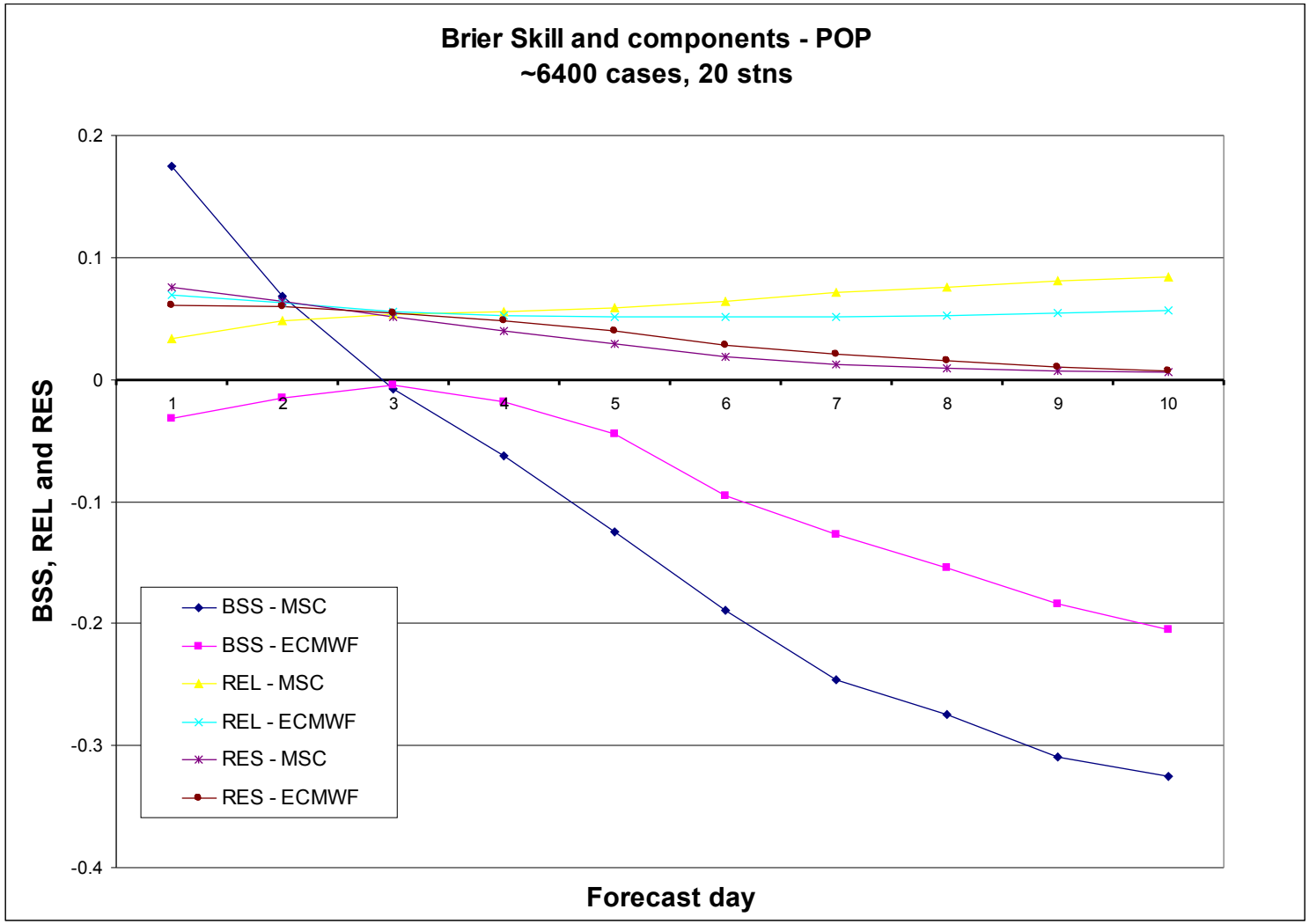
Precipitation verification project : methodology - Europe

- Upscaling:
  - 1x1 gridboxes, limit of model resolution
  - Average obs over grid boxes, at least 9 stns per grid box (Europe data)
  - Verify only where enough data
  - Matches obs and model resolution locally
  - Answers questions about the quality of the forecasts within the capabilities of the model
  - Most likely users are modelers.

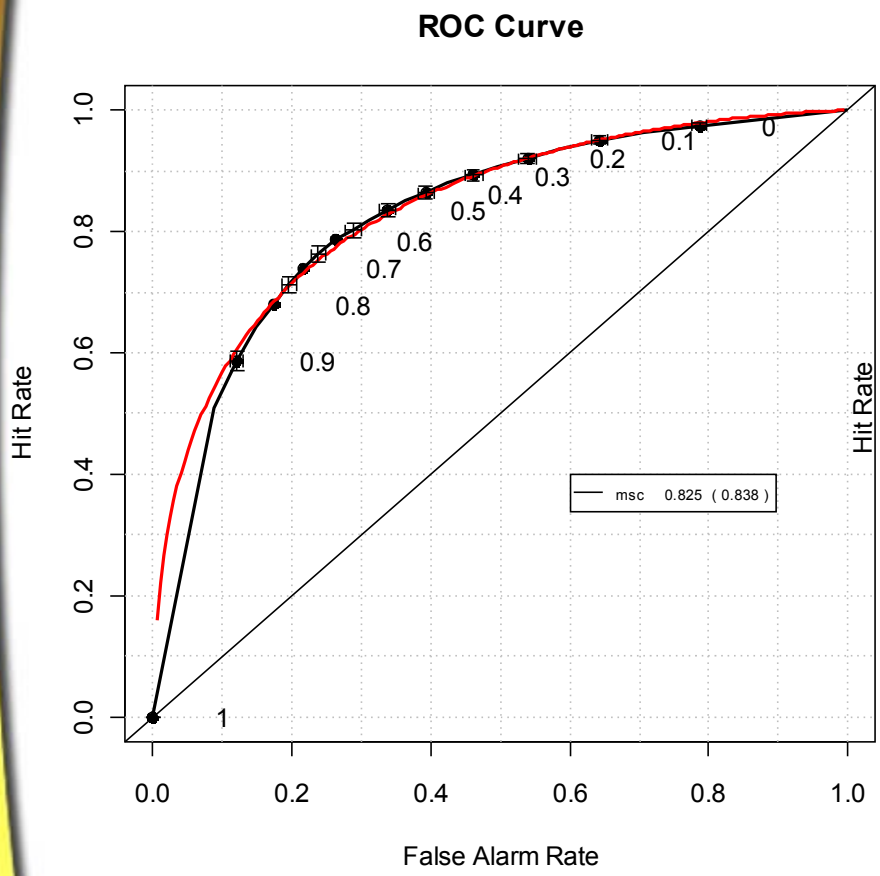
Precipitation verification project: methodology - Canada

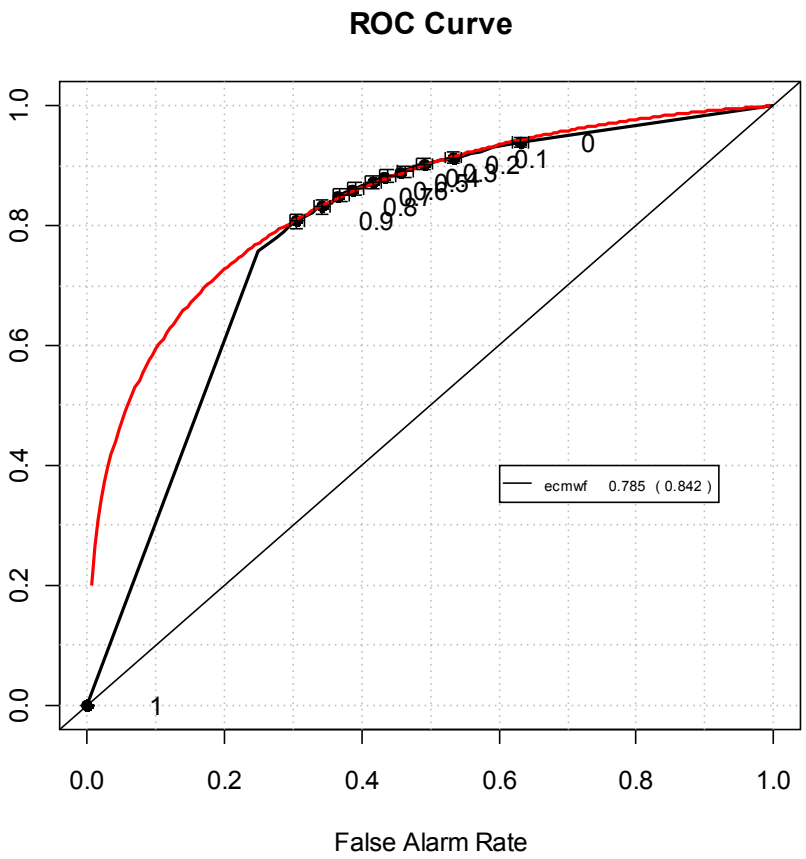
- Single station verification
  - Canadian verification over 20 widely-spaced stations, only one station per gridbox; comparison of nearest gridpoint fcst to obs
  - Pointwise verification, does not (we cannot) upscale properly because don't have the data density necessary.
  - Valid nevertheless as absolute verification of model predictions

Results – Canada – Brier Skill, Resolution and Reliability

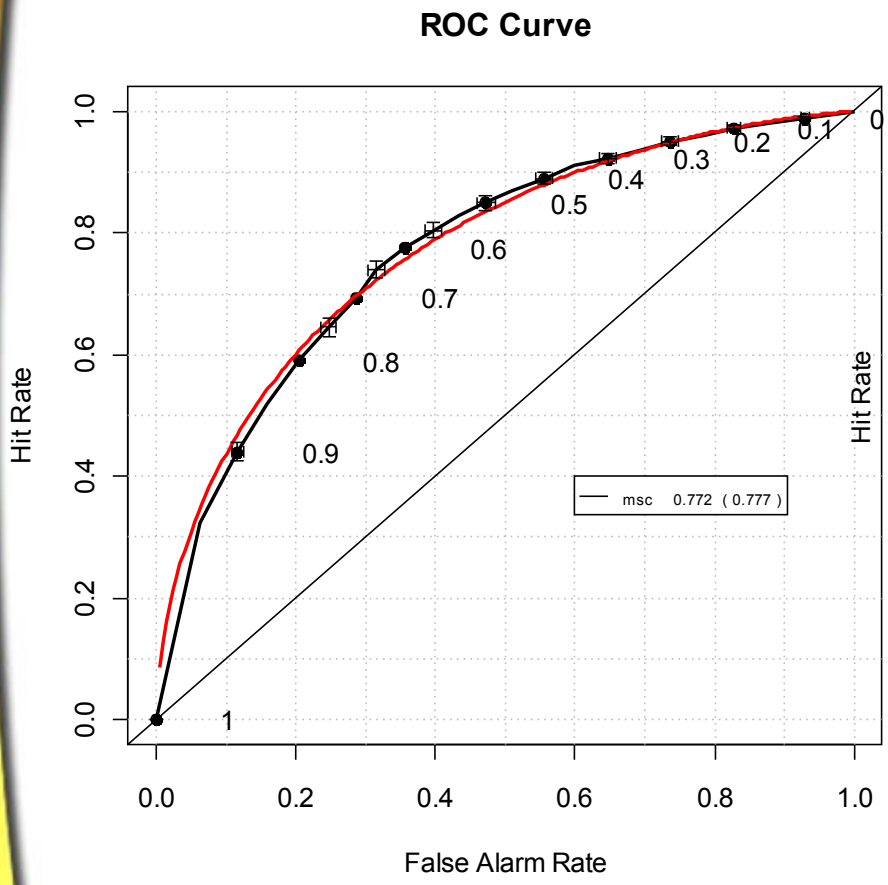


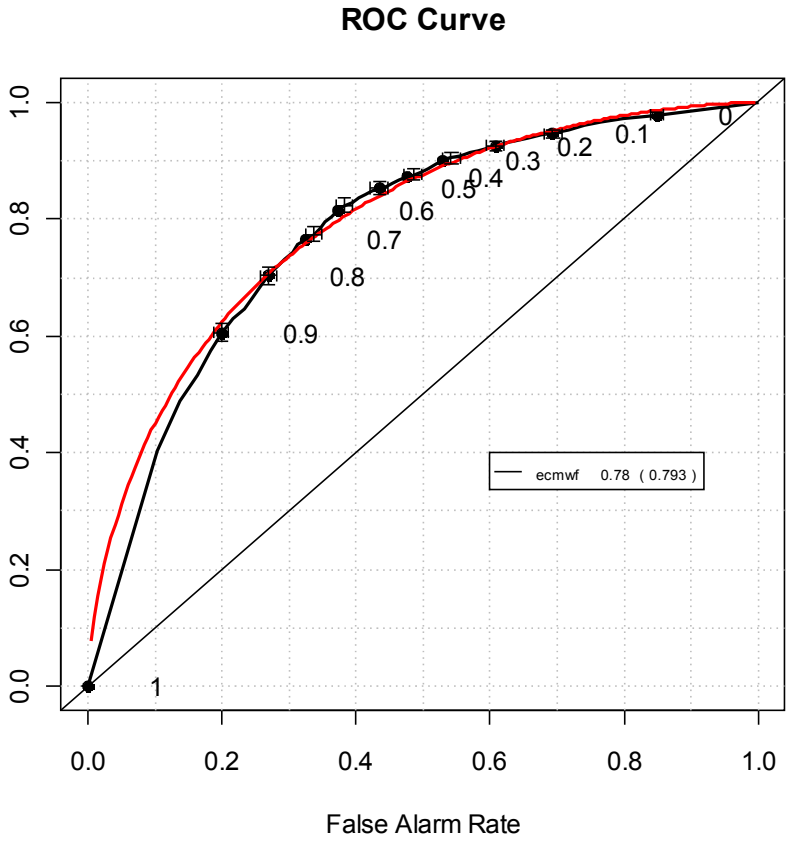
Results – Canada – ROC curves – 24h



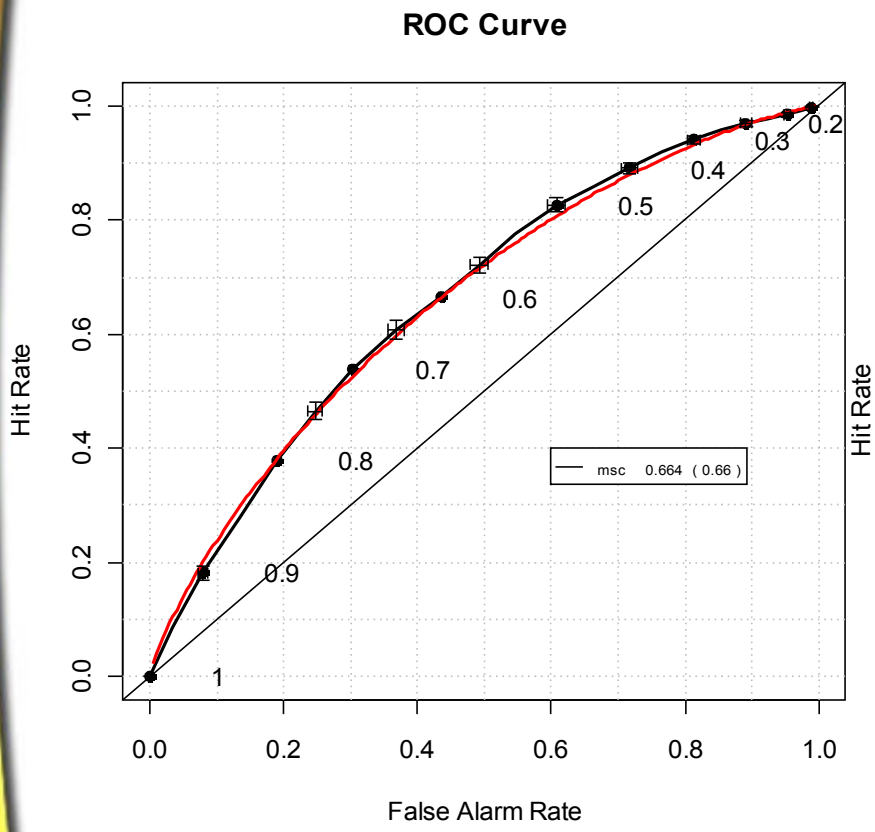


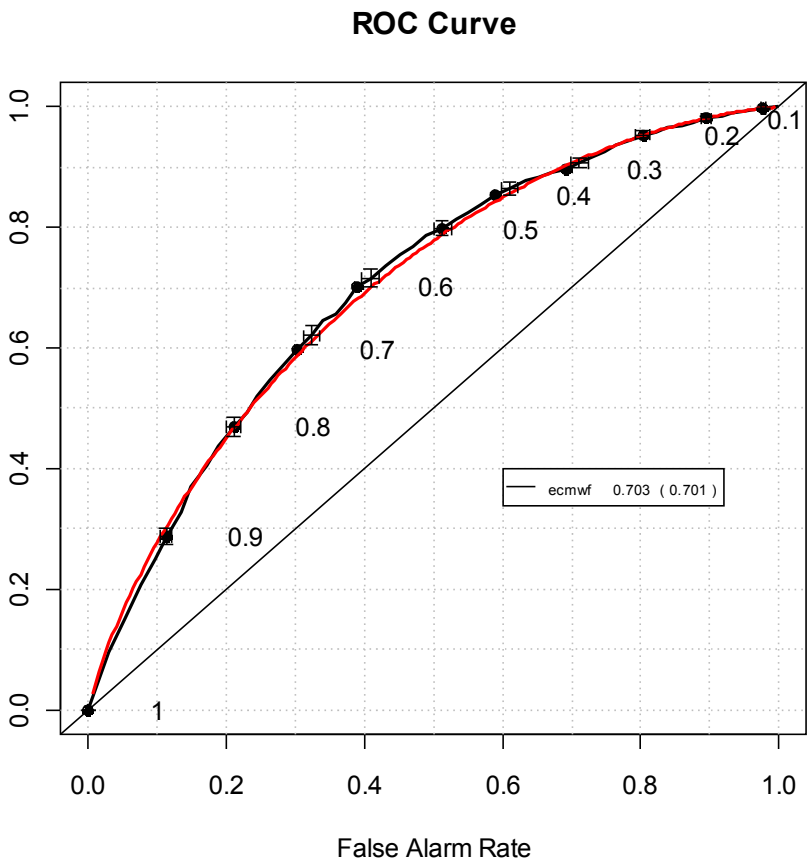
Results – Canada – ROC Curves – 72h



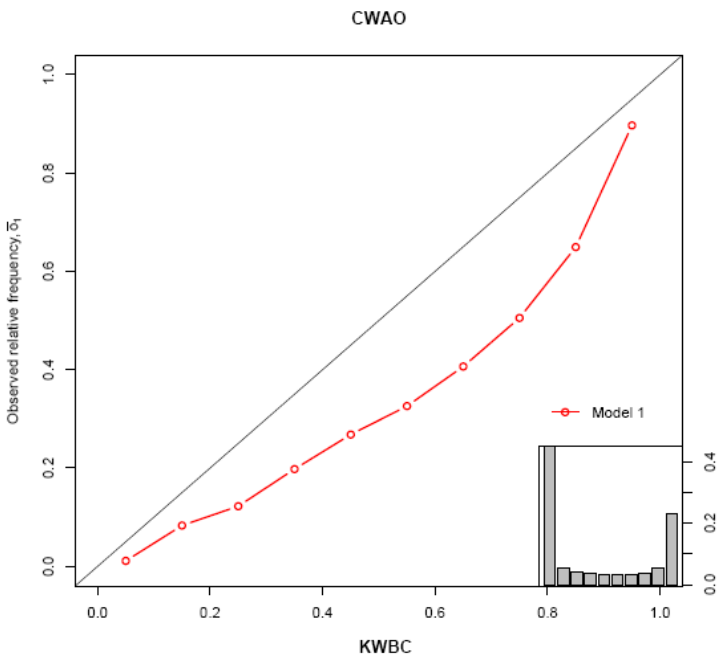


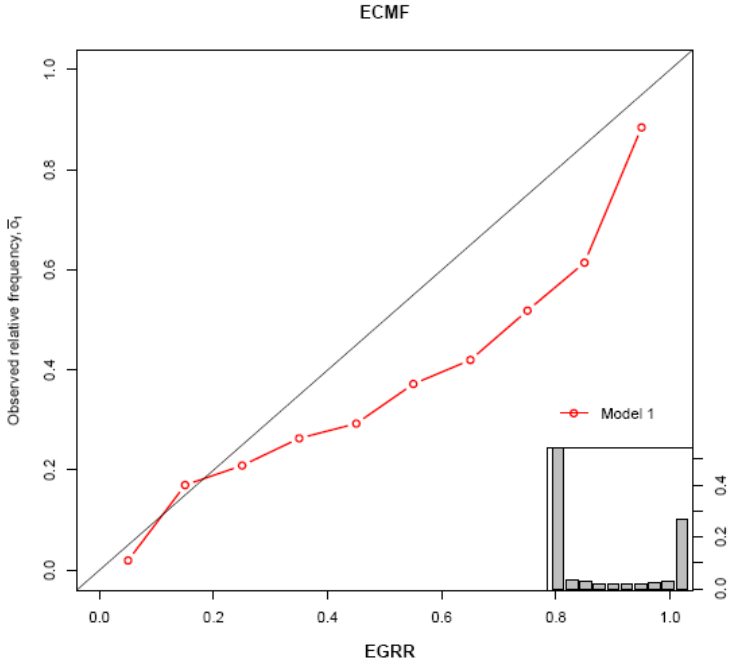
Results – Canada – ROC Curves – 144h

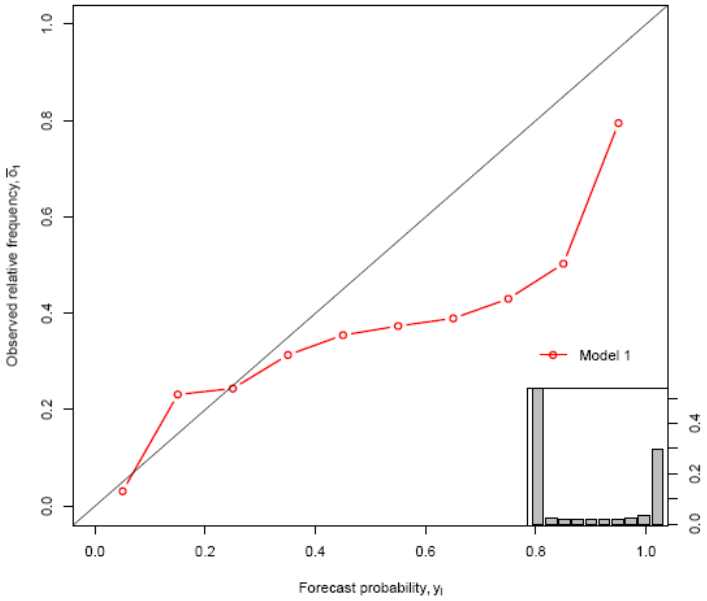


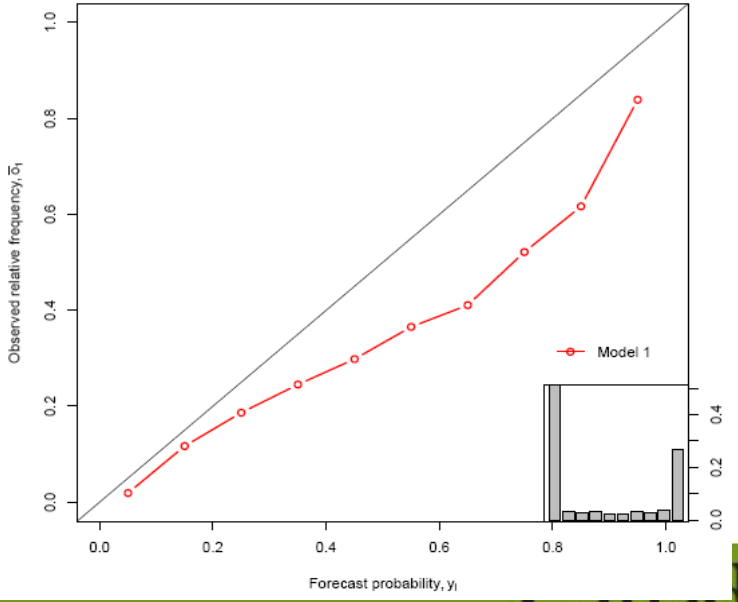


Results – Europe – Reliability – 42h

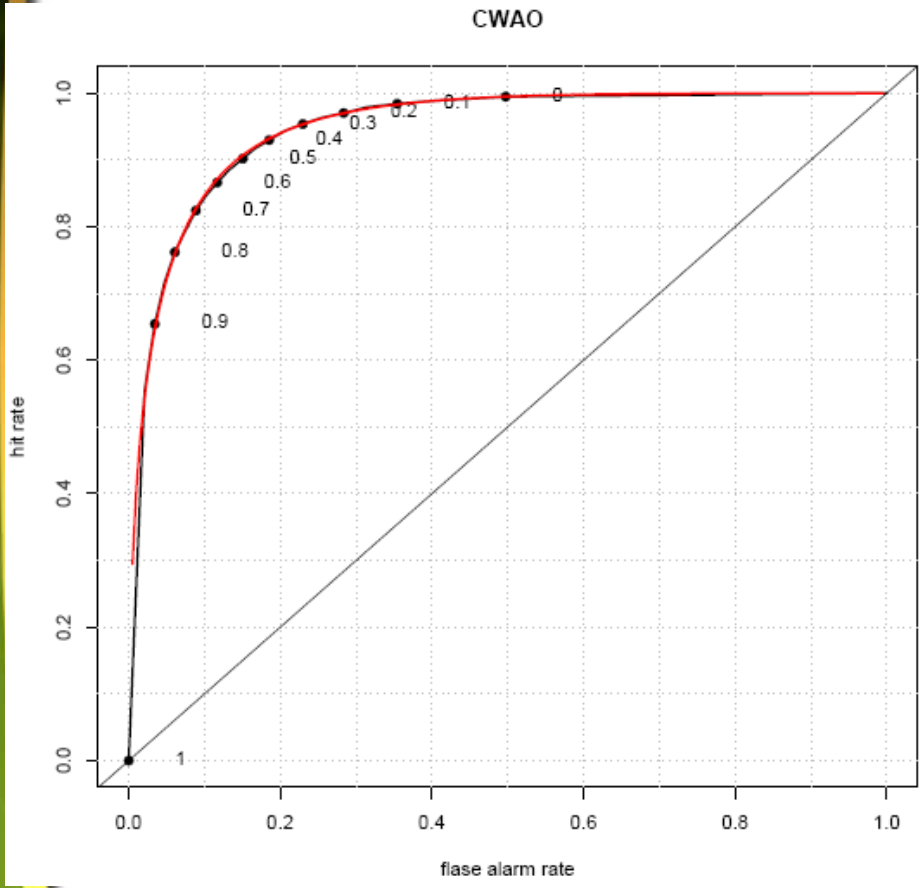


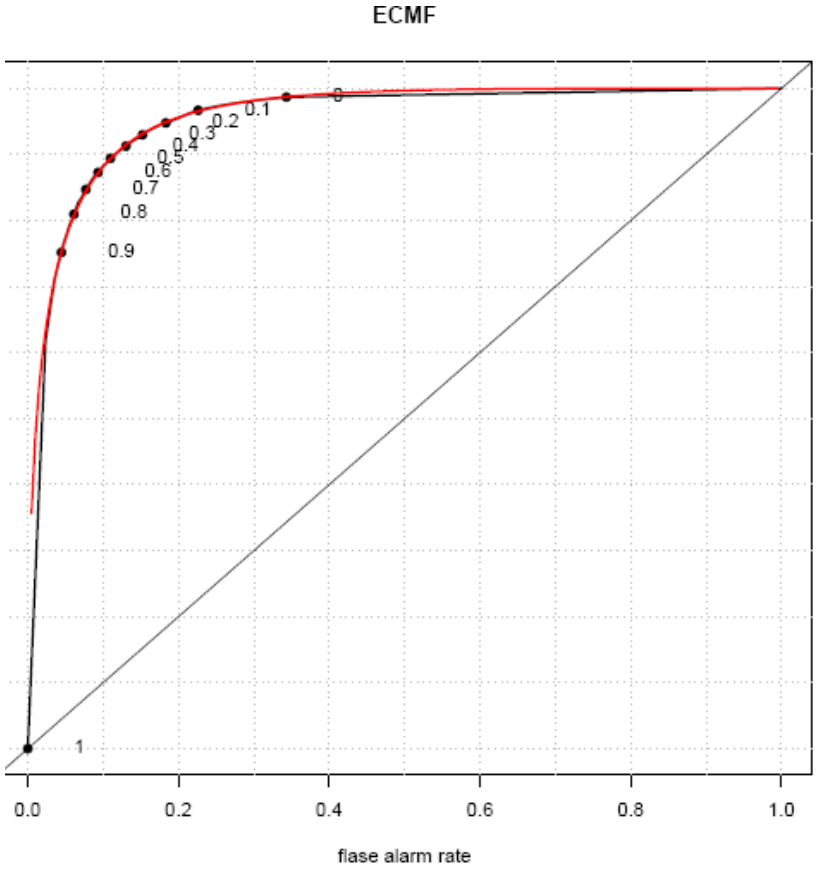




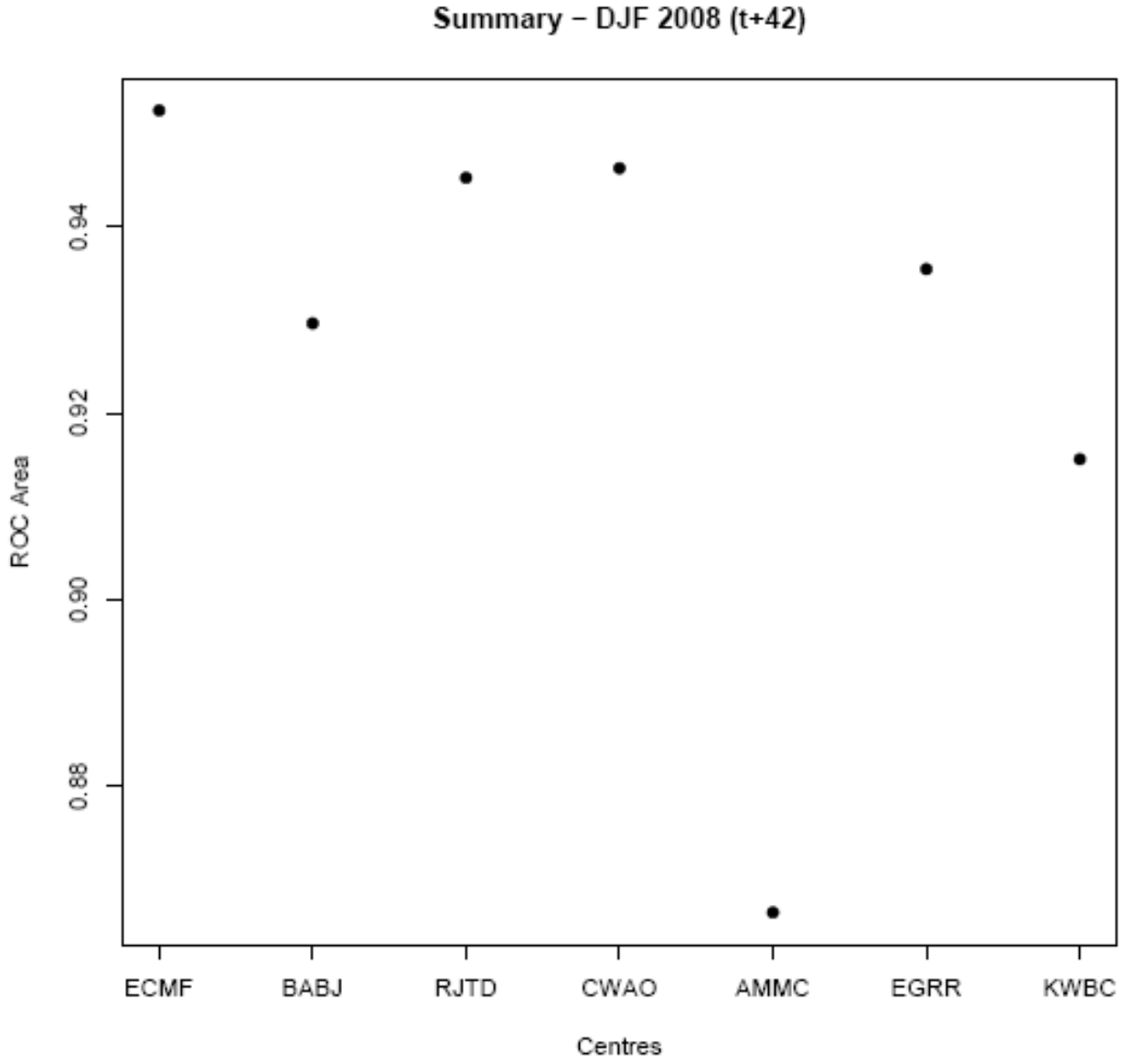


Results – Europe – ROC Curves – 42h





Results – Europe – ROC area – 42h



Summary

- Early results from TIGGE precipitation verification project
  - Not as positive as we would like, more checking needed.
  - But this is surface verification of lowish resolution forecasts.
- To do:
  - More models, more areas of the world – Africa next? Where there is good data…
  - Distribution fitting/assessment of extremes