Verification of hydrological ensemble forecasts

Emmanuel Roulin

Royal Meteorological Institute of Belgium

Joint COST Action 731 and NetFAM
Workshop on

Uncertainty in High-Resolution Meteorological and Hydrological Models

Vilnius, Lithuania, 26-28 April 2006

Overview

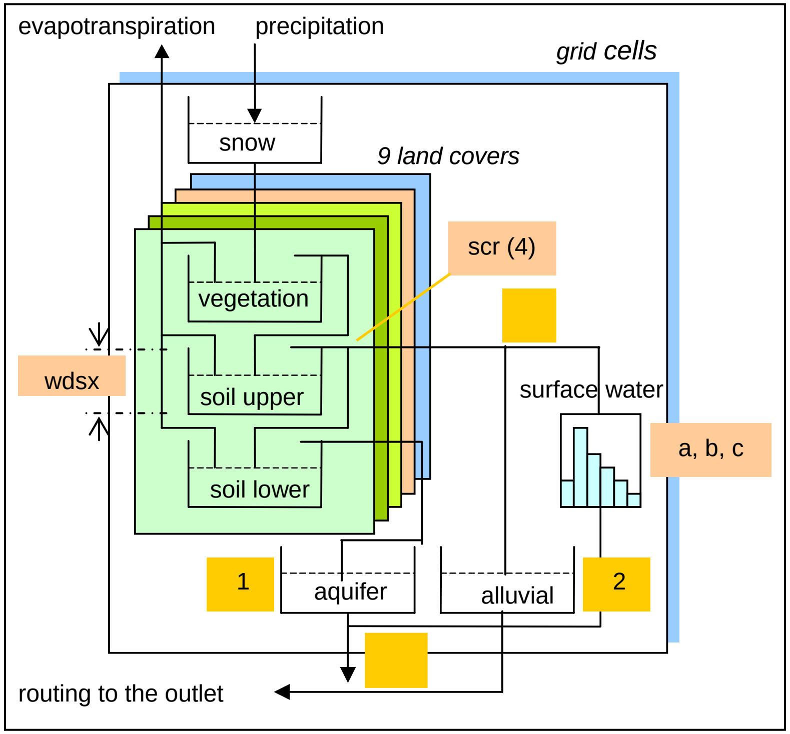
- Hydrological ensemble forecasts
- Verification methods
- Operational procedure
- Conclusions

Hydrological ensemble forecasts

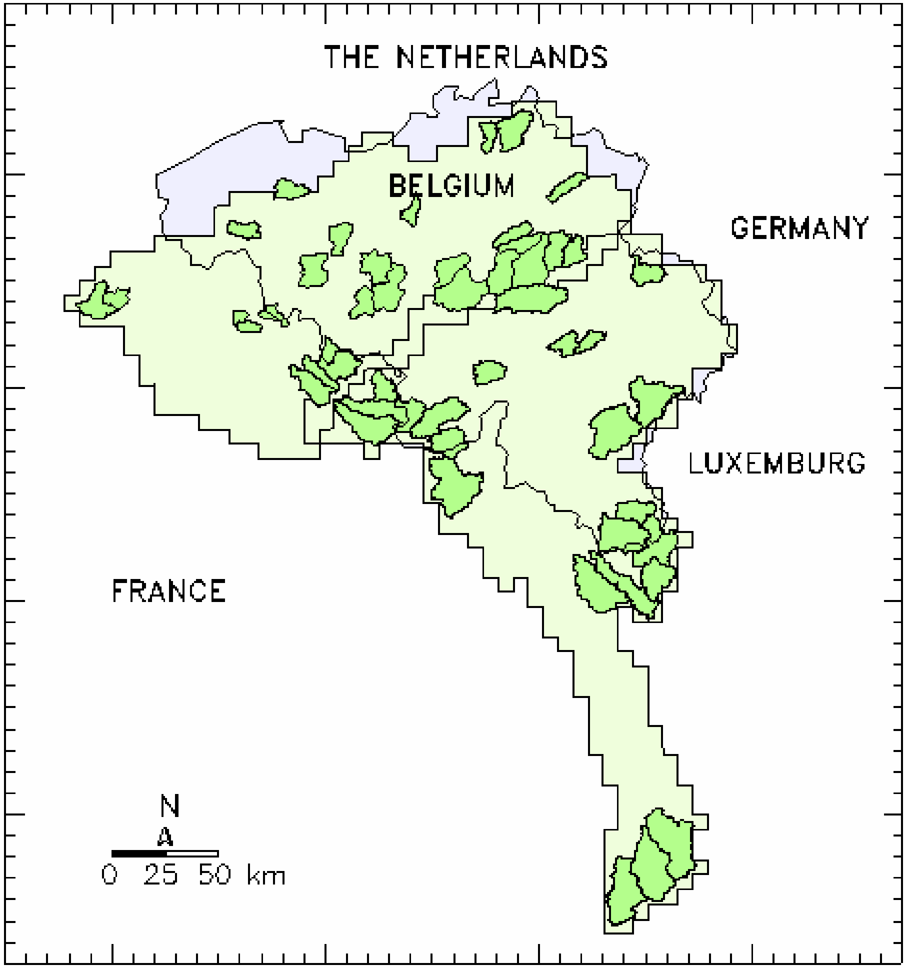
- Hydrological model « SCHEME » (SCHElde & MEuse River Basins)
- Ensemble Prediction System ECMWF
- Probability forecasts
- Two test catchments

- Hydrological model « SCHEME »
  - Based on “IRMB” conceptual model (Bultot and Dupriez, 1976) applied on a grid with 7 7 km^2
  - Routing module based on “width function”
  - Optimization of model parameters (SCE-UA)
  - Regionalization: ANN models between parameter values and physiographic indices
  - Daily time step

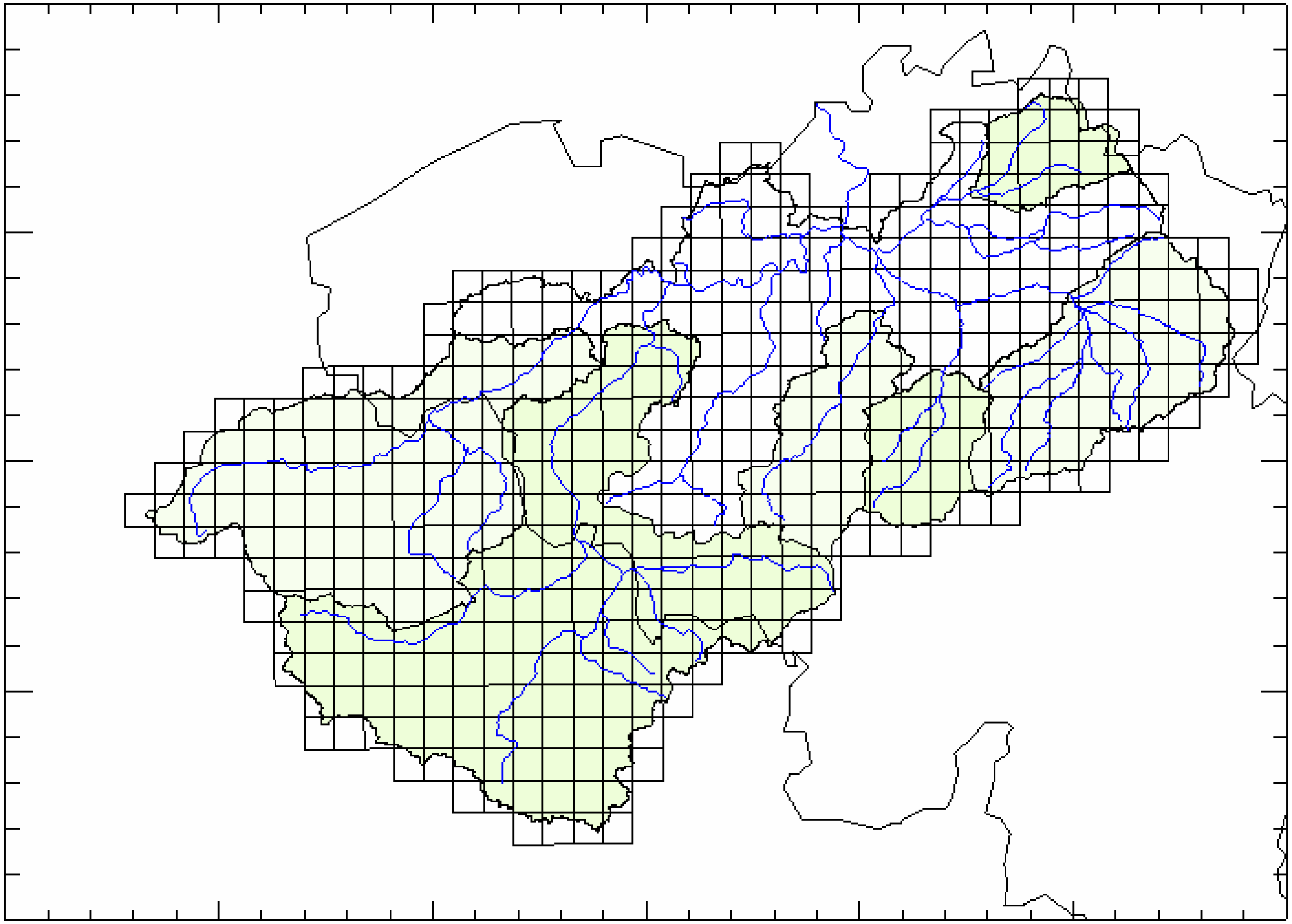
– “IRMB” conceptual model applied on 7 7 km²



– Model parameters optimized & regionalized

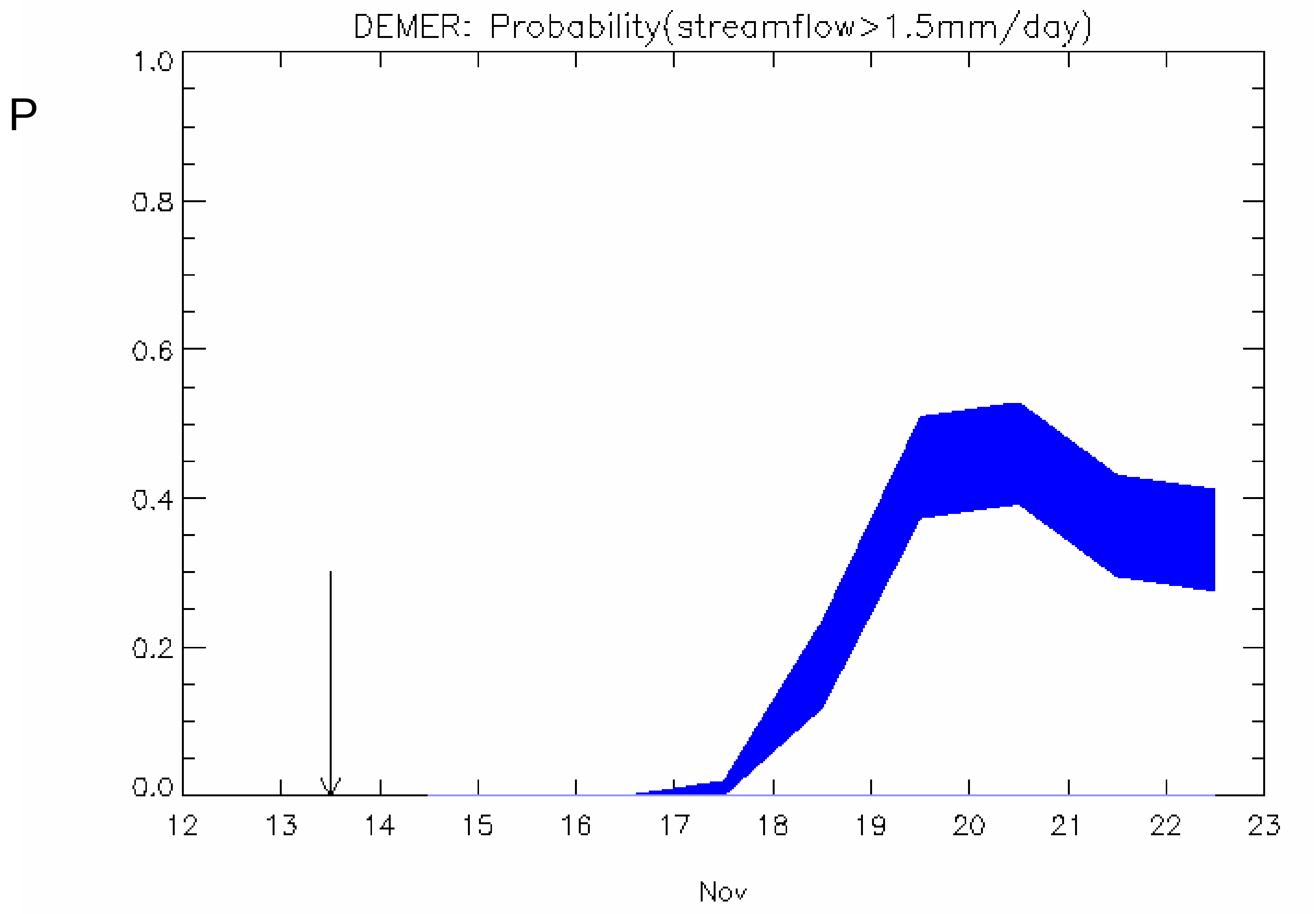


– Routing parameters on catchments > 500 km^2

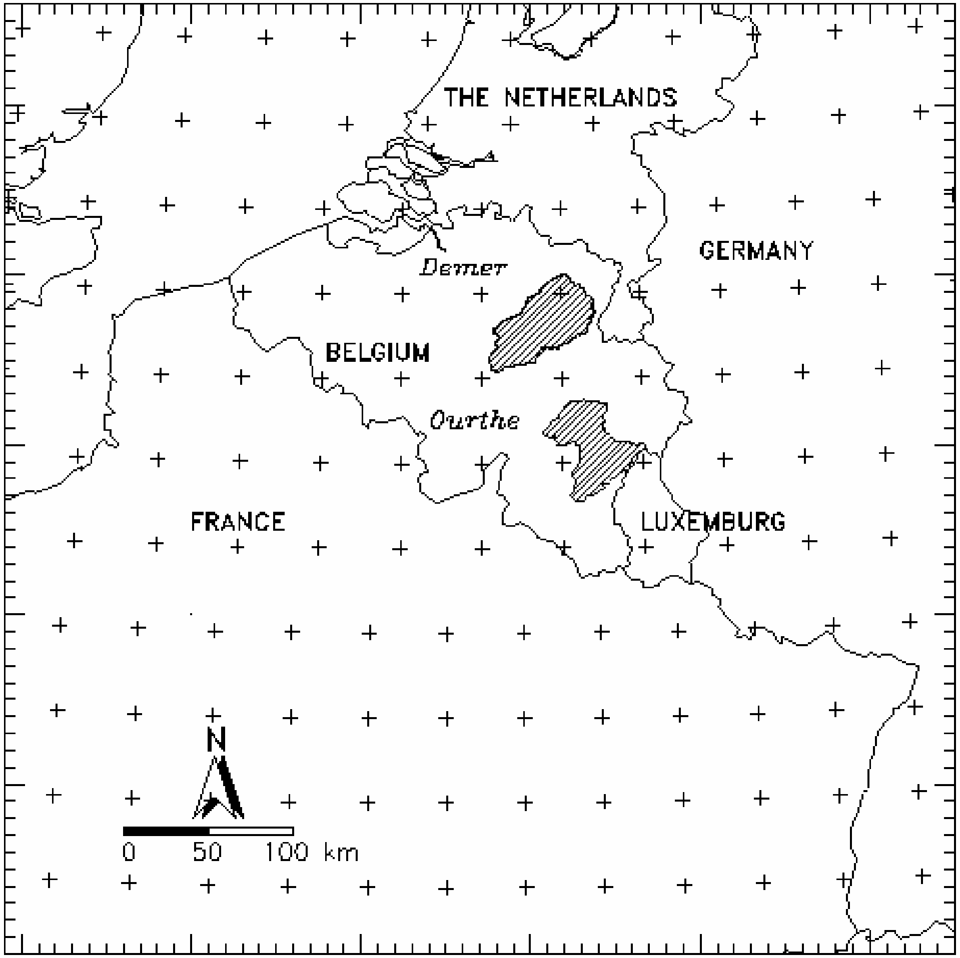


- Ensemble Prediction System ECMWF
  - Use of archived forecasts since December 1996
  - Horizontal resolution: 120 km, 80 km (Nov. 200), 50 km (Feb. 2006)
  - Direct use of precipitation

- Probability forecast
  - EPS + SCHEME = 51 hydrograms   probability to exceed a threshold (e.g. P95)



- Two test catchments
  - Démer at Diest (1775 km^2), Ourthe at Tabreux (1616 km^2)



Verification methods

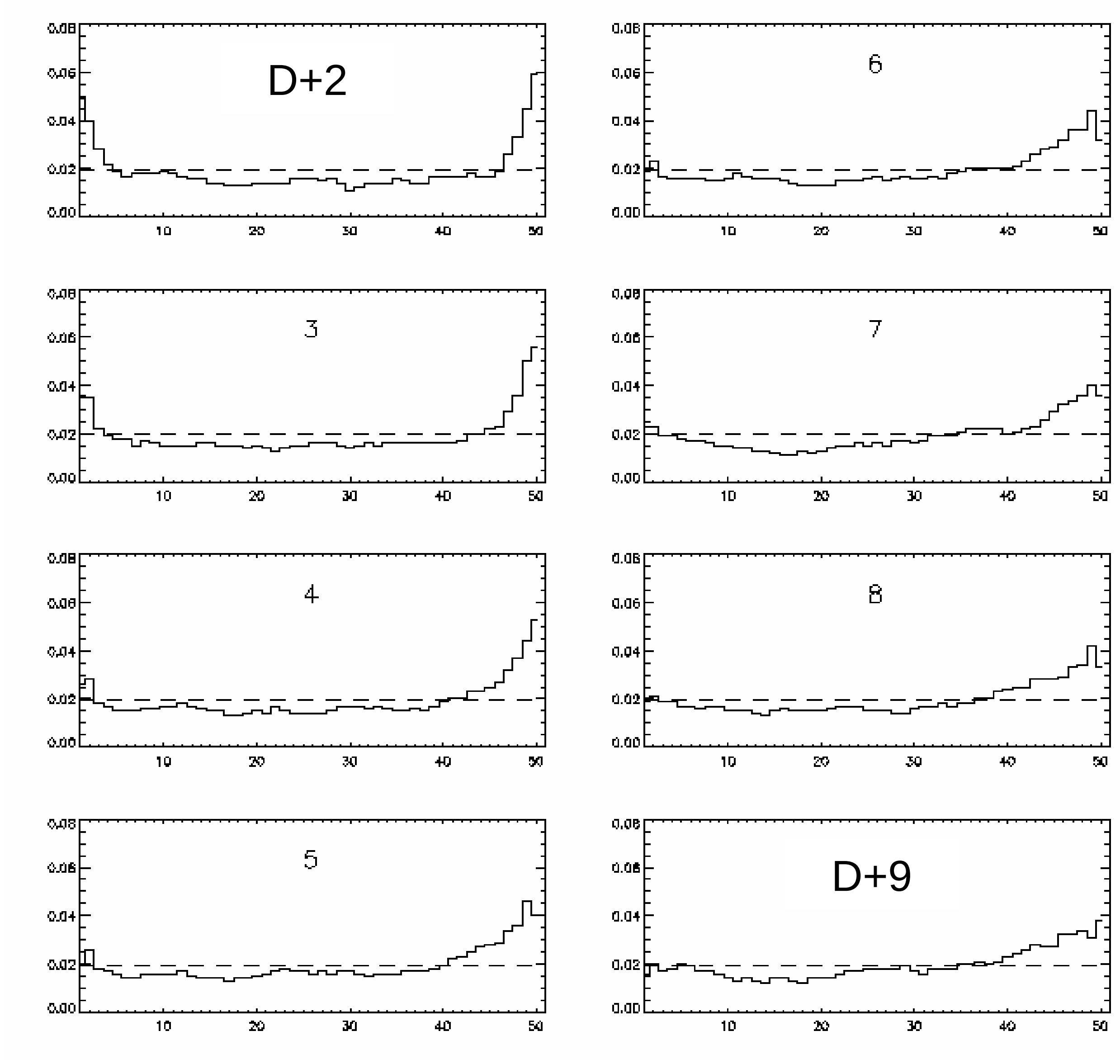
- Rank histogram
- Distribution Oriented approach
- Relative economic value

- Rank histogram
  - Rank histogram (e.g. Talagrand *et al*., 1997)
    - event x and each forecast f_i equally plausible*
  - Correction (Hamill and Colucci, 1997)
    - $P (x < f_i) = \quad R_j$

      j=1 i
    - below the lowest ensemble member : uniform
    - beyond the highest ensemble member : Gumbel distribution
  - Example: Ourthe (Nov 2000 – Mar 2004)

* Here, the reference streamflow is simulated using observed precipitation.

– Example: Ourthe, threshold 6 mm, winter



- Distribution Oriented approach
  - Joint distribution of forecast and observations
  - Brier Skill Score
  - Calibration-refinement factorization
  - Likelihood-base rate factorization

- Joint distribution of forecast and observations
  - f : forecast = probability that P or Q > threshold = (# members with P or Q > threshold) / size of ensemble
  - x : observation = 1 if P or Q > threshold, otherwise, x = 0
  - Thresholds based on [1971-2000], in mm day^{-1}

	P80	P90	P95
P Demer	3.6	6.5	9.8
P Ourthe	4.7	8.3	12.2
Q Demer	0.8	1.1	1.5
Q Ourthe	1.8	2.8	3.9

– Brier Skill Score

- Brier Score or mean square error

  $$MSE(f,x) = E(f - x)^2$$

- Brier Skill Score relative to the “sample climate” or $\ ^{2}_{x}$ = p (1 - p)

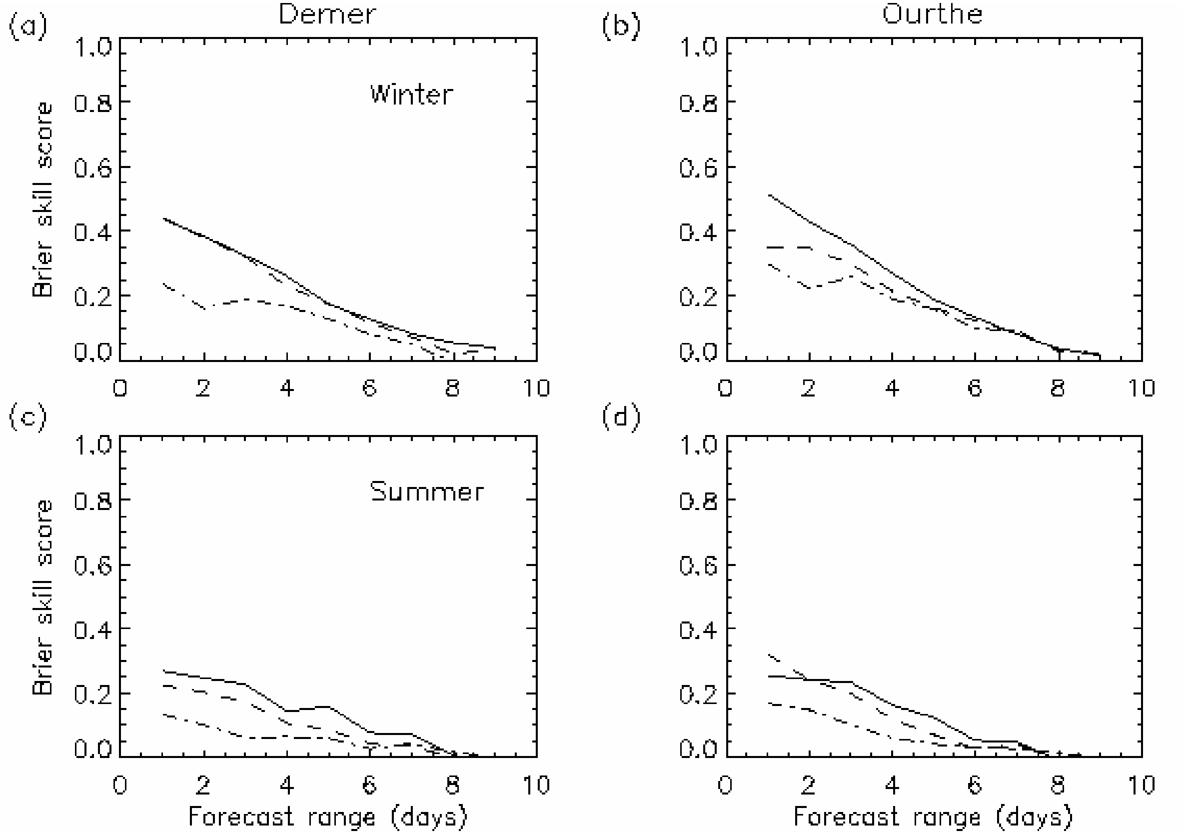
  $$BSS = 1 - MSE(f,x) / \ ^{2}_{x}$$

– BSS: EPS precipitation

P80

P90

P95

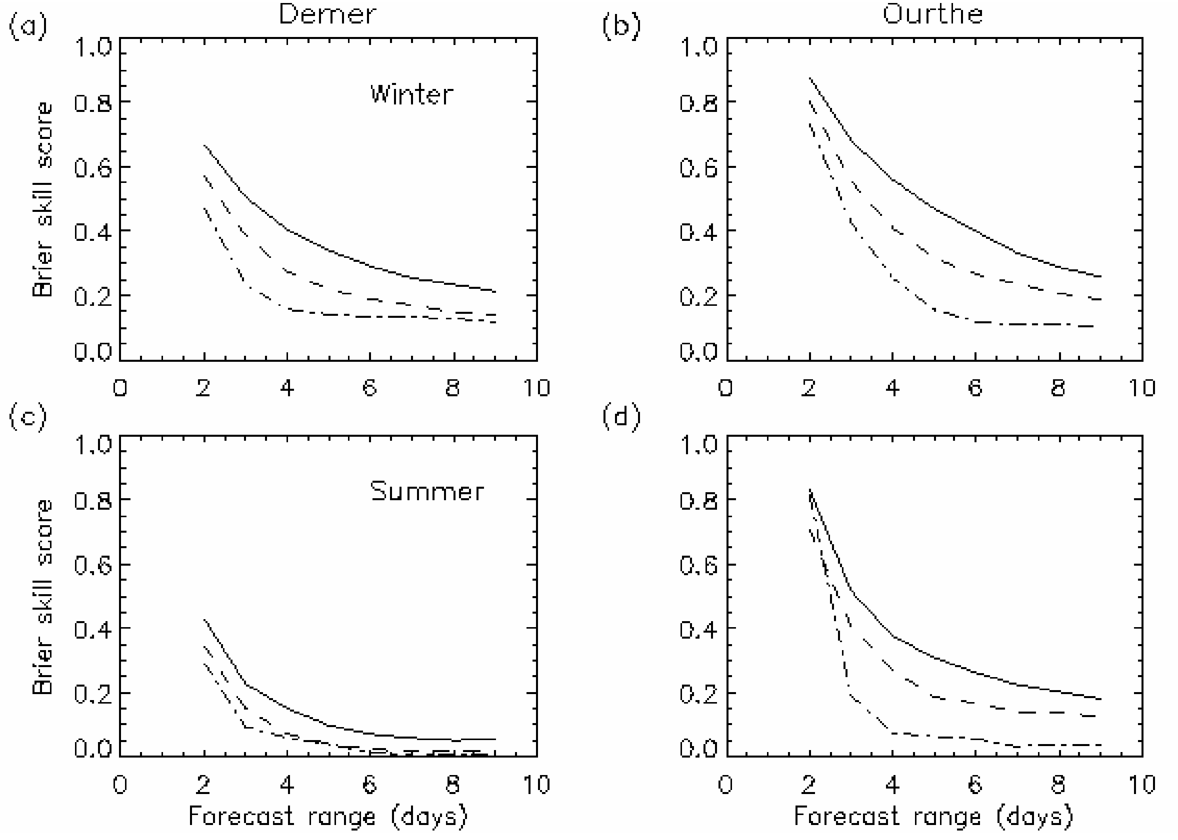


– BSS: Streamflow with precipitation from climatology

P80

P90

P95

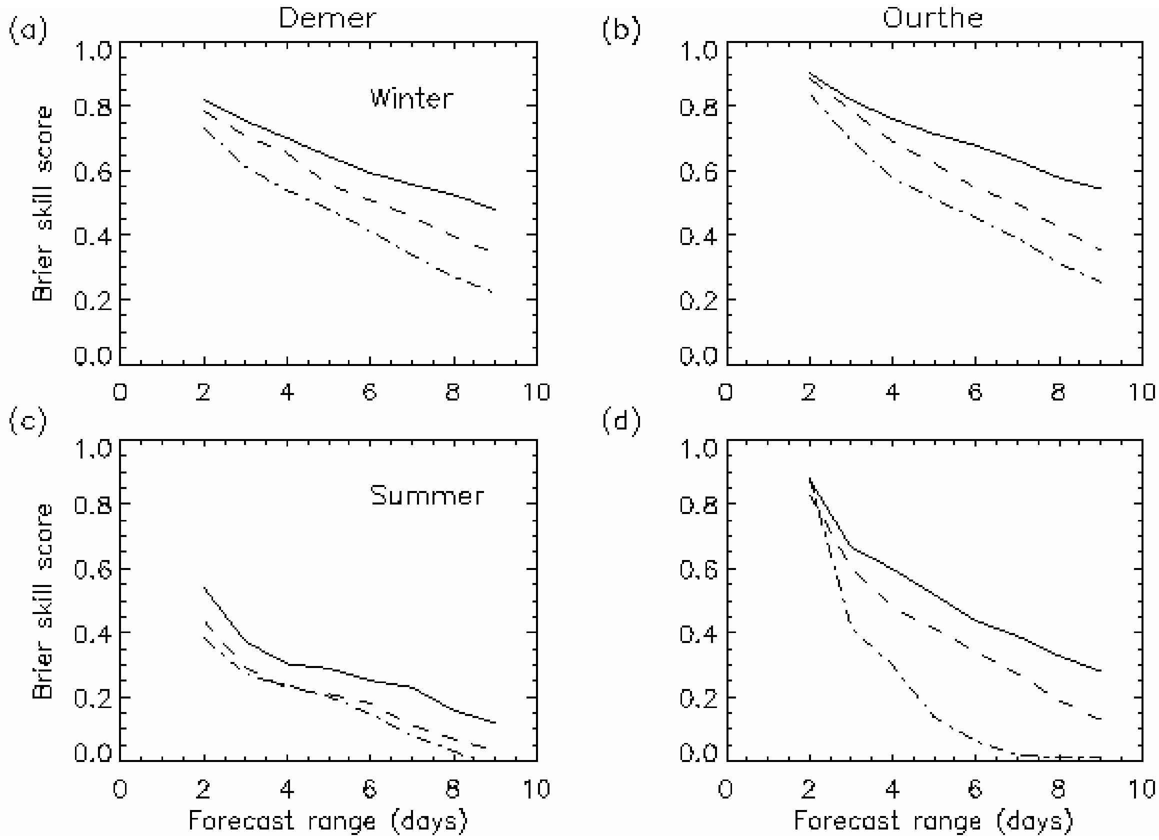


– BSS: Streamflow with EPS precipitation

P80

P90

P95



– Calibration-refinement factorization

- $p(f,x) = p(x|f)\ p(f)$
- Decomposition of the mean square error:

$$MSE(f,x) = E(f - x)^2 = \ {}^2{}_x + E(\ {}_{x|f} - f)^2 - E(\ {}_{x|f} - \ {}_x)^2$$

  - $^2{}_x$: *uncertainty*
  - $E(\ {}_{x|f} - f)^2$ = REL : *reliability (conditional bias)*
  - $E(\ {}_{x|f} - \ {}_x)^2$ = RES : *resolution*

- Skill score relative to the “sample climate” ($^2{}_x$)
  - BSS = 1 - MSE(f,x) / $^2{}_x$
  - BSS = RRES - RREL

- Likelihood-base rate factorization
  - $p(f,x) = p(f|x)\ p(x)$
  - Decomposition of the mean square error:

$$MSE(f,x) = E(f - x)^2 = \ {}^2{}_f + E(\ {}_{f|x} - x)^2 - E(\ {}_{f|x} - \ {}_x)^2$$

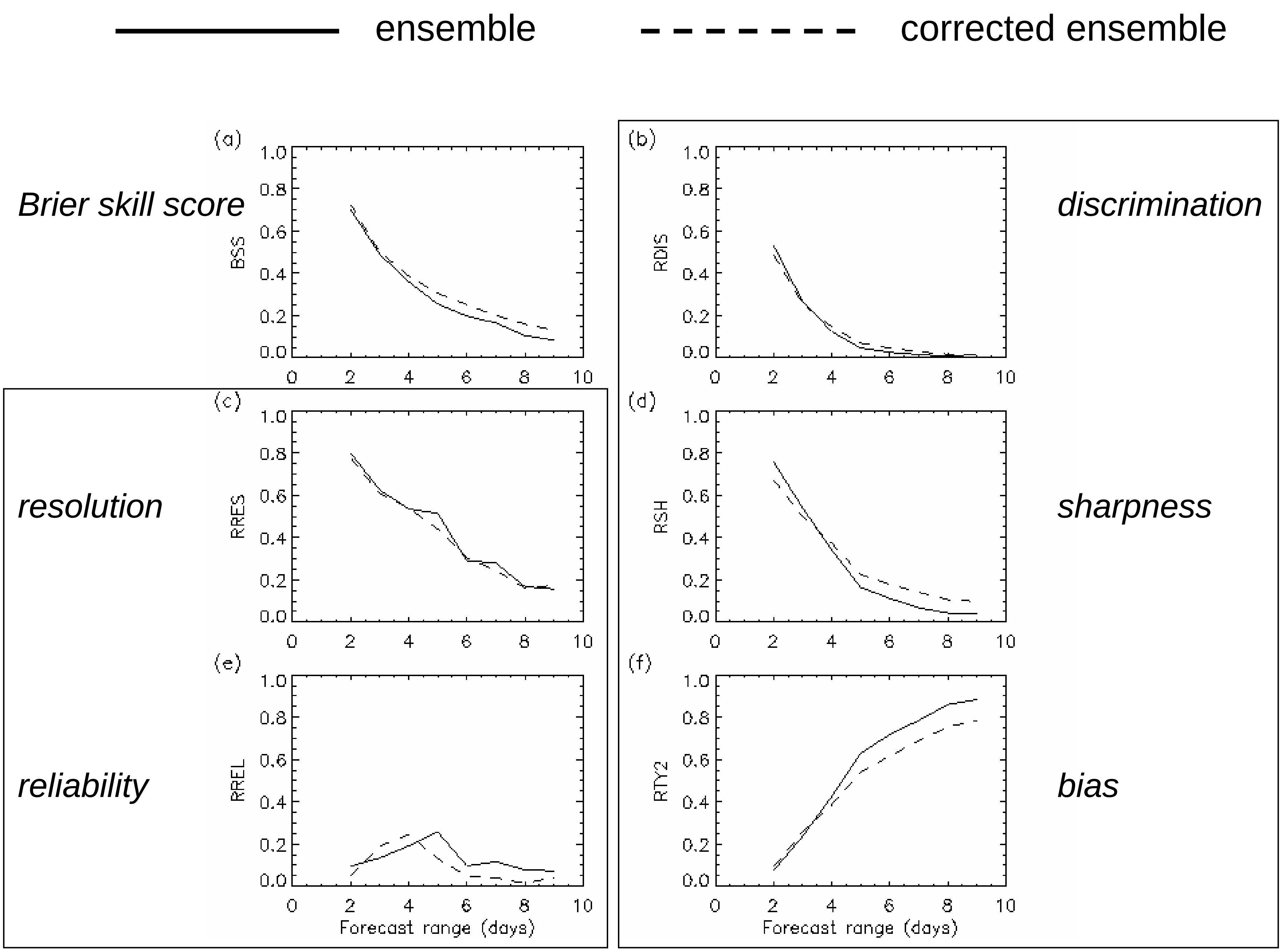
- $\ {}^2{}_f$ = SH : *sharpness*
- $E(\ {}_{f|x} - x)^2$ = TY2 : *type II conditional bias*
- $E(\ {}_{f|x} - \ {}_x)^2$ = DIS : *discrimination*

- Skill score relative to the “sample climate” ($^2{}_x$)
  - $BSS = 1 - MSE(f,x) / \ {}^2{}_x$
  - BSS = 1 + RDIS - RSH - RTY2

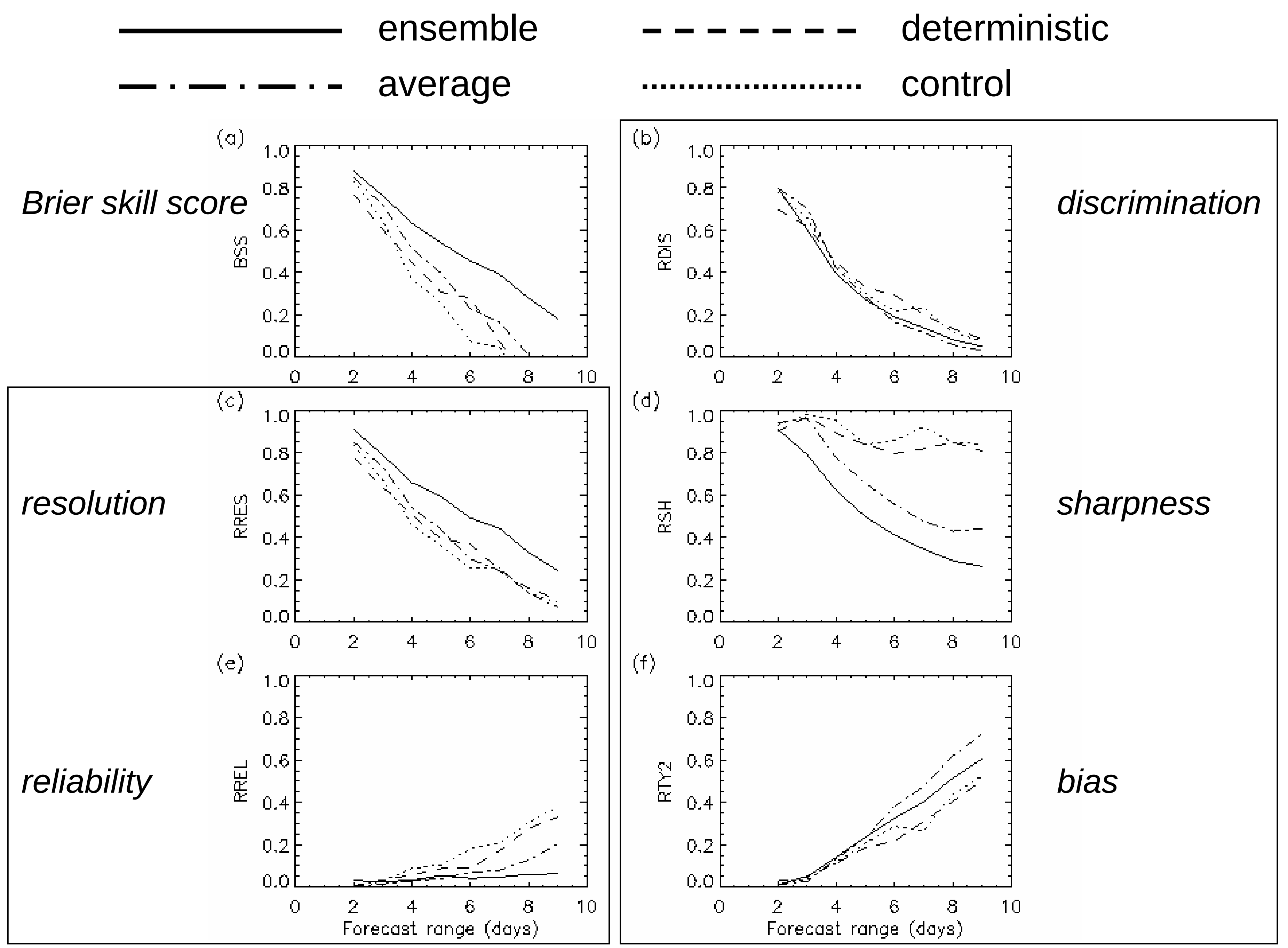
- Example 1: corrected ensemble (see above, Ourthe during winter, from Nov 2000 to Jul 2004, threshold: streamflow above 6 mm)

- Example 2: comparison of ensemble streamflow with streamflow forecast using precipitation from the ECMWF deterministic, the EPS ensemble average and the EPS control (threshold 4 mm)

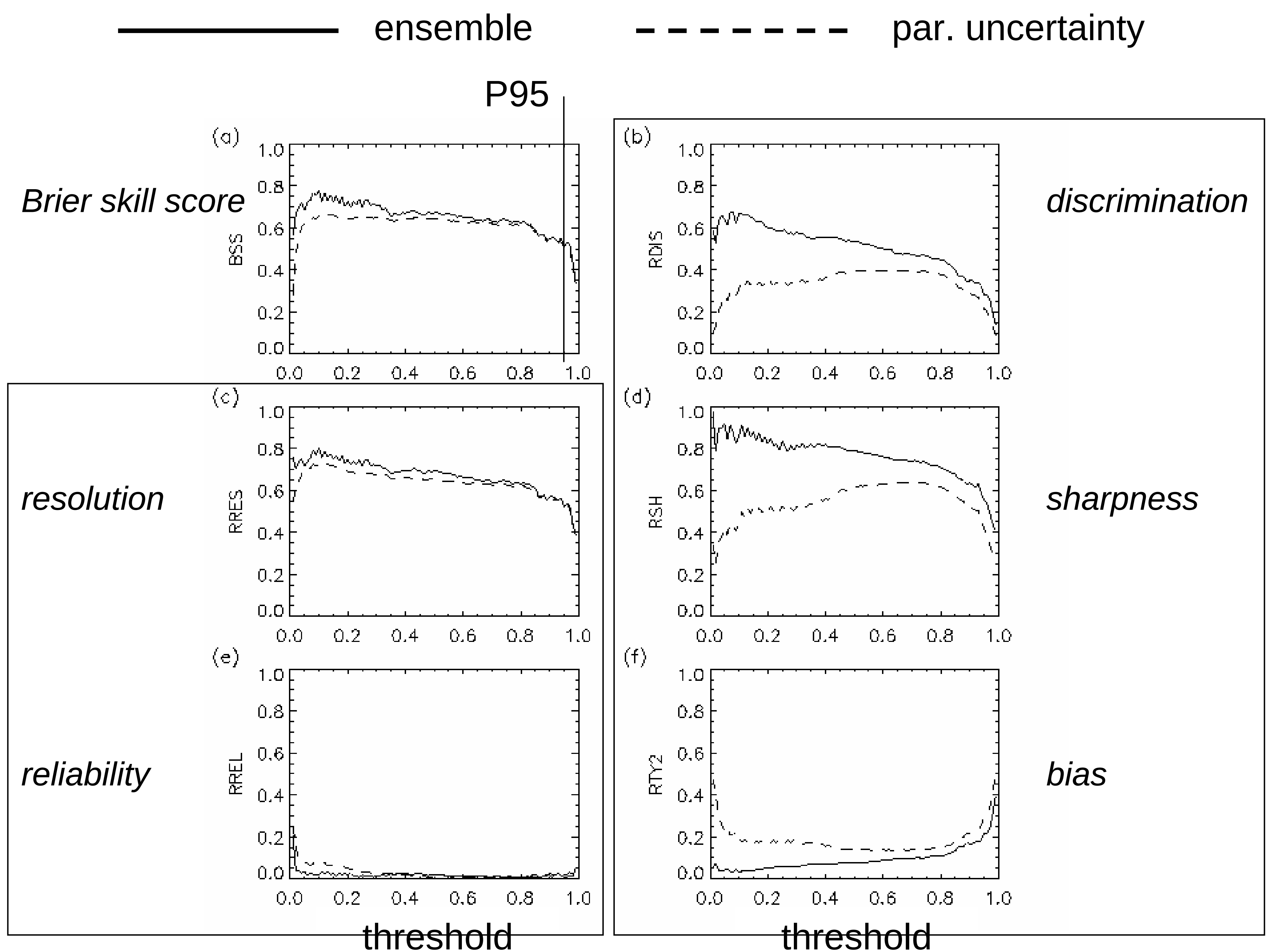
- Example 3: uncertainty associated to the value of hydrological parameters (in progress: uncertainty about only 1 parameter) (Demer)



ensemble
corrected ensemble
(a)
Brier skill score
BSS
(b)
discrimination
RDIS
(c)
resolution
RRES
(d)
sharpness
RSH
(e)
reliability
RREL
(f)
bias
RTY2
Forecast range (days)
Forecast range (days)



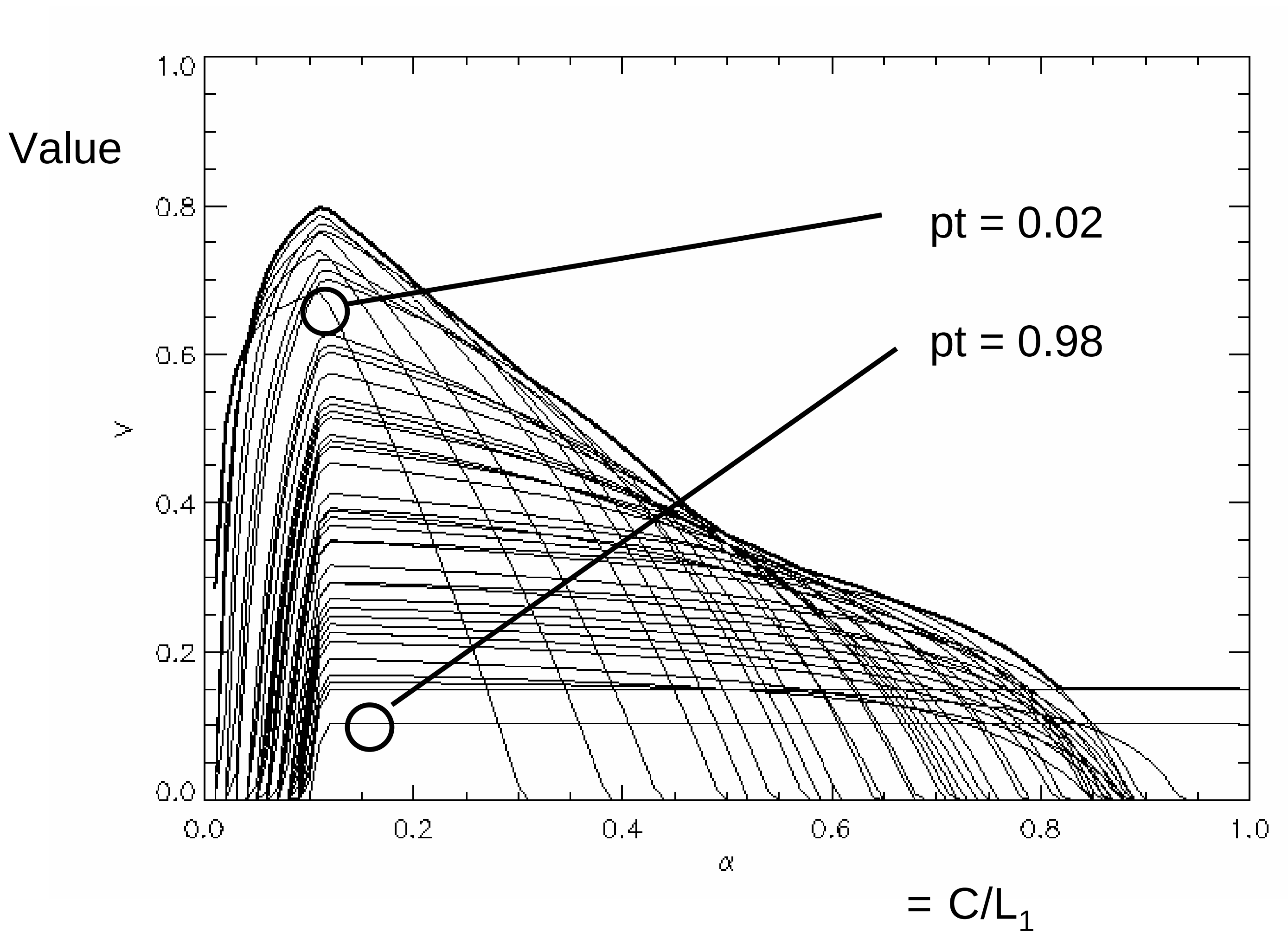
ensemble
deterministic
average
control
(a)
Brier skill score
BSS
(b)
discrimination
RDIS
(c)
resolution
RRES
(d)
sharpness
RSH
(e)
reliability
RREL
(f)
bias
RTY2
Forecast range (days)
Forecast range (days)


ensemble
par. uncertainty
P95
(a)
Brier skill score
BSS
(b)
discrimination
RDIS
(c)
resolution
RRES
(d)
sharpness
RSH
(e)
reliability
RREL
(f)
bias
RTY2
threshold
threshold

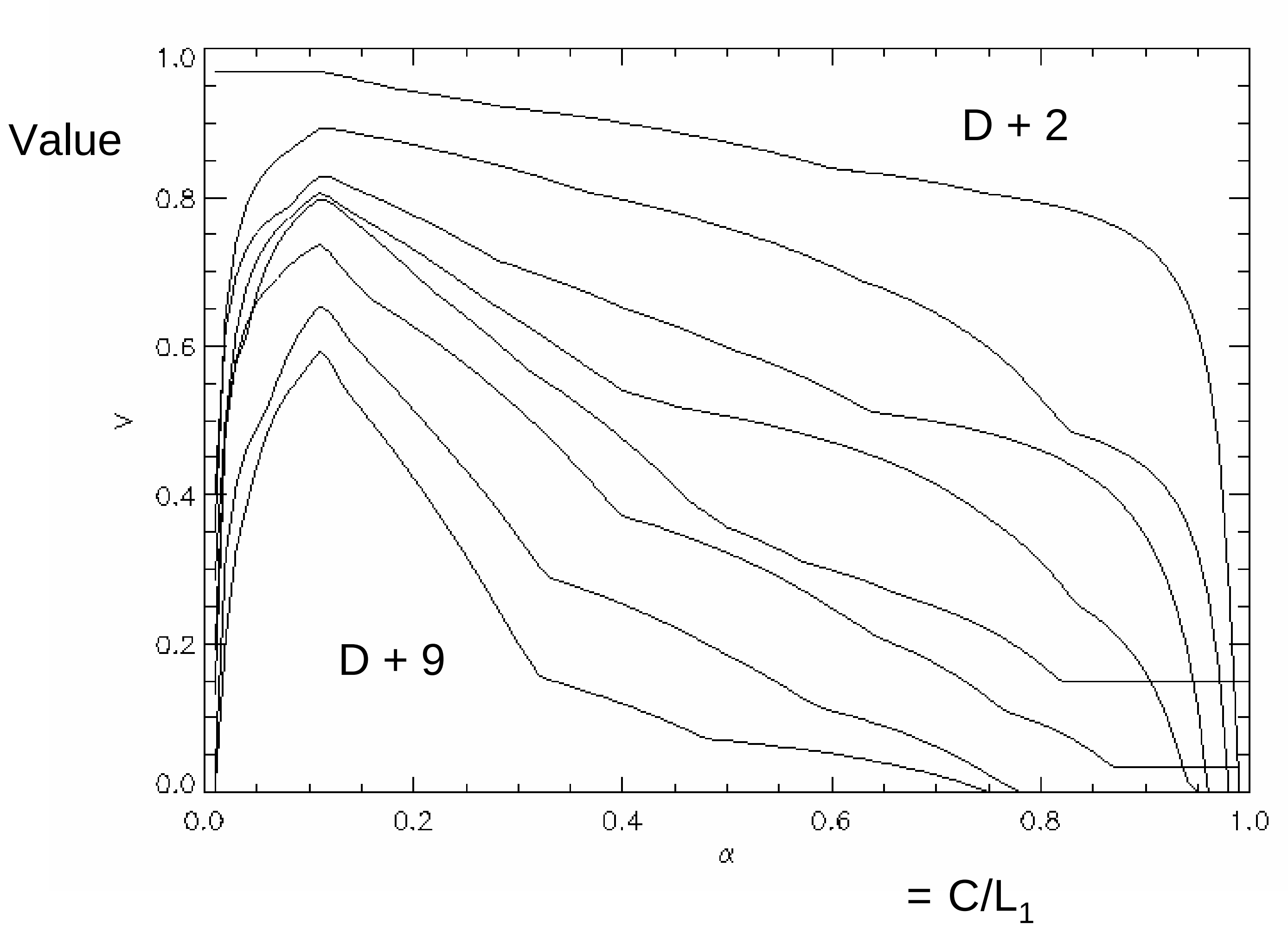
- Relative economic value
  - Static cost-loss model (Richardson, 2000)
    - C: cost of action
    - L: loss if the event occurs and no action is taken
    - L1: part of the loss prevented by taking action
    - o: fraction of occasions when the event occurs
  - Minimize the expense E:
    - Eclim = min { C + o (L – L1), o L }
    - Eperf = o (C + L – L1)
  - Relative value:
    - V = (Eclim – Efor) / (Eclim – Eperf)

- Relative economic value
  - Deterministic forecast system
    - H : hit rate
    - F : false-alarm rate
    - = C / L1, cost-loss ratio
  - Relative value
    - V = {min (, o) – F (1 – o) + H o (1-) – o }
      / (min (, o) – o)
  - Probabilistic forecast
    - pt : threshold probabilty

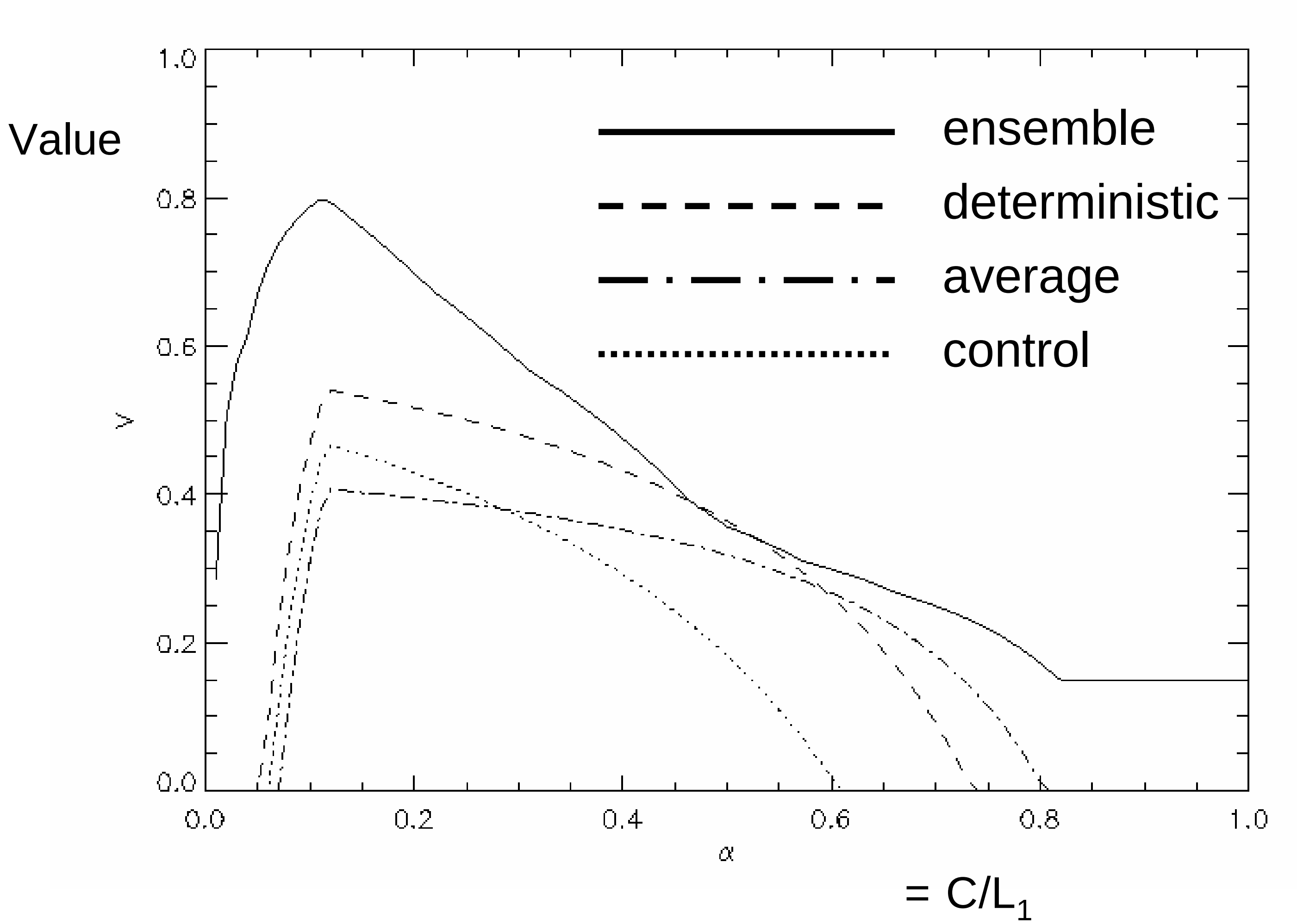
– Hydrological ensemble, 4 mm (P95), D+6



– Hydrological ensemble, 4 mm (P95), D+2 D +9



– 4 mm (P95), D+6

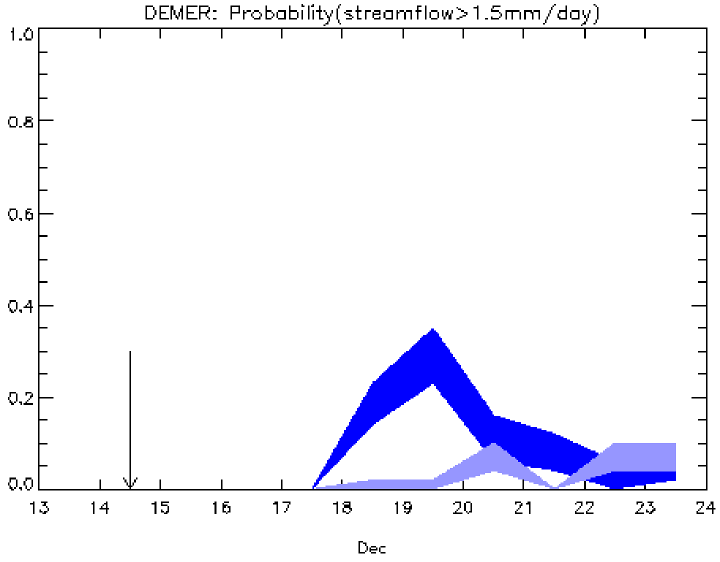

Value
ensemble
deterministic
average
control
1.0
0.8
0.6
0.4
0.2
0.0
V
0.0
0.2
0.4
0.6
0.8
1.0
α
= C/L1

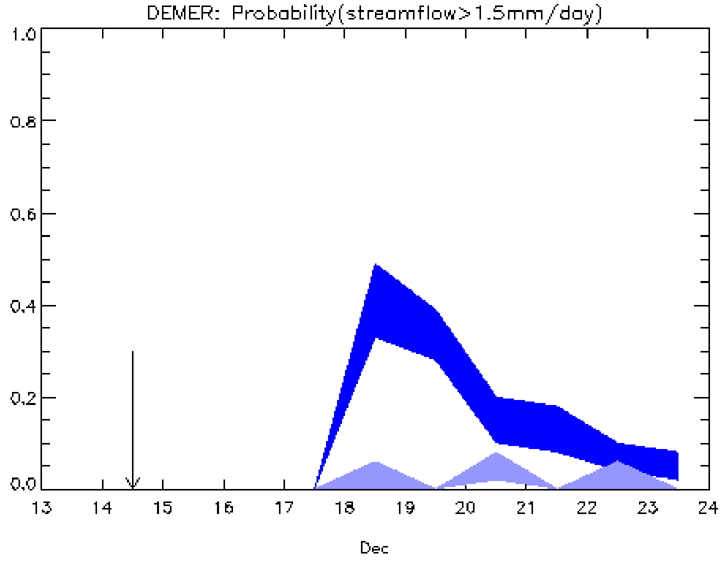
Operational procedure

- HEPDO project
- Validation
- Verification

- HEPDO project
  - “Hydrological Ensemble Prediction for the Demer and the Ourthe”: setting-up an automatic operational procedure
  - EPS precipitation but also temperature, wind speed etc. to estimate potential evapotranspiration and account for snow accumulation and melting
  - Use of precipitation data available operationally: weather radar and automatic weather stations to update the water content of the conceptual reservoirs

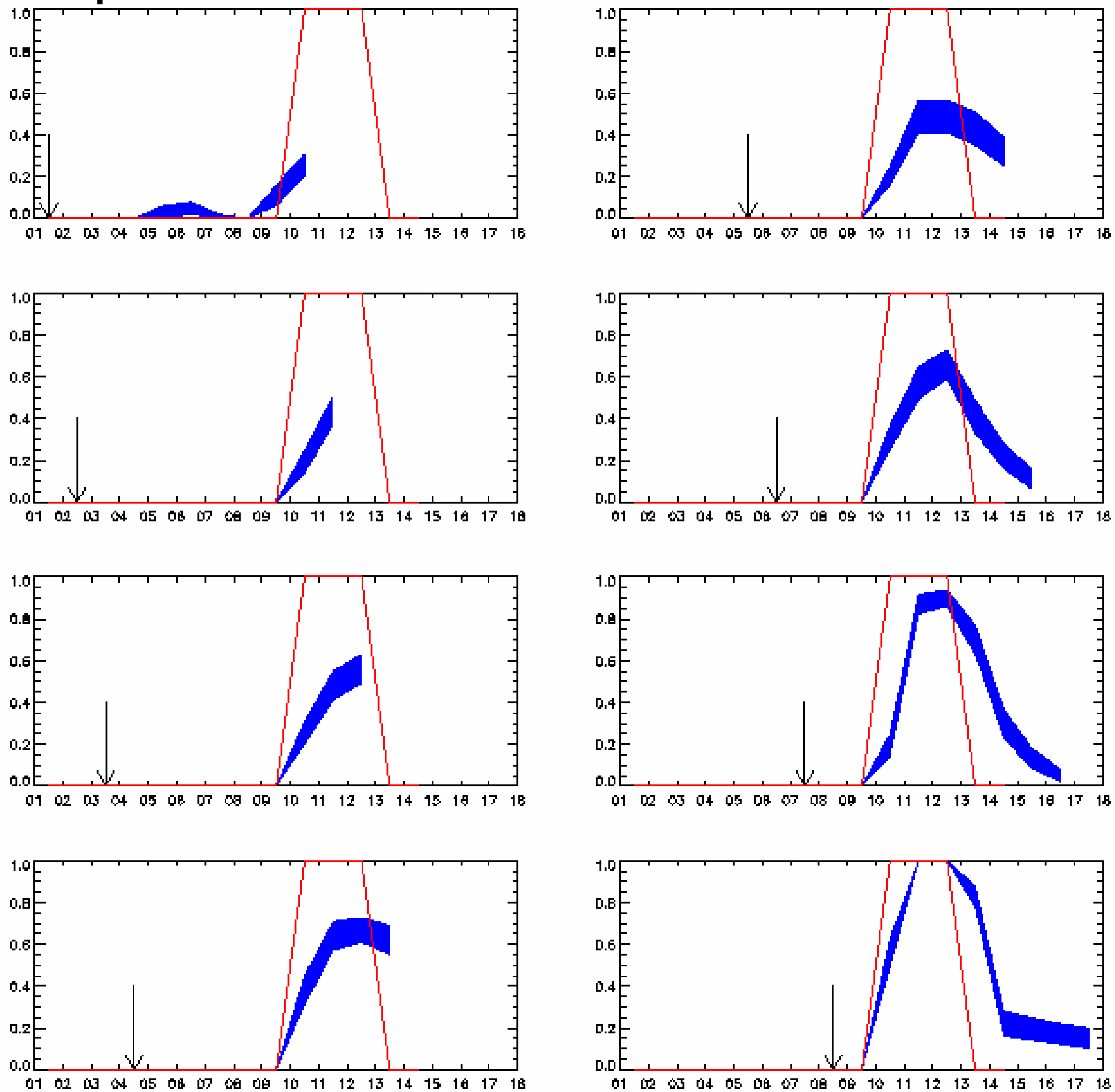
– Demer in December 2005; left: with radar data only; right: with raingauge data during November and then, with radar data; in progress: with combined radar data and available raingauge data





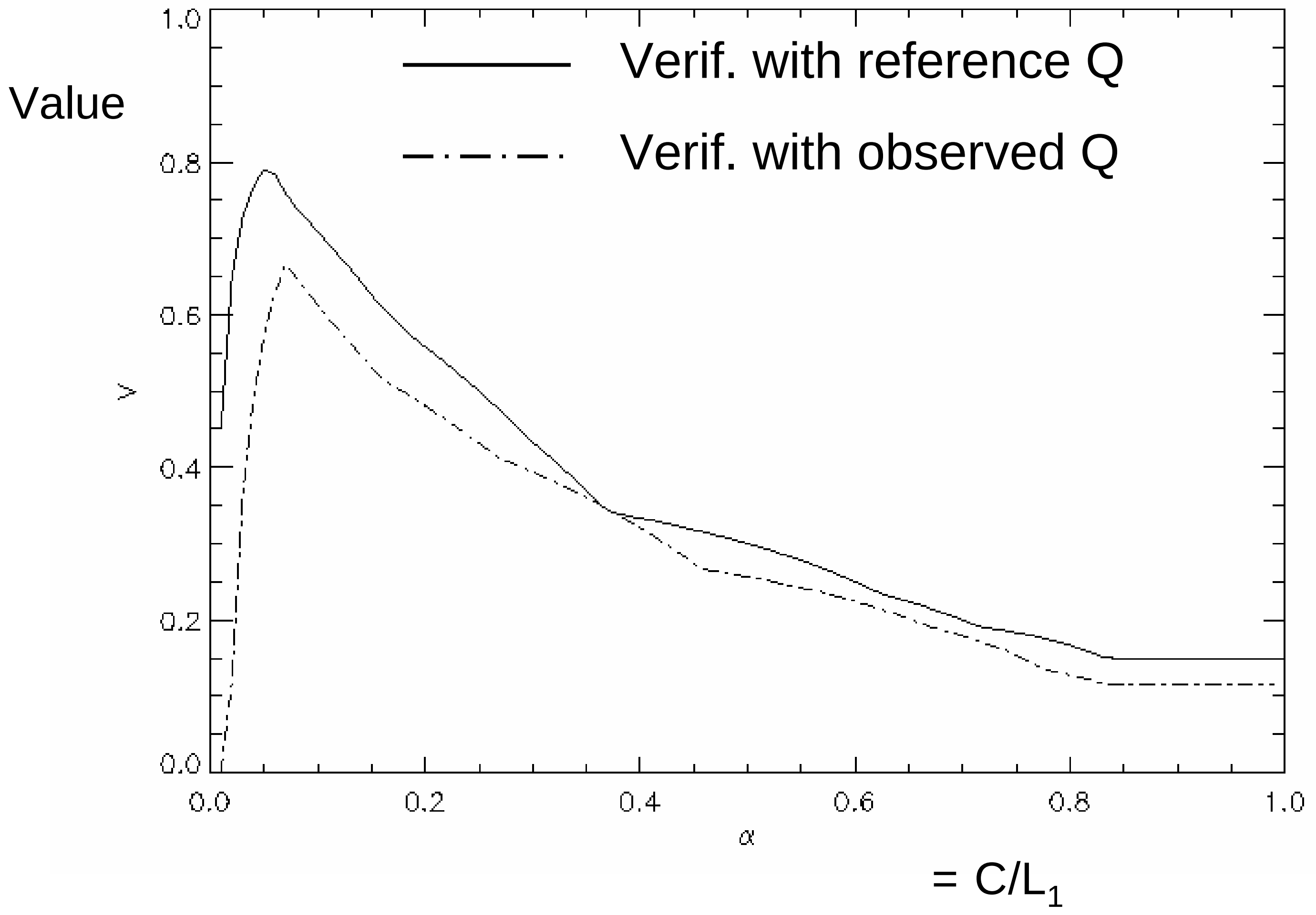
- Proposed for validation to:
  - Weather forecasters : maps with probability of precipitation
  - Regional authorities in charge of water management : web page with ensemble precipitation and probability of streamflow > P95

– Example: Ourthe, March 2006

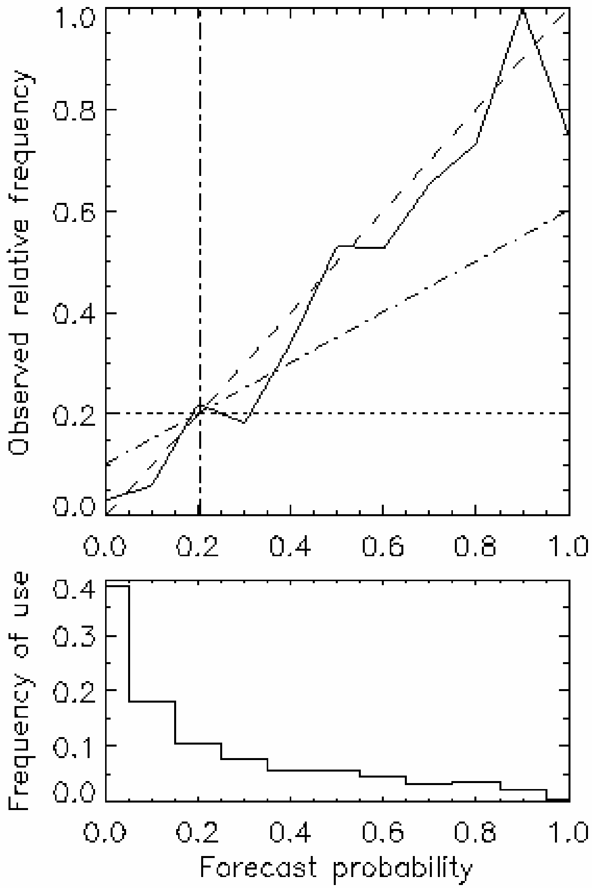


- Verification
  - Comparison with measured streamflow
  - New EPS resolution: small sample size

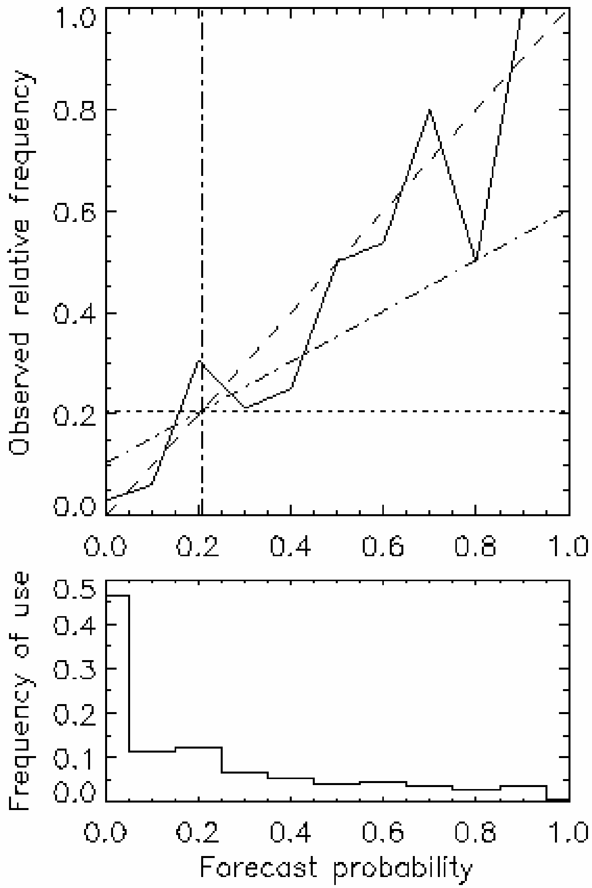
– Threshold 5mm, D+6, Nov. 2000 – Dec 2002



– Reliability diagram for precipitation (Demer), P80



Winter 2001-2005



Feb-March 2001-2005

Conclusions

- Hydrological ensemble predictions have skill and value for early warning
- Need to define the probability threshold for a specific management situation
- Improve verification methods for rare events and small data sample
- Investigation on effect of uncertainties on the aspects of forecats quality