Discharge ensemble forecasts based on the COSMO-LEPS quantitative precipitation forecasts

Tommaso Diomede[1,2], Chiara Marsigli[1], Andrea Montani[1], Tiziana Paccagnella[1]

(1) Regional Hydro-Meteorological Service ARPA-SIM, Bologna, Italy

(2) Centro Interuniversitario di Ricerca in Monitoraggio Ambientale (CIMA), Università degli Studi di Genova e della Basilicata, Savona

Purpose

It is investigated the usefulness and the skill of the meteorological mesoscale ensemble prediction system COSMO-LEPS to supply operational quantitative precipitation forecasts driving a meteo-hydrological coupled system aimed at providing reliable real-time discharge ensemble forecasts

The coupled atmospheric-hydrological modelling system

The meteorological forecasting systems

• **COSMO-LEPS** is a Limited-area Ensemble Prediction System based on the non-hydrostatic limited-area model COSMO, daily running (12 UTC) at ECMWF since November 2002.

The different model runs are nested on some selected members of the ECMWF Ensemble Prediction System (EPS), chosen by means of an ensemble-size reduction technique based on a Cluster Analysis algorithm.

The system has been developed for the late-short to early-medium forecast range (48-120 h).

• The deterministic model COSMO operational at ARPA-SIM (**COSMO-LAMI**) is used as term of comparison to evaluate the added value of the probabilistic system.

The configurations (for the autumn seasons 2003-2005)

Name	Boundary conditions	Initial conditions	Moist convection	Prognostic precipitation	Horizontal resolution	Vertical resolution	Forecast range	Number of members
COSMO-LEPS	EPS forecasts	EPS analyses	Tiedtke or Kain-Fritsch (randomly selected)	yes	10 km	32 layers	132 h	10
COSMO-LAMI	DWD-GME forecasts	LAMI mesoscale assimilation (nudging)	implicit (Tiedtke)	no	7 km	35 layers	72 h	1

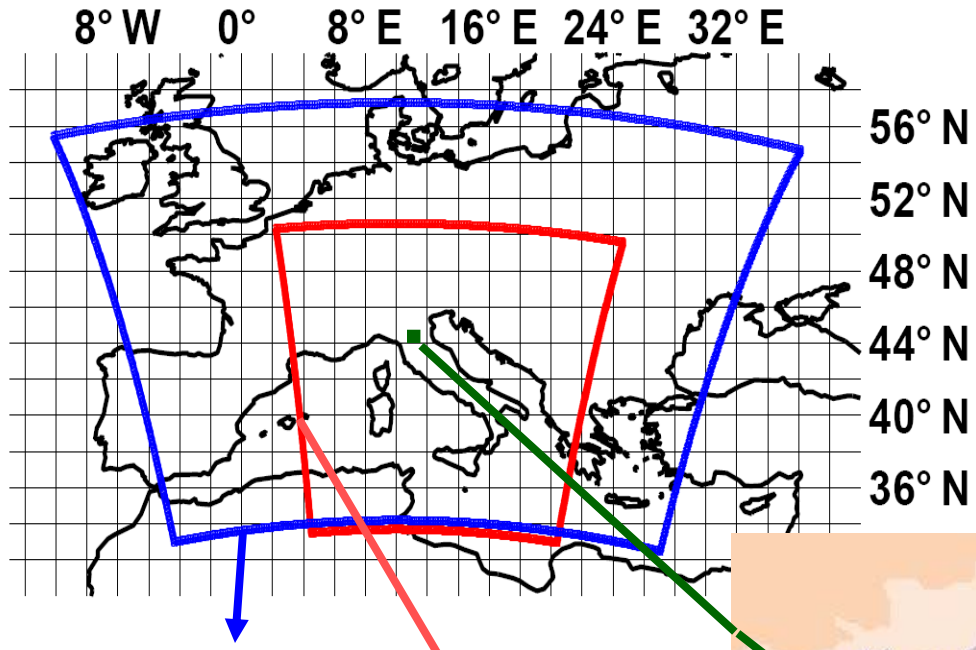
nb: for the COSMO-LEPS system of the year 2003 the forecast range is 120 h, the number of ensemble members is 5, the adopted moist convection scheme is Tiedtke and the prognostic treatment of rain and snow is not added.

The hydrological model

TOPKAPI (TOPographic Kinematic APproximation and Integration)

physically-based distributed rainfall-runoff model

Spatial Domains and Study Area



spatial domain of COSMO-LEPS

spatial domain of COSMO-LAMI

Emilia-Romagna Region

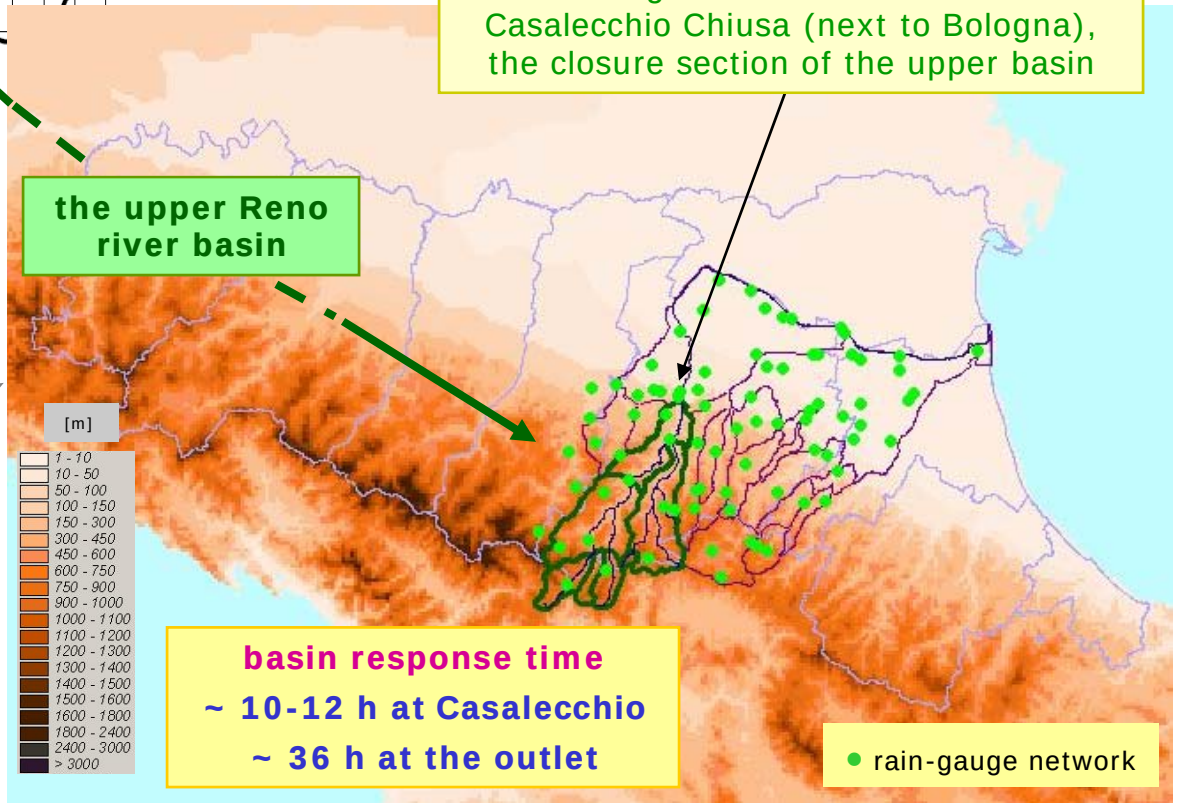
main river total length : 210 km

dimension : entire basin ~ 5000 km^2
upper basin ~ 1000 km^2

Alert threshold:
0.8 m (~ 80 m^3/s) ***warning***
1.6 m (~ 630 m^3/s) ***alarm***

the discharge forecasts are evaluated at Casalecchio Chiusa (next to Bologna), the closure section of the upper basin

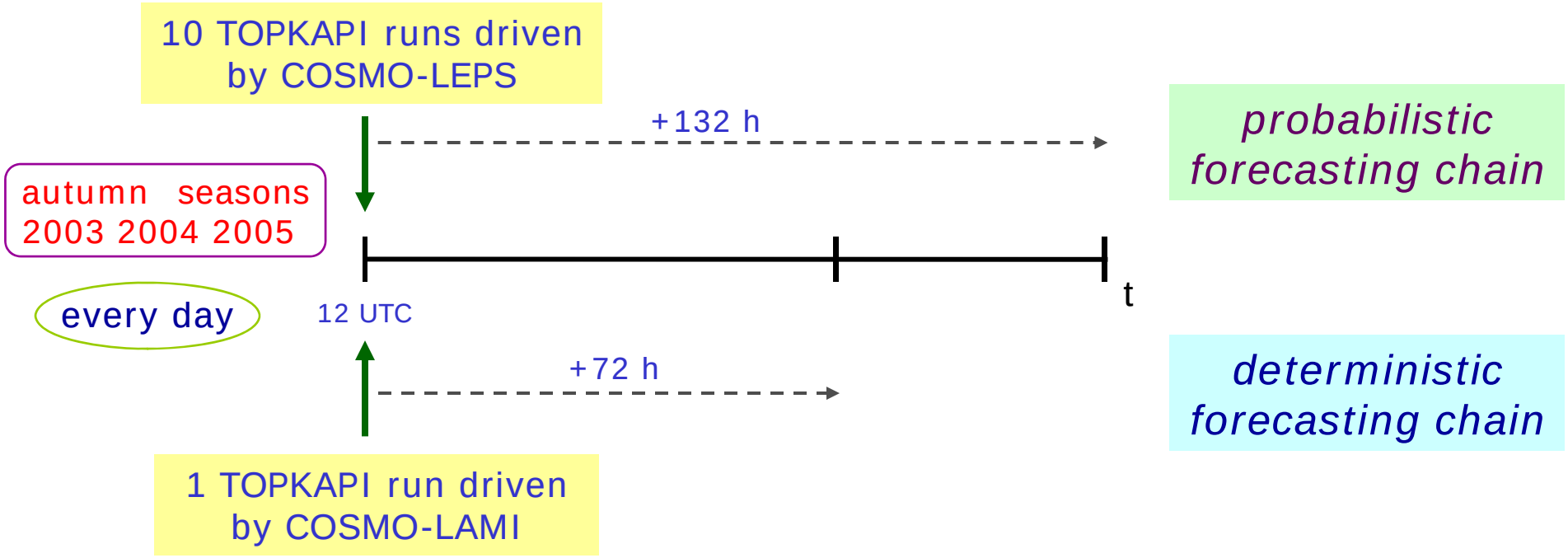
the upper Reno river basin



basin response time
~ 10-12 h at Casalecchio
~ 36 h at the outlet

● rain-gauge network

The coupling strategy

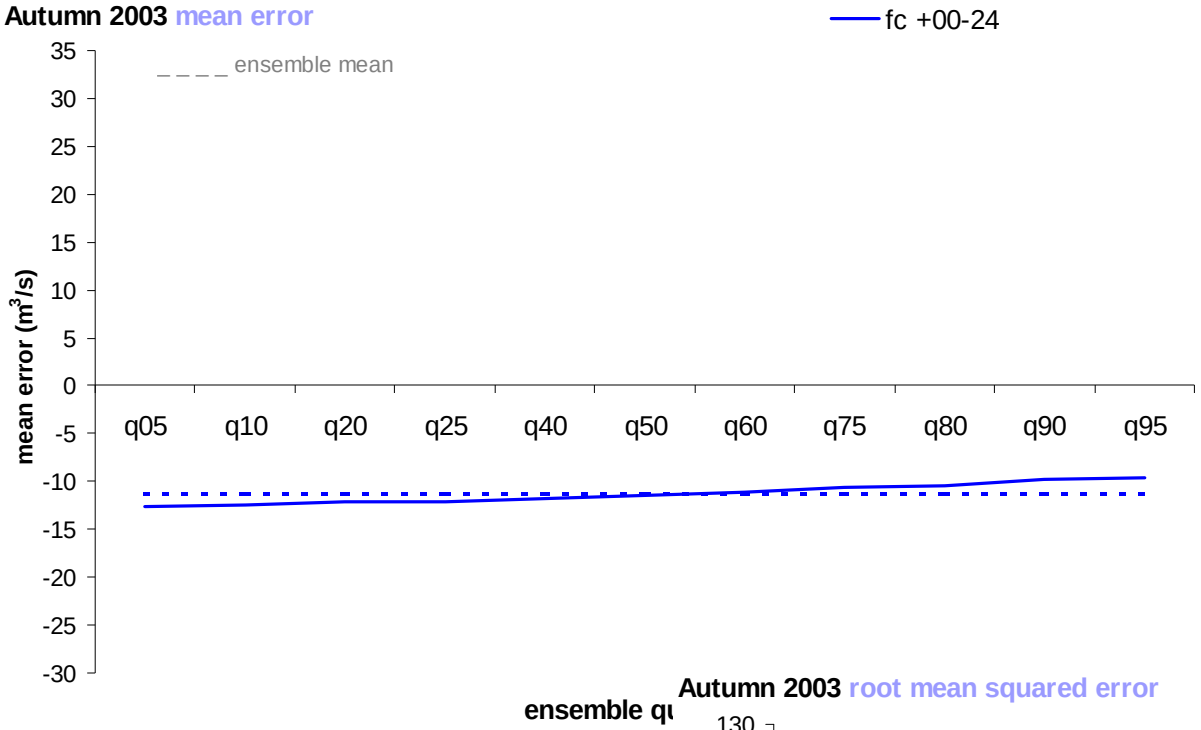


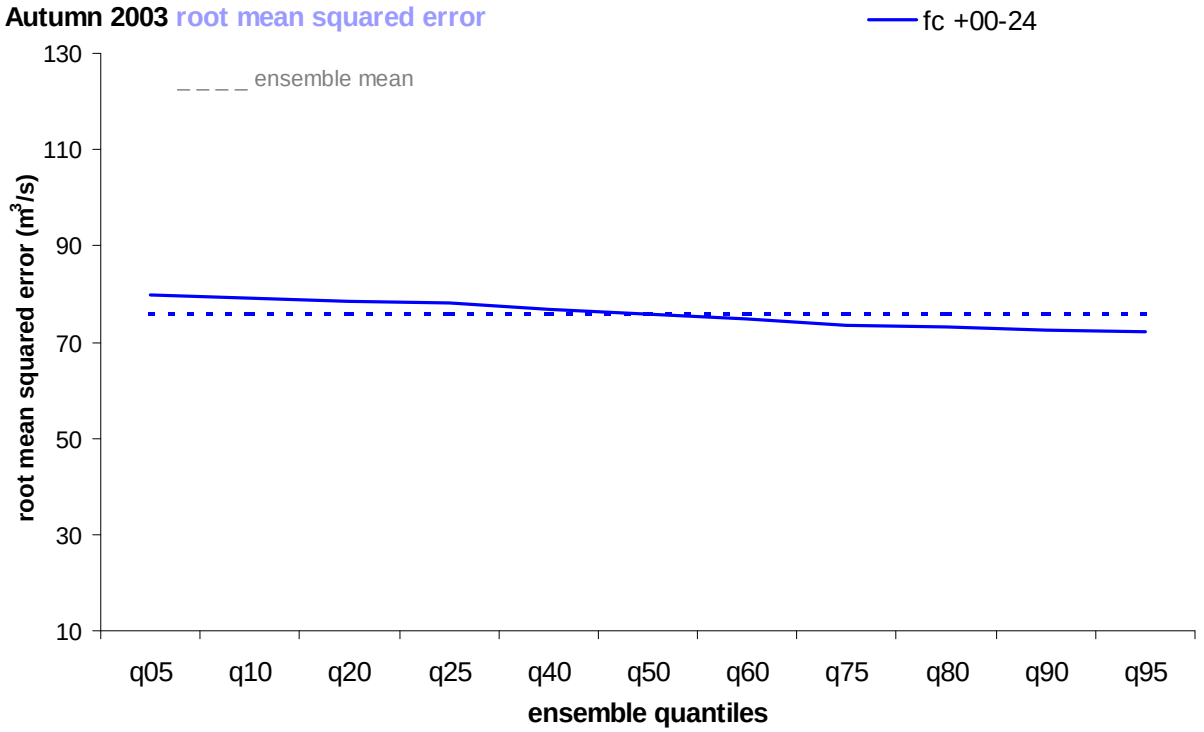
- forecast rainfall fields down-scaled to the hydrological model grid resolution (500 x 500 m) by assigning to each cell the QPF value provided on the nearest COSMO-LEPS grid point
- no stochastic procedure for rainfall downscaling
- no bias correction
- COSMO-LEPS QPF provided every 3 hours (6 hours for autumn 2003)
- all members equally probable in the direct coupling with the TOPKAPI

Preliminary results

Which quantile is more suitable to represent the ensemble forecast?

autumn 2003



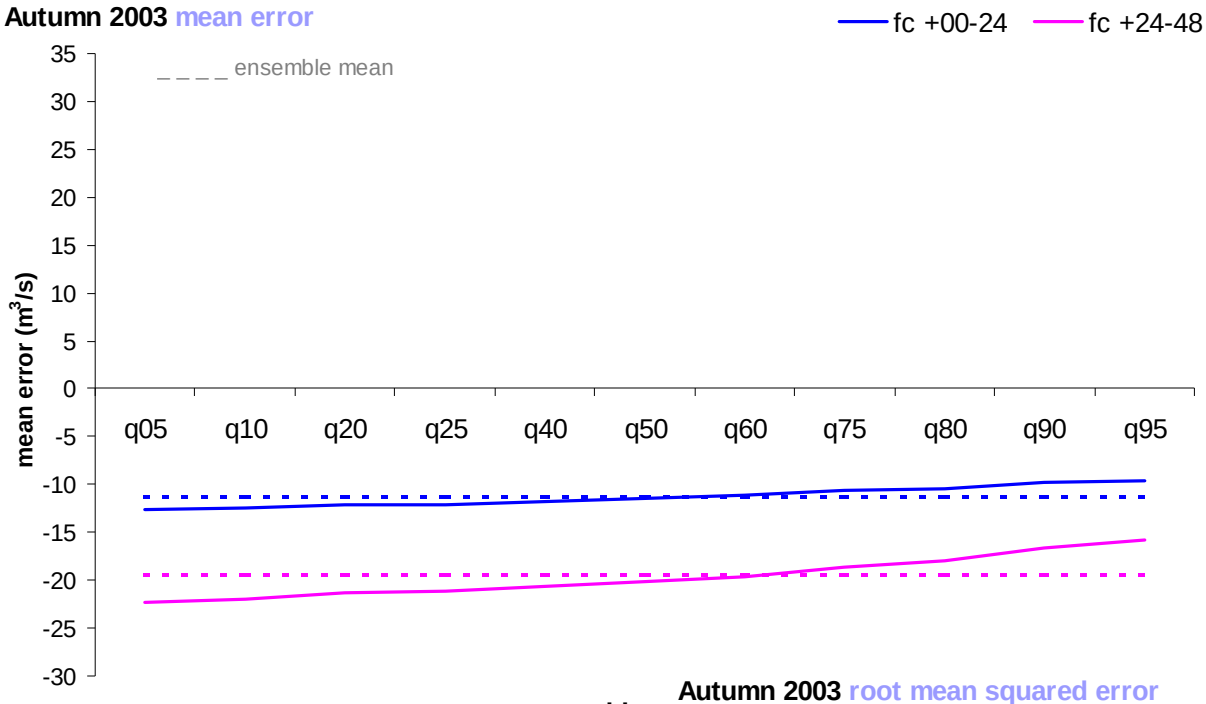


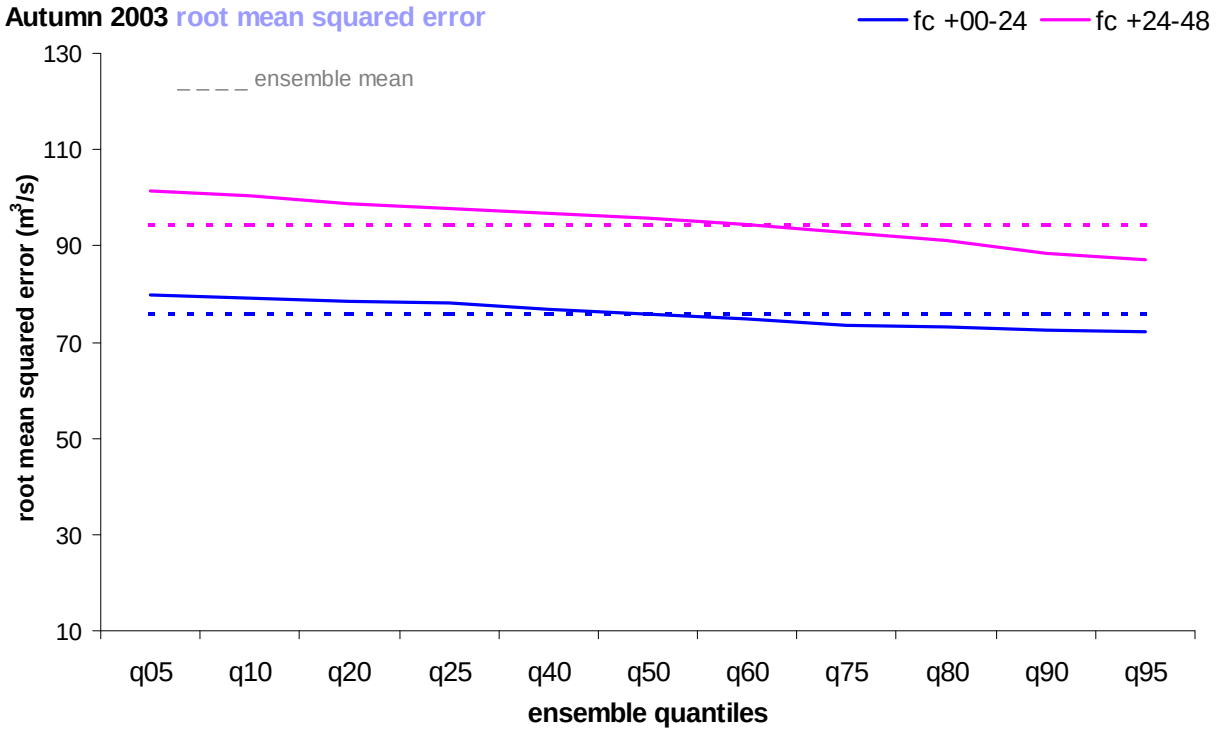
- The 80% and 90% quantiles perform generally better with respect to the ensemble mean as the lead-time increases. This trend is more evident when the autumn season is characterised by higher streamflow values (years 2003 and 2004).
- The performance decay of the probabilistic coupled system is evident up to the first 48-hour forecast range; for longer lead-times, the decay is partly compensated by the ensemble skill.

Preliminary results

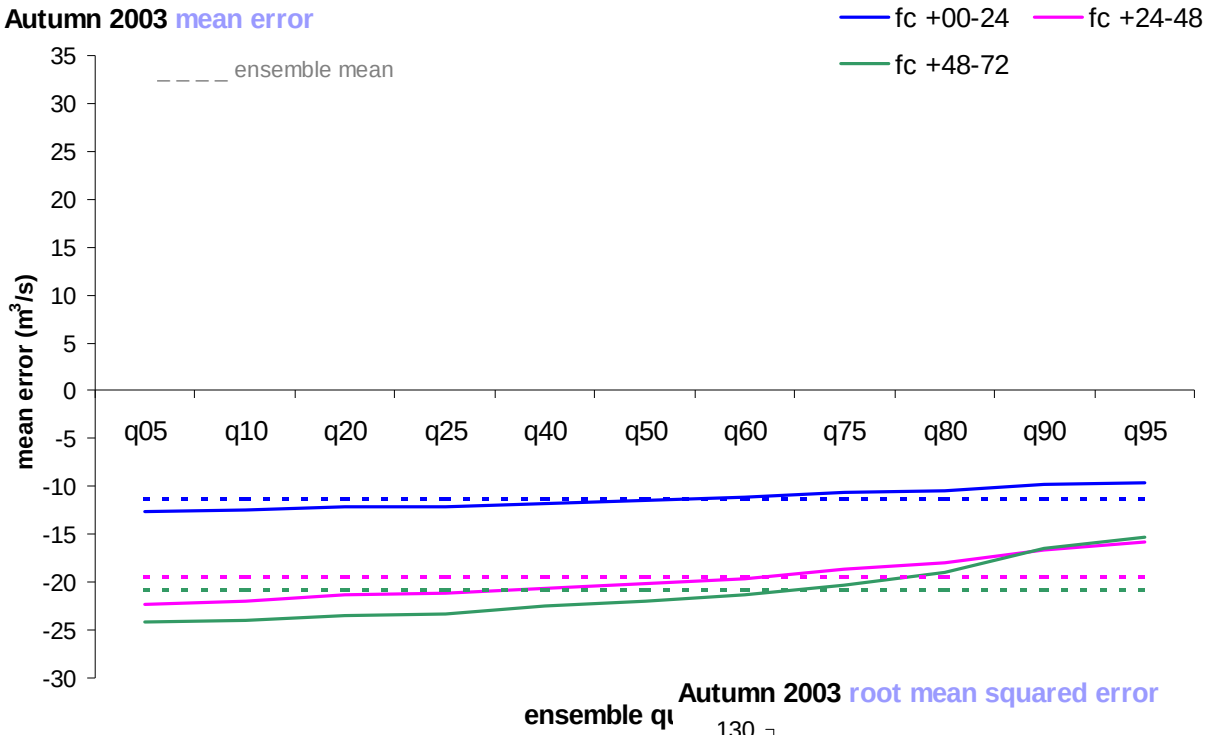
Which quantile is more suitable to represent the ensemble forecast?

autumn 2003





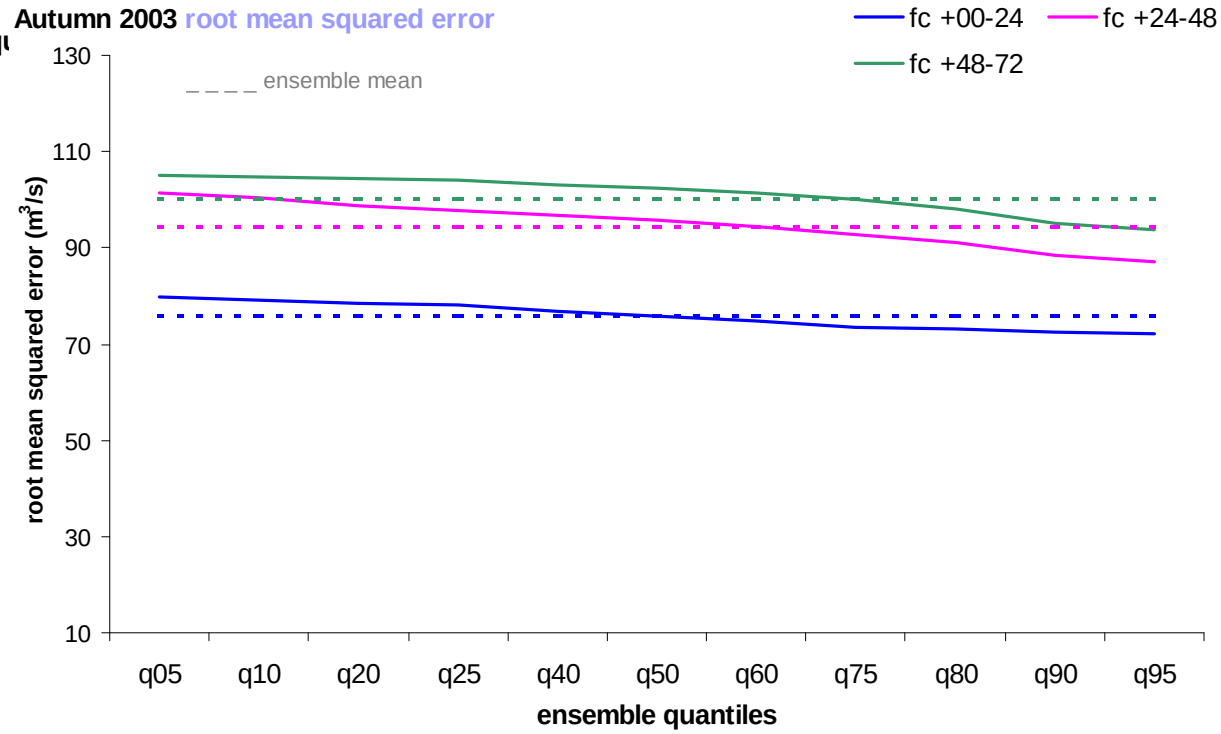
- The 80% and 90% quantiles perform generally better with respect to the ensemble mean as the lead-time increases. This trend is more evident when the autumn season is characterised by higher streamflow values (years 2003 and 2004).
- The performance decay of the probabilistic coupled system is evident up to the first 48-hour forecast range; for longer lead-times, the decay is partly compensated by the ensemble skill.



Preliminary results

Which quantile is more suitable to represent the ensemble forecast?

autumn 2003

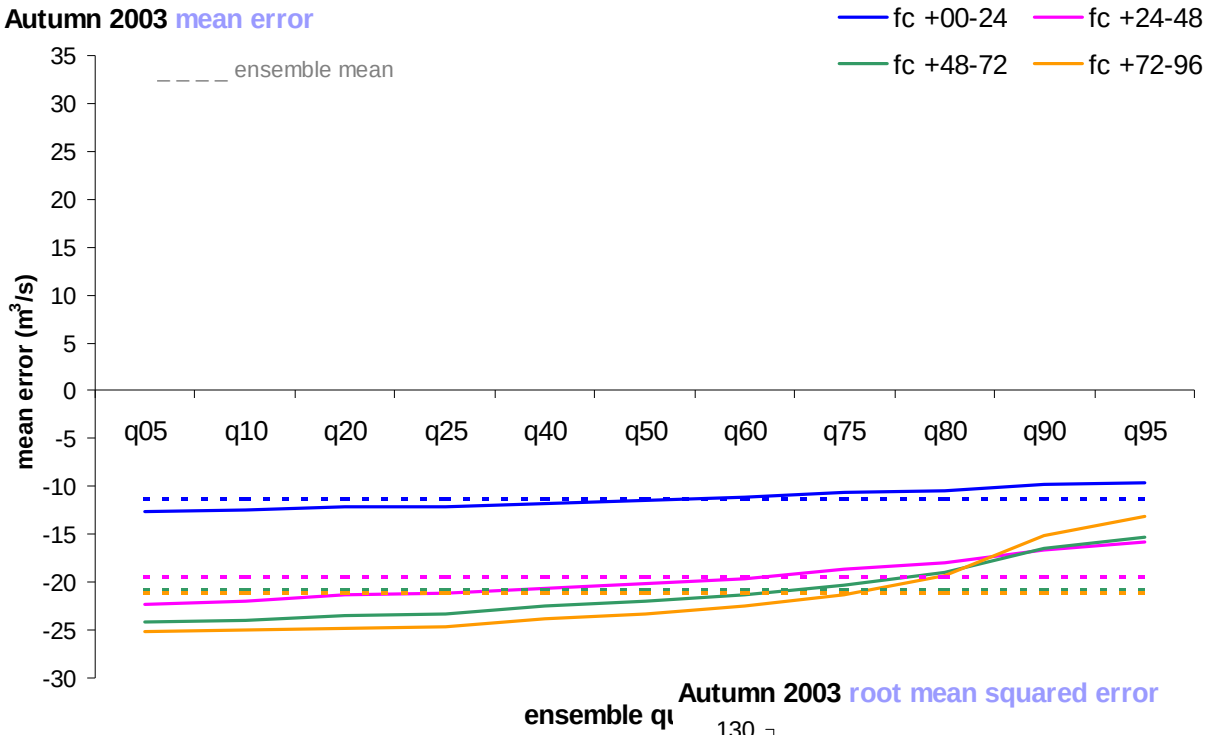


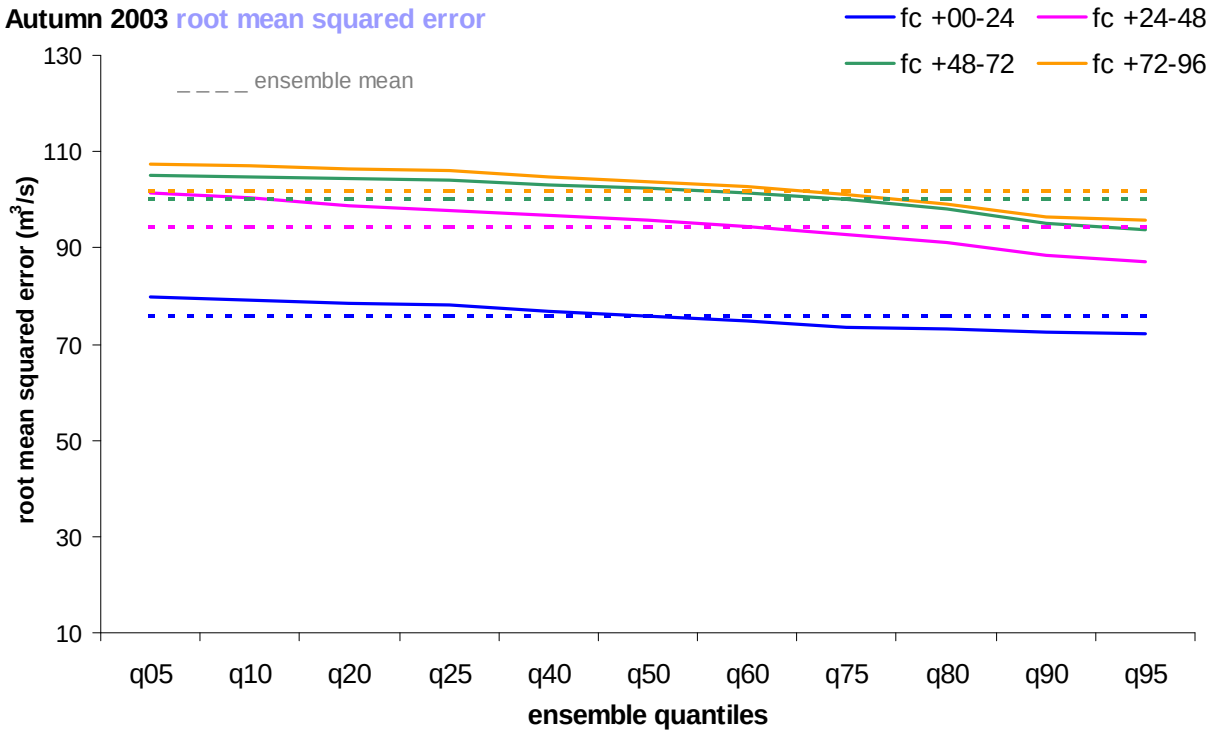
- The 80% and 90% quantiles perform generally better with respect to the ensemble mean as the lead-time increases. This trend is more evident when the autumn season is characterised by higher streamflow values (years 2003 and 2004).
- The performance decay of the probabilistic coupled system is evident up to the first 48-hour forecast range; for longer lead-times, the decay is partly compensated by the ensemble skill.

Preliminary results

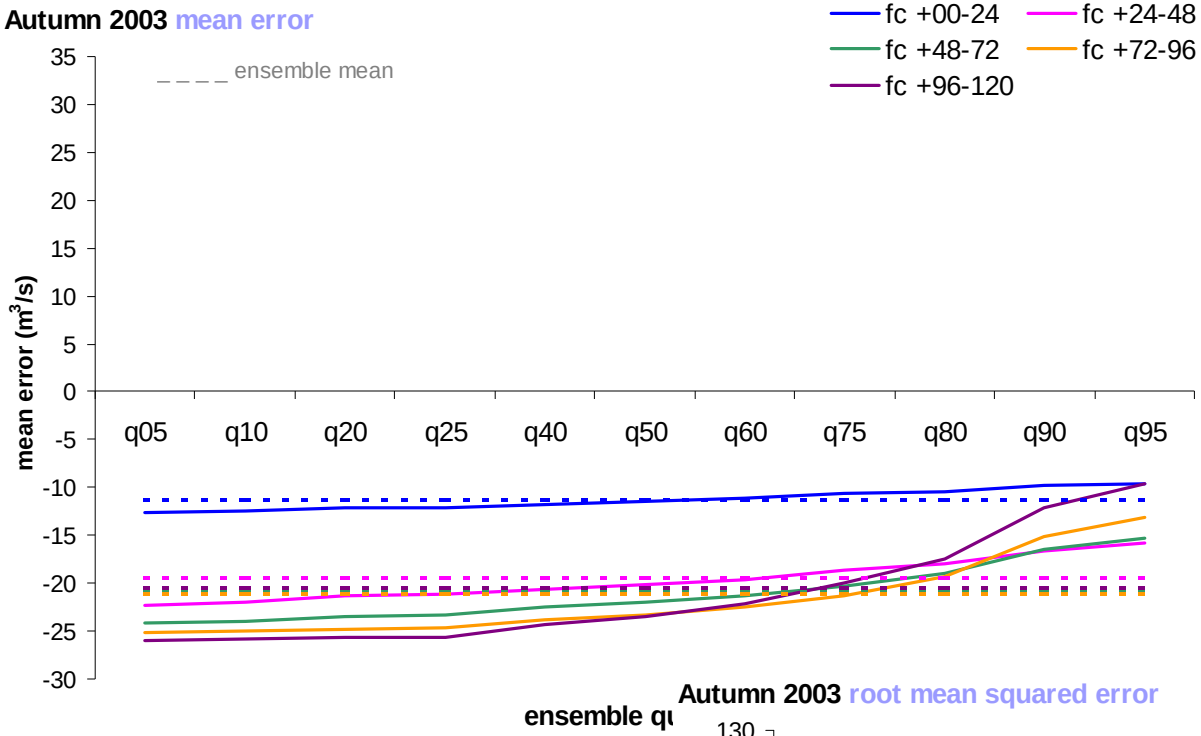
Which quantile is more suitable to represent the ensemble forecast?

autumn 2003





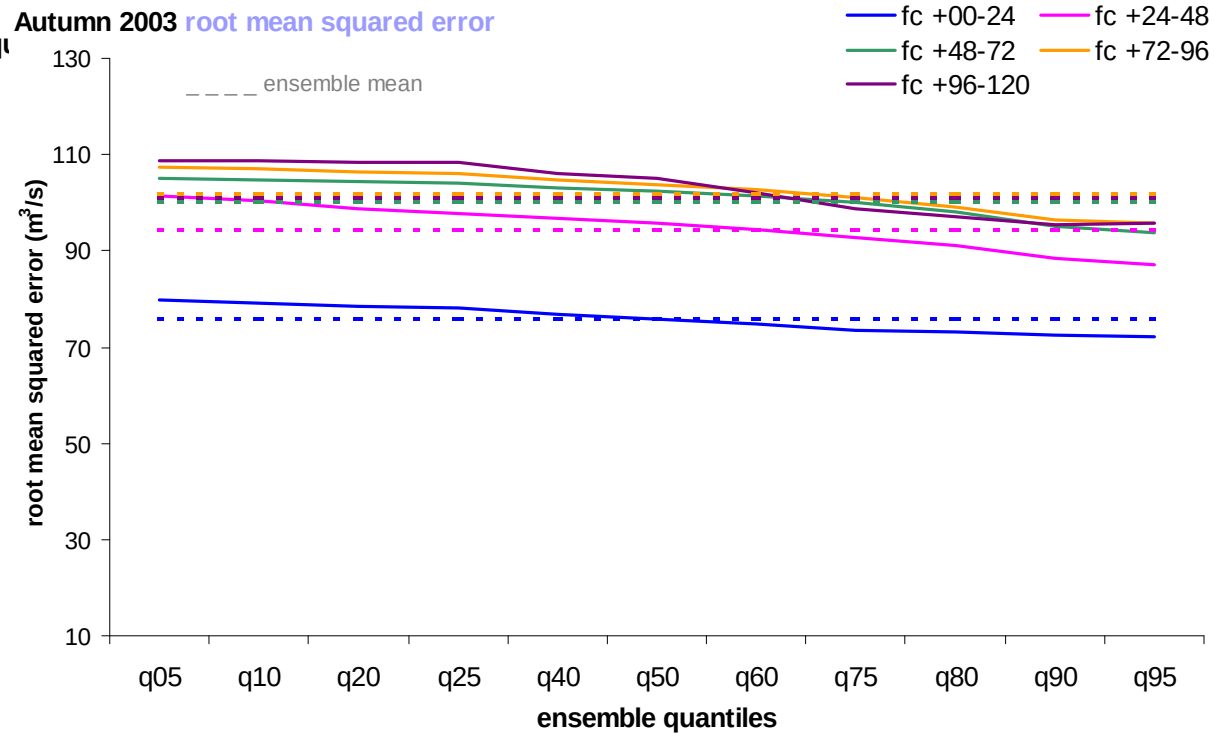
- The 80% and 90% quantiles perform generally better with respect to the ensemble mean as the lead-time increases. This trend is more evident when the autumn season is characterised by higher streamflow values (years 2003 and 2004).
- The performance decay of the probabilistic coupled system is evident up to the first 48-hour forecast range; for longer lead-times, the decay is partly compensated by the ensemble skill.



Preliminary results

Which quantile is more suitable to represent the ensemble forecast?

autumn 2003

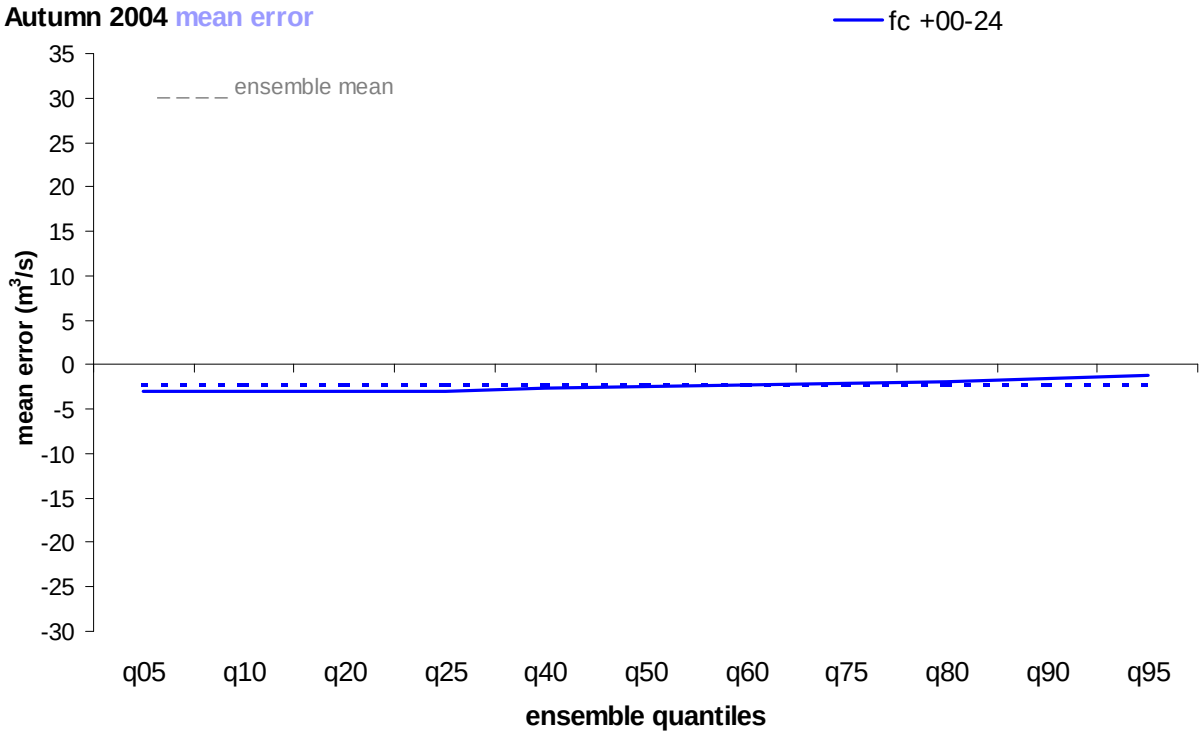


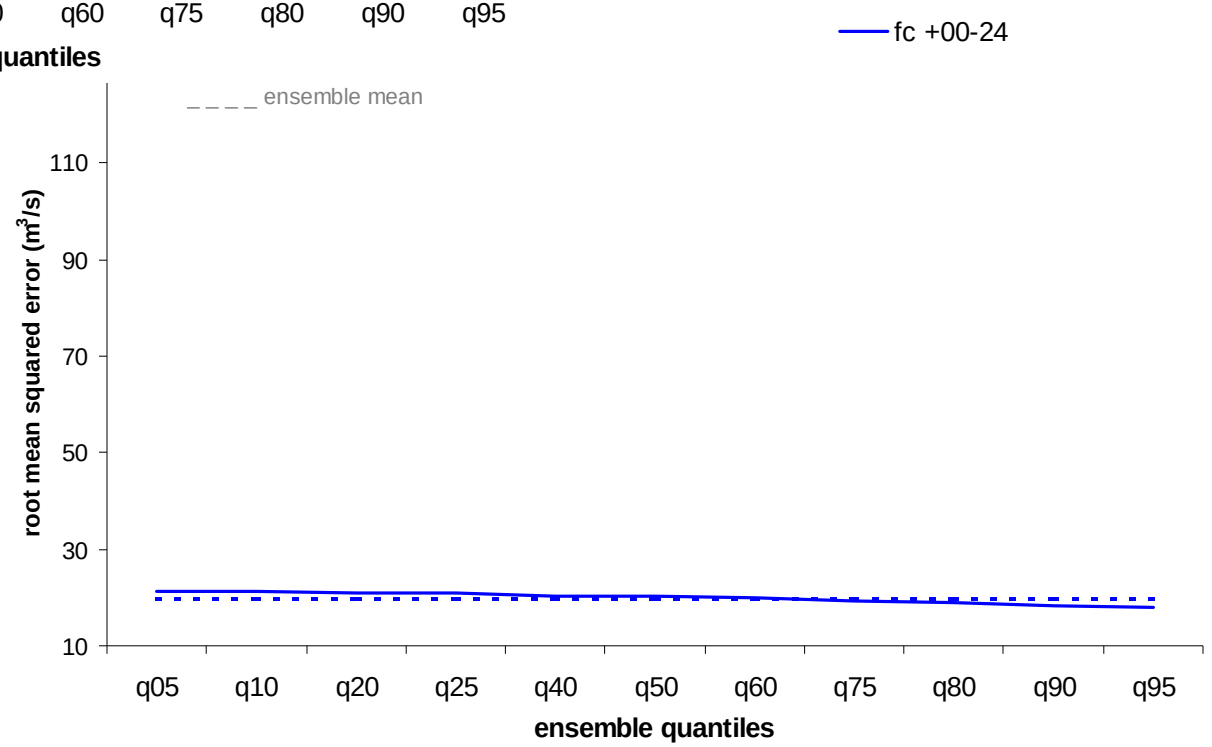
- The 80% and 90% quantiles perform generally better with respect to the ensemble mean as the lead-time increases. This trend is more evident when the autumn season is characterised by higher streamflow values (years 2003 and 2004).
- The performance decay of the probabilistic coupled system is evident up to the first 48-hour forecast range; for longer lead-times, the decay is partly compensated by the ensemble skill.

Preliminary results

Which quantile is more suitable to represent the ensemble forecast?

autumn 2004



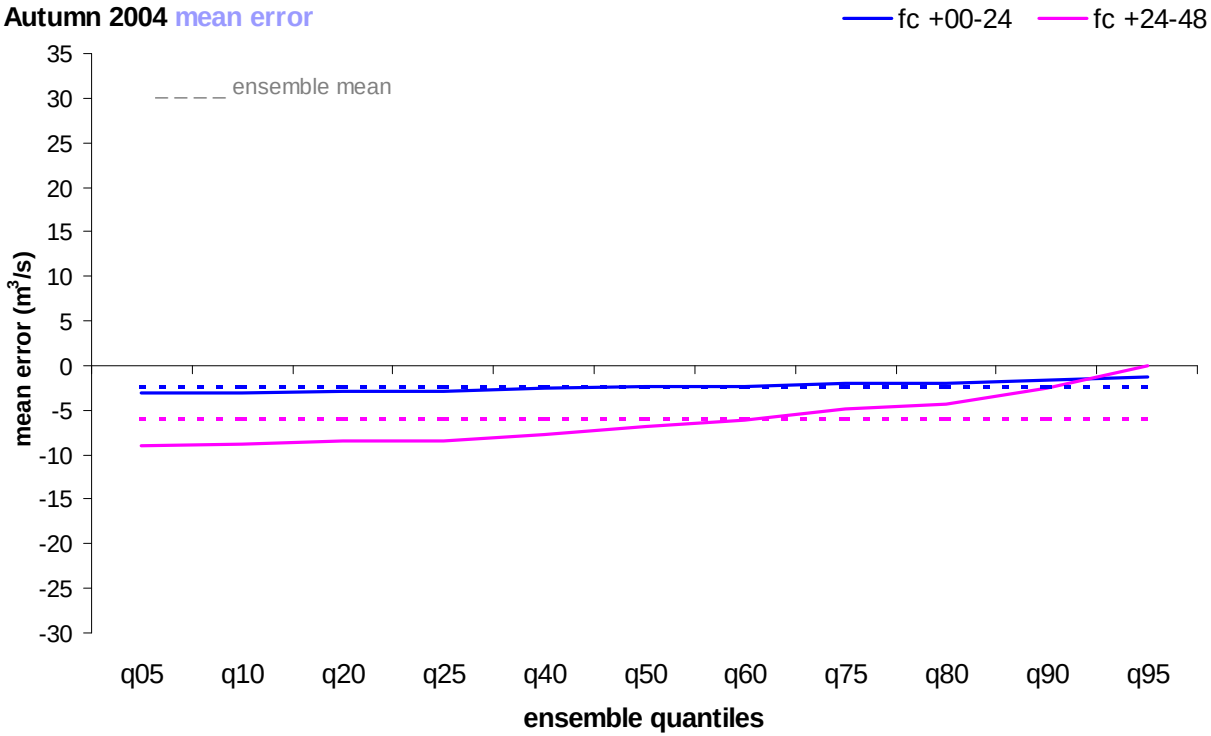


- The 80% and 90% quantiles perform generally better with respect to the ensemble mean as the lead-time increases. This trend is more evident when the autumn season is characterised by higher streamflow values (years 2003 and 2004).
- The performance decay of the probabilistic coupled system is evident up to the first 48-hour forecast range; for longer lead-times, the decay is partly compensated by the ensemble skill.

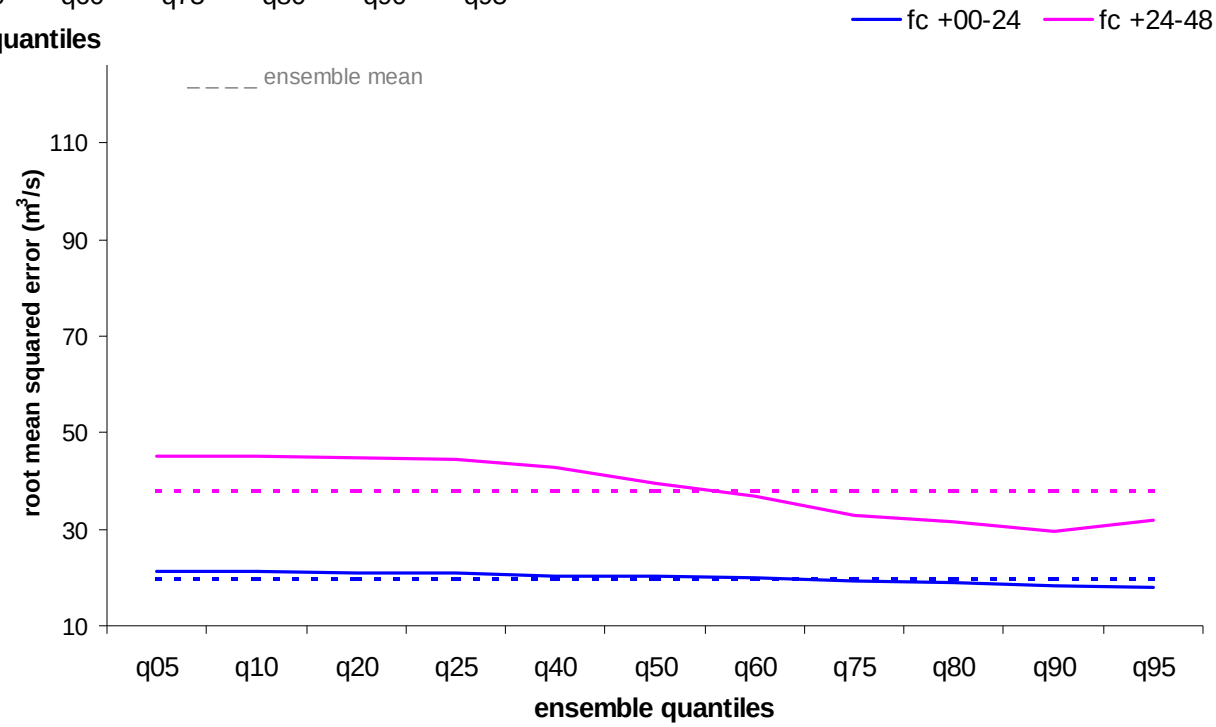
Preliminary results

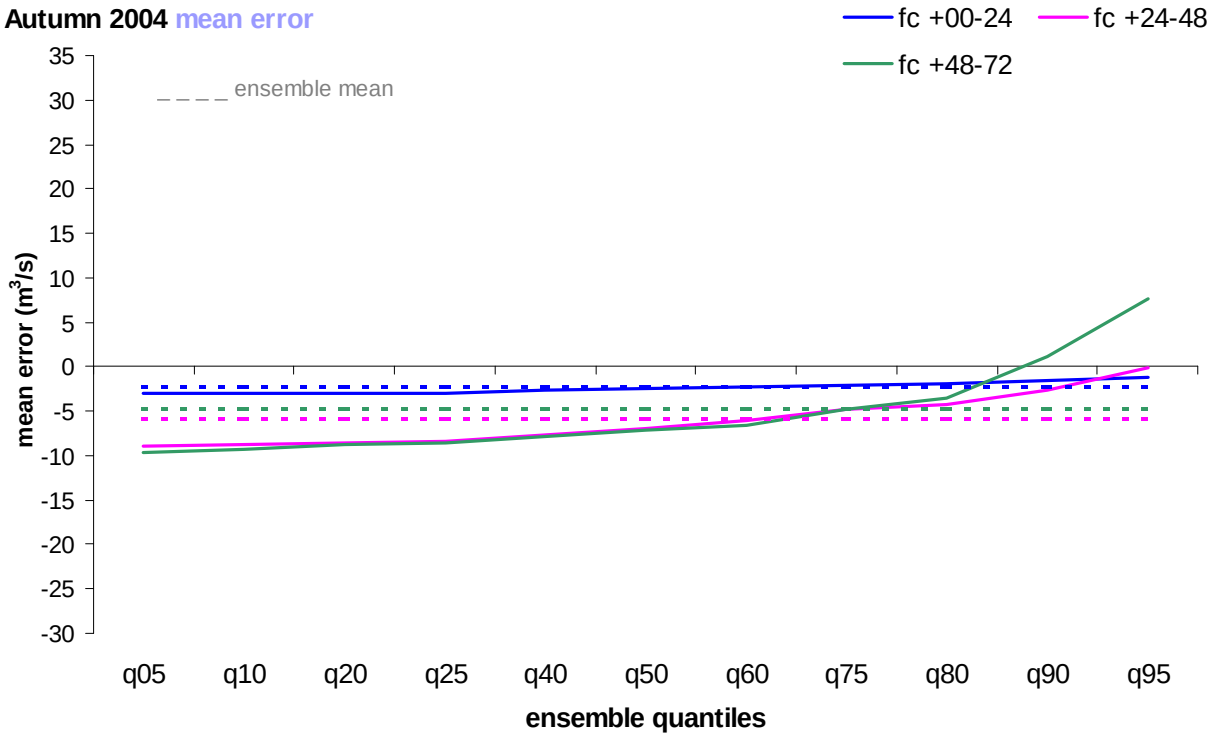
Which quantile is more suitable to represent the ensemble forecast?

autumn 2004



- The 80% and 90% quantiles perform generally better with respect to the ensemble mean as the lead-time increases. This trend is more evident when the autumn season is characterised by higher streamflow values (years 2003 and 2004).
- The performance decay of the probabilistic coupled system is evident up to the first 48-hour forecast range; for longer lead-times, the decay is partly compensated by the ensemble skill.



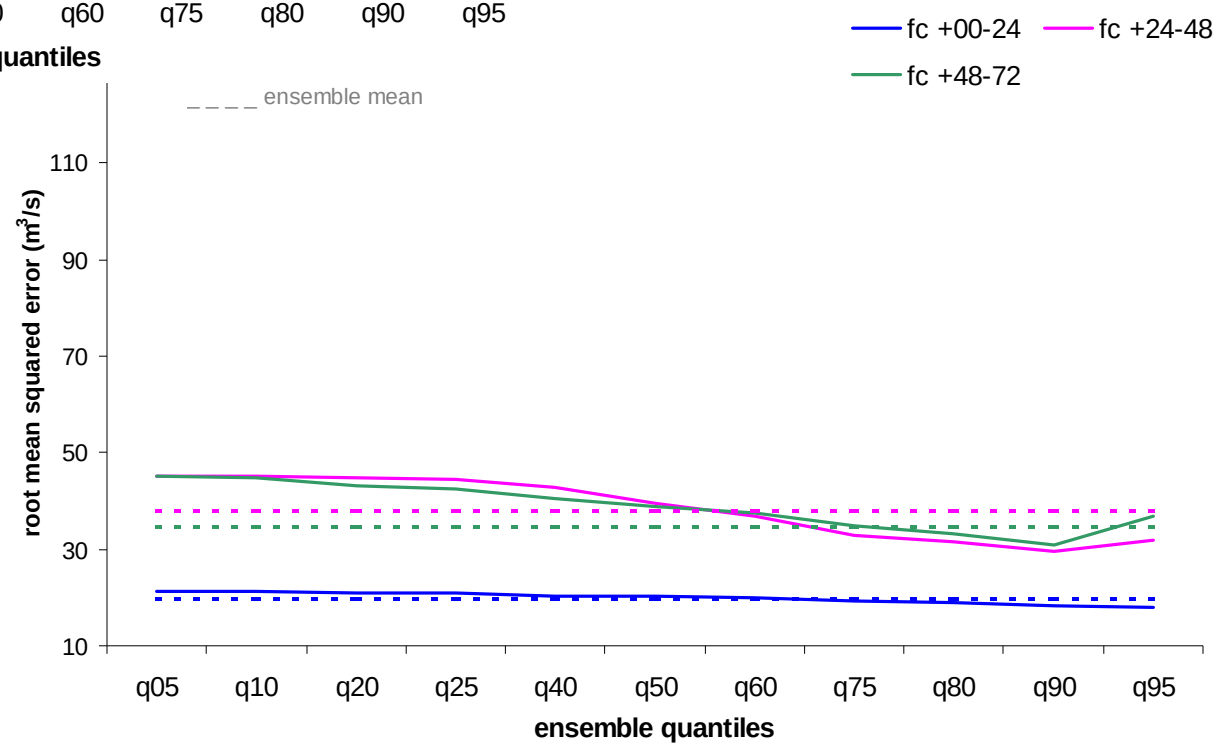


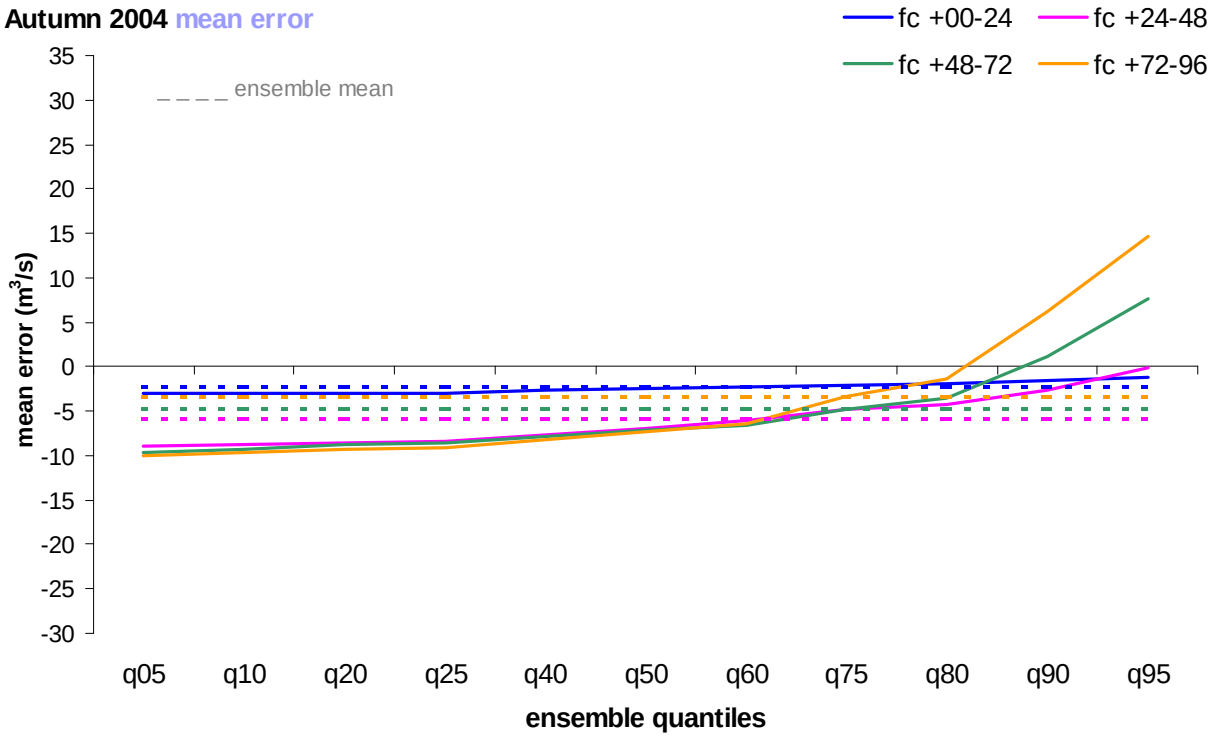
Preliminary results

Which quantile is more suitable to represent the ensemble forecast?

autumn 2004

- The 80% and 90% quantiles perform generally better with respect to the ensemble mean as the lead-time increases. This trend is more evident when the autumn season is characterised by higher streamflow values (years 2003 and 2004).
- The performance decay of the probabilistic coupled system is evident up to the first 48-hour forecast range; for longer lead-times, the decay is partly compensated by the ensemble skill.



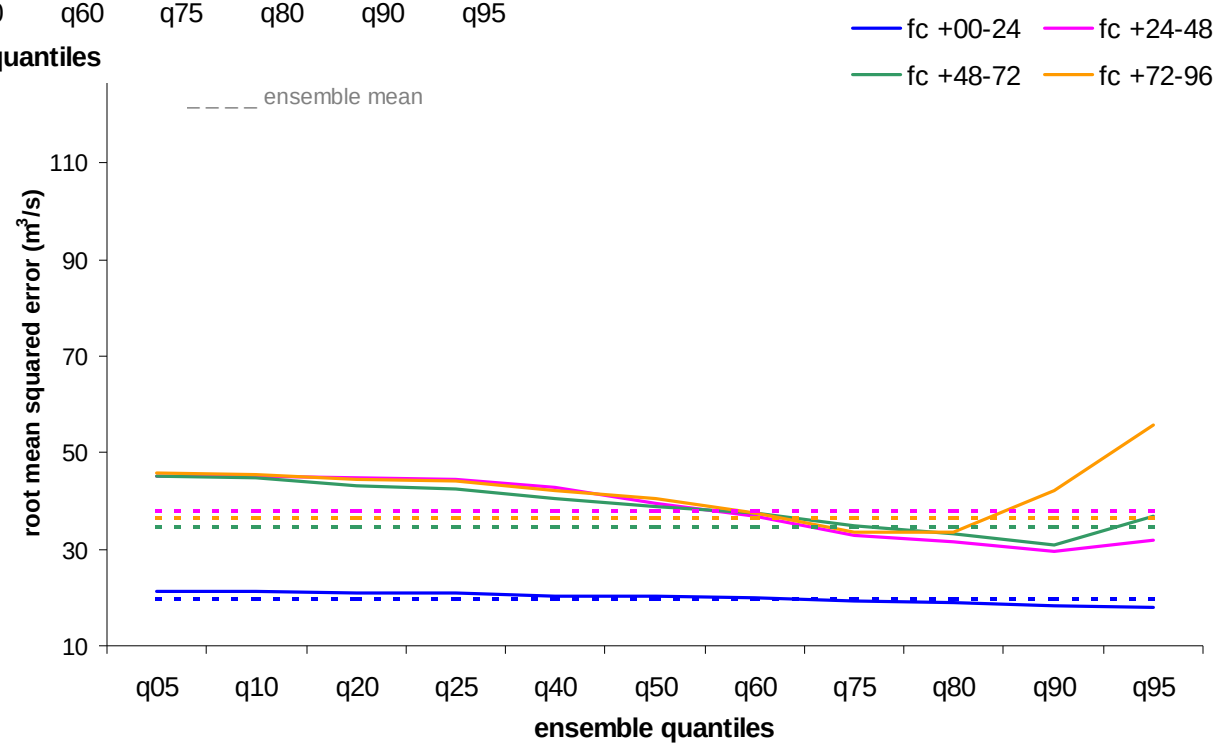


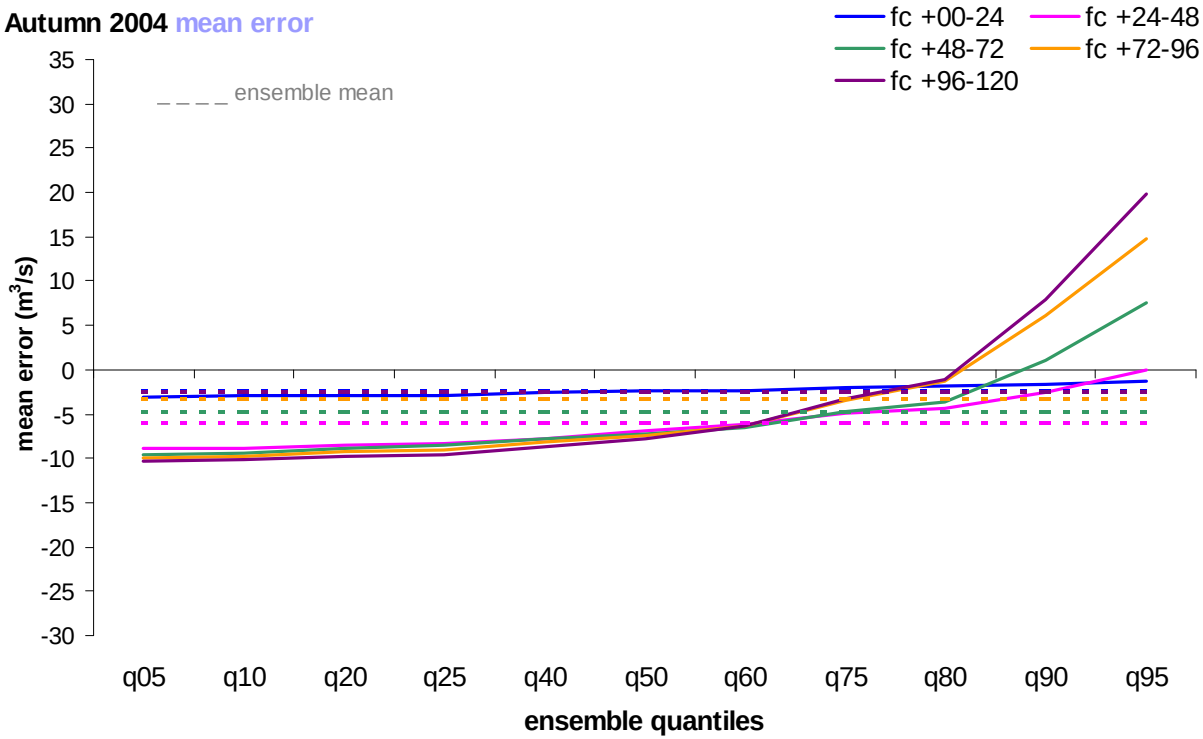
Preliminary results

Which quantile is more suitable to represent the ensemble forecast?

autumn 2004

- The 80% and 90% quantiles perform generally better with respect to the ensemble mean as the lead-time increases. This trend is more evident when the autumn season is characterised by higher streamflow values (years 2003 and 2004).
- The performance decay of the probabilistic coupled system is evident up to the first 48-hour forecast range; for longer lead-times, the decay is partly compensated by the ensemble skill.



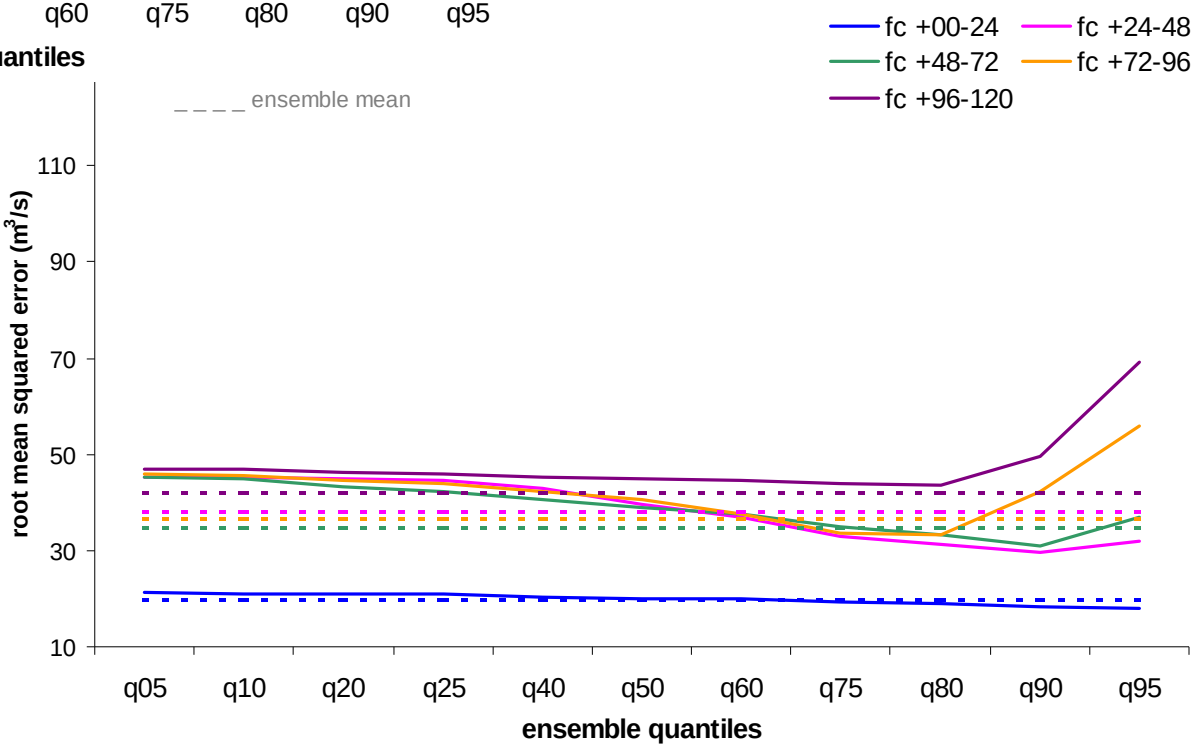


Preliminary results

Which quantile is more suitable to represent the ensemble forecast?

autumn 2004

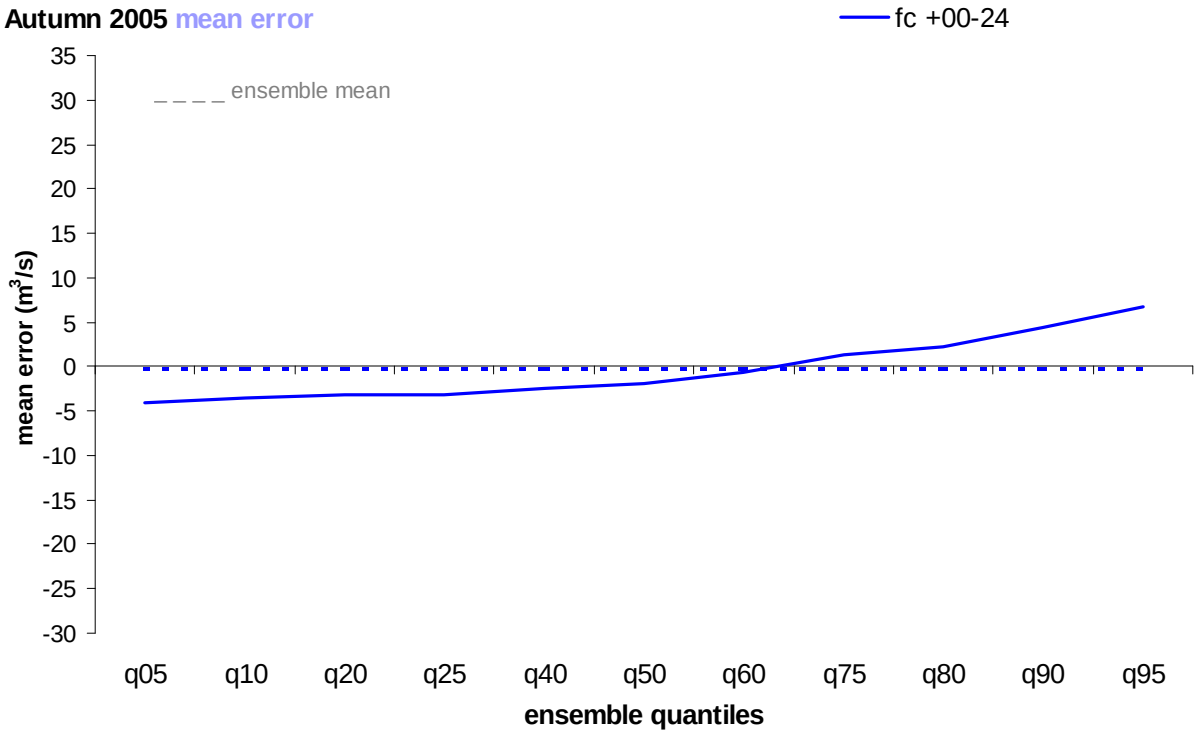
- The 80% and 90% quantiles perform generally better with respect to the ensemble mean as the lead-time increases. This trend is more evident when the autumn season is characterised by higher streamflow values (years 2003 and 2004).
- The performance decay of the probabilistic coupled system is evident up to the first 48-hour forecast range; for longer lead-times, the decay is partly compensated by the ensemble skill.



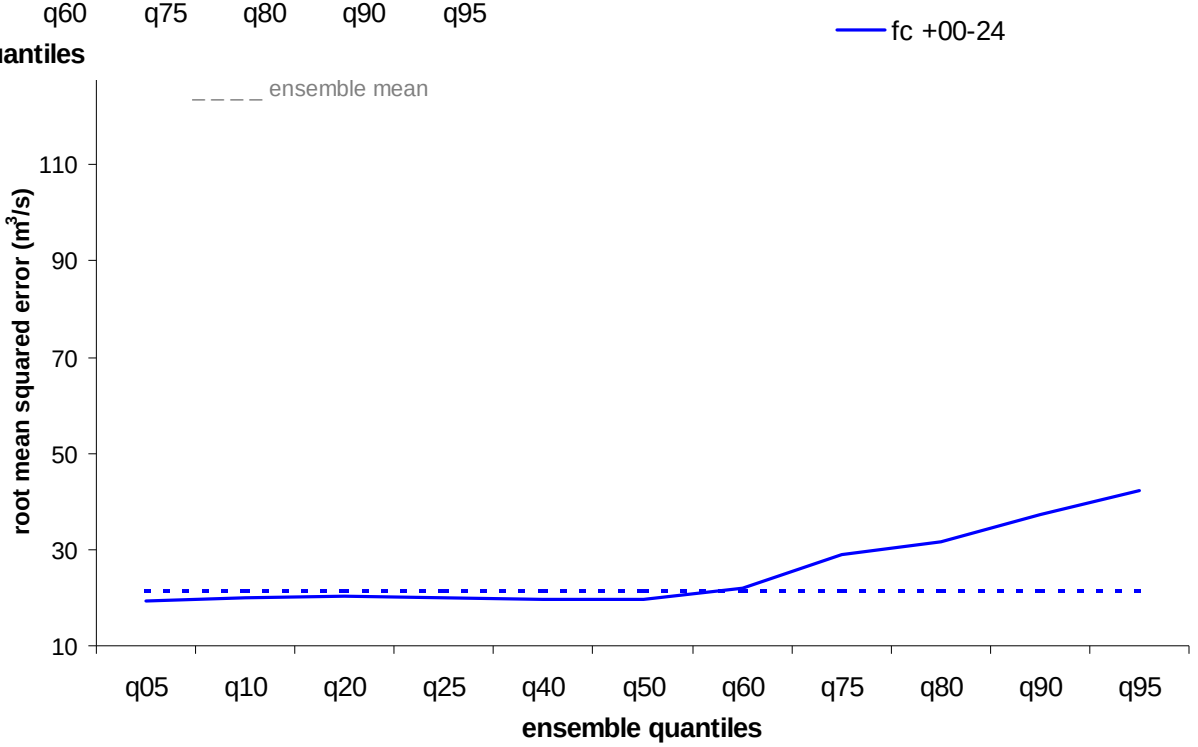
Preliminary results

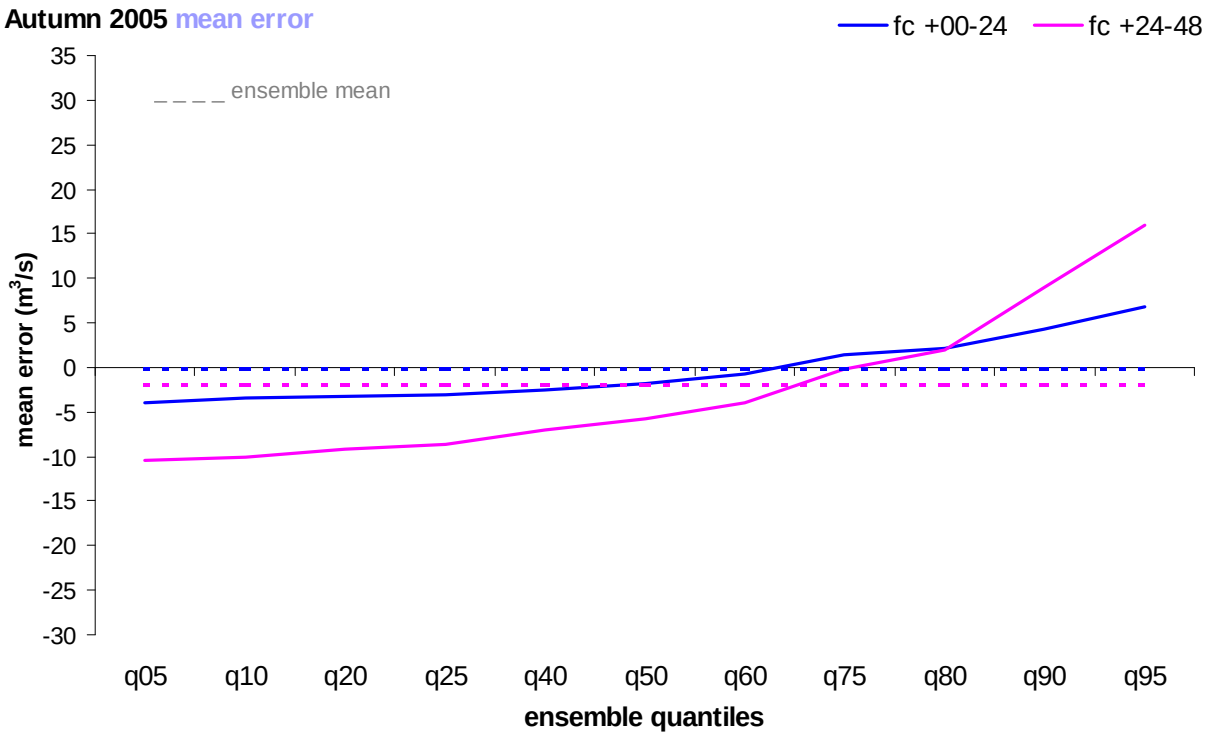
Which quantile is more suitable to represent the ensemble forecast?

autumn 2005



- The 80% and 90% quantiles perform generally better with respect to the ensemble mean as the lead-time increases. This trend is more evident when the autumn season is characterised by higher streamflow values (years 2003 and 2004).
- The performance decay of the probabilistic coupled system is evident up to the first 48-hour forecast range; for longer lead-times, the decay is partly compensated by the ensemble skill.



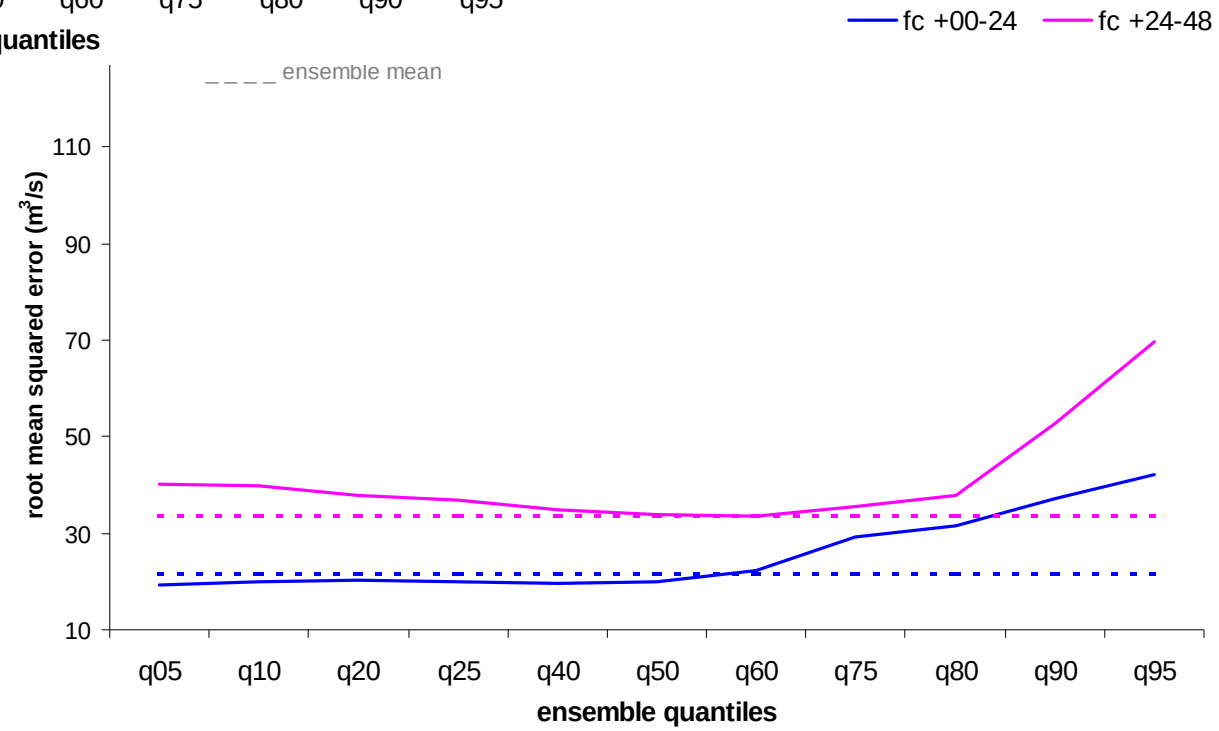


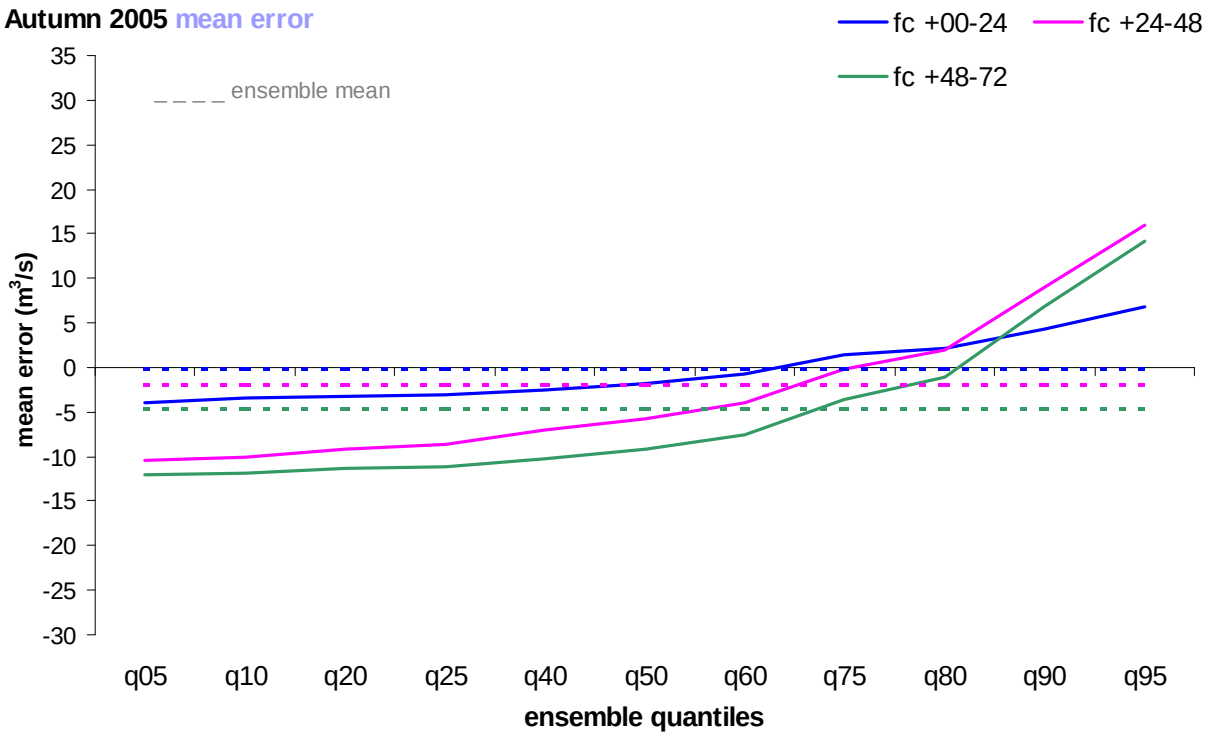
Preliminary results

Which quantile is more suitable to represent the ensemble forecast?

autumn 2005

- The 80% and 90% quantiles perform generally better with respect to the ensemble mean as the lead-time increases. This trend is more evident when the autumn season is characterised by higher streamflow values (years 2003 and 2004).
- The performance decay of the probabilistic coupled system is evident up to the first 48-hour forecast range; for longer lead-times, the decay is partly compensated by the ensemble skill.



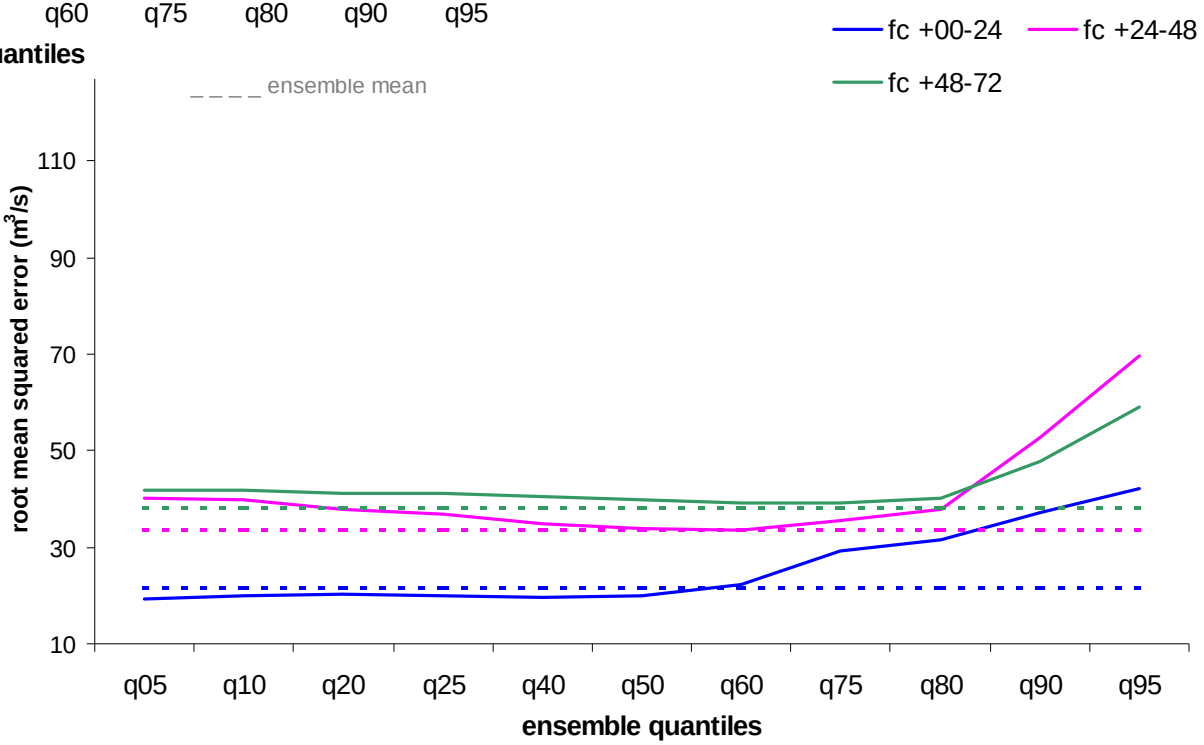


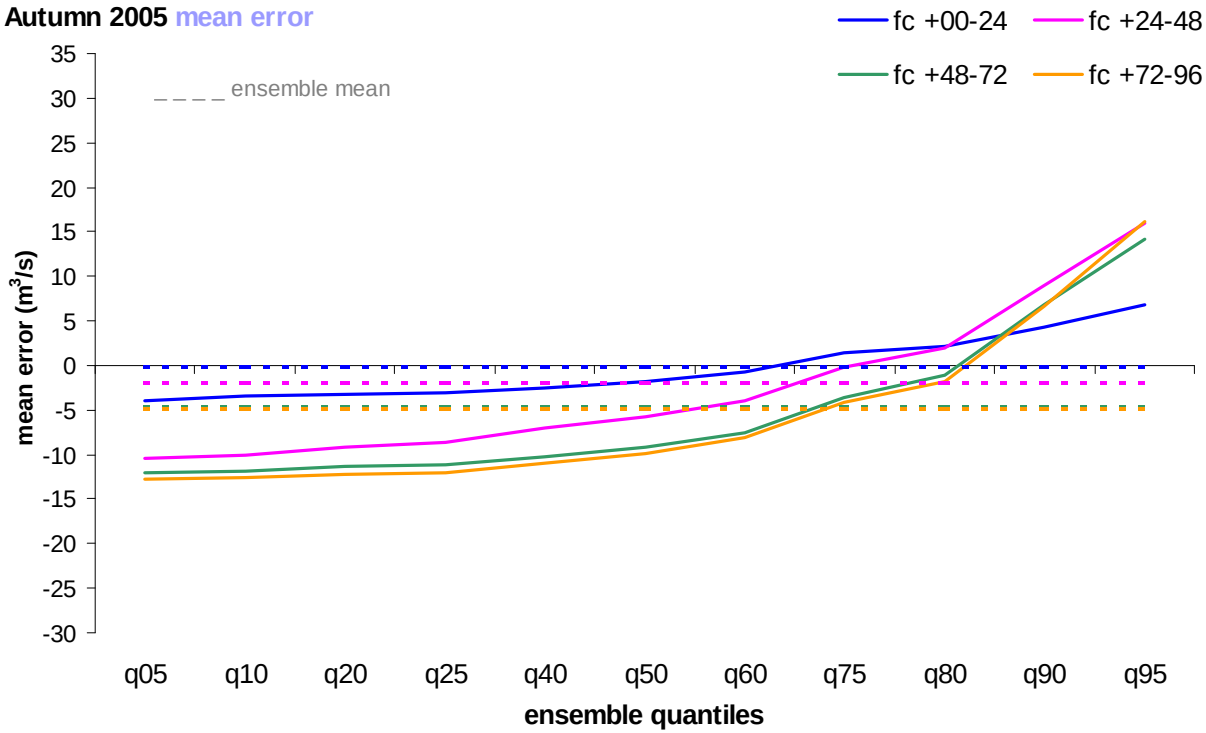
Preliminary results

Which quantile is more suitable to represent the ensemble forecast?

autumn 2005

- The 80% and 90% quantiles perform generally better with respect to the ensemble mean as the lead-time increases. This trend is more evident when the autumn season is characterised by higher streamflow values (years 2003 and 2004).
- The performance decay of the probabilistic coupled system is evident up to the first 48-hour forecast range; for longer lead-times, the decay is partly compensated by the ensemble skill.



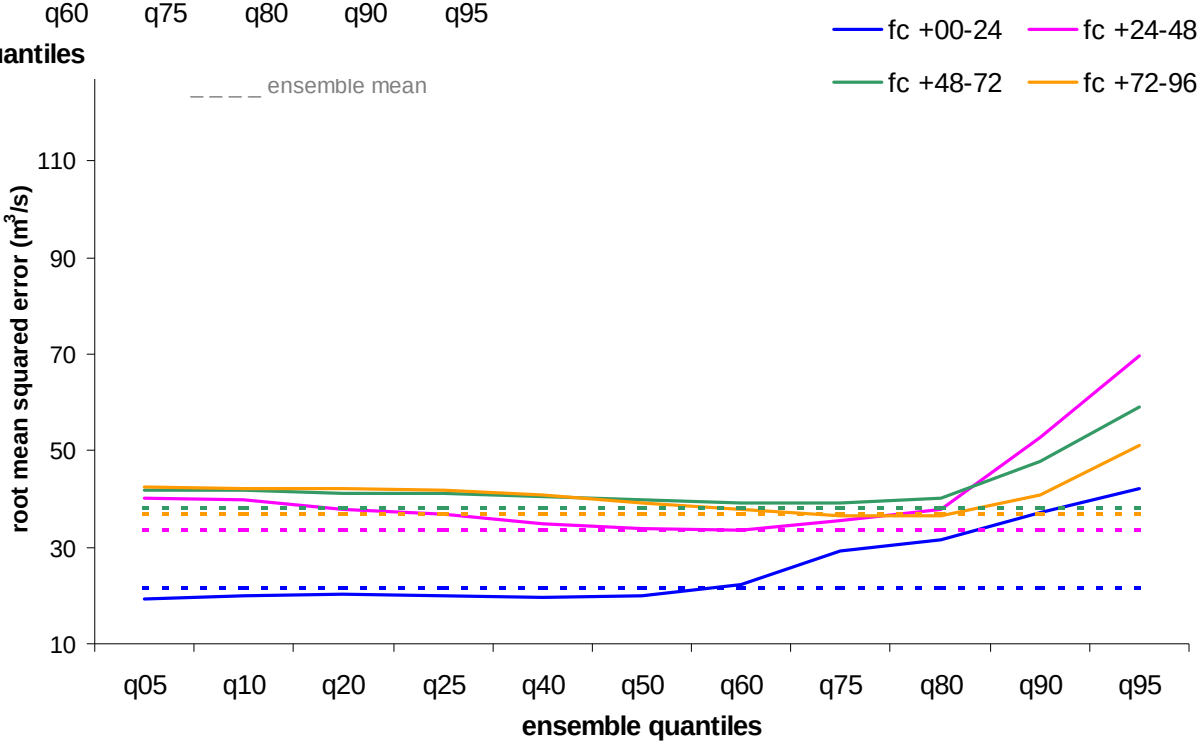


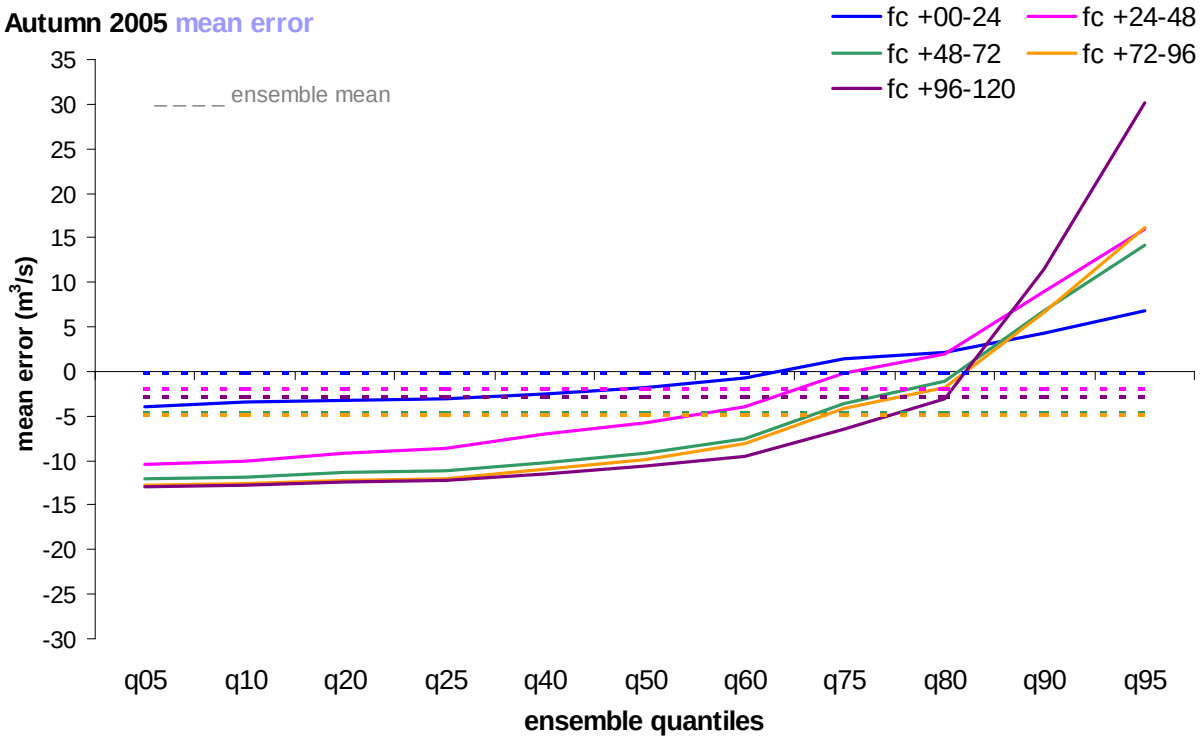
Preliminary results

Which quantile is more suitable to represent the ensemble forecast?

autumn 2005

- The 80% and 90% quantiles perform generally better with respect to the ensemble mean as the lead-time increases. This trend is more evident when the autumn season is characterised by higher streamflow values (years 2003 and 2004).
- The performance decay of the probabilistic coupled system is evident up to the first 48-hour forecast range; for longer lead-times, the decay is partly compensated by the ensemble skill.



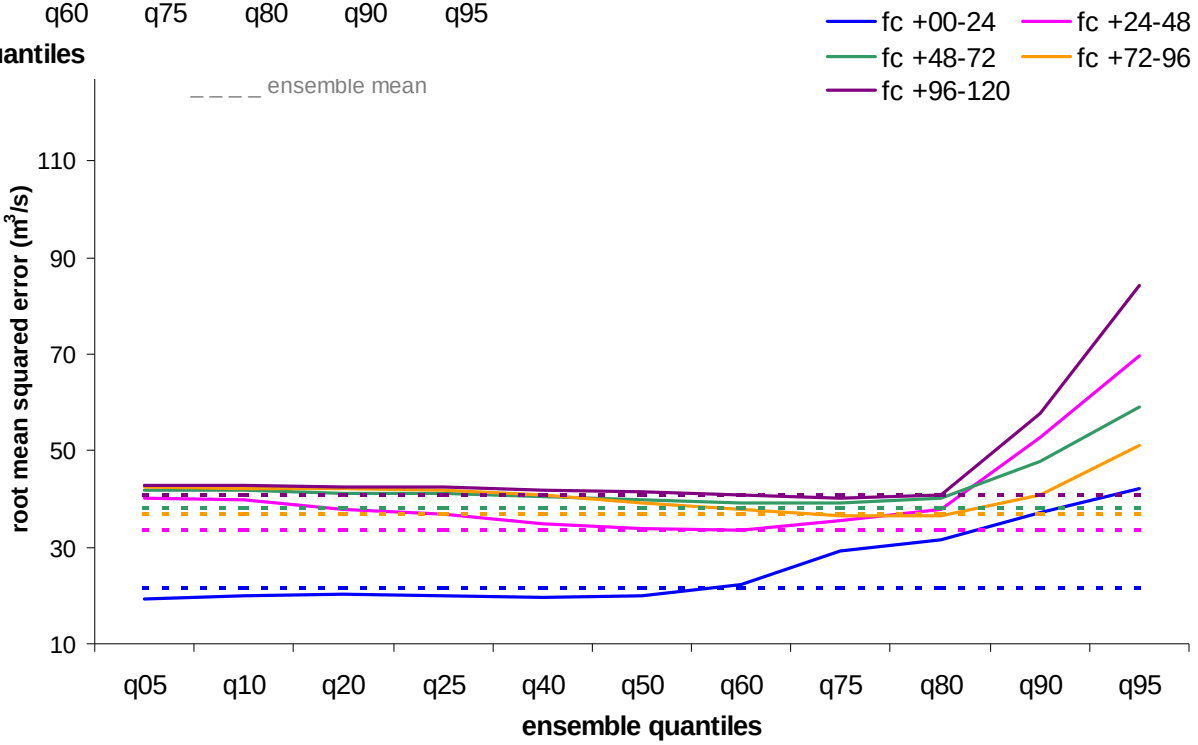


Preliminary results

Which quantile is more suitable to represent the ensemble forecast?

autumn 2005

- The 80% and 90% quantiles perform generally better with respect to the ensemble mean as the lead-time increases. This trend is more evident when the autumn season is characterised by higher streamflow values (years 2003 and 2004).
- The performance decay of the probabilistic coupled system is evident up to the first 48-hour forecast range; for longer lead-times, the decay is partly compensated by the ensemble skill.



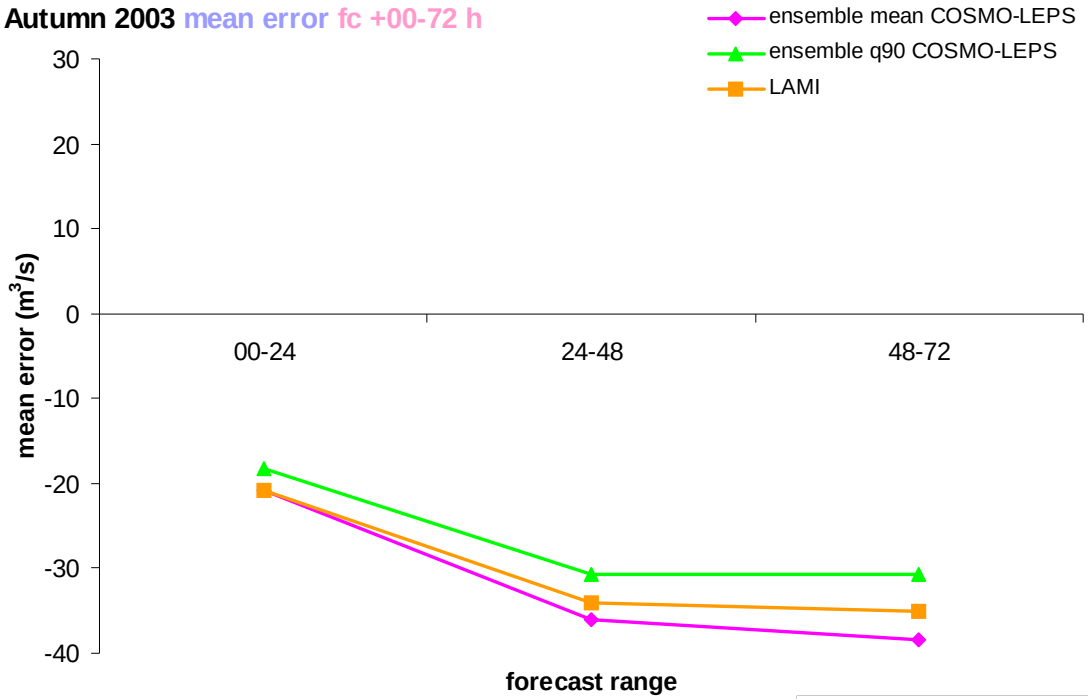
Preliminary results

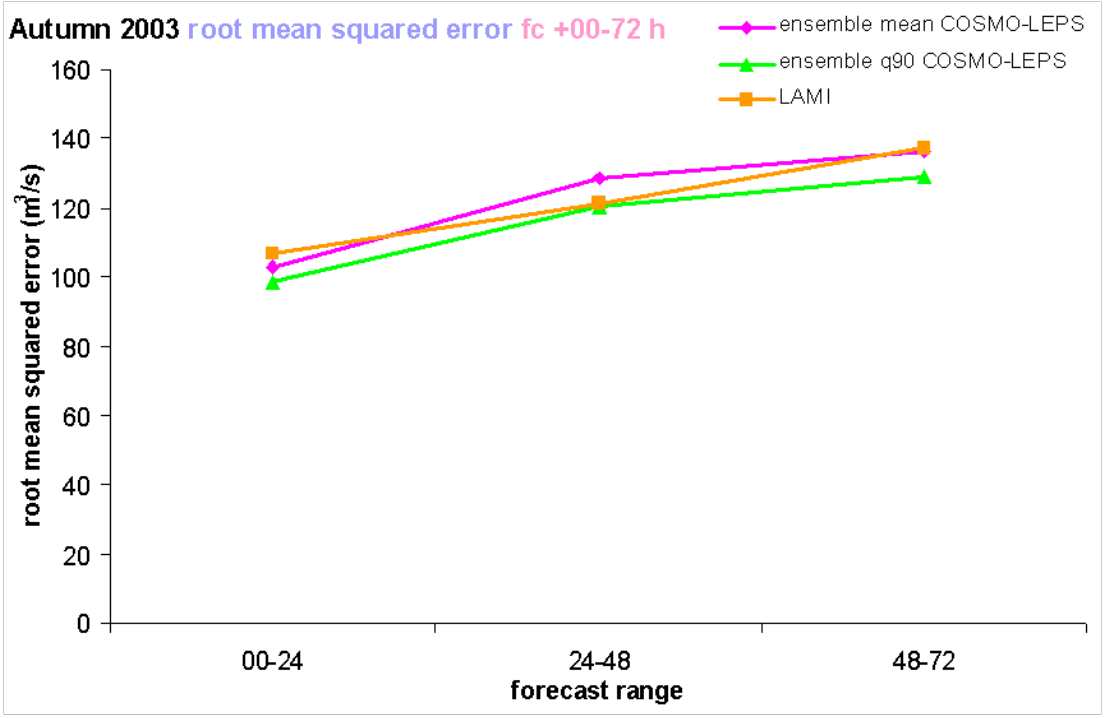
probabilistic chain

vs

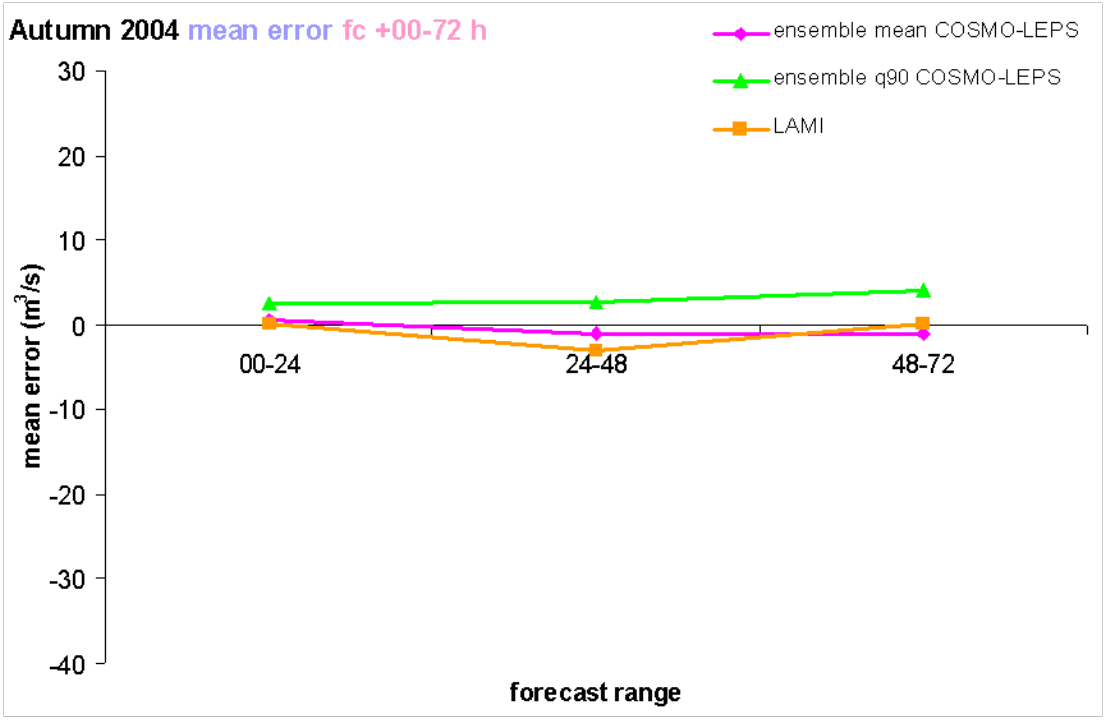
deterministic chain

autumn 2003





- For the first 24-h forecast range, the ensemble mean of the discharge simulations driven by COSMO-LEPS shows similar performance with respect to the deterministic forecast.
- The added value of the probabilistic system comes out if the 90% quantile is considered.



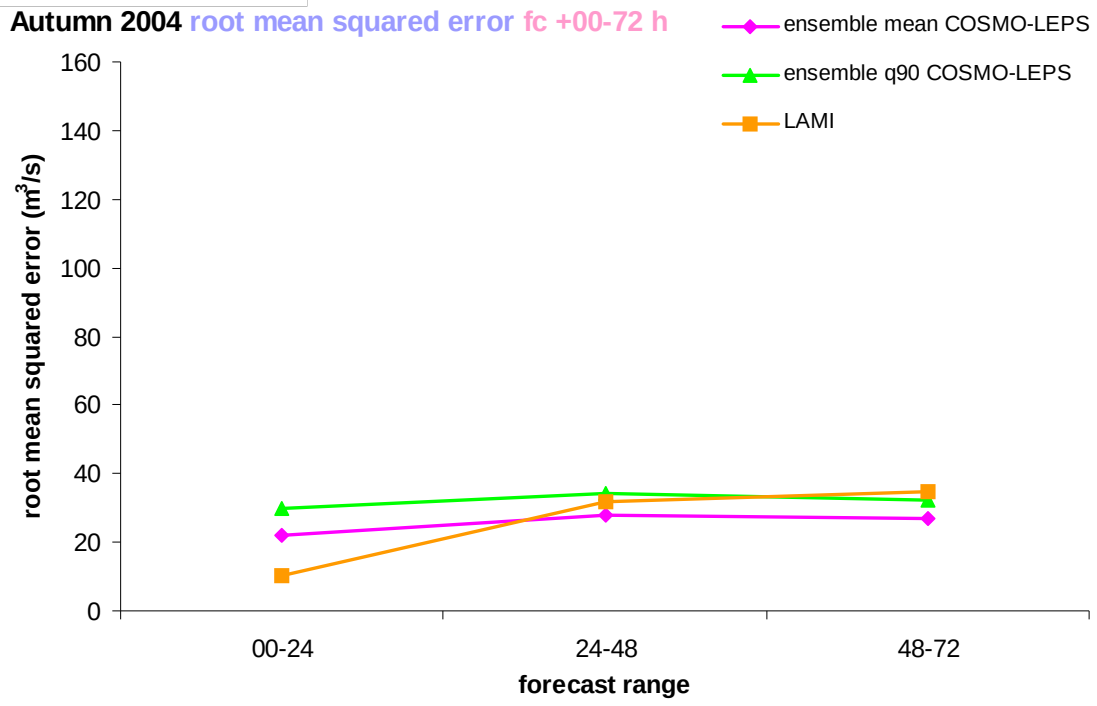
Preliminary results

probabilistic chain

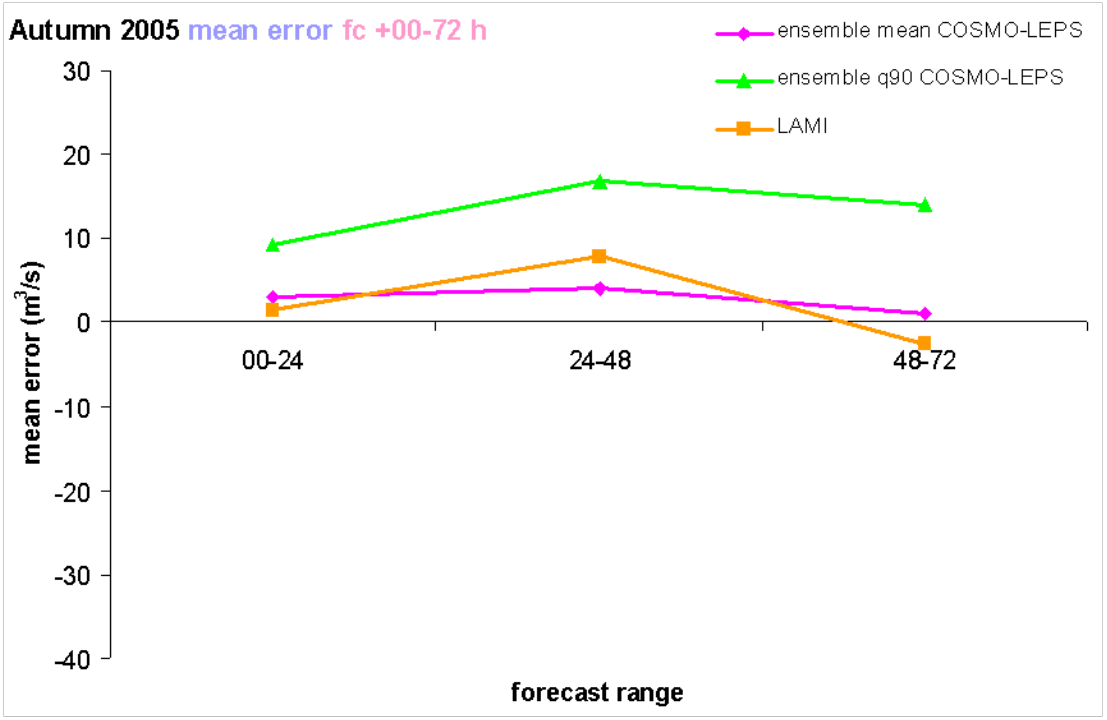
vs

deterministic chain

autumn 2004



The discharge simulations based on the deterministic precipitation forecast perform slightly better for the first 24-hour range. The added value of the probabilistic system is more evident for longer forecast ranges.



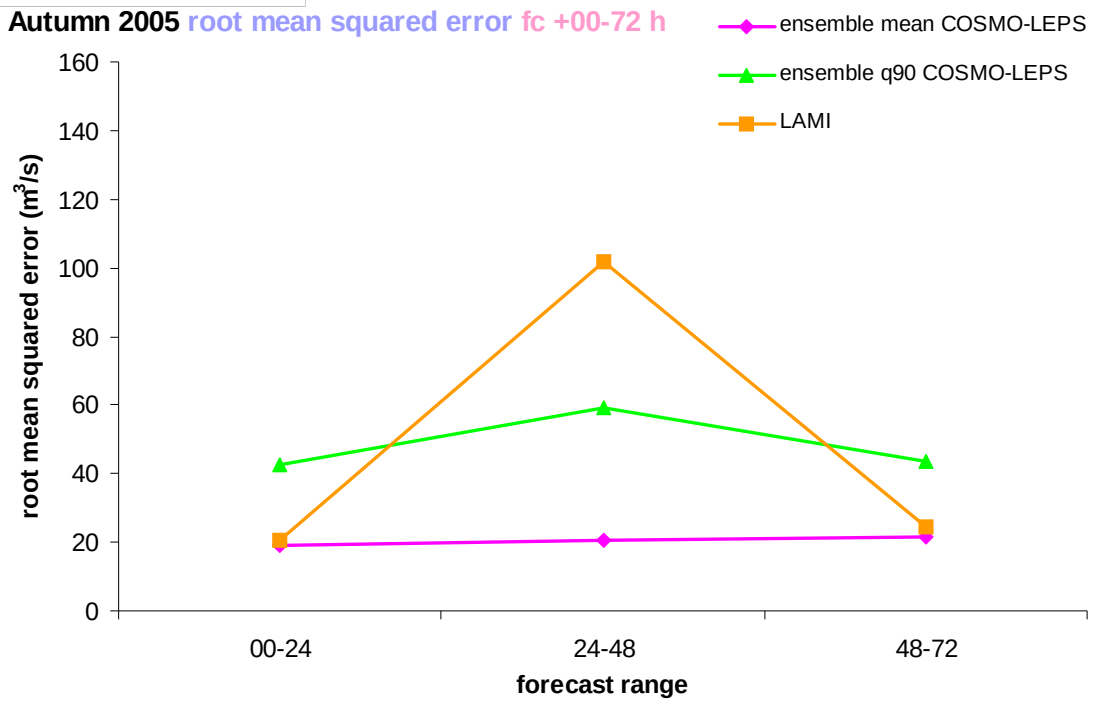
Preliminary results

probabilistic chain

vs

deterministic chain

autumn 2005



- For every forecast range, the ensemble mean performs better than the 90% quantile.
- The statistics related to the discharge forecast driven by COSMO-LAMI for the autumn 2005 are strongly influenced by one very high streamflow overestimate occurred within the time range +24-48 h.

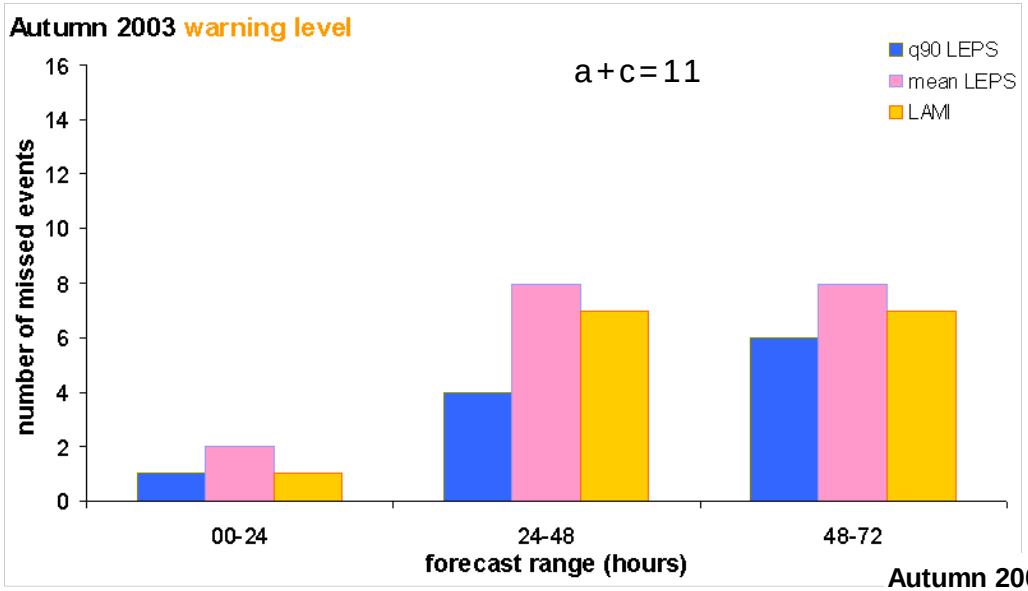
Preliminary results

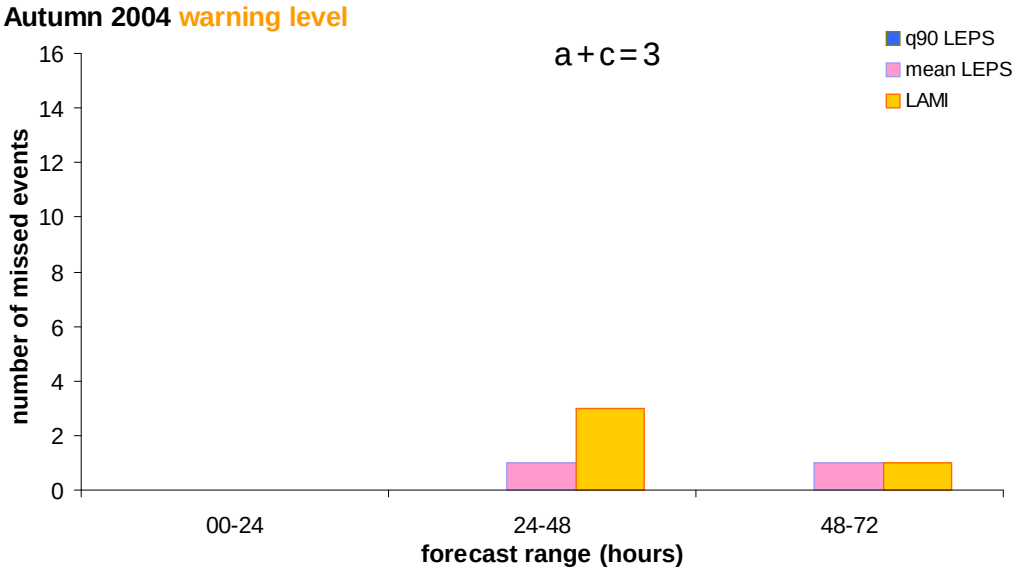
Potential use for early warnings and alarms

warning threshold

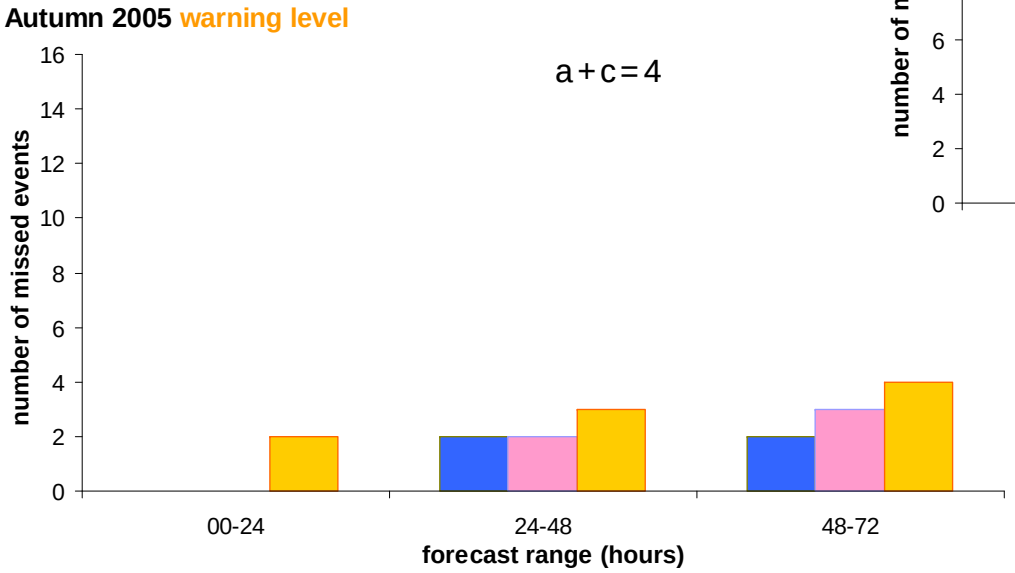
misses

forecast \ observed	yes	no
yes	a	b
no	c	d





for both systems, missed events and false alarms increase for longer lead-times



if the 90% quantile is considered, the missed events decrease with respect to those obtained with the ensemble mean or the deterministic forecast, but ...

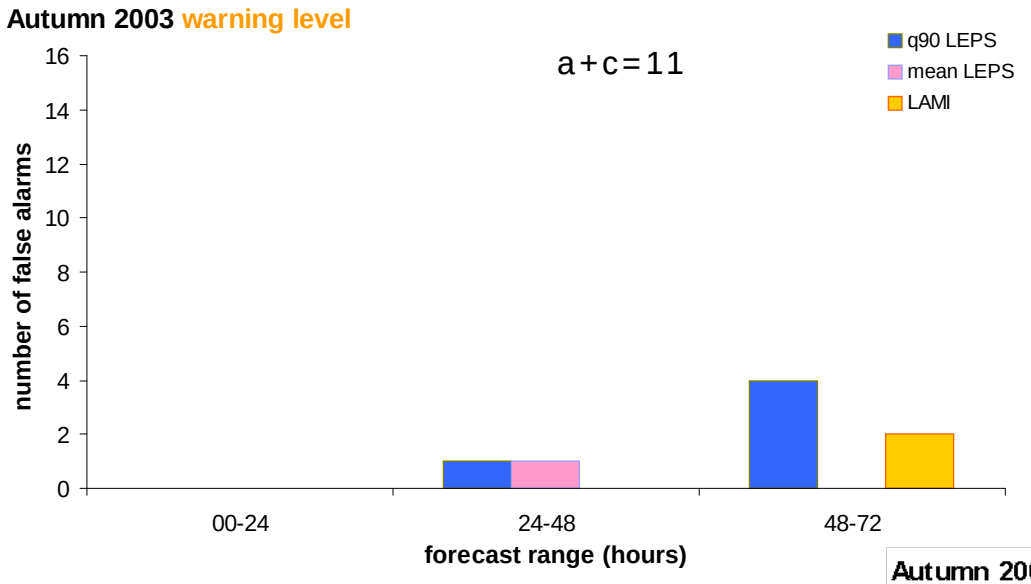
Preliminary results

Potential use for early warnings and alarms

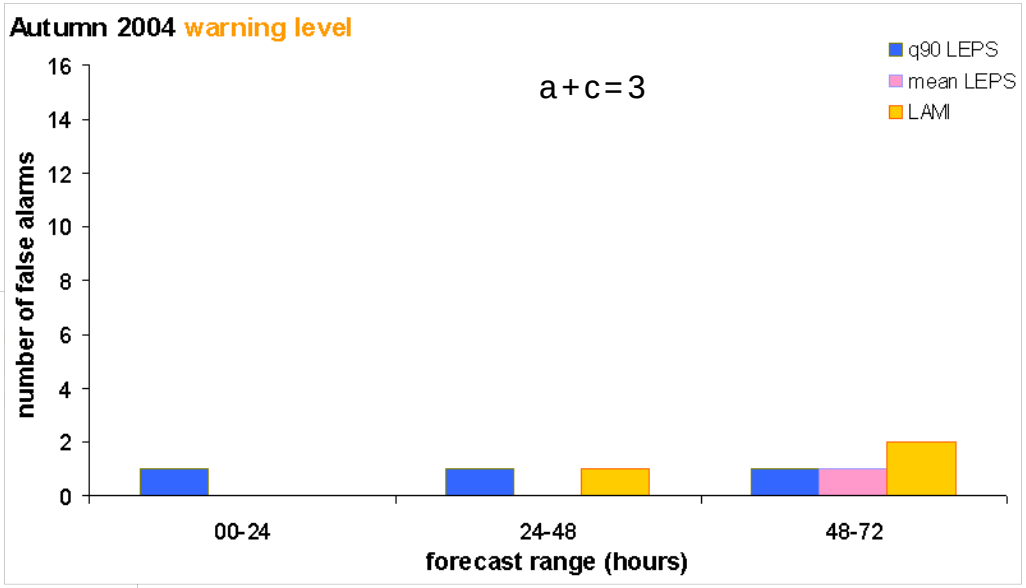
warning threshold

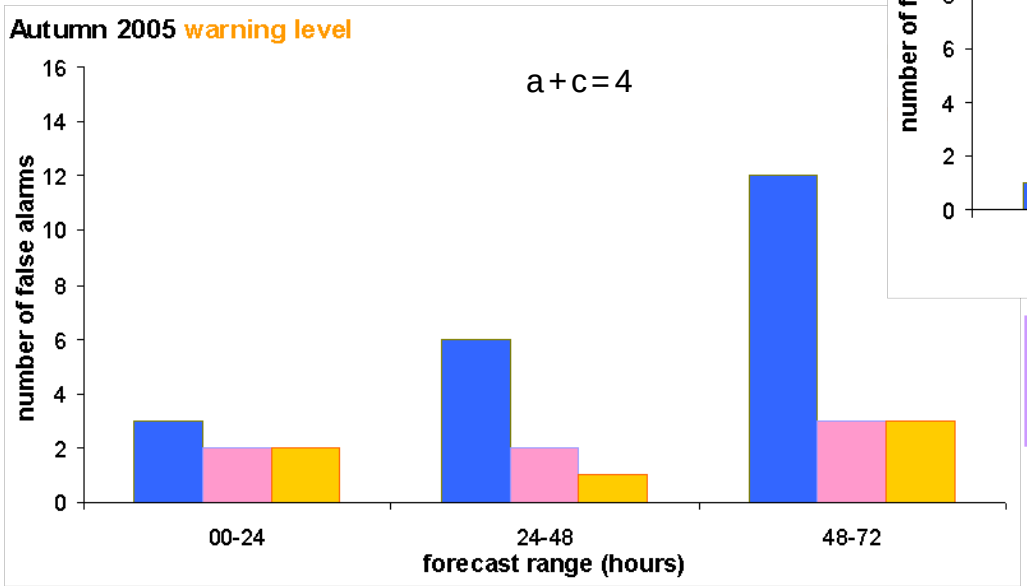
false alarms

forecast \ observed	yes	no
yes	a	**b**
no	c	d



for both systems, missed events and false alarms increase for longer lead-times





… false alarms increase, even if not considerably (except for autumn 2005).

Preliminary results

Potential use for early warnings and alarms

alarm threshold

misses

both forecasting chains fail the forecast of the three events occurred (two in autumn 2003, one in autumn 2004) for every 24-h forecast range, up to the +72 h lead-time.

false alarms

autumn seasons 2003 and 2004: both systems do not provide false alarms for every 24-h forecast range, up to the +72 h lead- time.

autumn 2005: no false alarms would have been issued by the probabilistic forecasting chain (regardless of the statistical measures used to represent the ensemble forecast), whereas one false alarm would have been issued by the deterministic chain (for the forecast range +24-48 h)

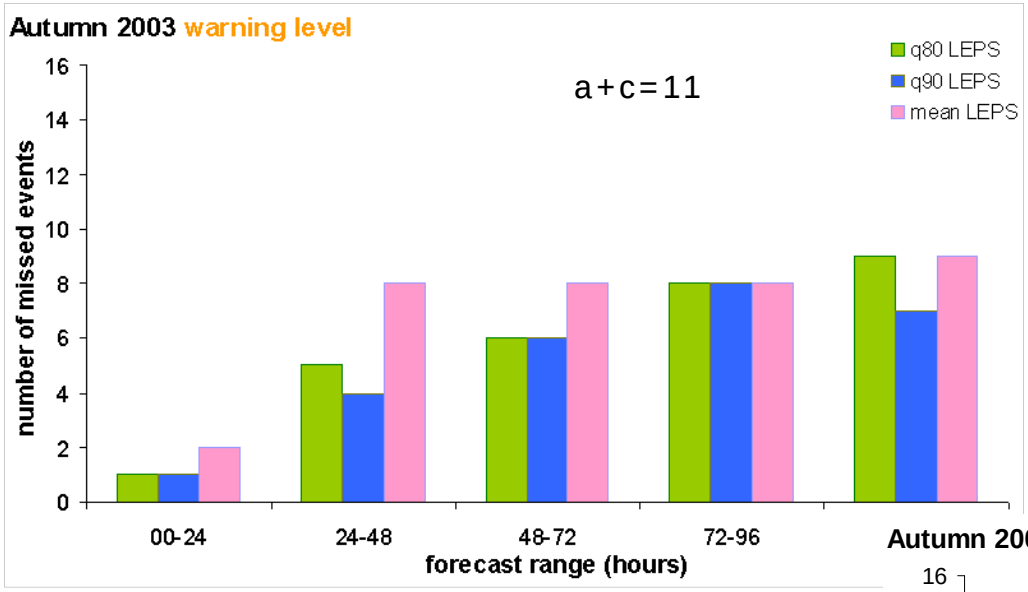
Preliminary results

Potential use for early warnings and alarms

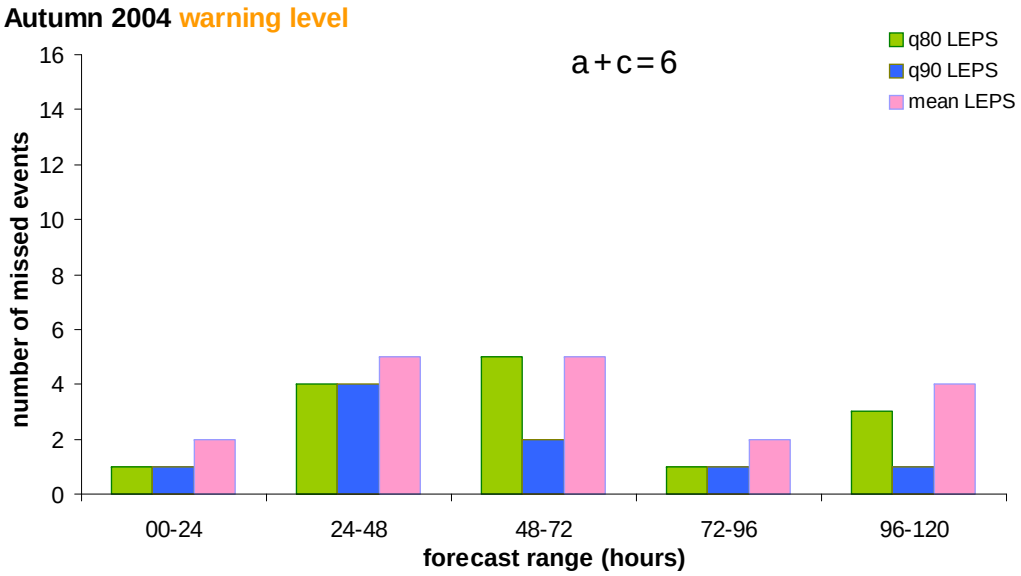
warning threshold

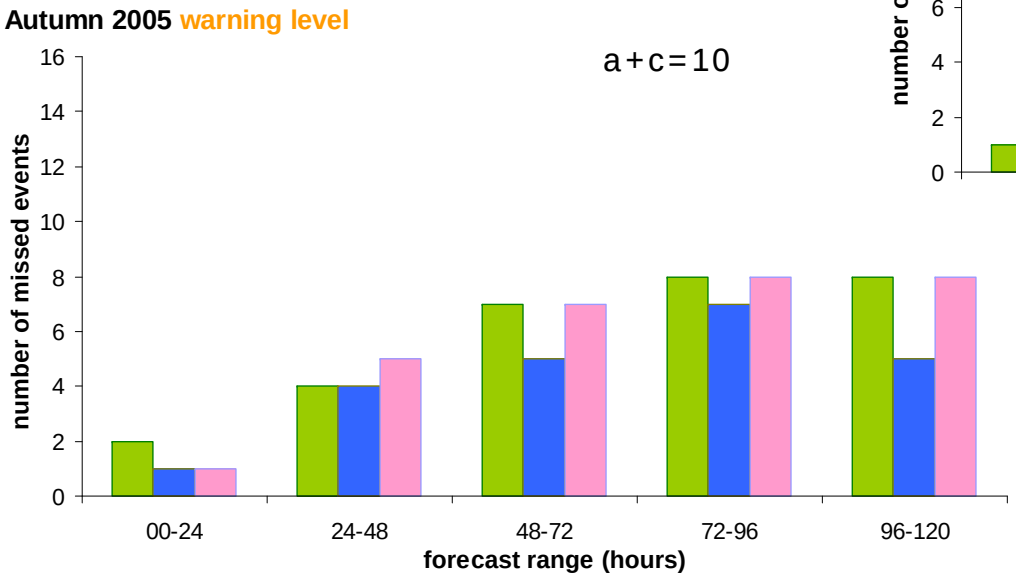
misses

forecast \ observed	yes	no
yes	a	b
no	c	d



Starting from the +48-72 h forecast range, the performance decay is attenuated...





... and the 90% quantile performs slightly better than the ensemble mean or the 80% quantile

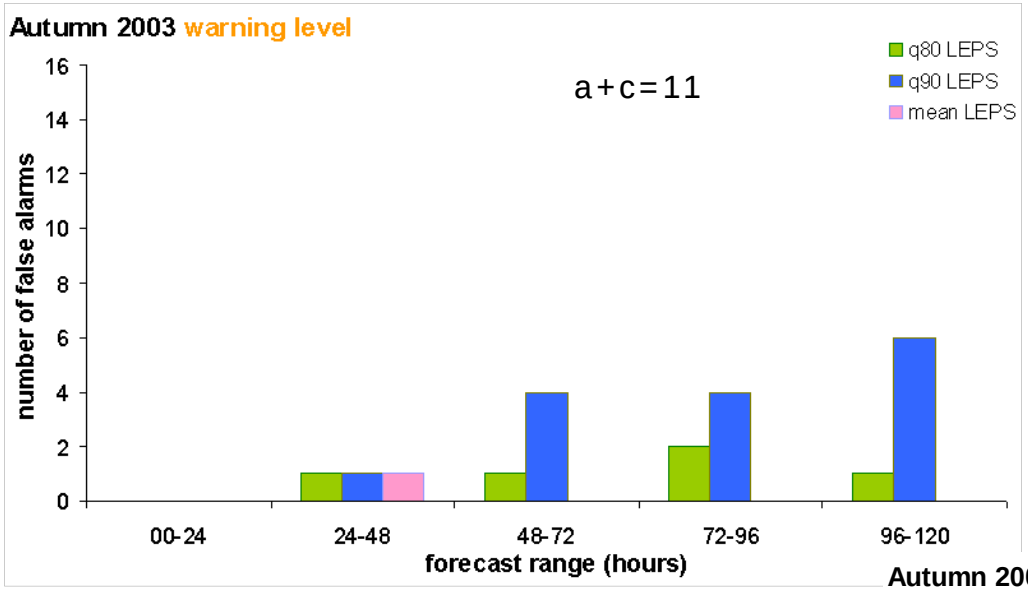
Preliminary results

Potential use for early warnings and alarms

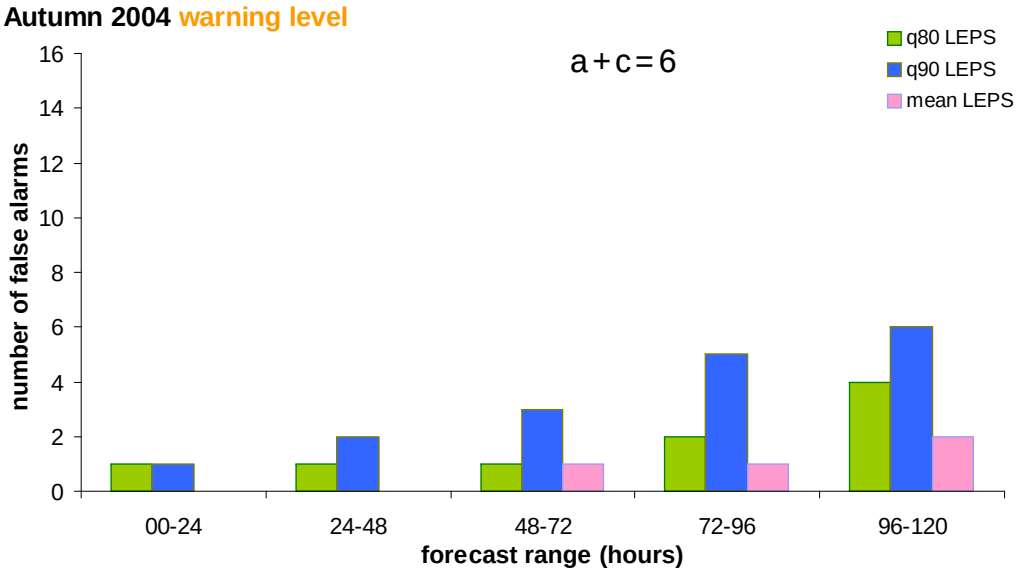
warning threshold

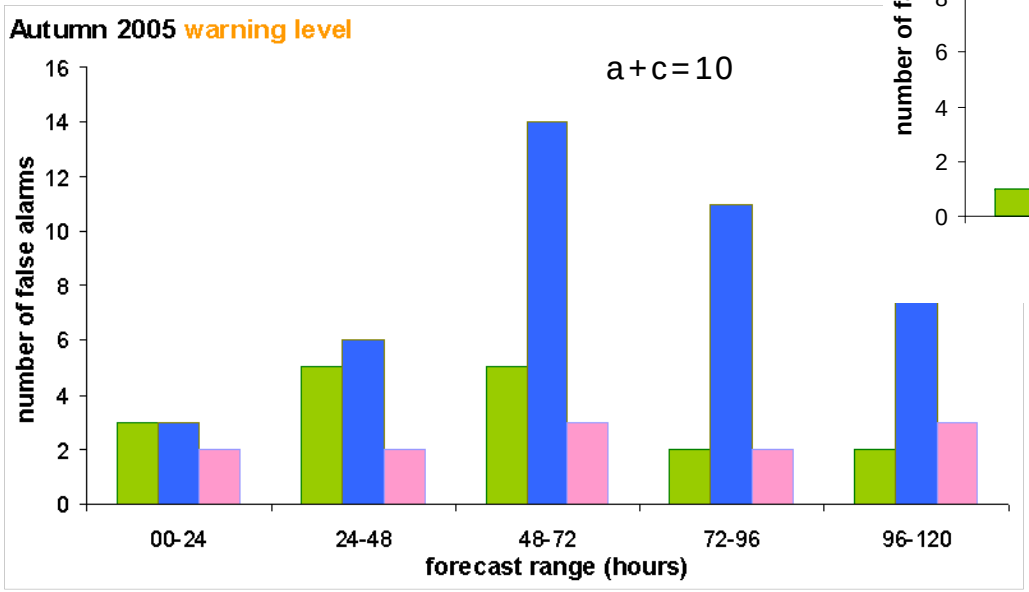
false alarms

forecast \ observed	yes	no
yes	a	**b**
no	c	d



if the 90% quantile is considered, false alarms tend to increase with the lead-time ...





... even if after the +48 h forecast range this trend does not outgrow excessively (except for autumn 2004).
This trend is not evident if the ensemble mean or the 80% quantile is considered.

Preliminary results

Potential use for early warnings and alarms

alarm threshold

misses

the forecast of the three events occurred (two in autumn 2003, one in autumn 2004) is failed by the ensemble mean, 80% and 90% quantile for every 24-h forecast range, up to the +120 h lead-time.

false alarms

autumn seasons 2003, 2004 and 2005: no false alarms would have been issued by the ensemble mean, the 80% and 90% quantile for every 24-h forecast range, up to the +120 h lead-time.

Conclusions and future developments

- The discharge predictions based on the COSMO-LEPS ensemble precipitation forecasts show performance which are comparable to the single-valued forecast driven by COSMO-LAMI for the first 24-hour forecast range. The added value of the probabilistic system comes out with increasing lead-times, especially from the +48-72 forecast range.
- COSMO-LEPS turns out to be a promising forecasting tool to drive hydrological predictions: the coupled system provides appropriate forecast guidance for early warnings of flood event on the upper Reno river basin.
- Which probability should be assigned to each COSMO-LEPS member? The debate in the scientific community is still open; this is even more true if these members are used as input to a hydrological model. It is planned to evaluate the impact of weighting each member differently (by tagging each member with a probability measure representing the relative size of the cluster resulting from the cluster analysis of ECMWF EPS).

- future works:

- to determine which confidence interval is more adequate to convey the forecast for operational purposes/applications and to support end-users in their decision-making processes

- objective criterion to evaluate the ensemble spread and the outliers percentage (Talagrand diagram)

- to evaluate the system performance in the remaining seasons

- comparison with other meteorological probabilistic systems (ECMWF EPS? analogues? MAP D-PHASE campaign?)