Conditional Exceedance Probabilities and the Prediction of Extreme Events
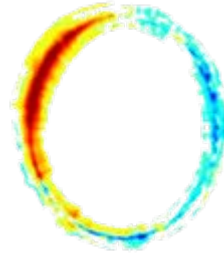
Simon J. Mason[1], Andreas Weigel[2], Michael K. Tippett[1], Anna Ghelli[3]

simon@iri.columbia.edu

1. International Research Institute for Climate and Society
The Earth Institute of Columbia University
2. Federal Office of Meteorology and Climatology (MeteoSwiss)
3. European Centre for Medium-Range Weather Forecasts

Eleventh Workshop on Meteorological Operational Systems

ECMWF, Shinfield Park, England, 12–16 November, 2007

Can we forecast this kind of extreme event?

In advance, preferably.

Data example

- *Variable*: weekly precipitation totals.

- *Observations*: Swiss network of automatic meteorological stations (ANETZ) – Zurich, Switzerland.

- *Forecasts*: ECMWF monthly ensemble prediction system at lead-times of 1 to 26 days.

- *Verification period*: 09-2005 until 09-2007

- *Climatology*: estimated from 12 years of re-forecasts and observations

Model Skill

The seasonal signal is removed from the time-series of measured and simulated precipitation
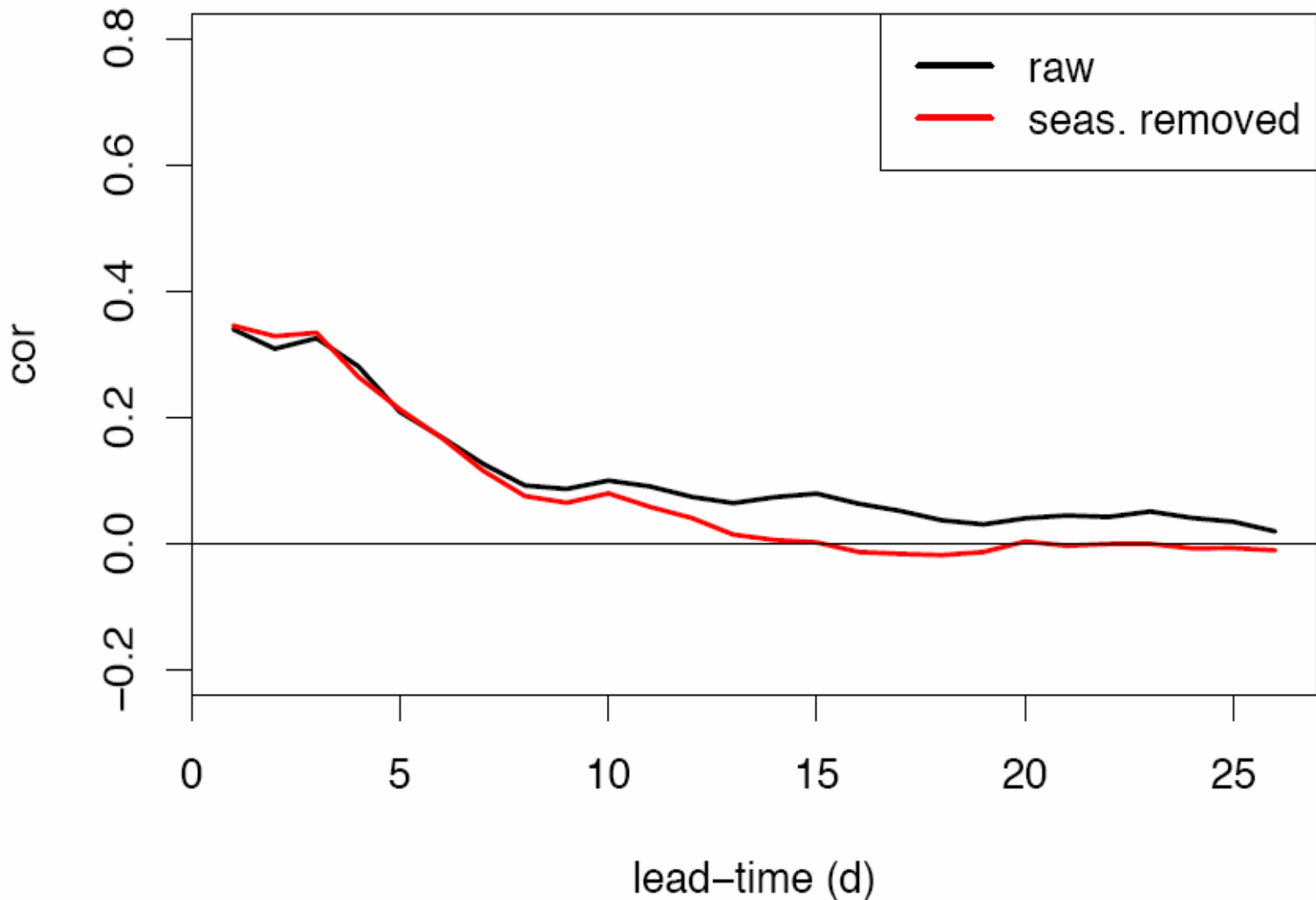
Introduction

- An EPS is designed to:

  - improve the skill of the best-guess forecast;

  - indicate the uncertainty in the forecast.

- We therefore need to measure the

  - accuracy

  - reliability

  of the forecast.

- Unfortunately most skill metrics measure only the average accuracy and average reliability. Can we say anything about the expected accuracy and reliability of any specific forecast (especially of the current forecast)?
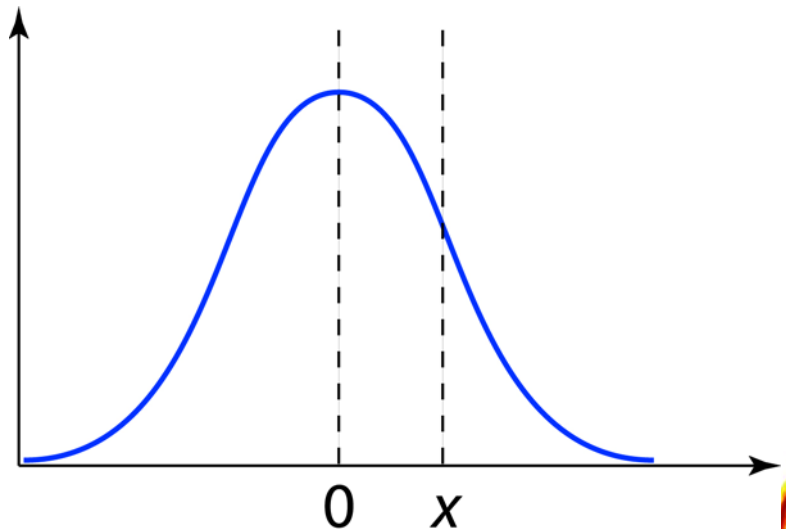
Accuracy

- Imagine a forecast, X, whose climatology is well-calibrated.

$$E(\varepsilon) = 0$$

- But if we know that the forecast has no skill then:

$$E(\varepsilon \mid X = x) = x$$

- The forecast should be recalibrated so that we forecast climatology.

Reliability

A consistency between the a priori stated probabilities of an event and the a posteriori observed relative frequencies of this event.

Murphy, A. H., 1973: A new vector partition of the probability score. *Journal of Applied Meteorology*, **12**, 595–600.

Murphy, A. H., 1993: What is a good forecast? An essay on the nature of goodness in weather forecasting. *Weather and Forecasting*, **8**, 281–293.

If the proportion of times that the event occurs is the same as the prior stated probability for all values of the prior probability, the system is reliable (or well calibrated).

This definition will prove to be inadequate …

Mason, S. J., J. S. Galpin, L. Goddard, N. E. Graham, and B. Rajaratnam, 2007: Conditional exceedance probabilities. *Monthly Weather Review*, **135**, 363–372.

Reliability

The climatological median defines the amount of rainfall that will be exceeded in 50% of cases.

$$P(v > x) = 0.5 \qquad P(v > x) = c$$

rainfall

If the forecast is completely reliable, the probability of exceeding the forecast median is 50%.

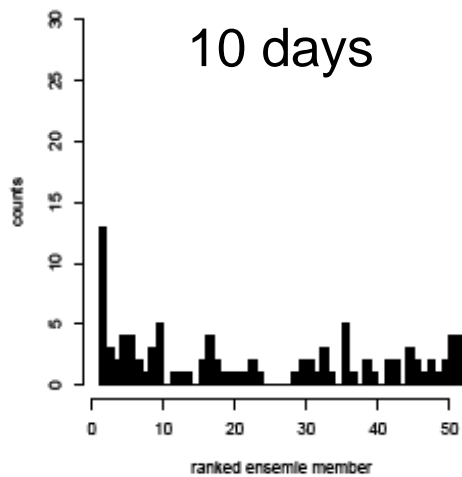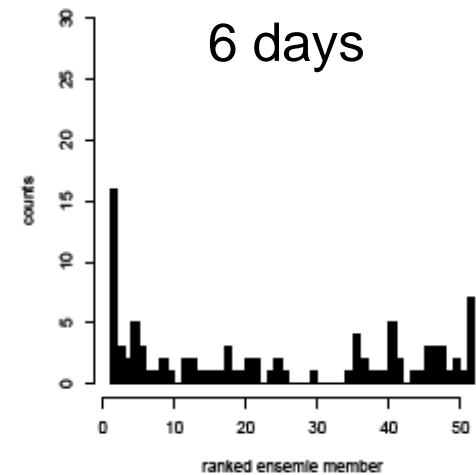If the forecast is completely useless, the probability of exceeding the forecast median is the climatological probability.
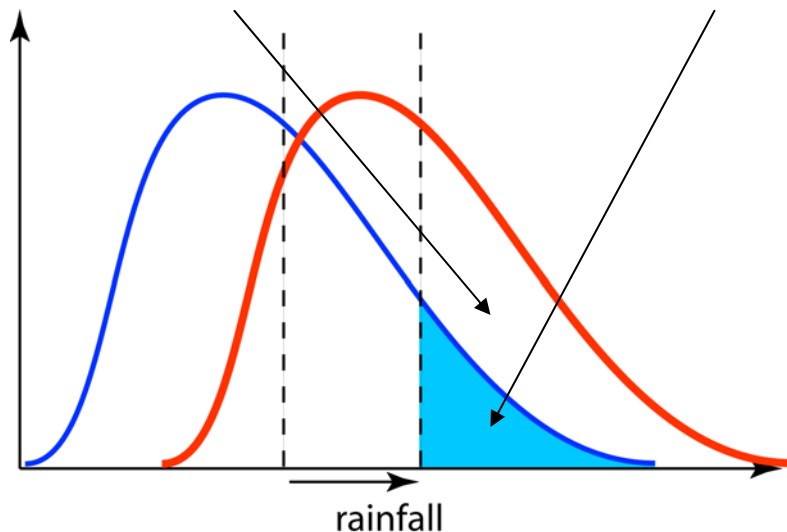
A forecast may indicate an increase in the probability of exceeding the median: maybe to 75%.

Measuring Reliability

A standard way to measure the reliability of the EPS is with *Talagrand Diagrams* or *Ranked Histograms*:
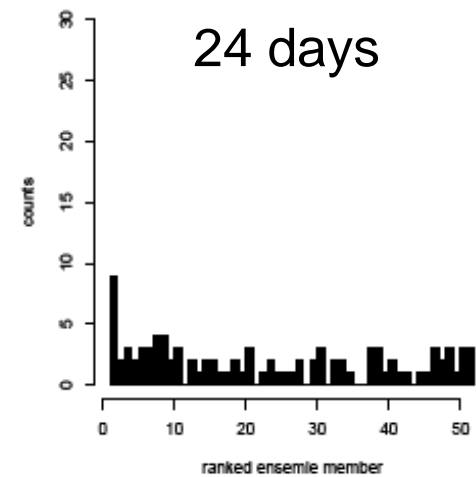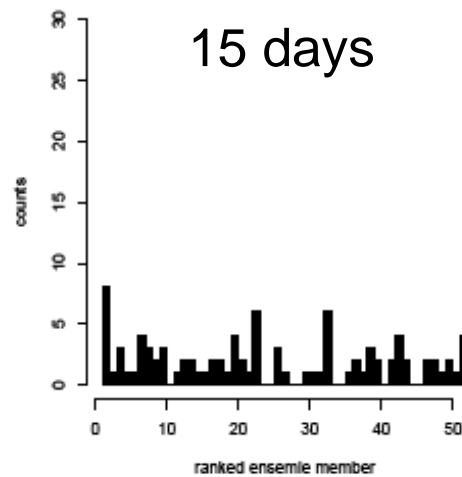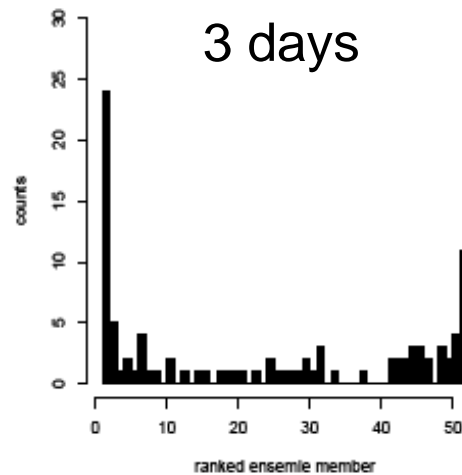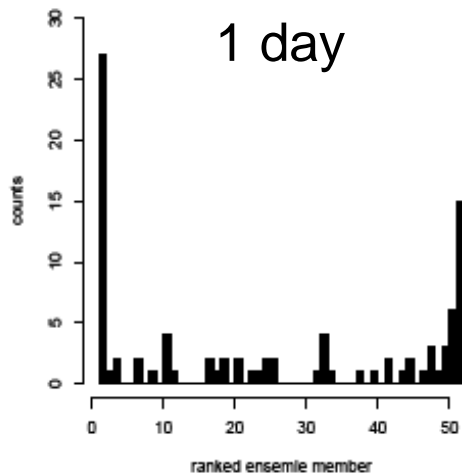
Measuring Reliability

Problems:

It is possible to design a useless system that has good reliability:

- persistent climatological distribution;

- random forecasts with "correct" ensemble spread.

For example, assuming only one ensemble member, if

$$E(x) = \text{CDF}^{-1}[0.5] \qquad \text{assuming symmetry}$$

$$P(v > E(x)) = 0.5$$

Measuring Reliability

If the system contains no useful information, the probability of being in the bottom bin is equal to the climatological probability of being less than the bin threshold.

The probability of this bin increases as the ensemble distribution shifts towards wet.

The probability of this bin increases as the ensemble distribution shifts towards dry.
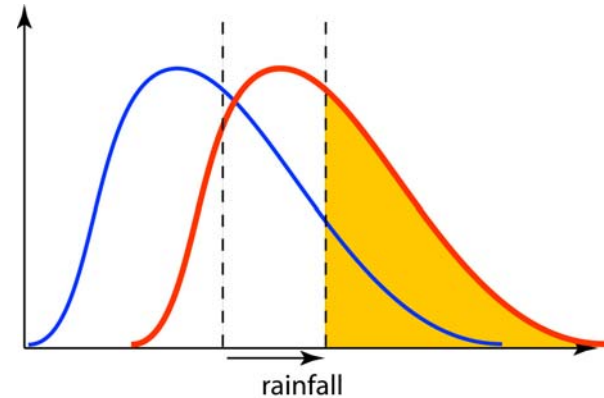
6 days
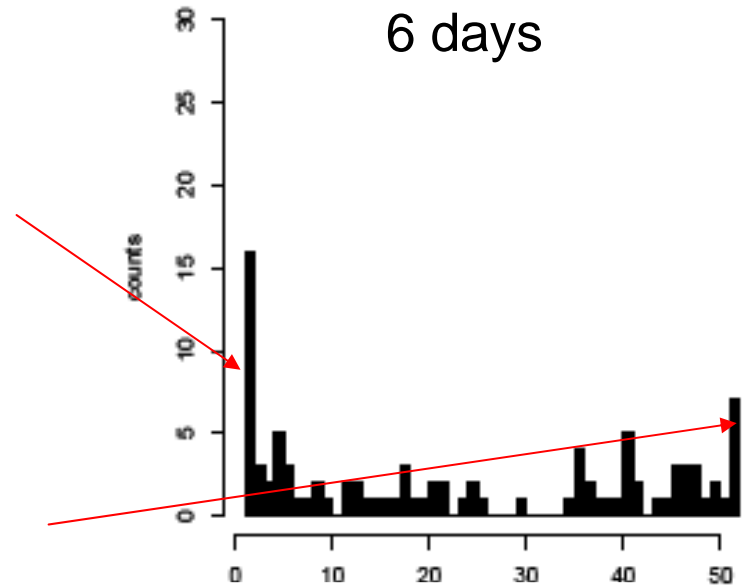
Measuring Reliability

But we want

$$P(v > x) = 0.5$$

rainfall

not

$$P(v > E(x)) = 0.5$$

rainfall
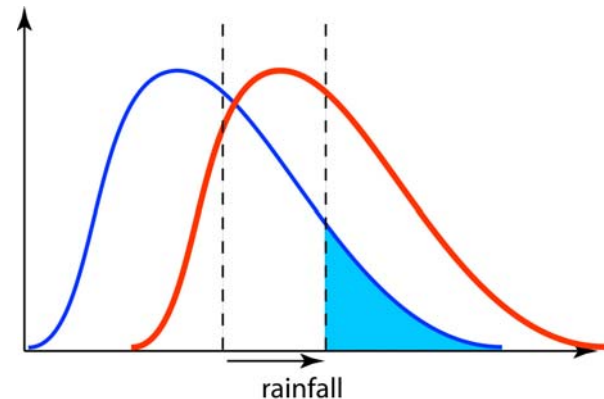
Specifically, if the forecast distribution is shifted far to the right, is there still a 50% chance of the observed value exceeding the median?

Conditional Exceedance Probabilities

The probability that the observed value exceeds the ensemble median depends upon the skill of the model.

If the model is perfect, this probability will be constant (50%). If it is imperfect, it will depend on the ensemble median.

Identify whether the exceedance probability is conditional upon the forecast.

The conditional exceedance probability (CEP) defines the probability that the observed value will exceed the forecast, conditioned upon the value of the forecast.

Conditional Exceedance Probabilities

Generalized linear models with binomial errors can be used:

$$P\left(v \geq X_{i:m}\,\middle|\,X_{im} = x\right) = 1 - \exp\left(-\exp\left(\beta_0 + \beta_1 x\right)\right)$$

β_0 measures unconditional bias, and

β_1 measures conditional bias

"Completely" reliable (β_1=0)   Unconditional bias (β_1=0)
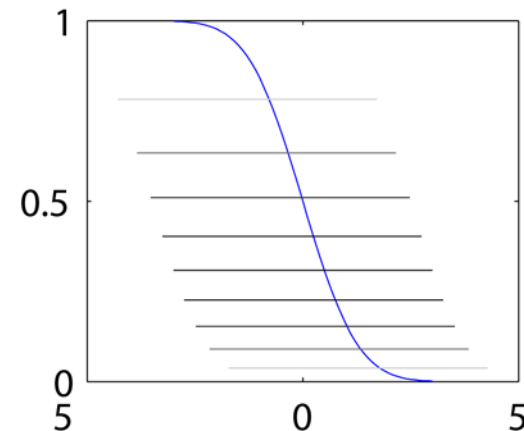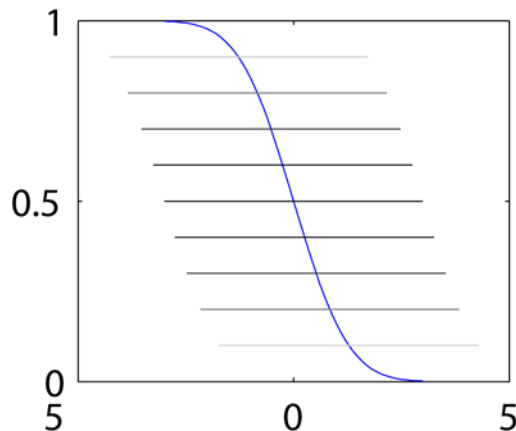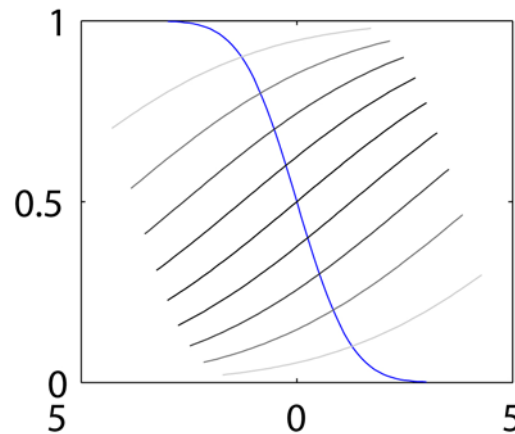
Conditional Exceedance Probabilities
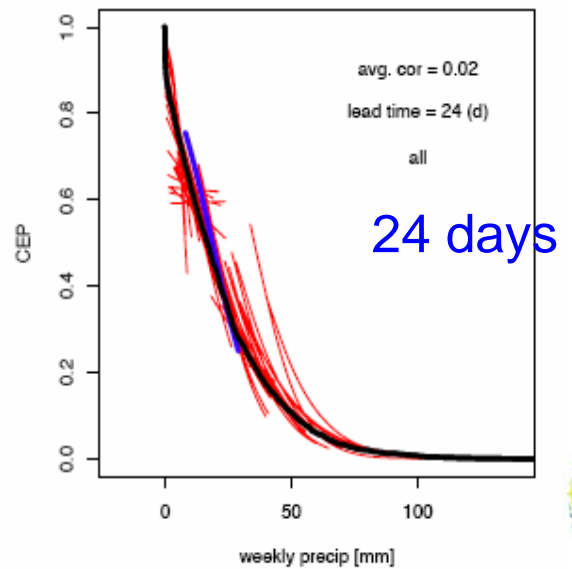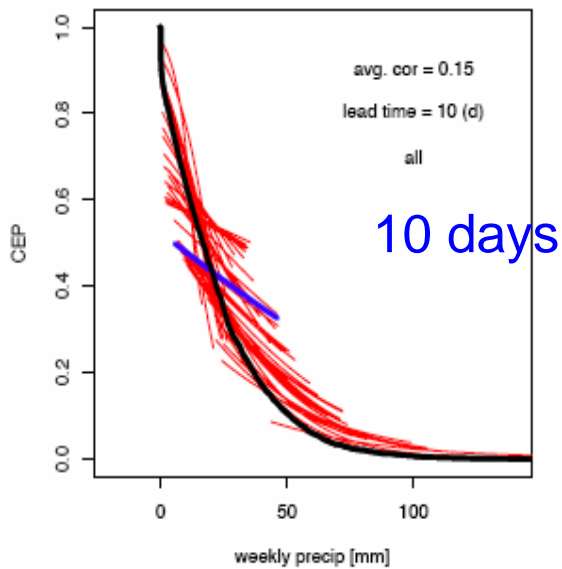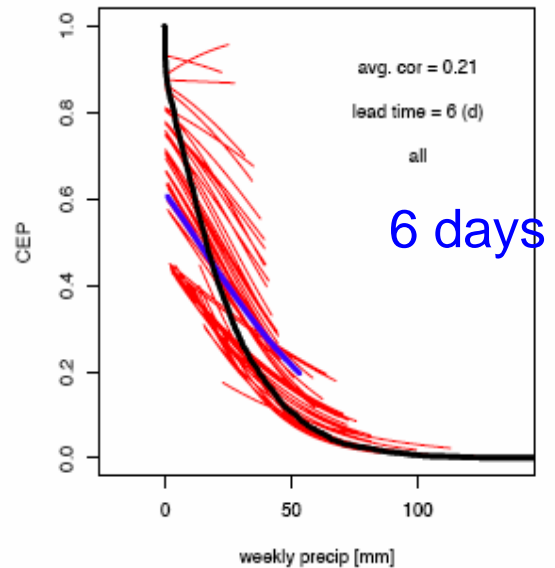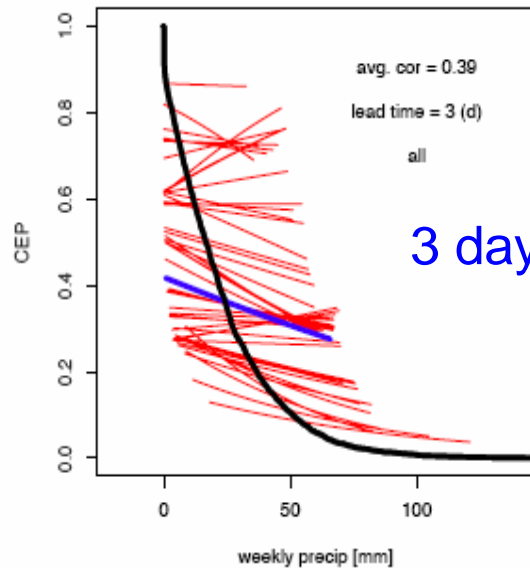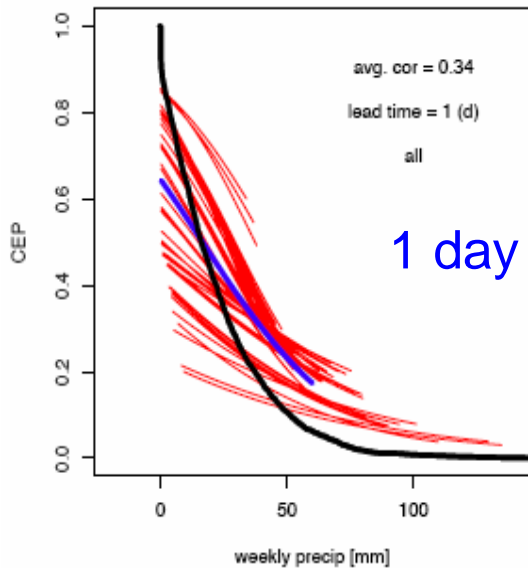
Conditionally biased ($\beta_1<0$)
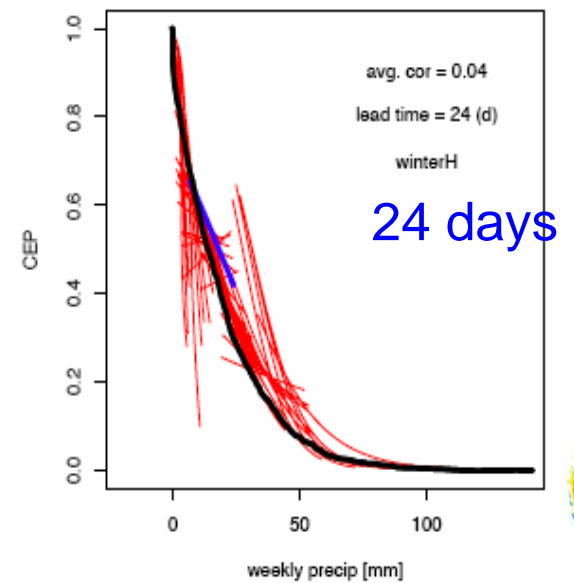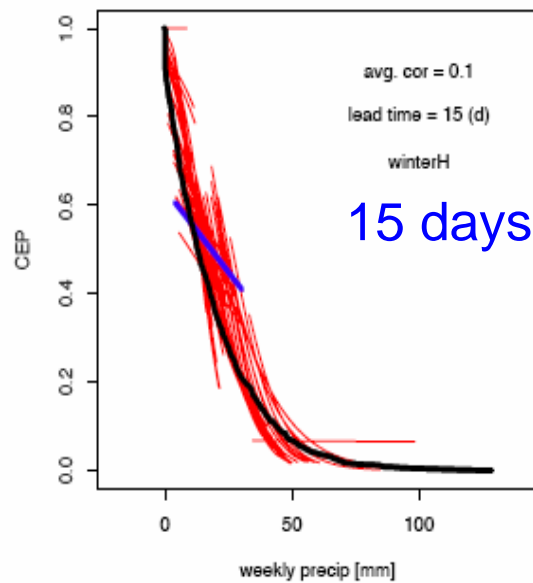Positive but weak skill
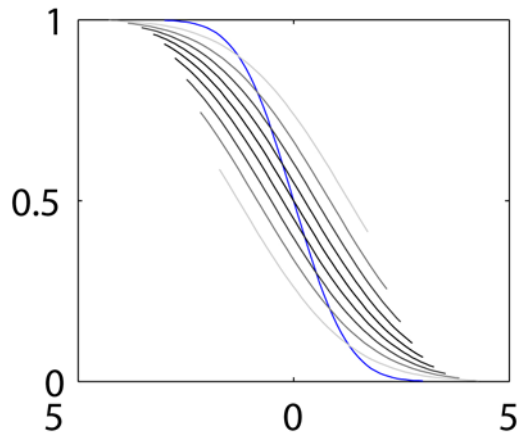
Conditionally biased ($\beta_1<0$)
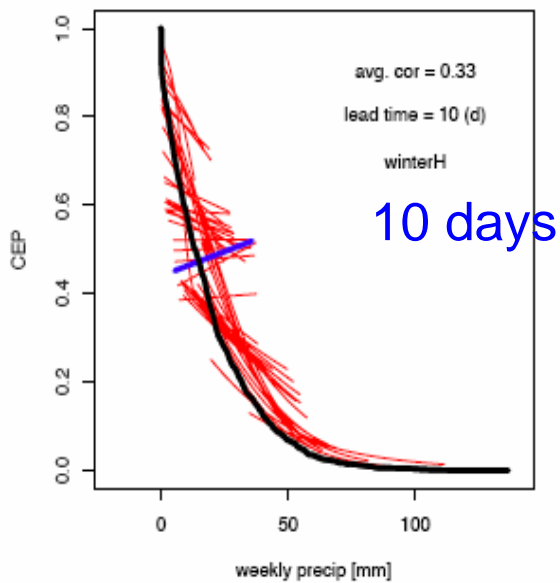Negative skill
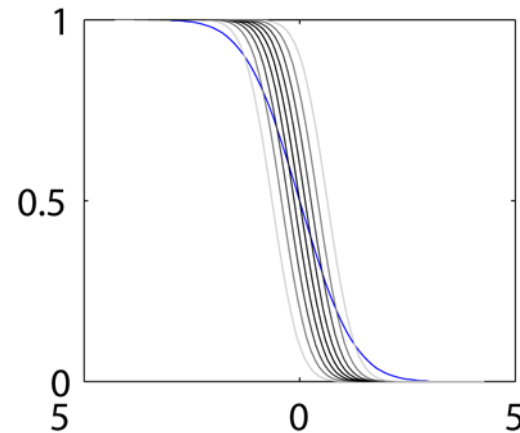
Conditionally biased ($\beta_1>0$)
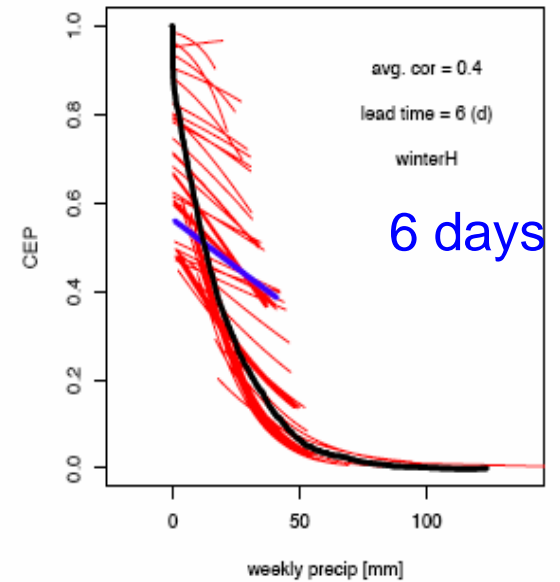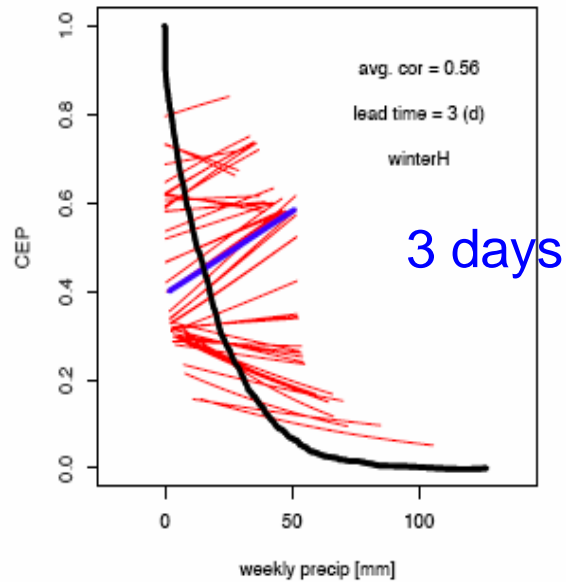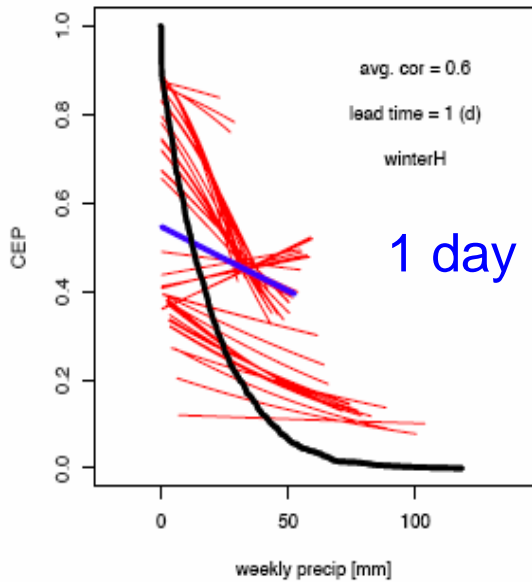Positive skill, weak signal

Conditional Exceedance Probabilities
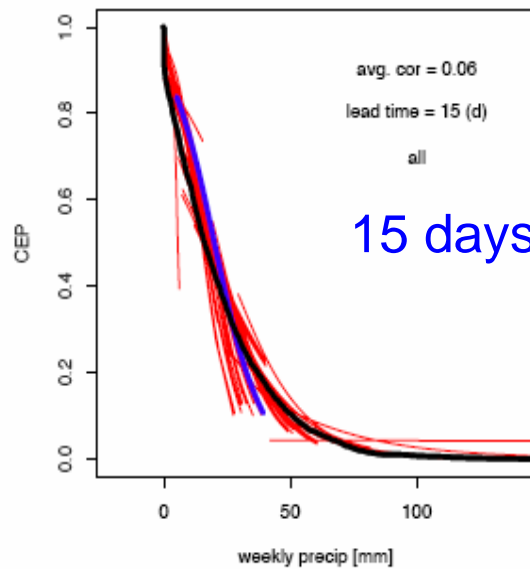
CEPs ("winter" only forecasts)

Conditional Exceedance Probabilities

What do "good" CEPs tell us?

That when the EPS forecasts an increase in the probability of very heavy rainfall, that increase in probability is reliable.

But that does not tell us whether conditions of very heavy rainfall are forecast reliably. Can we condition the exceedance probability on the observed rainfall?
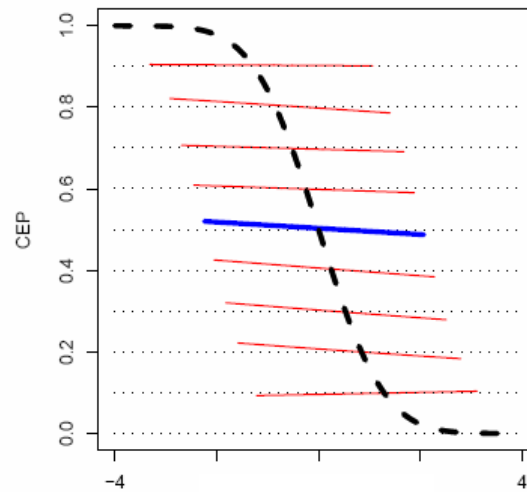
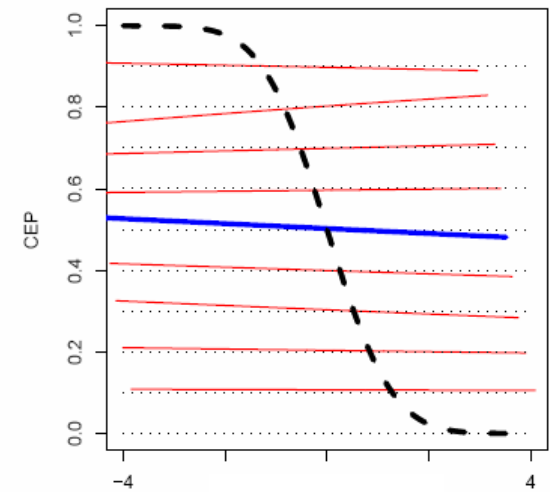# Conditional Exceedance Probabilities

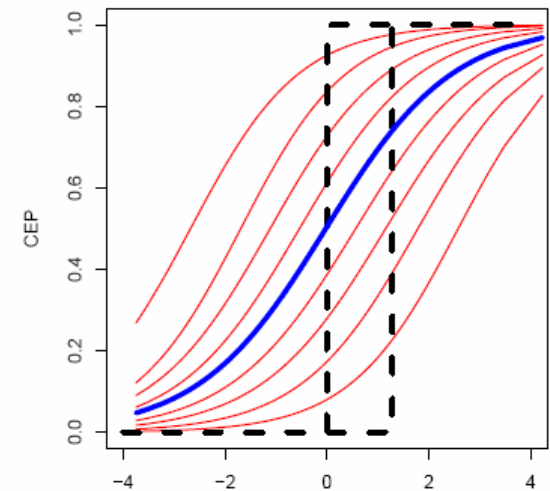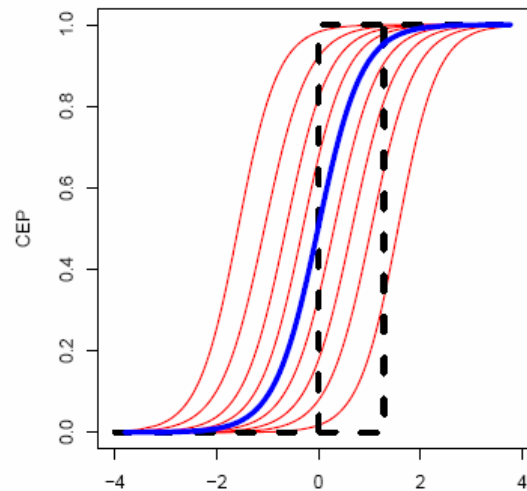CEPS | forecast (correct ensemble spread and variance):
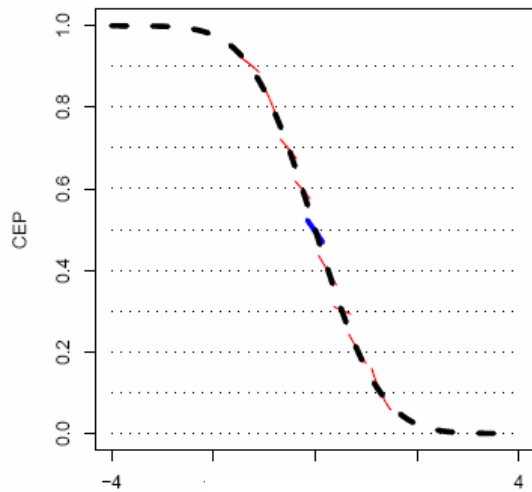
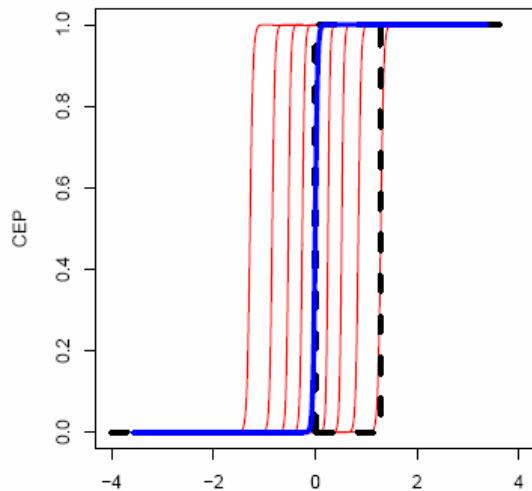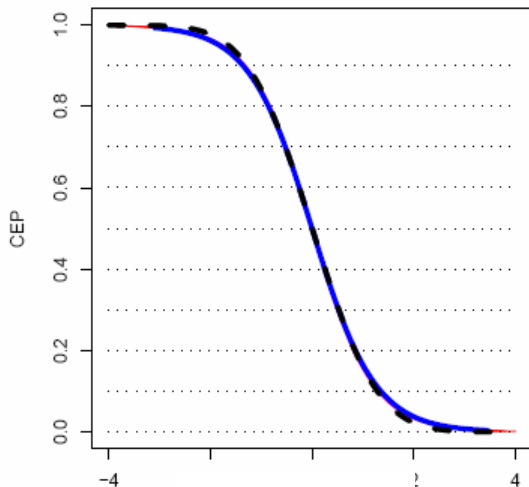r = 0.0   r = 0.3   r = 0.9
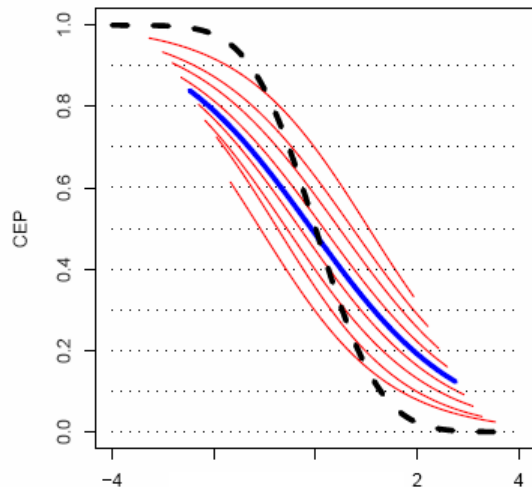
CEPS | observations:

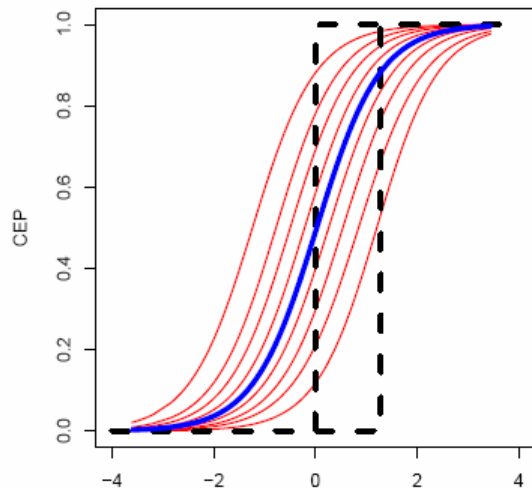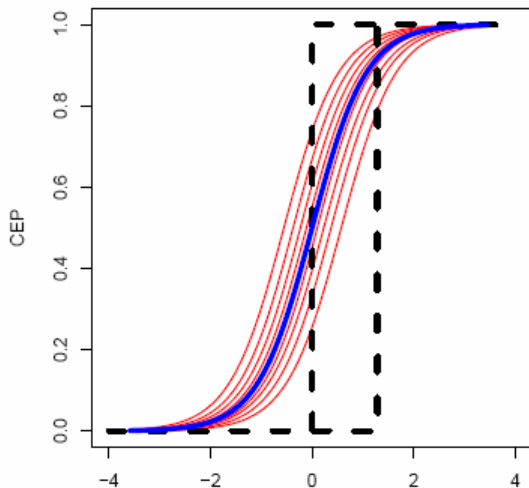# CEPs conditioned on observations

CEPS | forecast:

r = 0.0   r = 0.3

CEPS | observations:

By increasing the variance of the forecasts the CEPS conditioned on the observations look like the CEPs given completely reliable forecasts, but imperfect skill.

Conditional Exceedance Probabilities

So, like the ranked histograms:

Bad CEP | obs. indicates bad forecasts

Good CEP | obs. indicates good or bad forecasts

Because the forecasts are calibrated, given imperfect skill, the variance of a single ensemble member is less than of the observations. So an extreme observation is likely to exceed the forecast. The only ways to reduce the exceedane probability are to increase the skill, or increase the over-confidence.

But:

Bad CEP | for. indicates bad forecasts

Good CEP | for. indicates good forecasts

Conclusions

- THE ECMWF EPS provides reasonably reliable indications of increases in the probability of 7-day averaged heavy rainfall over Zurich.

- Most skill is close to the ensemble median.

- However, extremely heavy rainfall events are rainfall events that are in the far right tail of possibilities, and so if they are forecast reliably, they will be forecast with low (but increased) probability.

- Extreme events are extreme!

Conclusions

τάχ᾽ ἄν τις εἰκὸς αὐτὸ τοῦτ᾽ εἶναι λέγοι,
βροτοῖσι πολλὰ τυγχάνειν οὐκ εἰκότα.

Aristotle

which, like most things by Aristotle, translates into something that is just as incomprehensible in English as it is in Greek, so a rather liberal paraphrase is:

It is in the very nature of probability that improbable things will happen.