ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Probability forecast: quantifying uncertainty in forecasts

Marko Laine
Finnish Meteorological Institute
NMM31 2018-06-18 Reykjavík

Contents

- The concept of probability
- Probability forecasts
- Uncertainty in forecasts
- Predictability and chaos
- How to verify probabilities
- Calibrating ensemble forecasts

What is probability

- ”The probability that it rains tomorrow is 20%”
- **Classical interpretation** as long run frequencies. Relevant for simple, symmetric, repeatable (and deterministic) events, like a tossing of coin or gambling.
- Probability as a subjective measure of degree of belief, aka the **Bayesian interpretation**.
- When talking about a single future event, there is no direct frequentistic interpretation. In most cases, **we use probability to quantify uncertainty**.
- Weather and climate are complicated phenomena. We need the notions of **chaos** and **predictability**.
- Mathematically, probability is finite and additive measure, defined for a set of events. No philosophical disputes here.

Probability and statistical theory to quantify uncertainty

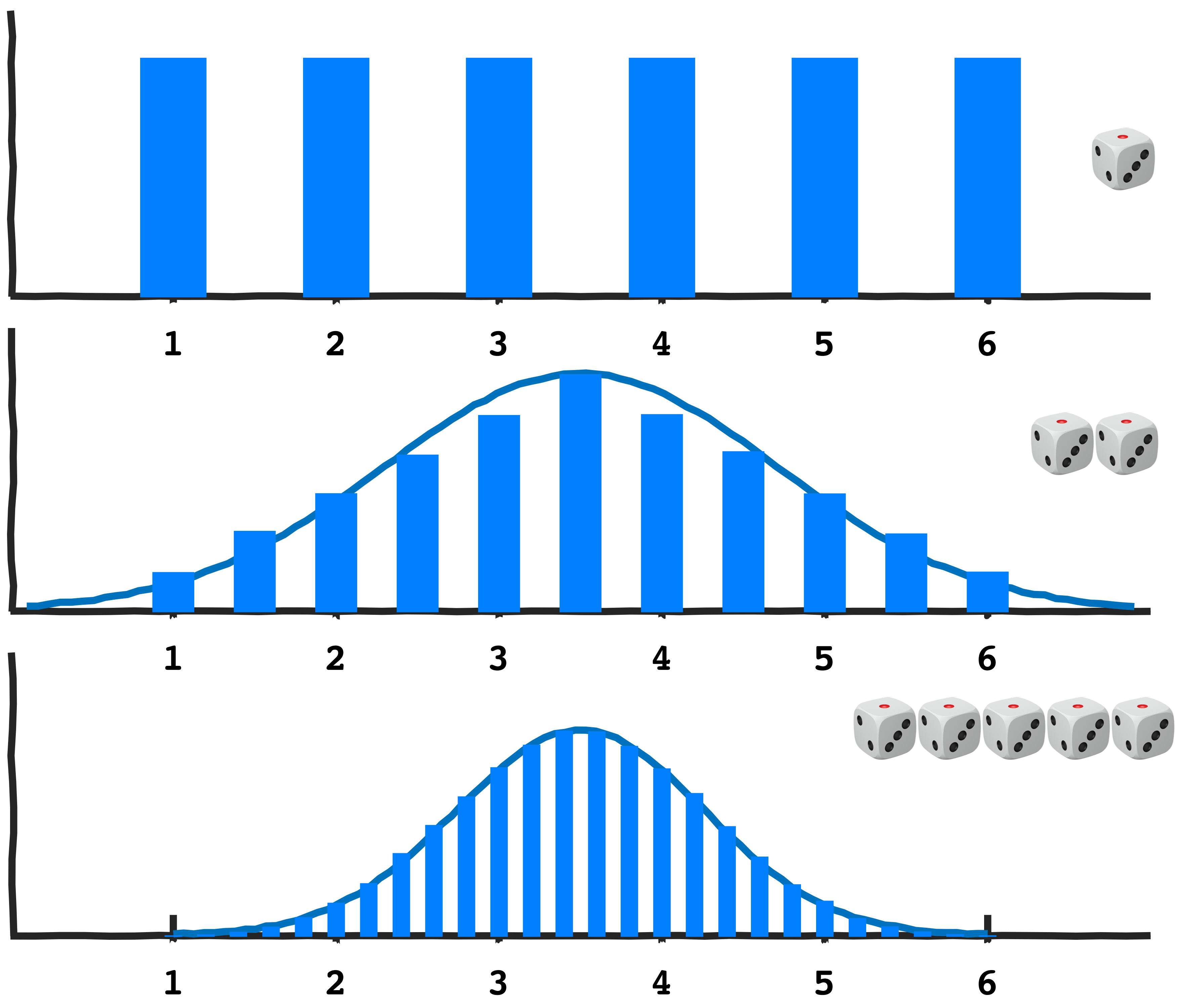
- Short history

- Origins in gambling theory. Probabilities for symmetric repeatable events, like throwing a dice (17. century, Fermat, Pascal).

- Statistical theory of distributions, central limit theorem (18. century, Gauss, Laplace).

- Mathematical statistics, statistical inference (20. century, Kolmogorov, Wald, Fisher).

Average when throwing dice



Probabilistic weather forecasts

”There is 20% probability for rain exceeding 10 mm, tomorrow between 8 – 12 AM, at Kumpula, Helsinki.”

- The meteorologist best opinion (but he/she might fear feedback for false negatives).
- Of 50 ENS forecast members, 20% had heavy rain (but ensemble system might not be well calibrated).
- Of 5 different deterministic models, 1 forecasted rain (but they all use the same observations).
- In October, it usually rains 20% of the days in Helsinki (no skill).

World Cup probabilities

≡ SECTIONS **FINANCIAL EXPRESS** READ TO LEAD हिन्दी

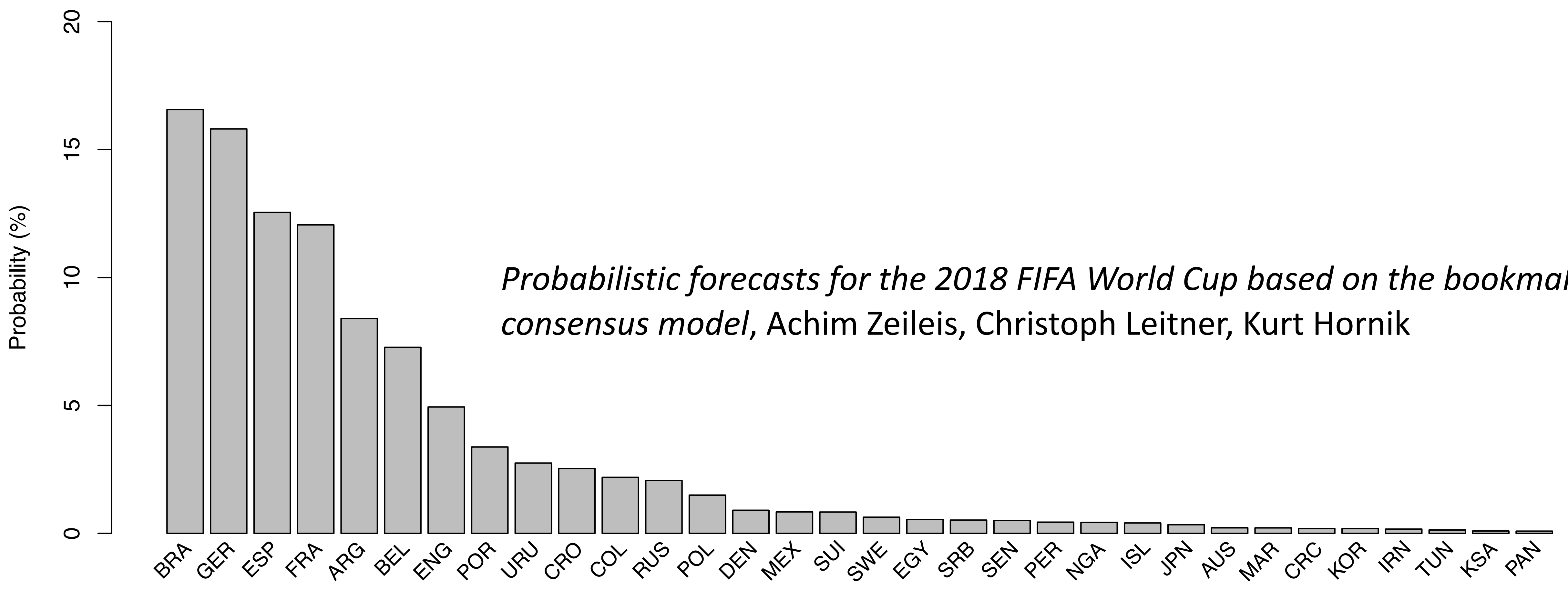
Home / Sports / FIFA World Cup 2018 prediction: Iceland have outside chance to win, find out who lifts trophy

FIFA World Cup 2018 prediction: Iceland have outside chance to win, find out who lifts trophy

FIFA World Cup 2018 prediction: With World Cup less than a month away, predictors have come to life to predict who will lift the trophy at Russia on July 15.

By: FE Online | New Delhi | Published: May 17, 2018 9:17 PM

8 SHARES SHARE



Probabilistic forecasts for the 2018 FIFA World Cup based on the bookmaker consensus model, Achim Zeileis, Christoph Leitner, Kurt Hornik

Figure 1: 2018 FIFA World Cup winning probabilities from the bookmaker consensus model.

Investing and football, UBS Chief Investment Office GWM Investment Research, May 2018

And the winner is…

Simulated likelihood of each team to advance through the tournament (in %)

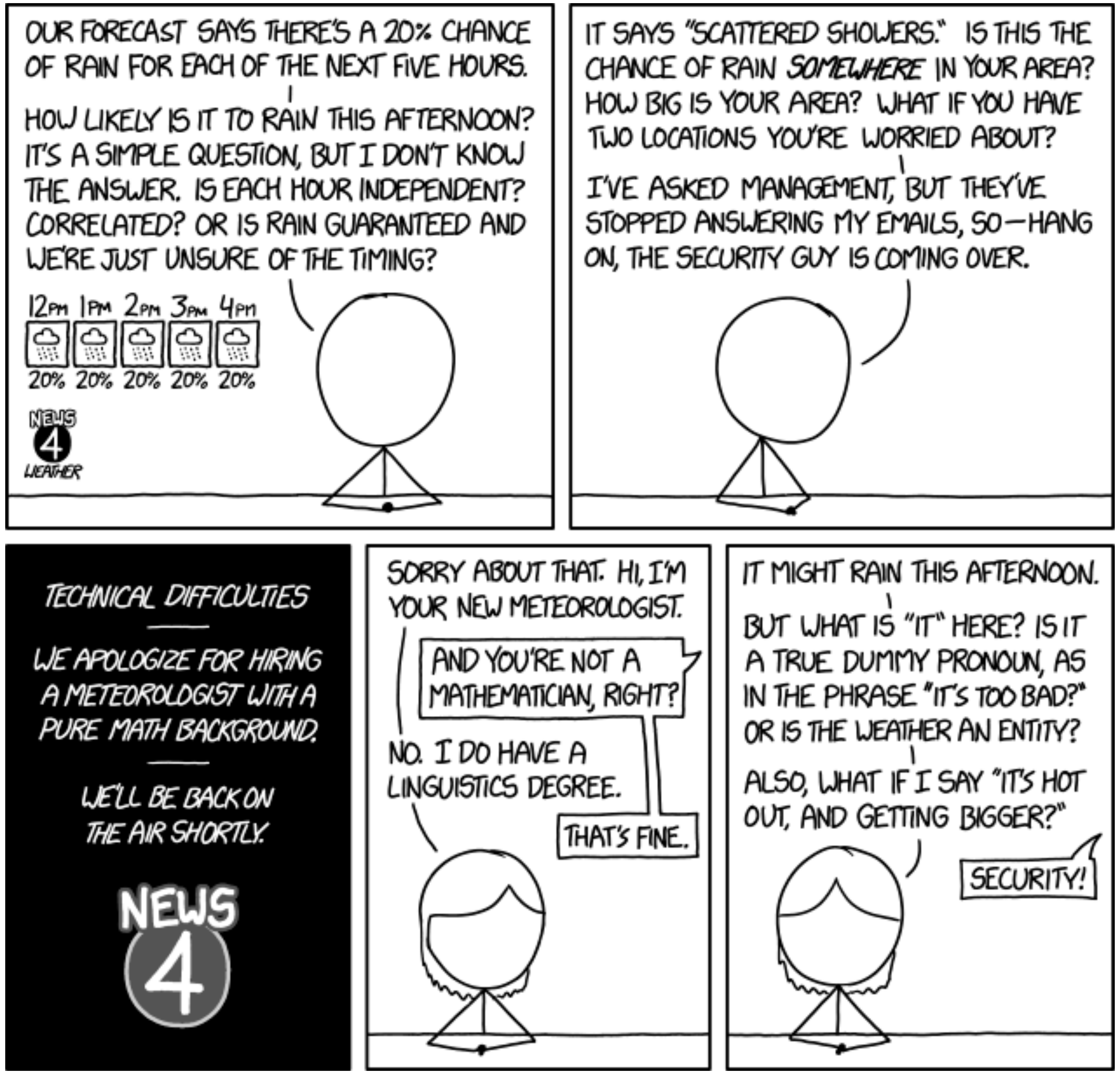
	Winner	Runner-Up	Semi-Finalist	Quarter-Finalist	Winner Group Stage	Second Group Stage
Germany	**24.0**	**36.7**	**51.3**	**66.7**	**68.6**	**22.0**
Brazil	**19.8**	**31.9**	**44.1**	**60.5**	**66.8**	**23.1**
Spain	**16.1**	**28.0**	**50.5**	**68.5**	**60.6**	**26.5**
England	8.5	18.7	31.4	66.2	53.7	33.6
France	7.3	16.1	35.1	59.5	60.1	24.6
Belgium	5.3	11.6	23.8	56.9	38.3	43.7
Argentina	4.9	11.3	26.9	51.8	54.7	26.4
Portugal	3.1	8.0	21.8	39.8	25.2	38.2
Uruguay	1.8	5.5	15.8	32.0	42.5	34.3
Switzerland	1.8	5.0	11.5	22.9	19.7	39.6
Mexico	1.8	5.3	10.9	22.5	17.2	36.6
Italy	1.6	4.4	10.1	19.4	15.3	31.0
Russia	1.6	4.6	14.4	30.5	41.4	33.6

How to interpret probability statements

- Probability forecast is tied to the estimated **probability distribution of event**. The distribution contains information on the likelihood of all possible events.
- The width of the distribution tells about the predictability.
- Easiest to interpret are single event probabilities.
- They need to be tied to time, location, duration and to a threshold.

- We can not combine probabilities without knowledge on dependence and correlation.

- P(A or B) = P(A) + P(B) - P(A and B)
- P(A and B) = P(A|B)P(B)



https://xkcd.com/1985/

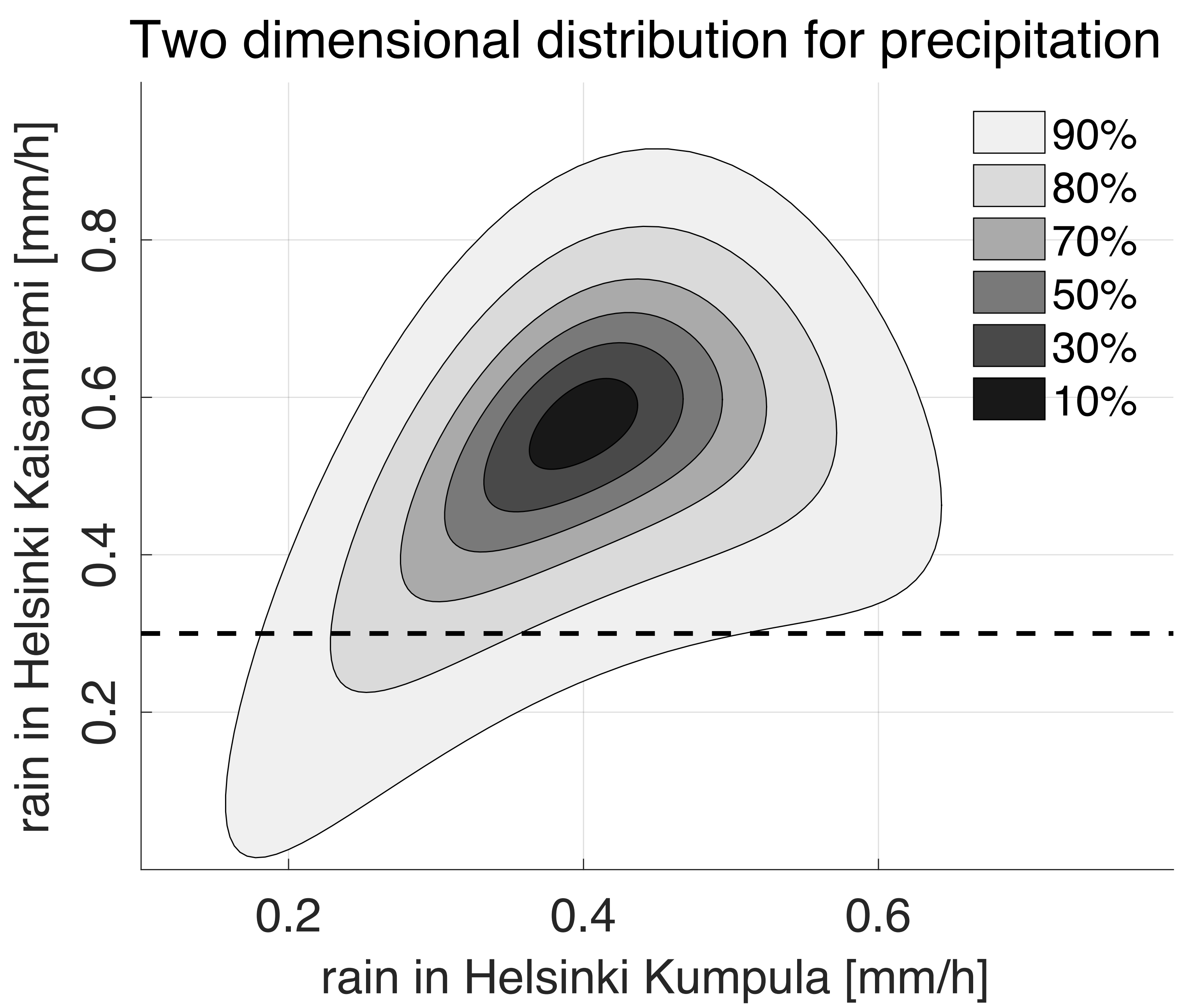
Going beyond single event probability forecasts

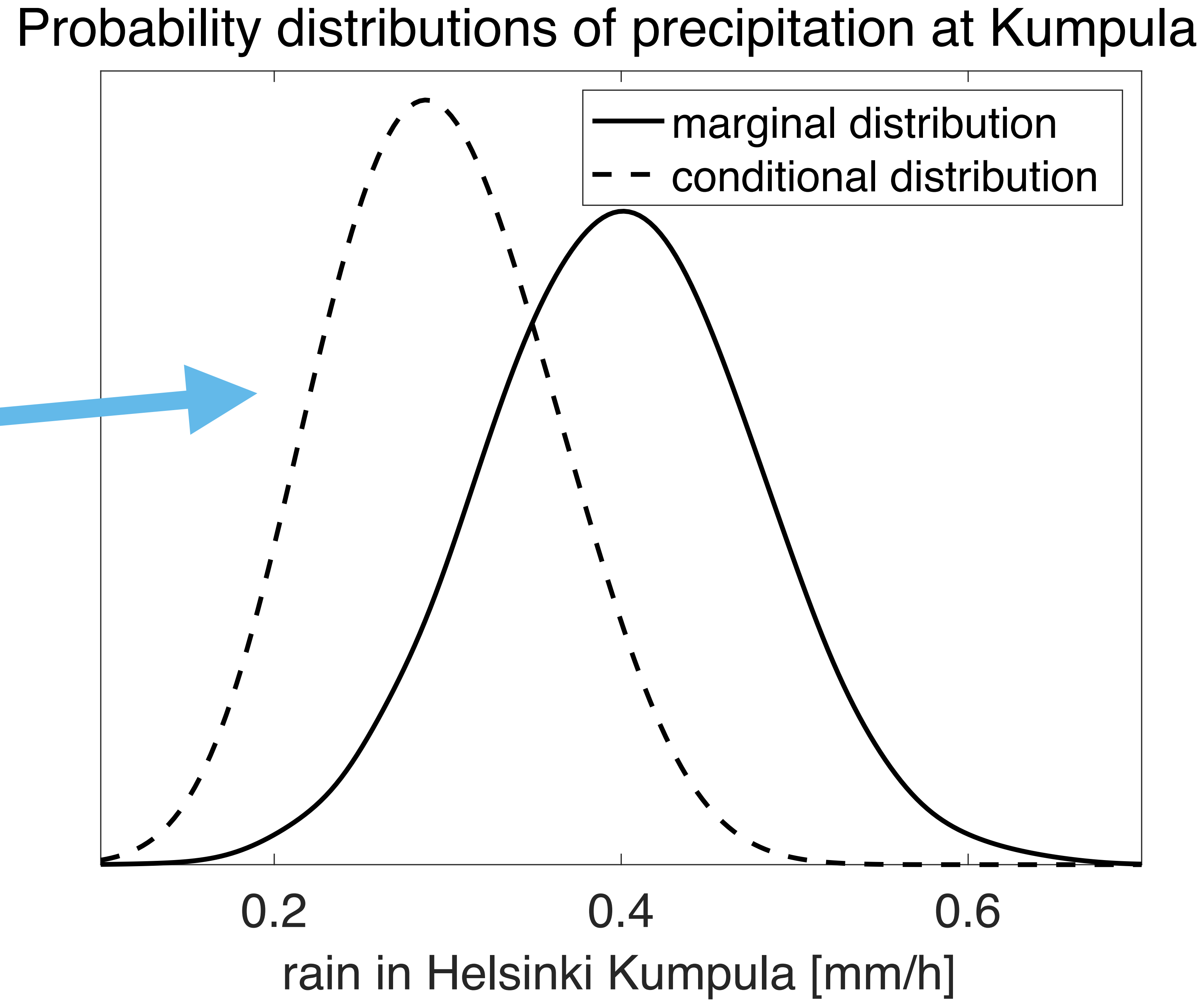
- Two dimensional distribution of precipitation simultaneously at two locations.

- **Marginal distribution** at location 1: no matter what happens at locations 2.
- **Conditional distribution**: conditional on some event at 2.

- Dashed line is the conditional distribution at location 1 given that the precipitation at 2 will be < 0.3 mm/h.

- Multi dimensional probabilities are easy to calculate from model ensembles, but their consistent calibration is a challenge.





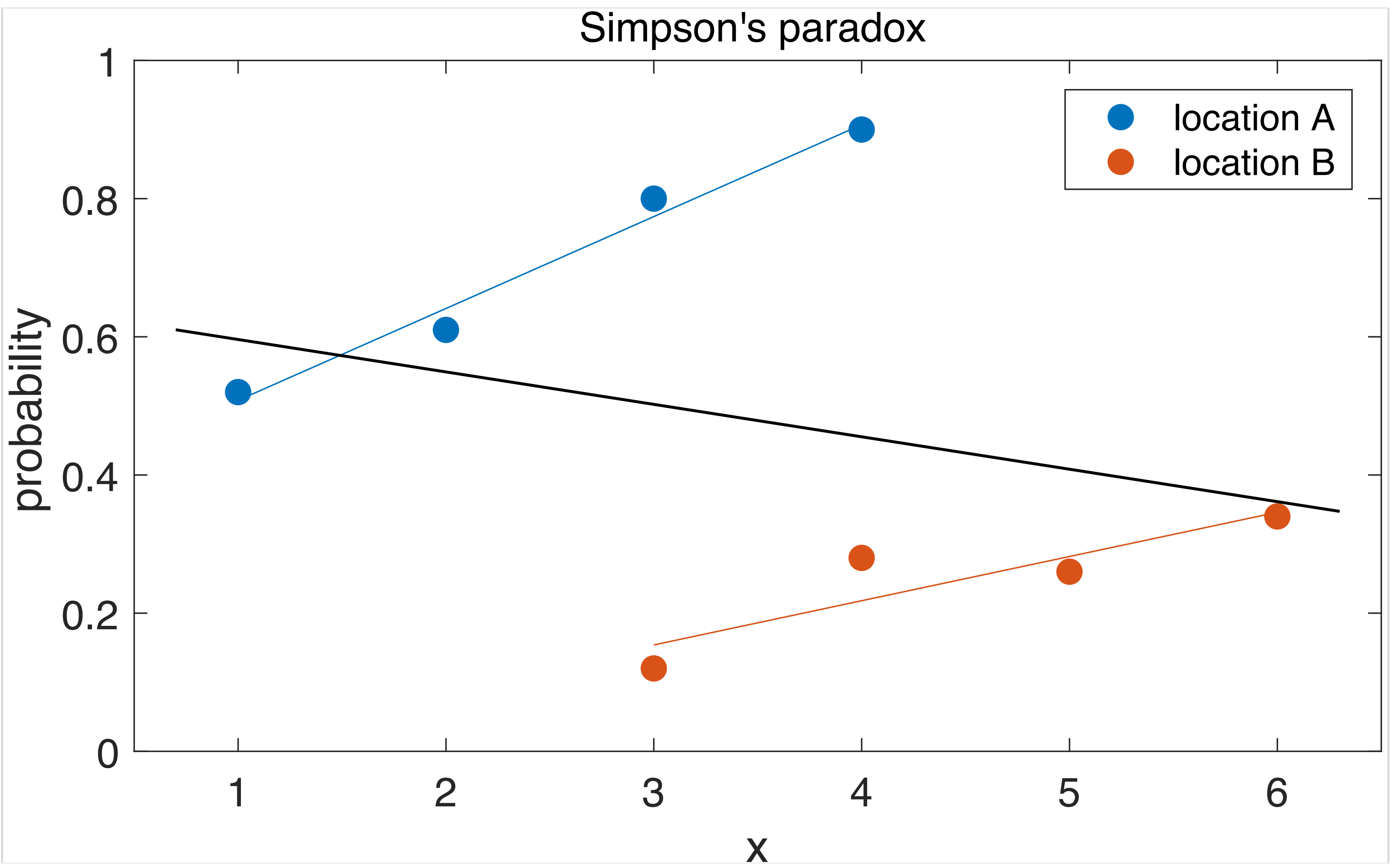
Difficulties with probabilities

- Probabilities, especially conditional probabilities, can fool our intuition.
- *Thinking, Fast and Slow* by Daniel Kahneman:
  - People overestimate rare probabilities.
  - Adding more information, makes the scenario more plausible in our minds.
  - Risk policies are difficult, as we tend to avoid immediate losses.

- Simpson's paradox. Change in the background assumptions, e.g. different climatologies.

Probability of thunderstorm in Helsinki tomorrow at 9 AM while I am cycling to work.

Choose between:
A. sure gain of $ 240
B. 25% chance to gain $ 1,000
and 75% chance to gain nothing



More discussion on probability

- Uncertainty is about the model and our beliefs, not about the Nature.

- Individual probability statements of single future events are measures of subjective belief.
- They can be based on objective facts and must be consistent.
- If P(”rain tomorrow”) = 30 % then P(”no rain tomorrow”) = 70%.
- It is very hard to assign consistent subjective probabilities to complex events.
- Algorithmic, model based forecasts can be verified against observation and tuned to be consistent.

How to make probability forecasts

- Probabilities for an event based on **an ensemble of forecasts** from NWP models.
- **Statistical post-processing** of NWP output from a single model run or the output of ensemble-based NWP.
- By **analysis of historical weather and climate data** to yield statistical relationships between currently observable predictors and the future observations of interest.
- Meteorologist **subjective interpretation** of NWP forecasts and other information.

Consistent terminology is important

- **Confidence** and **likelihood** in the IPCC Fifth Assessment Report.

Likelihood Terminology	Likelihood of the occurrence/ outcome
Virtually certain	> 99% probability
Extremely likely	> 95% probability
Very likely	> 90% probability
Likely	> 66% probability
More likely than not	> 50% probability
About as likely as not	33 to 66% probability
Unlikely	< 33% probability
Very unlikely	< 10% probability
Extremely unlikely	< 5% probability
Exceptionally unlikely	< 1% probability

Confidence Terminology	Degree of confidence in being correct
Very high confidence	At least 9 out of 10 chance
High confidence	About 8 out of 10 chance
Medium confidence	About 5 out of 10 chance
Low confidence	About 2 out of 10 chance
Very low confidence	Less than 1 out of 10 chance

FMI POP terminology

probability of precipitation	change for rain or showers
less than 10 %	dry weather
10 – 30 %	small chance
30 – 70 %	medium chance
70 – 90 %	high chance
over 90 %	overall change

Deterministic, stochastic, chaotic

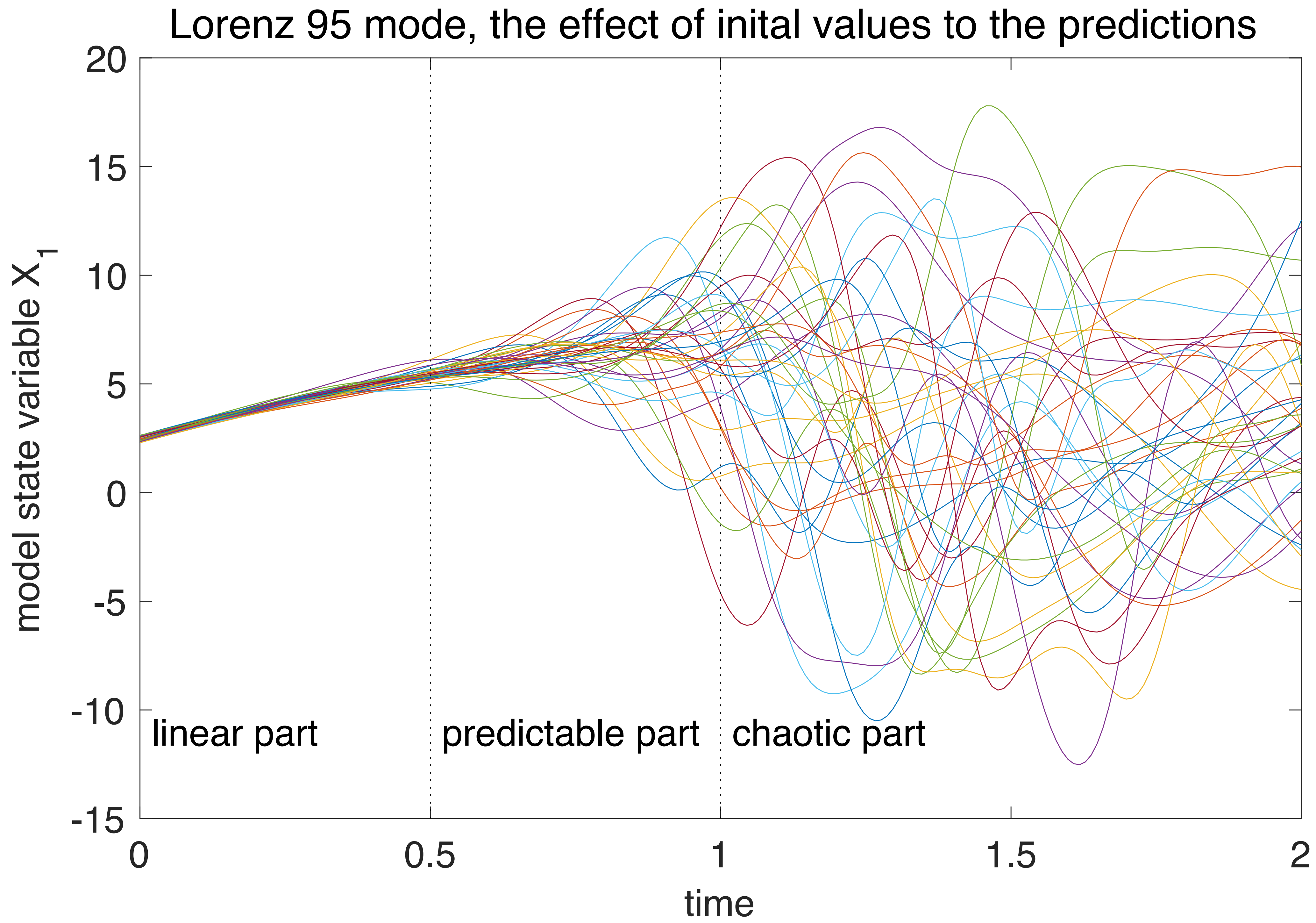
- A phenomena is **deterministic**, if its final state can be predicted form initial conditions.
- A phenomena is **stochastic** or random if there are several possible final states from the same initial state, but there is systematic statistical behaviour in the distribution of outcomes.
- A phenomena is **chaotic**, if a small change in initial conditions leads eventually to non predictable state.

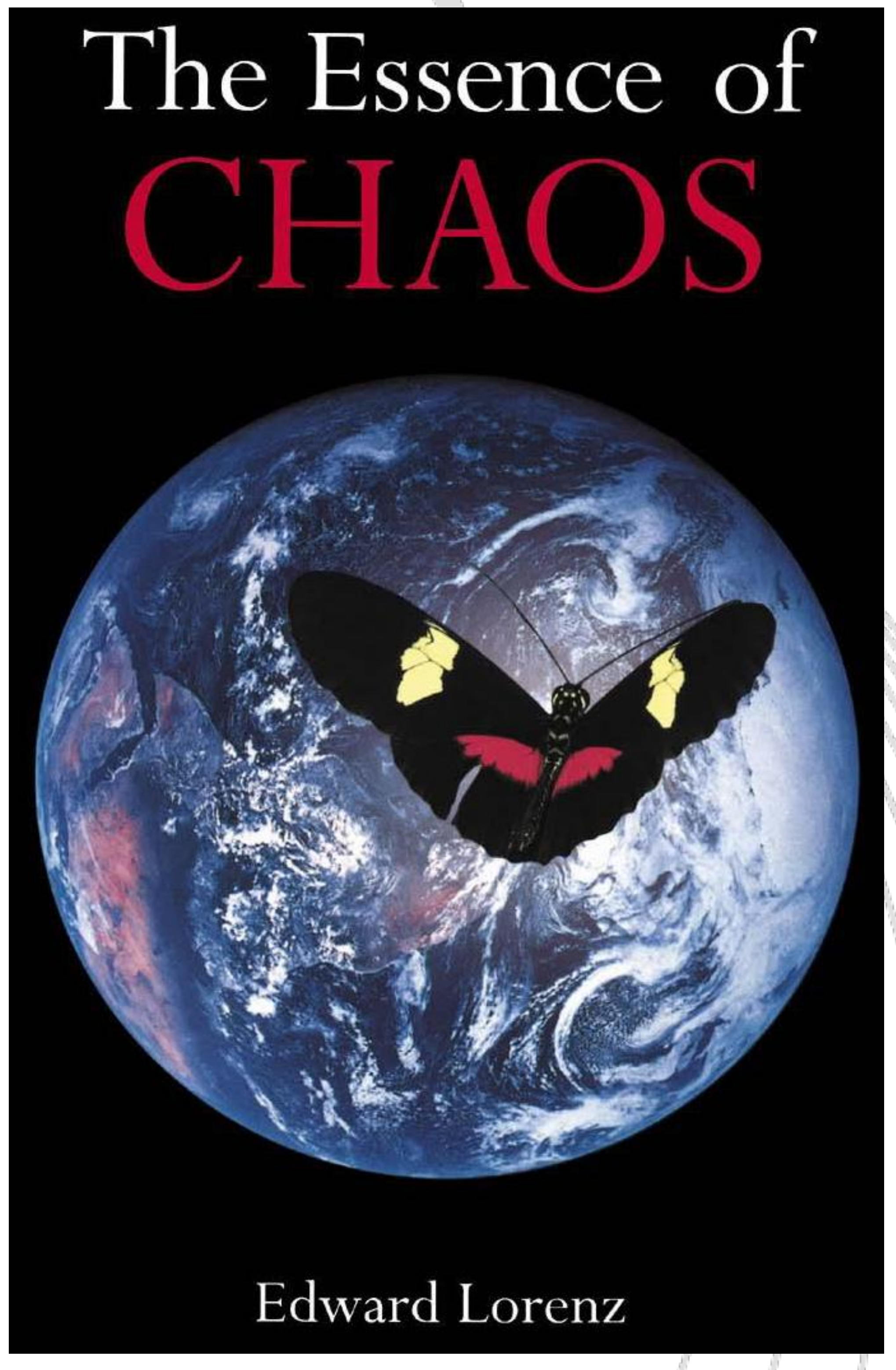
- The weather system: stochastic and chaotic.
- Numerical weather model: deterministic and chaotic.

Predictability and chaos

- Numerical models describing weather are chaotic: a small perturbation in the initial conditions accumulates and makes the system eventually **non predictable**.
- Small change = one bit in computer representation.

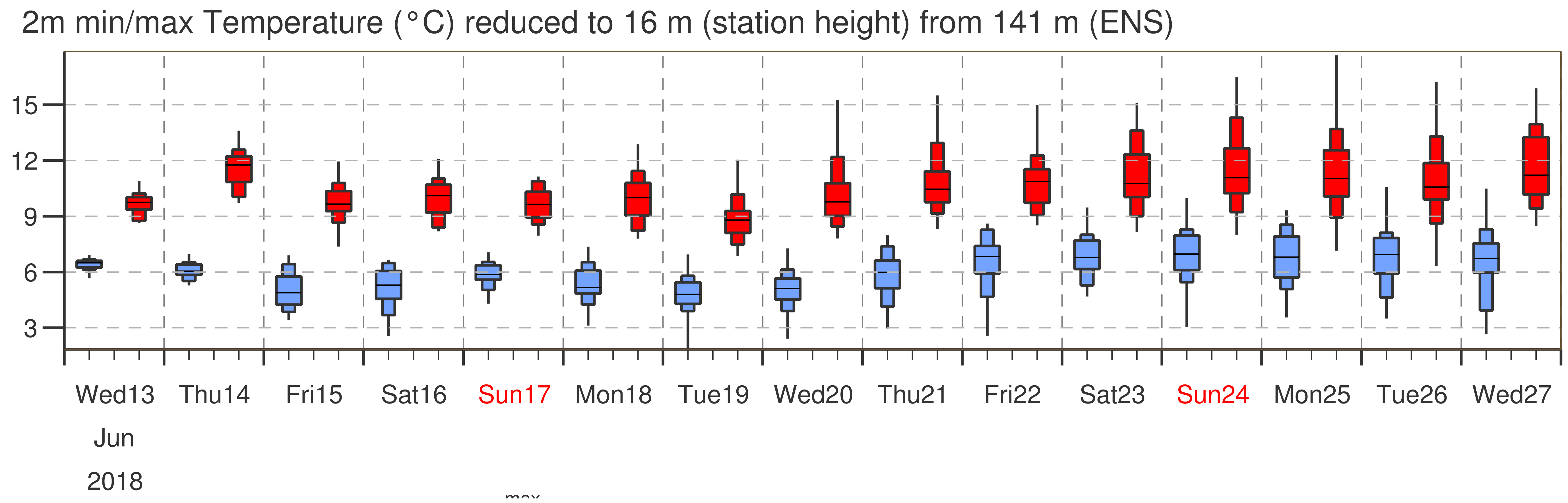
- By **perturbing model initial values** we can evaluate the predictability!

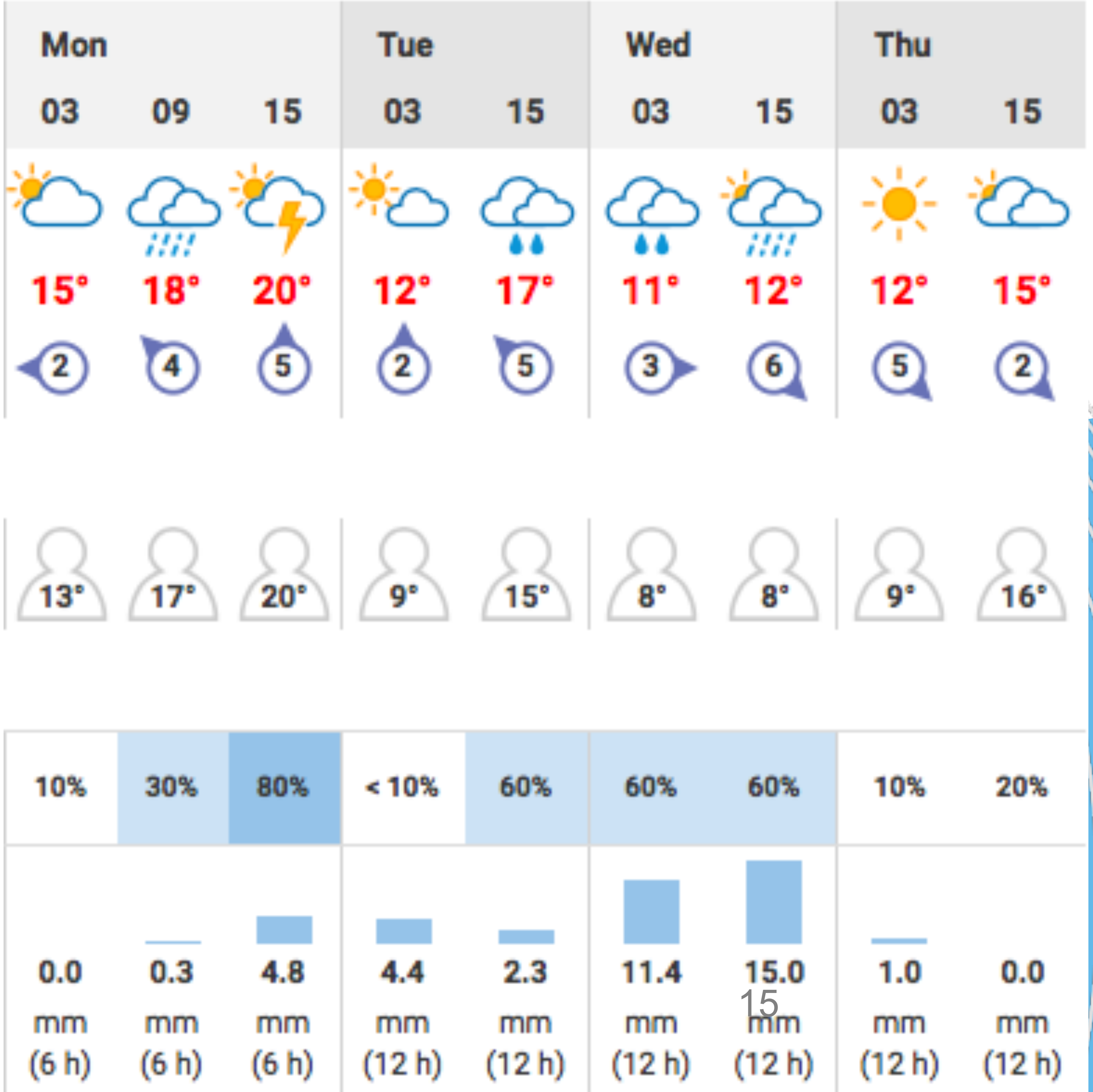




Quality of probability forecasts

- No matter how probability forecasts are made we want them to have good performance in the long run.
- Every time the forecaster says rain with 60% probability, we assume that in 6 out of 10 times it rains.
- So, from one forecast no quality statements are possible.



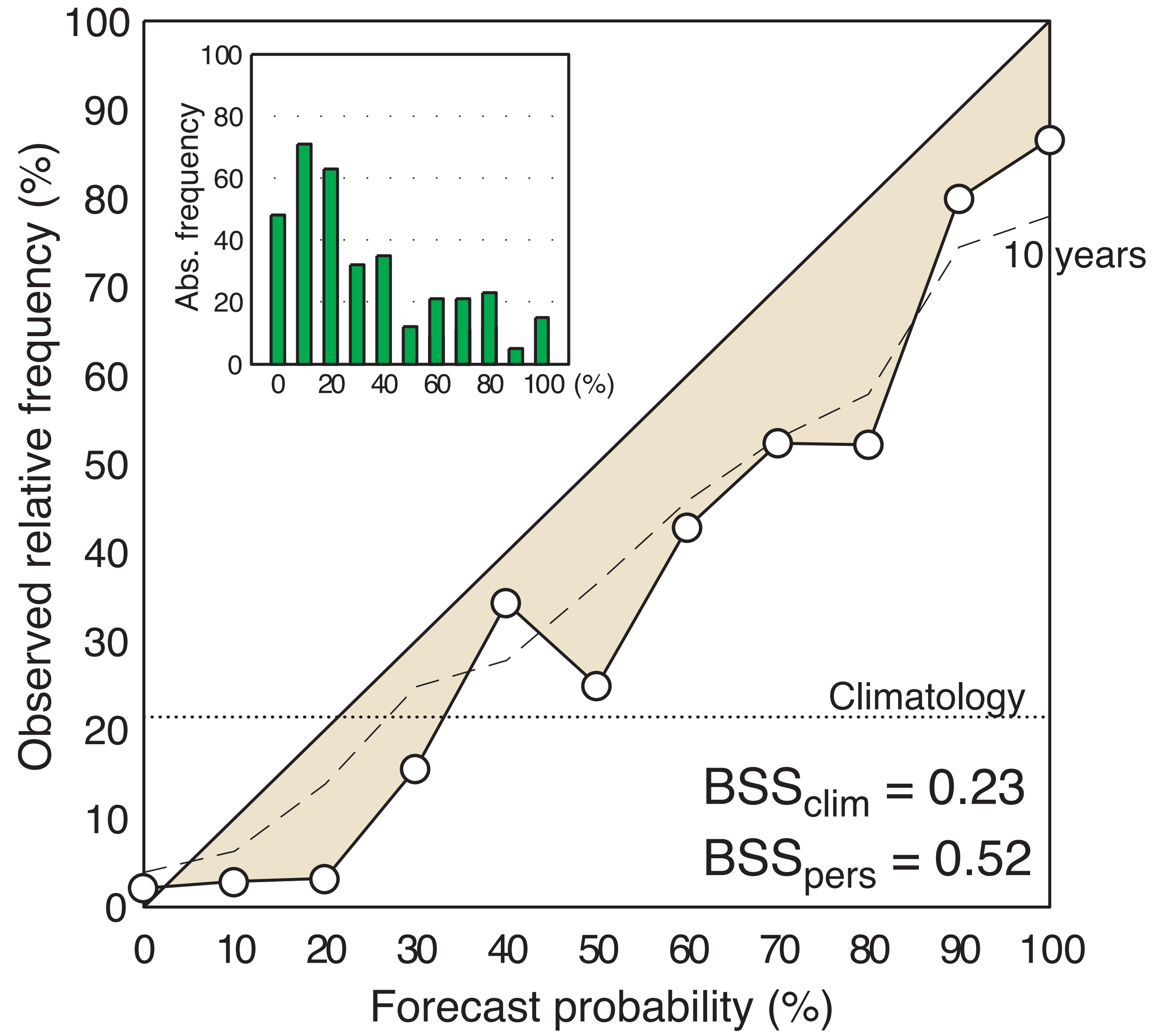


How to verify probabilities

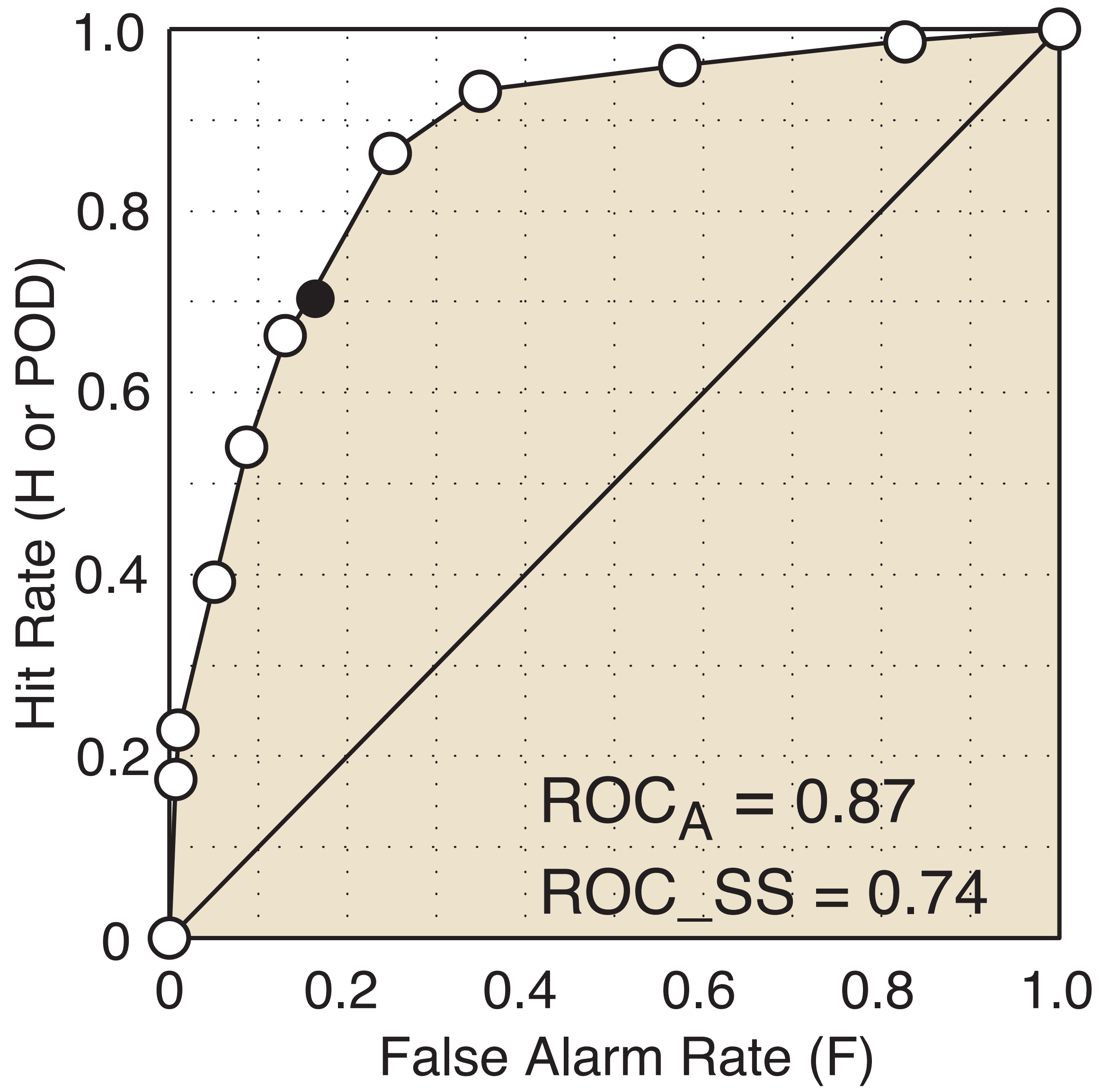
- When we do **repeated** probability statements, they can be verified by using actual observations. The forecasted probabilities have to match the observed frequencies (reliability). Several statistics and diagrams are used.

Reliability and ROC diagrams of one year of Probability of Precipitation forecasts. The reliability curve (with open circles) indicates strong over-forecasting bias throughout the probability range.

Reliability diagram



ROC

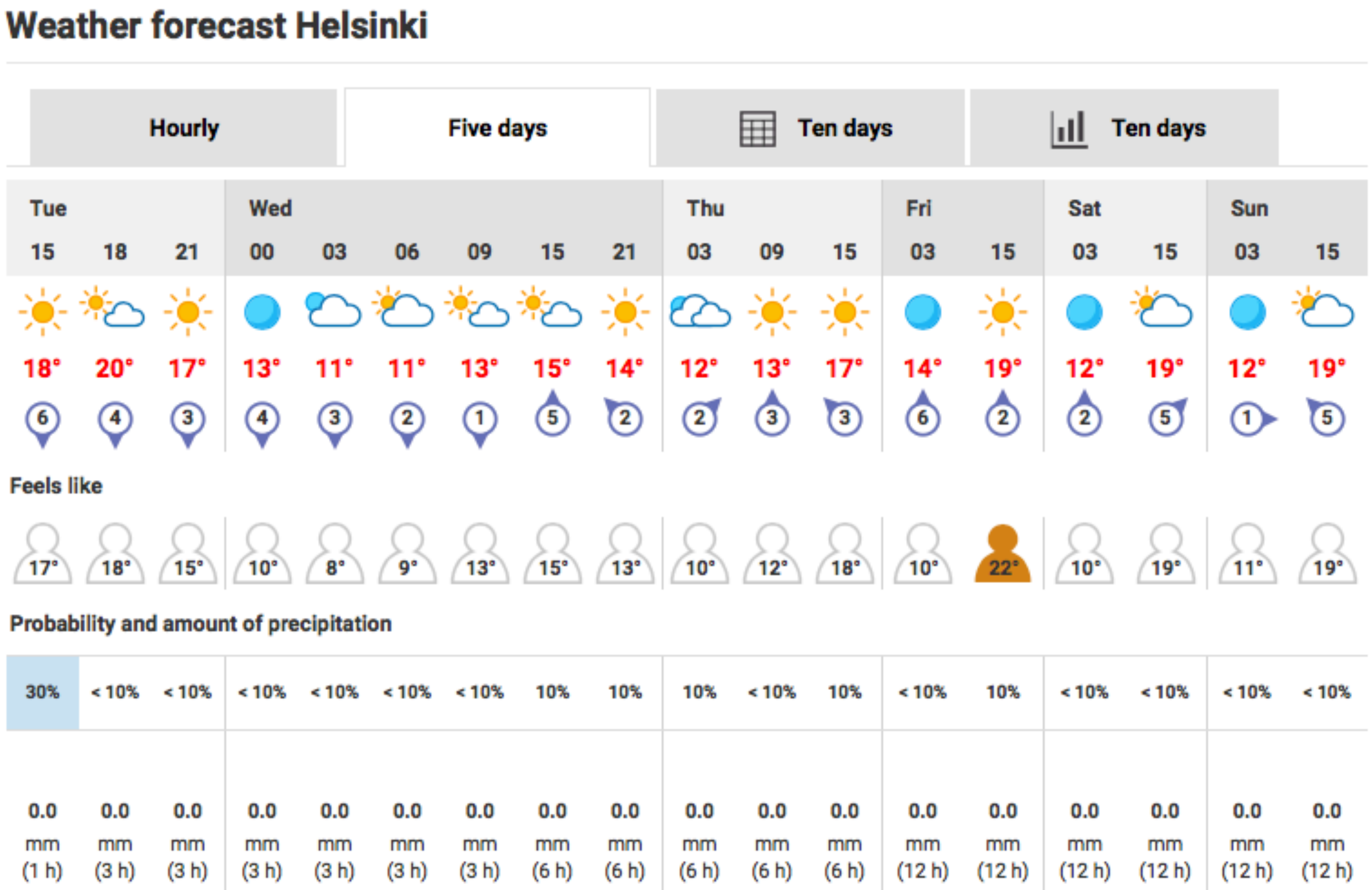


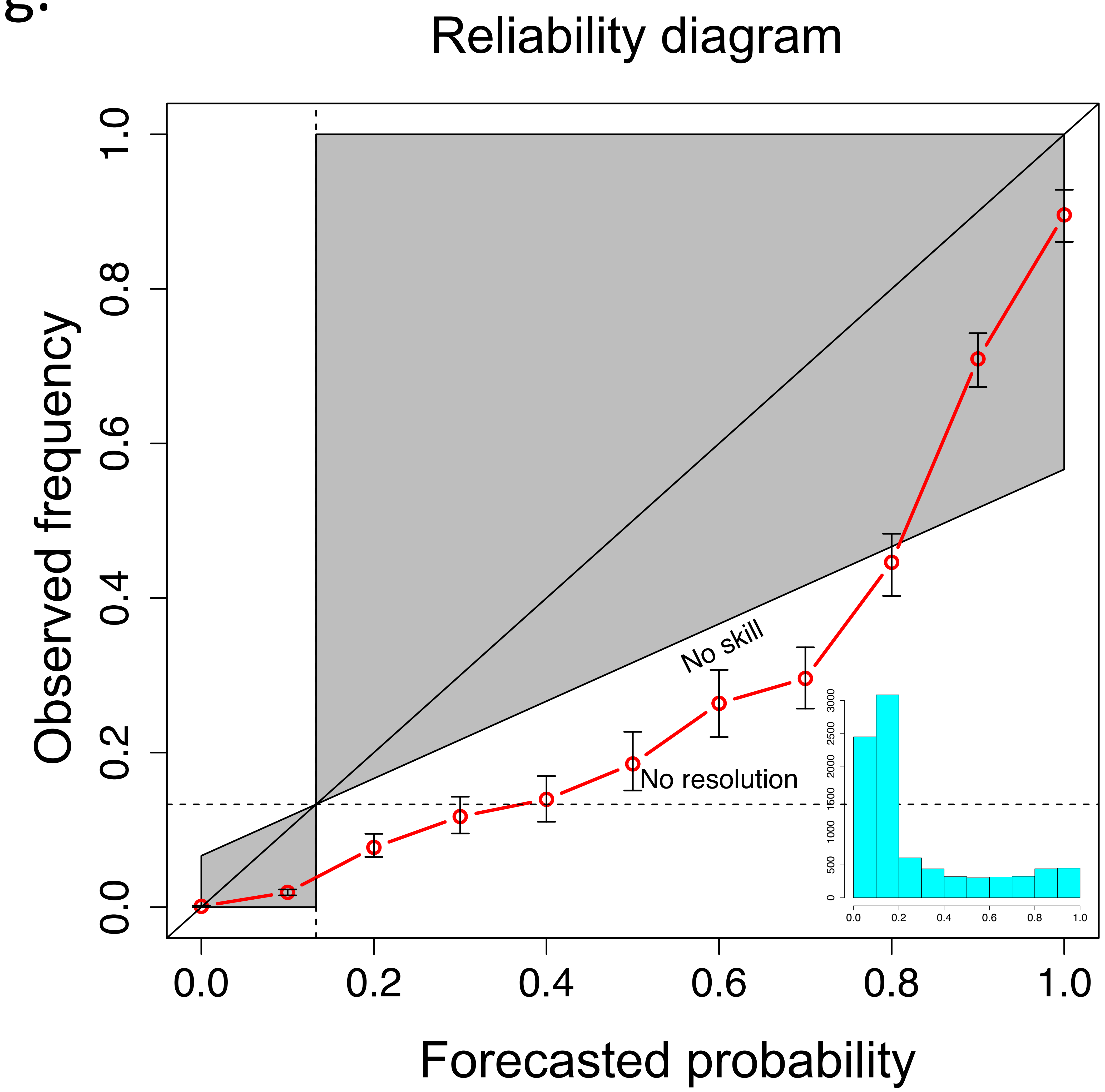
The ROC curve is constructed on the basis of forecast and observed probabilities leading to different potential decision thresholds. The black dot represents the single value ROC when using 50% treshold (H=0.7; F=0.17).

Figures by Pertti Nurmi.

Example: POP at FMI

- POP at FMI by multi model neighbourhood processing.
- Helsinki Kaisaniemi stations,
  P(prec>0.1 mm/h) 24 h fcst, all of 2017.





Ensemble forecasts

- Conceptually, the best way to make probability forecasts for complicated events.
- Run the same forecast model with **perturbed initial conditions**.
- Probability 20% means that 10 out of 50 ensemble members predict the event to happen at the specified location in the defined time window.

- ENS systems have to be tuned to match **predictability** and account model's inaccuracies.
- To be useful, ensembles have to be **calibrated** to correct the spread and remove biases.

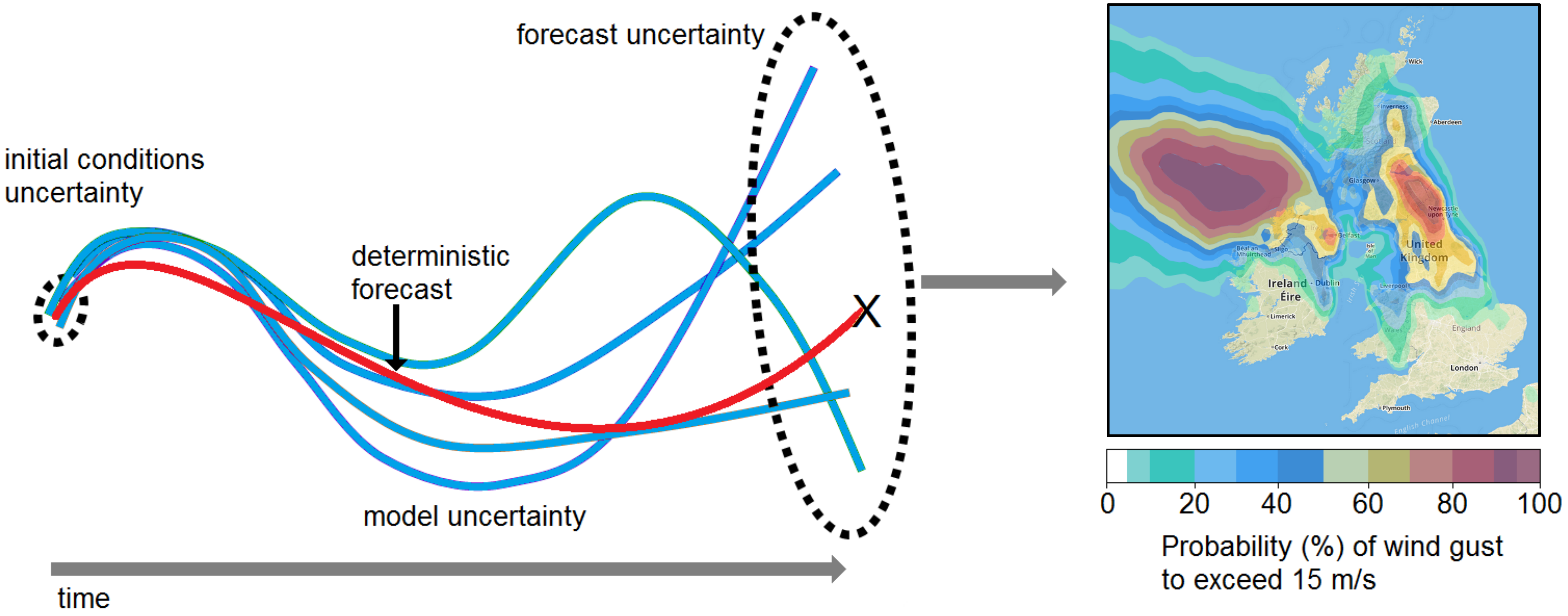
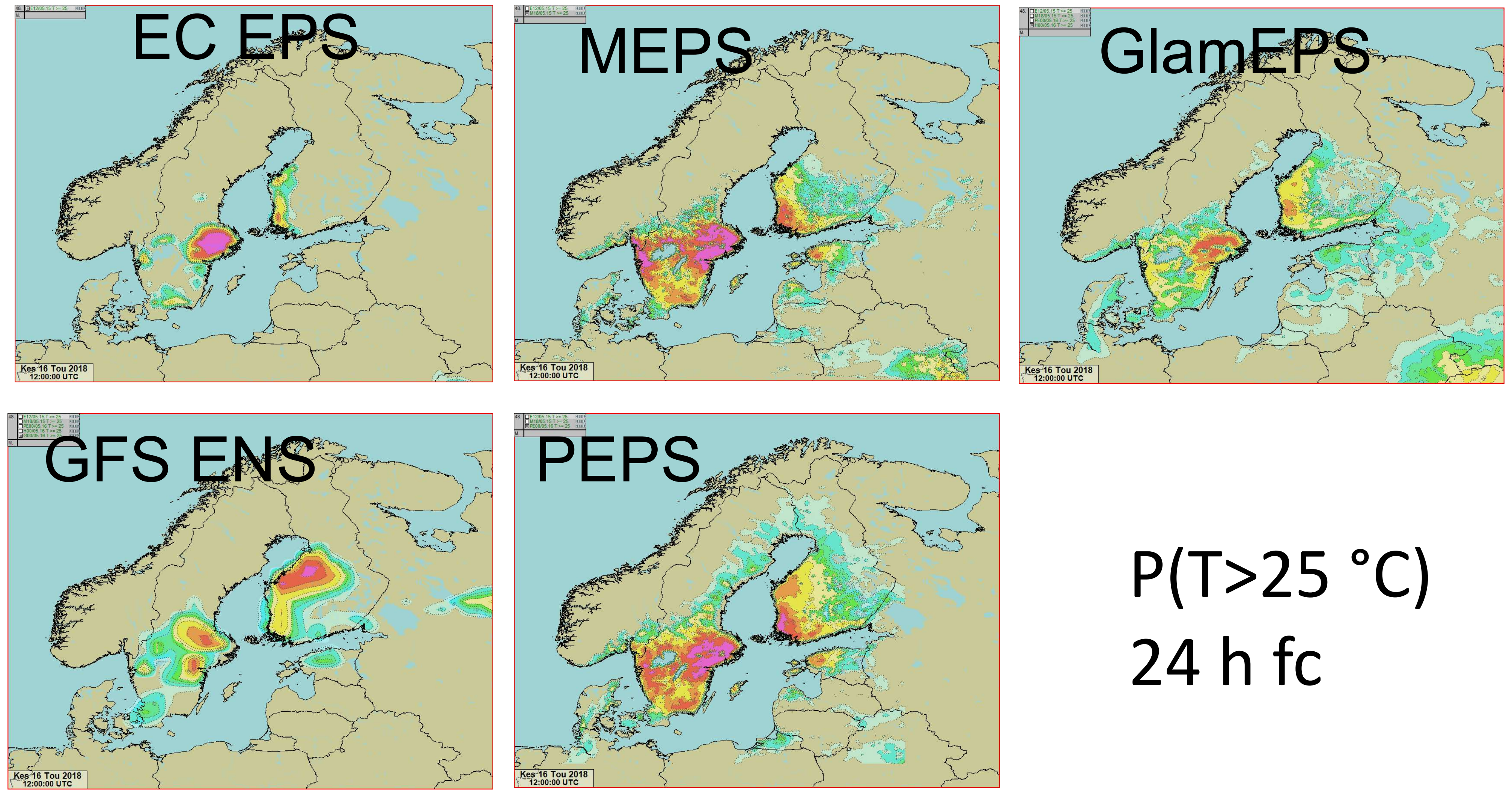


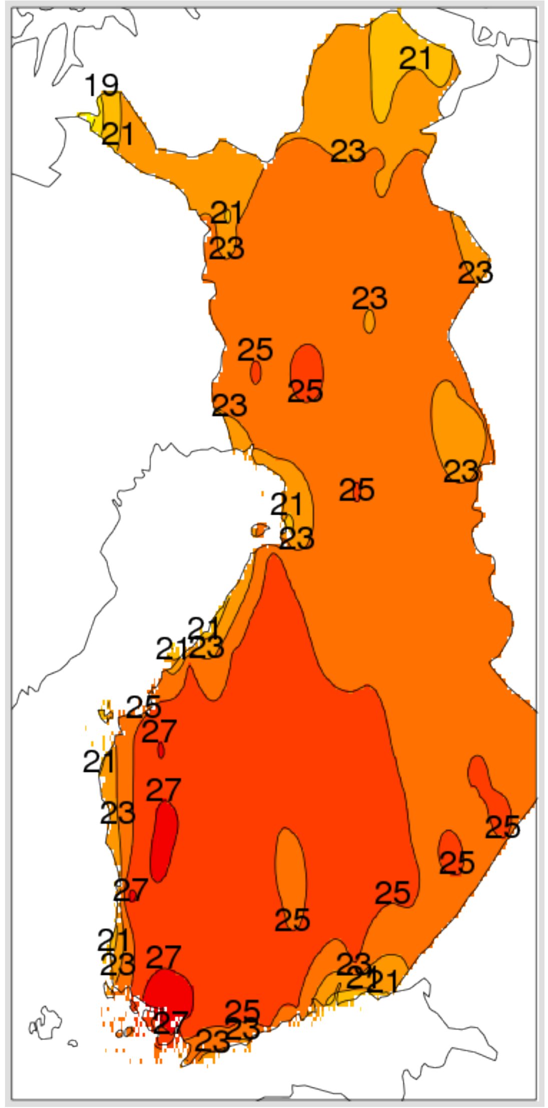
figure: Kaisa Ylinen FMI

Example: warm May in Finland

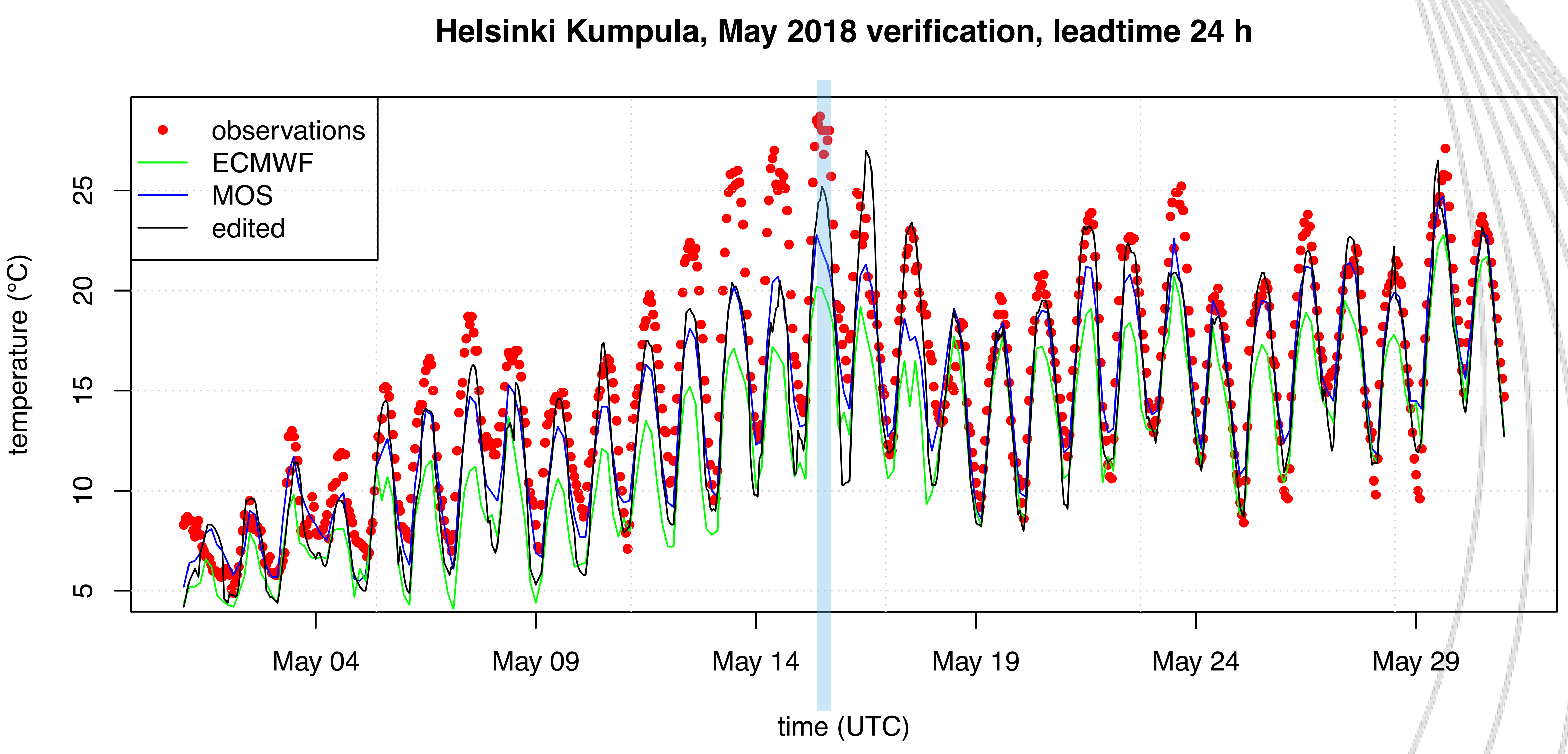
- Probabilities for warm weather for 16.5.2018 12 UTC with 25 h lead time.
- Pink colour in the model maps on left means P(Temp > 25 °C) > 90%
- All the models are too cold, there is bias, and the ensemble spreads are too narrow.
- There is a strong need for post-processing and ensemble calibration.



P(T>25 °C)
24 h fc

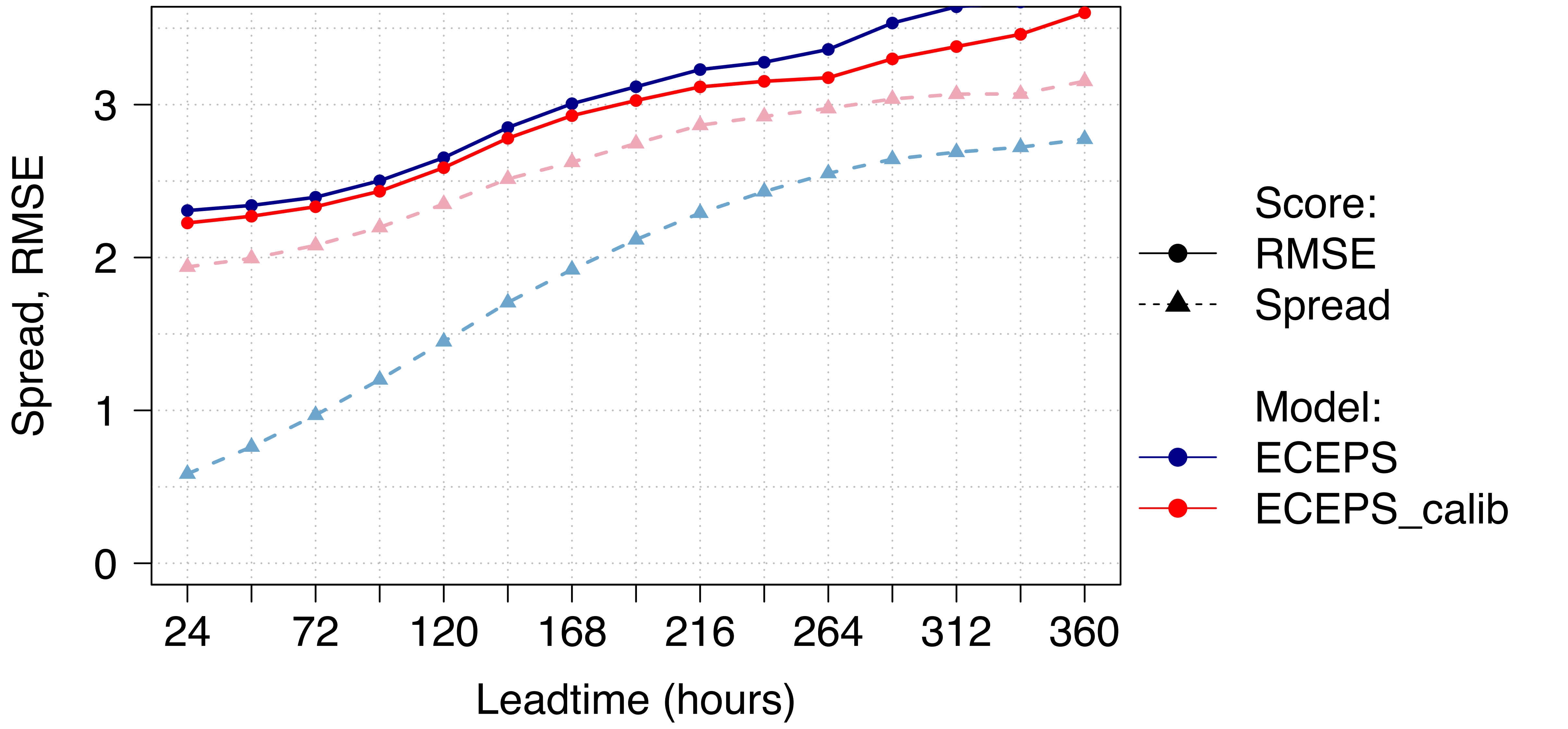


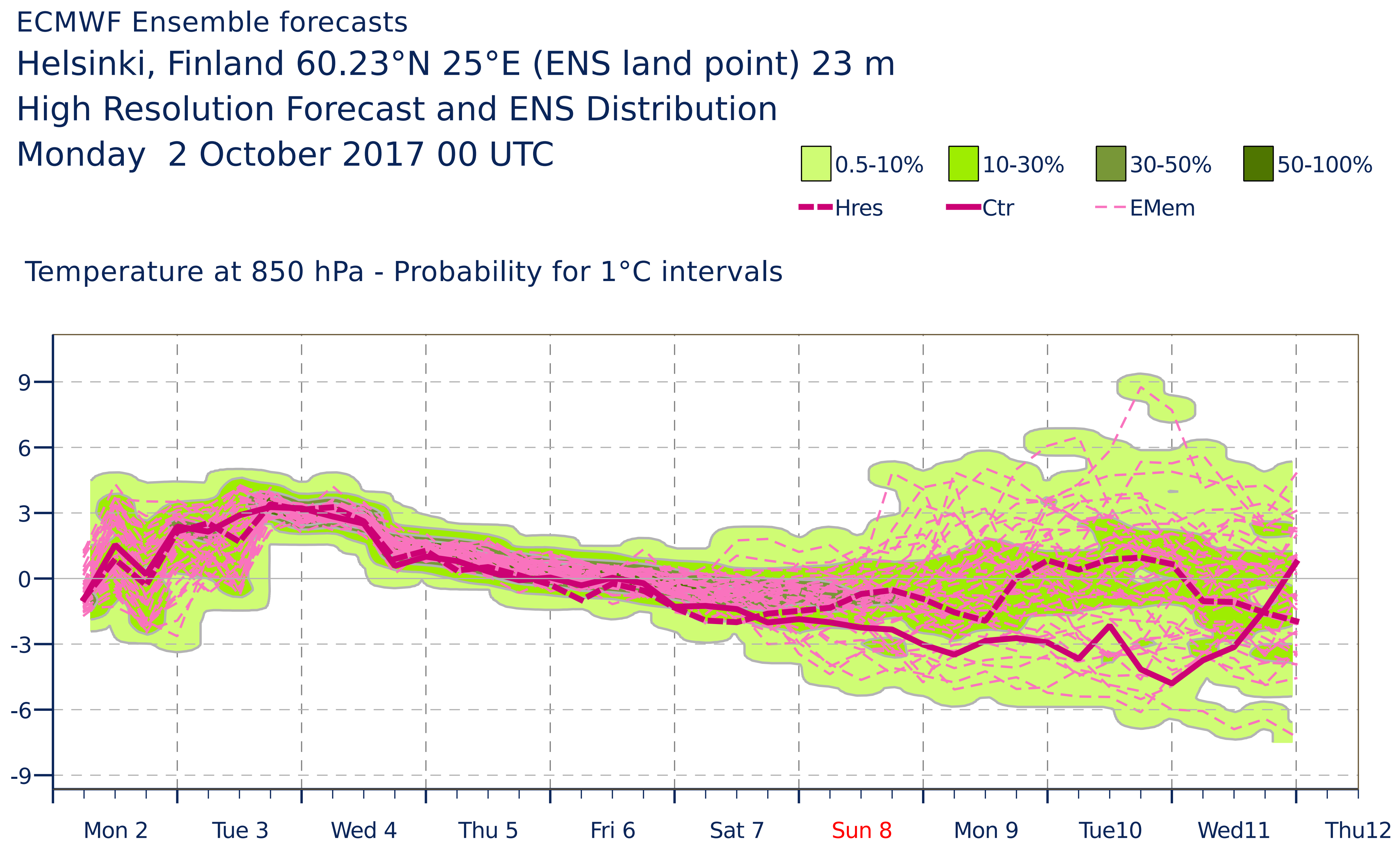
Tmax 16.5.2018

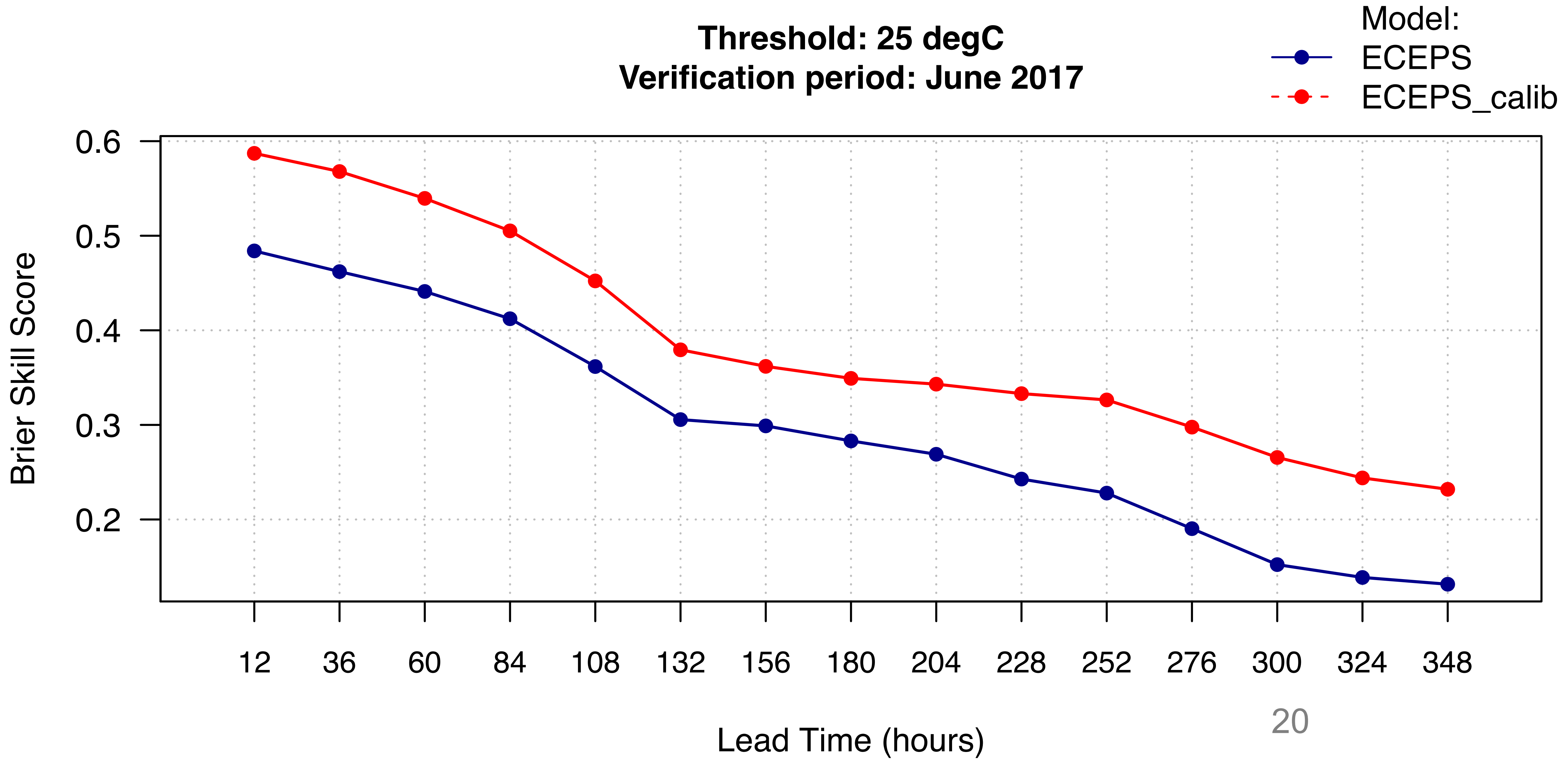


Calibrating ensemble forecasts by past observations

- ECMWF EPS with 51 members.
- Harp tool from the Hirlam group.
- Ensemble MOS with 30 days history.
- EU H2020 I-REACT project.







Conclusion — why probability forecasts

- By quantifying the uncertainties related to forecasts we give more information than by a single deterministic forecast.
- They allow better handling of risks associated with different actions.
  ”We want to be 95% sure that in the next 30 years the water level will rise more than 1 m from the average less that 2 times.”
- They allow for better verification measures, i.e. which account for the predictability.
- There are still no perfect systems for probability forecasts, work to be done on EPS tuning and post-processing.

Several people at FMI contributed this talk, including:
Leila Hieta, Kaisa Ylinen, Juha Kilpinen, Marja-Liisa Tuomola, Carl Fortelius, Jussi Ylhäisi

This is the last slide! Thank you!