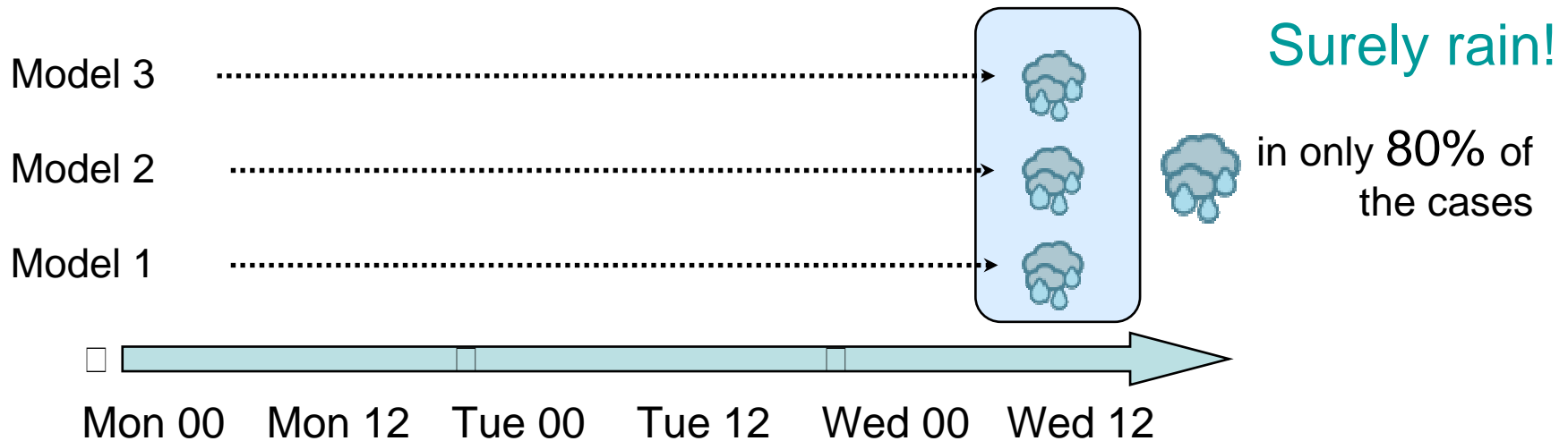
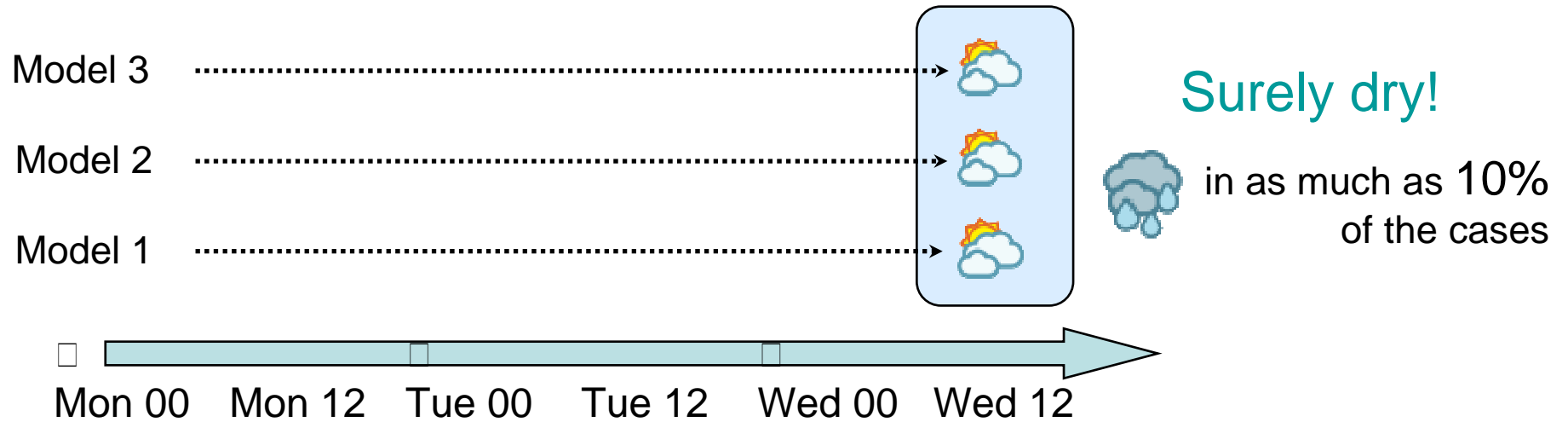


V The psychology of probabilities

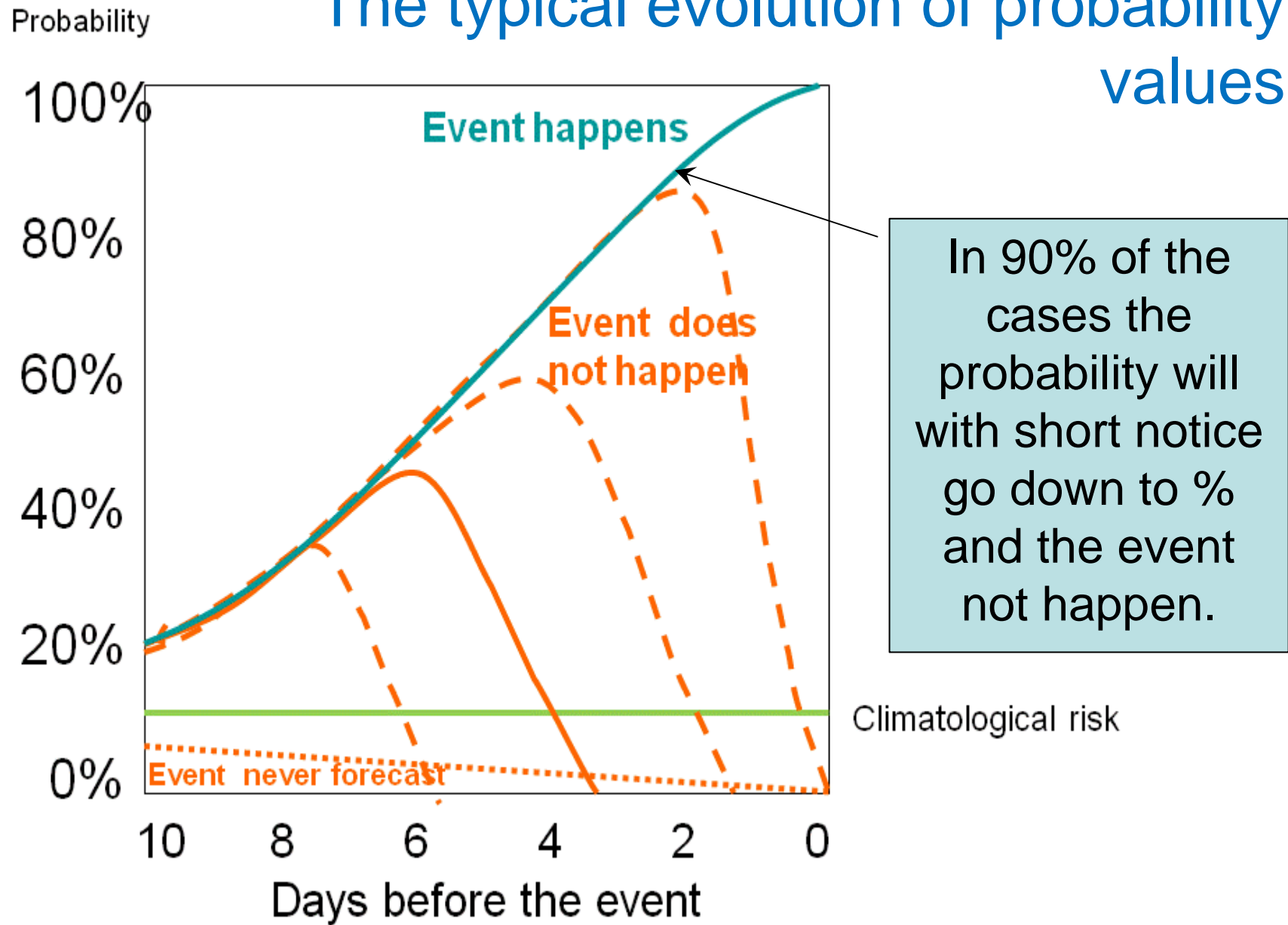
1. Common pitfalls

V.1.1 Over-confidence

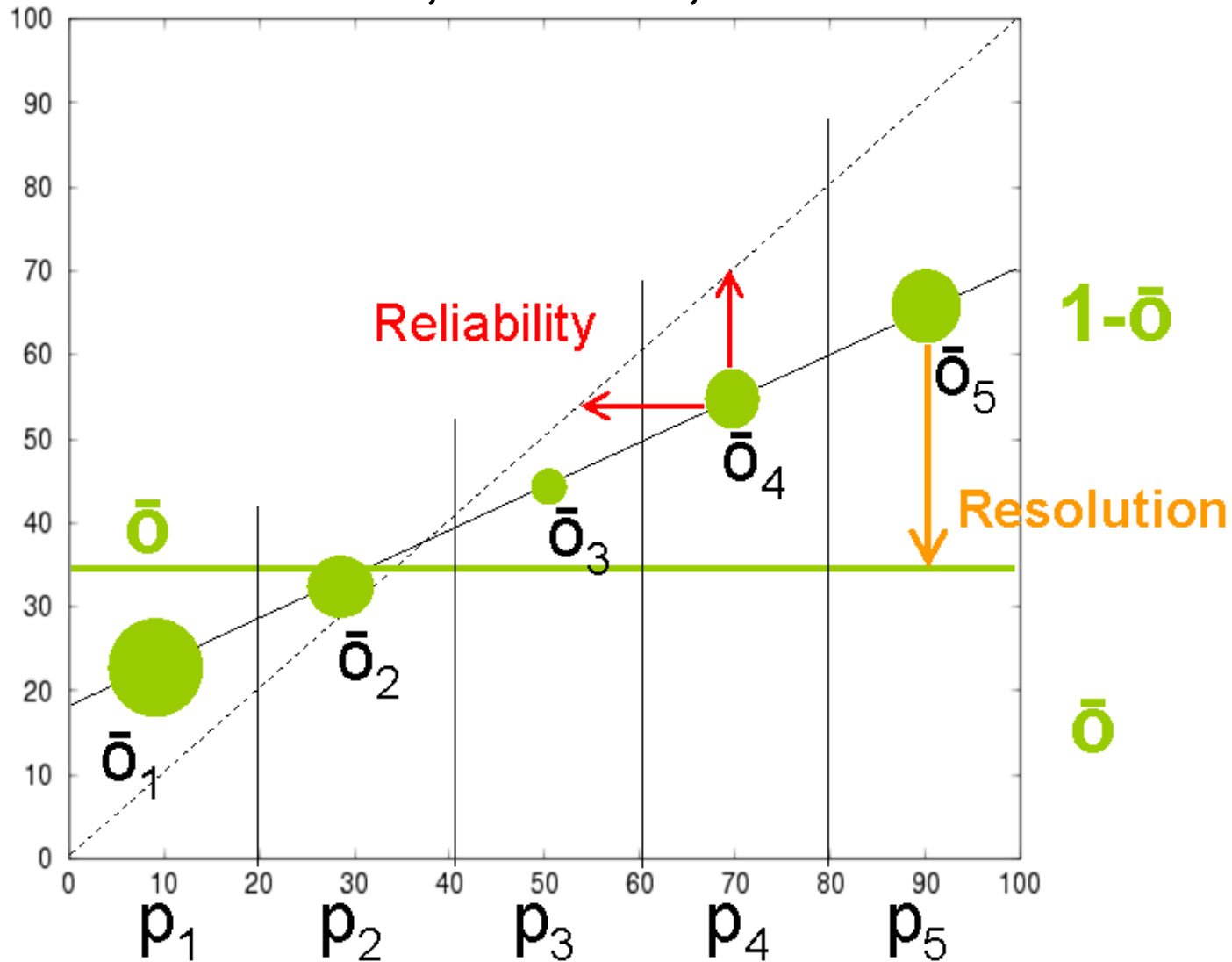
Three forecasts from different NWP models valid at the same time



The typical evolution of probability values

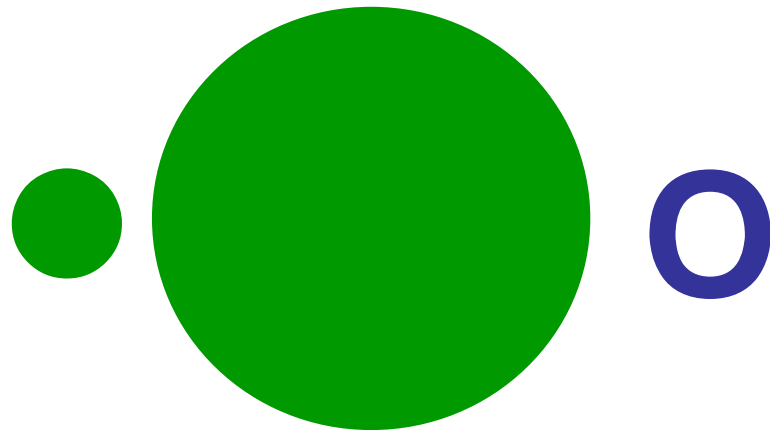


A typical over-confident reliability diagram from the 1950's, 1970's, 1990's and now



V.1.2 Halo effect

The halo effect (at the UKMO)



ECMWF

UKMO

GFS

UKMO! ! *Best model in the world!*

In the 1970's, when I was a young forecaster, there were three major NWP centres in
Washington, Bracknell and Offenbach

- The Norwegian forecasters favoured the American NWP
- The Danish forecasters favoured the British NWP
- The Finnish forecasters favoured the German NWP

And what did the Swedes, “neutral” in WWII, do??

-USE ALL OF THEM!

V.1.3 Representative effect

Representative bias



What is more
probable?

1. Danny is an accountant
2. Danny is an accountant and a skilful rugby player

Representative bias



People illogically tend to
chose 2.

A case of **representative
bias**

He appears more ***typical*** for
group 2. than for 1. because
of the added realistic detail.

People tend to chose 2. without considering that there are much
more accountants than skilful rugby players (see ***base rate*** later).

It is a common mistake, also in our science, to
confuse “**typical**” and “**probable**”

More on “typical” and “probable”

A coin is tossed 10 times showing **Heads** or **Tails**

Which outcomes are more probable?
And which are less?

HHHHHHHHHH

HHHHHTTTTT

HHTHTHTTTH

Least probable?

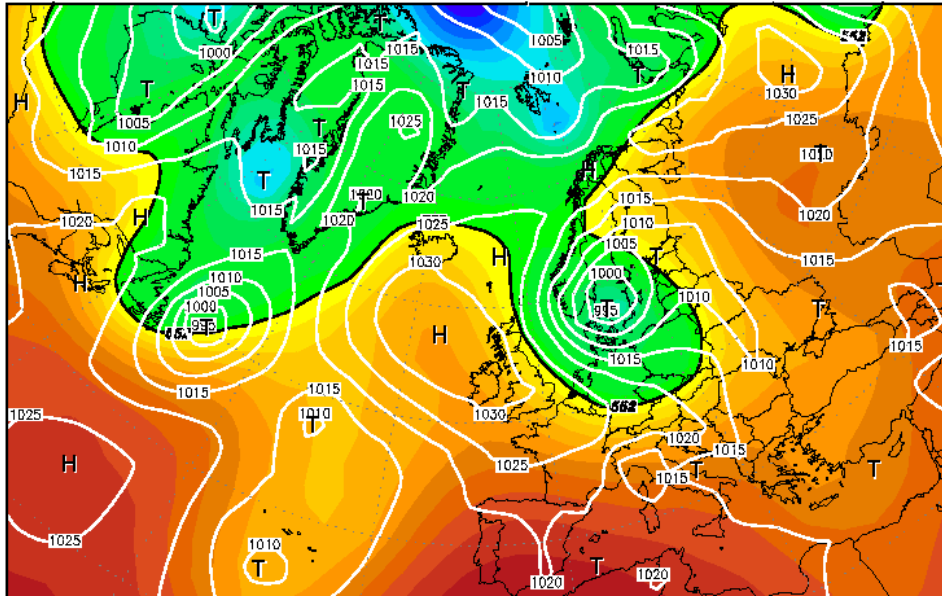
Most probable?

They are all equally probable

Humans again confuse what is “typical” with “probable”

This is another “representative error”

Init : Fri,04MAY2012 00Z Valid: Sat,12MAY2012 00Z
500 hPa Geopot. (gpm) und Bodendruck (hPa)

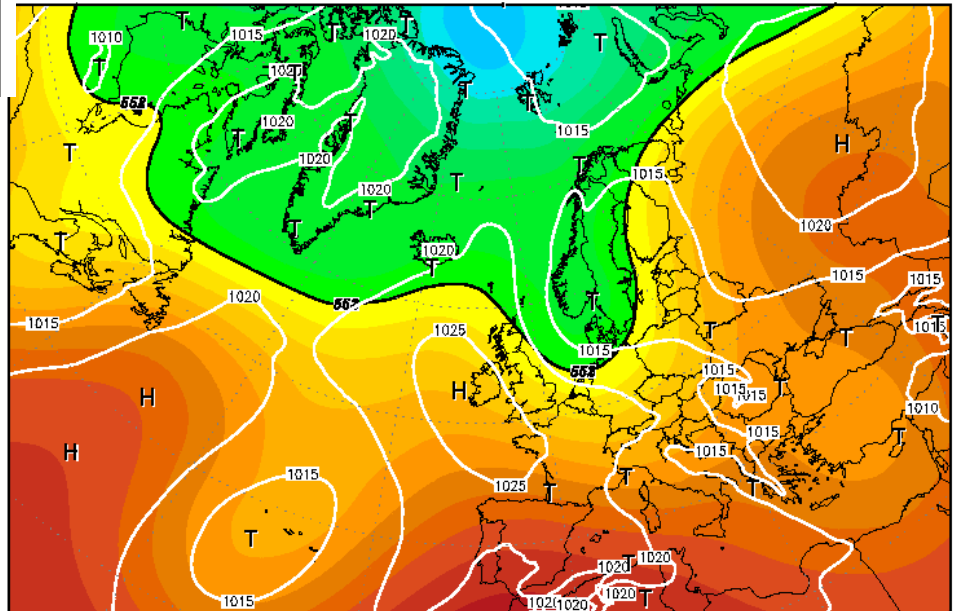


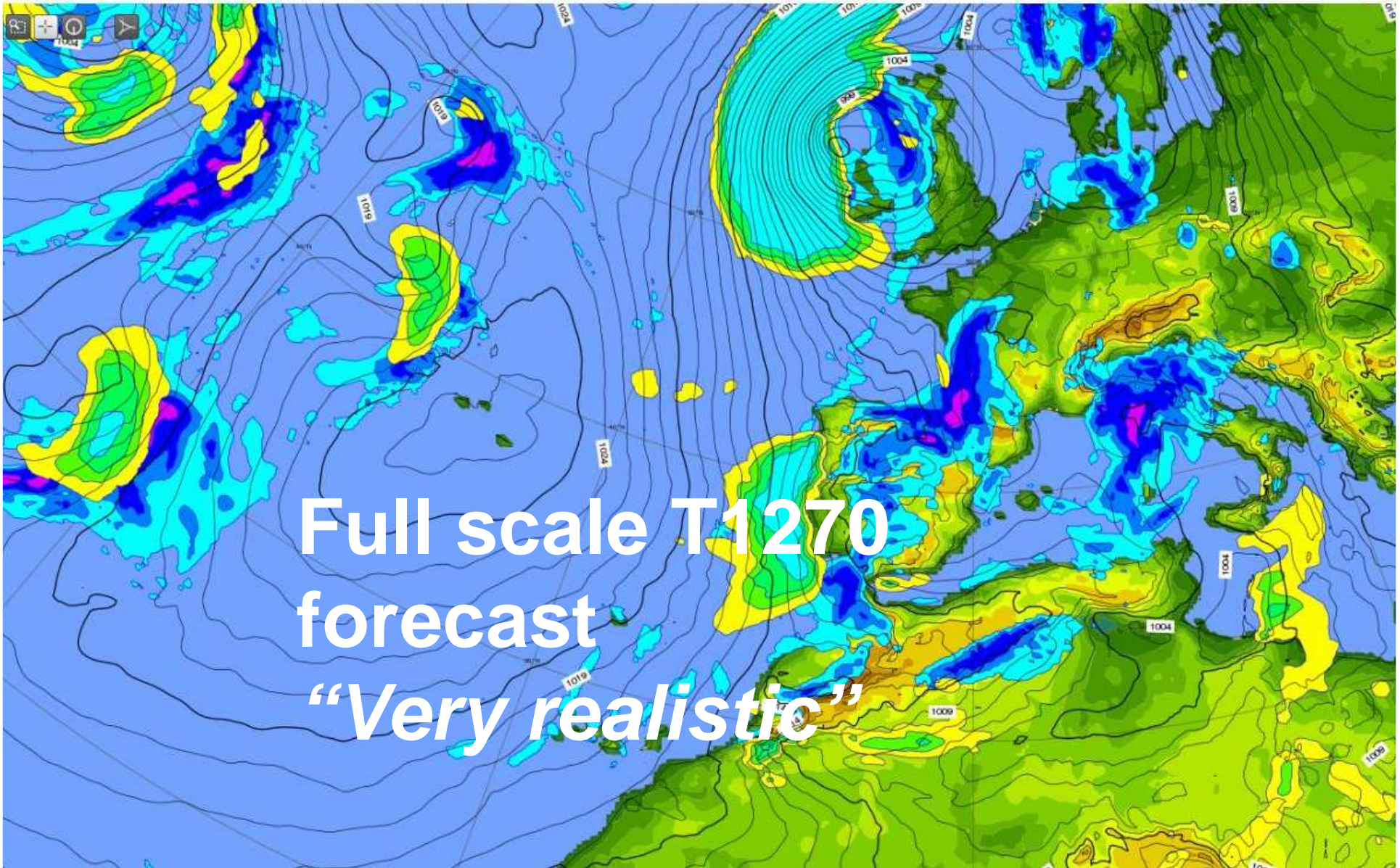
Full scale T1270 forecast
“Very realistic”

The same, but the ensemble mean
“The atmosphere cannot look like this”

We meet representativeness error in the tendency to provide too detailed, and therefore more realistically looking, forecasts.

nit : Fri,04MAY2012 00Z Valid: Sat,12MAY2012 00Z
500 hPa Geopot. (gpm) und Bodendruck (hPa)



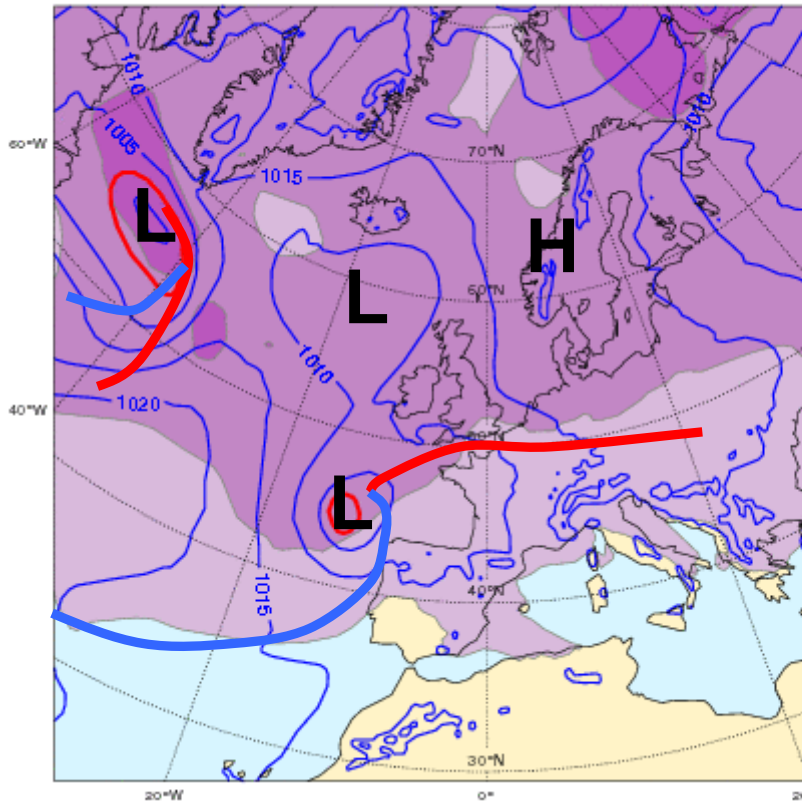


ECMWF mistake:

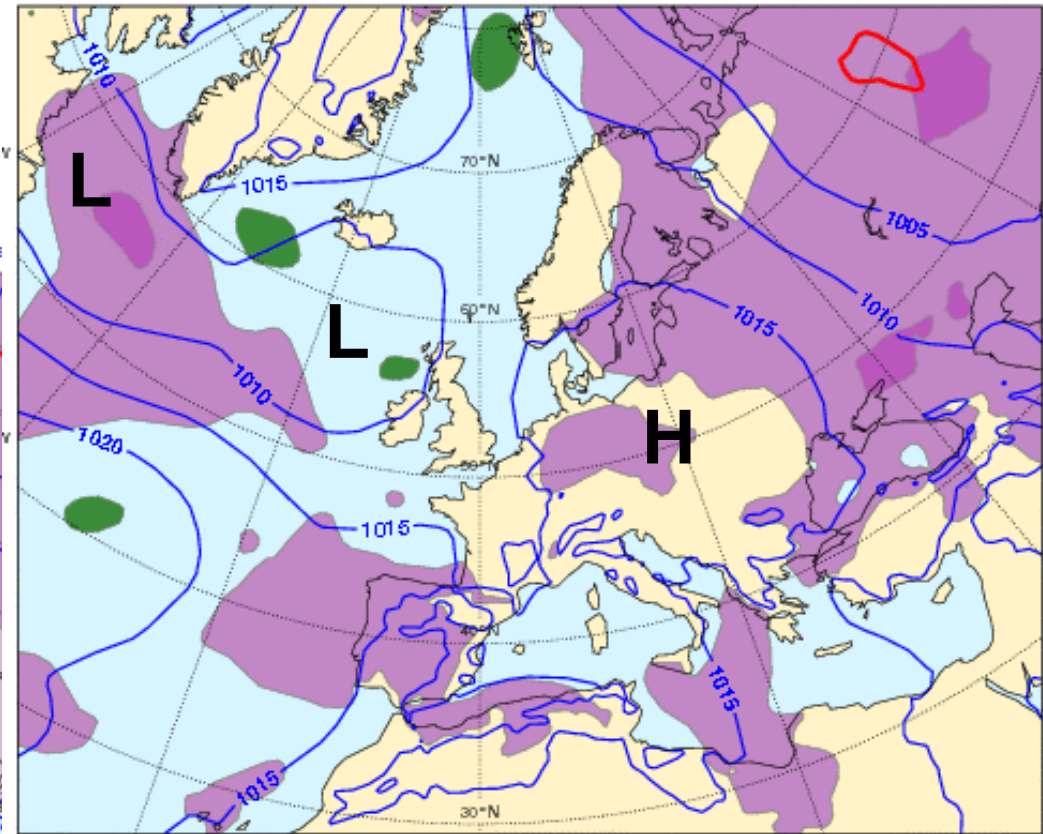
- 1. Using their skill and resources to paint the T1279 in the most wonderful colours, making it more attractive and “available”**

In weather forecasting we meet the representativeness error in the tendency to prefer detailed, and therefore more realistically looking, forecasts.

Wednesday 15 August 2012 12UTC ECMWF Forecast +168 VT: Wednesday 22
 Mean sea level pressure (MSLP) Deterministic Forecast and Standard Devi



Wednesday 15 August 2012 12UTC ECMWF Forecast +168 VT: Wednesday 22 August 2012 12UTC
 Mean sea level pressure (MSLP) Ensemble Mean and Normalised Standard Deviation (shaded)



The full scale ECMWF forecasts looks more realistic than the ensemble mean

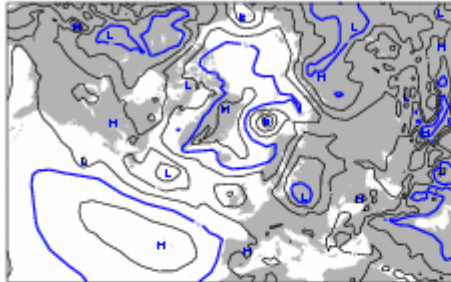
Clustering 2010 -

Wednesday
Ri

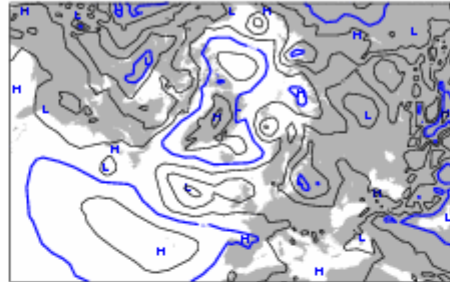
“Typical member”

1Pa Geopotential
clustering

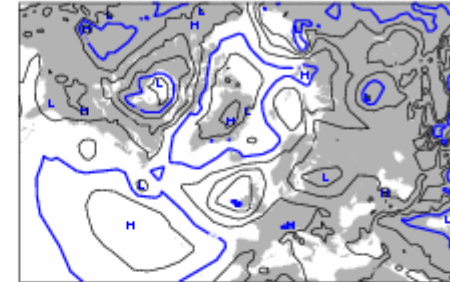
forecast t+120 VT:Monday 4 July 2011 00UTC
Cluster: 1(of 3), population: 22, repres. member: 6



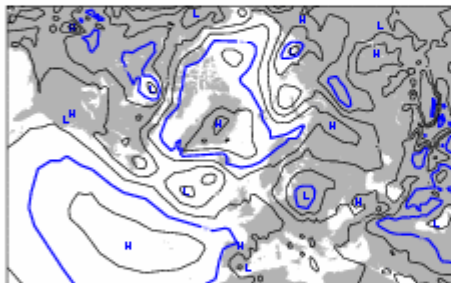
forecast t+144 VT:Tuesday 5 July 2011 00UTC
Cluster: 1(of 3), population: 22, repres. member: 6



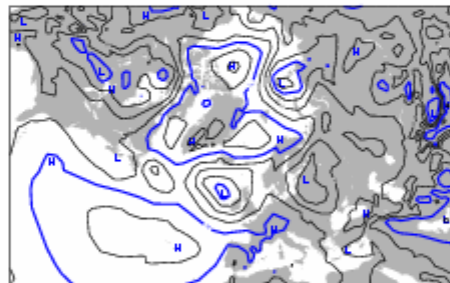
forecast t+168 VT:Wednesday 6 July 2011 00UTC
Cluster: 1(of 3), population: 22, repres. member: 6



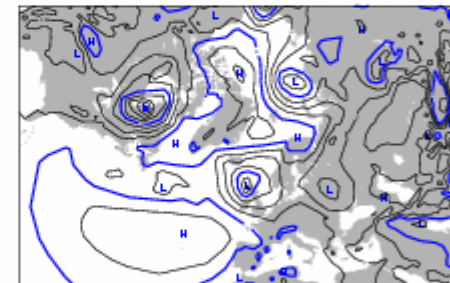
forecast t+120 VT:Monday 4 July 2011 00UTC
Cluster: 2(of 3), population: 15, repres. member: 1



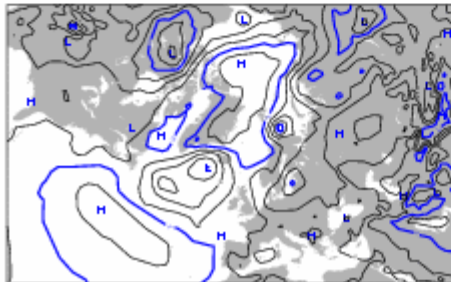
forecast t+144 VT:Tuesday 5 July 2011 00UTC
Cluster: 2(of 3), population: 15, repres. member: 1



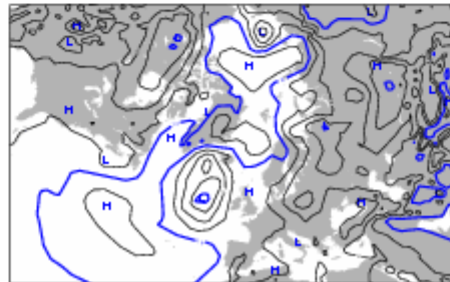
forecast t+168 VT:Wednesday 6 July 2011 00UTC
Cluster: 2(of 3), population: 15, repres. member: 1



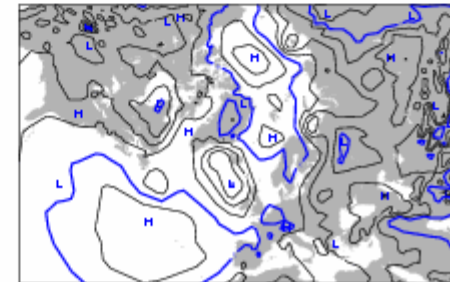
forecast t+120 VT:Monday 4 July 2011 00UTC
Cluster: 3(of 3), population: 14, repres. member: 28



forecast t+144 VT:Tuesday 5 July 2011 00UTC
Cluster: 3(of 3), population: 14, repres. member: 28



forecast t+168 VT:Wednesday 6 July 2011 00UTC
Cluster: 3(of 3), population: 14, repres. member: 28



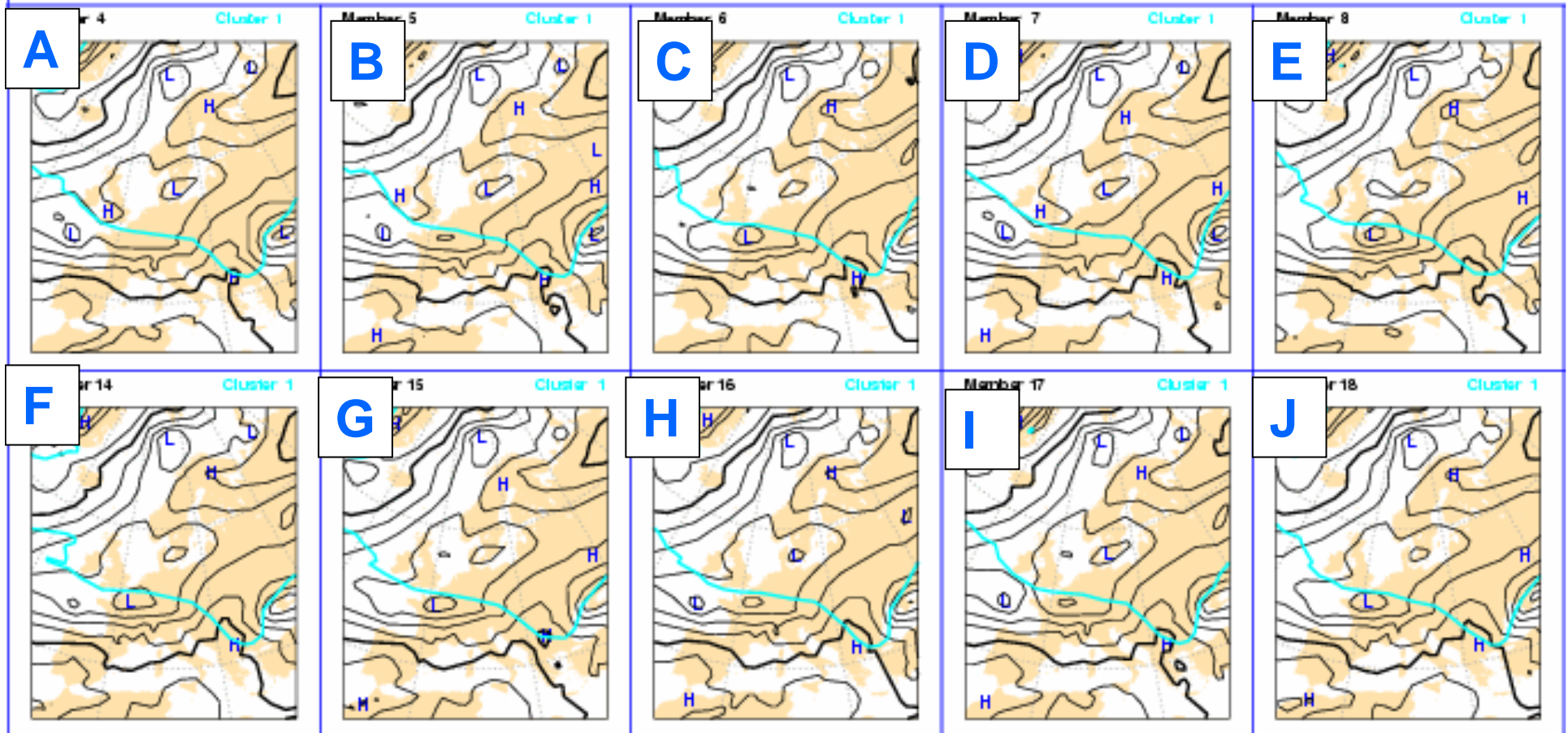
ECMWF mistakes:

1. Using their skill and resources to paint the T1279 in the most wonderful colours, making it more attractive and “available”
- 2. The play on the human weakness to confuse what is “typical” with what is “probable”**

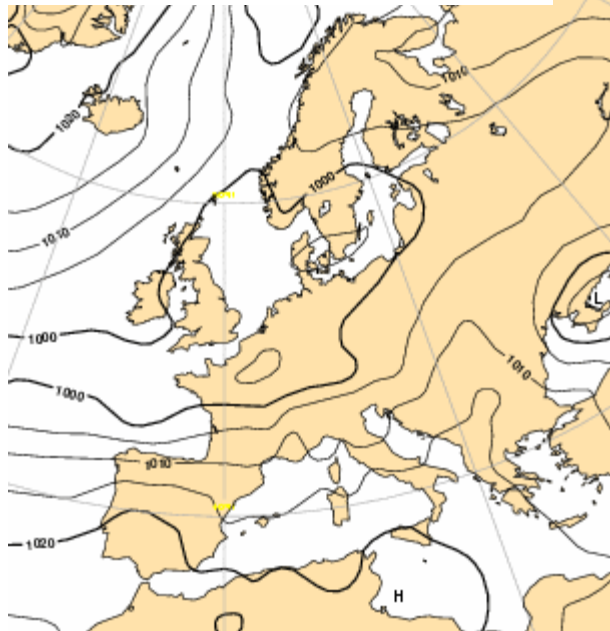
“Model of the Day?”

Can we pick the “model of the day” by judging from its performance during the first 12 or 24 hours?

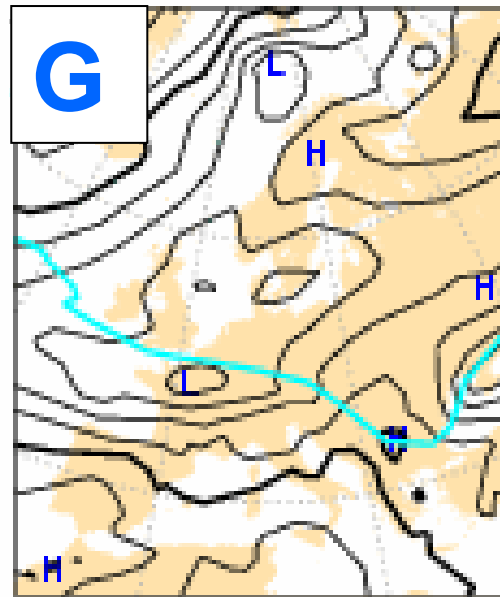
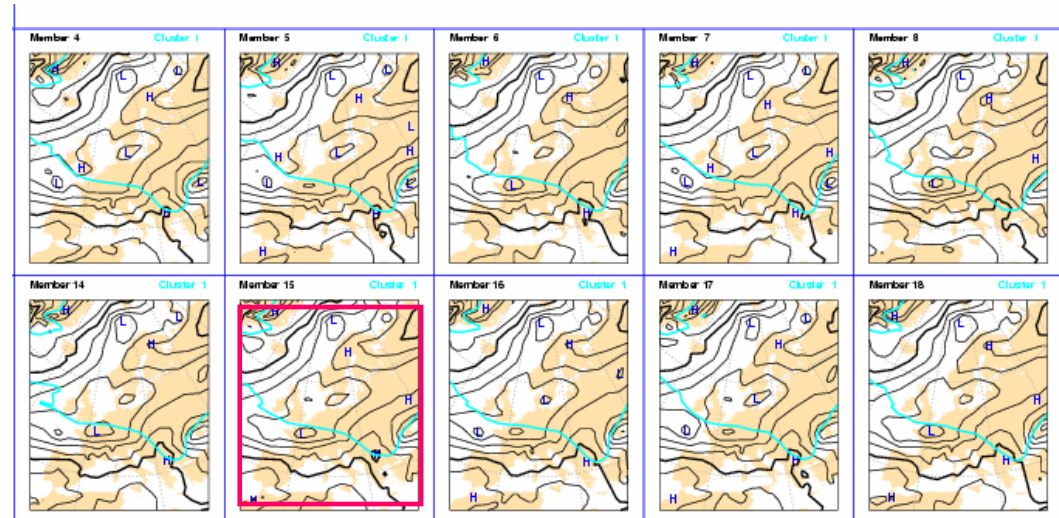
Imagine a set of 10 NWP 00 UTC +12 h forecasts from ten different centres



Verifying analysis at 12



12 hours into the forecast, model G seems to have succeeded best!

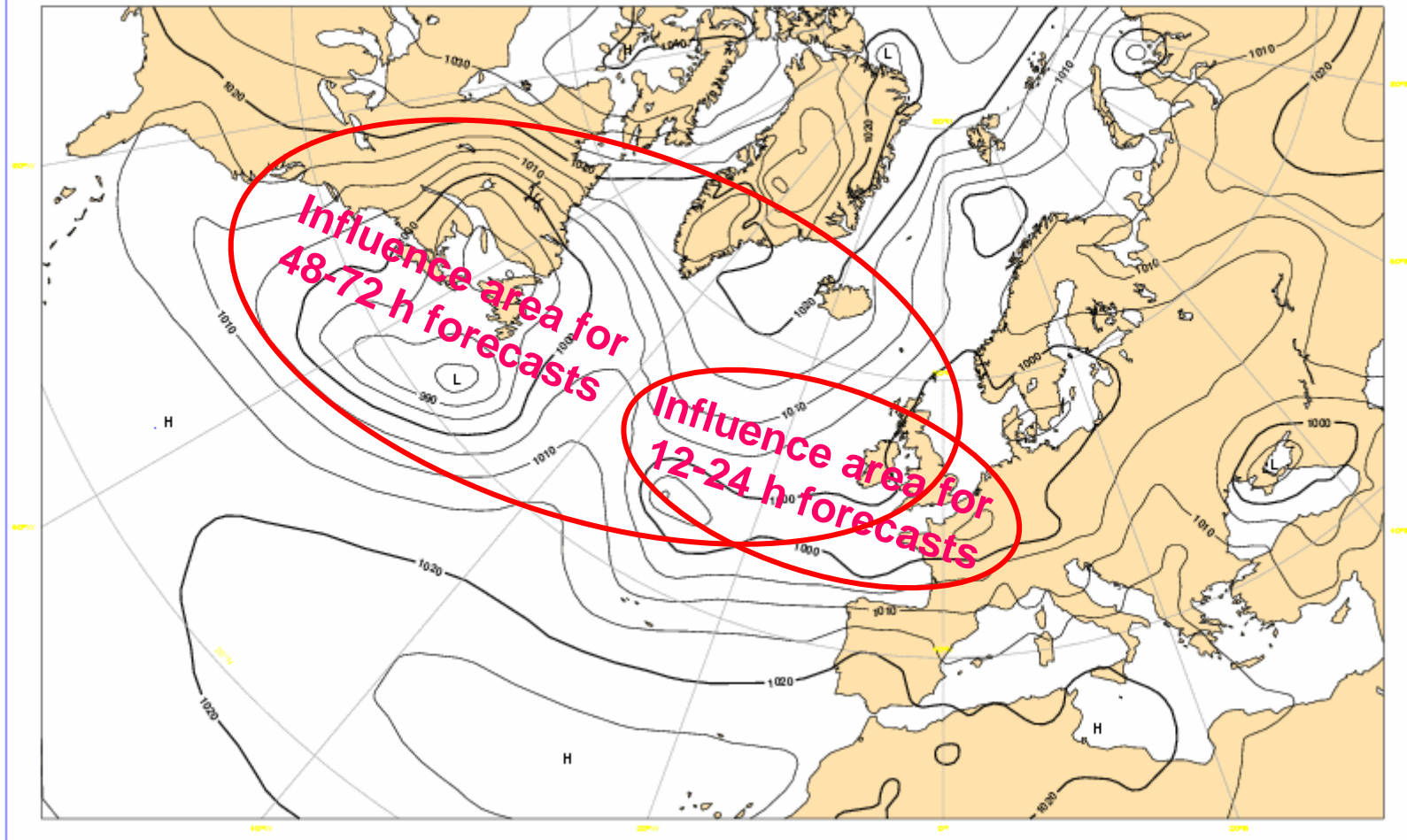


But this can only be used 12-18 hours into the future!

But the influences stretch far back upstream

Why?

ECMWF Analysis VT:Friday 3 March 2006 12UTC Surface: mean sea level pressure

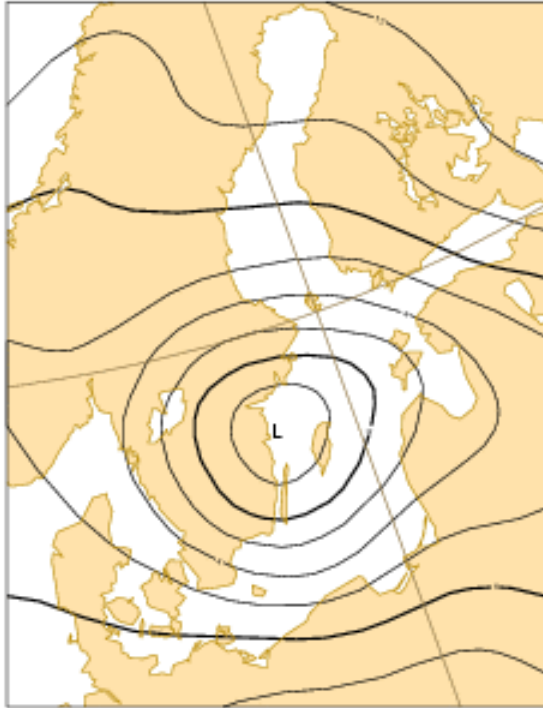


ECMWF mistakes:

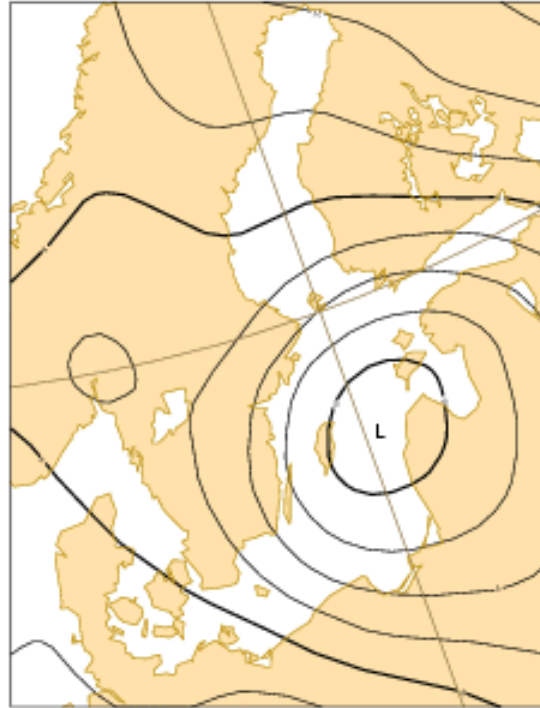
1. Using their skill and resources to paint the T1279 in the most wonderful colours, making it more attractive and “available”
2. The play on the human weakness to confuse what is “typical” with what is “probable”
- 3. To play on the misconception that it is possible to select the “model of the day” or the “best member”.**

Sometimes a 10-day forecast can be as good as a 1-day – as in 1991

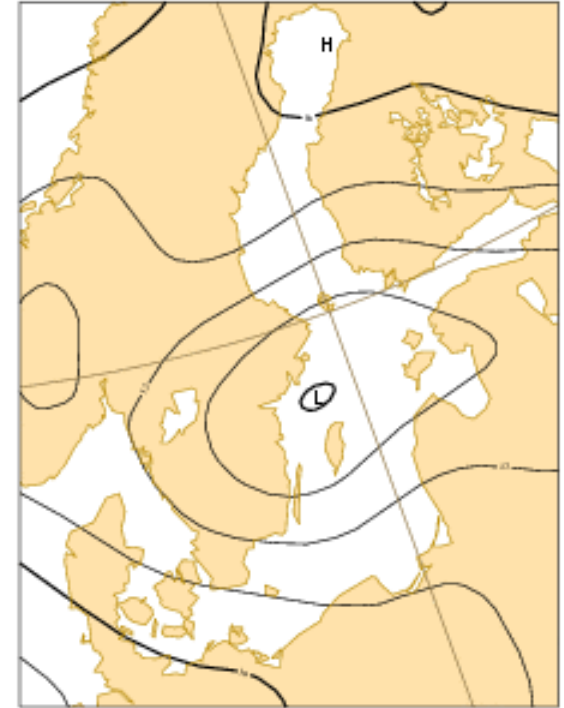
Z 1000 hPa 18 June 1991 18 UTC



Z 1000 hPa 17 June 1991 12z+30h

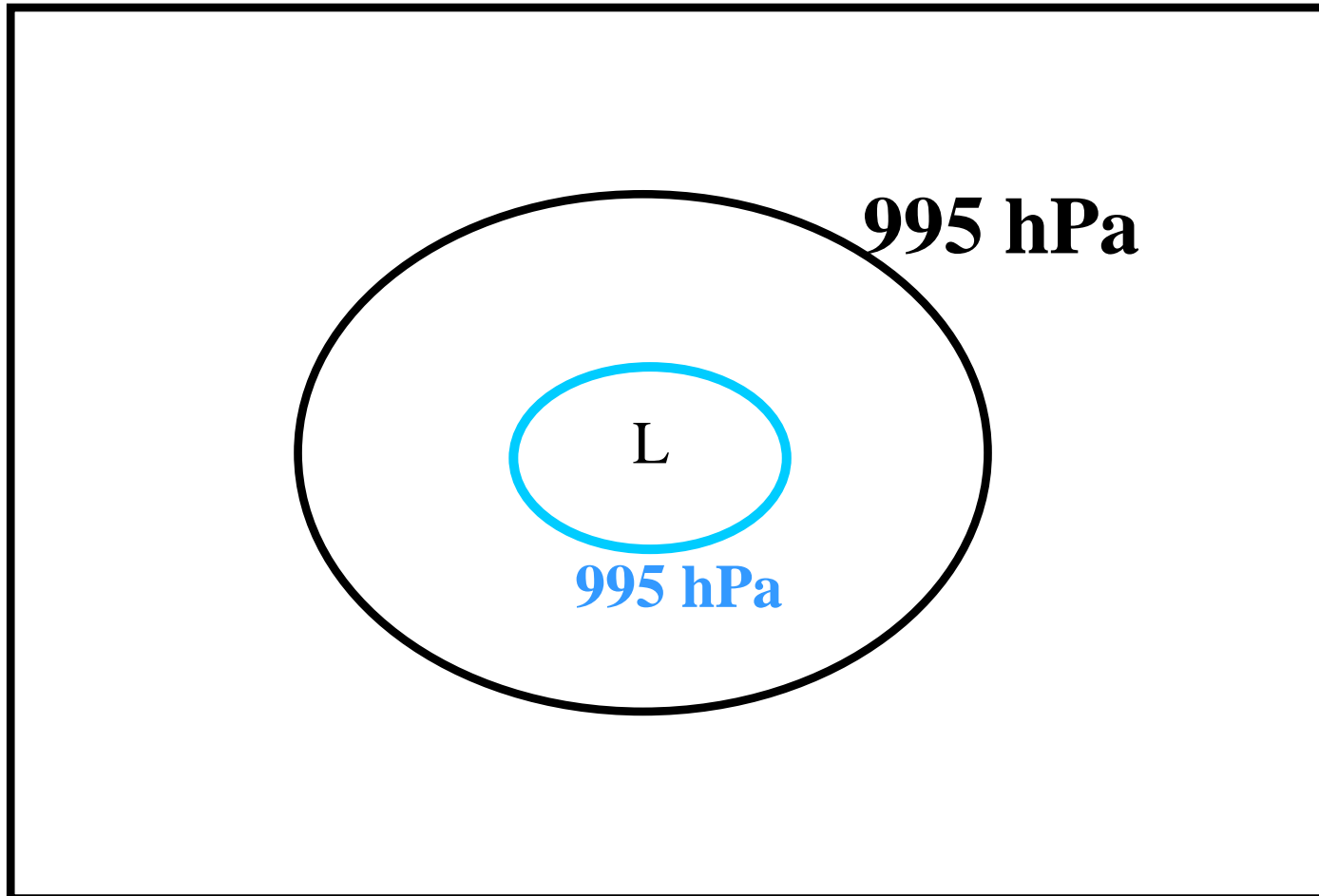


Z 1000 hPa 8 June 1991 12z+240h



















V.1.4 The mean - again

Are we systematically underestimating the Icelandic low?
The mean of the forecasts look weaker than the analysis.

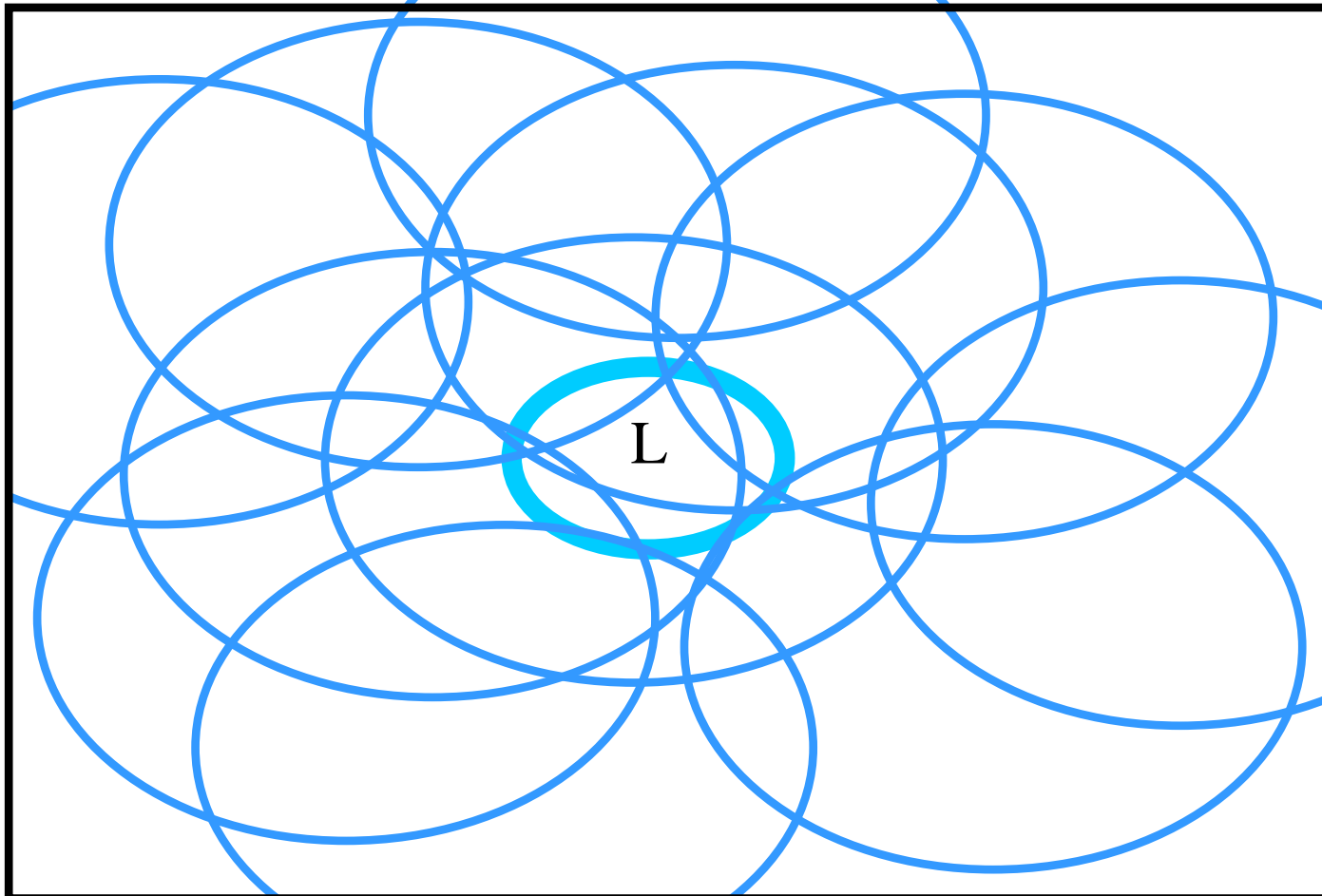


Are we systematically underestimating the Icelandic low?

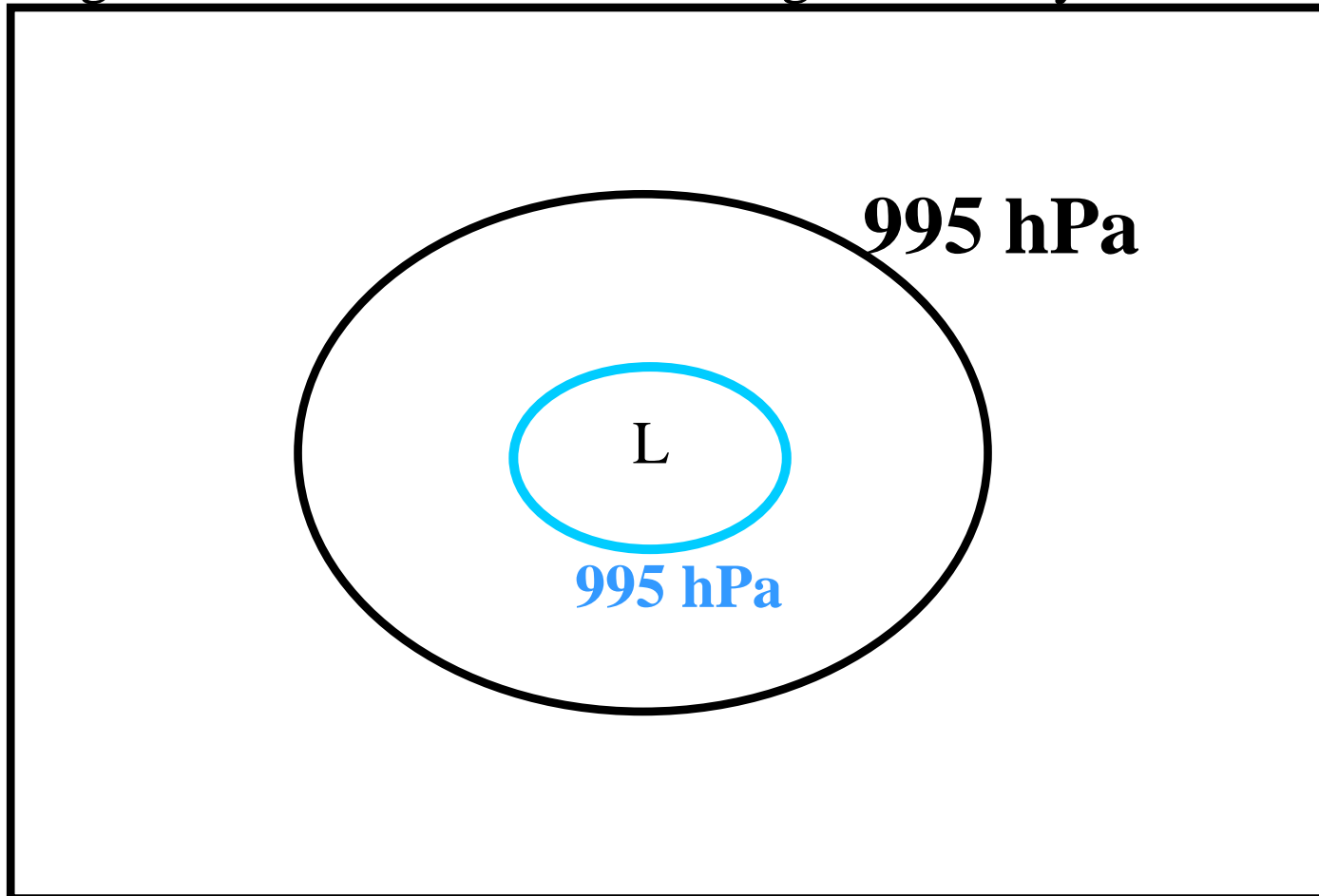
Fifteen different forecast get the intensity right, but not position

Analysis 			
			
			
			

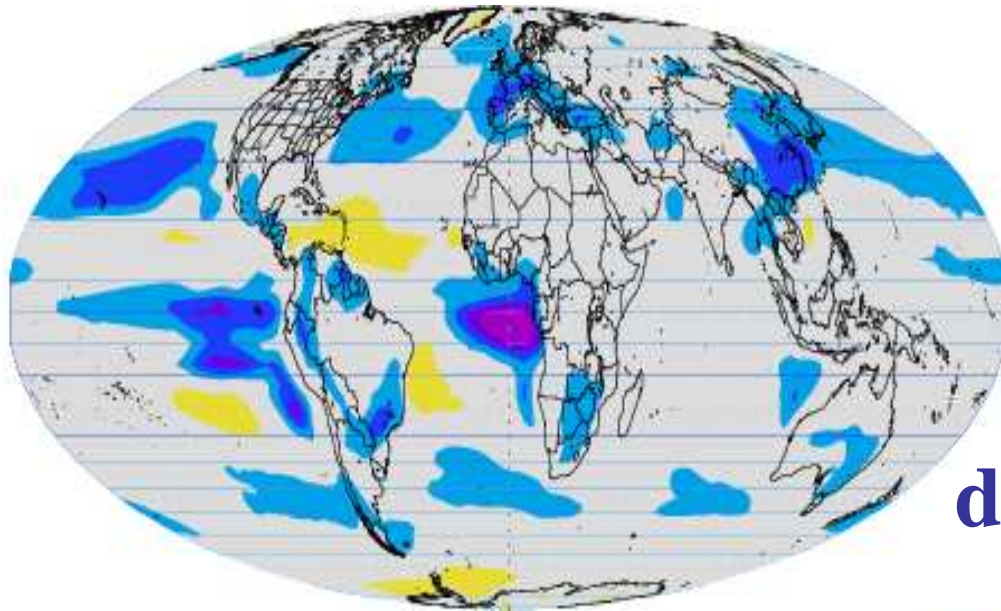
The mean of the forecasts look weaker than the analysis although the forecast all had the right intensity



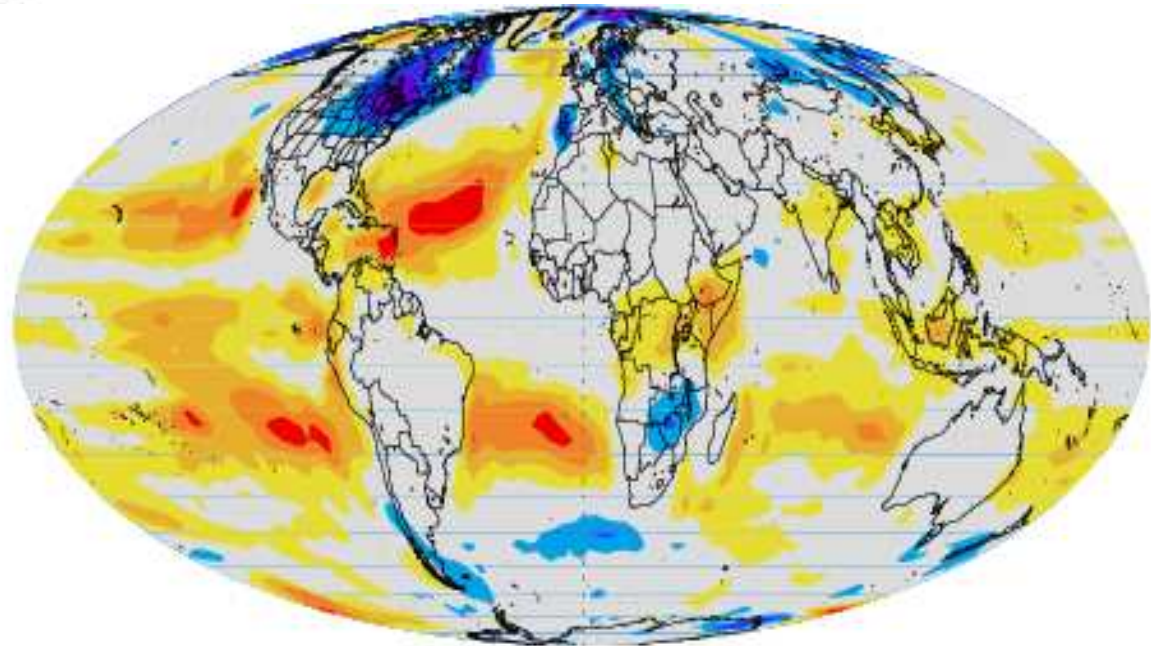
Because of the (non-systematic) positions errors the mean of the forecasts looks (systematically) weaker than the analysis although all the forecasts had the right intensity



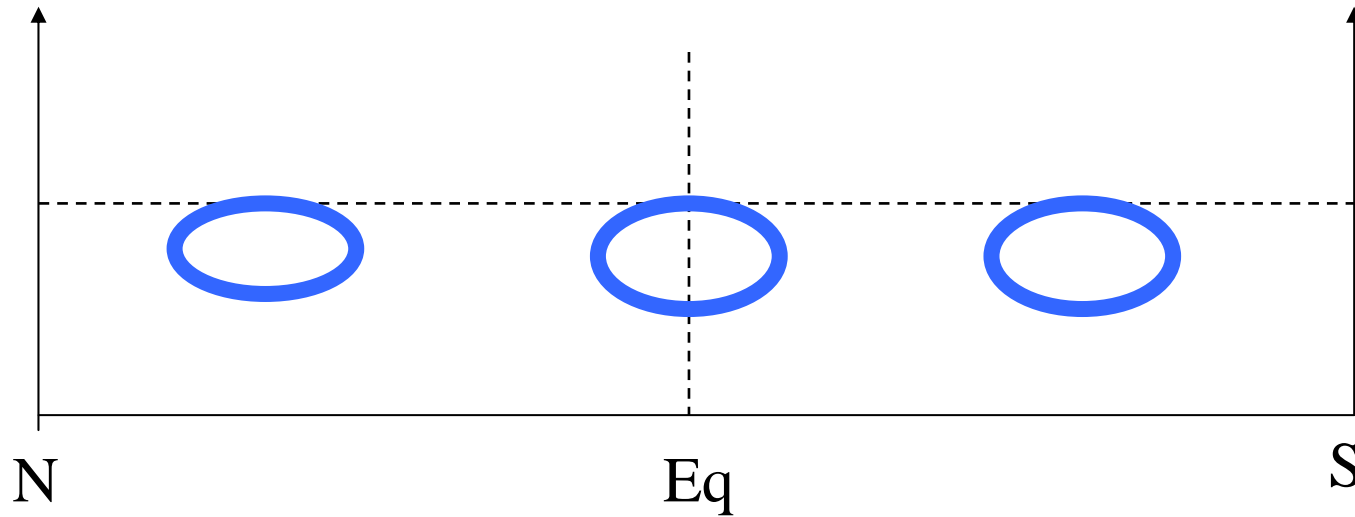
V.1.5 See how a “bias” can make a difference



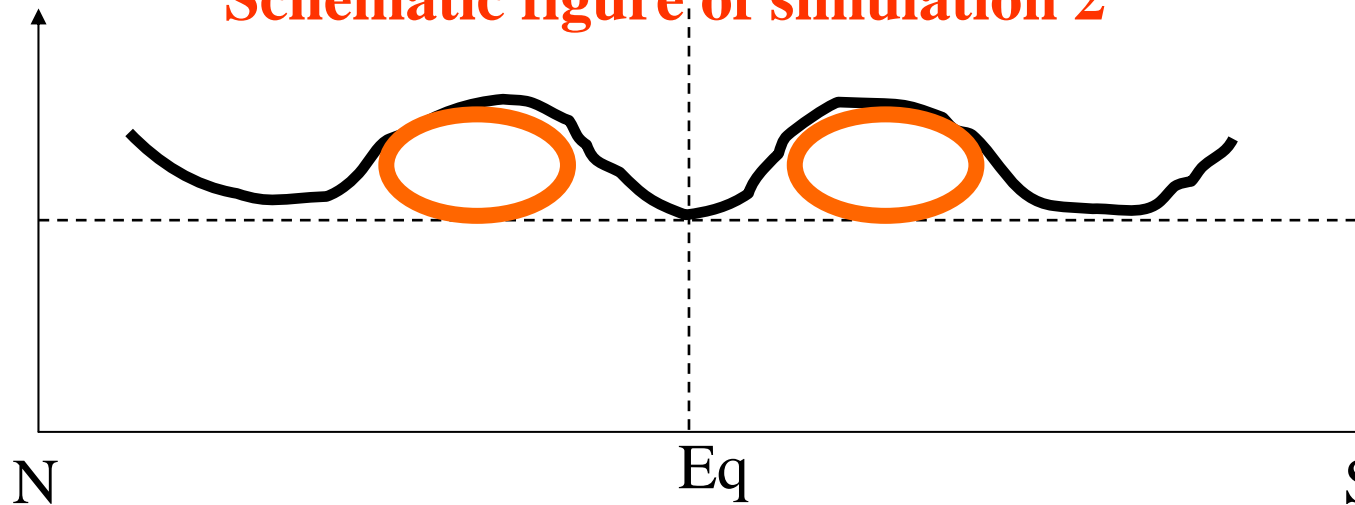
**Two global cloud
impact simulations
which look quite
different – or do they?**

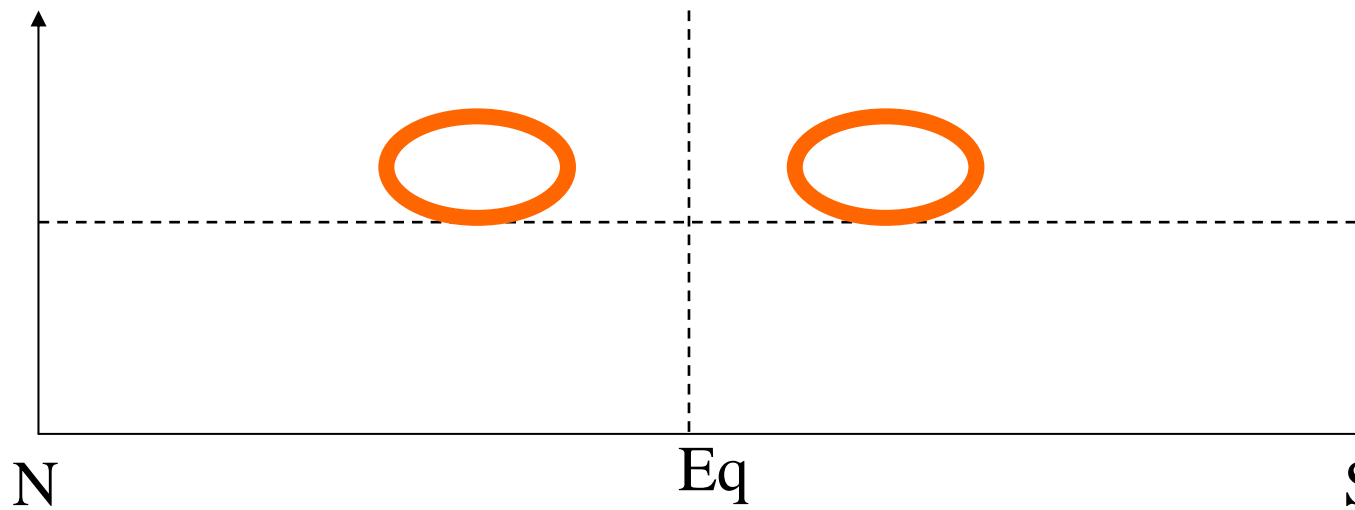
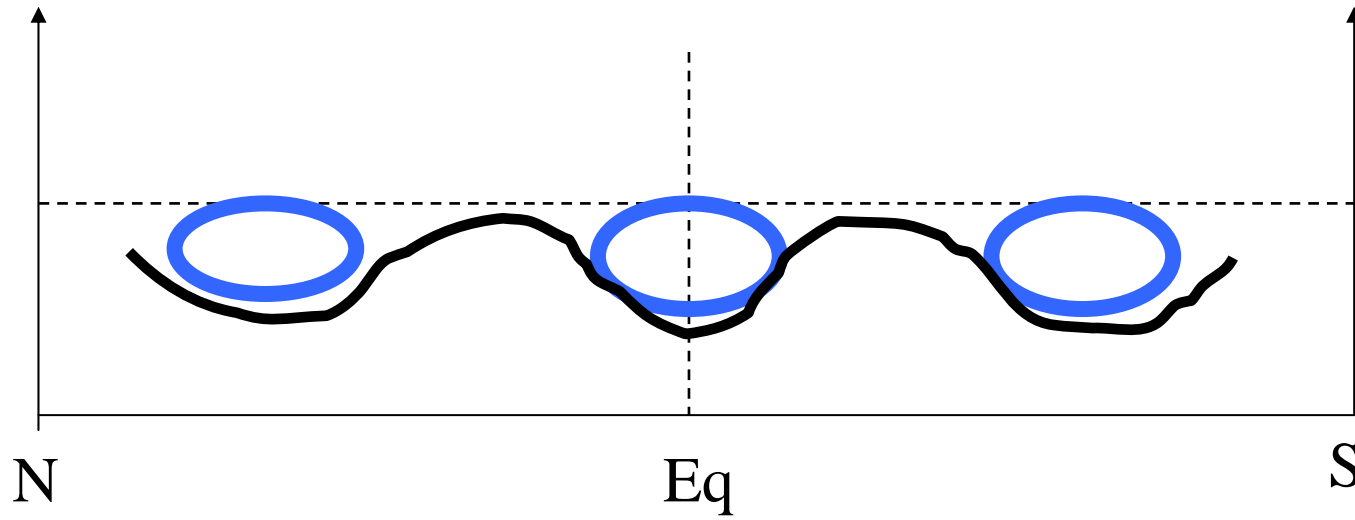


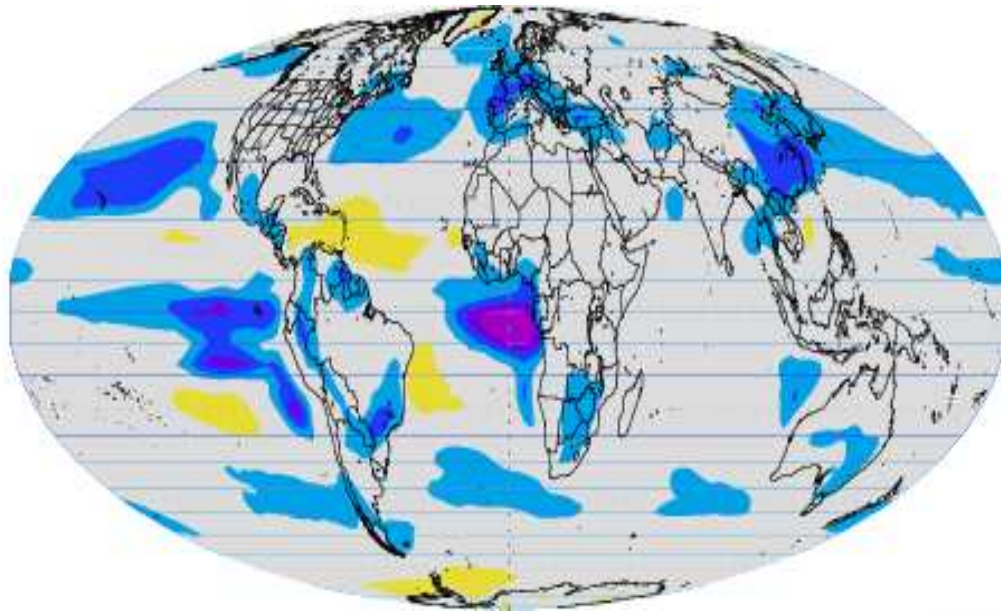
Schematic figure of simulation 1



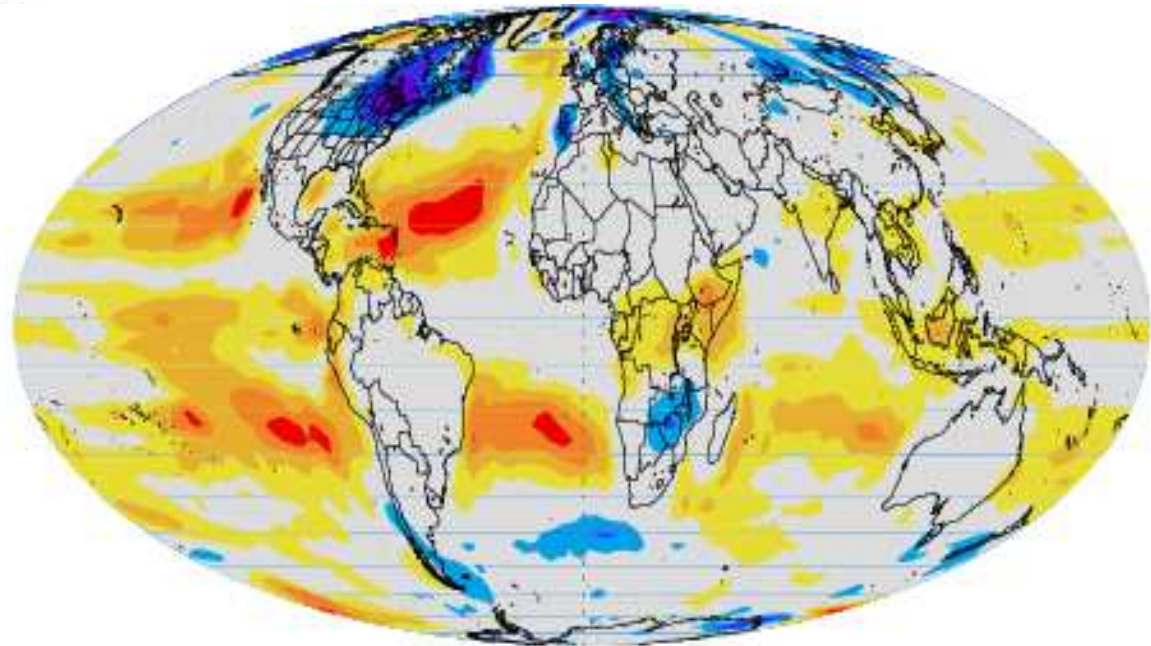
Schematic figure of simulation 2





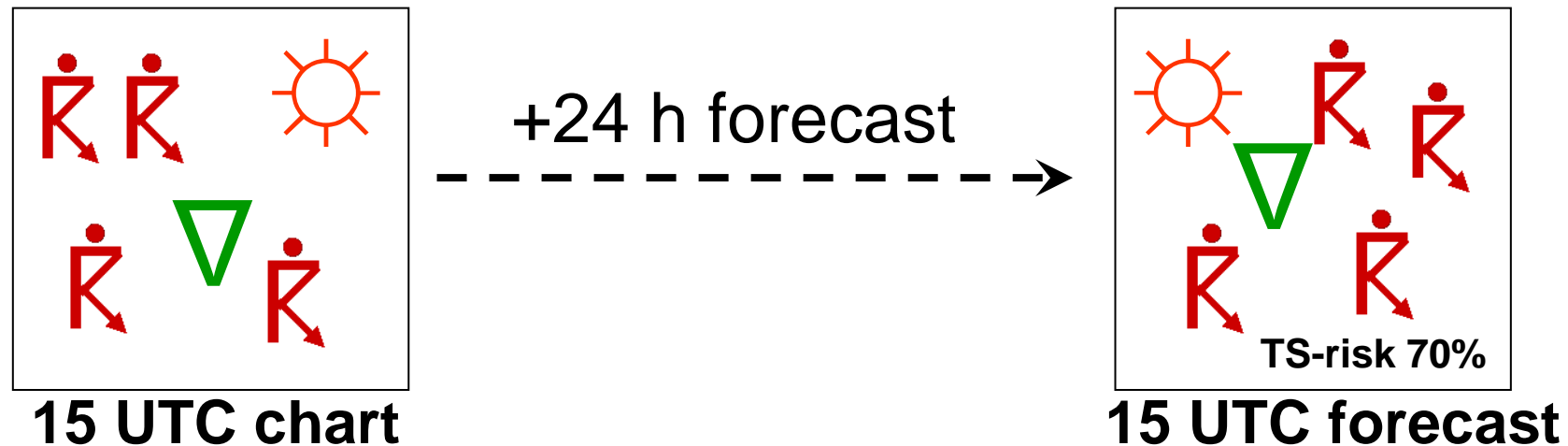


**No, they differed
just by a “bias”, a
true bias for once**



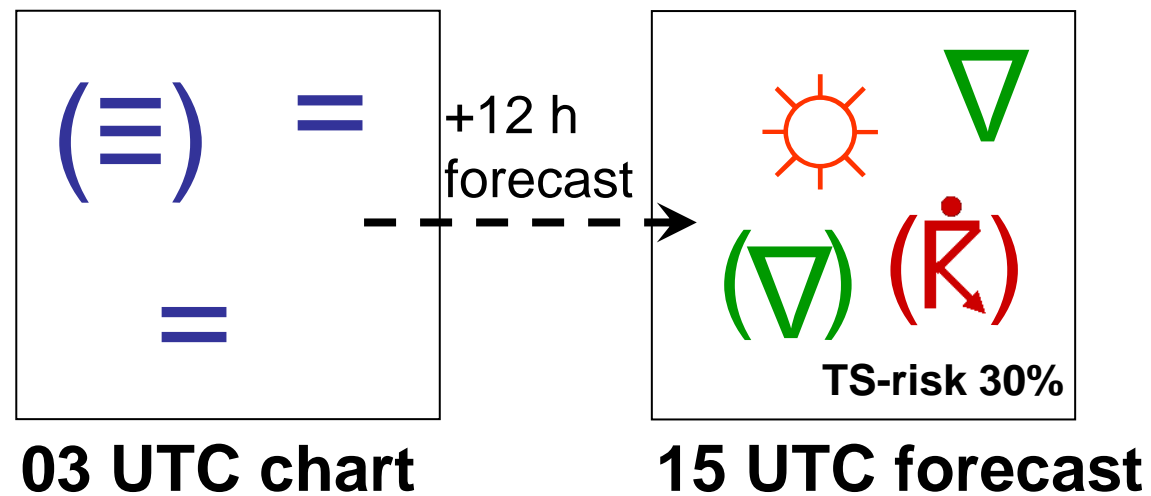
V.1.6 Availability effect

3. More about estimating uncertainty



Thunderstorms forecast 24 hours ahead are well forecast because they are “available” on the afternoon weather maps or radars

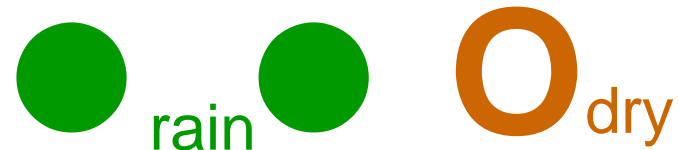
Thunderstorms forecast only 12 hours ahead are not as well forecast because they are not “available” on the early morning weather maps or radars



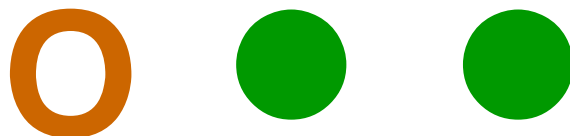
V.1.7 The primacy effect

The primacy effect

When you receive the NWP's in this order



..you might be more inclined to forecast rain than if you receive them in this order



V.1.8 Consistency, jumpiness

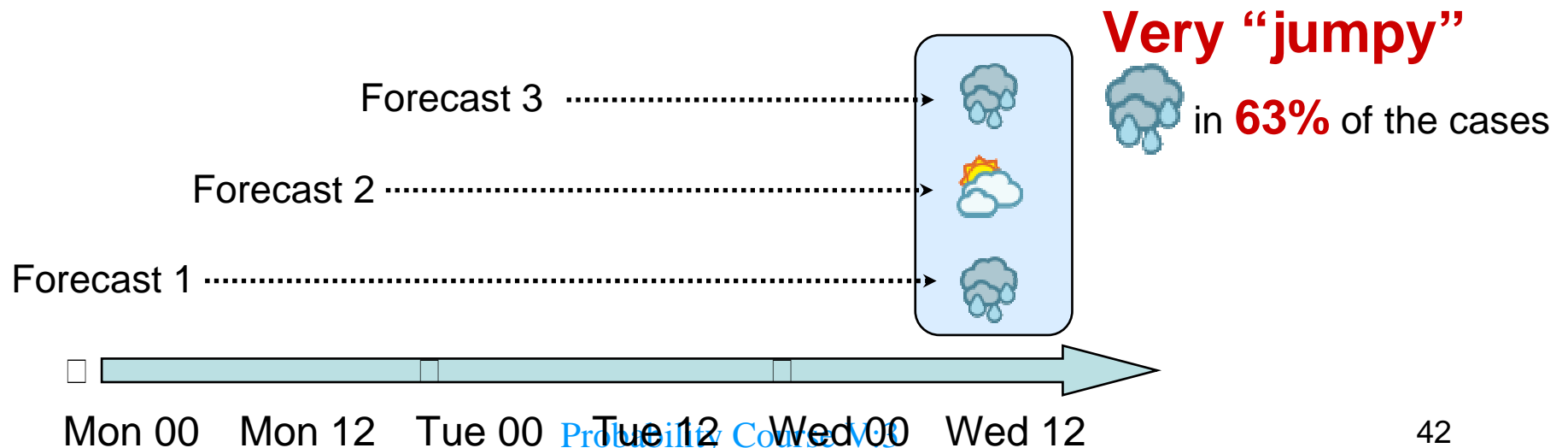
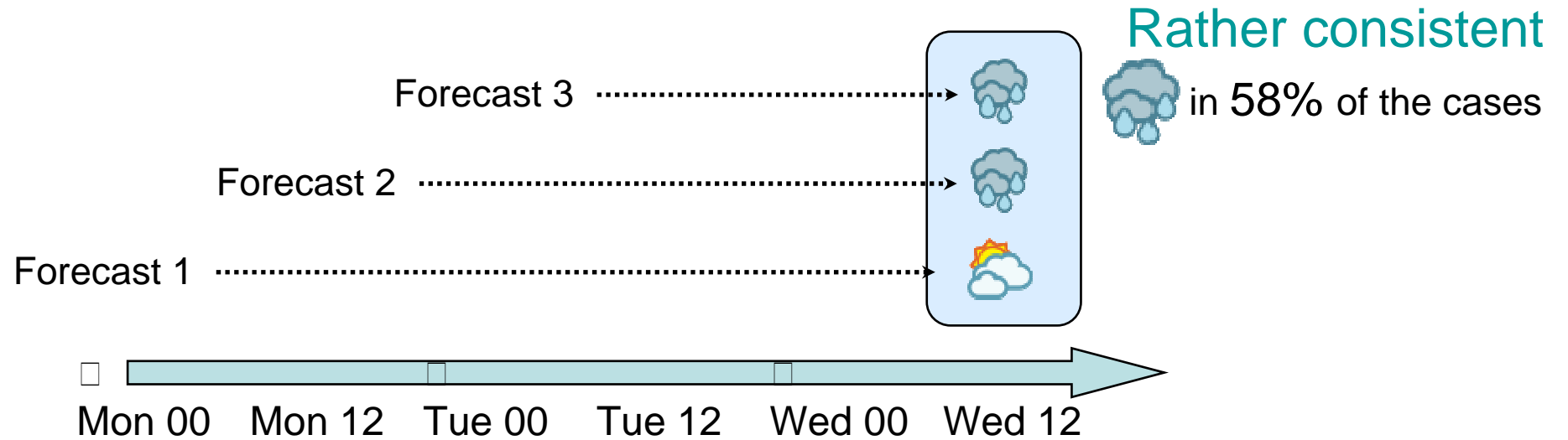
Over reliance in consistency

Tests conducted during ECMWF training courses 1993-99:

Consistency and skill	1993	1994	1999	1999	mean
D+3/4	-4%	-6%	29%	13%	8%
D+4/5	-14%	16%	20%	2%	6%
D+5/6	7%	-20%	3%	5%	0%
D+6/7	15%	-5%	-9%	-9%	-2%
D+7/8	-8%	-12%	-9%	-18%	-11%

Correlation $\pm 10\%$

Misinterpreted inconsistency in three consecutive runs from the same model

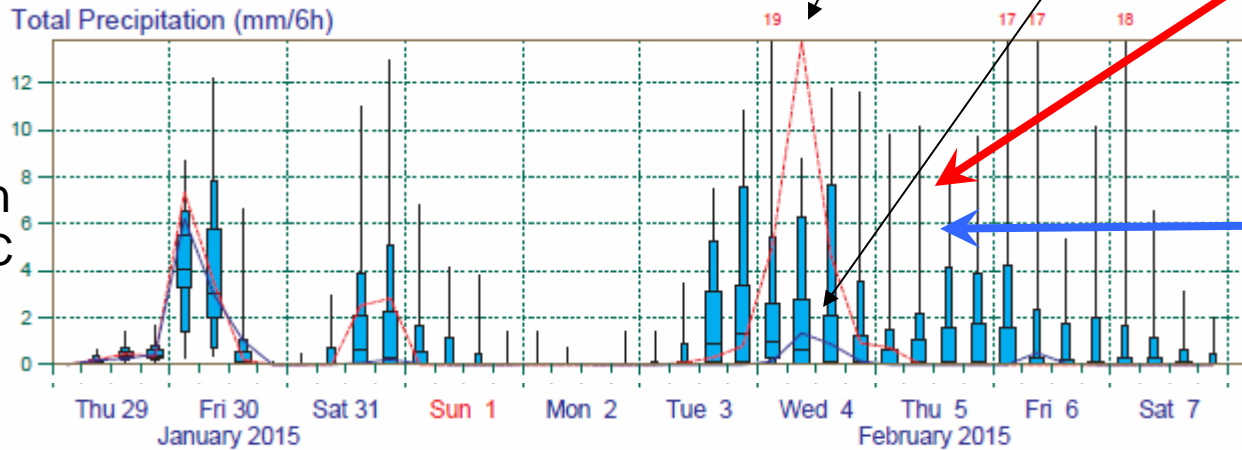


V.1.9 Confirmation bias

EPS gram for Bologna covering this week

In the 29 Jan forecast EPS Control has 14 mm, Ops only 1 mm

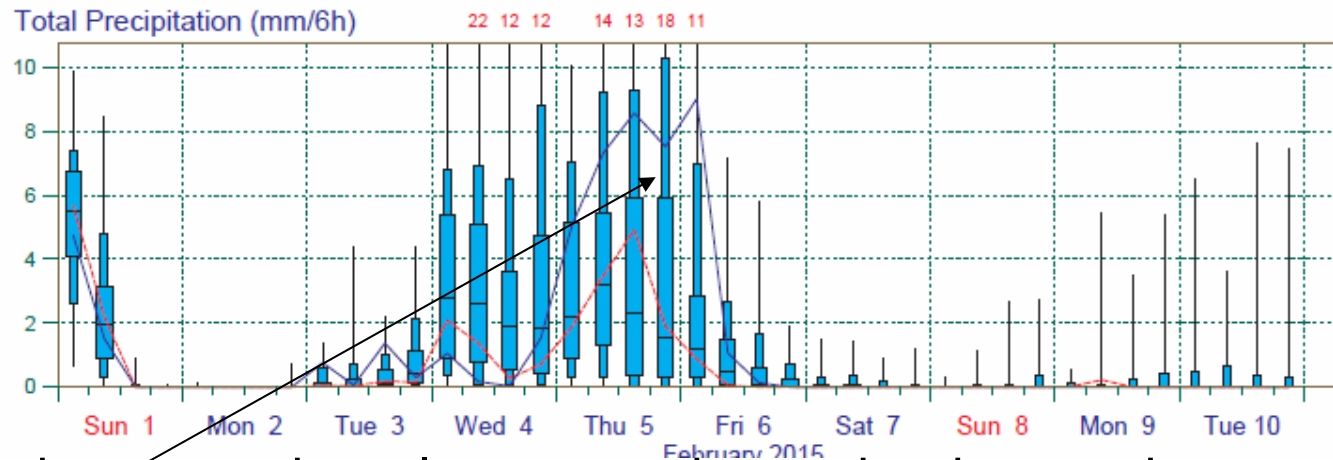
29 Jan
00 UTC



Can the EPS help us to **confirm** which one to trust?

There is a probability of rain which the Control and Ops may or may not support

1 Feb
00 UTC



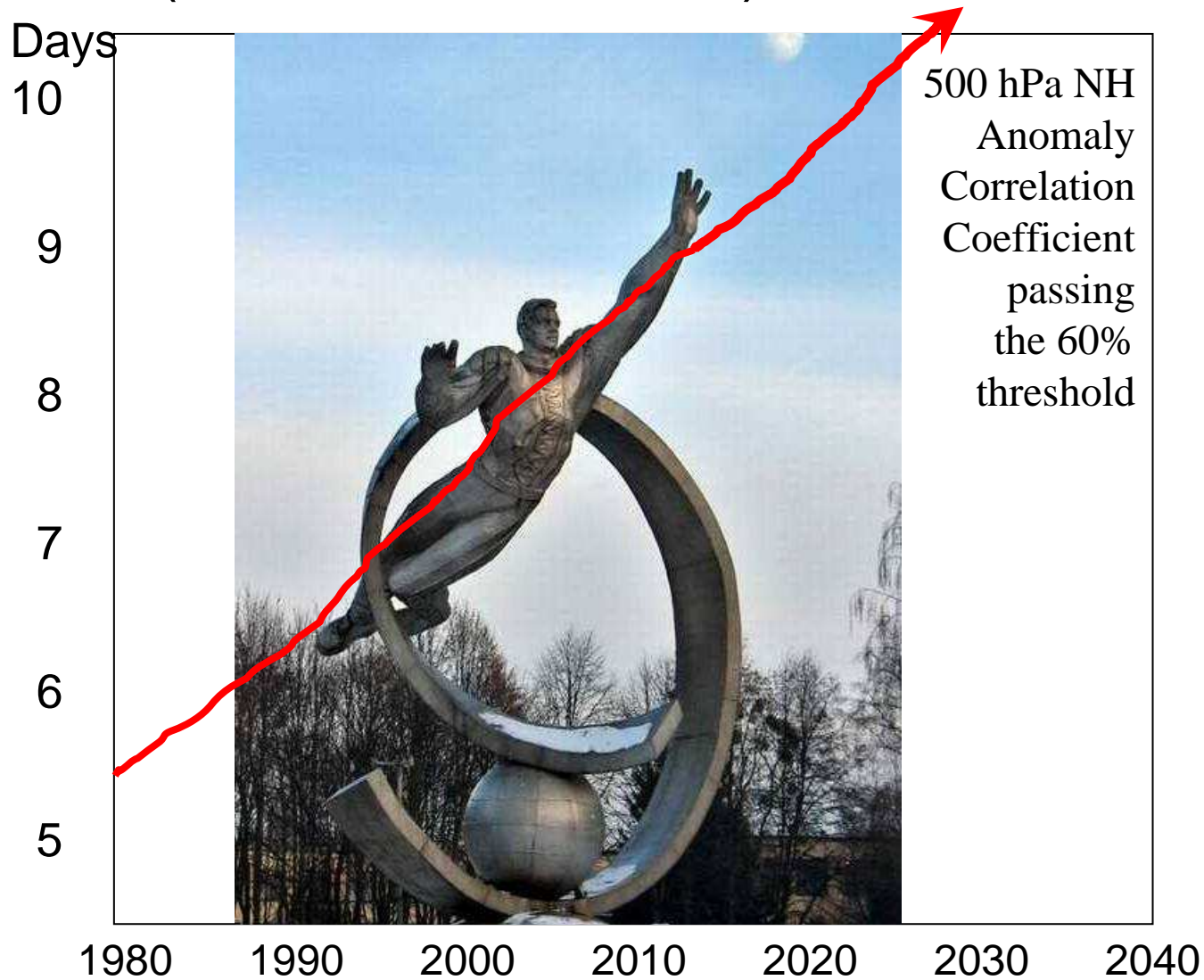
The EPS continues consistently to warn about rain whereas the Control and Ops have “jumped”, which they are “entitled” to do!

ECMWF mistakes:

1. Using their skill and resources to paint the T1279 in the most wonderful colours, making it more attractive and “available”
2. The play on the human weakness to confuse what is “typical” with what is “probable”
3. To play on the misconception that it is possible to select the “model of the day” or the “best member”.
- 4. The rationale when EPS started in 1992 was that it should be used only to estimate the credibility of the operation model**

V.1.10 Substitution

Substitution: If it is difficult to A, you forecast B which is easier (A=rain, B=500 hPa).



ECMWF mistakes:

1. Using their skill and resources to paint the T1279 in the most wonderful colours, making it more attractive and “available”
2. The play on the human weakness to confuse what is “typical” with what is “probable”
3. To play on the misconception that it is possible to select the “model of the day” or the “best member”.
4. The rationale when EPS started in 1992 was that it should be used only to estimate the credibility of the operation model
- 5. To put too much emphasis on the ACC of 500 hPa deterministic forecasts**

END