We are surprisingly skillful yet they call us liars: On accuracy versus skill in the weather forecast production process.

Martin Göber, Thomas Kratzsch
Deutscher Wetterdienst
Basic Services

Acknowledgement:
K. Balzer, I. Scholz, G. Schweigert, S. Tremmel

(Outline) **Summary**

(1) *Given the **risk**,*

(3) *meteorologists have more **skill**,*

(2) *than is **perceived**.*

*Perception = a priori difficulty * skill (+ eps)*

Basis of weather forecasting

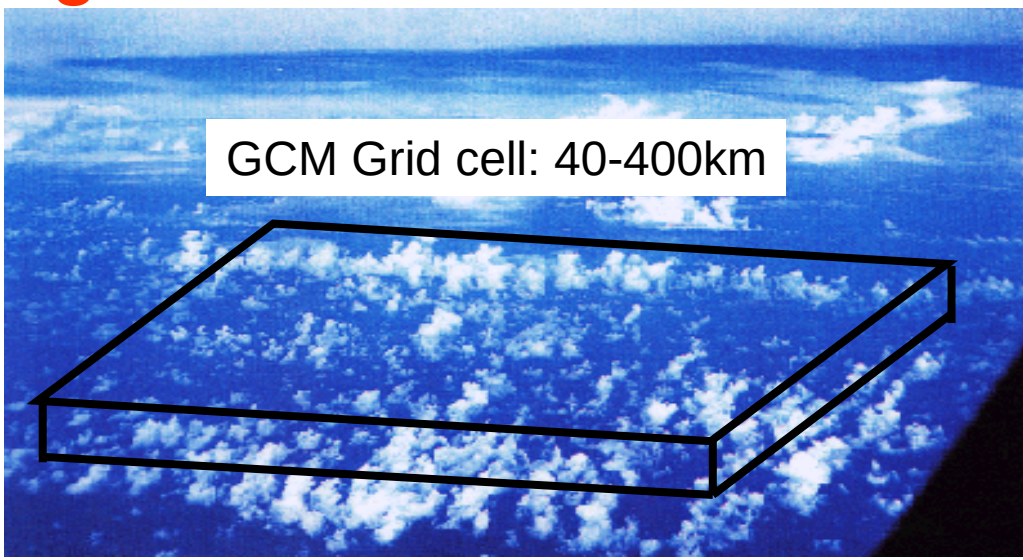
1) Knowledge

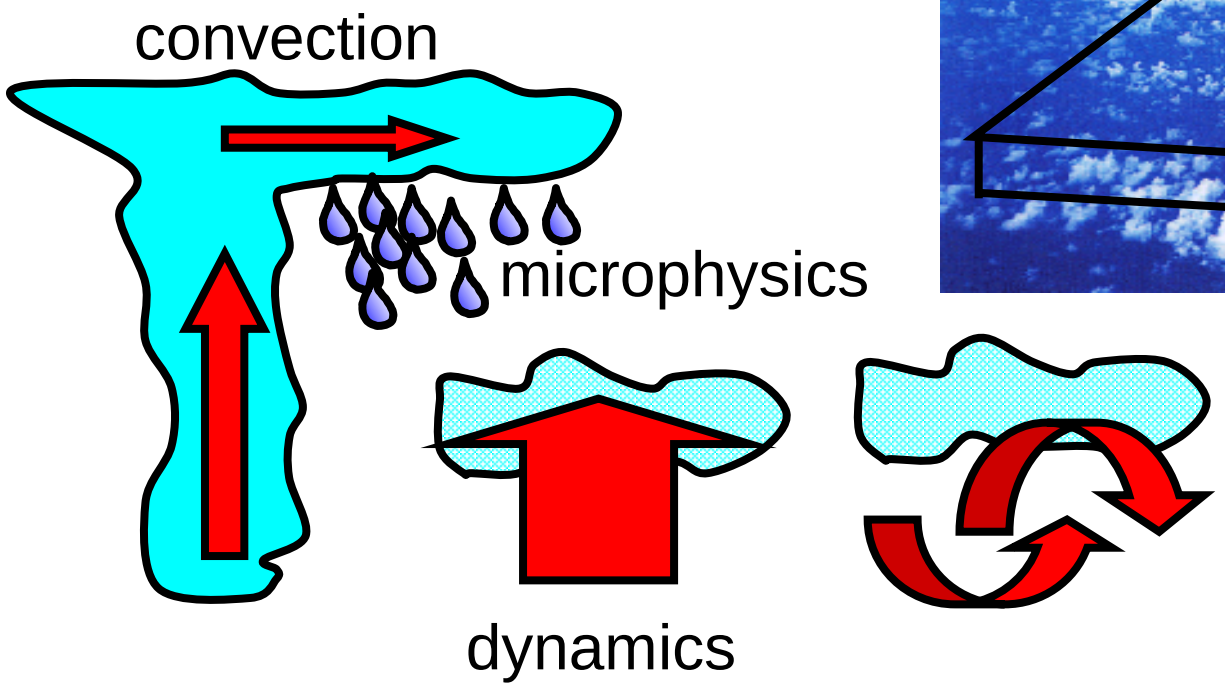
2) Observations

3) Computer programs

Basis of weather forecasting (1)

Knowledge





Courtesy Adrian Tompkins, 2004

Basis of weather forecasting (1)

Knowledge

a : cloud coverage

i)
$$a = \begin{cases} 0 : q_v < q_s \\ 1 : q_v \geq q_s \end{cases}$$
Fowler, Randall, 1996

ii)
$$a = 1 - \sqrt{\frac{1 - RH}{1 - RH_{critical}}}$$
Sundquist, 1989

iii)
$$a = f(RH, q_L)$$
Xu, Randall, 1996

iv)
$$a = \int_{q_s}^{\infty} G(q_t)\, dq_t$$
X, Y, Z, >1990

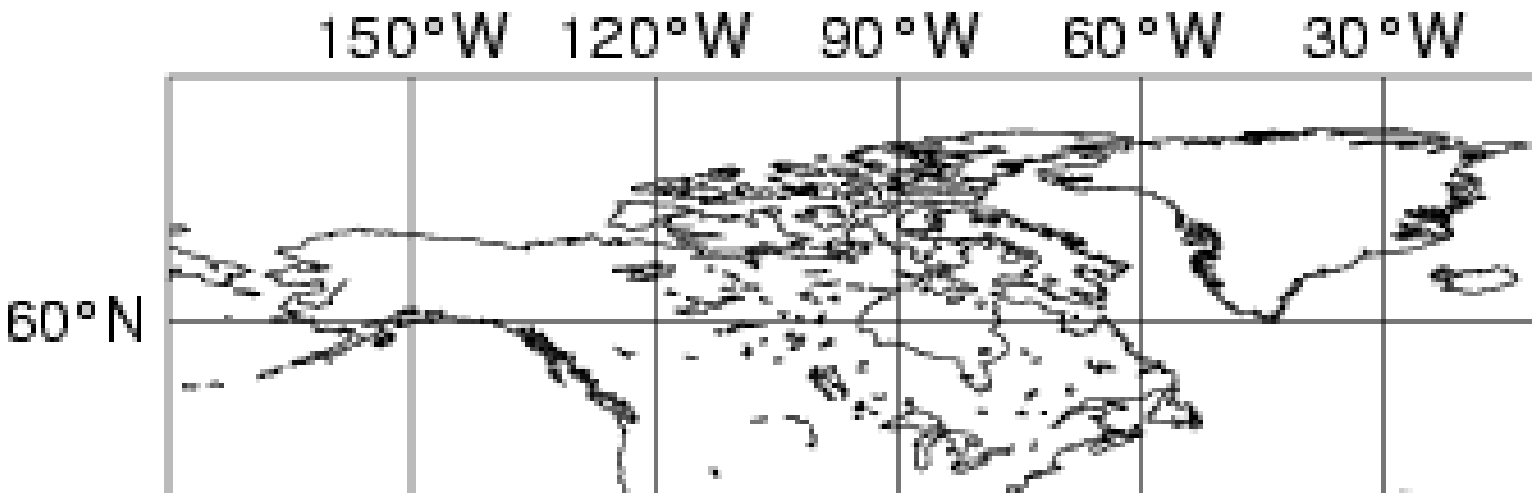
v)
$$a_m = \begin{cases} 0 & \omega \geq 0 \\ a_m^{*}\, \omega / \omega_{crit} & \omega_{crit} \leq \omega < 0 \\ a_m^{*} & \omega < \omega_{crit} \end{cases}$$

$$a_m^{*} = \left[\max\left(\frac{RH - RH_{crit}}{1 - RH_{crit}}, 0 \right) \right]^2$$
Slingo, 1990

After Tompkins, 2004

Basis of weather forecasting (2)

Observations



18 UTC Radiosondes

Basis of weather forecasting (3)
Programs

```
            PROGRAM Perfect_weather_forecast
            PARAMETER i_country=1
            CHARACTER country=[„U.K“, „Italy“]
            CHARACTER default_weather=[„bad“, „nice“]

1           IF obs=„nice“ THEN GOTO 999
            IF obs=„bad“ THEN GOTO 42
999         CONTINUE
42          PRINT *,“The weather will be “, default_weather
10          GOTO 42
            END
```

Basis of weather forecasting

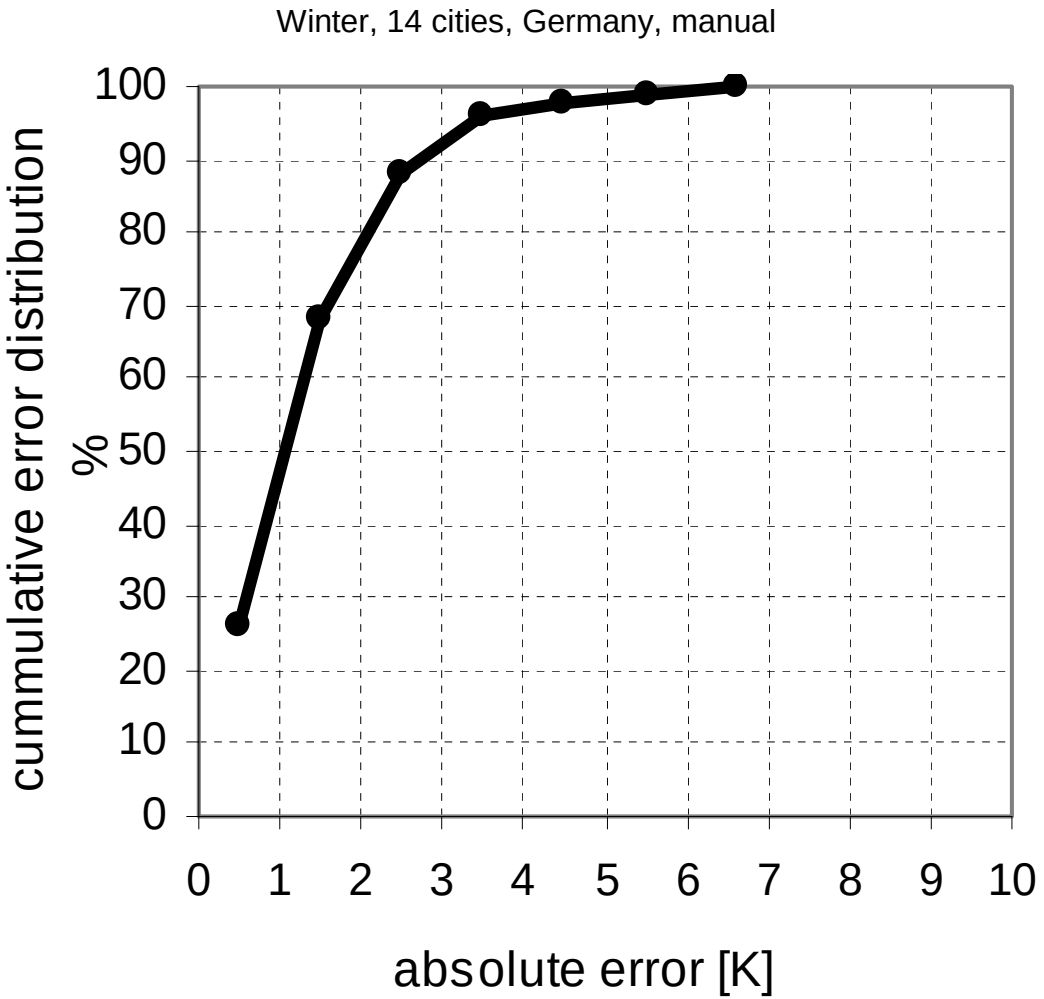
1) (almost) **no** knowledge

2) (almost) **no** observations

3) (almost) **only bugs**

What can we do on **that basis** ?

Error maximum temperature, next day

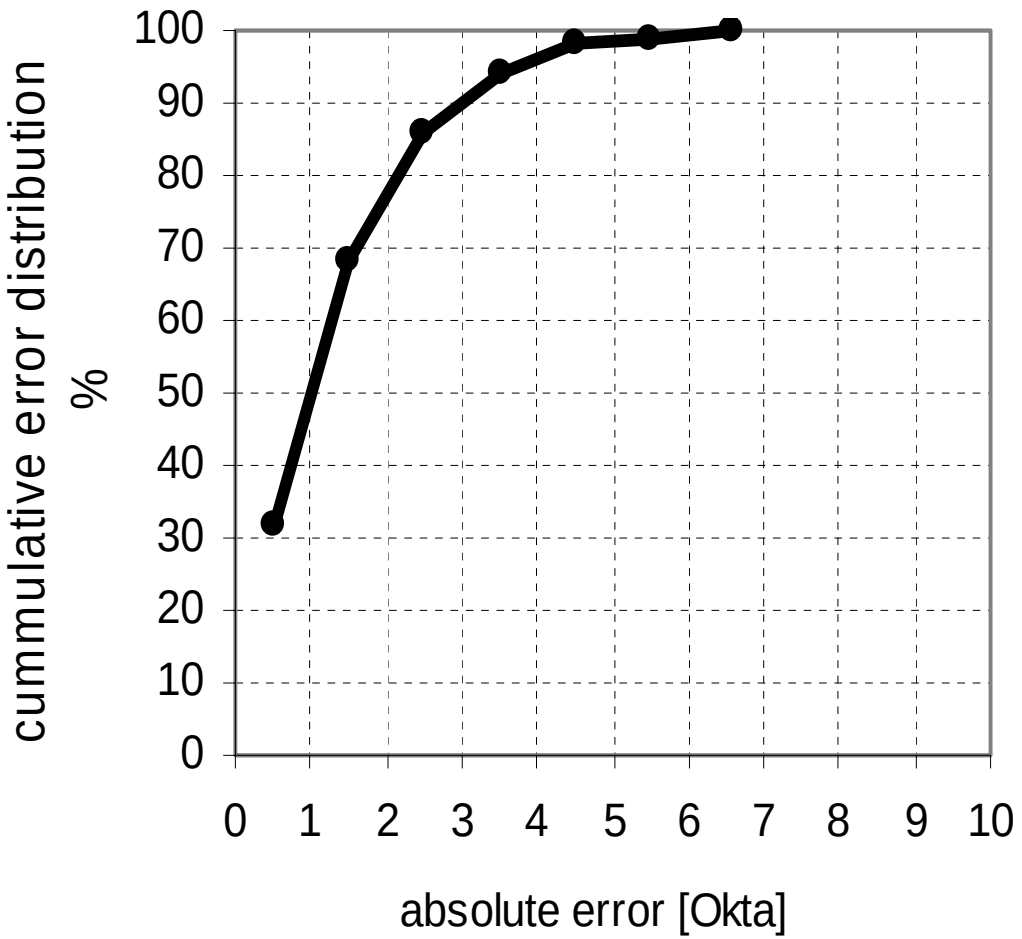


88% : ± 2 K
1% : > 5 K

320 days: content
3 days: angry

Error cloud amount, next day

Winter, 14 cities, Germany, manual

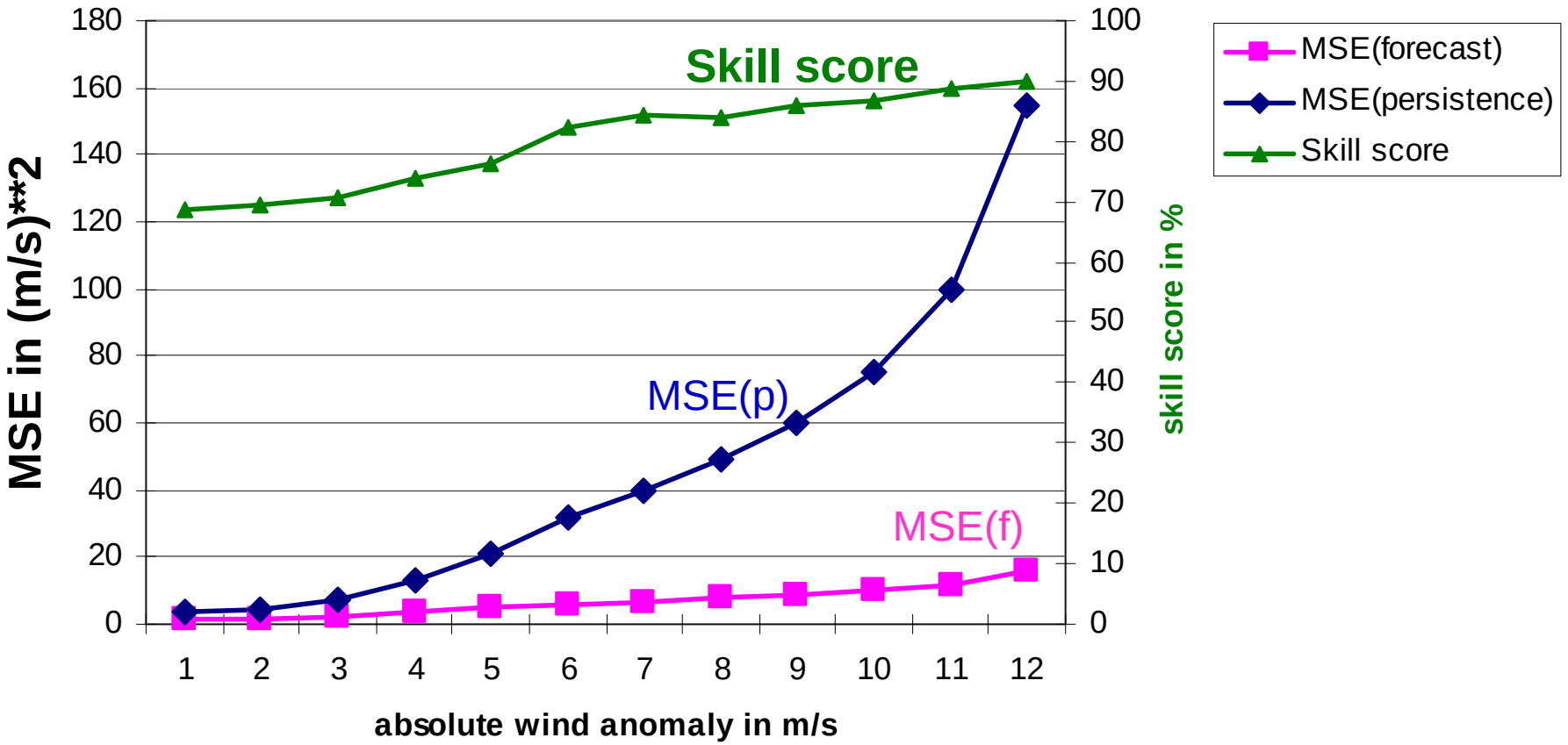


86% : ± 2 Okta

6% : > 3 Okta

315 days: content

20 days: angry


Skill score
MSE(p)
MSE(f)
MSE in (m/s)**2
skill score in %
absolute wind anomaly in m/s
MSE(forecast)
MSE(persistence)
Skill score
0
20
40
60
80
100
120
140
160
180
0
10
20
30
40
50
60
70
80
90
100
1
2
3
4
5
6
7
8
9
10
11
12

Lies, damn lies, statistics, ...weather forecasting?

Medical Screening

instrument („verifikation“)

- H=Hit rate= p (positive test | disease) = 70%
- F=false alarm rate= p (positive test | no disease) = 20%

Characteristic of disease („climate“)

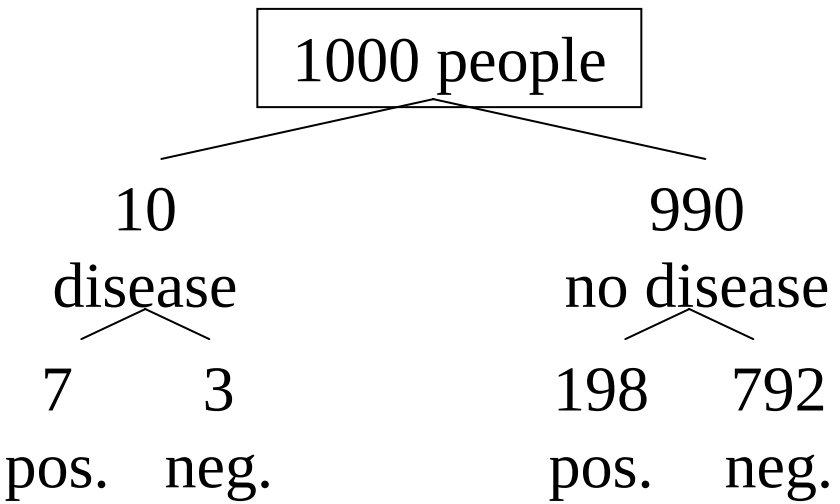
- base rate= p (disease) = 1%

Patient („Forecast“)

- detection rate= p (disease | positive test) = **???**

Instrument: H=70% F=20%

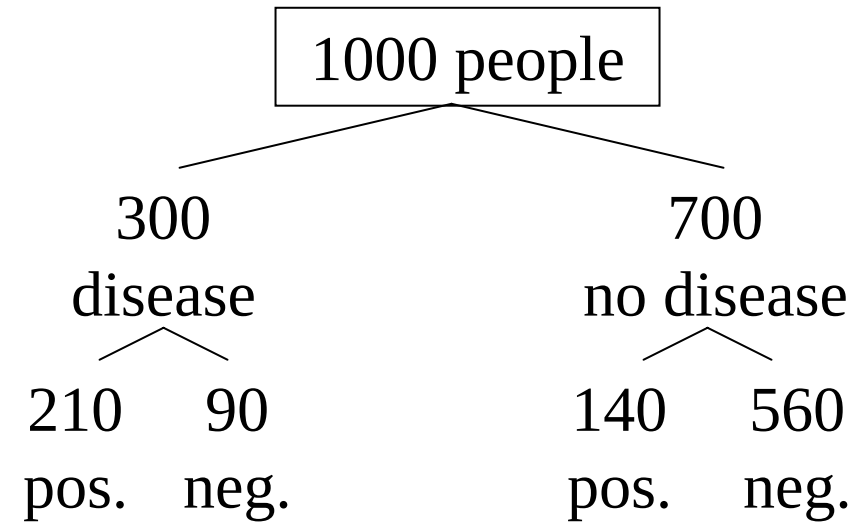
Rare disease
p(E)=**1**%



p(disease | positive test)= 7 / (7+198)

~ 3%

Frequent disease
p(E)=**30**%



p(disease | positive test)= 210 / (210+140)

~ 60%

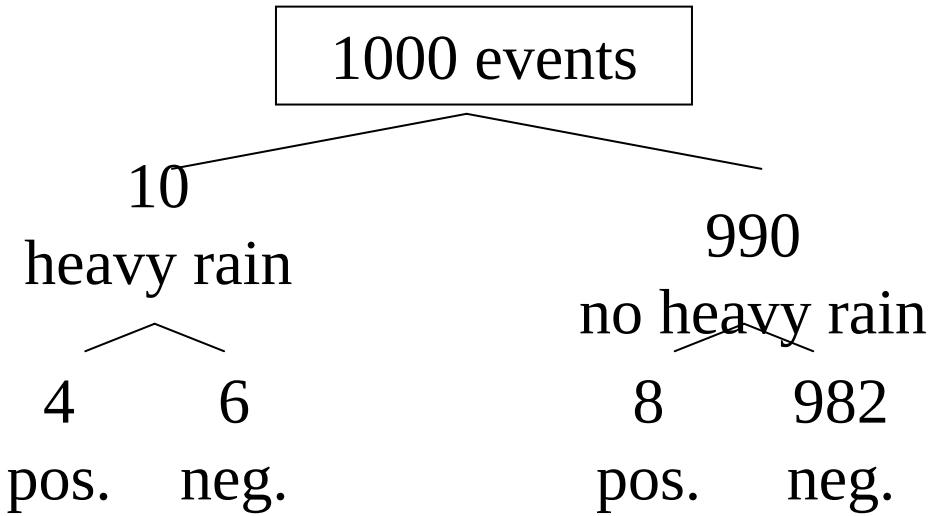
Weather-Screening

- Disease --> rain
- instrument --> weather forecasting process (**observation**+model...)
- Patient --> user

H=40% F=0.8%

Rare event

P(heavy rain)=1%



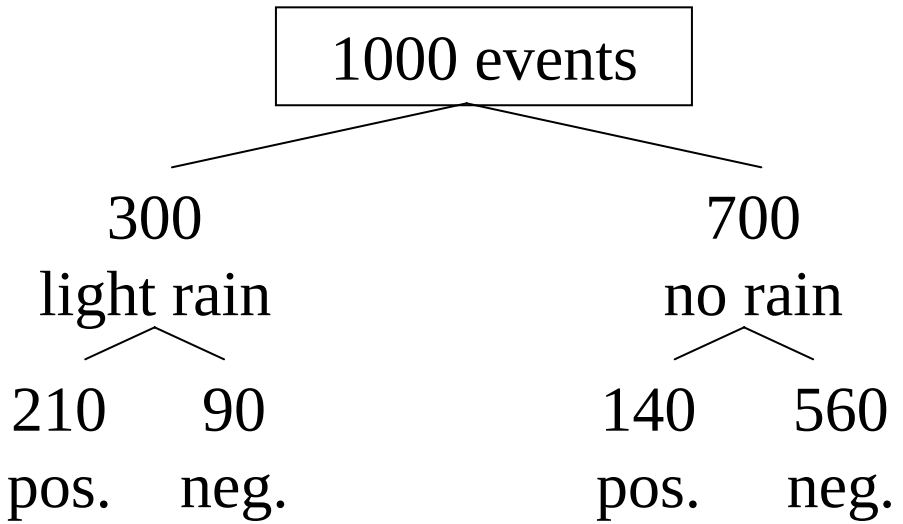
p(o=heavy rain | f=heavy rain)= 4 / (4+8)

~ 33%

H=70% F=20%

Frequent event

P(light rain)=30%



p(o=light rain | f=light rain)= 210 / (210+140)

~ 60%

$$\text{Odds (event)} = \frac{p(\text{event})}{p(\text{non-event})} = \frac{p(\text{event})}{1 - p(\text{event})}$$

$$\text{Odds (o=rain | f=rain)} = \text{climatological odds(rain)} * \frac{H}{F}$$

$$\text{Odds (o=no rain | f=no rain)} = \text{climatological odds (no rain)} * \frac{1-F}{1-H}$$

„Rain“ forecast

	odds (o=rain | f=rain)	=	Climatol. odds (rain)	*	H / F
Light rain	1:1	=	1:2	*	**2**
Heavy rain	1:3 (DMO) 1:2 (MOS) 1:5 (manual)	=	1:100	*	**30** **50** **20**

„Dry“ - forecast

	odds (o=no rain | f= no rain)	=	Climatol. odds (no rain)	*	(1-F) / (1-H)
Light rain	8:1 (DMO) 14:1 (MOS) 16:1 (manual)	=	2:1	*	**4** **7** **8**
Heavy rain	200:1	=	100:1	*	**2**

„Rain and dry“ forecast

	Odds ratio	**=**	**Climatological odds**	*****	**H*(1-F) / [F*(1-H)]**
Light rain	8:1 (DMO) 14:1 (MOS) 16:1 (manual)	=	1:1	*	**8** **14** **16**
Heavy rain	60:1 (DMO) 100:1 (MOS) 40:1 (manual)	=	1:1	*	**60** **100** **40**

Summary

(1) *Given the **risk**,*

(3) *meteorologists have more **skill**,*

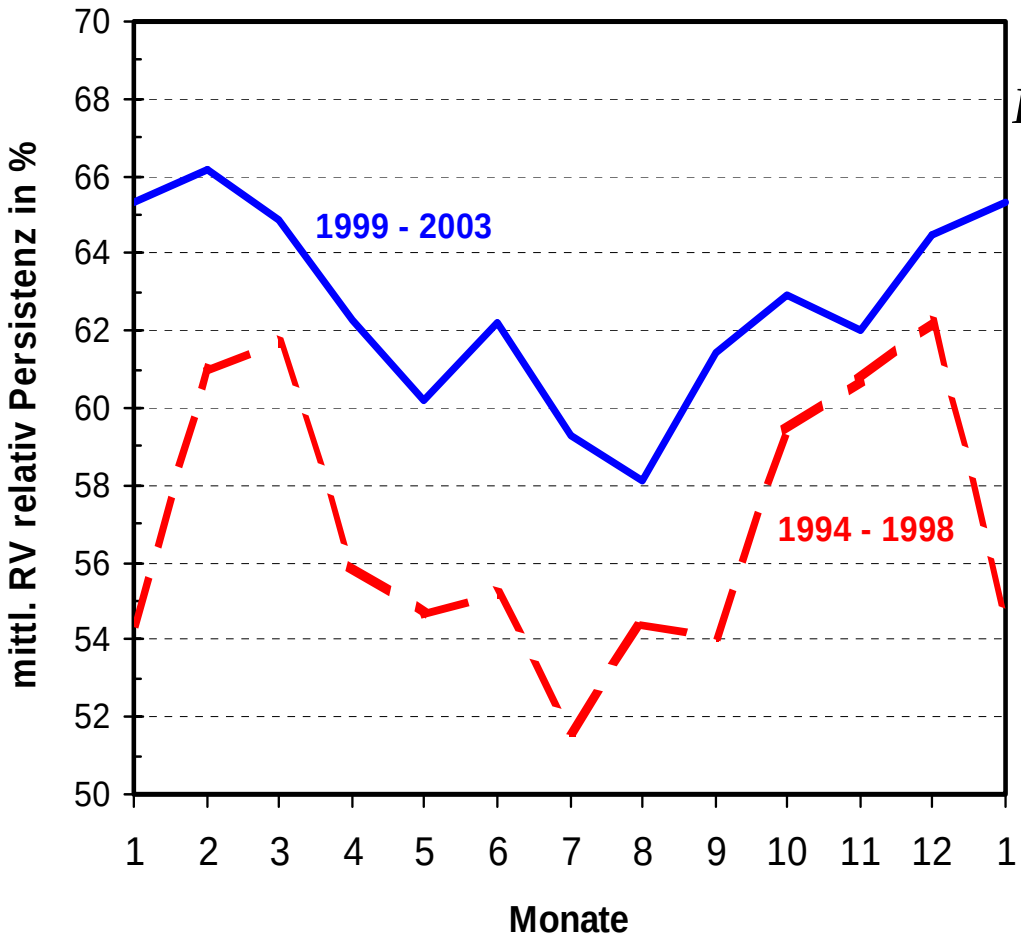
(2) *than is **perceived**.*

*Perception = a priori difficulty * skill (+ eps)*

Relevance:

- user-forecaster relationship
- forecaster-guidance relationship
- weather „independent“ verification
- model testing strategy

Averaged skill (RV) relative to persistence



$$RV = \left(1 - \frac{mse_{forecast}}{mse_{persistenz}}\right) * 100\%$$

		Beobachtung		
		Nein	Ja	Summe
Vorhersage	Nein	Korrekte Zurück-weisung	Verpaßt	$p(\overline{f})$
	Ja	Falscher Alarm	Treffer	$p(f)$
	Summe	$p(\overline{o})$	$p(o)$	1

Zufälliges Ereignis

Eintrittswahrscheinlichkeit $p(o) \equiv \varepsilon$

		Beobachtung		
		Nein	Ja	Summe
Vorhersage	Nein	$(1-\varepsilon)^2$	$\varepsilon * (1-\varepsilon)$	$1-\varepsilon$
	Ja	$\varepsilon * (1-\varepsilon)$	ε^2	ε
	Summe	$1-\varepsilon$	ε	1

$$hit\_rate = \frac{\varepsilon^2}{\varepsilon} = \varepsilon$$

$$false\_alarm\_rate = \frac{\varepsilon * (1-\varepsilon)}{1-\varepsilon} = \varepsilon$$

$$false\_alarm\_ratio = \frac{\varepsilon * (1-\varepsilon)}{\varepsilon} = 1-\varepsilon$$

		Time/Space/user scale			
		Short range	Medium range	TAF	SWIS
Type of forecast	Climate	-	!	-	-
	Persistence	!	-	!	-
	DMO Direct Model Output	GME LM	GME EZMW EZ-EPS	-	-
	Statistical and physical post-processing	GMOS LM/Kalman „weather interpreation“	GMOS EZMOS MOSMIX EPS/Kalman AFReg-GME AFReg-EZ AFReg-MIX	AutoTAF	EBM MOS
	Forecaster	Regional office	Central Forecast office	airport	MMO