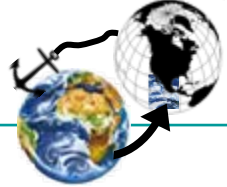
Ceilometer evaluation – Probabilistic verification of deterministic forecasts and observations

Felix Ament
Universität Hamburg
Motivated and supported by Christoph Selbach and Thorsten Reinhardt!

Observations and Models



OBS

- Two sites: Hamburg and Cabauw
- Measurements every 15s or 30s, respectively
- Only clouds bellow 3000m are considered
- Period: 2007 and 2008

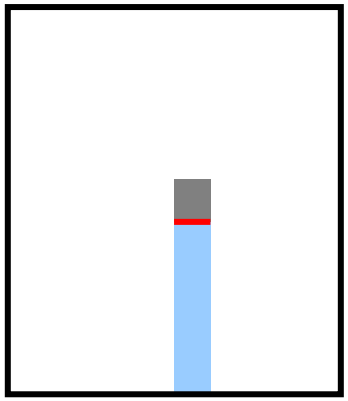
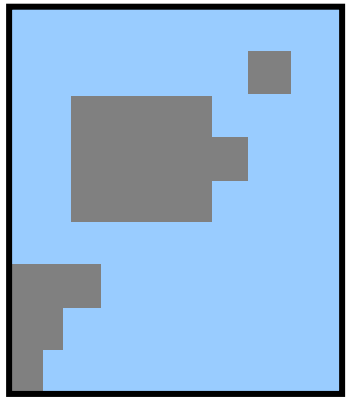
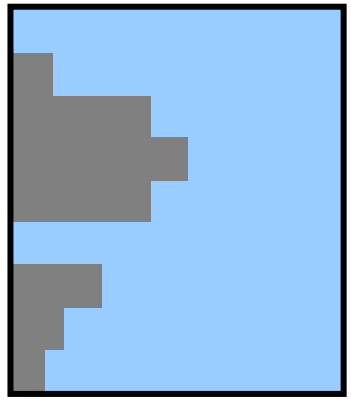
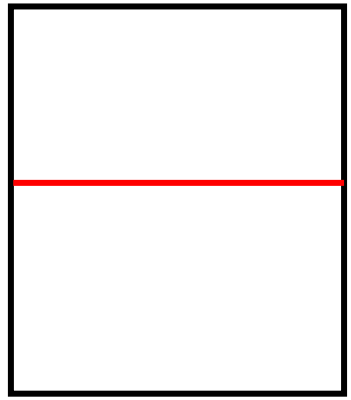
Model

- Two models: COSMO-EU and COSMO-DE
- Runs started at 00UTC (03UTC) up to +24h (+21h), respectively
- Output every 60min (15min), respectively.
- Period: 2007 and 2008

COSMO-EU at Hamburg	COSMO-DE at Hamburg	COSMO-EU at Cabauw	COSMO-DE at Cabauw

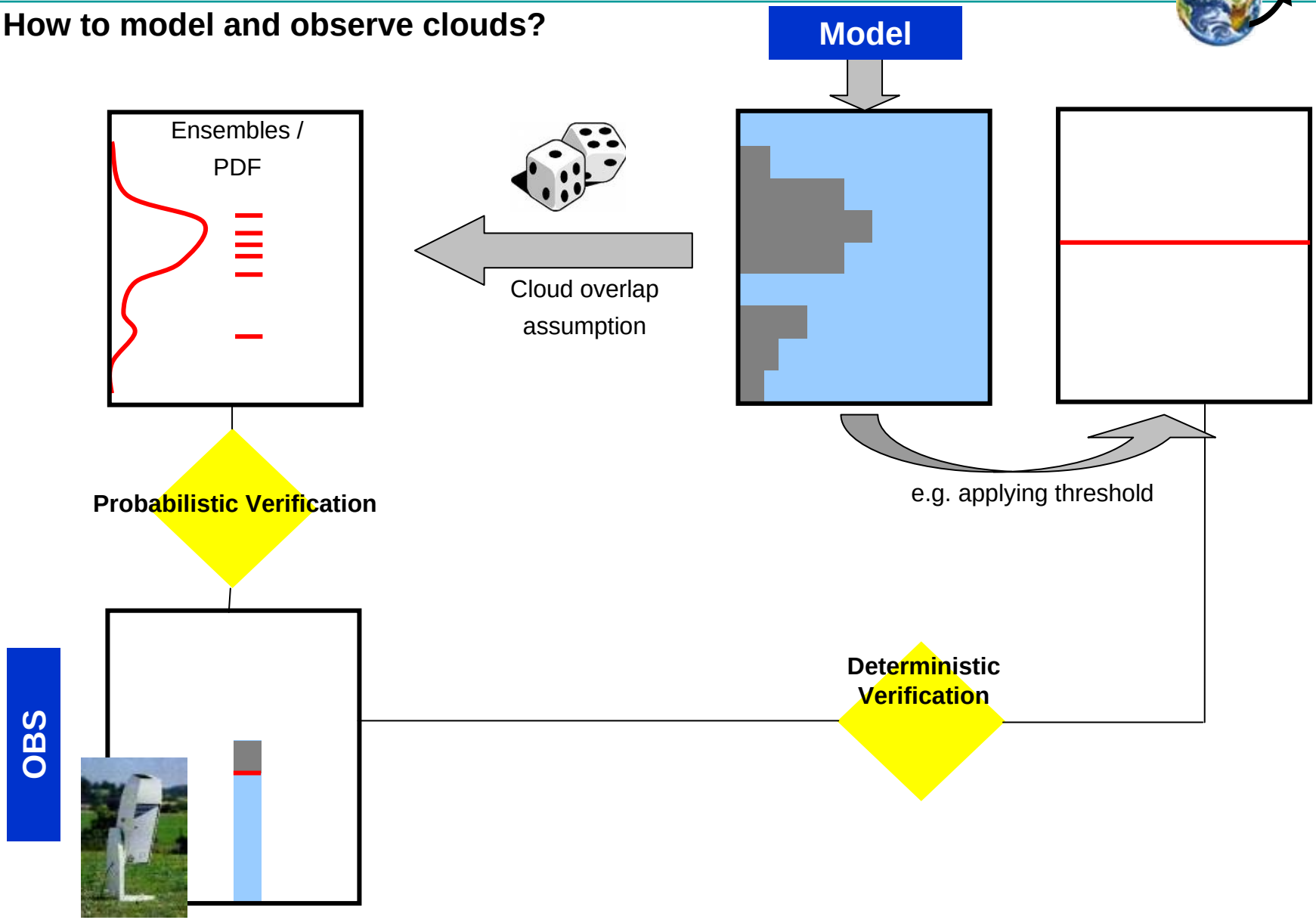
Setting the scene I

How to describe the cloud base?

Point value	3d cloud field	Cloud fraction	Characteristic numbers
			
• e.g. single point observations	• all information about cloud within one grid box	• Cloud cover in various layers (no information about location)	• e.g. mean cloud cover defined by a cloud fraction threshold

Setting the scene II

How to model and observe clouds?

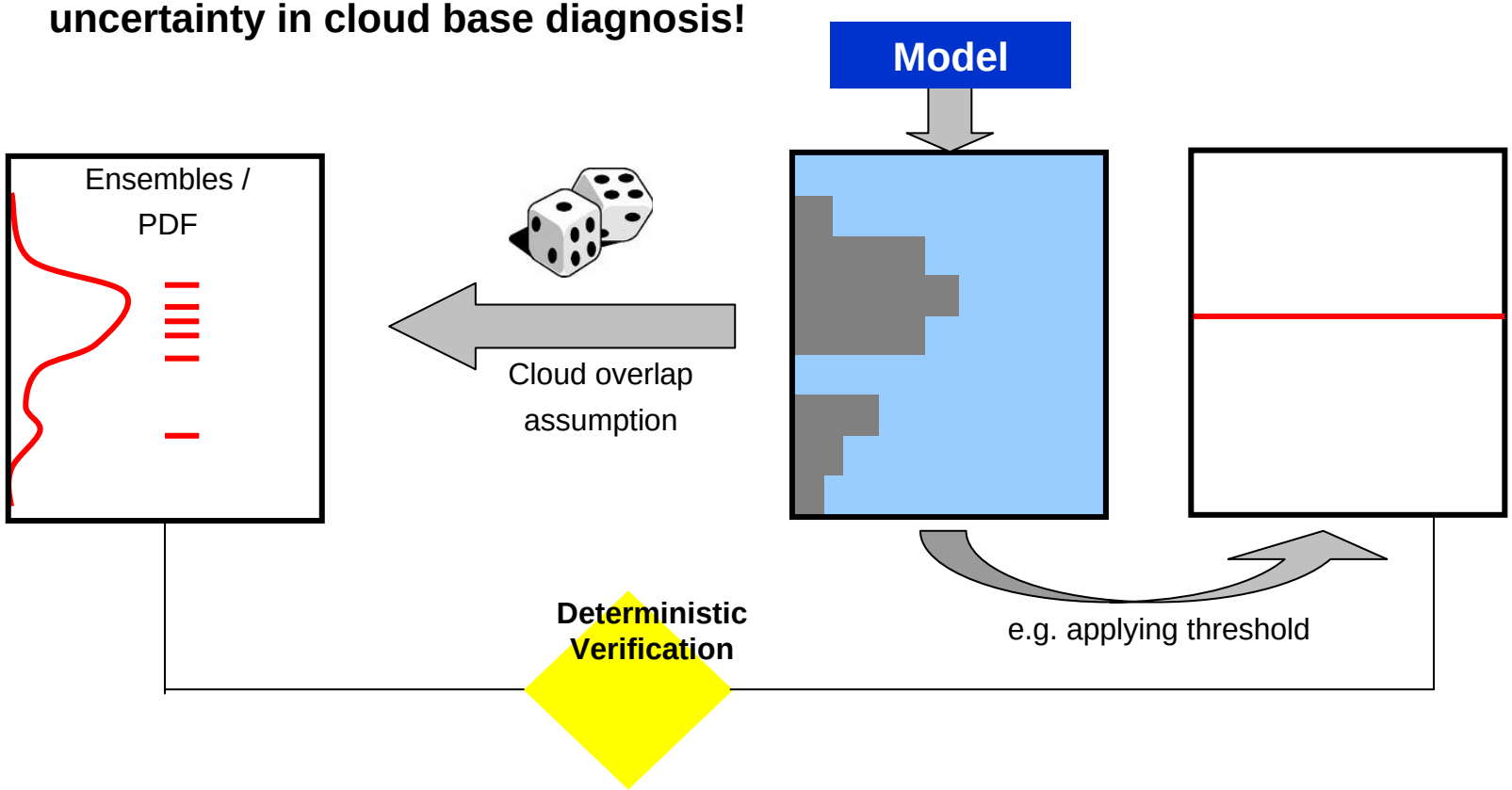


Perfect model approach I

Recipe

1. Assume that the model is perfect!
2. Generate an ensemble of virtual observations, which are in perfect agreement with forecast
3. Verify them deterministically.

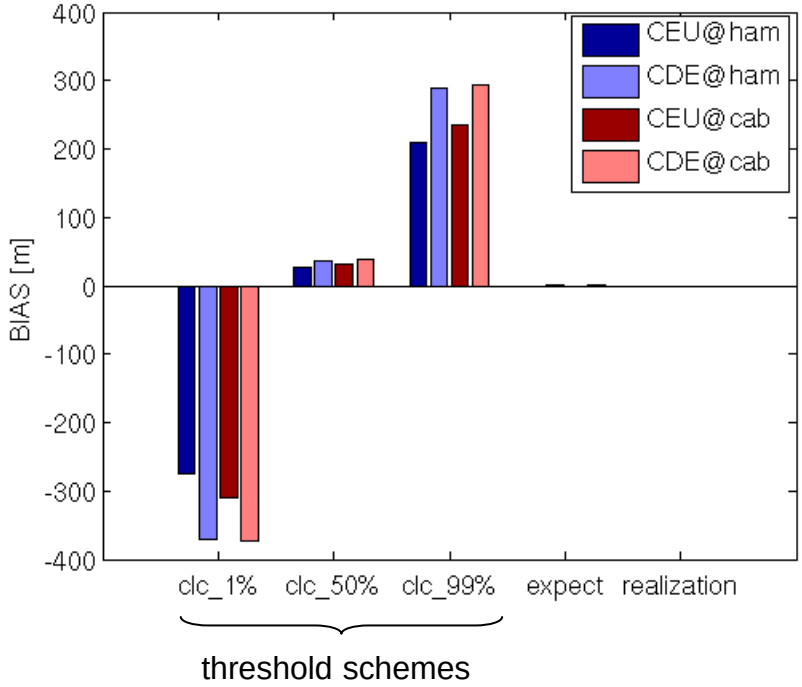
➔ **Any imperfect verification result is due to uncertainty in cloud base diagnosis!**



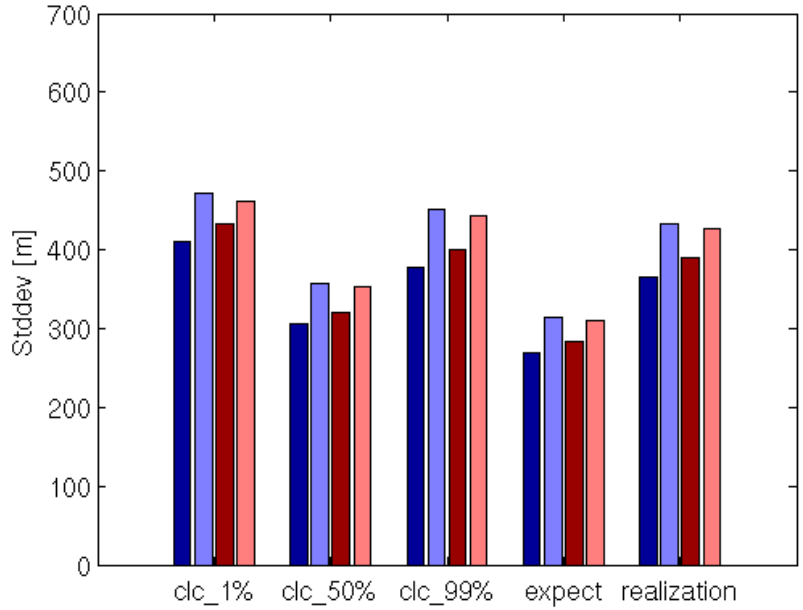
Perfect model approach I

Systematic and random errors of various cbase diagnoses

BIAS



STDDEV

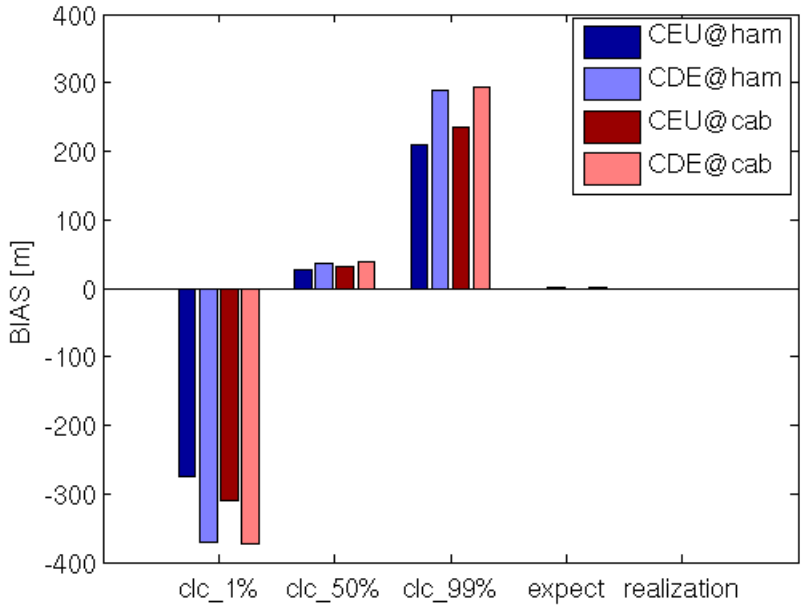


- Threshold schemes cause systematic errors.
- Expected value is BIAS-free (by construction) and results in smallest random error.

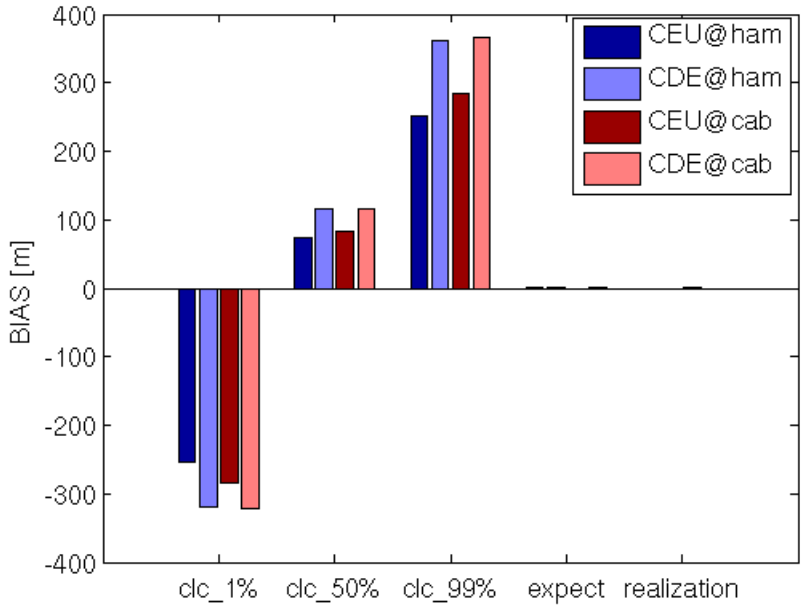
Perfect model approach II

Does the overlap assumption matter?

COSMO overlap (Maximum Random)



Random overlap

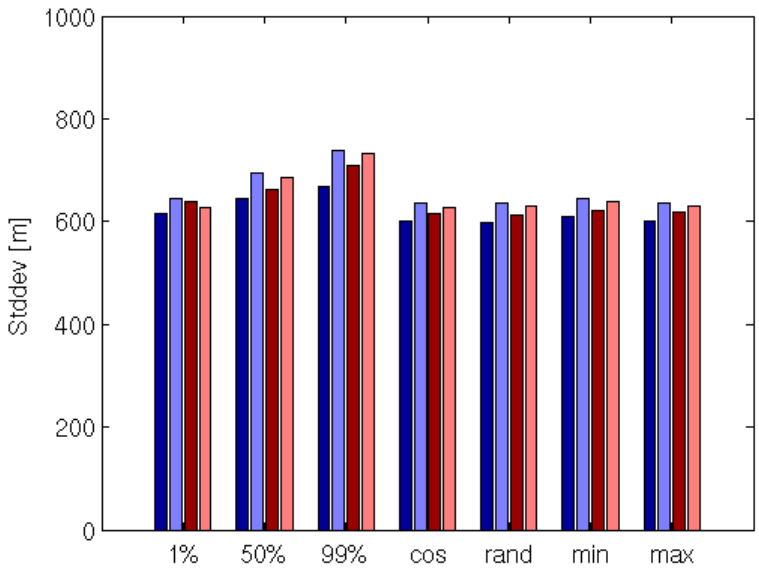


- Deviations of other assumptions (Minimum and maximum overlap) is even smaller.
- Impact of overlap assumption is small.

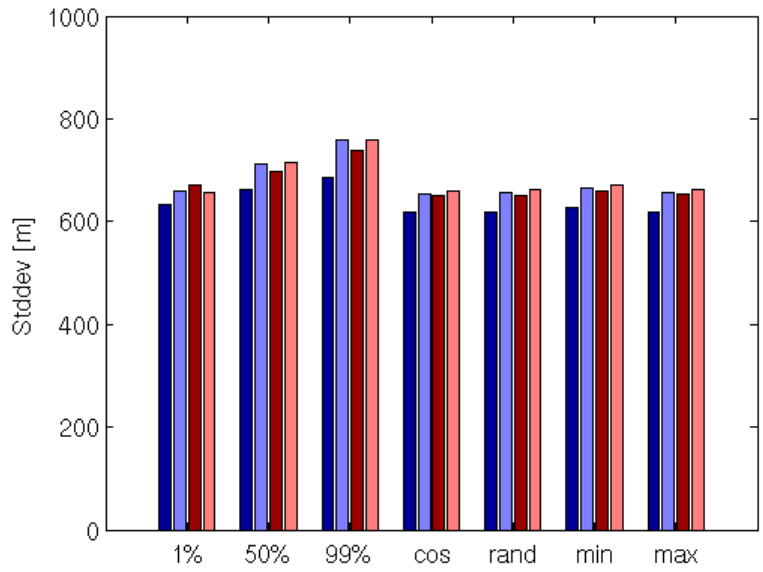
Perfect model approach III

Does the temporal averaging of observation matter?

Instantaneous



10min average

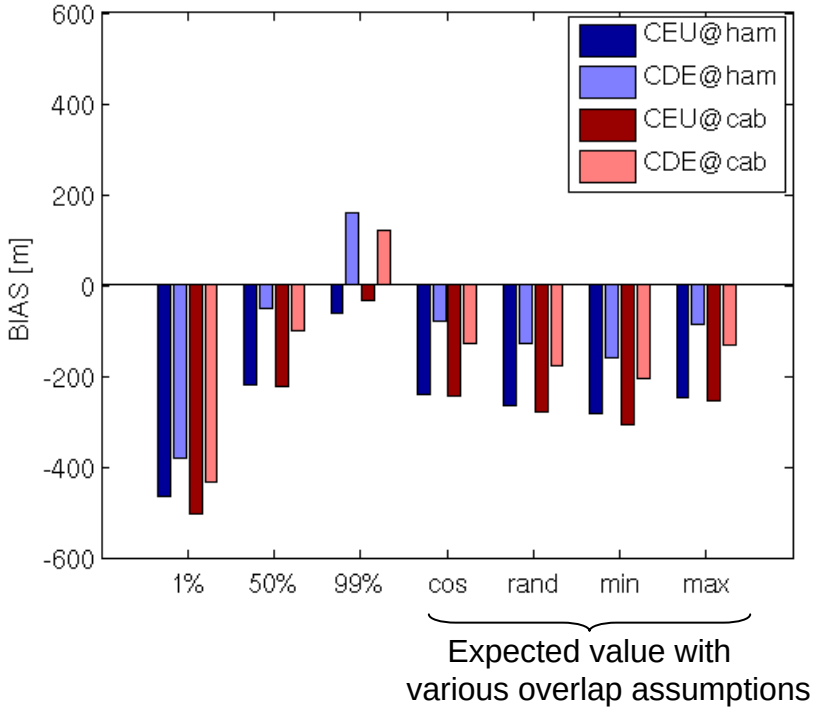


- Impact is surprisingly small. (Finding needs to be confirmed!)

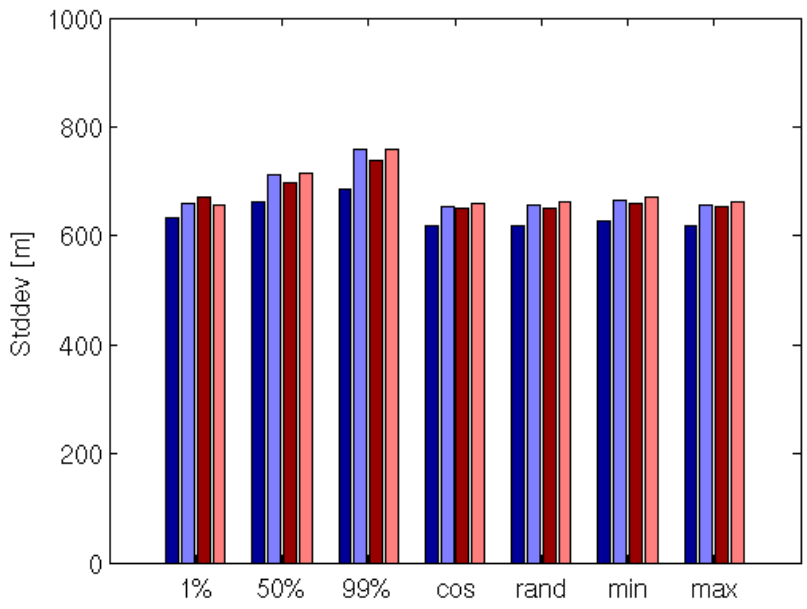
Verification against OBS

Overall error of cloud base forecasts

BIAS



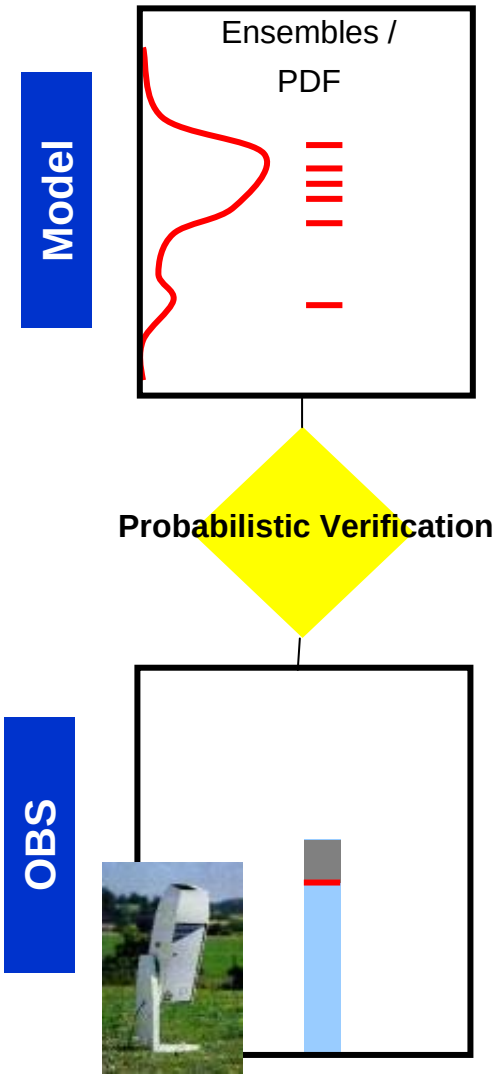
STDDEV



- All models predict too low clouds.
- Error is smaller for COSMO-DE
- Random error (STDDEV) is in same order of magnitude as the uncertainty derived by the perfect model approach.

Probabilistic Verification

Methodology



1.) Correct climatology?

➔ Frequency distribution (already down by C. Selbach).

2.) Correct at individual time steps?

➔ Brier skill score

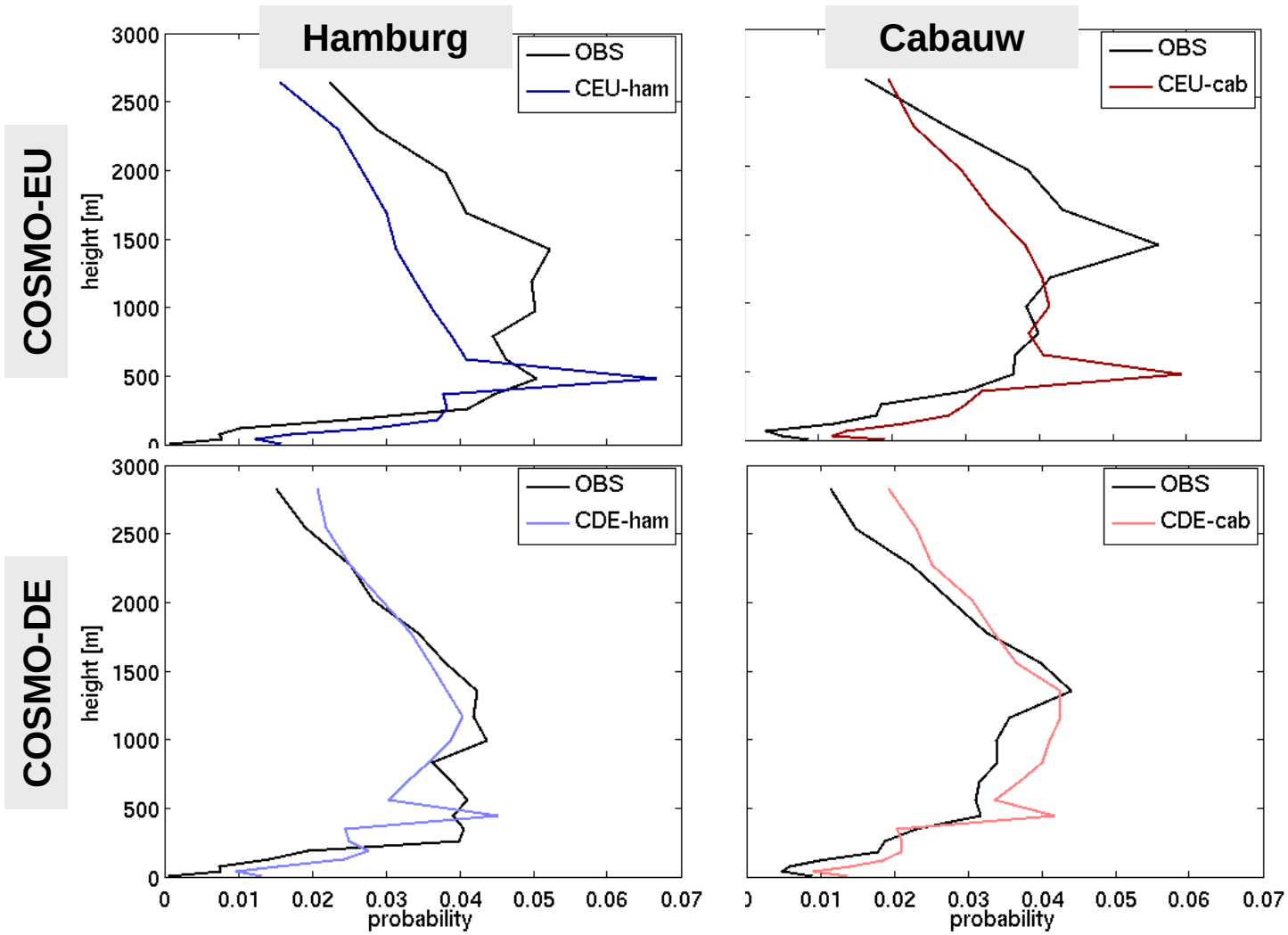
3.) Verifying the forecasted probabilities?

➔ Reliability diagrams

4.) Assessing the overall performance

➔ Expected RMSE

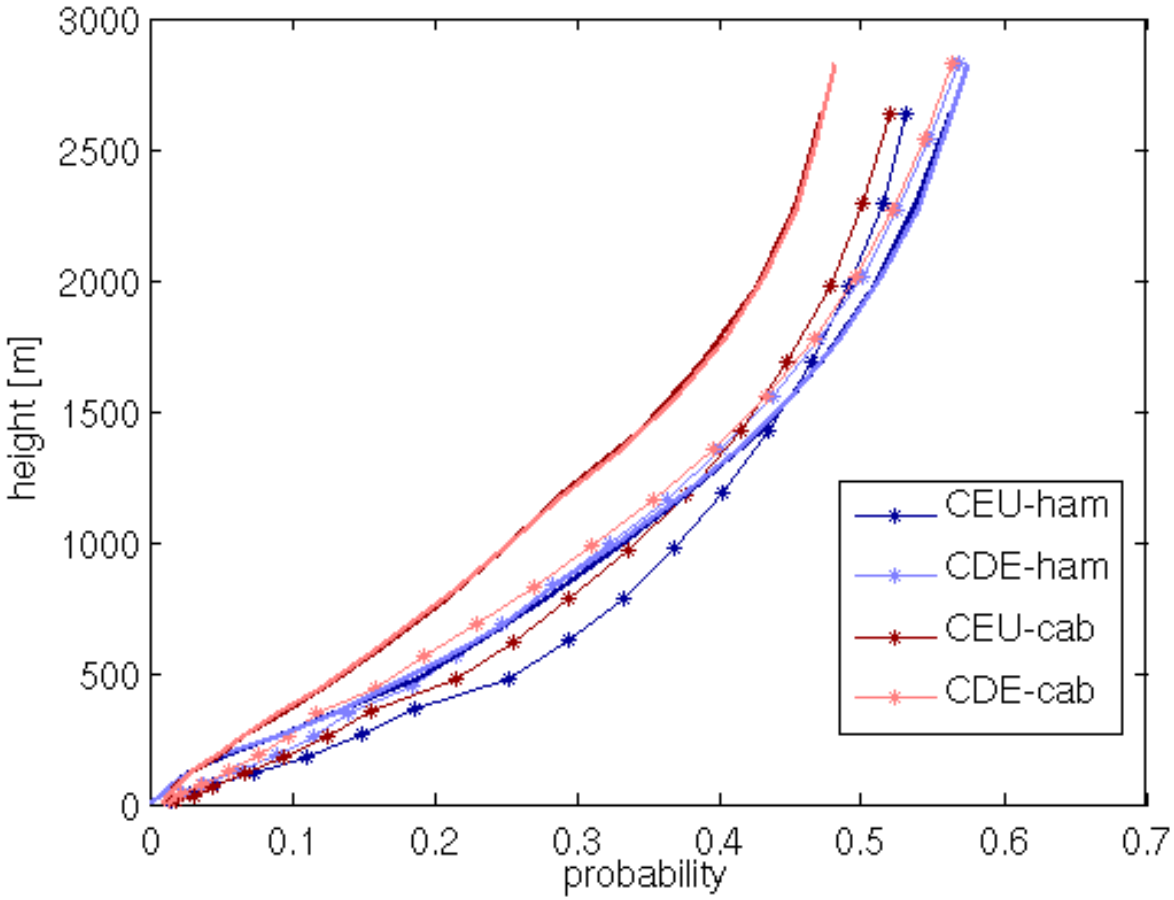
Frequency distribution I



- Overall distribution represented well by both models.
- Model predictions peak at ~500m (specially COSMO-EU). This may effect accuracy at higher altitudes.

Frequency distribution II

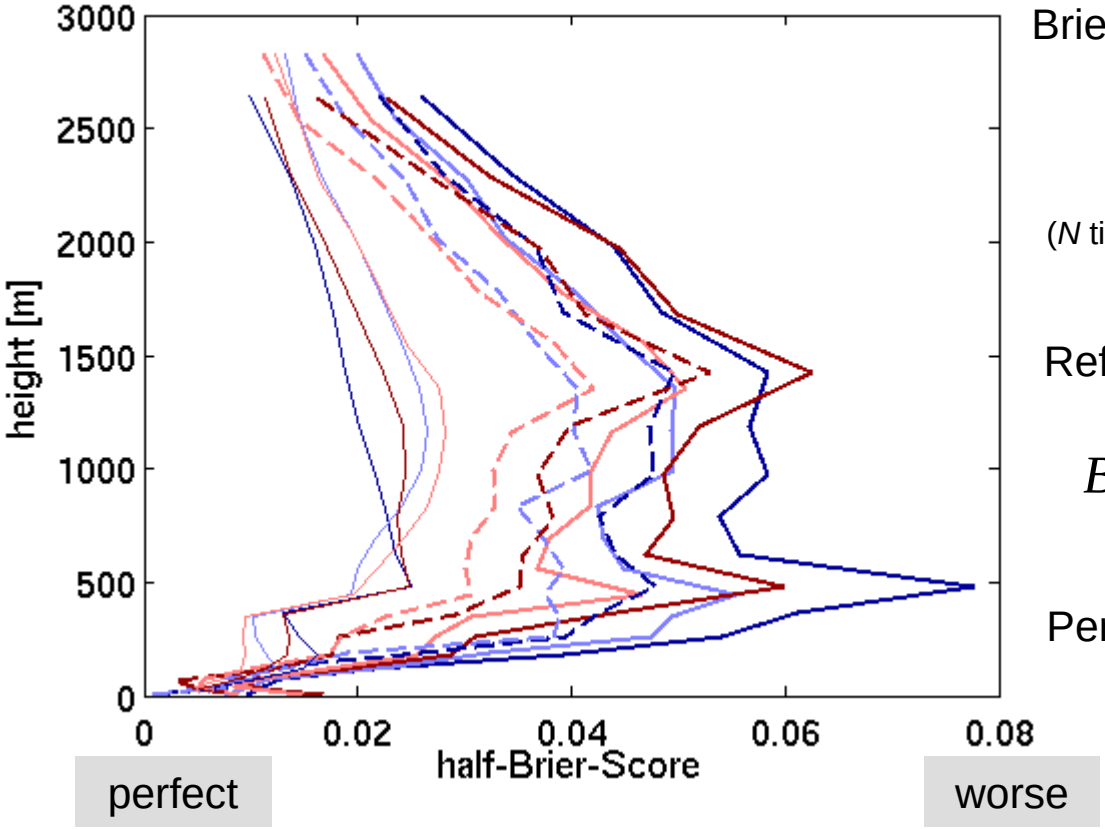
Cumulative frequency distribution
= probability to have cloud below height h



- Clouds are more frequent in Hamburg than in Cabauw. Models do not reflect this difference.
- Highres models have more high cloud basis – probably caused by the smaller peak at 500m.

Brier score I

~ RMSE in terms of probabilities



Brier score (thick solid line)

$$BS_k = \frac{1}{N}\sum_{i=1}^{N}\left(f_{i,k} - o_{i,k}\right)^2$$

(*N* timesteps; forecasted prob *f*; observed prob. *o*)

Reference Brier Score (dashed)

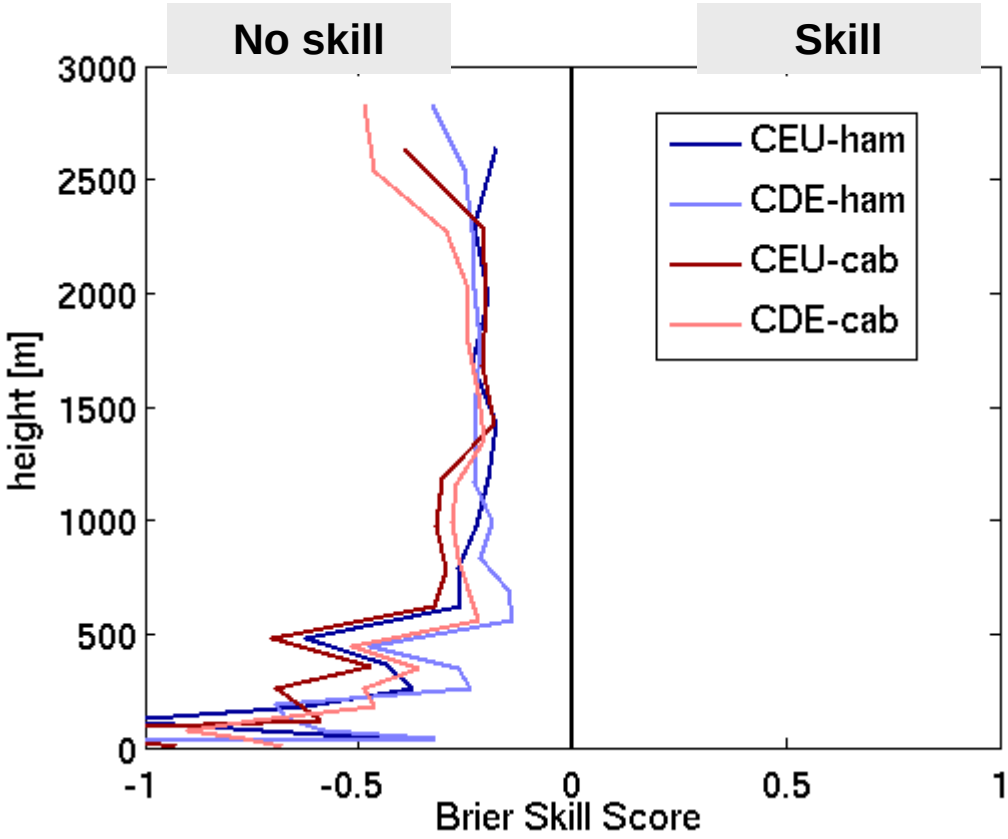
$$BS_{k,ref} = \frac{1}{N}\sum_{i=1}^{N}\left(\bar{o}_k - o_{i,k}\right)^2$$

Perfect Model Brier Score (thin solid)

- Forecasts cannot beat the climatology!

Brier Score II

Brier skill score

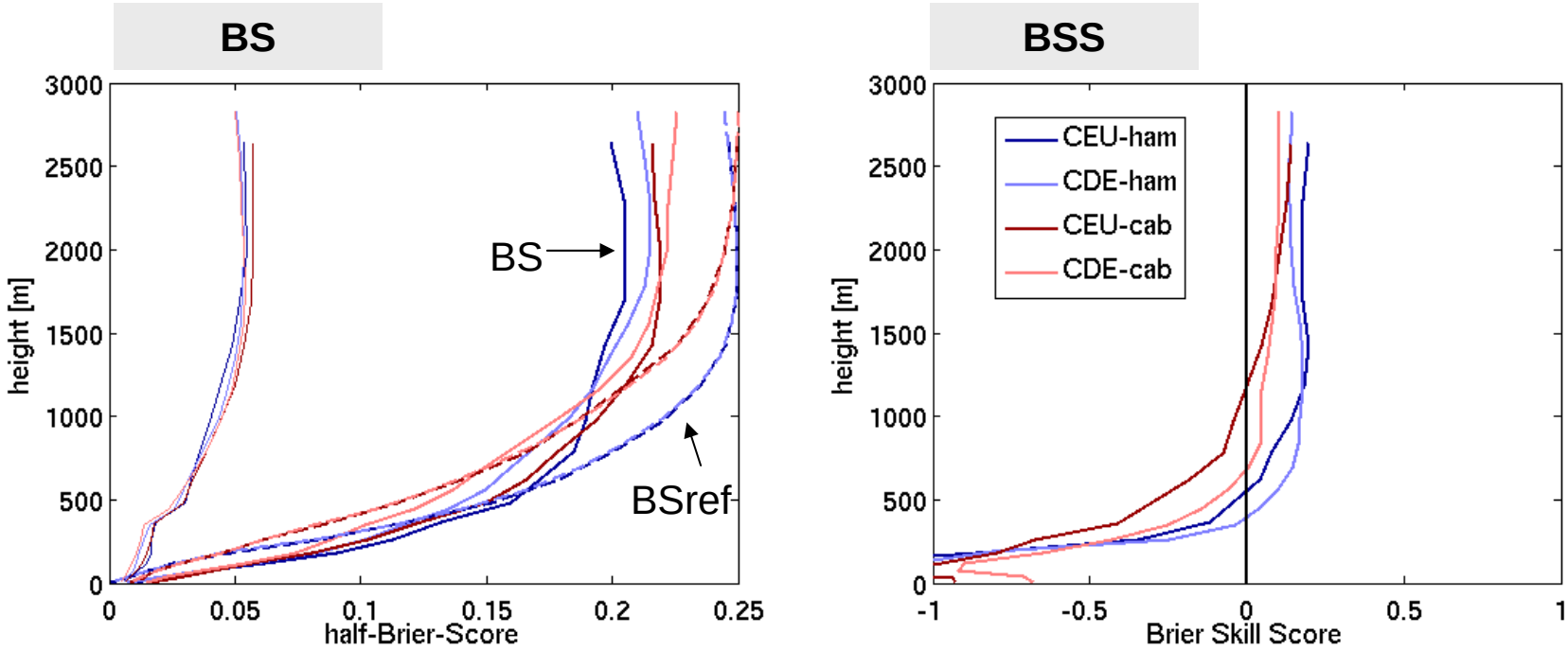


$$BSS = 1 - \frac{BS_k}{BS_{k,ref}}$$

- Models have no skill to predict probability of occurrence of cloud base height at individual levels – task is too difficult!

Brier Score III

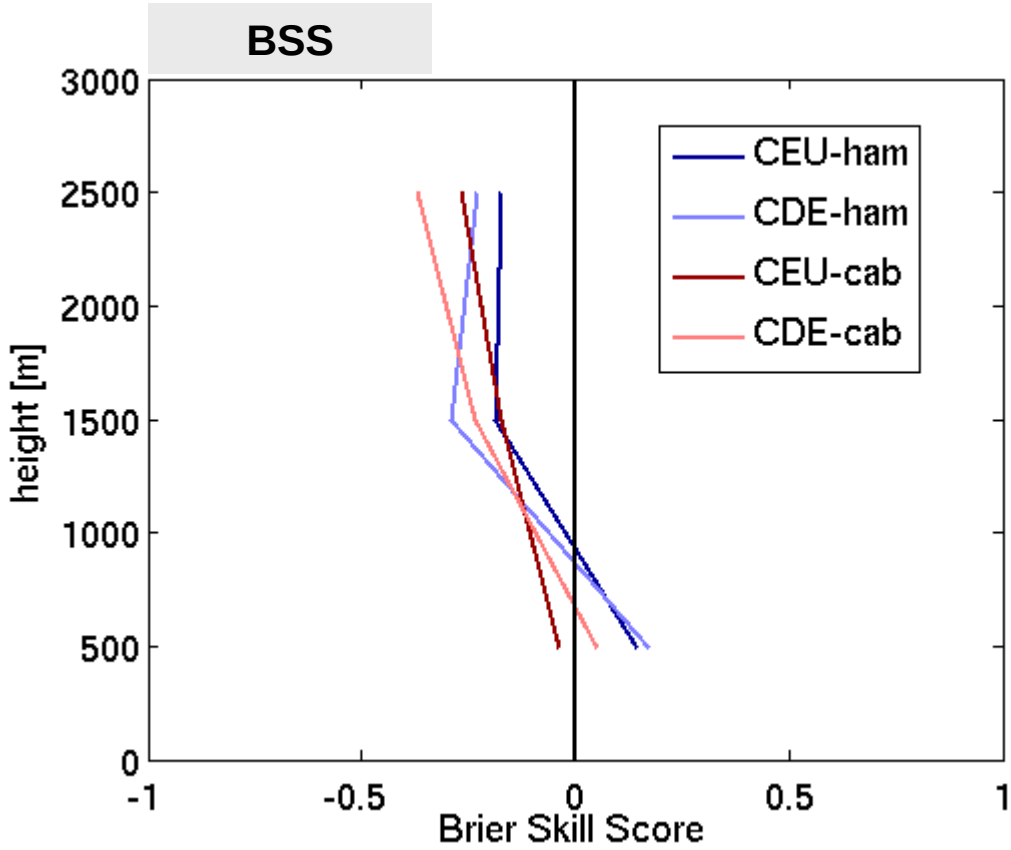
Cumulative cloud base height
= Is there a cloud with base height below height h?



- Models are skillful to predict the occurrence of cloud base height in layer which start at the surface and reach up to ~500m and more …
- COSMO-DE outperforms COSMO-EU

Brier Score IV

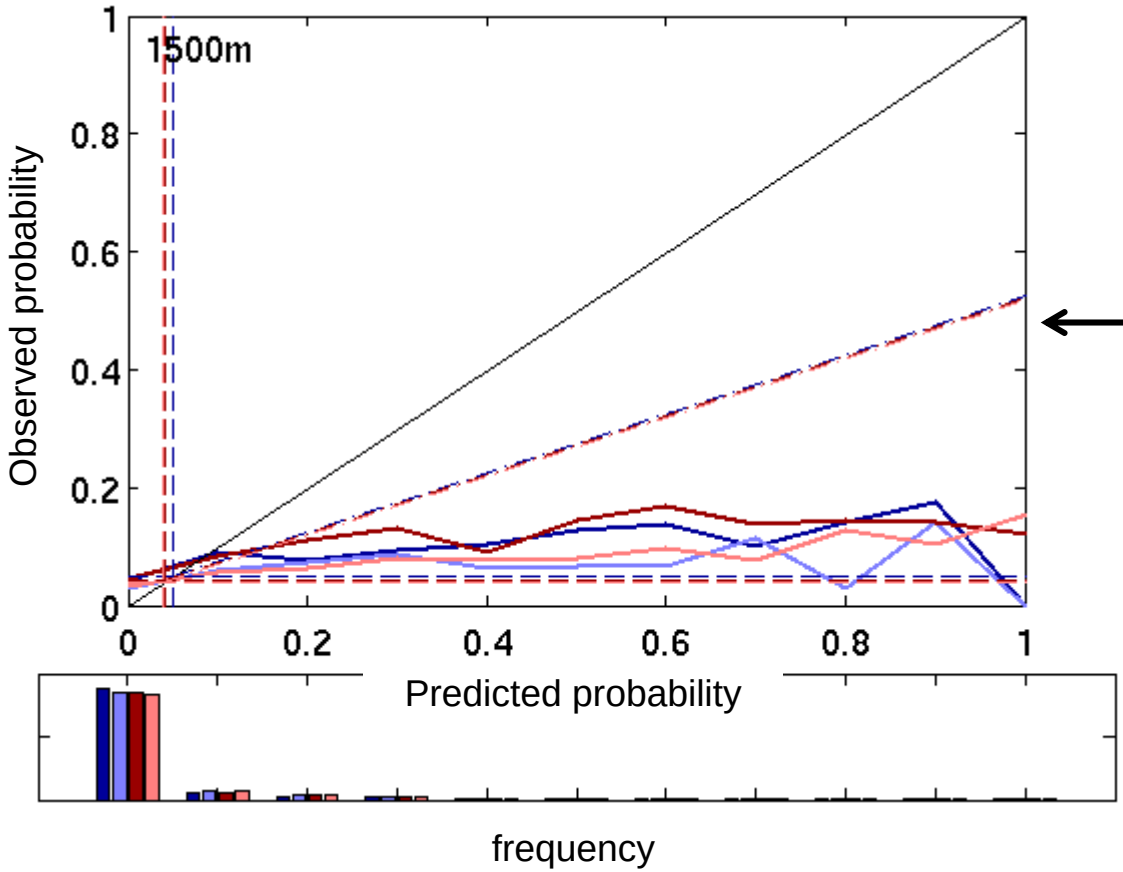
Layered averages: 0-1000m, 1000-2000m and 2000-3000m

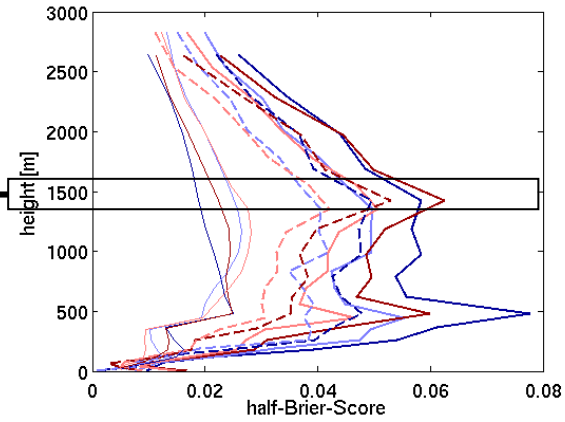


- Models are skillful only for the lowest layer.
- Note: Errors in lower layers accumulated in higher layers!

Reliability diagram I

Why do we have a poor BSS for individual layers?

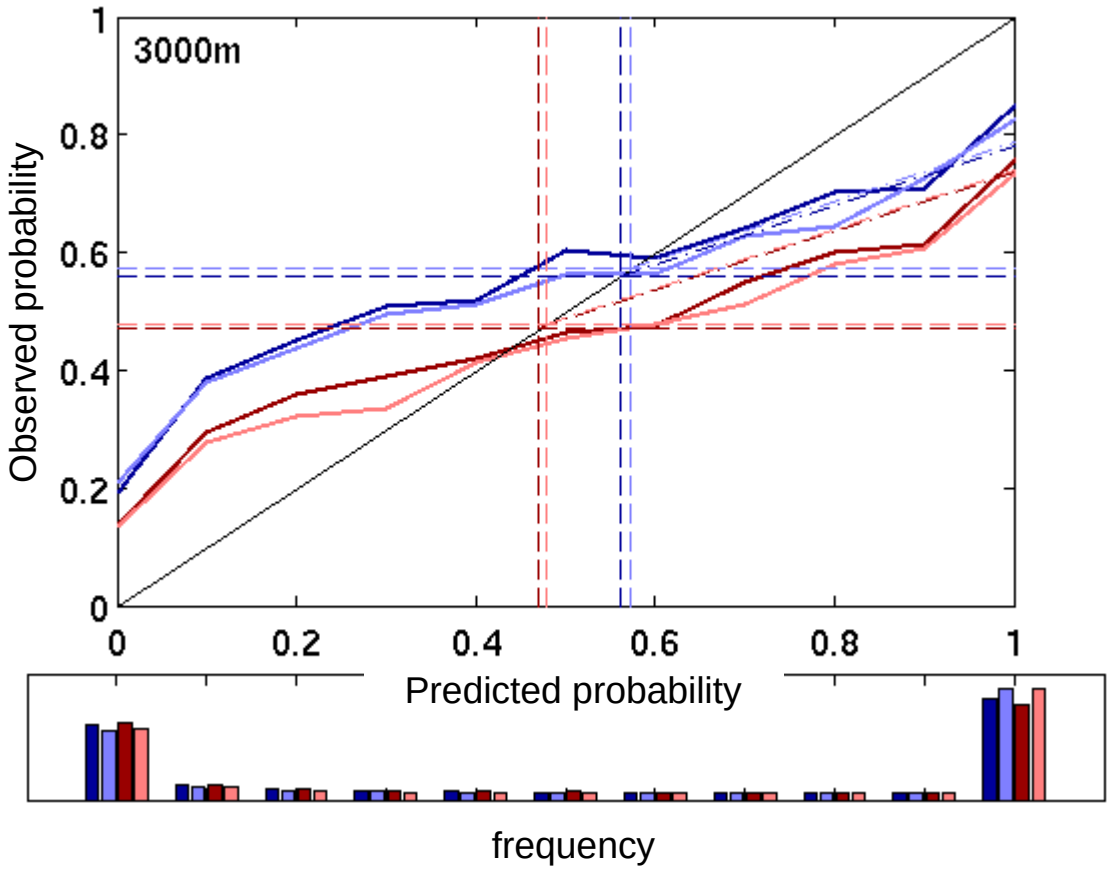




- Predicted probability is not related to observed frequency of occurrence!

Reliability diagram II

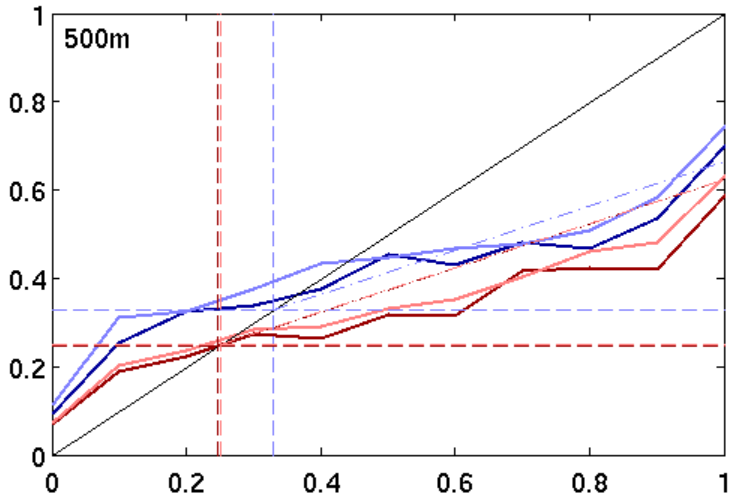
Some story – but for forecast “any cloud between 0 and 3000m?”

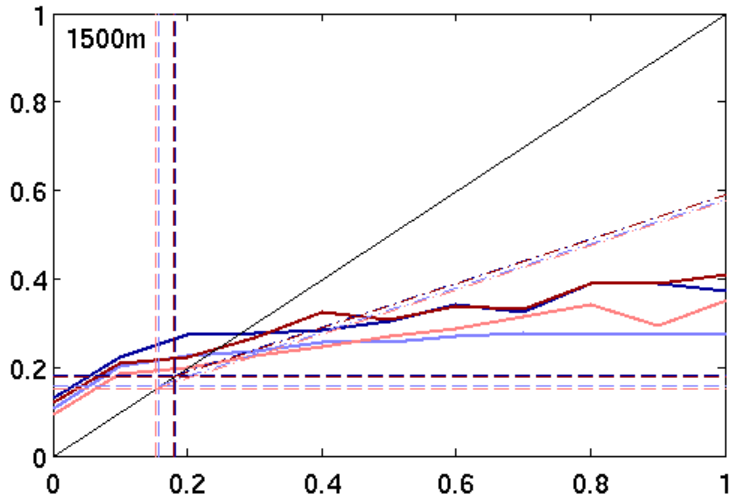


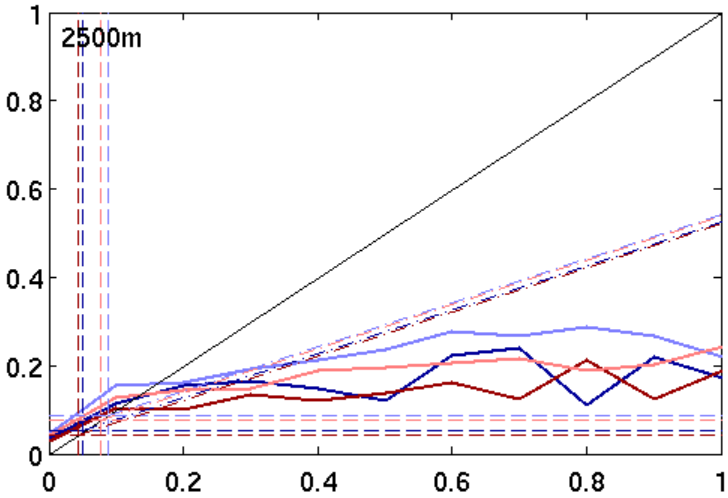
- Models are slightly overconfident and exhibit very small BIASes.

Reliability diagram III

Layered averages: 0-1000m, 1000-2000m and 2000-3000m







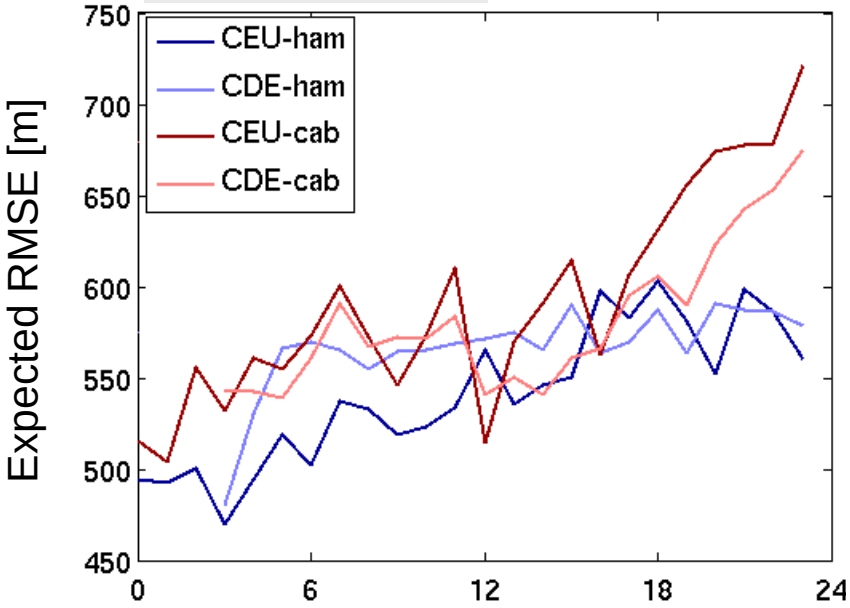
- Clear degradation of reliability with increasing layer height.

Expected RMSE

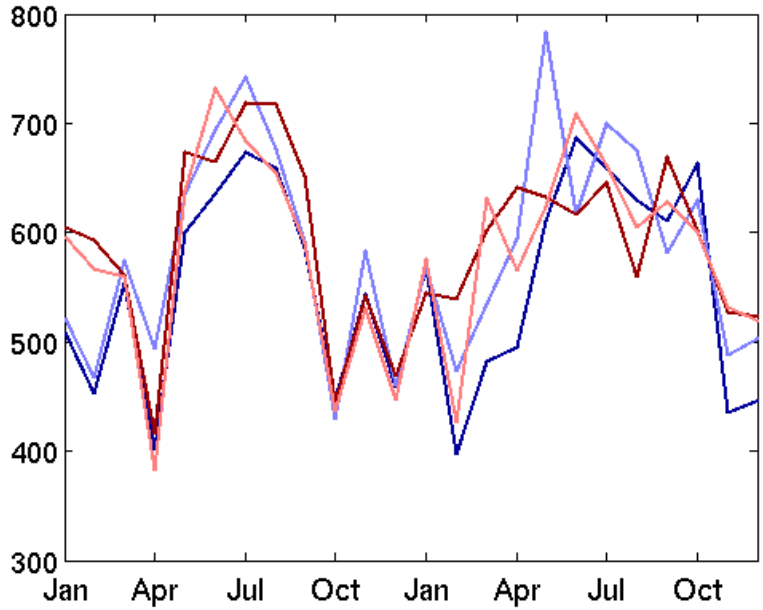
Coming again to one number:

$$ERMSE_i = \frac{\sum_{K=1}^{K} f_{i,k}\left(z_{obs} - z_k\right)^2}{\sum_{K=1}^{K} f_{i,k}}$$

Diurnal cycle

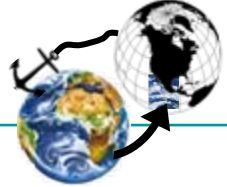


Monthly means



- Error increases with forecast time.
- Error is larger in wintertime than in summer.
- Despite a long verification period of two years, results are still quite noisy.

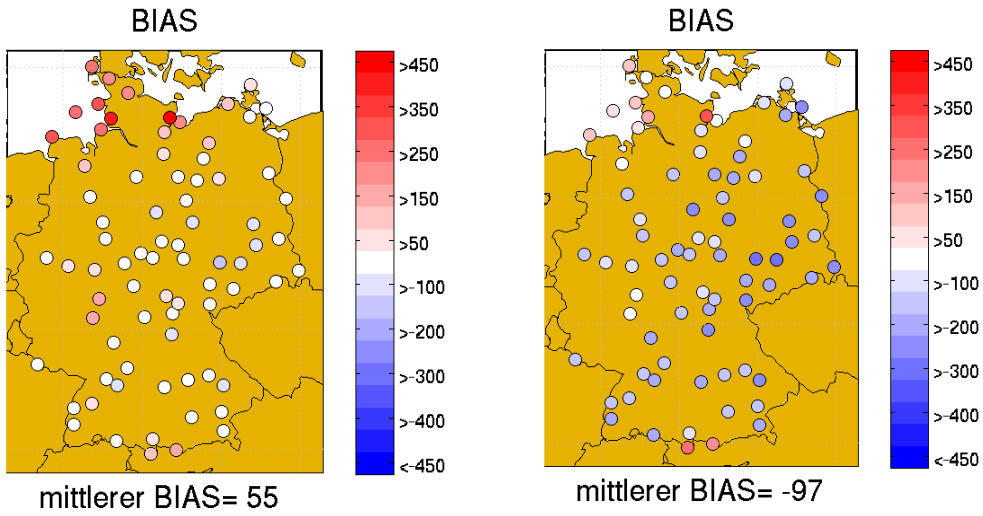
Outline of paper???



1. Introduction
2. Describing the problem
   (stuff from setting the scene)
3. Model and data
4. Evaluation of deterministic cloud base estimates (perfect model approach)
5. Probabilistic verification
   - Methodology
   - Applications: Which kind of model forecasts are skillful
6. Discussion and Conclusions

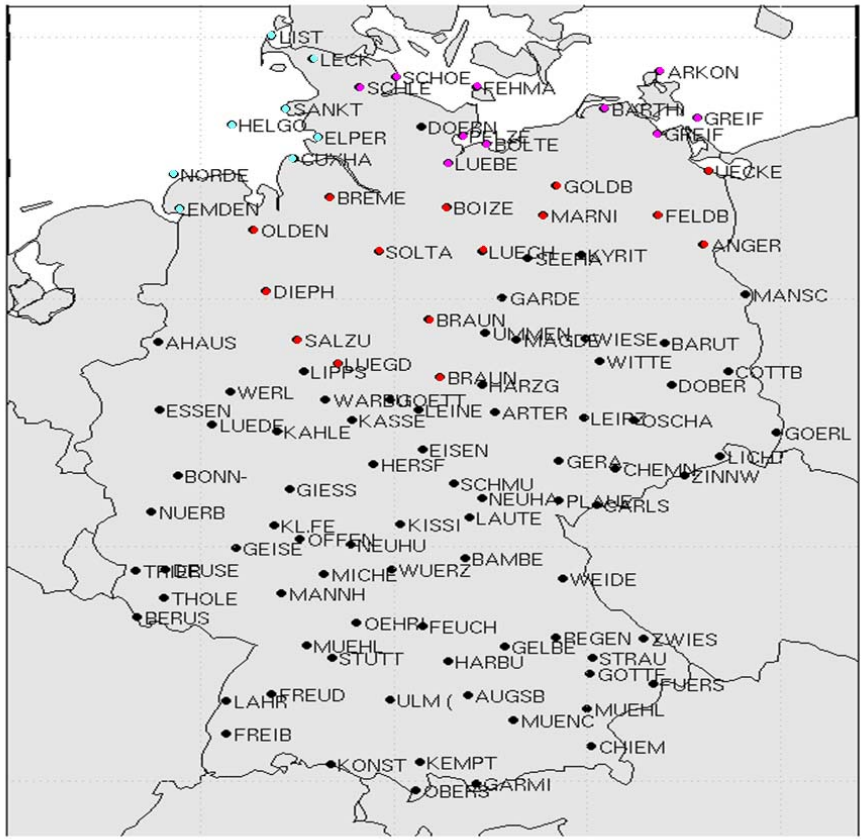
Your ideas are very welcome!

Probabilistic perspective



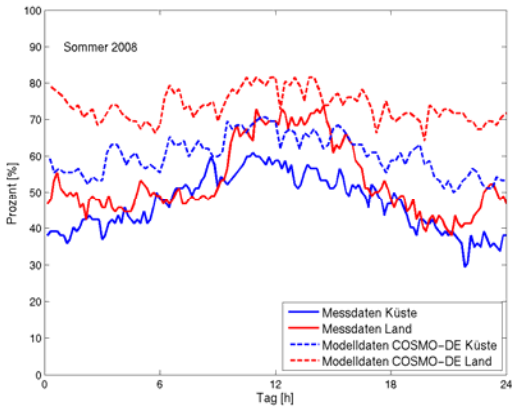
Mittlerer Bias des COSMO-DE und COSMO-EU im Sommer 2008 (Quelle:Dipl.-Arb. C. Selbach).

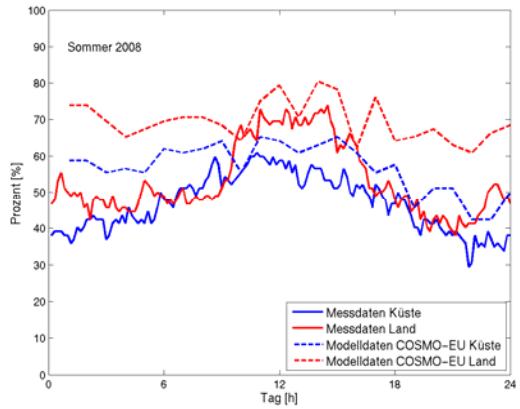
Probabilistic perspective



Mittlerer Bias des COSMO-DE und . C. Selbach).

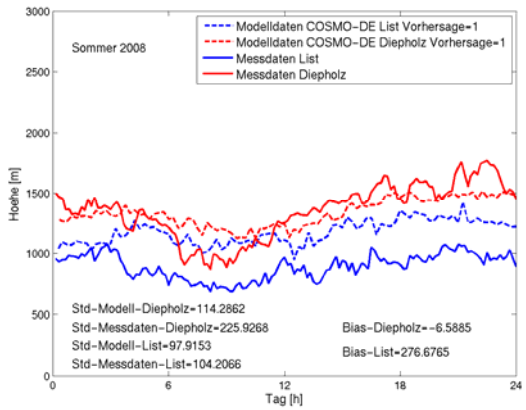
Probabilistic perspective

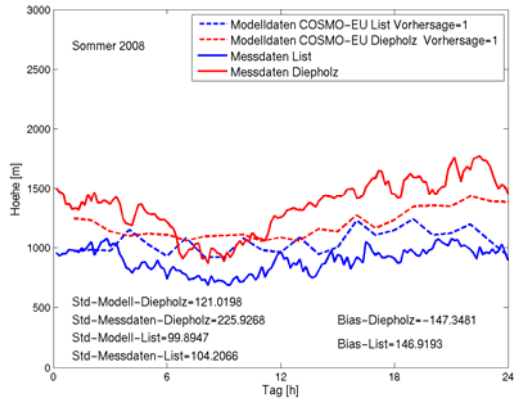




Mittlerer Bias des COSMO-DE und COSMO-EU im Sommer 2008 (Quelle:Dipl.-Arb. C. Selbach).

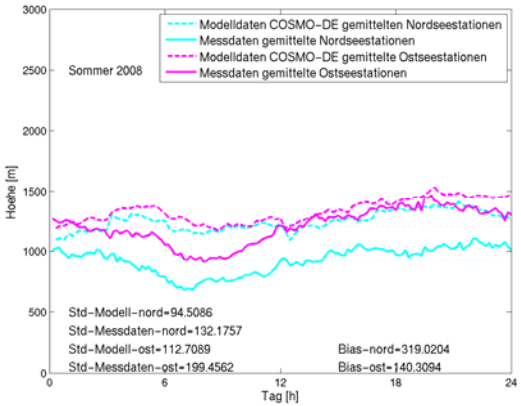
Probabilistic perspective

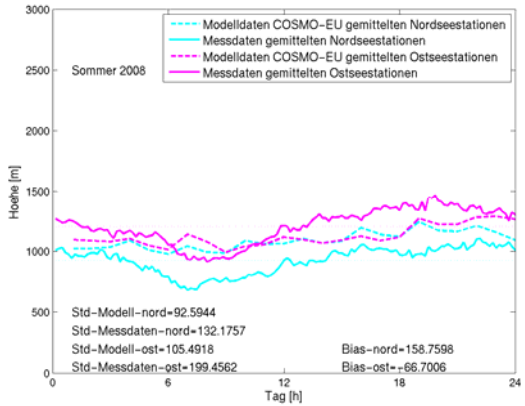




Mittlerer Bias des COSMO-DE und COSMO-EU im Sommer 2008 (Quelle:Dipl.-Arb. C. Selbach).

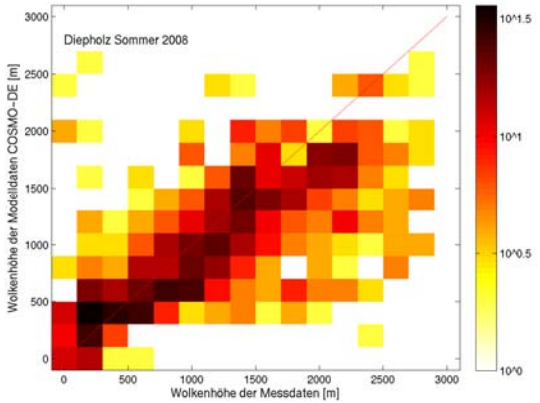
Probabilistic perspective

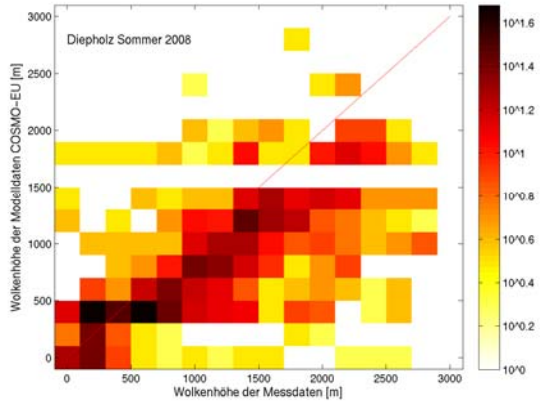


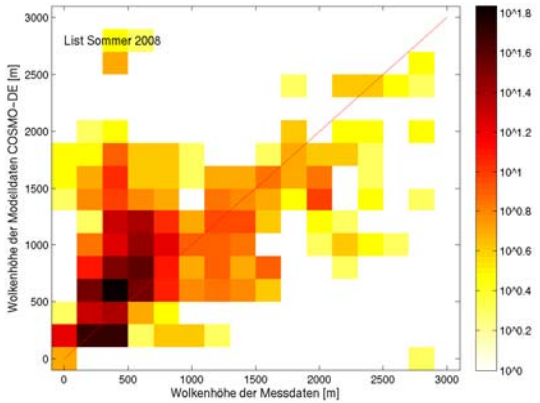


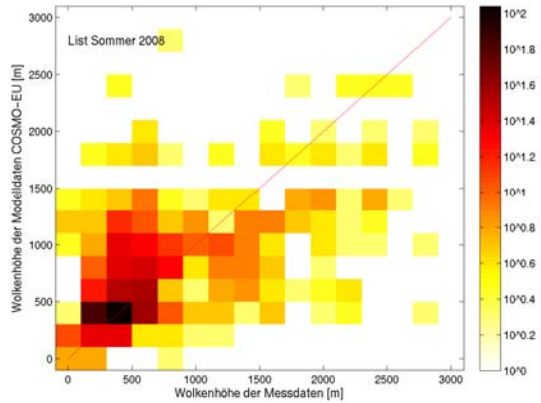
Mittlerer Bias des COSMO-DE und COSMO-EU im Sommer 2008 (Quelle:Dipl.-Arb. C. Selbach).

Probabilistic perspective



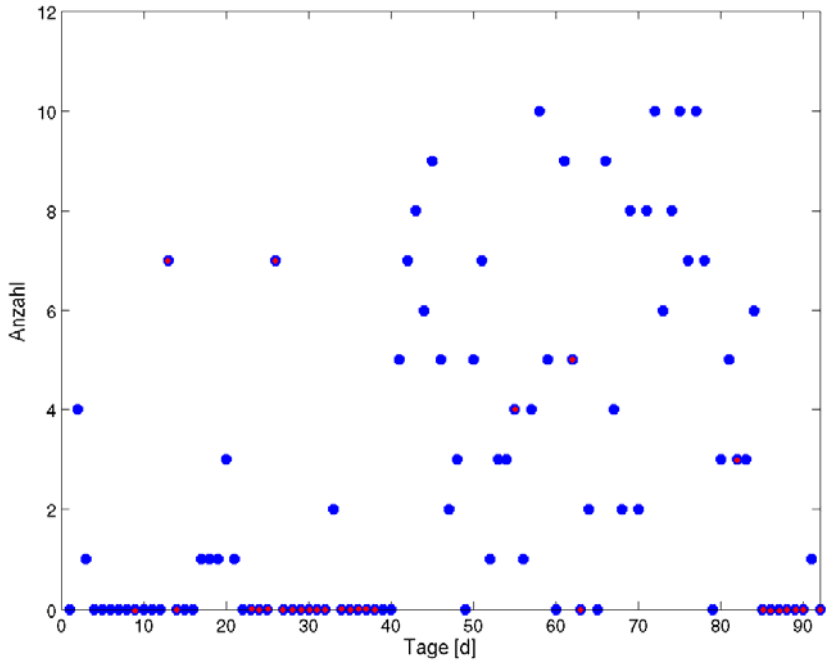


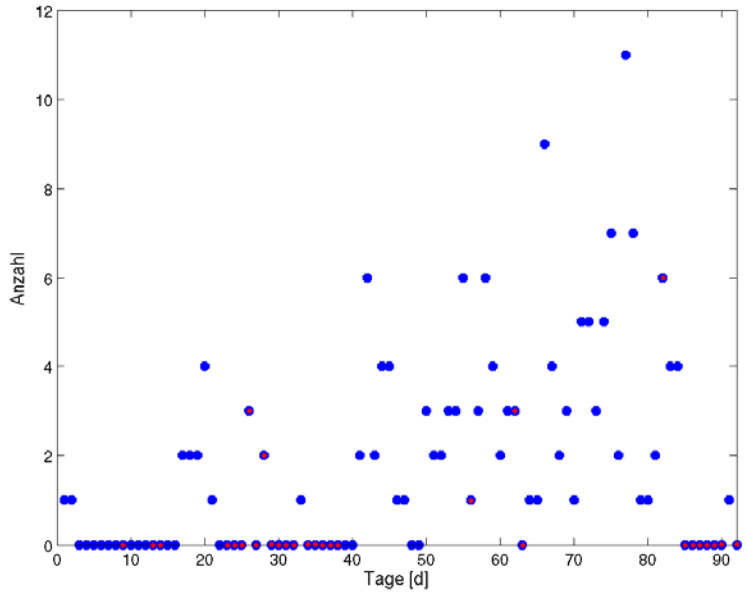




Mittlerer Bias des COSMO-DE und COSMO-EU im Sommer 2008 (Quelle: Dipl.-Arb. C. Selbach).

Probabilistic perspective

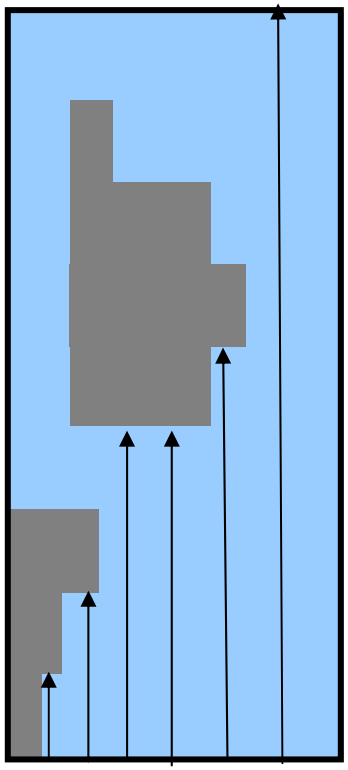




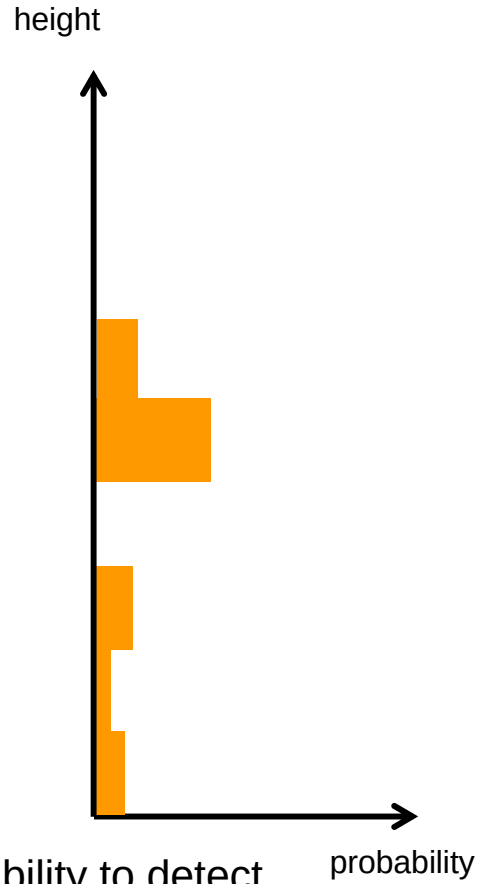
Mittlerer Bias des COSMO-DE und COSMO-EU im Sommer 2008 (Quelle:Dipl.-Arb. C. Selbach).

Probabilistic perspective

Model predicts a profile of cloud cover +
uses a certain cloud overlap assumption:

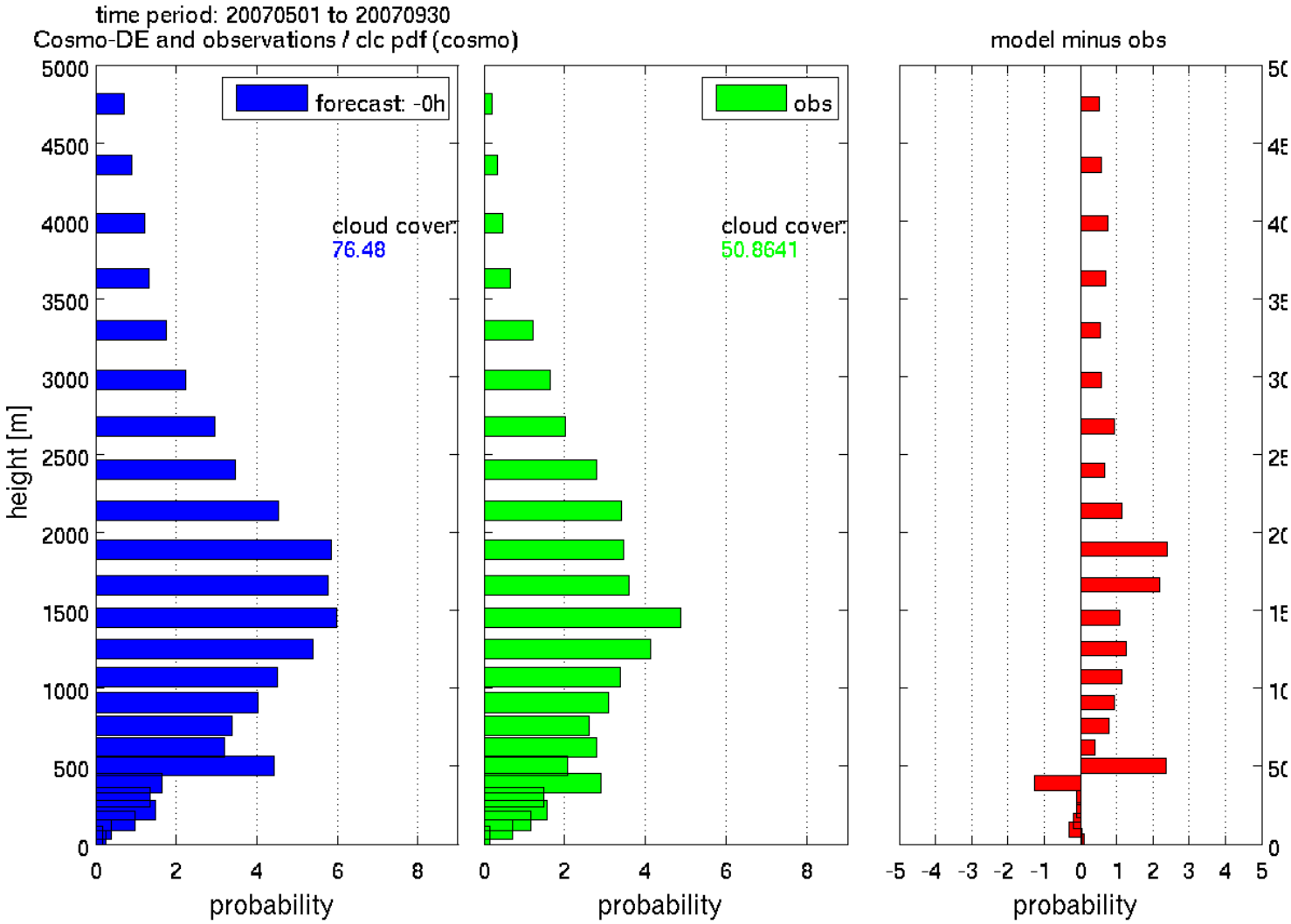


Ceilometer (or model clouds) are located arbitrarily inside the model column



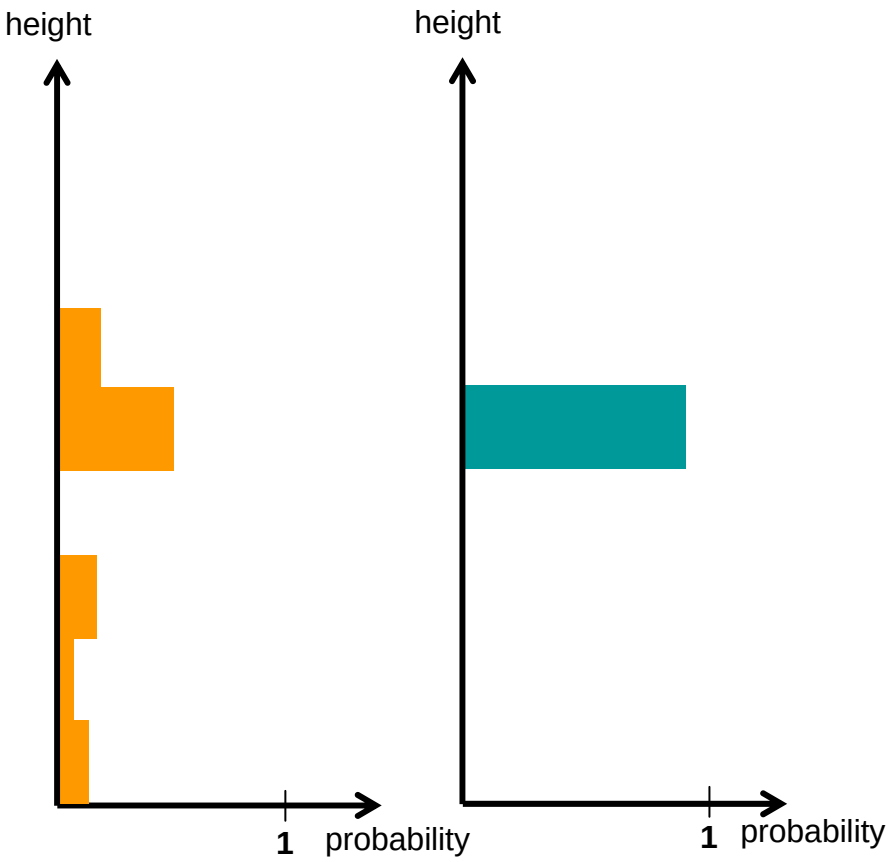
Profile of probability to detect cloud base at a specific height

Frequency bias



(by C. Selbach)

Brier Score (BS)



Brier Score for every height level k:

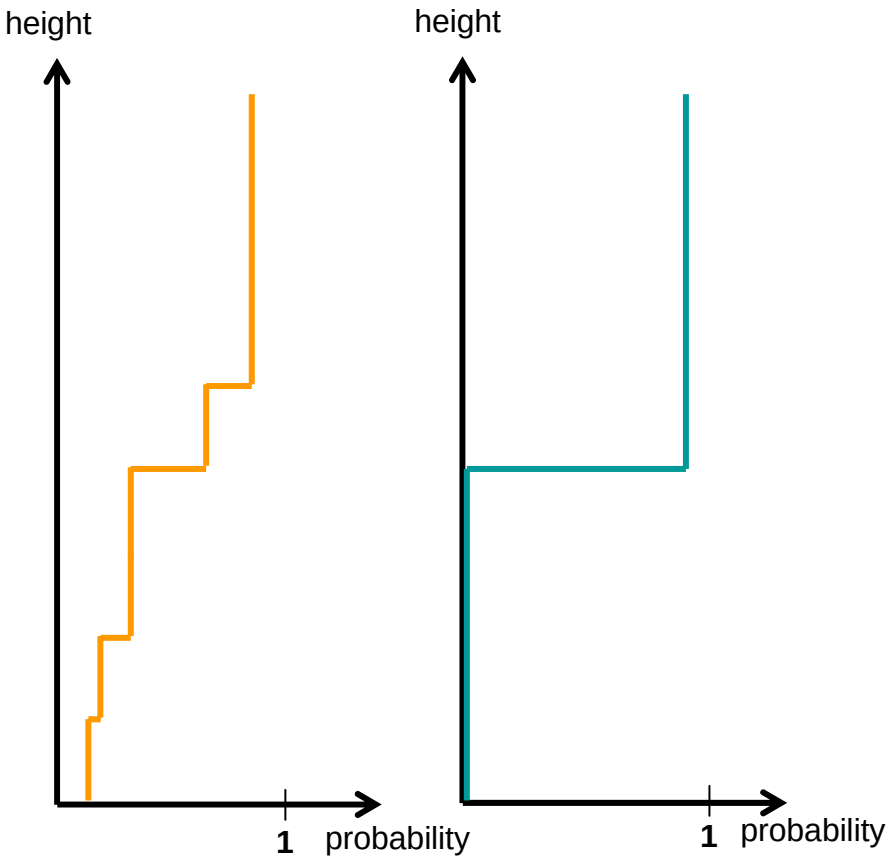
$$BS_k = \frac{1}{N}\sum_{i=1}^{N}\left(f_{i,k} - o_{i,k}\right)^2$$

Ranked probability Score (RPS)

Cumulative probability functions:

forecast
$F_{i,k}$

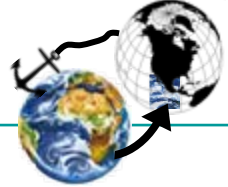
observation
$O_{i,k}$



$$RPS_i = \frac{1}{K}\sum_{k=1}^{K}\left(F_{i,k} - O_{i,k}\right)^2$$

$$RPS = \overline{RPS_i}$$

Outline of the paper



1. Introduction
2. Data and Method
3. Evaluation of deterministic cloud base estimates
4. Probabilistic verification
5. Sensitivity studies to optimize RPS
6. Discussion and Conclusions

Data:
- Ceilometers in Hamburg and Cabauw
- 2 model runs (CEU 00UTC+24h; CDE 03UTC+21h)
- Observations derived from perfect model forecasts
- Period: 2007-2008

Deterministic estimates:
- Consider various estimates: 5%, 50%, 100%, mean of PDF, median of PDF …
- … and compare them with instant observations, temporal mean and synthetic observations.
- BIAS, RMSE, Korrelation et al.

Probabilistic verification:
- BS and RPS
- Extent BS to skill score by using “perfect observations” as reference (probably analytically)

Sensitivity studies:
- What kind of modification increases RPS: stretching of CLC profile, amplification of CLC, …