CM SAF
Climate Monitoring

SMHI

Extending error characterization of cloud masking: Exploring the validity and usefulness of the SPARC-type and Naïve Bayesian probabilistic cloud masking methods

Fourth Cloud Retrieval Evaluation Workshop
4-7 March 2014, Grainau, Germany

Karl-Göran Karlsson
Contributions from Erik Johansson, Abhay Devasthale and Adam Dybbroe
SMHI, Norrköping, Sweden

BACKGROUND

- **Estimation of various geophysical parameters from satellite data still relies heavily on accurate cloud screening**

- **Ways of estimating the uncertainty in cloud masking and the implications for down-stream derivetion of geophysical quantities are urgently needed**

- **Different applications have different demands on cloud screening – especially, some do not accept any residual clouds (e.g. SST estimations)**

➡ **STRONG DEMAND OF PROBABILISTIC FORMULATIONS!**

Probabilistic theory

- **Well established - Thomas Bayes published first ideas on this already 1763 leading later to the famous Bayes Theorem:**

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

P(A|B) = (Posterior) Probability of event A given event B
P(B|A) = (Likelihood) Probability of event B given event A
P(A),P(B) = (A priori) Probability of event A and B

For image cloud masking appliations A will be ”Cloudy” and B an image feature vector of N dimensions

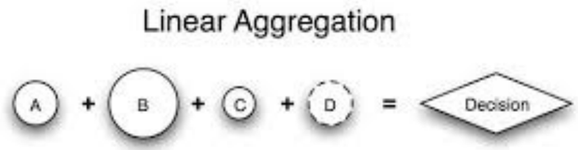
Theory established – so why the limited progress in applications?

- To calculate probabilities or likelihood you need to know the truth!
- Cloud occurrence is stochastic meaning that it cannot easily be modelled
- Current methods have used limited regional datasets for training
- Training often based on supervised methods – leading to subjective or inconsistent treatment

.but….recent access to A-train data (especially CALIPSO -CALIOP cloud mask) offers an objective or at least consistent training!

Results from two different approaches to be evaluated here:

- **Cloud probabilities derived as a weighted sum of conditional cloud probabilities (SPARC-like approach, here called PPS-Prob SPARC)**



(*Khlopenkov and Trishchenko, 2007, JAOT*)

Weights calculated as integrated ability to provide probabilities away from 50 % summed over the whole feature domain

- **Cloud probabilities calculated from a simplified Bayesian classifier (Naïve Bayesian – here called PPS-Prob Naïve)**

$$P(C_{yes} \mid F) = \frac{P(C_{yes}) \prod_{i=1}^{N} P(F_i \mid C_{yes})}{P(F)}$$

(*from Heidinger et al., 2012, JAMC*)

Assuming independent features allows probabilities to be calculated as multiplied conditional probabilities

Training aspects for the classifiers:

Training dataset with CALIPSO data:

- 99 matched full global orbits 2007-2009 for NOAA-18/CALIPSO

- SNO time difference less than 15 seconds leading to a maximum time deviation of approximately 2 minutes

- Combining 1 km and 5 km CALIPSO cloud layer datasets (method described by Karlsson and Johansson, 2013, AMT)

- Only using CALIPSO clouds with cloud optical thickness above 0.2! (<u>You should avoid training against noise!!!</u>)

- Still an imperfect truth: Some very thin but CALIOP-detected clouds will be missed by AVHRR-based method
  ➔ Derived cloud probabilities will be too high!

Experimental setup – Training philosophy:

Training philosophy:

Since image feature values depend on a long range of factors

(e.g., solar zenith, angle, satellite viewing angle, azimuth difference angles, total atmospheric moisture content, surface reflectivities and emissivities, snow/ice-cover, etc)

try to piggyback ride on NWC SAF multispectral thresholds (already accounting for these factors) when training the probabilistic classifier!

Probabilities are estimated against the feature difference from PPS thresholds!

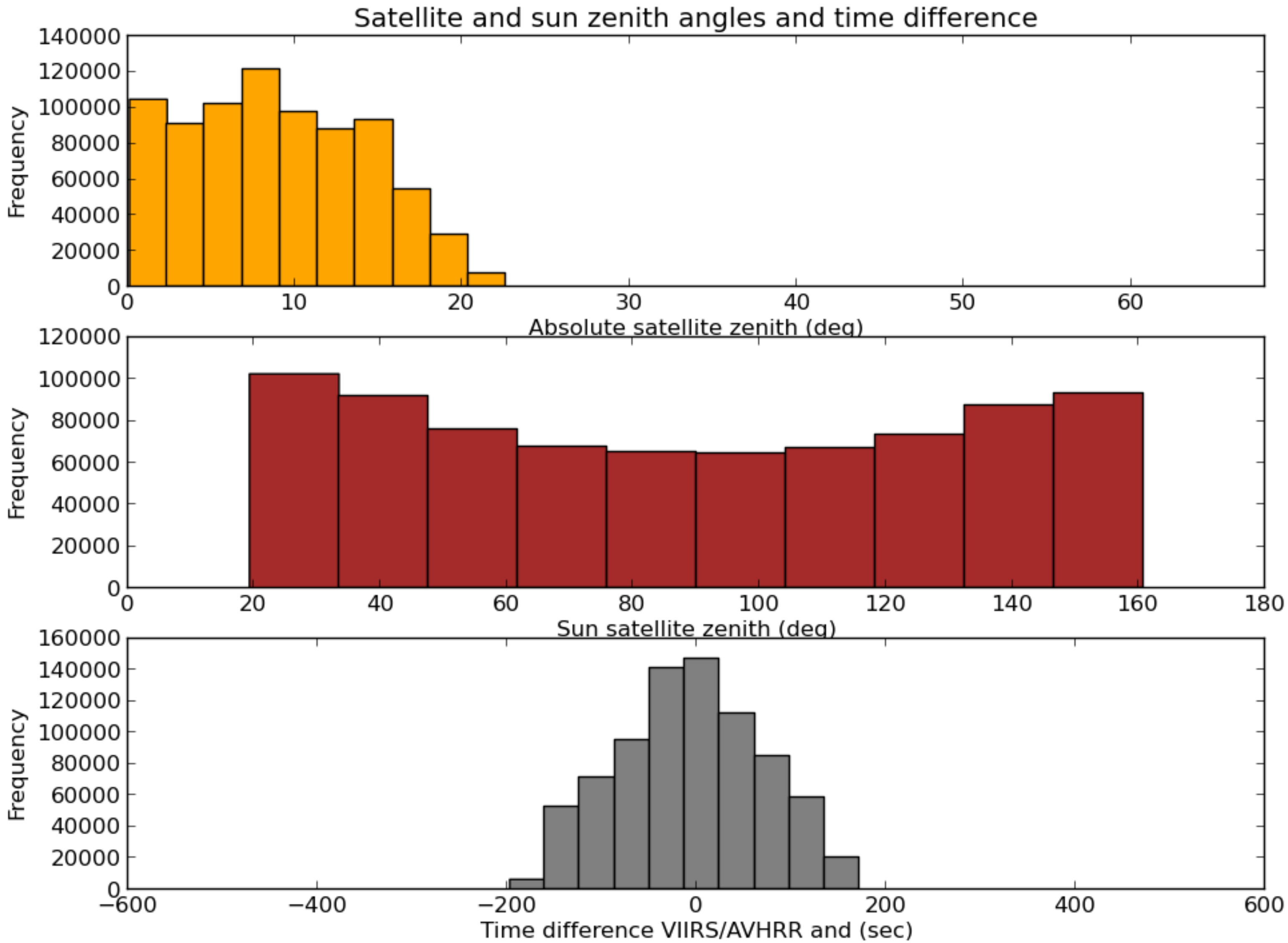
Validation of results:

Two approaches:

1. Evaluate results from derived cloud masks (probability thresholding at 60 %) against independent CALIPSO data (2010 dataset with 79 matched NOAA-18 and NOAA-19 orbits)

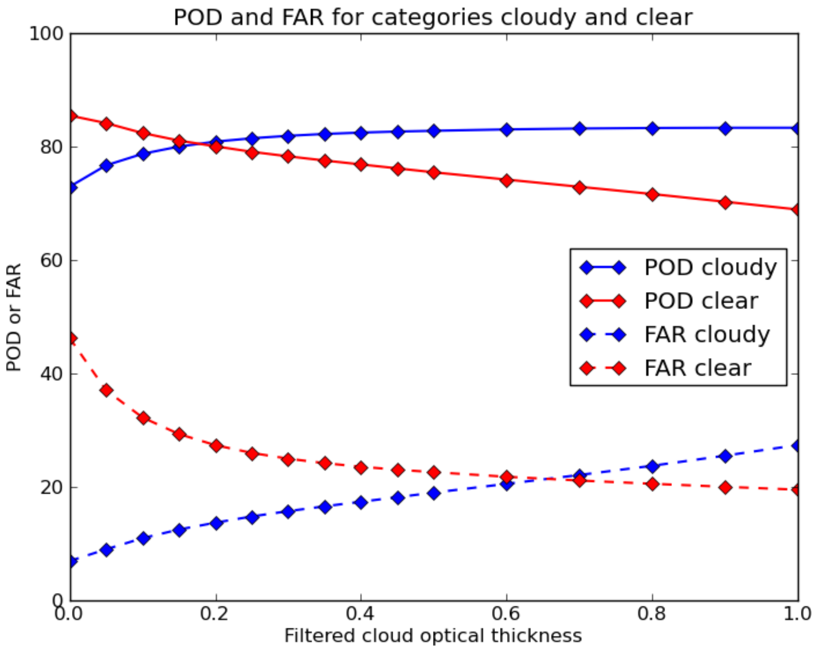
2. Evaluate results against surface observations (merged remote sensors Cabauw) and SEVIRI cloud masks (NWC SAF MSG Cloud Software version 2012) for checking validity at higher viewing angles (trained classifiers based on near-nadir data).

Viewing and solar angles and time difference for training dataset

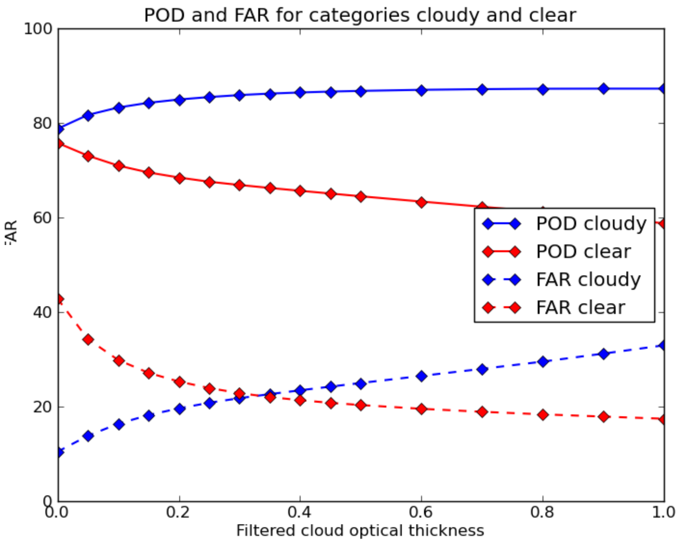


Plot by Adam Dybbroe, SMHI

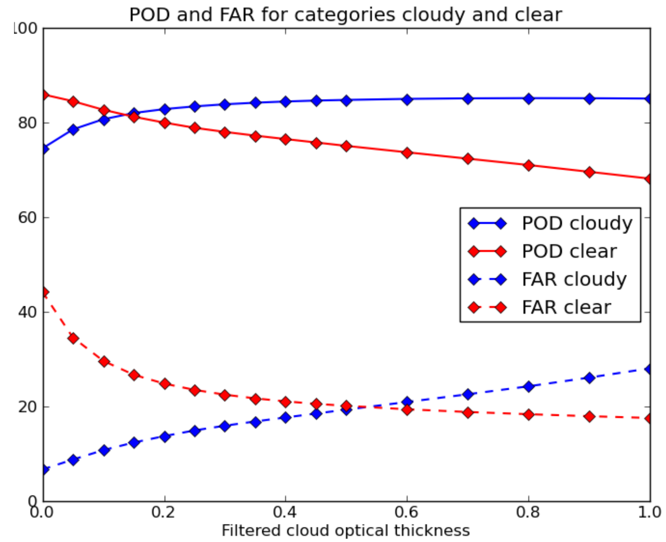
CALIPSO results in 2010 for POD and FAR:
- All cases, indep. dataset, prob threshold 60 %



PPS v2010+3 patches
(Version in CLARA_A1)

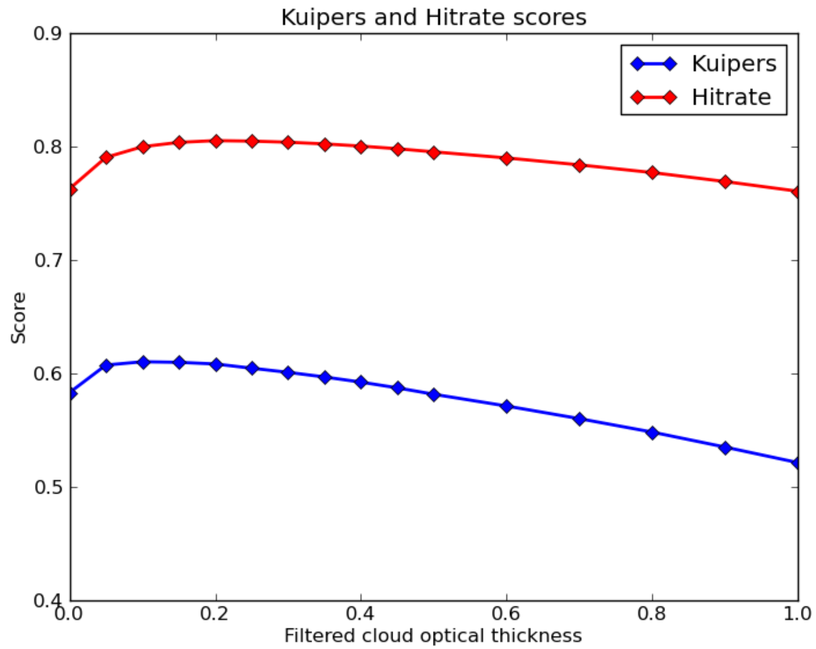


PPS_prob indep:
SPARC 60 %

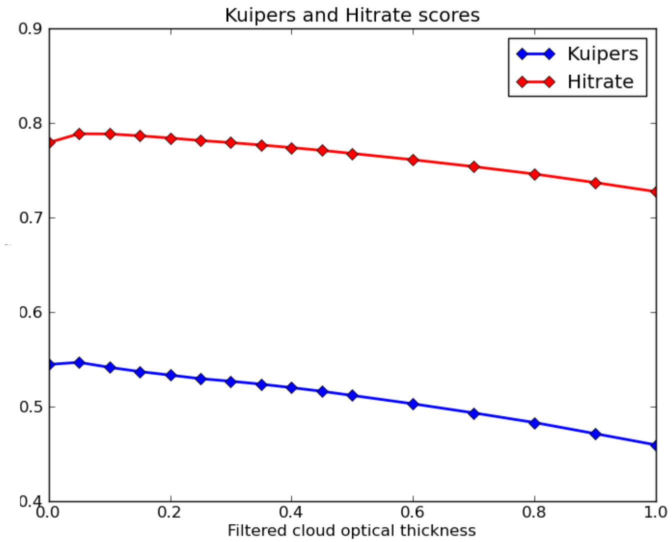


PPS_prob indep:
Naive 60 %

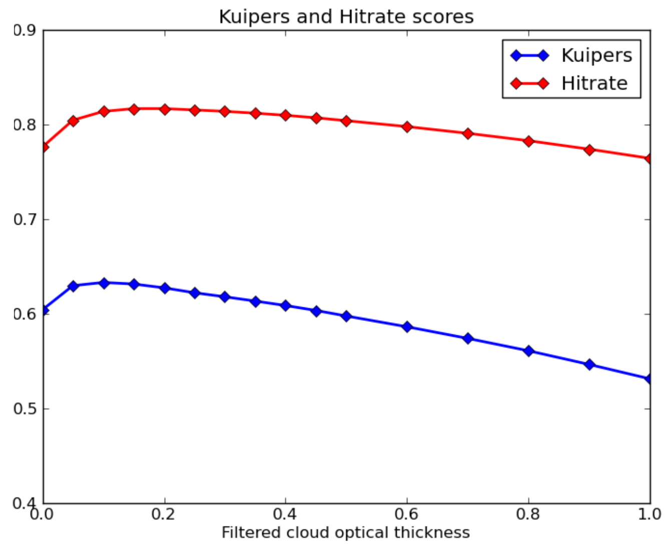
Results for Hitrate and Kuiper's skill score:
- All cases, indep. dataset, prob threshold 60 % bugfixed



PPS v2010+3 patches
(Version in CLARA_A1)



PPS_prob indep: SPARC 60 %



PPS_prob indep: Naive 60 %

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 115, D24116, doi:10.1029/2010JD014661, 2010

Optimized fractional cloudiness determination from five ground-based remote sensing techniques

R. Boers,[1] M. J. de Haij,[1] W. M. F. Wauben,[1] H. Klein Baltink,[1] L. H. van Ulft,[1] M. Savenije,[1] and C. N. Long[2]

Received 22 June 2010; revised 30 September 2010; accepted 18 October 2010; published 23 December 2010.

[1] A 1 year record of fractional cloudiness at 10 min intervals was generated for the Cabauw Experimental Site for Atmospheric Research (CESAR) (51°58′N, 4°55′E) using an integrated assessment of five different observational methods. The five methods are based on active as well as passive systems and use either a hemispheric or column remote
sensin ... ainst a
30 ye ... In the
interm ... ods, report
lower ... es from the
30 ye ... he 1–2 octa
range ... clouds.
Durin ... the
instruments. A reference algorithm was designed to derive a continuous and optimized record of fractional cloudiness. Output from individual instruments were weighted according to the cloud base height reported at the observation time; the larger the height, the lower the weight. The algorithm was able to provide fractional cloudiness observations every 10 min for 99.92% of the total period of 12 months (15 May 2008 to 14 May 2009).

- 2 ceilometers (low and high range)
- CLOUDNET radar/lidar
- NubiScope - longwave scanning radiometer
- Infrared pyrgeometer – APCADA
- Total Sky Imager

Citation: Boers, R., M. J. de Haij, W. M. F. Wauben, H. K. Baltink, L. H. van Ulft, M. Savenije, and C. N. Long (2010), Optimized fractional cloudiness determination from five ground-based remote sensing techniques, *J. Geophys. Res.*, *115*, D24116, doi:10.1029/2010JD014661.

Validation against Cabauw merged remote sensing observations 15 May 2008-14 May 2009

Two inter-comparison datasets:

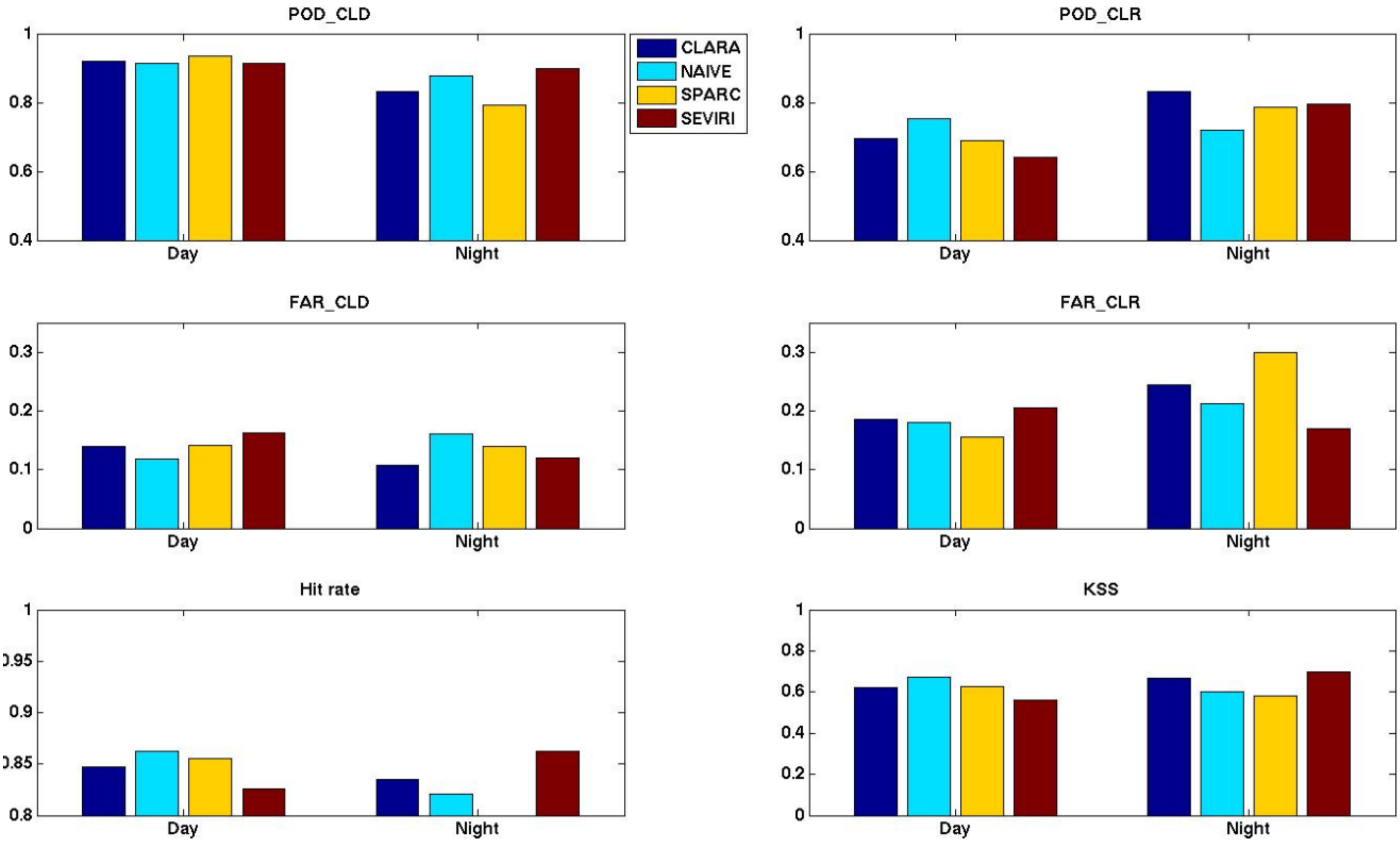
1. **NOAA-18 PPS-Prob Naïve and PPS-Prob SPARC probabilistic cloud masks**

   - One afternoon and one night observation (orbit) per day ➔ 730 cases
   - Results for 3x3 pixels over Cabauw position studied

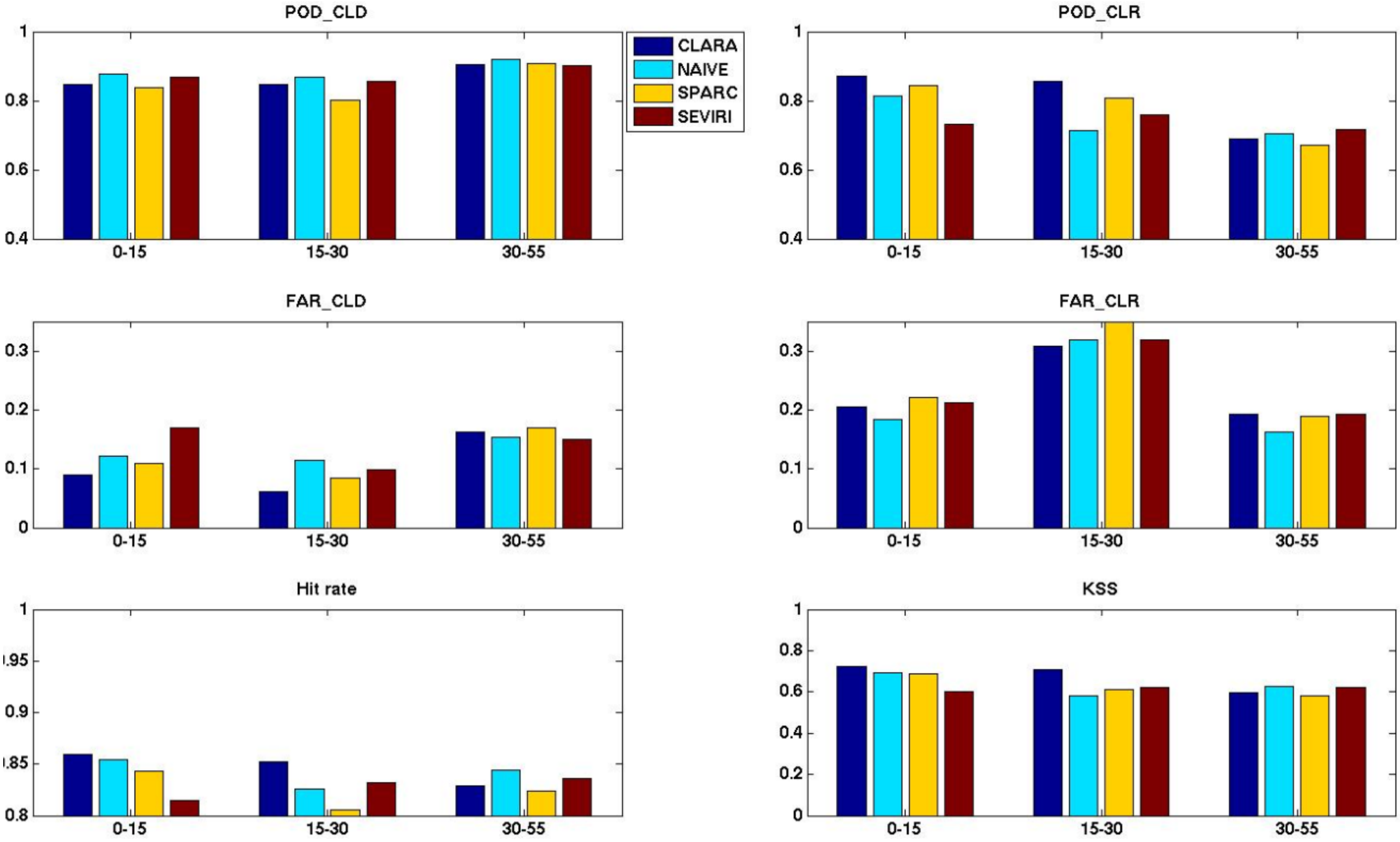
2. **SEVIRI cloud masks from NWC SAF MSG cloud software version 2012**

   - For each of the 730 cases, the closest in time (within 15 min) SEVIRI observation (3x3 pixels centred over Cabauw) was selected

Summary of Cabauw validation results separated into day and night portions



Summary of Cabauw validation results separated into three viewing angle categories



CONCLUSIONS

- **Reliable cloud amount estimations are possible with probabilistic approaches!**

- **The Naive Bayesian approach gives best results**

- **Results appear comparable and even sligthly superior to existing multispectral thresholding schemes during daytime**

- **Night-time results still slightly inferior to multispectral schemes for Cabauw study but not for global CALIOP (comparable or better)**

- **Only weak decrease of results with satellite viewing angle**
  ➔ PPS-Prob training concept with pre-calculated thresholds holds!

- **Next steps: Adding PPS-Prob Naïve products as complementary products to CLARA-A2 and PPS Version 2016.**