



Koninklijk Nederlands  
Meteorologisch Instituut  
*Ministerie van Verkeer en Waterstaat*

# Probability of Cb and Tcu occurrence based upon radar and satellite observations

Paul de Valk and Rudolf van Westrhenen

De Bilt, 2010 | Scientific report; WR 2010 – 04



# **Probability of Cb and Tcu occurrence based upon radar and satellite observations**

Paul de Valk  
Rudolf van Westrhenen



## Contents

0. Summary	1
1. Introduction	2
2. Theory	3
2.1 Convective clouds.	3
2.2 Cb/Tcu detection	4
2.3 Verification	6
3. Convection detection algorithms	8
3.1 operational algorithm-2007 and possible improvements	8
3.2 Convective cloud in satellite imagery.	9
3.3 Algorithm development	10
3.4 Logistic regression	11
4. Data use and predictor selection	14
4.1 Dataset Choices and restrictions	14
4.2 Dependent and independent data	15
4.3 Predictor selection	16
4.4 Conclusions from predictor selection	20
5. Detection results and evaluation	21
5.1 Results	21
5.2 Summary of results	26
6. Conclusions and future	30
6.1 Conclusions	30
6.2 Future	30
Literature	32
Appendix 1 Recommendations towards application.	34
Appendix 2 ICAO METAR	35
Appendix 3 Attributes diagrams and histograms	37
Appendix 4 POD and FAR diagrams and applied coefficients	49
Appendix 5 Verificatie autometar	52
Appendix 6 Nota AVW L beoordeling autometar	61

## Acknowledgements.

This work is accomplished by the valuable contribution of our colleagues. Their knowledge and feedback improved the obtained results and report. So we like to thank, Maurice Schmeits, Janet Wijngaard, Rob Roebeling, Cintia Carbajal-Henken, Kees Kok, Joop Konings, Pieter Luijendijk, Pieter Arts, Wiel Wauben, Jan Hemink, and Martin Stam. Han The performed a study in 2006 over the results in 2005, which is included in the appendix and used as a reference.



## 0. Summary

The detection of Cumulonimbi (Cb) and towering cumuli (Tcu) is relevant for aviation as they are associated with hazardous flight conditions. Their detection is therefore a requirement by ICAO. Since 1-8-2007 an operational algorithm, referred to as algorithm-2007, is used at the airports EHBK and EHGG to detect Cb and Tcu. It uses the radar reflection observations and lightning observations as input. The performance of the algorithm-2007 is poor in terms of probability of detection (POD) and false alarm ratio (FAR). At KNMI this study was initiated to develop an improved algorithm.

An automated Cb-Tcu detection algorithm based on the synergy between radar and satellite observations is developed. The algorithm uses logistic regression to determine the probability of Cb-Tcu occurrence. Within logistic regression a forward stepwise approach is applied. The predictors selected by the forward stepwise regression method are related to the highest radar contour occurring in the 15 and 30 km radii collocation area, and to the satellite observations, reflection range of the high resolution visible channel, the cloud temperature and its standard deviation. The latter three all in the 15 km radius collocation area.

The obtained results show in general an improvement in performance of the developed algorithm in comparison to the algorithm2007 results.

The performance of the developed algorithm is dependent on season and day-night conditions. The best performance is achieved in the Summer day category followed by the winter day category, with the summer defined from April till October. Surprisingly the summer night category shows the worst performance.

Although the algorithm is developed for EHBK and EHGG no year round evaluation of the performance of the newly developed algorithm was possible for those airports because of the lack of sufficient Cb occurrences, which are required for a statistical analysis. Especially for the EHBK airport data was lacking. This hampers a successful operational application of the developed algorithm for EHBK.

Note that since there is no other observation which covers both the required spatial and time dimensions a future assessment of the performance of the algorithm is disabled. The METARs are the most reliable source of Cb and Tcu observation, but they are terminated at EHGG and EHBK. At EHAM and EHRD they are still continued.

Based on the results an improved operational algorithm can be defined. The probability threshold selection will determine the performance of the developed algorithm. For the daytime categories a POD of 65 % and a FAR of 35 % appears feasible in the summer and winter day categories. For the night time a POD of 55 % and a FAR of 45 % appears achievable.

During the study it became clear that in the algorithm-2007 the evaluation area with a radius of 30 km on June 30, 2008 is decreased to an area with a radius of 15 km. The decrease in area leads to a lower FAR, but also to a more significant loss in POD when compared to the METAR. It is recommendable to evaluate the effect on the results of this radius change for the algorithm-2007 with a data set covering an area with a radius of 15 km.

## 1. Introduction

The occurrence of strong turbulence forms a hazard for aviation. Observations of turbulence are a prerequisite for safe aviation conditions, especially around airports. Here an unexpected vertical movement of an aircraft can have serious consequences. Direct observations of turbulence are not common in meteorology. There are indirect methods using for example the radiosondes to determine the stability and the likelihood of turbulence. The radiosonde network, however, has a drawback: it is too coarse in spatial - temporal resolution.

Another indicator of turbulence though indirect is the occurrence of convective clouds, towering cumuli and Cumulonimbi hereafter referred to as Tcu and Cb. Convective clouds may vary from fair weather cumuli or cumuli humulis, to tornado generating super cells. Embedded cumulus can grow from stratocumulus. For aviation at aerodromes in the Netherlands the embedded cumuli, the towering cumuli and Cumulonimbi are relevant. Not only the turbulence associated with these clouds can form a hazard to aviation, but also the associated precipitation, super cooled water occurrence and lightning can be a threat.

It is therefore a primary requirement by ICAO to include the occurrence of Cb or Tcu in the METAR (Meteorological Aerodrome Report or MÉTéorologique Aviation Régulière) of an airport to limit the risks for aviation. The METAR report is predominantly given by an observer.

In 2007 an automated Cb/Tcu detection system, hereafter referred to as the operational algorithm-2007, replaced the observers at two smaller airports in the Netherlands: Groningen airport and Maastricht Aachen airport. The algorithm-2007 uses radar and lightning observations. Its performance is not optimal, a study by The ( 2006), showed a probability of detection of 50 % and a False Alarm ratio of 70 % as averaged values over the whole year.

This report describes a study initiated at KNMI to develop an automated detection algorithm that will have a better performance than the algorithm-2007. The presently proposed algorithm is based on a synergy of both radar and satellite observations. The satellite information is provided by the SEVIRI imager on the Meteosat satellites operated by EUMETSAT. The radar information stems from the two operational radars used at KNMI.

The goal of the study is an algorithm that detects Cb/Tcu in all seasons with a relatively low false alarm and high probability of detection at four different airfields.

A master thesis study performed in the same period as this study within the weather research department overlaps with this work, (Carbajal-Henken et al.,2009). Carbajal-Henken studied the summer season for one airfield but for four years. The thesis study indicated logistic regression to be a successful approach in the classification of Cb-Tcu.

This report describes the background theory on convection, observation methods and verification methods. In the third section the Cb detection methods are described including the operational one, methods applied in the literature, and the developed algorithm based on logistic regression. The fourth section describes the data used in the algorithm development and the fifth section the obtained results. The last section gives the conclusion and considerations for future research.

## 2. Theory

This chapter describes the process of convection and the observation methods of convection. The chapter finalises with the verification scores to assess the performance of the detection methods.

### 2.1 Convective clouds.

Convection in fluid dynamics refers to the transport of heat and momentum. There is a wealth of literature on this subject, indicating its importance. For further reading we refer to S Petterssen (2008) or Mcintosh and Thom (1969).

Convection in the atmosphere occurs when a parcel of air is distorted or displaced. This can have an impact on the temperature of the parcel. The parcel has a temperature and a moisture content. The surrounding air may have a different temperature and moisture content. When the parcel becomes positively buoyant, it will accelerate upwards. The movement will initiate a number of processes. The rise of the parcel will lead to an expansion of the parcel due to the decreasing pressure. The expansion will cause a temperature decrease. The temperature decrease will continue until the moment that the moisture in the parcel experiences a phase change: condensation of moisture to droplets occurs. The condensation of moisture will produce heat, increasing the temperature of the parcel. The released heat is referred to as latent heat. This will contribute to the buoyancy of the parcel. A further drop in temperature may introduce another phase change: water to ice, again with the release of latent heat. The extent of the rise of the parcel is restricted by the amount of latent heat release and the temperature profile of the surrounding air.

The rise of the parcel will produce an exchange of air with the surrounding air. This exchange is referred to as either entrainment or detrainment. The amount of exchange will impact the energy content and temperature of the whole parcel.

The formation of rain is mainly due to (auto)coalescence when the temperature of the air parcel is higher than 273.15 K. When the temperature drops below 273.15 K mixed phase hydrometeors can occur where the Wegener Bergeron Findeisen(WBF), Pruppacher and Klett, (1980) process dominates the precipitation formation. In the WBF process ice crystals grow at the expense of liquid droplets due to the difference in saturation pressure of water vapour above ice and water surfaces. The difference in fall velocity enables an aggregation process resulting in growing hydrometeors.

When the fall velocity of hydrometeors is larger than the up-draft motion precipitation will start. The drag velocity of the hydrometeors will impact the energy content of the air parcel.

When ice particles become present in the air parcel, a process of charging can occur. This may lead to lightning. Lightning is also regarded as a hazard for aviation.

The spatial horizontal dimension of this process can vary from small scales (~100 m) to extensive scales of multi cells or squall lines(~ 10-100 km). The vertical scales at the latitude of the Netherlands vary from fair weather cumuli (~ 1km ) to mature Cb rising up to the tropopause at 12~13 km. The micro-physical scales involved vary from nanometre, smallest activated cloud condensation nuclei, to centimetres, hail stones, Pruppacher and Klett (1980). The life cycle of the process can vary from 20~30 minutes, for single cell thunderstorms to circa 6 hours for organised multi cell storms or squall lines. The whole process is non-hydrostatic, and includes non steady turbulence, Holton (1975). A description would require a full three-dimensional representation of all processes at all relevant scales. The complexity of the process is hard to capture in numerical weather prediction (NWP) models. The grid on which the NWP models perform the calculations to forecast the weather is (still) too coarse to capture all scales of convection.

NWP may describe the vertical temperature and humidity profile adequate, and NWP could forecast favourable conditions for convection but it will most likely fail to forecast the correct location of convection initiation; see for example Zbyněk Sokol and Petr Pešice (2009).

NWP may however provide valuable information for the algorithm of the level at which condensation occurs (the lifting condensation level: LCL). This LCL can not be observed from satellite or radar observations. In an algorithm development the LCL could be used as an estimation of cloud base height, there were ceilometer observations fail to observe the cloud base height.

## 2.2 Cb/Tcu detection

### *METAR*

The METAR (Meteorological Aerodrome Report) is produced every thirty minutes. It is issued at 25 and 55 minutes past the hour. The METAR reflects the weather conditions in the vicinity of the airport ten minutes previous to the moment of reporting.

*The AERODROME METEOROLOGICAL OBSERVATION AND FORECAST STUDY GROUP (AMOFSG) February 2010 states that: "VCTS Thunderstorm in the vicinity are a "Primary requirement (thunderstorm) (ICAO Standard), but requires remote sensing to provide this in automated reports, requiring substantial work from many States to comply."*

*It is recognized by this group that it is a challenge to automate the detection of Cb/Tcu.*

ICAO prescribes the format and the content of the METAR, see for details the appendix 2. The clouds part includes the vertical visibility, the coverage at several layers, when observable, and the cloud base height. Relevant and mandatory to report is the occurrence of Tcu or Cb. When a Tcu and a Cb occur at the same cloud level the observer shall report only Cb.

The time required to develop from a Tcu to a Cb is relatively short in comparison to the total life cycle of a Cb. The observation frequency of Tcu is therefore considerably lower compared to the Cb occurrence.

### *Radar (Radio detection and ranging)*

In the Netherlands two Doppler radars are operated primarily for precipitation detection. The C-band radar emits and receives pulsed 6 Ghz radio waves with a wave length of around 5 cm. The lowest inclination of the radar beam is 1 degree. Therefore the part of the atmosphere not observed by the radar increases with the distance to the radar position.

The observed reflections are obtained from a distance from the Radar site (varying from 0-320 km) and at moderate altitude (0.8-3 km) above surface of the earth. The reflection signal is proportional to the sixth power of hydrometeor diameter, when the particles are smaller than the wavelength, Holleman (2000). Due to the sixth power the variance of the reflectivity value is huge. Therefore a decibel or logarithmic unit is used to represent the signal. The radar reflections are projected on a grid with grid cells of 2.5 by 2.5 km.

<b>Z [dBZ]</b>	7	15	23	31	39	47
<b>R [mm/h]</b>	0.1	0.3	1	3	10	30

*Table 2.1 Relation between radar signal and rain rate*

In the table 2.1 a few examples of reflectivity Z values and corresponding precipitation rates R are given.

The KNMI uses the following equation to relate reflections to rain rate:

$$Z = 200 * R^{1.6}$$

with R in mm/hr.

It is relevant to note some considerations about radar observations in relation to Cb/Tcu detection:

- the operational radars are sensitive to precipitation and not to cloud occurrence. Therefore developing convection without precipitation can not be observed by the radar. Hence the probability is small that the radar will observe Tcu correctly.
- Additionally the radar cannot distinguish between heavy non convective precipitation or convective precipitation. This may lead to false alarms when for example strong frontal related precipitation occurs.

### *Satellites*

Meteorological satellites provide an instantaneous view of the atmospheric state. The geostationary satellites are an invaluable source of information for nowcasting. The latest generation of operational geostationary satellites provides an image each 15 minutes over Western Europe. They are operated by EUMETSAT. The Spinning Enhanced Visible and Infrared Imager (SEVIRI) on board the METEOSAT 8 and its follow-on observes the world since January 2004. SEVIRI is a passive instrument, it does not emit a signal, opposite to the radar. SEVIRI observes the reflection of the earth in spectral bands from 0.5  $\mu\text{m}$  to 3.9  $\mu\text{m}$  and the emission from the earth in spectral bands ranging from 3.9  $\mu\text{m}$  to 13.4  $\mu\text{m}$ . Next to the eleven spectral bands, there is a high resolution visible (HRV) channel, 0.4  $\mu\text{m}$  - 1.1  $\mu\text{m}$ . The sampling grid distance in the nadir point of the satellite is 3 km for the eleven channels and 1 km for the HRV channel.

The observation cycle consists of a 12.5 minutes scan of the earth from south to north. Then the scan mirror returns to its starting position and calibration occurs in 2.5 minutes remaining from the 15 minutes cycle.

Further details on the satellite platform and the SEVIRI instrument can be found at [www.eumetsat.int](http://www.eumetsat.int).

It is relevant to note here some consideration about satellite observations in relation to Cb/Tcu detection:

- Satellite view is obscured when higher cloud layers block the view to the lower atmosphere. Cirrus may hamper a correct interpretation of the satellite data.
- The lack of the HRV and other reflection channels in the night period, when there is no insolation, affects the detection of clouds.
- The satellite only observes the top layer of the cloud.
- The horizontal spatial resolution degrades when moving away from the nadir point. At the latitude of the Netherlands, the spatial resolution is approximately 3.5 km West East and 6 km North South, for the 11 channels and 1.2 by 2  $\text{km}^2$  for the HRV channel. Clouds smaller than the pixel size can not be classified correctly.
- One should correct for the slanted view of the satellite to collocate radar and satellite signals when both are used. A correction requires shifts up to several radar pixels.

### Other observation methods

Due to the strong discontinuity in appearance of convective clouds, point measurements will not contribute to a successful detection. However one could consider additional observations to obtain certain cloud properties, not observable by radar or from satellite platforms.

The ceilometer, based on lidar technology, provides information about cloud base height

and vertical visibility. But it can not classify clouds as convective or not. Hence the observed cloud base height does not always relate to the convection occurrence in a mixed cloud situation.

The 2 meter air temperature can give information for a threshold for cloud masking when using the brightness temperatures of the satellite observation.

The 2 meter air temperature combined with the dew point temperature can give an estimation of cloud base height.

The so called SAFIR network provides information about lightning. The lightning detection was shown to be a non significant contributor to Cb detection in the evaluation study done over 2005 on the operational algorithm-2007, The (2006). Lightning is also associated with significant convection, where this study also aims to detect early stages of convection.

### 2.3 Verification

Cb/Tcu occurrence is a dichotomous phenomenon. The frequency of Cb/Tcu occurrence is relatively low in comparison to the total number of METARs. The value of a forecast or classification can be assessed by comparison to an observation. Frequently used for assessment is the contingency table, table 2.2 (Wilks 1995). Here the occurrences of forecast/classification in comparison to observations are represented.

	observed yes	observed no
classified yes	hits	false alarms
classified no	misses	correct negatives

*Table 2.2. Contingency table, (Wilks 1995). Relationship between the number of observed and classified cases of a dichotomous phenomenon. The sample size is the sum of the hits, misses, false alarms and correct negatives.*

From the table a number of scores can be calculated. Given the large number of correct negatives for this specific Cb-Tcu classification this number is not incorporated in any of the scores used in this report. It may lead to an incorrect interpretation of the results.

Considered are, the Probability of Detection (POD), The False Alarm Ratio (FAR) the Critical success index (CSI) or threat score, and the BIAS.

$$POD = Hits / (Hits + Misses)$$

$$FAR = False Alarms / (Hits + False Alarms)$$

$$CSI = Hits / (Hits + Misses + False Alarms)$$

$$BIAS = (Hits + False Alarms) / (Hits + Misses)$$

The BIAS is a ratio of the observed events and the classified events. The bias is not an accuracy measure. It states whether the event is classified more (bias >1) or less (bias < 1) than observed. Bias is 1 only states that Hits and Misses are in balance with Hits and False Alarms.

#### Brier Score

To assess the performance of probabilistic forecasts it is convenient to capture it in a scalar number Wilks, (1995). The most commonly used is the Brier score (BS). The score

is the average of the squared differences between the forecast probability  $c$  and the dichotomous observable  $o$ .

$$BS = \frac{1}{n} \sum_{k=1}^n (c_k - o_k)^2$$

with  $n$  the sample size and  $k$  the index.

The Brier score is negatively oriented. The performance of an algorithm improves when the  $BS$  decreases.

From the Brier score a skill score  $BSS$  can be computed.

$$BSS = 1 - \frac{BS}{BS_{ref}}$$

with  $BS_{ref}$  the reference  $BS$ , usually the climatological occurrence frequency. When the  $BSS$  becomes negative the  $BS$  is worse than the climatological  $BS_{ref}$ . A  $BSS$  close to 1 indicates a very good performance. A  $BSS$  equal to 0 indicates that the  $BS$  and  $BS_{ref}$  have an equal score, so that the forecast does not contribute significantly relative to the climatological  $BS_{ref}$ .

### 3 Convection detection algorithms

This chapter describes the operational algorithm-2007 as developed in 2005 and implemented in 2007, the evaluation of its performance and the requirements by the end users. The second part summarises a literature study on Cb detection in applied algorithms. Based on the gained knowledge from the first two parts an algorithm is developed in this study that is described in the last two parts of this chapter.

#### 3.1 operational algorithm-2007 and possible improvements

In 2005 a detection algorithm for Cb/Tcu detection based on radar signals was rapidly implemented at KNMI, based on Kucharski 2005. The algorithm should detect the Cb and Tcu occurrences around the airport reference point (ARP) with a performance as good as the reports of the observers.

A decision table 3.1 is defined for a Cb/Tcu classification from a single image. In the applied algorithm two consecutive images are evaluated to come to a classification.

radar reflectivity in dBZ	sfr1>0	sfr2>0	no safir
41	Cb	Cb	Cb
33	Cb	Cb	Tcu
29	Cb	Cb	0
0	Cb	Cb	0
no significant signal	Cb	Cb	///

*Table 3.1 The decision table to come to a Cb/Tcu Classification. Radar signal should occur in a radius of 15 km around the station and at least at two connected radar pixels. Sfr1 denotes a safir lightning signal within the 15 km collocation area, sfr 2 is at a distance of 15-20 km to the station location. No safir means no lightning information near the station location.*

The implementation of the algorithm is correctly done in accordance with the thresholds given by Kucharski (2005). Both Kucharski and the KNMI algorithm obtain for Cb detection similar probability of detection (POD) and false alarm ratios (FAR) of 50 and 70 % respectively. For Tcu the scores were POD 25 % and FAR 99%. The latter scores are in line with the in section 2.2 described inability of radar to observe non precipitating clouds. Additional studies done to improve the algorithm did not lead to acceptable POD and FAR values, see The (2006) in Dutch in appendix 5.

During this study described here it became apparent that the applied algorithm in the 2005 version used a radius of 30 km for the collocation area. The radius was changed on June 30, 2008 to 15 km. No evaluation has been done on the performance of the algorithm-2007 with the new radius.

#### *Interviews*

Internal interviews were held at KNMI to elucidate the problem. The requirements of the end users is relevant. In the paper "Nota XAVW-L beoordeling Cb-Tcu in autometar-3", (2006), included in the appendix 6 (in Dutch) indicates that for safety a CSI of 90% would be desirable. A more realistic achievable CSI of 66 % related to a POD of 80 % and a FAR of 20 % as thresholds for an acceptable implementation is mentioned as a goal.

The internal interviews revealed that even looser thresholds would also be acceptable. Given the present performance of the operational Cb/Tcu autometar algorithm-2007 any significant improvement would be welcomed by the end user.

### 3.2 Convective cloud in satellite imagery

Cloud detection in satellites imagery is one of the major applications of satellite data interpretation. Tracking the motions of clouds from consecutive images arose as soon as such images became available, Fujita (1969). A significant number of studies is dedicated to cloud detection and identification, cloud work shop in Locarno, Thoss, (2009). The Satellite Application Facility on Now Casting (SAF-NWC, 2000) provides a cloud mask and cloud identification. The cloud masking is based on threshold technique applied to a selection of SEVIRI channels.

SAF-NWC also provides a Rapid Developing Thunderstorm (RDT,2000) product. Based on temporal analyses of the decrease rate of brightness temperatures, thresholds of the temperature and spatial growth of a cluster of cool pixels a thunder storm classification can be made. Lightning observations can be used to increase the discrimination between thunderstorms and other developing cloud systems. The RDT tracks the thunderstorms and predicts their future development and location.

Unfortunately no archived data of the SAFNWC products were available for 2005. Therefore the SAFNWC products were not included in this study, but they can be considered in future updates.

Severe or intense convection is a topic of many studies. V shaped patterns in well developed convective clouds were already discriminated in the imagery of the MVIRI the predecessor of SEVIRI, Levizzani V., Setvák M. (1996). The specific behaviour of the 3.7  $\mu\text{m}$  channel on the polar orbiter platform NOAA, AVHRR and its relation to convection and micro-physical processes was also described, Setvak (1989) .

Mecikalski and Bedka (2006) studied the precursor signals of convective initiation in day time imagery over the United States. They applied thresholds on the GOES (Geostationary Satellite) channels, on the difference between various channels and on the temporal development or trend of some channels. For three case studies a comparison to a radar network showed a correlation of 60-70 % in accuracy with radar signals larger than 35 dBZ.

Their method is only applied on day time imagery. For night time different methodologies are required which were not considered in their study.

For the next generation of geostationary satellites Mecicalsky (2007) wrote a report on expected performances with the observations. The next generation of satellites will become operational in 2017, so for the present autometar improvement it is not an option to look into the improved performances of this satellite generation.

Zinner et al. (2008) published a Cumulonimbus tracking and monitoring (Cb-TRAM) algorithm.

Their method identifies intense convection. It is based on thresholding the 6.2 and 10.8  $\mu\text{m}$  channels of SEVIRI. They also incorporate significant changes in reflection of the HRV channel into their analysis. The tracking algorithm determines the motion vectors on coarse pixel resolution. In an iterative process the pixel resolution is then stepwise increased improving the accuracy of the motion vector. They find an acceptable correlation with radar observations and recognize the ability of satellite observations to detect Tcu and Cb even before the precipitation formation process occurs.

Pattern recognition using neural networks requires for every specific study a well trained data set. The training of the dataset requires human supervision. It has been applied at the meteorological service in the UK, Pankiewicz (2001). As there are various atmospheric conditions in which Cb-Tcu convection occurs it is not straightforward to create a dataset for training and validation.

For more detailed background information on convection detection algorithms we refer to the master thesis of Carbajal-Henken (2009).

### 3.3 Algorithm development

Interviews with the KNMI R&D department of instrumentation revealed that there is no knowledge of instrumentation with proven ability to detect Cb/Tcu at time of writing available which can detect Cb/Tcu with a similar spatial coverage as an observer and at acceptable costs. The exploration into other instrumentation is therefore not pursued in this study. The consequence of this choice is that only radar and satellite observations can be used in this algorithm development.

The satellite based methods given in the literature all require at least a significant number of pixels to come to reliable statements on convection. Most of the discussed articles focus on severe or intense convection occurring frequently in the USA, Mecikalsk (2004), and mountainous areas in Europe, Zinner (2008). The early stages of convection are not captured by these algorithms.

The goal of this study is to detect both early and mature convection. The early convection will occur in a small number of pixels, with a low or no intensity in the radar signal. The detection of the early convection category is a larger challenge, in comparison to developed severe convection detection.

The algorithm presented here to meet this challenge, uses the synergy between radar and satellite information to come to classification between Cb/Tcu and non- Cb/Tcu cases. This implies that the radar information can not be used as a source for validation studies as done by others, e.g. Zinner (2008) and Mecikalski (2004). This a point of consideration for future evaluation.

The developed algorithm to detect the Cb/Tcu clouds is expected to be implemented in an operational environment. A direct interpretation of available observations from satellite and radar is preferred as it facilitates the communication to the end users on the behaviour of the classification algorithm.

This is in contrast with the work of Carbajal-Henken et al (2009) where a physical model is introduced which calculates cloud products from the observed SEVIRI radiances. These cloud products are used as predictors in her study.

The inclusion of cloud products can improve the detection performance but it also requires knowledge on the applied algorithm by the end user to interpret the classification.

As radar nor satellite observations can discriminate between Cb and Tcu both cloud types are treated as one category Cb/Tcu further used as the predictand in the algorithm development.

As the vicinity of the aerodrome is not uniquely interpreted two radii of collocation areas are considered, 15 and 30 km radii around the aerodrome point of reference. The data of both radar and satellite observations within the collocation areas are used in the algorithm.

Partly based on the literature studies a large number of predictors were determined from the original data. The data involved :

The original radar reflection given in dBZ

The satellite radiances expressed in reflection and brightness temperatures.

For two different radii 15 and 30 km of the collocation area the following variables were calculated as predictors:

-radar contours varying from 14 dBZ to 56 dBZ ( in 16 steps of 2.5 dBZ)

-satellite 10.8  $\mu\text{m}$  channel brightness temperature contours varying from 213 K to 258 K  
-satellite high resolution visible reflection contours corrected for the solar zenith angle varying from 59 % to 100 %.

From these distributions also the sum of the pixels, mean, median, minimum, and maximum values were determined, next to the maximum occurring contour, sum and a weighted sum of the occurring contours. The weighted sum here consists of the sum of the occurring contours multiplied by their order number, i.e. 1 x first contour + 2 x second contour + ...etc.

Additionally a rudimentary cloud mask was introduced for the smaller area (15 km radius). First the difference between the 12.0  $\mu\text{m}$  and the 10.8  $\mu\text{m}$  channel larger than -3K is evaluated to flag those pixels probably containing cirrus. For those pixels with a brightness temperature in the 10.8  $\mu\text{m}$  channel lower than the two meter air temperature minus 20 K and not flagged as cirrus contaminated the average temperature and its standard deviation is determined.

In a future version the SAF-NWC cloud mask could be implemented here leading to an improvement of both cloud mask and cirrus mask.

In the development study of the algorithm also other predictors derived from satellite observations were evaluated, e.g. difference between 6.2  $\mu\text{m}$  and 10.8  $\mu\text{m}$ , difference between 3.9 and 10.8  $\mu\text{m}$ , difference between 13.4 and 10.8  $\mu\text{m}$ . Also the difference between the reflection channels 0.6  $\mu\text{m}$  and 1.6  $\mu\text{m}$  was evaluated. Unfortunately these predictors did not show a correlation with the predictand of Cb-Tcu occurrence over the time period considered. The precursor signals as given by Mecikalski and Bedka, (2006), to study convective initiation were not found to have an explanatory power in this study. Presumably because the convection in their study is has a higher intensity than can occur in the mid latitude climate studied here. Convection with regular occurrence of super cells is a rare phenomenon in the Netherlands.

At the start of the study a hypothesis was that the rapid growing Tcu-Cb would give a clear signal in the development of the 10.8  $\mu\text{m}$  channel. A clear cooling of the cloud top would be detectable from consecutive images. Unfortunately this signal did not correlate significantly to Cb-Tcu occurrence in the study. Therefore the development of the 10.8  $\mu\text{m}$  channel was de scoped as predictor from the present version of the algorithm.

The lack of a clear development signal is possible related to the period in which a Tcu develops to a Cb. It is probably too short to be captured by a sampling frequency of 15 minutes.

Another explanation for the lack of successful classification by either the development or the differences as proposed by Mecikalski (2004) could be that the study period is not concentrated on the summer months July and August. The algorithms have to be applicable throughout the year, including the modest convection occurring in spring and winter. This limits the algorithm in the inclusion of predictors of severe but rare summer convection.

Pattern recognition in a neural network is considered as an applicable method. It requires a dataset for each airfield and each possible climate season. Also the number of Cb occurrences must be sufficient in each season.

Given the high dataset requirements of the neural network and the limited amount of data available this method is not pursued here in the development of this algorithm.

### **3.4 Logistic regression**

Nearly two hundred potential predictors are determined to classify the binary predictand: Cb/Tcu or non Cb/Tcu. A successful approach to come to binary results is the Logistic

regression, Wilks (1995) and Carbajal-Henken et al (2009). Logistic regression models result to a classification or prediction of a binary predictand while the predictor variables can be of any type. A non-linear equation can fit the predictand  $c$  using a multiple number of predictors  $x$ .

$$P(c) = \frac{1}{1 + \exp(-(b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n))}$$

With  $P(c)$  the probability that  $c$  occurs,  $b_i$  the regression parameters and  $x_i$  the predictor variables. The function is bounded between 0 and 1 due to its mathematical form allowing only for properly bounded probability estimates. The function drawn will always result in a S- shape curve.

Logistic regression is well known in social and medical sciences. In meteorological research it is commonly applied, e.g. for severe thunderstorm occurrence Schmeits et al (2008), or for contrail occurrence, Duda and Minnis (2009).

It is not possible a priori to indicate which predictors will lead to the best result in the desired classification. The dependencies and correlations between them are too complex. Commonly used is the forward stepwise regression, Wilks (1995). In each step a predictor is added to the equation and based on the statistical scores it is decided if the additional predictor contributes to the overall performance. It is up to the user to decide how many steps or predictors contribute significantly to the classification performance. Using all predictors may lead to an over-fit regression, Wilks (1995). In an over-fit regression too many predictors are used in the equation to describe the observations. The regression will fit to the used observations but the equation may fail to describe other observations not used for its determination.

To assess the performance of the obtained equation it is recommendable to split the data set in two parts: one part is referred to as dependent set, the other part is the independent set. By logistic regression predictor variables and regression parameters, also called coefficients, are determined on the dependent part. The performance of the derived predictors and coefficients are tested and evaluated on the independent set.

There are numerous statistical scores which can be determined to assess the overall goodness of fit. The Nagelkerke  $R^2$  is explained here, but there are more tests available: the Wald test, likelihood ratio test and the Hosmer-Lemeshow test for example Wilks (1995), Carbajal-Henken, (2009).

### Nagelkerke $R^2$ ( $NR^2$ )

In linear regression models the  $NR^2$  indicates the explained variance fraction. The  $NR^2$  is a modified Cox and Schnell coefficient and can be applied in the multiple regression used here. It indicates the proportion of the variance explained by the model (Nagelkerke, 1991). The  $NR^2$  coefficient can vary from 0 to 1. A higher value indicates a better performance.

### *Testing the continuous predictors*

After selection of a limited set of predictors with explanatory power an extra test can be applied on the continuous variables. The continuous variables should have a linearity in the exponential coefficient of equation, Hosmer and Lemeshow, (2000). To check this, the variables should be split into equal parts distributed over the value range of the variable, preferably evenly populated. The lowest value part serves as a reference state. For the other parts dummy variables equal to zero are introduced. The dummy variable will assume the value of one if the value of the independent variable (predictor) lies within the

range of the part of the associated dummy variable. The regression coefficients are determined for the dummy variables and plotted against the midpoints of the value parts. When the coefficients show a linear behaviour this is conform the theory that the values of the predictor have a linear explanatory power. The significance of the coefficient should be low. If the significance is too high the coefficient is not applicable. This may lead to coefficients only valid in a limited value range, e.g. a cloud top temperature can only be used as an predictor in the 240 to 270 K range. For non linear behaviour one may reconsider the relation of the predictor to the predictand: a square or root function of the predictor could give a linear behaviour. It is possible to expand this evaluation to combinations of predictors by multiplication.

## 4 Data use and predictor selection

This chapter describes the data and its limitations. The method used to come to an independent and dependent set is explained. From the data the predictors are derived and the obtained predictors are evaluated and discussed.

### 4.1 Dataset Choices and restrictions

To evaluate the classification of a relatively rare event a large dataset is preferable. In such a data set frequent sampling of different seasons occur. Previous evaluations of the autometar results for Cb detection, however, were based on data gathered in 2005. To enable a comparison to this evaluation it was decided to limit the dataset to 2005. At the end of the study the results of the algorithm-2007 had to be recalculated, so the comparison could have been done for a longer period. A comparison over a longer period, however, would require a substantial effort, for which the time was simply unavailable. Hence the studied period was kept to 2005 only.

As described in the algorithm development part the choice for a synergy of radar and satellite data as a classification method rules out the possibility to use the radar data as an information source for validation. Various sources for information applicable for verification were explored. This included METAR, NWP information, lightning, and soundings. For various reasons given below the METAR appeared as the best verification data set. E.g. NWP can only state that the conditions are favourable for convection but it can not forecast the location where the convection actually will occur. Lightning information was shown not to contribute significantly to classification results in the previous studies, The (2006), Carbajal-Henken (2009). Additionally lightning is mostly related to deep mature convection. Sounding information similar to NWP information, informs over the favourable conditions, and over the possible vertical extent of the convection but not over the actual position of the Cb. The actual position is important for aviation warnings.

#### *METAR*

For 2005 the METAR of four airfields were available. Amsterdam, EHAM, Rotterdam EHRD, Groningen EHGG, and Maastricht-Aachen EHBK, Unfortunately the night shifts at airport EHGG and EHBK were already automated and therefore no METAR was available from 23:00 till 07:00 GMT.

A study of years before 2005, to overcome this data gap, was disregarded as there was no complete SEVIRI data set available. In August 2007 all the METAR of EHGG and EHBK were replaced by automatic observations.

For a more extended evaluation of EHGG and EHBK between autometar and METAR only 2006 and a part of 2007 is additionally available, with the limitation that there are no night time METARs.

Vicinity is not uni-vocally interpreted by observers. It can range from a circle of 15 km radius around the airport to a range where both cloud top and cloud base can be observed. The latter is reported when the observed cloud is moving towards the aerodrome area. The distance to the observer varies with the height of the observer, and the height of cloud base.

#### *Radar.*

In early 2008 the resolution on which the radar data becomes available has been increased. The former radar signals were distributed on 2.5 x 2.5 km<sup>2</sup> grid. In early 2008 1x1 km<sup>2</sup> gridded data became available. The radar signals given in reflections were easily obtainable. Other radar observations like the echo top height were not readily available. In this study only the readily available data, the radar reflections were considered. For the operational adaption of the algorithm to the 1x1 km<sup>2</sup> gridded information the coefficients may need to be re-evaluated. But it is foreseen that the 2.5 x 2.5 km<sup>2</sup> gridded

products remain available in the future.

### *Satellite*

The SEVIRI observations are operationally available since January 2004 and are properly archived at KNMI since August 2004 onwards. As the satellite pixels are larger than the radar pixel size it was decided to project the satellite information on the radar grid using nearest neighbour method to facilitate comparisons and calculations. In this version no direct comparison is made between radar and satellite pixel values. Should this occur in the future than the slanted view of the satellite has to be accounted for.

## **4.2 Dependent and independent data**

A first evaluation of the data indicated that there are differences between summer and wintertime. Obviously there is also a difference for the satellite data between day and night, as during daytime the reflection channels and the HRV channel are available. It was therefore decided to subdivide the data set in four groups: winter day, winter night, summer day, summer night. Summer is defined as the months April till October, winter from October till April. Night is defined as those time slots where the maximum HRV reflection value within the studied area is less than 4% (summer) and 6% (winter). Day is defined as the remaining time slots. The difference in day-night threshold between the summer and winter night is introduced as the results for night time conditions showed a correlation between the HRV channel and the predictand of Cb occurrence when higher values of maximum HRV were used as day-night discriminator. Apparently the relation between HRV channel and Cb occurrence is so strong that even in weak twilight conditions they correlate.

All the data for the available time slots are distributed over three nearly equal parts in the following method. At the start the first three days with Cb/Tcu occurrence are distributed over the three parts. The next Cb/Tcu occurrences of one day are put into that part containing the lowest number of Cb/Tcu reports. This procedure is iterated until all Cb/Tcu occurrences are contained in three parts. All time slots per day containing non Cb/Tcu reports are evenly distributed added to the three parts. With this distribution method it is aimed to avoid dependencies which may occur when all time slots are randomly distributed over the three parts. This could result in a distribution of a day with many Cb reports over all three parts, which would introduce an undesired dependency between the three parts.

Although the distribution ensures the splitting of Cb situation during daytime, it may fail in night time conditions, where a Cb case may last long enough to pass the date denominator. This case will then be split over two parts. This should be considered when interpreting the results.

The chosen distribution can be redone, ensuring that the same data will end up in the same part. This would not occur with a random distribution.

In the study two of the obtained three parts will serve as the dependent data set while the remaining part serves as the independent part. Cycling between the three parts enables an assessment on the data to evaluate if there are other dependencies.

Given the data available and the distribution chosen a total number of 36 data set parts need to be evaluated summarised in table.4.1. Note that the datasets for EHGG and EHBK are not complete. EHGG and EHBK lack METAR information from 23:00 till 7:00 GMT. In the summer night this leaves hardly any data to perform a statistically analysis. Due to its climatology EHBK has a too low Cb number occurrence in winter time for a statistical interpretation. So wintertime will not be evaluated for EHBK with the 2005 data set.

	Winter day	Winter night	Summer day	Summer night
EHAM	3	3	3	3
EHRD	3	3	3	3
EHGG *	3	3	3	
EHBK *			3	

*Table 4.1 The number of dataset parts with a significant amount of Cb occurrence for an evaluation.  
\* Note that for EHBK and EHGG the night time METAR is missing from 23:00 till 7:00. Some 65,000 METAR observations distributed over a total number of 36 data sets are evaluated.*

### 4.3 Predictor selection

For predictor selection the **Statistical Package for the Social Sciences** SPSS package is applied. All 36 data sets are used in the package.

A forward stepwise regression selection method is applied. Starting with a constant-only model at each step a predictor is selected with the largest statistical score (likelihood ratio based) and a significance less than 0.05. The selection and inclusion is stopped when the significance of the remaining predictors is more than 0.05. Should during the inclusion a predictor obtain a significance of more than 0.10 then this predictor is excluded from the further steps of the evaluation. Forward stepwise regression selects the predictors purely on statistical criteria. The regression is capable to identify groups of predictors which individually contribute only weakly to moderately to the explanatory power but as a group contribute significantly.

It is unlikely that a unique set of predictors will be found describing all the occurrences in a perfect model. As the method does no physical interpretation the predictors should be scrutinized for their physical relation to the predictand. This could lead to the removal of predictors which have a high statistical correlation with the predictand but lack a physical explanation.

42 Different predictors were found to contribute to the Cb detection. Given that there could have been 180 predictors (5 x 36), one can conclude that there is a big overlap in the selected predictors. There were differences between seasons, day versus night, and stations. Frequently these different predictors have a similar information content, e.g. the range of the HRV value was selected for EHAM, and EHRD, where the maximum and minimum HRV value appeared for EHBK and EHGG for the summer day season. As the minimum HRV value always had a negative coefficient, the information content of the combined HRV maximum minus the HRV minimum is similar to the HRV range predictor. The hypothesis was that the combination of HRV maximum and the HRV minimum can be applied at all stations and can replace the HRV range as a predictor.

In other cases a single contour value of satellite or radar was selected as predictor. The predictors summarising the contour information can capture the single value information. The predictors summarising the contour information were expected to wrap up the information of a number of the single value contours. Therefore the contour summarising predictors were applied there were a single value contour appeared as a predictor.

By careful examination of the set of predictors the number could be reduced. To facilitate the interpretation and communication over the predictors for the different stations it is expected that a high degree of uniformity is beneficial both for the development and for the end-use. It facilitates the interpretation by the end-user.

Where it was acceptable the remaining set of predictors were reduced to comply with uniformity.

For the summer day uniformity was achieved. In table 4.2 the chosen predictors are summarised for each station and category. For the summer night the lack of sufficient data enabled uniformity only for EHAM and EHRD. In wintertime the EHGG predictors differ slightly from the EHAM and EHRD predictors. This may be due to a difference in climate or lack of sufficient night time data, please note the METAR from EHGG does not cover the full night.

It is remarkable that:

- the contours summarising the radar reflections are frequently selected as first by SPSS, although not apparent from table 4.2.
- in the winter night the weighted summation of contours is selected, where in the other categories the maximum radar contour is selected.
- for all cases the radar reflection within the 30 km radius area is selected as predictor.
- for daytime in winter and summer the difference between HRV maximum and HRV minimum (the HRV minimum always has a negative coefficient) is a selected predictor
- the average of the brightness temperature contributes significantly in the summer night.
- the standard deviation of the brightness temperature contributes significantly in the winter.

	Winter day	Winter night	Summer day	Summer night
EHAM	a,b,c,d,f	b,f,g	a,b,c,d	a,b,e
EHRD	a,b,c,d,f	b,f,g	a,b,c,d	a,b,e
EHGG *	b,c,d,f	b,f,g	a,b,c,d	
EHBK *			a,b,c,d	

*Table 4.2 Used predictors for each category with a) the maximum radar contour within the 15 km radius, b) the maximum radar contour within the 30 km radius, c) the minimum value of HRV within 15 km, d) the maximum value of HRV within 15 km radius e) the averaged brightness temperature within the cloud inside the 15 km radius, f) the standard deviation of the brightness temperature within the cloud inside the 15 km radius, g) the weighted sum of radar contours, which is related to the maximum radar contour within the 15 km radius. \*For EHGG and EHBK there was not sufficient data to make a statistical analysis for all the cases.*

To elucidate the relationship, linear or otherwise, between the predictors and the predictand a subsequent study is performed. The variable range of each predictor is subdivided. This subdivision should be done carefully. A simple subdivision in four quartiles each containing 25 percent of the data was not possible. Due to the highly uneven distribution of the data bins contained more than 25 percent of the data. Adding to the complexity of subdivision is that the data of some variables are affected by non Cb related influences. E.g. the high values of HRV maximum can be affected by the correction for the solar zenith angle, which can introduce artefacts at high solar zenith angles, occurring in the twilight period. The standard deviation of the brightness temperature within a cloud can be affected by cloud edges. Cloud edges may lead to high standard deviations, whereas high deviations may also be related to Cb occurrence. The averaged brightness temperature of clouds may be affected by surface temperatures, again as the cloud edge is a fuzzy defined entity. So too high values of HRV, of averaged cloud brightness temperature, and of standard deviation of brightness temperature should be excluded from the analysis on the relationship between predictors and predictands.

In Figures 4.1 to 4.5 the regression coefficients of the most frequently used predictors are given as function of their binned values.

The HRV range coefficients are shown in figure 4.1 as a function of the values of HRV bins. The value of the coefficient first bin is set as a reference point equal to zero. The

coefficient for the summer case shows a linear behaviour with increasing HRV range value. This is in accordance with the results of Carbajal-Henken et al (2009). For the winter case the linearity is apparent after the second bin. The significance of the coefficient of the second HRV bin is 0.055. This could indicate a limited applicability of the HRV range as a predictor to values lower than 70 in the winter day time.

The coefficients based on the radar signals, in figure 4.2 and 4.3 both from the 15 and 30 km collocation area show an increase with increasing dBZ. The variation in behaviour in winter time is relatively small. In summer there is a steep increase in the coefficients for the 30 km radius collocation area when radar signal is over 24 dBZ.

The coefficients for the averaged cloud top temperature have a different behaviour between summer and winter, figure 4.4. In summer both night and day time coefficients show a decrease with increasing cloud top temperatures, please note the first bin with the to zero set coefficient is at the right side of the figure, at 270 K. Whereas in the winter there is an increase with increasing temperatures. However the significance of the winter coefficients is too high. The summer coefficient behaviour is in agreement with the results of Carbajal-Henken et al (2009). It corresponds to an increase of Cb occurrence with decreasing cloud top temperature. The behaviour of the cloud top temperature in the winter is possibly related to a different type of Cb occurrences in winter in comparison to the summer. In winter the Cb do not have a high cloud top height and therefore relatively high cloud top temperatures. The cloud top temperature is not considered as a reliable predictor in the wintertime. It is only used in the summer night category.

The standard deviation from the cloud top temperature in figure 4.5 the only one with non linear distributed bins. High standard deviations occur relatively seldom, so to come to equally distributed population over all bins the highest bin had to be larger in comparison to the other bins. Most likely the high standard deviations are related to cloud edges and not to Cb occurrence. This limits the applicability of this predictor to moderate standard deviation values.

All the discussed predictors have a clear relationship to the predictand. Apart from the exclusion of the cloud top temperature as a predictor in the wintertime no other choices were made for the relationship between predictors and predictand.

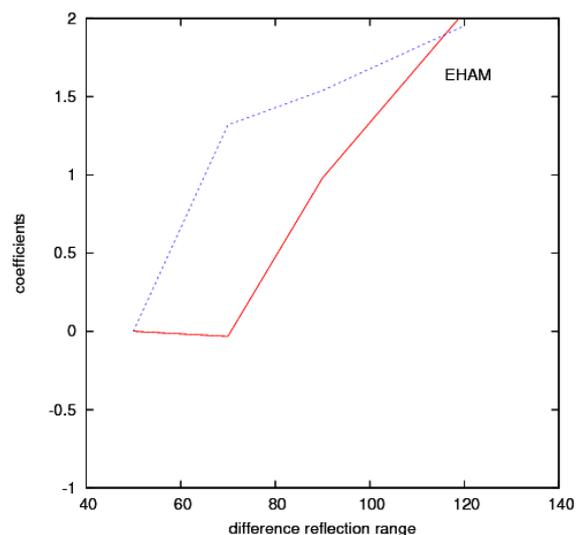


Figure 4.1. Check on behaviour of the predictors with respect to the regression coefficients of the defined bins for HRV difference range for EHAM. Summer day in dashed blue and winter in red line.

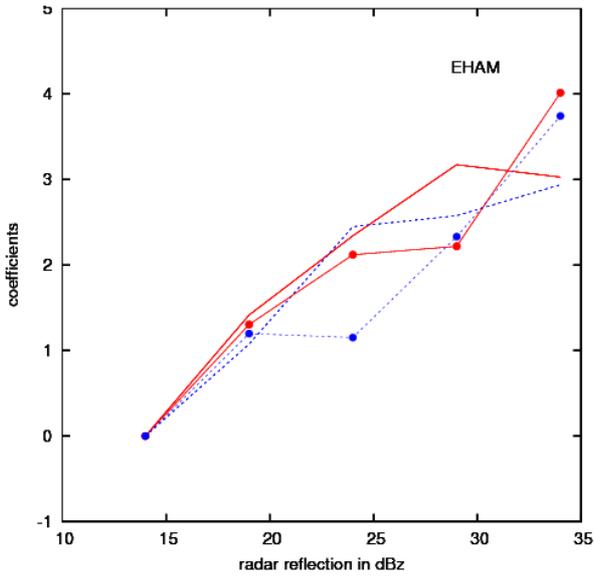


Figure 4.2 As Figure 4.1 but for radar dBZ for EHAM 15 km collocation area. Winter night in dashed blue and winter day in red line, summer day in dashed blue with bullets and summer night in red line with bullets.

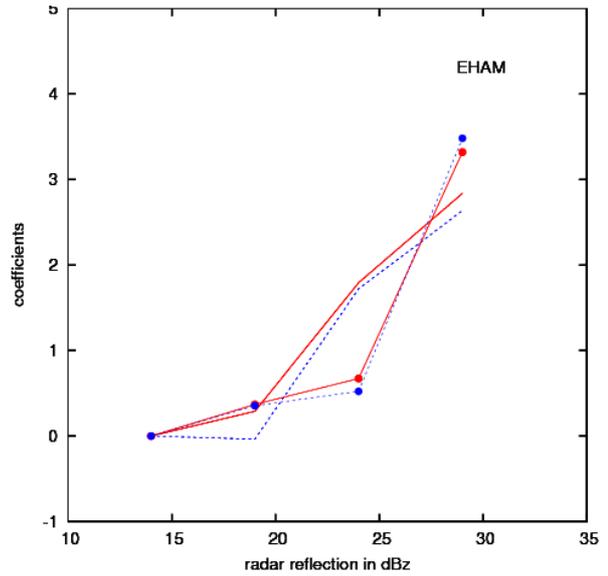


Figure 4.3 As Figure 4.1 but for radar dBZ for EHAM 30 km collocation area. Winter night in dashed blue and winter day in red line, summer day in dashed blue with bullets and summer night in red line with bullets.

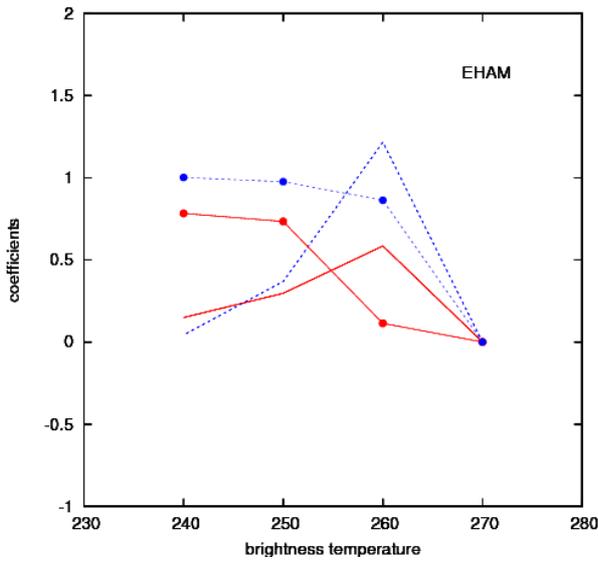


Figure 4.4 As Figure 4.2 but for Cloud top brightness temperature EHAM 15 km collocation area. Note that the bins are from high to low temperatures. The first bin is at 270 K.

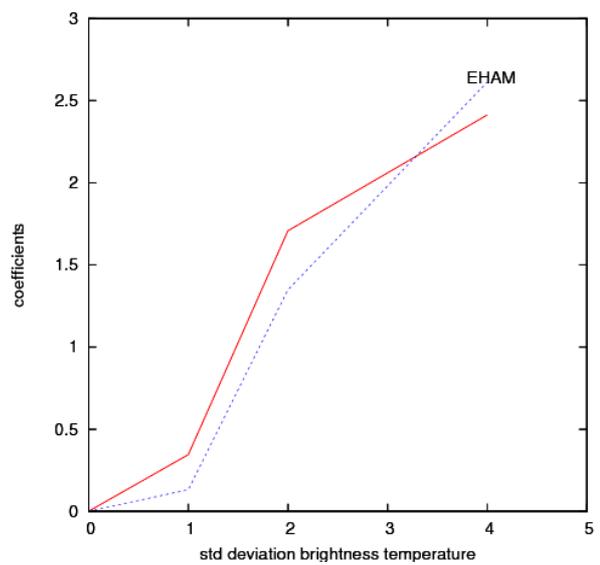


Figure 4.5 As Figure 4.1 but for Cloud top brightness temperature standard deviation for EHAM 15 km collocation area. Winter night in dashed blue and winter day in red line.

#### 4.4 Conclusions from predictor selection

The maximum radar contour or weighted sum of contours is very frequently selected by logistic regression indicating that these are significant predictors for Cb occurrence. The maximum occurring contour value will vary with atmospheric conditions. A Cb occurrence can therefore not be linked to a fixed threshold in radar reflectivity observation as is done in the present operational algorithm-2007.

This result corresponds to the frequently reported experience of users of failure of the operational algorithm-2007 to detect Cb. The user recognises a pattern of Cb occurrence in the radar image which could be missed by the autometer, because the threshold value was not reached. The user will only focus on the pattern and not on the maximum occurring value. Therefore the user will recognize the Cb occurrence despite the fact that the threshold value is not reached and he will conclude that the operational algorithm-2007 results are poor.

The appearance of 30 km based predictors from the radar observations within the selection can be an indicator that the METAR includes information outside the 15 km target area. An algorithm neglecting the signals outside the 15 km radius collocation area will never be able to account for all the METAR reports and hence will always have a poorer performance when compared to the METAR.

The occurrence of HRV difference range as indicator is presumably linked to illumination of convective clouds with high reflective sides and tops and sharp shadows. Especially in the winter with a lower solar elevation angle the difference range will be more apparent.

The averaged and standard deviation of the brightness temperature become significant in the night as no reflection information is available. In the summer the Cb tops can reach high altitudes resulting in low cloud top temperatures. In the winter the relation between cloud top temperature and Cb occurrence is less clear in order to distinguish Cb. Within a cloud a large variation in cloud top temperature in winter and a low average cloud top temperature in the summer may be indicative for Cb, but it depends very much on the scale of the cloud relative to the area under study. The cloud top temperature variation has less explanatory power compared to the HRV range and cloud top temperature. But it becomes relevant when the signal of the latter two is weak or non existing.

## 5 Detection results and evaluation

The selected predictors, described in the previous section, and their coefficients form the basis to a probability of Cb occurrence.

The results are discussed and summarised. This chapter describes the statistical results and shows a selection of figures. More results and figures are given in the appendix 3.

### 5. 1 Results

The contribution of each predictor can be assessed by performing logistic regression in a forward stepwise selection method and study the impact on  $NR^2$ . Creating nested models shows that adding predictors will lead to an increase in the explained variance reflected in the  $NR^2$  score (Nagelkerke, 1990).

In tables 5.1 two examples are given of the development of  $NR^2$  scores.

In a number of categories the radar contour predictor causes the largest increase in the  $NR^2$  score. The subsequent predictor causing the second largest reduction in the forward stepwise selection is different for each station and category. It can even change if a different part of the data is chosen to be the independent data. For example for EHAM in the summer day category and EHRD in the winter night both show different orders for predictors. The increase of  $NR^2$  with the addition of predictors is clearly visible.

$NR^2$ (1)	$NR^2$ (2)	$NR^2$ (3)	$NR^2$ (1)	$NR^2$ (2)	$NR^2$ (3)
0.470 [a]	0.472 [a]	0.460 [a]	0.293[d]	0.458 [a]	0.433 [a]
0.519 [b]	0.496 [b]	0.490 [d]	0.397[a]	0.505 [e]	0.486 [e]
0.561 [c]	0.527 [d]	0.514 [c]	0.430[e]	0.512 [d]	0.493 [d]
0.584 [d]	0.552 [c]	0.550 [b]			

*Table 5.1 Example the increase of  $NR^2$  with increasing number of included predictors for left: EHAM summer day and right : EHRD winter night. The number in brackets in the top line indicate the data set used as independent data set. The variation visible is due to cycling of the independent part between the three data parts where the two remaining data sets serve as dependent data set. When the cycling causes a change in the selection order of the predictors this is reflected in the letter order in the square brackets, which denote the used predictors a: Contour radar 30 km radius b: HRV minimum, c: HRV maximum d: Contour radar 15 km radius, e: Standard deviation T.*

For a comparison between the various categories the final results obtained are shown in the table 5.2 here below. The cycling between the three datasets has been applied, leading to three numbers for each category.

The scores are summarised in a table given the  $BSS$  with the sample climatology as reference as well as the Nagelkerke  $NR^2$  score. An attempt to determine the  $BSS$  with the persistence as reference was less successful as persistence of the observation 30 minutes earlier is a predictor with a high performance.

From the table one can conclude that for the summer night cases the performance expressed in  $BSS$  is lower than the summer day cases.

EHGG has a small  $BSS$  value in the winter night indicating that the climatological performance is only slightly worse, but this dataset is not complete as it partly lacks METARS during the night.

The variation occurring due to the cycling of data parts as independent data is probably

caused by differences in Cb occurrence within the three parts.

At all aerodromes the summer day time scores are good. The summer night has on average the lowest BSS values, when excluding the EHGG WN from the comparison.

$NR^2$ SD	$NR^2$ (1)	$NR^2$ (2)	$NR^2$ (3)	BSS(1)	BSS(2)	BSS(3)
EHAM	0.58	0.55	0.55	0.36	0.44	0.43
EHRD	0.56	0.46	0.53	0.22	0.51	0.34
EHGG	0.55	0.56	0.58	0.44	0.44	0.43
EHBK	0.54	0.52	0.54	0.36	0.45	0.38

SN	$NR^2$ (1)	$NR^2$ (2)	$NR^2$ (3)	BSS(1)	BSS(2)	BSS(3)
EHAM	0.41	0.48	0.49	0.41	0.29	0.21
EHRD	0.39	0.47	0.53	0.42	0.29	0.14

WD	$NR^2$ (1)	$NR^2$ (2)	$NR^2$ (3)	BSS(1)	BSS(2)	BSS(3)
EHAM	0.65	0.64	0.61	0.45	0.49	0.54
EHRD	0.43	0.51	0.49	0.59	0.25	0.56
EHGG	0.43	0.42	0.33	0.37	0.17	0.30

WN	$NR^2$ (1)	$NR^2$ (2)	$NR^2$ (3)	BSS(1)	BSS(2)	BSS(3)
EHAM	0.51	0.51	0.51	0.32	0.33	0.34
EHRD	0.43	0.51	0.49	0.40	0.23	0.23
EHGG*	0.41	0.38	0.31	0.22	0.04	0.27

*Table 5.2 Summarising all the results of  $NR^2$  and BSS for all the categories and airports, summer day (SD) summer night (SN) winter day (WD) winter night (WN) for the three possible independent datasets, indicated by 1,2,3 in the  $NR^2$  columns.*

*\* EHGG WN is not a complete dataset as part of the night METAR is lacking.*

Please note that there are no EHBK results for winter season. The difference in climatological conditions between EHBK and the other airports disables a meaningful application of the predictor coefficients at EHBK.

The results are evaluated per category: summer, winter, day, night per aerodrome. For each case a set of three graphs are determined from the independent data: the *POD*, *FAR*, *CSI*, and *BIAS*, as function of the probability threshold, summarised in one figure, the histogram distribution, and the attributes diagram both as function of the predicted probability. A graph of one of the better and one of the worse obtained results is included in this section to illustrate the variation in the results.

From the *POD*, *FAR*, *BIAS* and *CSI* diagrams a threshold can be determined on which the classification can be based. The wide variety occurring within the graphs shows that one fixed value of threshold for all categories cannot be determined. For each season and day/night situation a threshold can be derived.

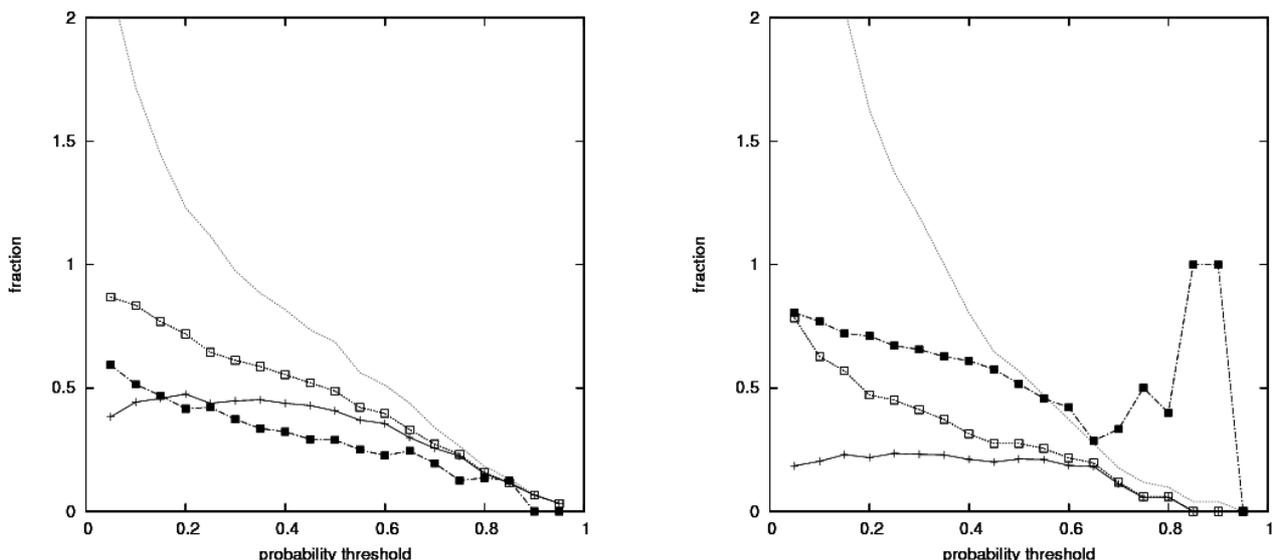


Figure 5.1 Scores for EHAM summer day (left) and EHRD summer night (right). POD (open squares), FAR (black squares), CSI (pluses) and BIAS (dashed lines) as a function of the probability threshold.

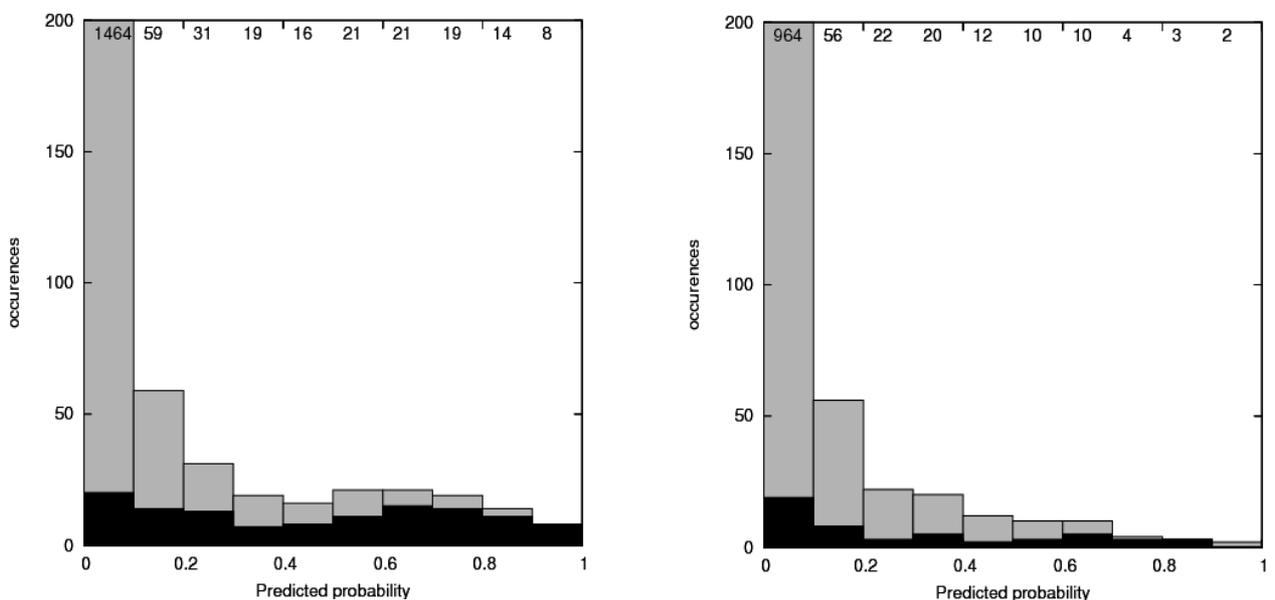


Figure 5.2 Frequency histogram of Cb distribution for probability threshold ten percent bins for EHAM summer day (left) and EHRD Summer night (right). In the top of the figure the number of cases in the bin are given. The light grey indicates the non-events, the black bars indicate the observed Cb occurrences. Note the large population of the first bin, out of the scale of the figure.

In figure 5.1 the POD, FAR, CSI and BIAS are given for a summer day at EHAM and for a summer night at EHRD. The detection during day light conditions is more successful than in the night. This is reflected in the relative high CSI scores of the shown EHAM case versus the EHRD case.

Where the CSI for the EHAM case peaks to 0.5 at a probability threshold of 0.2, the CSI at EHRD remains more or less constant from probability threshold 0.1 to 0.6. For EHRD the FAR score exceeds the POD.

The distribution given in Figure 5.2 shows the lower population for the EHRD case in comparison to EHAM in all the bins. Also the Cb occurrence is less at EHRD night case in

comparison to the day time EHAM case. At EHAM the higher value bins have a high percentage of correct Cb detections. At EHAM the highest occurrence ratio of Cb is in the last bin.

The attributes diagram compares the predicted probability to observed relative frequency. The predicted values are binned into 10 percent bins, the occurrence number is given in the diagram and also in the histogram. The no-resolution line relates to the climatology, in this study the number of Cb occurrences compared to all reported METARS in the category. The no-skill line is halfway the no-resolution line and the perfect reliability line, which is represented by the diagonal.

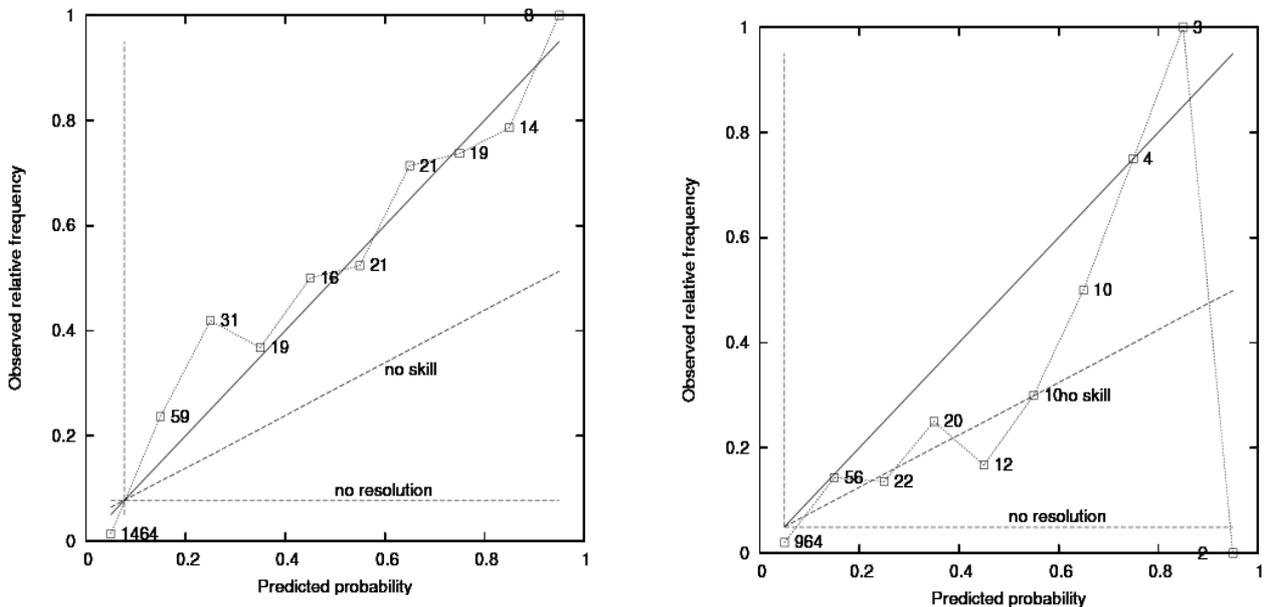


Figure 5.3 The attributes diagrams for EHAM summer day (left) and EHRD summer night (right) as a function of the predicted probability. The numbers in the figure indicate the number of cases per 10 percent bin. The no resolution line relates to the climatological Cb occurrence, different for each station and season. The perfect reliability line is the diagonal. The no skill line is halfway the diagonal and the no resolution line.

In Figure 5.3 for EHAM a significant number of the results contribute to the skill of the model. Points to the left of the perfect reliability line indicate a too low predicted probability in comparison to the observed relative frequency. And vice versa for the points to the right of the diagonal. For EHRD a significant number of points have a large distance to the perfect reliability line, and are closer to the no-resolution line. These points contribute marginally to the skill of the model.

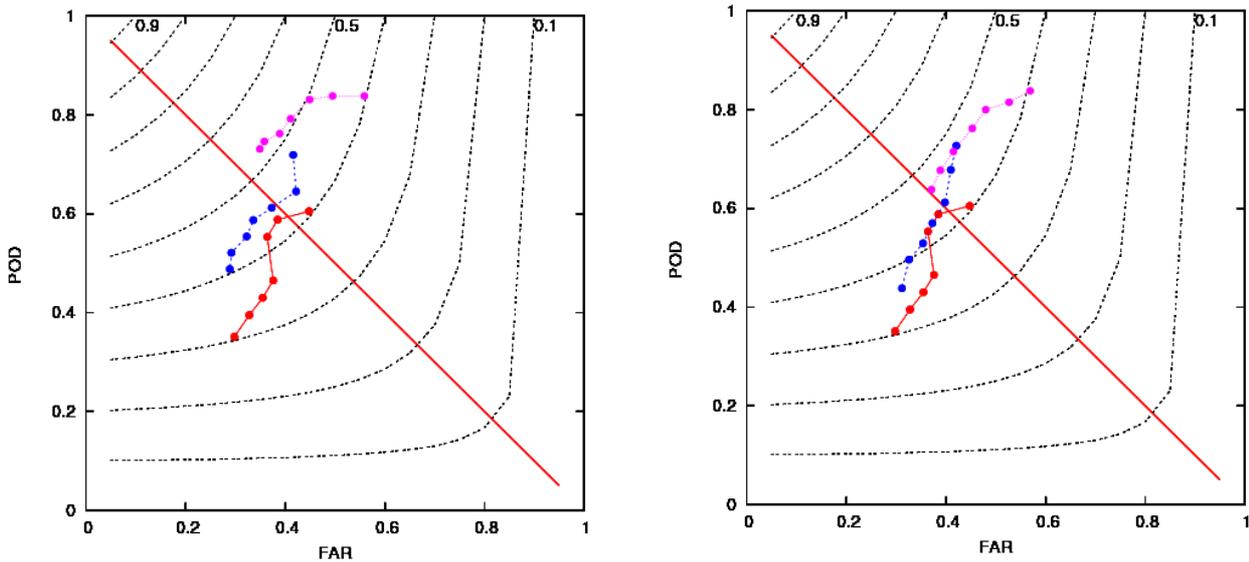
In the appendix 3 all attributes diagrams and histograms of 36 categories are given.

The attributes diagrams show that the majority of the results show a good performance. The summer night performances at EHAM and EHRD are relatively poor in comparison to other performances in the other categories. A significant number of cases of the EHRD show a relative poor performance with results close to the no-skill line. Also in winter time for EHGG there are number of cases with limited skill. For the night time at EHGG the dataset, however, is limited, as METARs are lacking. The lack of sufficient data in the night causes a spiky behaviour in the attributes diagram as is also visible in the figure above for EHRD.

Most summer day cases resemble the given summer case, only one of the EHRD cases has a score close to the no-skill line, see appendix 3.

Night cases have a relative poorer performance in comparison to the day cases. The example of EHRD given here above is one of the worse cases.

Another way of presenting the information of POD, FAR, CSI and BIAS given in Figure 5.1 is given in Figure 5.4. Here the CSI values, obtained for the probability thresholds ranging from 0.2 to 0.5, are shown as a function of FAR and POD. The chosen representation facilitates the interpretation, but the results have to be related to the attribute diagrams described above and given in appendix 3 to come to balanced conclusions.



*Figure 5.4 The lines with bullets indicate the variation in performance due to the variation of the probability thresholds of the three cycling independent data sets for EHAM summer day. Left part: applied predictors are made uniform for all three sets. Right part: the first five selected predictors by the forward method for each set. Note that there will be different predictors used for each coloured line in the right side figure. The red bullet line is by coincidence in both figures based on the same predictors.*

*The performance curves are given for probability threshold values ranging from 0.2, upper right, to 0.5, lower left, in steps of 0.05 as a function of FAR and POD. Note that an increase of probability threshold will result in a lower POD and a lower FAR. Isolines of CSI in dotted black varying in steps of 0.1 from 0.9, the values of 0.1, 0.5, and 0.9 are indicated in the top of the figure. The red line denotes the BIAS is 1, right to this line are higher values of bias, left lower values.*

The chosen presentation allows for an evaluation of the impact of the choice of uniform predictors in comparison to the predictors selected by the forward stepwise regression method. Three examples are shown in figures 5.4 and 5.5 for summer day at EHAM and for summer night and winter day at EHRD.

The predictors chosen to accomplish uniformity might cover a broader spectrum of possible Cb/Tcu occurrences, because the choice is based on a good performance at different locations, therefore capturing more different Cb occurrences.

The uniformity choice impact on the scores for the summer day at EHAM are marginal in terms of variation in POD and FAR. The uniformity choice impact for the summer night at EHRD are relatively more significant. The difference in performance can partly be attributed to the difference in applied predictors.

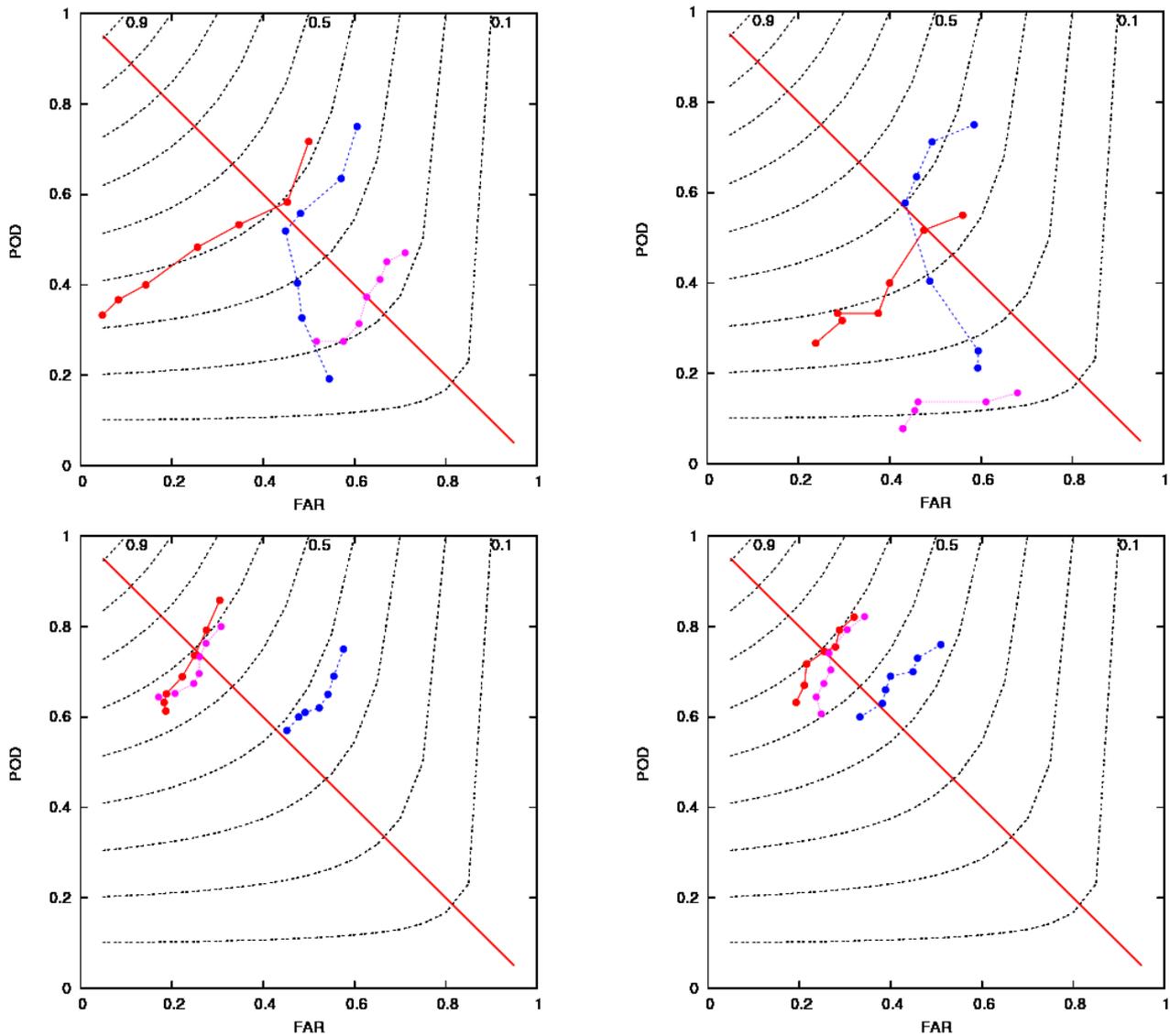


Figure 5.5 As figure 5.4 but for the summer night (upper row) and winter day (lower row) at EHRD with uniform applied predictors (left), and for the first five selected predictors (right). Note that there might be different predictors used for each coloured line in the right part of the figure.

Occasionally the selection directly derived from the forward method is better, e.g. for EHRD winter day. But in the majority of the cases the uniformly applied predictors lead to a better performance, i.e. in the figures closer to the upper left corner where POD equals 1 and FAR equals 0 and lesser variation between the results of the three datasets, than the performance of the first five selected predictors by the forward stepwise regression method. This justifies partly the approach of selecting uniform predictors. An additional benefit is that the approach facilitates the interpretation and communication both by developers and users at the different airports.

## 5.2 Summary of results

The final results are summarised in Figure 5.6 to 5.9 for the uniform applied predictors. Here the difference between the performance curves of three datasets is relatively larger for the EHRD and EHGG categories in comparison to the EHAM category.

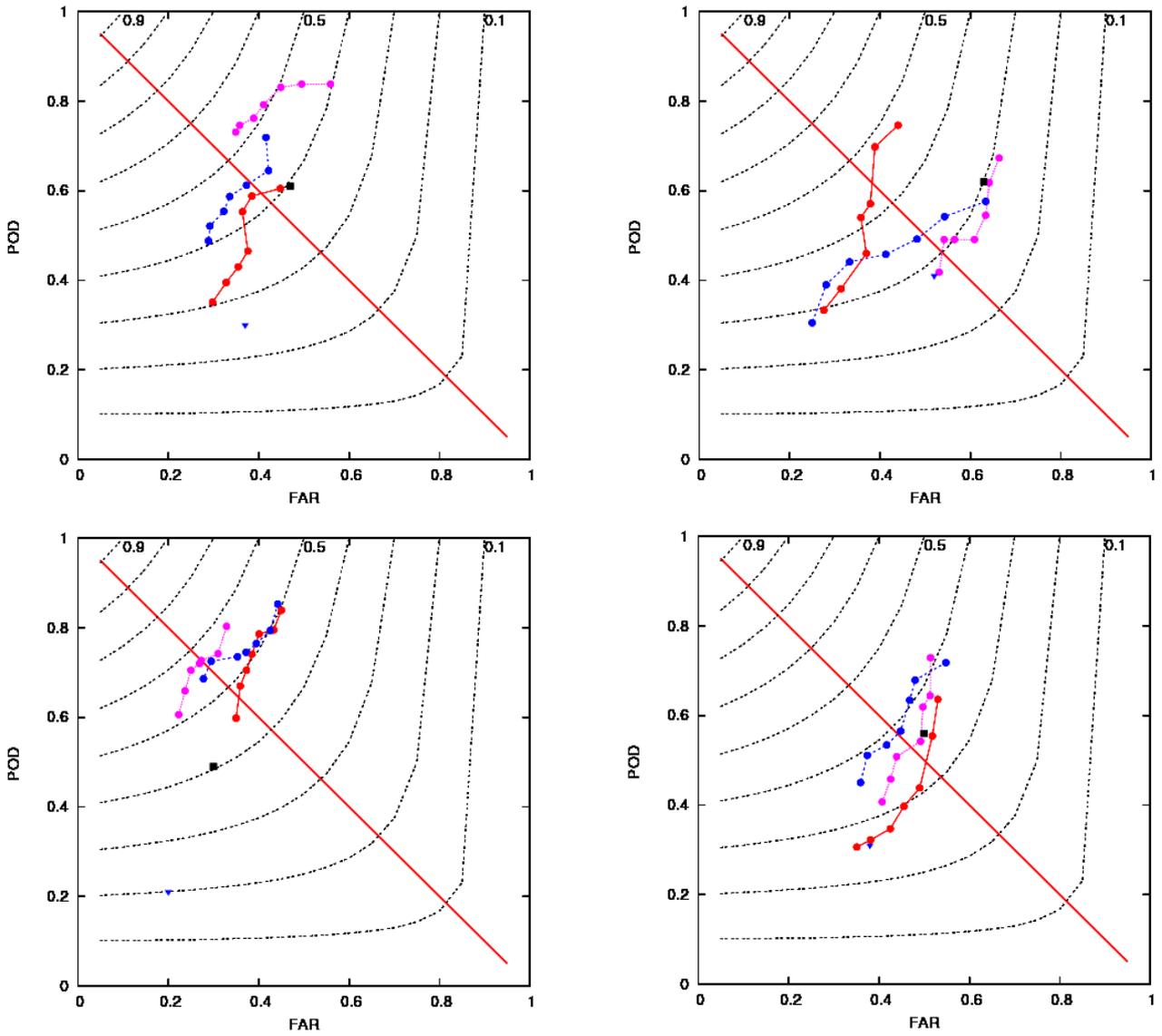


Figure 5.6 As Figure 5.4 for all cases at EHAM summer day (upper left), summer night (upper right), winter day (lower left) and winter night (lower right). The black square represents the autometar score for Cb-Tcu for a 30 km radius of the collocation area, the triangle gives the autometar result for a 15 km radius of the collocation area.

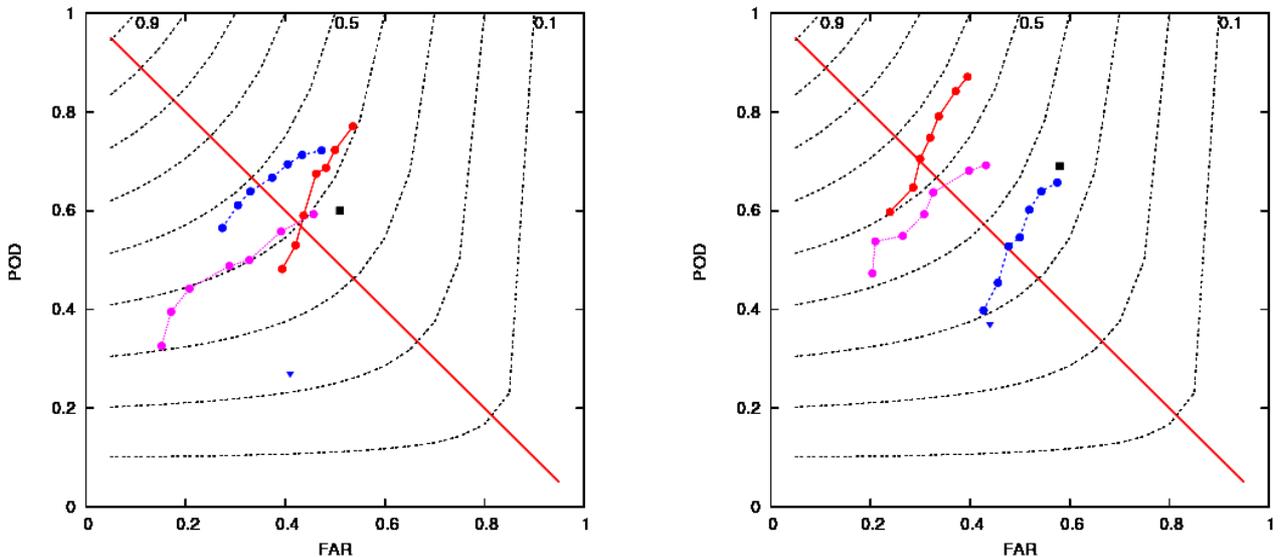


Figure 5.7 as Figure 5.6 for EHBK (left) and EHG (right) summer day.

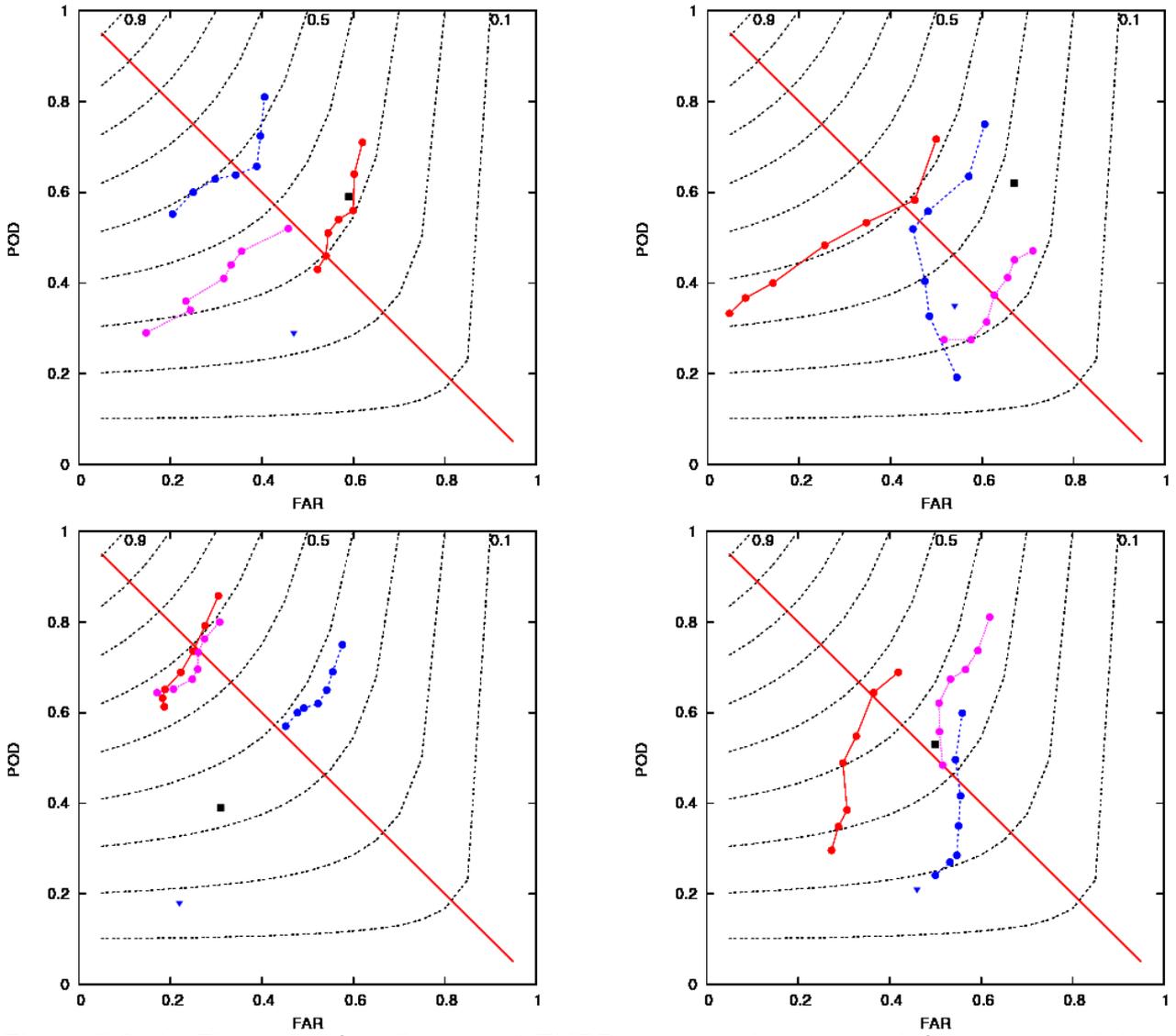


Figure 5.8 As Figure 5.4 for all cases at EHRD summer day (upper left), summer night (upper right), winter day (lower left) and winter night (lower right).

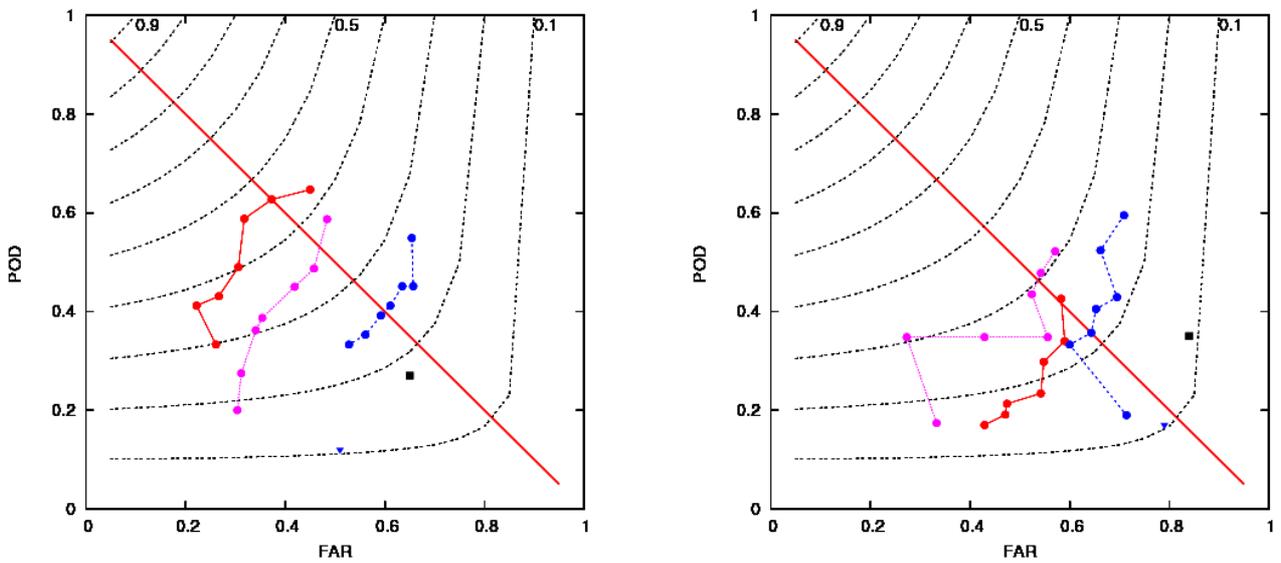


Figure 5.9 As Figure 5.6 for all cases at EHG winter day (left) and winter night (right).

In winter the results show less variation in the FAR dimension compared to the summer cases. The results show the smallest variation in the POD dimension in the winter day category. The variation in the results may vary with a different distribution of the Cb over the three subsets. Also a variation in METAR reports could explain the difference in variability at the different airports.

Also for the complete datasets, consisting of the three parts together, the coefficients are determined. These coefficients are used for the developed operational algorithm. The curves based on these coefficients are given in appendix 4.

The majority of the obtained results show a (much) better performance both in POD and FAR compared to the results of the present operational algorithm-2007 denoted by the black squares in the figures. Depending on the probability threshold some performance curves of the developed algorithm show lower POD and higher FAR values in comparison the performance of the algorithm-2007. But the developed algorithm results have far better CSI values compared to the Cb/Tcu results of the operational algorithm-2007. Depending on the choice of the probability threshold there are cases which have a CSI in the order of the 0.60 for the developed algorithm. Next to a high CSI a BIAS of close to 1 is preferable in the results.

The presented evaluation can not be considered as complete. What is excluded from the present evaluation is the cases of embedded Cbs which will not always be included in the METAR. Here both the operational and the developed algorithms may detect correctly Cbs but this can not be assessed on the used METAR dataset. The METAR has been used as a reference set. It is based on human observations, so mistakes remain possible. There is only a modest exchange of personnel between the various aerodrome locations. The observers are all trained in a similar way. Still it may occur that subtle differences in METAR reports can occur between the various locations. A new shift will certainly be aware of the previous METARS and may take them into account. The new shift will be less interested to what other locations report. These subtle differences will have an impact on METARS of the various airports.

## **6 Conclusions and future.**

The chapter summarises the conclusions and gives an outlook on future research.

### **6.1 Conclusions.**

Since 1-8-2007 an operational algorithm-2007 is implemented at the airports EHBK and EHGG to detect Cb and Tcu. It uses the radar reflection observations and lightning observations as input. The detection of Cb Tcu is relevant for aviation and therefore a requirement by ICAO. The performance of the algorithm-2007 is evaluated and considered as poor in terms of POD and FAR. This study was initiated to develop an improved algorithm.

An automated Cb-Tcu detection algorithm based on the synergy between radar and satellite observations is developed. The algorithm uses logistic regression to determine the probability of Cb-Tcu occurrence. Within logistic regression a forward stepwise approach is applied. The predictors selected by the forward stepwise regression method are related to the highest radar contour occurring in the 15 and 30 km radii collocation area, and to the satellite observations, reflection range of the high resolution visible channel, the averaged cloud temperature and its standard deviation. The latter three all in the 15 km radius collocation area.

The obtained results show in general an improvement in performance of the developed algorithm in comparison to the operational algorithm results.

The performance of the developed algorithm is dependent on season and day-night conditions. The best performance is achieved in the Summer day category followed by the winter day category, with the summer defined from April till October. Surprisingly the summer night category shows the worst performance, not significantly better than the operational algorithm-2007, see appendix 4.

Although the algorithm is developed for EHBK and EHGG no year round evaluation was possible for those airports because of the lack of sufficient Cb occurrences, required for a statistical analysis. Especially for the EHBK airport data was lacking. This hampers an operational application for EHBK.

Note that since there is no other observation which covers both the required spatial and time dimensions a future assessment of the performance of the algorithm is not possible. The METARs are the most reliable and continuous source of Cb and Tcu observation, but they are terminated at EHGG and EHBK. However at EHAM and EHRD they are still continued.

Based on the results of this study an improved operational algorithm can be defined. The probability threshold selection will determine the performance of the algorithm. For the daytime categories a POD of 65 % and a FAR of 35 % appears feasible in the summer and winter day categories. For the night time a POD of 55 % and a FAR of 45 % appears achievable.

In appendix 1 some recommendations are given towards an operational implementation.

### **6.2 Future**

Although the POD and FAR are improved to maximum values of 65 percent and 35 percent, respectively, they still do not comply with the values of 80 % POD and 20 % FAR mentioned in the NOTA from HWA of July 2006, included in appendix 6 (in Dutch). Due to a lack of time, a number of improvements could not be explored in detail to study

their impact upon the results. It is recommended to consider them in a study.

During the study it became clear that the evaluation area with a radius of 30 km on June 30, 2008 is decreased to an area with a radius of 15 km. The impact of the decrease in area is given in Figures 5.6 to 5.9 denoted by the blue triangles. It leads to a lower FAR, but also to a more significant loss in POD when compared to the METAR. It is recommended to evaluate the impact of this radius change on a dataset based on the same area.

In January 2008 the radar spatial resolution has been improved from 2.5x2.5 to 1x1 km. The operational algorithm-2007 requires at least three spatially connected pixels to come to a classification. This number should be reconsidered with the introduction of higher spatial resolution, as there is a factor of 6.25 in spatial resolution between the previous and present radar resolution. It will cause a better detection of smaller convective clouds but will also increase the noise. The spatial resolution improvement comes with a significant increase in clutter. The increase in clutter combined with the smaller radar pixels can cause an increase in false alarms. The operational algorithm-2007 and developed algorithm will be affected by clutter leading to a decrease in performance.

Suggestions to improve the autometar:

- possible predictors are not exhaustively explored: e.g. radar echo top height development, 3.9 and 1.6 microphysical satellite information. In individual summer months the predictor related to 3.9 micrometer channel was regularly selected by SPSS, it may be linked to intense summer convection, probably missed in the present predictor model set.
- collocation area scanned by the observer can show a variation: It may be different for night versus day, summer versus winter. Research should be done on this subject.
- Independent evaluation study of the data set on Cb occurrence and unobserved Cb occurrence to explore the caveats of both METAR and the algorithm.
- Explore the persistence of the probability as a predictor.
- Explore if there are reliable other sources for instability information: NWP, SAFIR, AMDAR, soundings
- The predictor selection appears to be applicable for all the airports. There is agreement for the summer day and winter day and night. So despite the lack of data predictors for EHBK and EHGG could be selected. As the necessity for an independent dataset is then no longer required it could be possible to determine coefficients despite limited data.
- The back-up MSG satellite, which will replace the operational one in case of maintenance or failure, is available for "rapid scan". It observes western Europe with an observation cycle of 5 minutes. The higher observation frequency makes it interesting for Cb detection. It is however not continuous. Two days a month it is certainly not available, next to the other moments when the spare satellite replaces the operational one.
- Air mass classification from the satellite: This gives an indication whether the air mass over the airport is potentially unstable or stable.
- SAFNWC (satellite application facility on Nowcasting, an EUMETSAT initiative) has products like Rapid Developing Thunderstorm (RDT) product, and a cloud mask. The cloud mask can be used to determine cloud top temperature and its standard deviation. The RDT product could be a significant indicator of Cb, but it only recognizes mature Cbs of several pixels. It does however give a forecast of the Cb occurrence up to two hours.
- Evaluate 2009 for EHAM and EHRD as an extra independent data set. Here the radar has a higher spatial resolution. Both the high and the low resolution radar are available for a comparison study.

## Literature.

Carbajal Henken, C., Schmeits, M., Wolters, E. and Roebeling, R. Detection of Cb and Tcu clouds using MSG-SEVIRI cloud physical properties and weather radar observations. KNMI WR 2009-04

Duda DP, Minnis P (2009) Basic Diagnosis and Prediction of Persistent Contrail Occurrence Using High-Resolution Numerical Weather Analyses/Forecasts and Logistic Regression. Part II: Evaluation of Sample Models. *Journal of Applied Meteorology and Climatology* 48(9): 1790.

Fujita, T., Watanabe, K. and Izawa, T. 1969, Formation and Structure of Equatorial Anticyclones Caused by Large-Scale Cross-Equatorial Flows Determined by ATS-I Photographs, *Journal of Applied Meteorology*, Volume 8, Issue 4 (August 1969) pp. 649–667

Harrison, D. L., 2001, Johnson, C. J. Harrison, D. L., 2001, Pankiewicz, G. S.; ; Improving radar observations of precipitation with a Meteosat neural network classifier *Meteorology and Atmospheric Physics*, Volume 76, Issue 1-2, pp. 9-22 (2001).

Holleman, I. Introduction to Radar internal webpage at KNMI.  
<http://info.knmi.nl/~holleman/index.html>

Holton, An introduction to meteorology, 1975, Academic Press.

David W. Hosmer, Stanley Lemeshow - 2000 ,Applied logistic regression, wiley

Kucharski, C., A. Pinault, en D. Tzanos, 2005: Utilisation de l'imagerie Radar et des données Foudre pour la detection d'activité convective (CB et TCU) pour compléter les METAR AUTO, Météo France, DSO/ENR/AMI-NI094

Levizzani V., Setvák M., 1996: Multispectral, High-Resolution Satellite Observations of Plumes on Top of Convective Storms. *J. Atmos. Sci.*, 53, 361-369.

Mecikalski, John R. and Bedka, Kristopher M.. Forecasting convective initiation by monitoring the evolution of moving cumulus in daytime GOES imagery. *Monthly Weather Review*, Volume 134, Issue 1, 2006, pp.49-78.

Mecicalski (2007) -based Convective Initiation Nowcasting System Improvements Expected from the MTG FCI Meteosat Third Generation Capability, EUM/CO/07/4600000405/JKG Technical Report  
[http://www.eumetsat.int/groups/pps/documents/document/pdf\\_mtg\\_rep35.pdf](http://www.eumetsat.int/groups/pps/documents/document/pdf_mtg_rep35.pdf)

Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, Vol. 78, No. 3: 691-692.

Petterssen S., 2008 Introduction to meteorology. 3rd edn. New York: McGraw-Hill.

Pruppacher, H.R. and Klet, J.D., 1980, Microphysics of clouds and precipitation, Reidel, Netherlands.

Sokol Z. and Pesice, P. 2009, Comparing nowcastings of three severe convective events by statistical and NWP models, Atmospheric Research Vol. 93. p397-407.

Schmeits, M.J., C.J. Kok, D.H.P. Vogelesang and R.M. van Westrhenen, *Probabilistic forecasts of (severe) thunderstorms for the purpose of issuing a weather alarm in the Netherlands* Weather and Forecasting, 2008, 23, 6, 1253-1267,

Setvák M., 1989: Convective storms - the AVHRR channel 3 cloud top reflectivity as a consequence of internal processes. Proc. 5th WMO Conference on Weather Modification and Applied Cloud Physics, Beijing, China, WMO/TD #269, 109-112.

The H. (2006), Verificatie AUTOMETAR CB/TCU- detectie (herziene versie) (in appendix 5)

A. S. Thom, D.H. McIntosh 1969, Essentials of Meteorology Wykeham series

Thoss, A. 2009, Algorithm inter comparison results of the Locarno Cloud Workshop, EUMETSAT Meteorological Satellites Conference, Bath, UK, 21-25 September 2009.

Wilks, D.S. , (1995) Statistical Method in the atmospheric sciences, academic press, New York

T. Zinner, H. Mannstein and A. Tafferner (2008) Cb-TRAM: Tracking and monitoring severe convection from onset over rapid development to mature phase using multi-channel Meteosat-8 SEVIRI data, Meteorology and Atmospheric Physics, Volume 101, Numbers 3-4 / October

SPSS [www.spss.com/nl/](http://www.spss.com/nl/)

SAFNWC 2000 <http://nwcsaf.inm.es/>

RDT,2000

[http://nwcsaf.inm.es/Scientific%20documentation/SAF-NWC-CDOP-MFT-SCI-VR-11\\_v1.3.pdf](http://nwcsaf.inm.es/Scientific%20documentation/SAF-NWC-CDOP-MFT-SCI-VR-11_v1.3.pdf)

[http://www.meteorologie.eu.org/RDT/doc/SAF-NWC-CDOP-MFT-SCI-ATBD-11\\_v1.3.pdf](http://www.meteorologie.eu.org/RDT/doc/SAF-NWC-CDOP-MFT-SCI-ATBD-11_v1.3.pdf)

## Recommendations towards application

The following steps have to be considered and/or implemented:

- 1.Applications: for the radar data notice the difference in resolution, the 2.5 km resolution radar product is still produced in 2010. This means that the coefficients determined in this study can be applied.
- 2.Satellite data need to be projected on the used radar resolution.
- 3.Predictors selection: Only 45 variables are proven to contribute positively to a Cb detection; 20 radar contours on 15 and 30 km; HRV maximum; HRV minimum; cloud average temperature; standard deviation of cloud top temperature, need to be calculated.
- 4.Concern question: What to do with the categories without coefficients for EHBK and EHGG? See above for the future plans.
- 5.Considerations for the newly developed Production chain: radar products, satellite products are still considered separately within the collocation area. So there is no problem with the pixel shift between radar and satellite observations. When in the future radar and satellite pixels are merged for evaluation this problem has to be solved. Project the satellite on radar resolution. Take into account the pixel shift (averaged two radar pixels (at  $2.5 \times 2.5 \text{ km}^2$  resolution) North South) due to satellite viewing geometry. Perform the required calculations, make the classification.
- 6.Please note that the present algorithms only provide a dichotomous classification of Cb occurrence. There is no unique classification of coverage. The predictors vary between day, night winter summer. Suggestion to come to a coverage value: the number of pixels enclosed by the highest radar contour or when it is exceeded the 33 dBZ contour.
- 7.The cloud base height of the Cb layer is difficult to assess. Suggestion to use the lowest layer from NWP derived radiosonde data which indicates an instability layer. Or use the thumb rule based on  $T_{\text{air}}$  and  $T_{\text{dewpoint}}$ , both at 2 meter.
- 8.Detected obscured Cb can occur in a covered situation. The developed autometar algorithm can accommodate this.

## Appendix 2.

<http://dcaa.slv.dk:8000/icaodocs/Annex%203%20-%20Meteorological%20Service%20for%20International%20Air%20Navigation/Cover%20sheet%20to%20AMDT%2074.pdf>

### ICAO

INTERNATIONAL STANDARDS  
AND RECOMMENDED PRACTICES  
METEOROLOGICAL SERVICE FOR  
INTERNATIONAL AIR NAVIGATION

PART I — CORE SARPS

PART II — APPENDICES AND ATTACHMENTS

ANNEX 3

TO THE CONVENTION ON INTERNATIONAL CIVIL AVIATION

SIXTEENTH EDITION — JULY 2007

INTERNATIONAL CIVIL AVIATION ORGANIZATION

Appendix 3 Annex 3 — Meteorological Service for International Air Navigation

APP 3-15 75/11 1/08 7

No. 74

4.5 Clouds

#### 4.5.1 Siting

Recommendation.— When instrumented systems are used for the measurement of the cloud amount and the height of cloud base, representative observations should be obtained by the use of sensors appropriately sited. For local routine and

special reports, in the case of aerodromes with precision approach runways, sensors for cloud amount and height of cloud

base should be sited to give the best practicable indications of the height of cloud base and cloud amount at the middle

marker site of the instrument landing system or, at aerodromes where a middle marker beacon is not used, at a distance of

900 to 1 200 m (3 000 to 4 000 ft) from the landing threshold at the approach end of the runway.

Note.— Specifications concerning the middle marker site of an instrument landing system are given in Annex 10,

Volume I, Chapter 3 and at Attachment C, Table C-5.

#### 4.5.2 Display

Recommendation.— When automated equipment is used for the measurement of the height of cloud base, height of

cloud base display(s) should be located in the meteorological station with corresponding display(s) in the appropriate air

traffic services units. The displays in the meteorological station and in the air traffic services units should relate to the same

sensor, and where separate sensors are required as specified in 4.5.1, the displays should clearly identify the area monitored

by each sensor.

#### 4.5.3 Reference level

Recommendation.— The height of cloud base should normally be reported above aerodrome elevation. When a precision approach runway is in use which has a threshold elevation 15 m (50 ft) or more below the aerodrome elevation,

local arrangements should be made in order that the height of cloud bases reported to arriving aircraft should refer to the

threshold elevation. In the case of reports from offshore structures, the height of cloud base should be given above mean sea

level.

#### 4.5.4 Reporting

4.5.4.1 In local routine and special reports and in METAR and SPECI, the height of cloud base shall be reported in

steps of 30 m (100 ft) up to 3 000 m (10 000 ft). Any observed value which does not fit the reporting scale in use shall be

rounded down to the nearest lower step in the scale.

4.5.4.2 Recommendation.— In local routine and special reports and in METAR and SPECI:

a) cloud amount should be reported using the abbreviations “FEW” (1 to 2 oktas), “SCT” (3 to 4 oktas), “BKN” (5 to

7 oktas) or "OVC" (8 oktas);

b) **cumulonimbus clouds and towering cumulus clouds should be indicated as "Cb" and "TCU",** respectively;

c) the vertical visibility should be reported in steps of 30 m (100 ft) up to 600 m (2 000 ft);

d) if there are no clouds of operational significance and no restriction on vertical visibility and the abbreviation "CAVOK" is not appropriate, the abbreviation "NSC" should be used;

e) when several layers or masses of cloud of operational significance are observed, their amount and height of cloud

base should be reported in increasing order of the height of cloud base, and in accordance with the following criteria:

1) the lowest layer or mass, regardless of amount to be reported as FEW, SCT, BKN or OVC as appropriate;

2) the next layer or mass, covering more than 2/8 to be reported as SCT, BKN or OVC as appropriate;

3) the next higher layer or mass, covering more than 4/8 to be reported as BKN or OVC as appropriate; and

4) cumulonimbus and/or towering cumulus clouds, whenever observed and not reported in 1) to 3);

f) when the cloud base is diffuse or ragged or fluctuating rapidly, the minimum height of cloud base, or cloud fragments, should be reported; and

g) when an individual layer (mass) of cloud is composed of **cumulonimbus and towering cumulus clouds** with a

common cloud base, the type of cloud should be reported as cumulonimbus only.

Note.— Towering cumulus indicates cumulus congestus clouds of great vertical extent.

4.5.4.3 In local routine and special reports:

a) the units of measurement used for the height of cloud base and vertical visibility shall be indicated; and

b) when there is more than one runway in use and the heights of cloud bases are observed by instruments for these

runways, the available heights of cloud bases for each runway shall be reported and the runways to which the values

refer shall be indicated.

4.5.4.4 Recommendation.— In automated METAR and SPECI:

a) when the cloud type cannot be observed by the automatic observing system, the cloud type in each cloud group

should be replaced by "///";

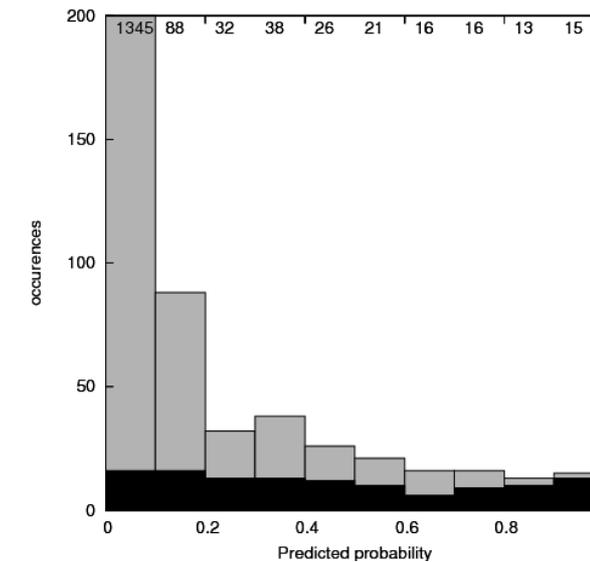
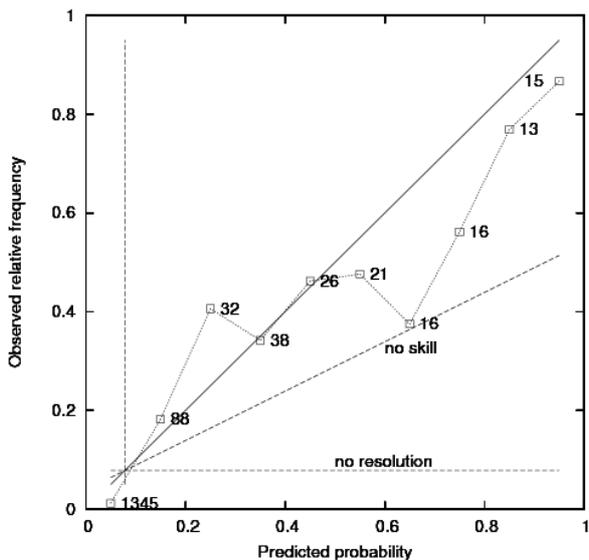
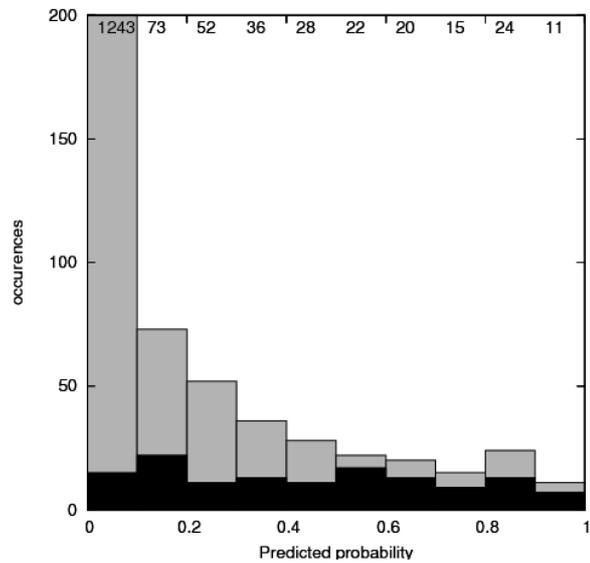
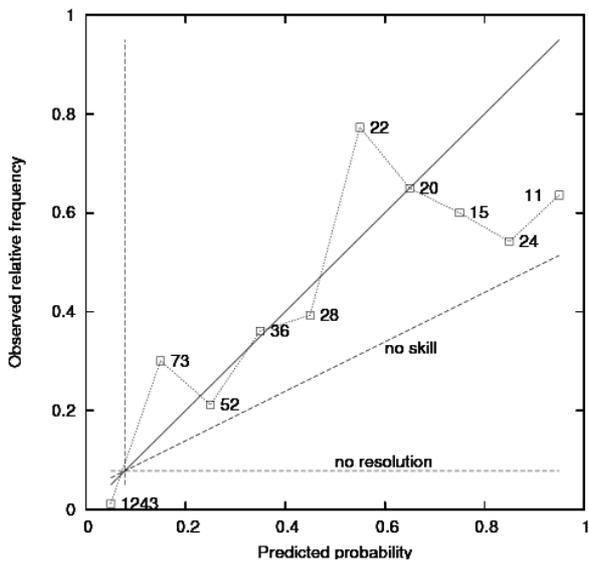
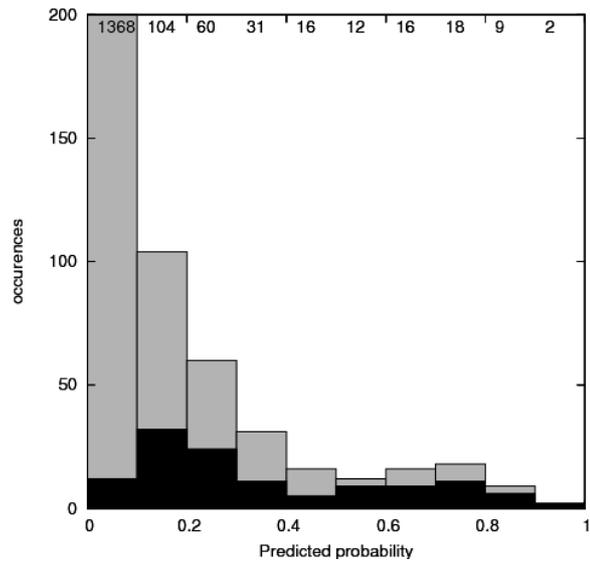
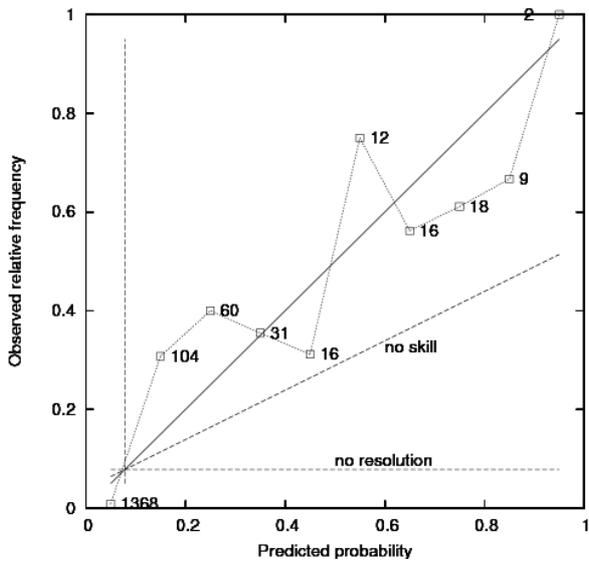
b) when no clouds are detected by the automatic observing system, it should be indicated by using the abbreviation

"NCD"; and

c) when cumulonimbus clouds or towering cumulus clouds are detected by the automatic observing system and the

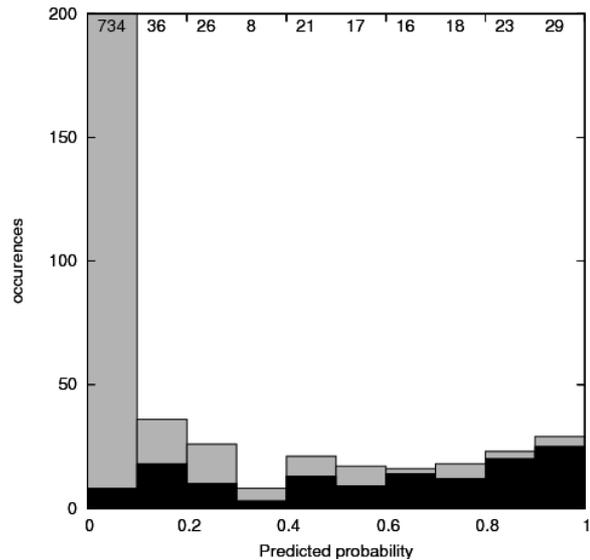
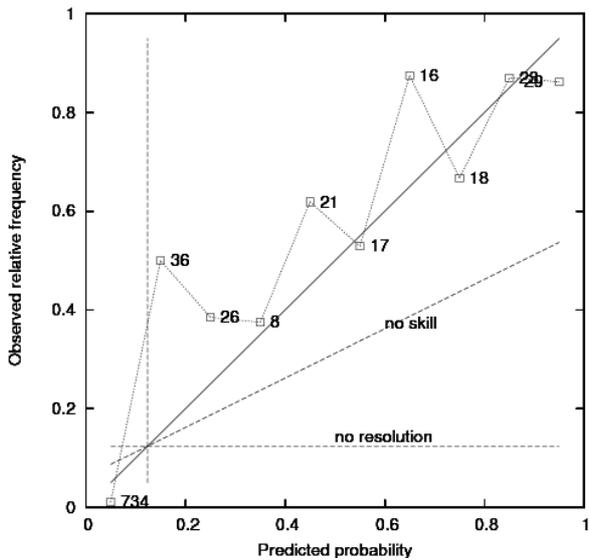
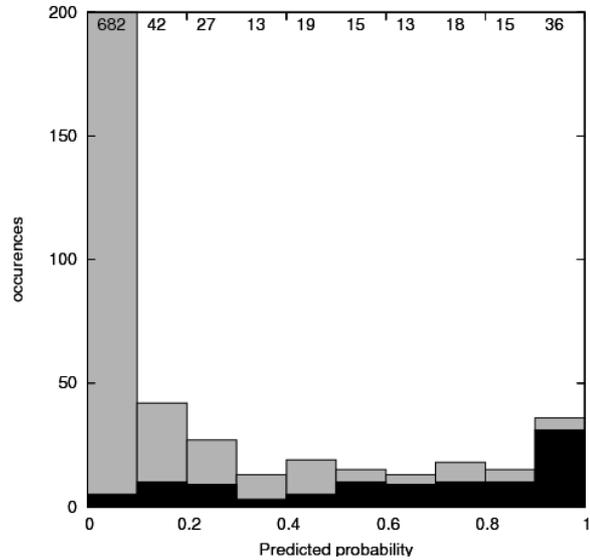
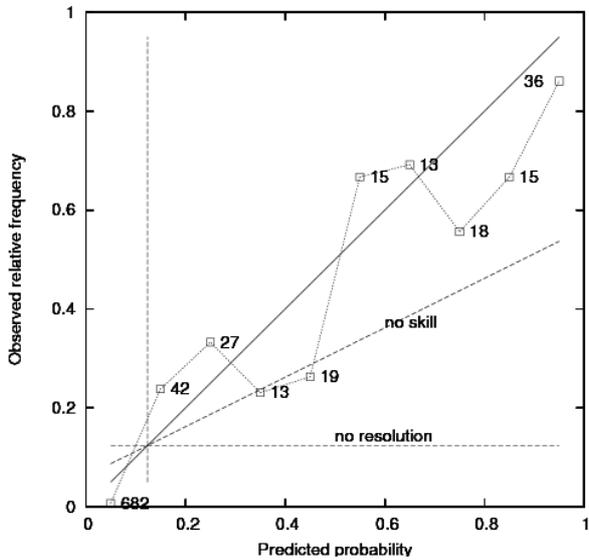
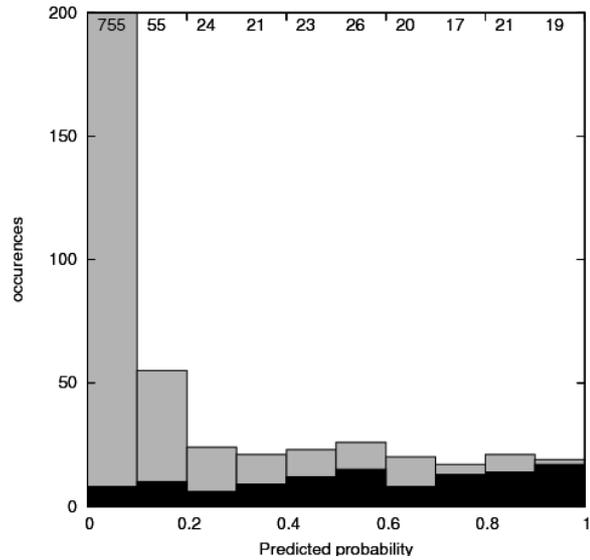
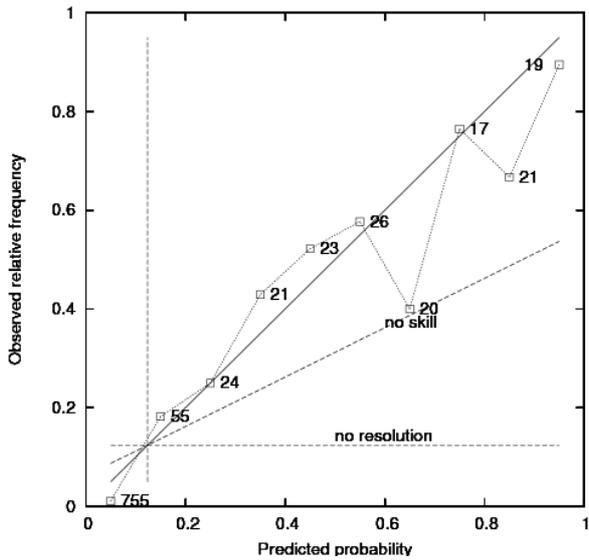
cloud amount and the height of cloud base cannot be observed, the cloud amount and the height of cloud base should be replaced by "/////".

Appendix 3: attributes diagrams and histograms



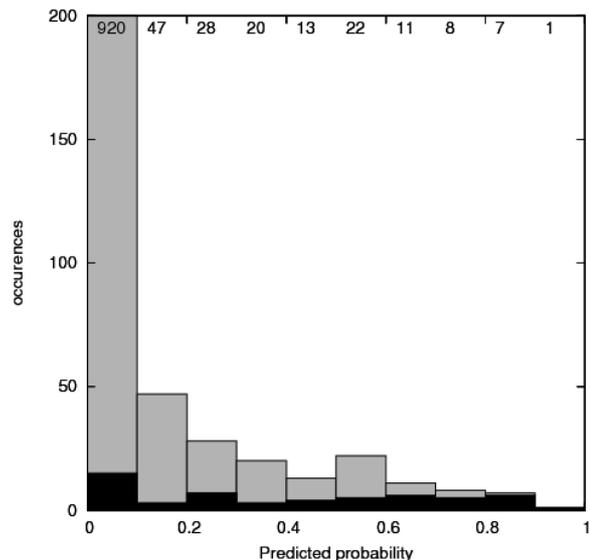
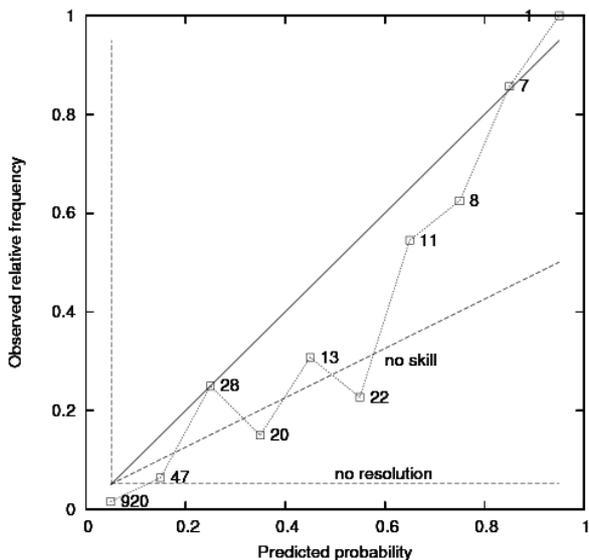
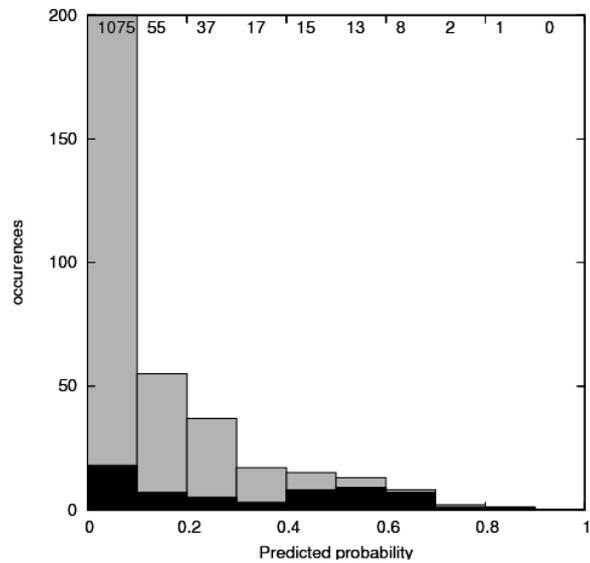
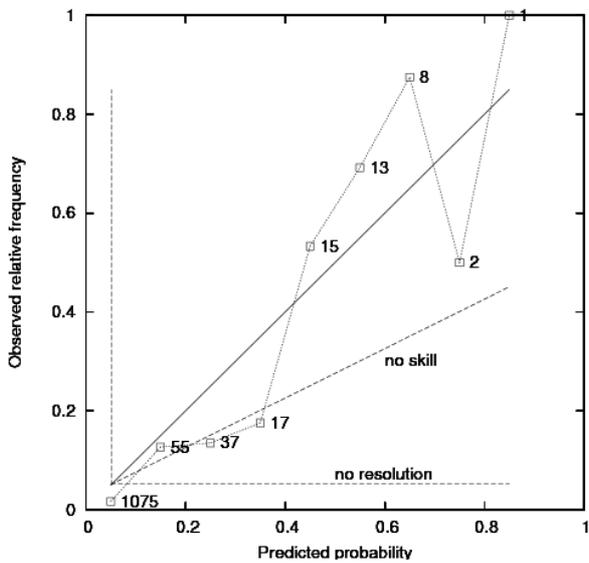
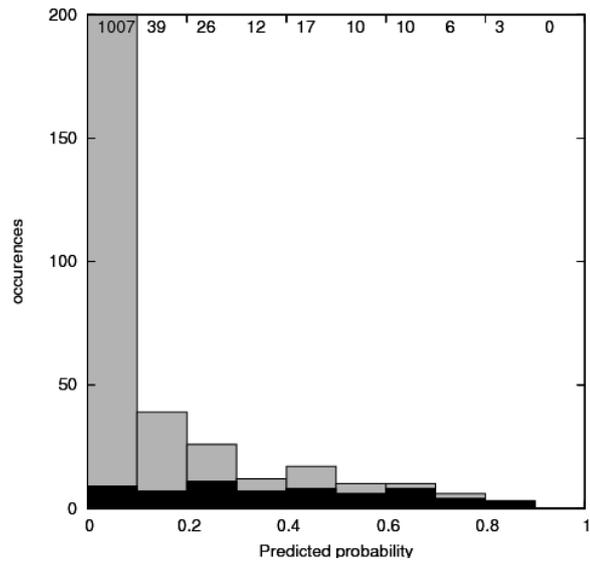
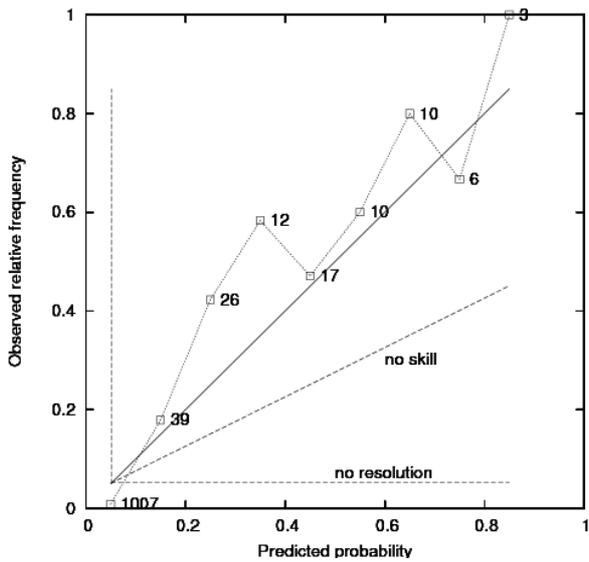
*EHAM winter night attributes diagram for the three permutations of dependent and independent data parts.*

*EHAM winter night histogram distribution for the three permutations of dependent and independent data parts. Grey all metar reports, black Cb-Tcu observations. Note that the first bin is out of scale.*



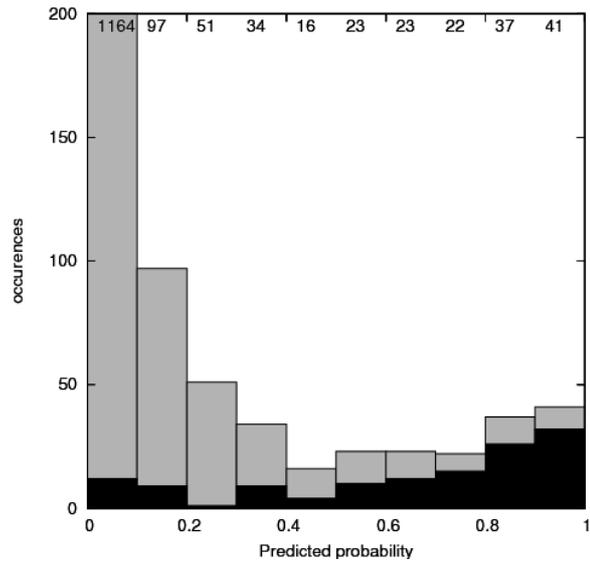
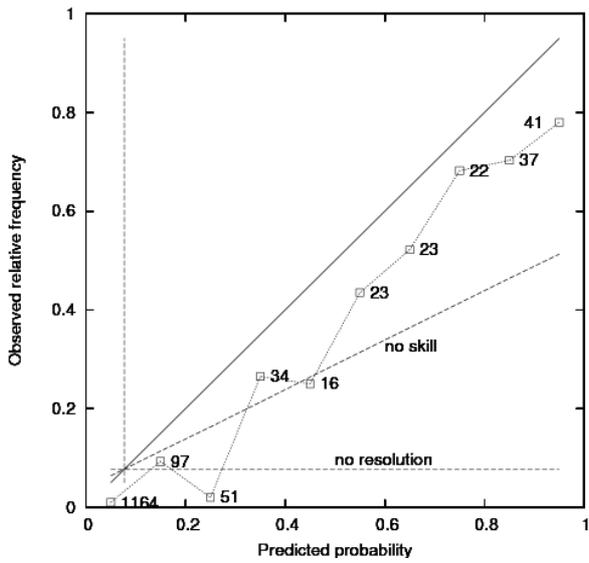
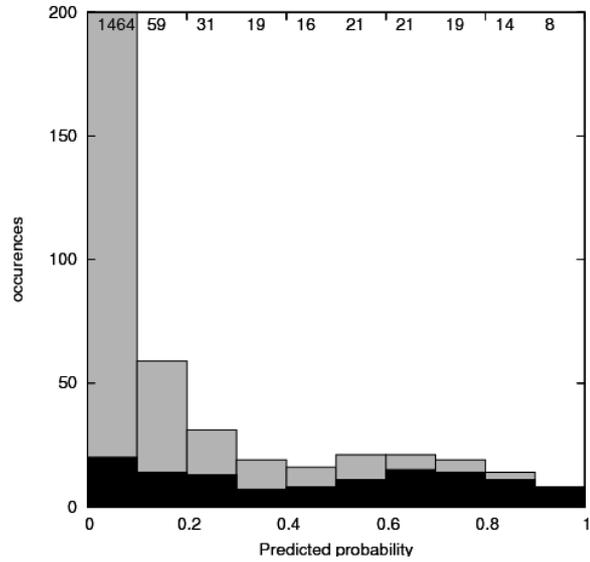
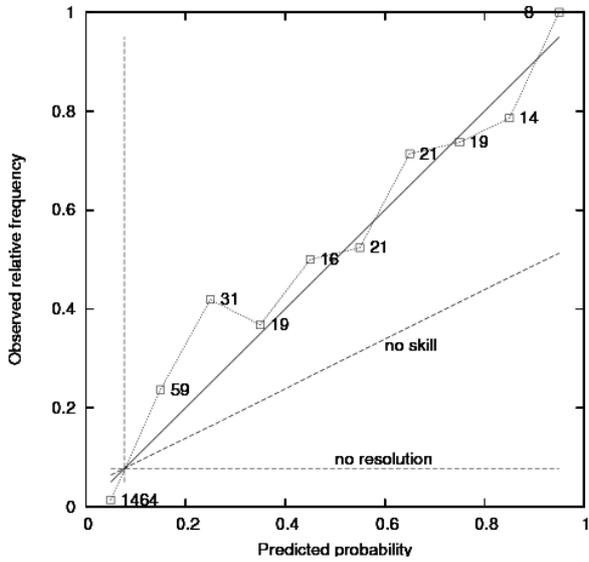
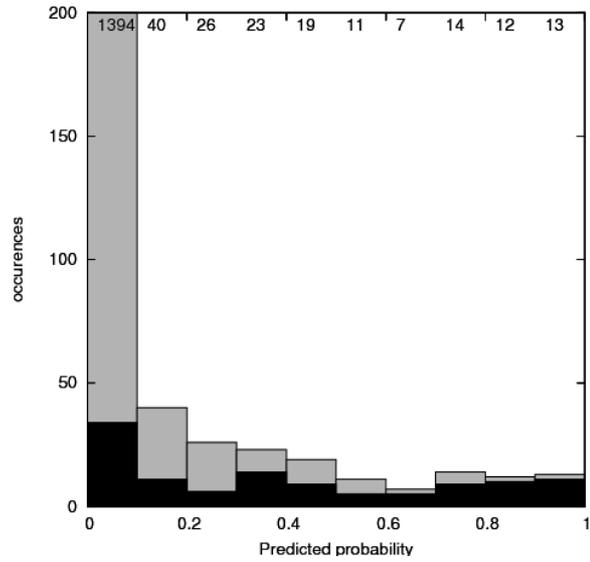
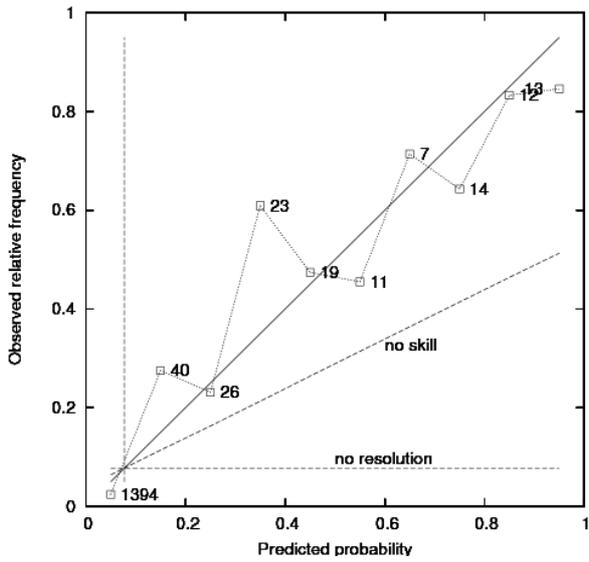
*EHAM winter day attributes diagram for the three permutations of dependent and independent data parts.*

*EHAM winter day histogram distribution for the three permutations of dependent and independent data parts. Grey all metar reports, black Cb-Tcu observations. Note that the first bin is out of scale.*



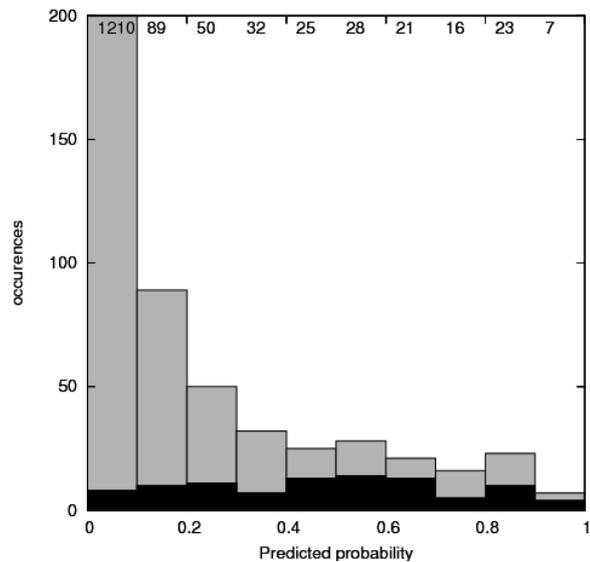
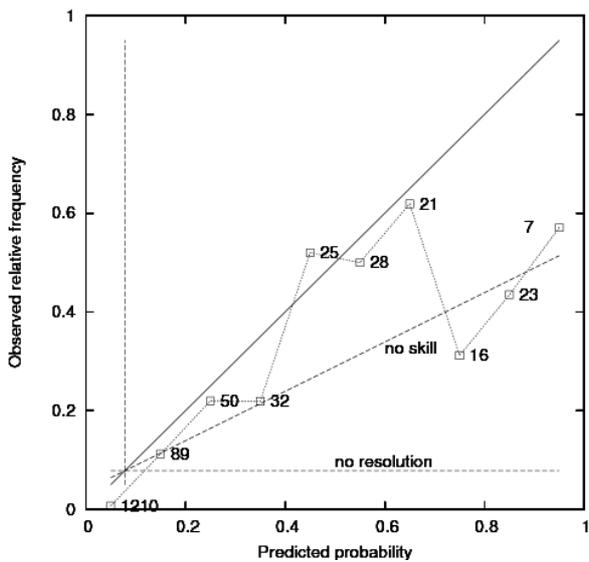
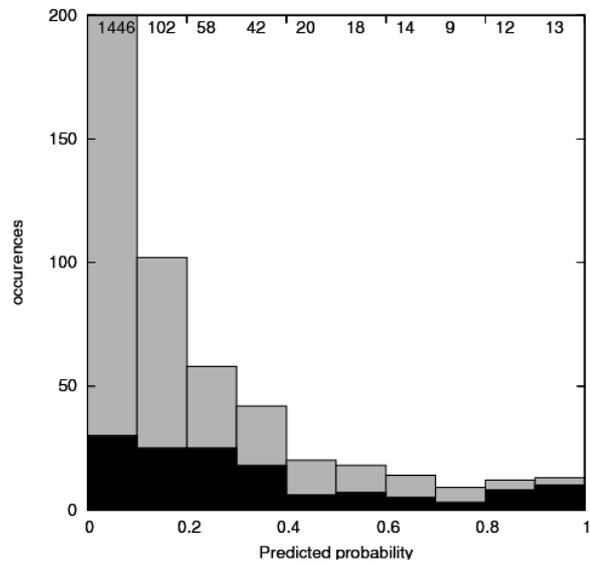
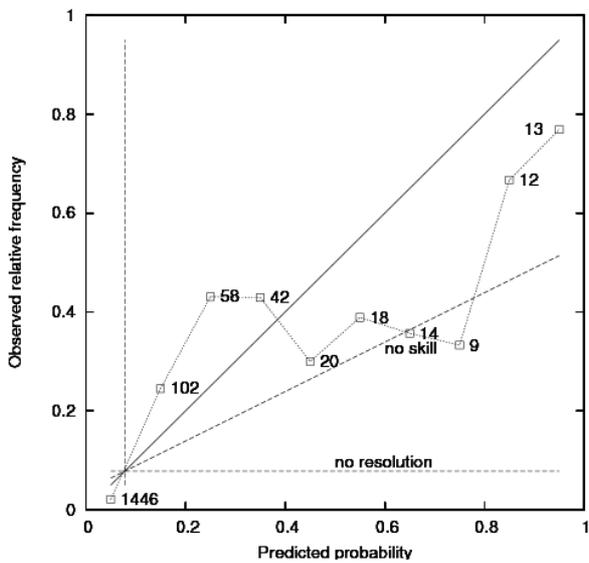
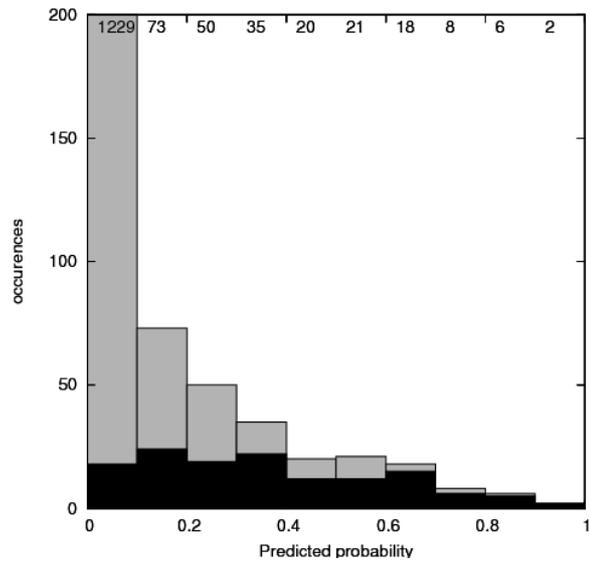
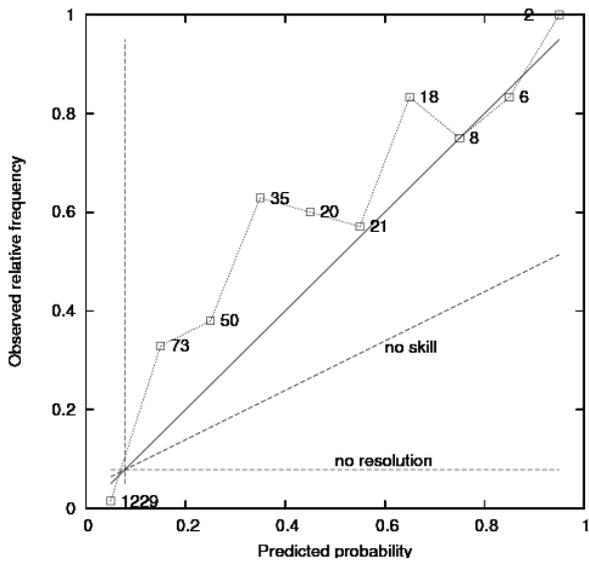
*EHAM summer night attributes diagram for the three permutations of dependent and independent data parts.*

*EHAM summer night histogram distribution for the three permutations of dependent and independent data parts. Grey all meteor reports, black Cb-Tcu observations. Note that the first bin is out of scale.*



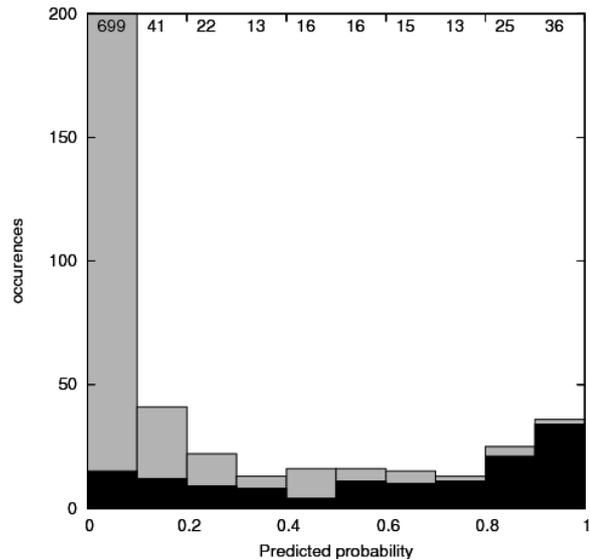
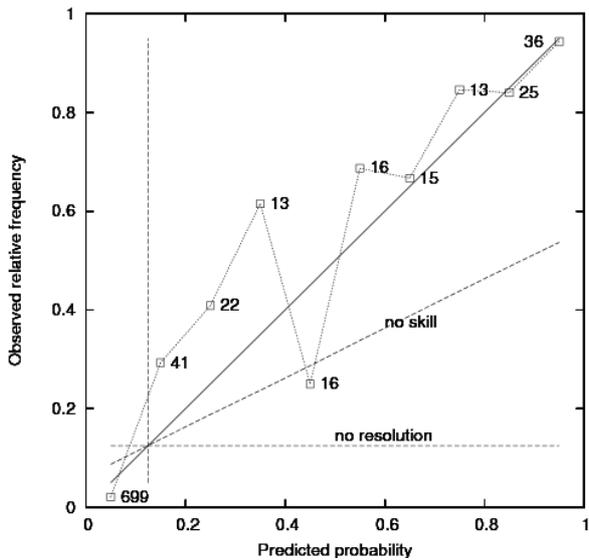
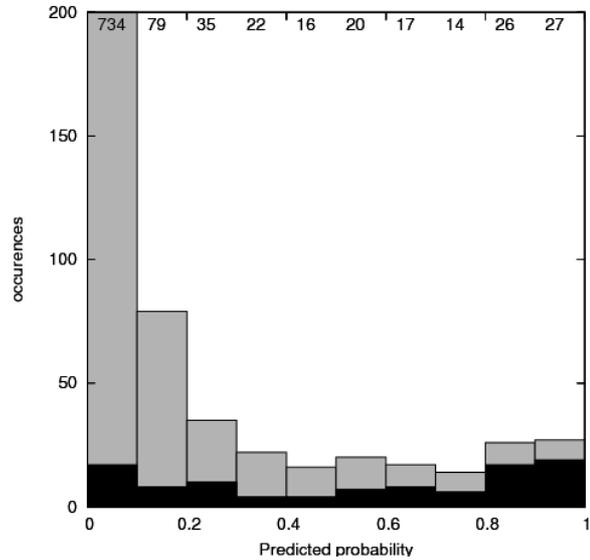
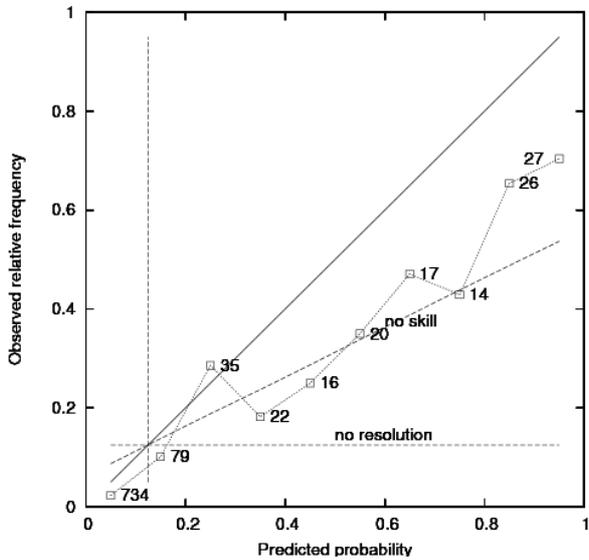
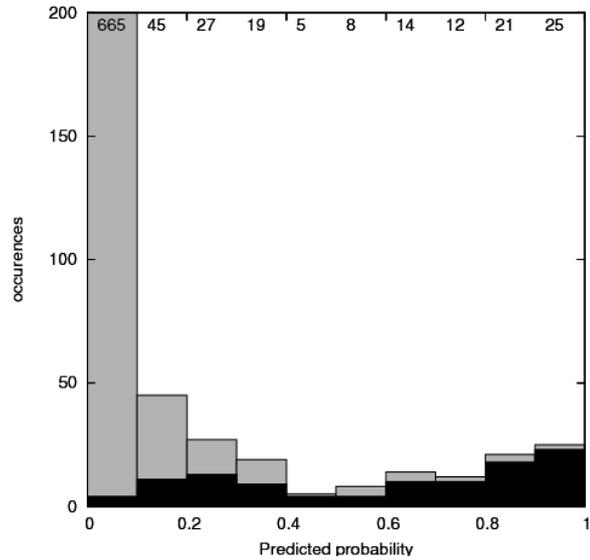
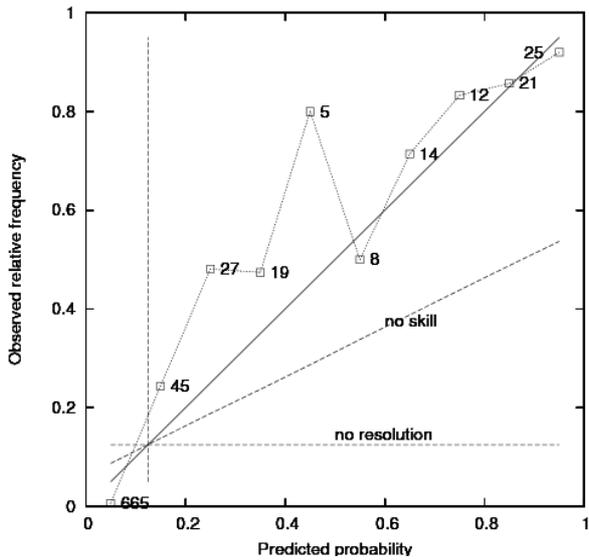
*EHAM summer day attributes diagram for the three permutations of dependent and independent data parts.*

*EHAM summer day histogram distribution for the three permutations of dependent and independent data parts. Grey all meteor reports, black Cb-Tcu observations. Note that the first bin is out of scale.*



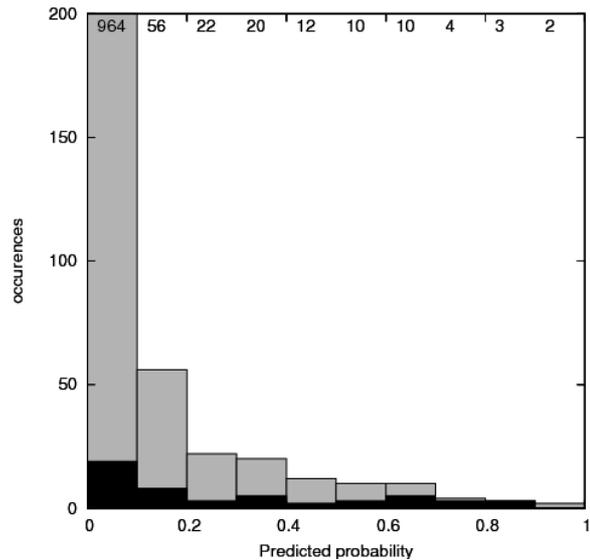
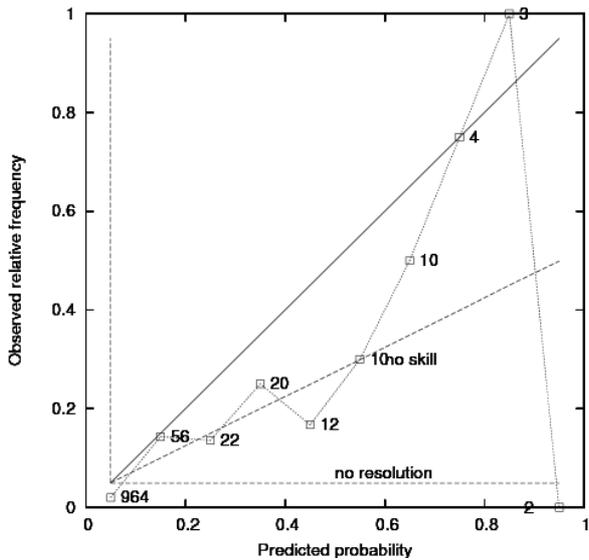
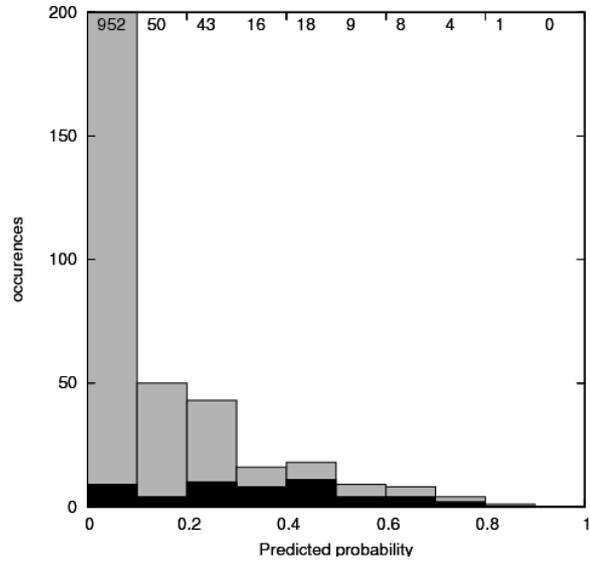
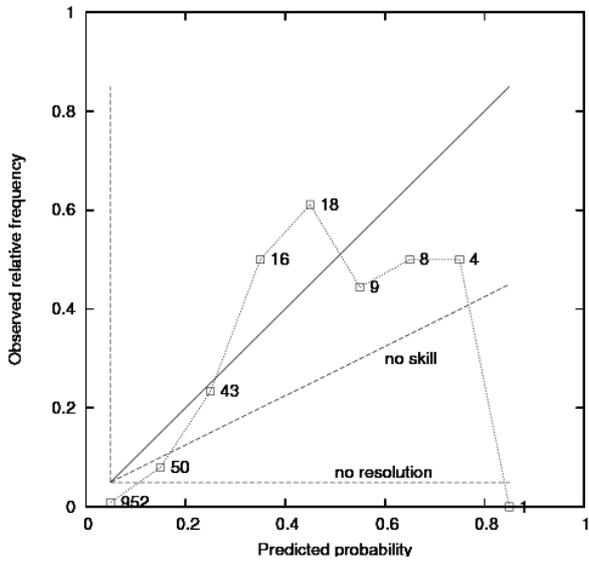
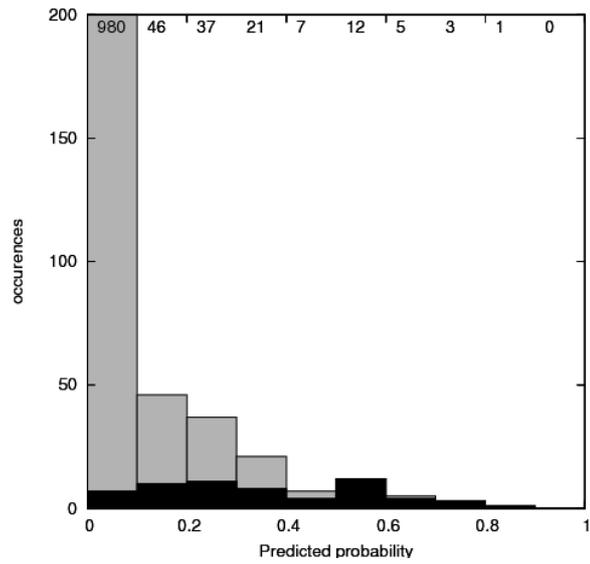
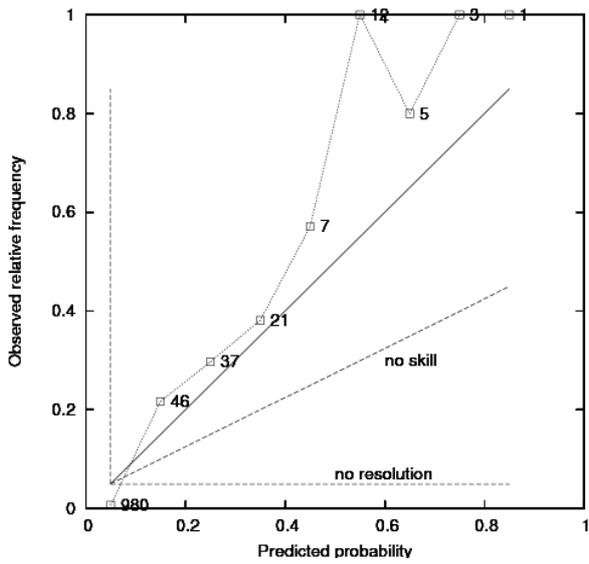
*EHRD winter night attributes diagram for the three permutations of dependent and independent data parts.*

*EHRD winter night histogram distribution for the three permutations of dependent and independent data parts. Grey all metar reports, black Cb-Tcu observations. Note that the first bin is out of scale.*



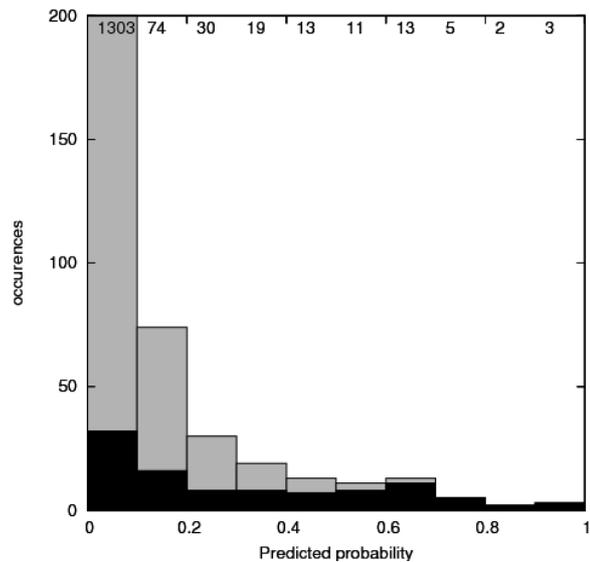
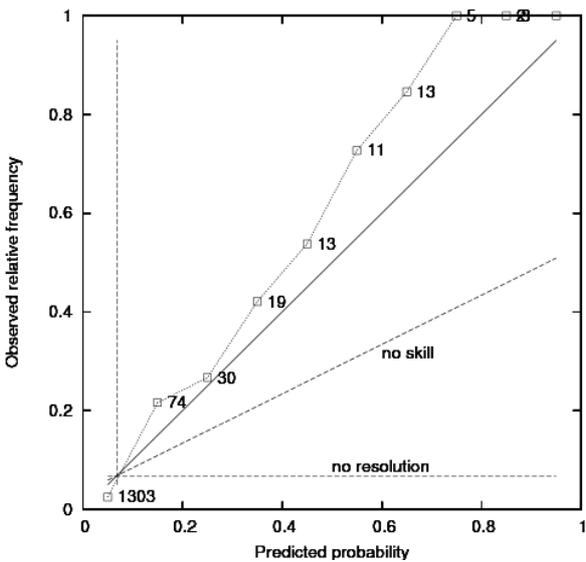
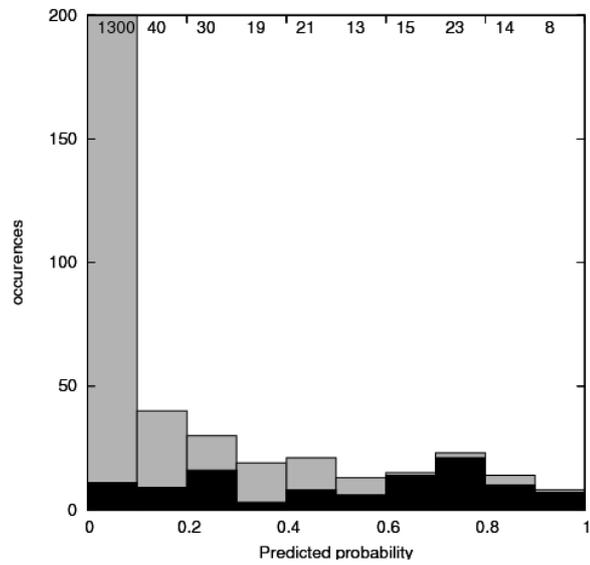
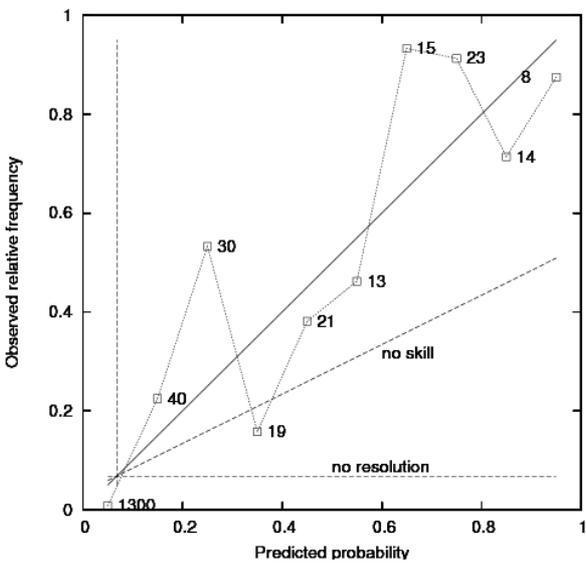
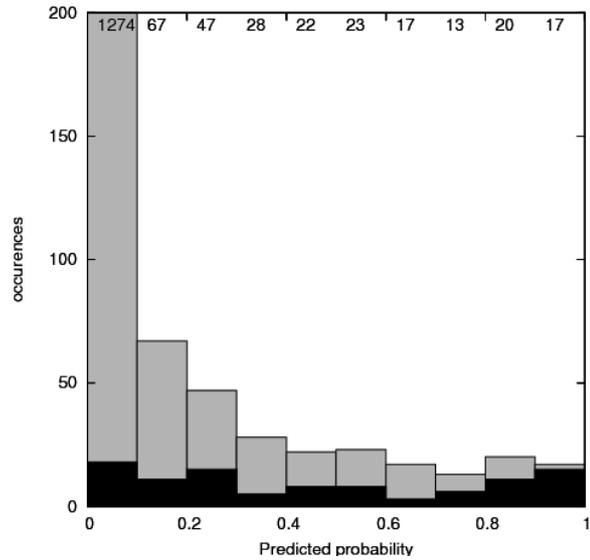
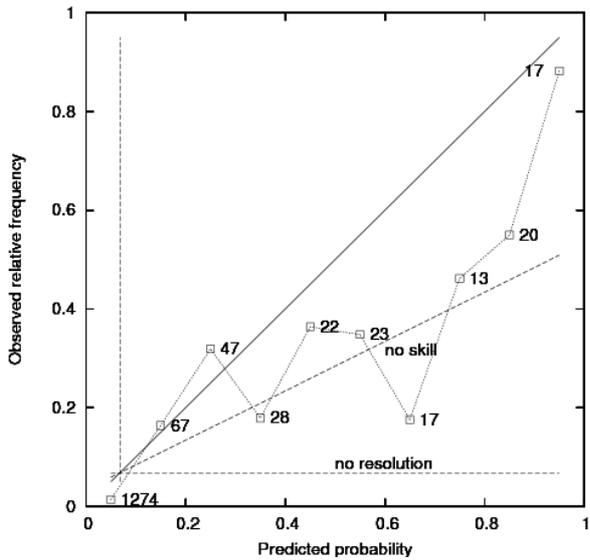
*EHRD winter day attributes diagram for the three permutations of dependent and independent data parts.*

*EHRD winter day histogram distribution for the three permutations of dependent and independent data parts. Grey all metar reports, black Cb-Tcu observations. Note that the first bin is out of scale.*



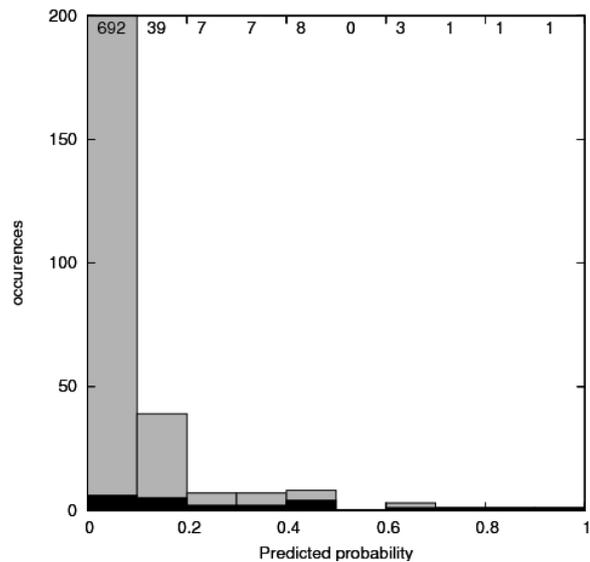
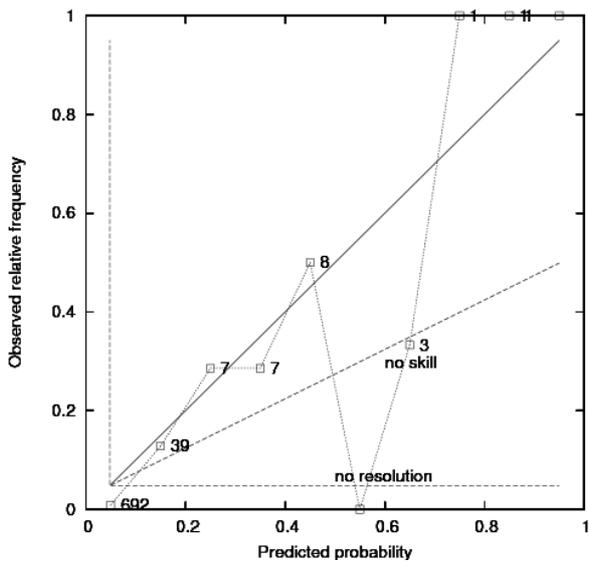
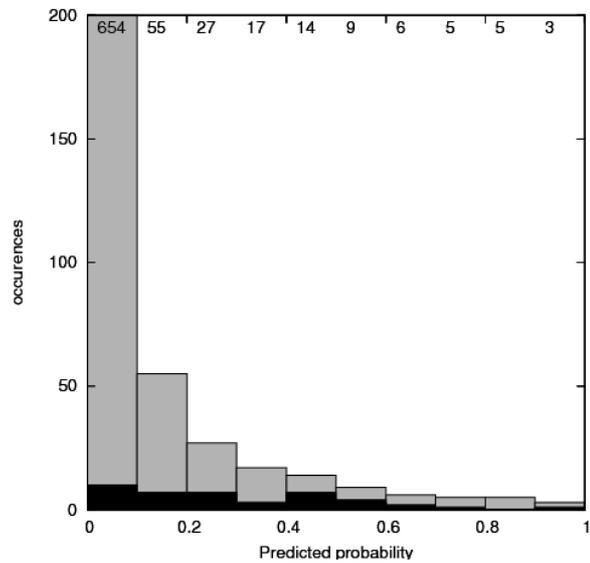
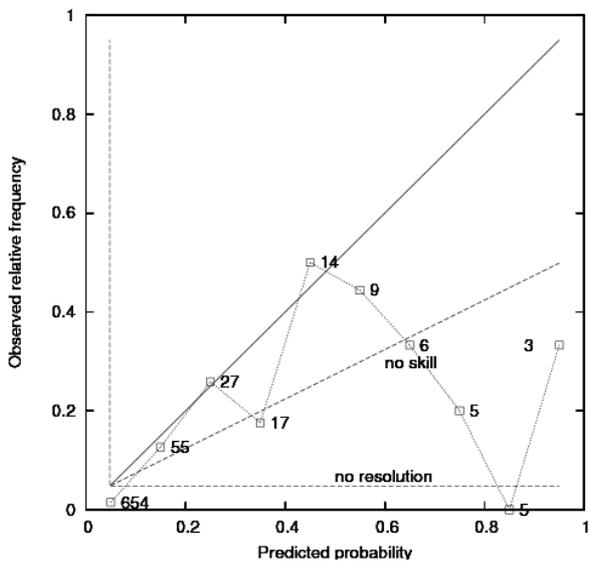
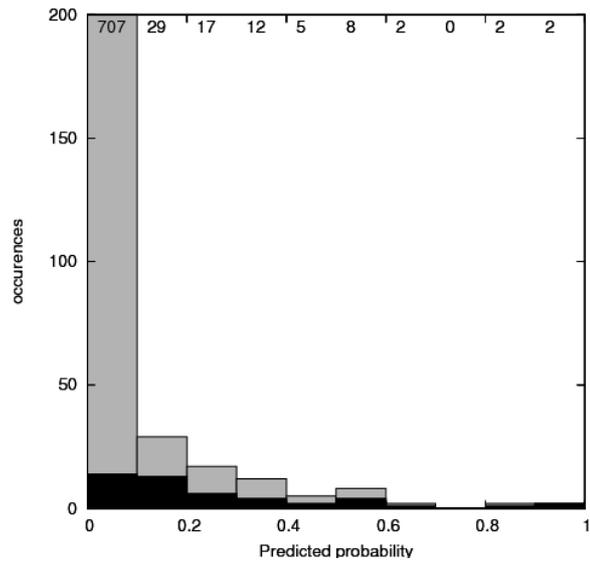
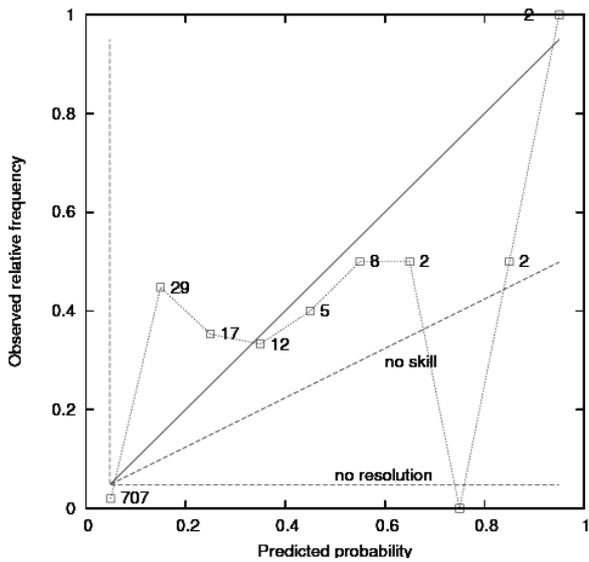
*EHRD summer night attributes diagram for the three permutations of dependent and independent data parts.*

*EHRD summer night histogram distribution for the three permutations of dependent and independent data parts. Grey all meteor reports, black Cb-Tcu observations. Note that the first bin is out of scale.*

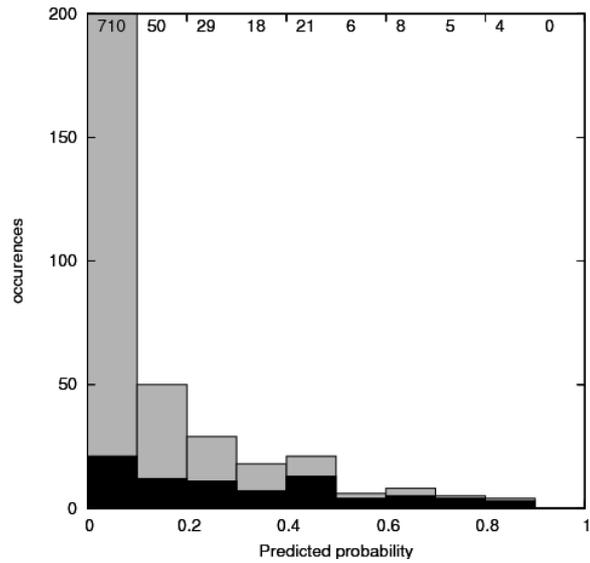
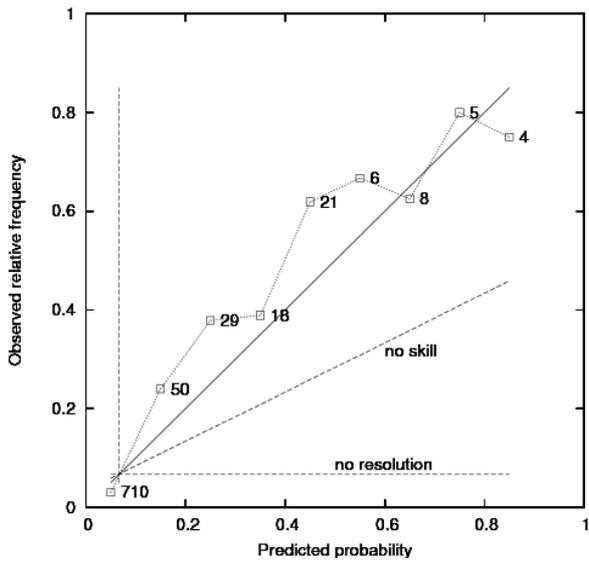
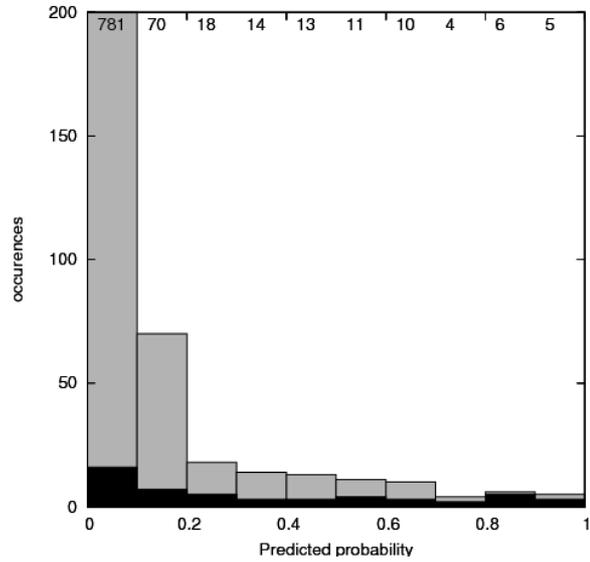
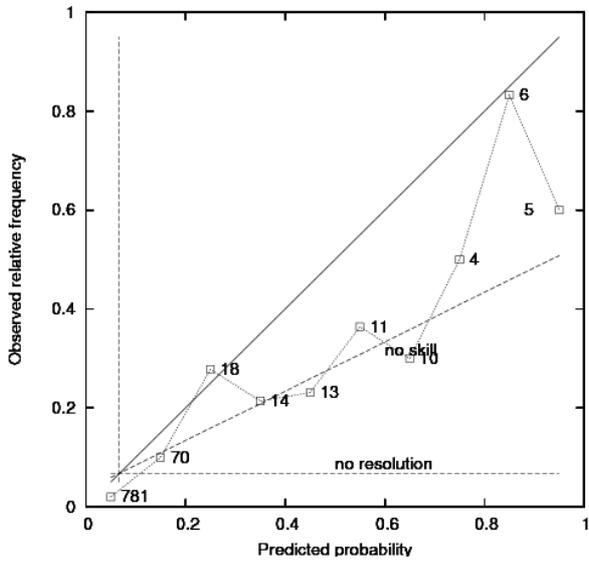
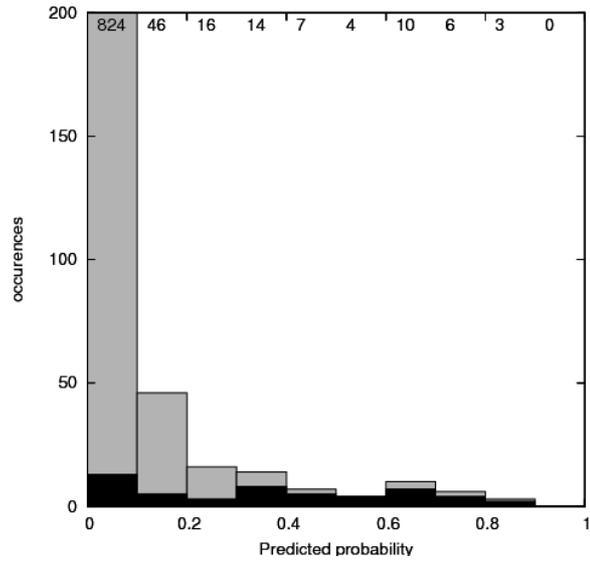
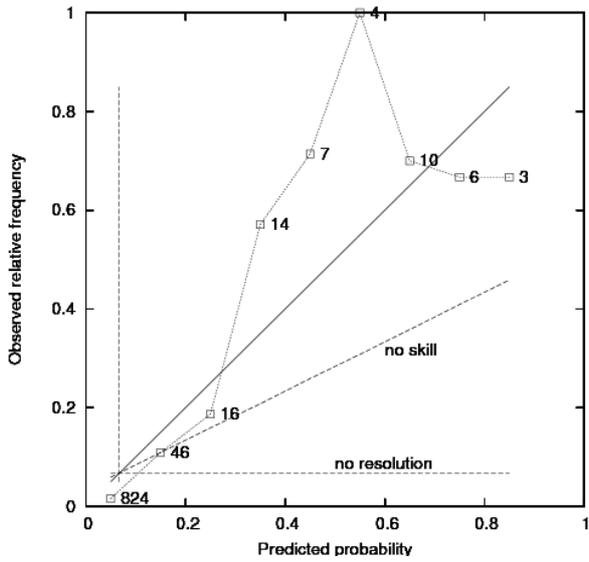


*EHRD summer day attributes diagram for the three permutations of dependent and independent data parts.*

*EHRD summer day histogram distribution for the three permutations of dependent and independent data parts. Grey all meter reports, black Cb-Tcu observations. Note that the first bin is out of scale.*

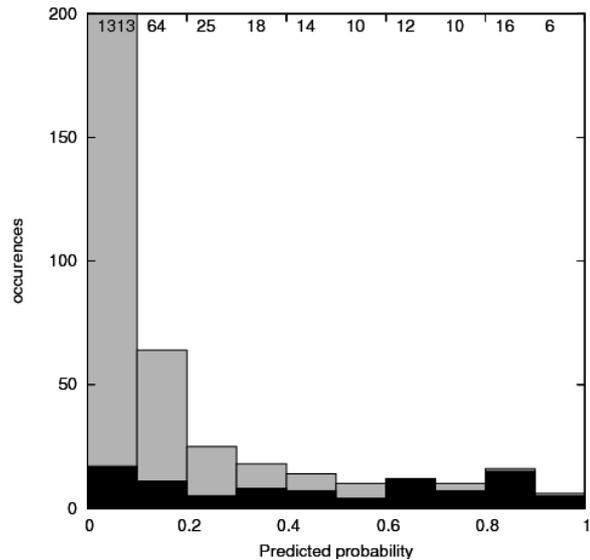
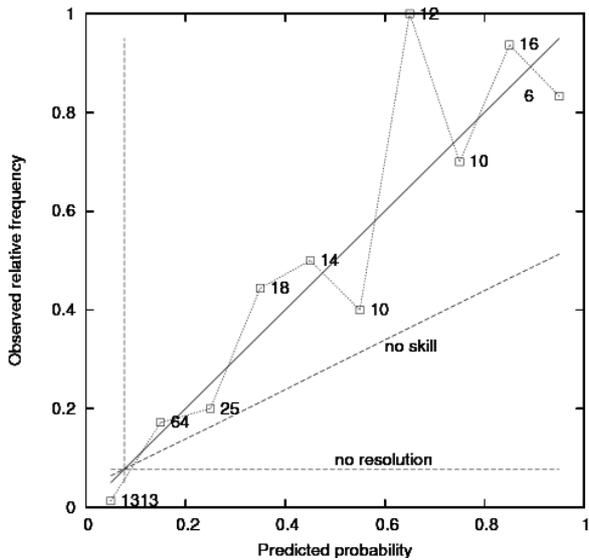
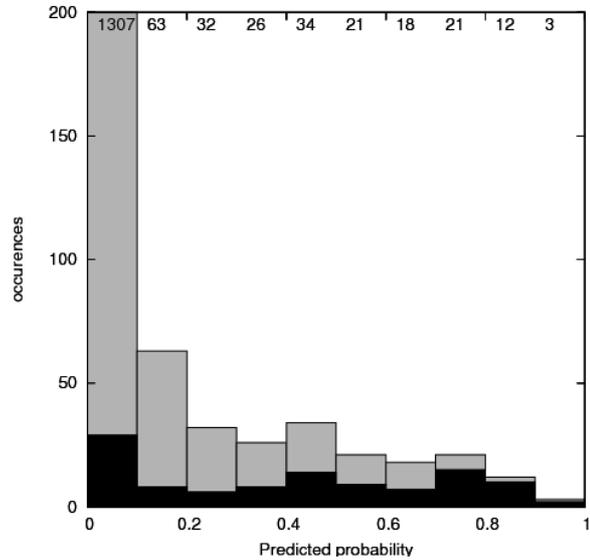
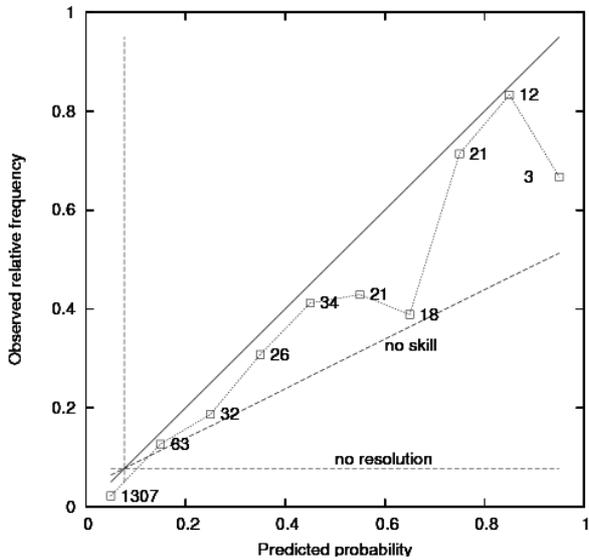
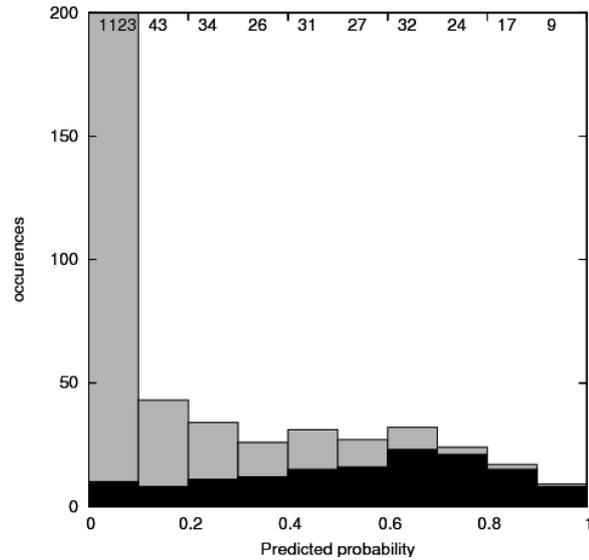
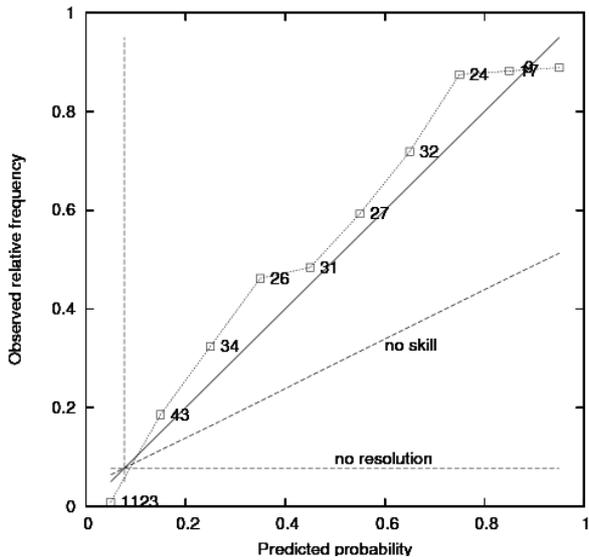


*EHGG winter night attributes diagram for the EHGG winter night histogram distribution for three permutations of dependent and independent data parts. Grey all metar reports, black Cb-Tcu observations. Note that the first bin is out of scale.*



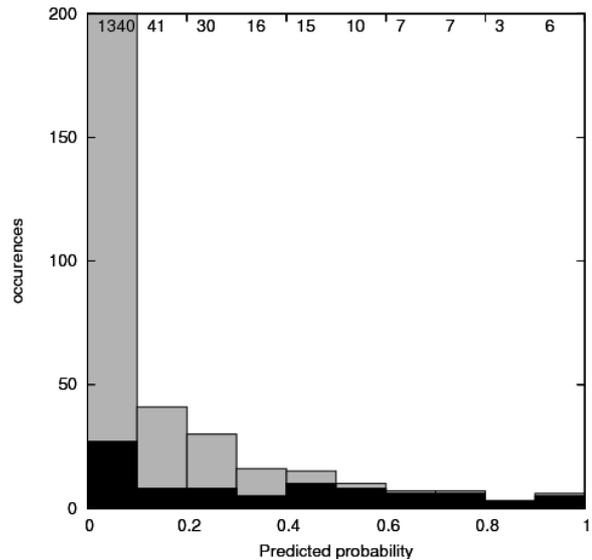
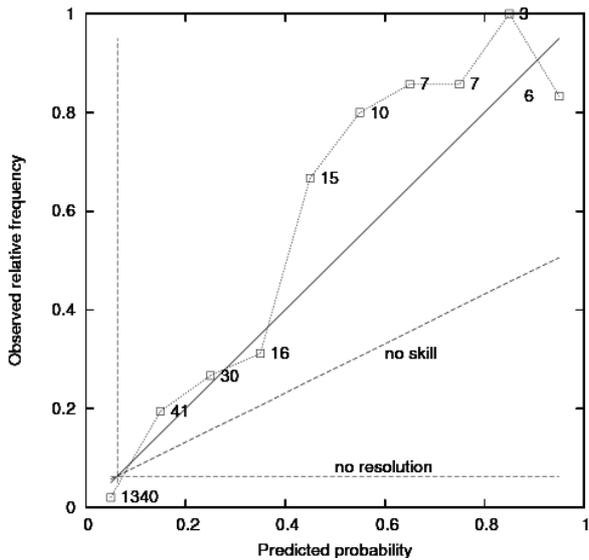
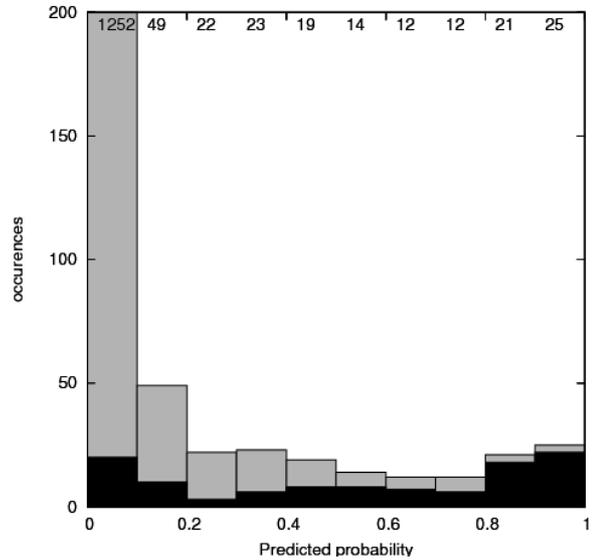
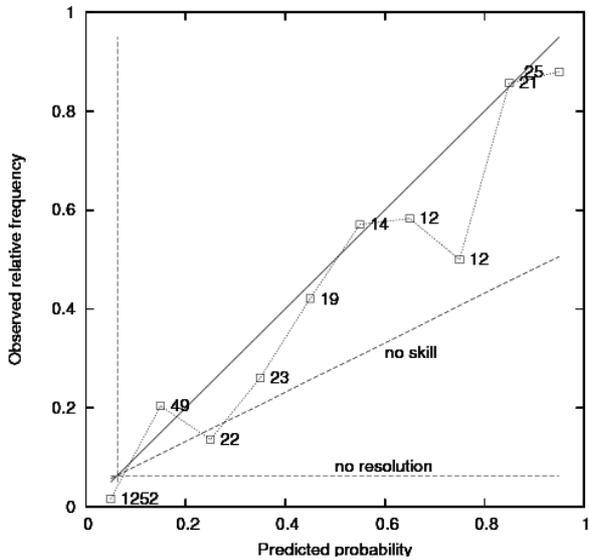
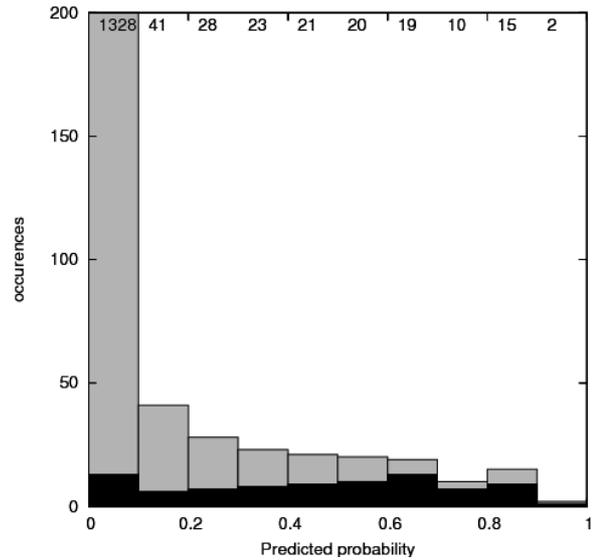
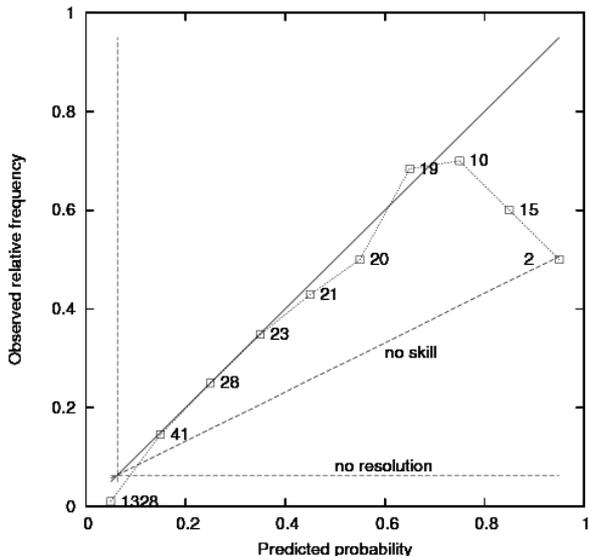
*EHGG winter day attributes diagram for the three permutations of dependent and independent data parts.*

*EHGG winter day histogram distribution for the three permutations of dependent and independent data parts. Grey all metar reports, black Cb-Tcu observations. Note that the first bin is out of scale.*



*EHGG summer day attributes diagram for the three permutations of dependent and independent data parts.*

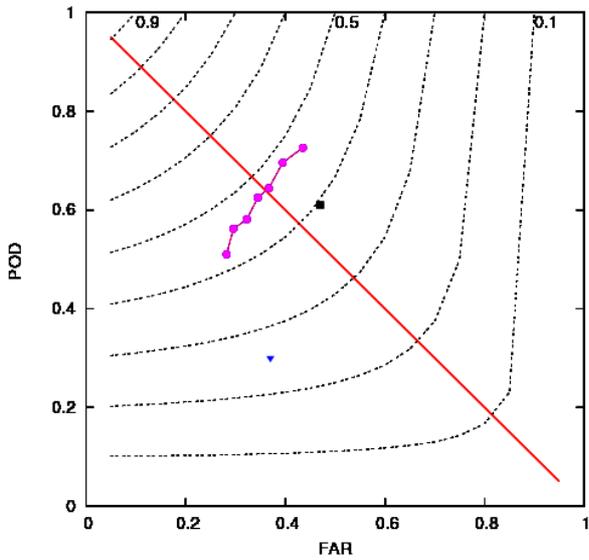
*EHGG summer day histogram distribution for the three permutations of dependent and independent data parts. Grey all metar reports, black Cb-Tcu observations. Note that the first bin is out of scale.*



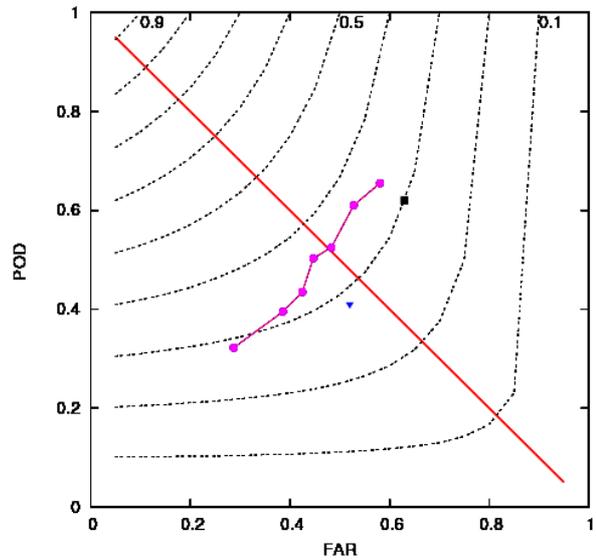
*EHBK summer day attributes diagram for the EHBK summer day histogram distribution for three permutations of dependent and independent data parts. Grey all meteor reports, black Cb-Tcu observations. Note that the first bin is out of scale.*

Appendix 4 POD FAR diagrams for the complete 2005 data set. No independent data are used here.

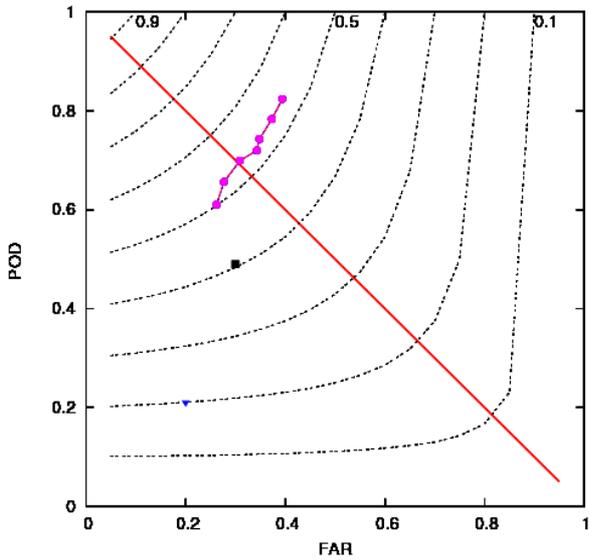
*Black square operational algorithm-2007 at 30 km, blue triangle operational algorithm-2007 at 15 km.*



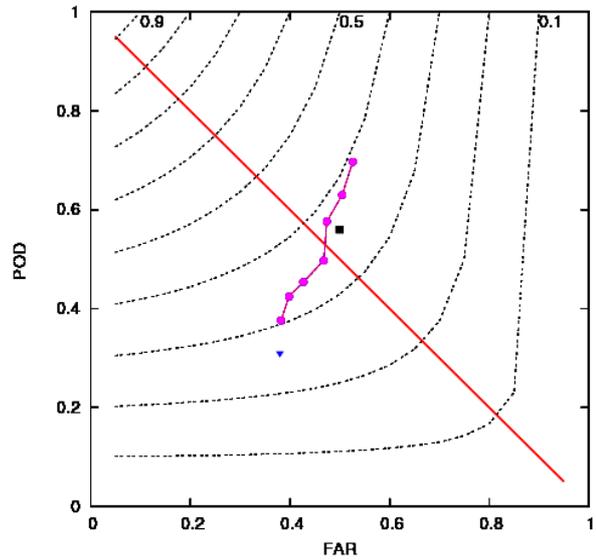
*EHAM summer day BSS 0.44*



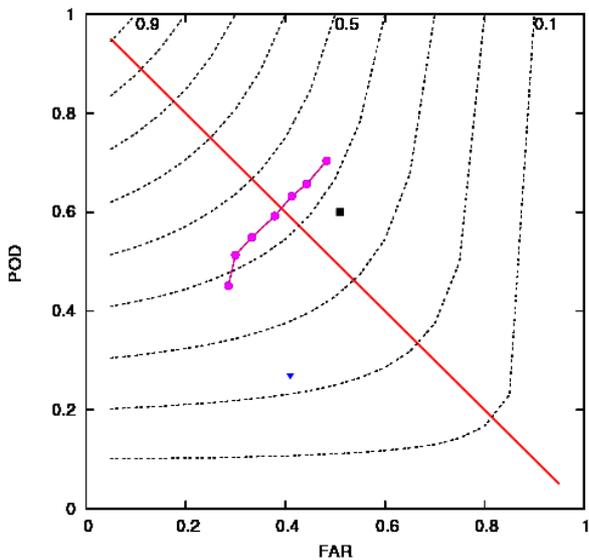
*EHAM summer night BSS 0.39*



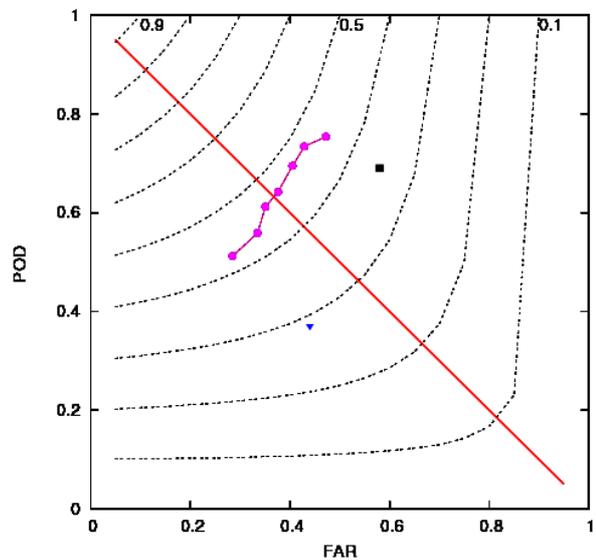
*EHAM winter day BSS 0.52*



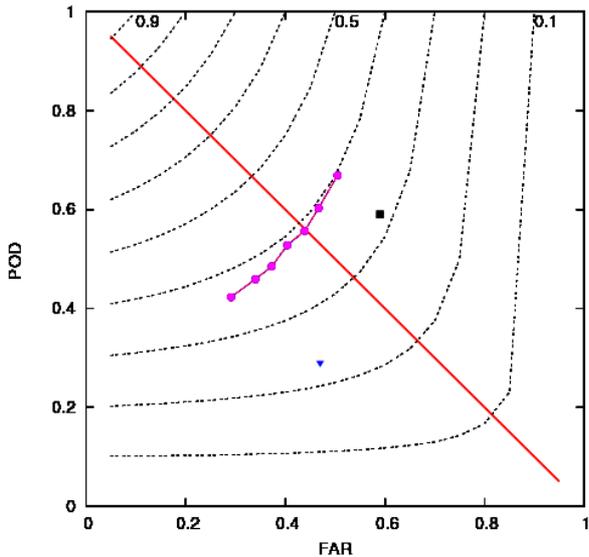
*EHAM winter night BSS 0.34*



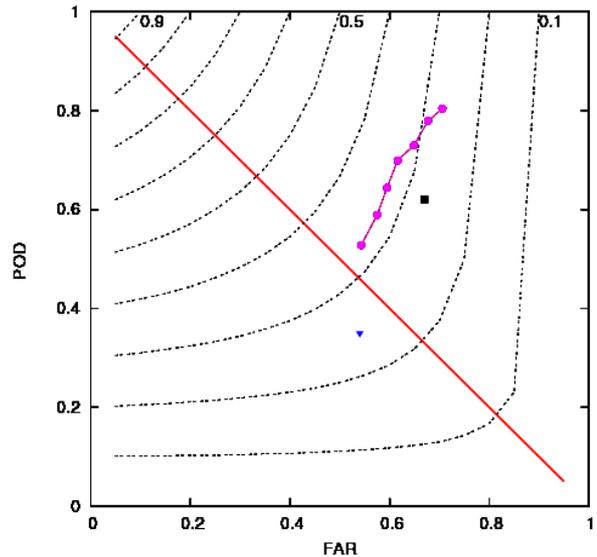
*EHBK summer day BSS 0.41*



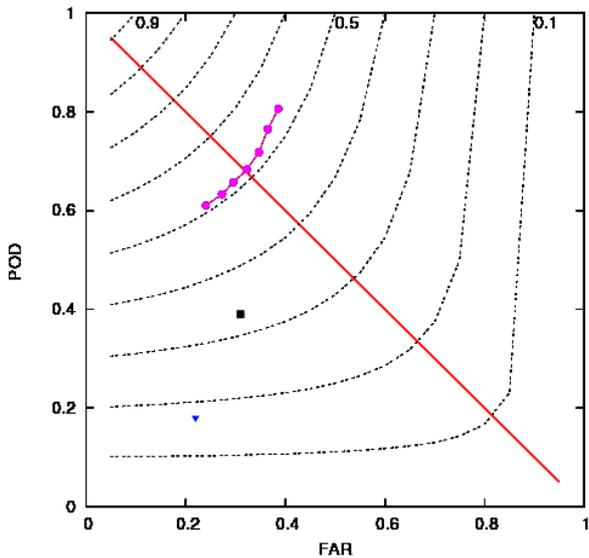
*EHGG summer day BSS 0.44*



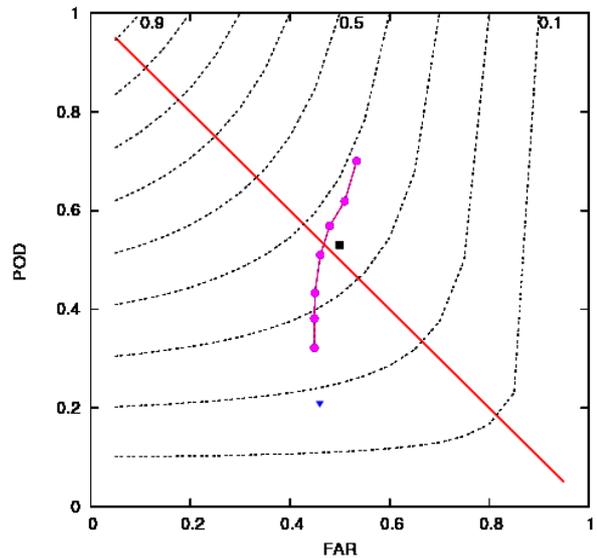
*EHRD summer day BSS 0.38*



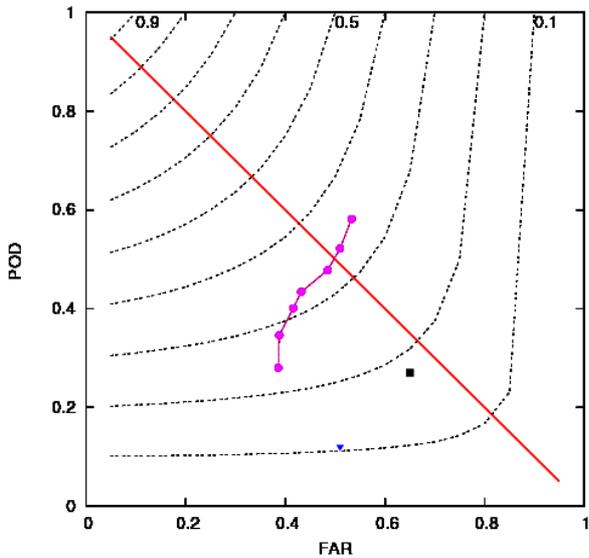
*EHRD summer night BSS 0.19*



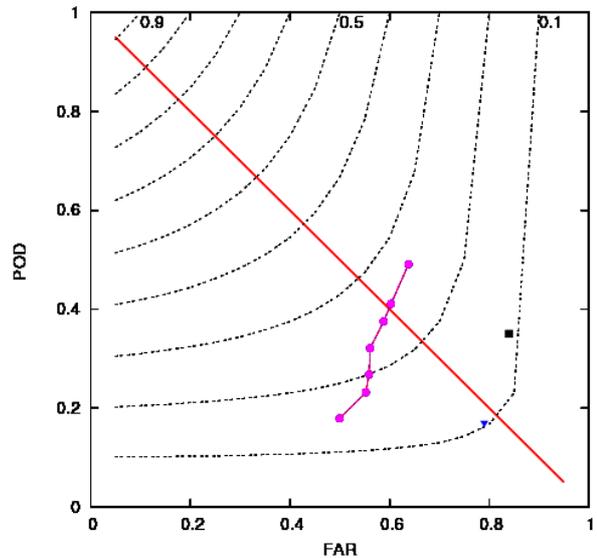
*EHRD winter day BSS 0.51*



*EHRD winter night BSS 0.31*



*EHGG winter day BSS 0.28*



*EHGG winter night BSS 0.20 (not complete dataset)*

The applied predictors and their coefficients and significance

#ams wd		# eelde wd	
Rad Cont 15 km	.190 .000	Rad Cont 30 km	.357 .000
Rad Cont 30 km	.399 .000	Std dev T cloud average 15 km	.271 .000
Std dev T cloud average 15 km	.254 .000	HRV min 15 km	-.026 .000
HRV min 15 km	-.023 .000	HRV max 15 km	.014 .000
HRV max 15 km	.009 .000	Constant	-5.112 .000
Constant	-4.439 .000	#eelde wn	
#ams wn		Rad cont sum 15 km	.028 .000
Rad cont sum 15 km	.022 .000	Rad Cont 30 km	.319 .000
Rad Cont 30 km	.343 .000	Std dev T cloud average 15 km	.279 .000
Std dev T cloud average 15 km	.316 .000	Constant	-4.878 .000
Constant	-4.809 .000	# eelde sd	
#ams sd		Rad Cont 15 km	.136 .000
Rad Cont 15 km	.246 .000	Rad Cont 30 km	.224 .000
Rad Cont 30 km	.301 .000	HRV min 15 km	-.049 .000
HRV min 15 km	-.026 .000	HRV max 15 km	.022 .000
HRV max 15 km	.013 .000	Constant	-6.133 .000
Constant	-5.622 .000	#eelde sn	
#ams sn		Rad Cont 15 km	.128 .027
Rad Cont 15 km	.160 .000	Rad Cont 30 km	.235 .000
Rad Cont 30 km	.259 .000	Constant	-4.285 .000
T cloud average 15 km	.004 .001	# rot wd	
Constant	-5.265 .000	Rad Cont 15 km	.206 .000
#beek wd		Rad Cont 30 km	.324 .000
Rad Cont 30 km	.309 .000	Std dev T cloud average 15 km	.332 .000
HRV min 15 km	-.022 .000	HRV min 15 km	-.025 .000
HRV max 15 km	.008 .001	HRV max 15 km	.012 .000
Std dev T cloud average 15 km	.251 .000	Constant	-4.424 .000
Constant	-4.512 .000	#rot wn	
#beek w		Rad cont sum 15 km	.019 .000
Rad Cont 30 km	.394 .000	Rad Cont 30 km	.356 .000
Std dev T cloud average 15 km	.213 .003	Std dev T cloud average 15 km	.275 .000
Constant	-6.134 .000	Constant	-4.764 .000
#beek sd		#rot sd	
Rad Cont 15 km	.145 .000	Rad Cont 15 km	.272 .000
Rad Cont 30 km	.291 .000	Rad Cont 30 km	.217 .000
HRV min 15 km	-.029 .000	HRV min 15 km	-.022 .000
HRV max 15 km	.013 .000	HRV max 15 km	.014 .000
Constant	-4.906 .000	Constant	-5.688 .000
# beek sn		# rot sn	
Rad Cont 15 km	.143 .027	Rad Cont 15 km	.130 .000
Rad Cont 30 km	.254 .000	Rad Cont 30 km	.277 .000
Constant	-4.596 .000	T cloud average 15 km	.007 .000
		Constant	-6.143 .000

## Appendix 5

### Verificatie AUTOMETAR CB/TCU-detectie (herziene versie)

Han The 2 november 2006

Het algoritme voor Cb/TCu-detectie in AUTOMETAR is gebaseerd op een koppeling van het vóórkomen van Cb's aan de radar signaalsterkte. Dit is equivalent met het gelijkstellen van de kans op het voorkomen van Cb's aan een bepaalde neerslagintensiteit. De drempelwaarden voor de signaalsterkte voor Cb-detectie zijn opgenomen in tabel 1. Op basis van deze drempelwaarden is een beslistabel opgesteld (tabel 2 en 3).

niveau 1	29.0 dBz
niveau 2	33.0 dBz
niveau 3	41.0 dBz

Tabel 1. De drie niveaus op basis waarvan tabel 2 is samengesteld.

dBz-klasse	sfr1 > 0	sfr 2 > 0	sfr 1 en sfr 2: 0 of ongeldig
<b>3</b>	Cb	Cb	Cb
<b>2</b>	Cb	Cb	TCu
<b>1</b>	Cb	Cb	0
<b>0</b>	Cb	Cb	0
<b>afwezig</b>	Cb	Cb	///

Tabel 2. Beslistabel op basis waarvan een enkel radarbeeld wordt geclassificeerd. Signaalsterkte moet voorkomen binnen een straal van 30 km rond de locatie. sfr1 en sfr2 zijn twee niveaus van safirwaarnemingen: <15 km en 15-20 km tot waarneemlocatie.

Beeld 1	Beeld 2			
	Cb	TCu	0	///
<b>Cb</b>	Cb	Cb	Cb	Cb
<b>TCu</b>	Cb	TCu	TCu	TCu
<b>0</b>	Cb	TCu	0	///
<b>///</b>	Cb	TCu	///	///

Tabel 3. Beslistabel op basis van twee opeenvolgende radarbeelden (5 minuten-basis).

Tabel 4 geeft een overzicht van het voorkomen van ontlading gecombineerd met het voorkomen van Cb/TCu. Gezien het geringe aantal waargenomen ontladingen zal dit criterium een geringe rol spelen. Dit wordt bevestigd door de verificatie. Het onderscheid tussen sfr1 en sfr2 (tabel 2) is hypothetisch en niet opgenomen in de verificatie.

	Safir			Alle waarnemingen
	geen Cb/TCu	TCu	Cb	Cb
<b>EHAM</b>	346	7	906	6059
<b>EHRD</b>	154	31	853	5417
<b>EHGG</b>	84	42	532	3306
<b>EHBK</b>	363	21	220	1992

Tabel 4. Voorkomen van ontladingen binnen een straal van 20 km van de locatie bij waargenomen wolkentype (aantal bins van 5 minuten. Cb/TCu voorkomen op basis van METAR, geïnterpoleerd naar 5 minuten. Periode 2005).

Tabel 5 geeft het aantal waargenomen TCu in vergelijking tot het aantal Cb. Op EHAM en EHRD wordt relatief het minste TCu in vergelijking met Cb waargenomen.

	TCu	Cb
<b>EHAM</b>	26	253
<b>EHRD</b>	40	196
<b>EHGG</b>	70	190
<b>EHBK</b>	35	100

Tabel 5. Voorkomen van TCu in vergelijking tot Cb (totaal aantal events in 2005<sup>1</sup>).

<sup>1</sup> Een event is een reeks van één of meerdere opeenvolgende METAR-meldingen.

De verificatie is weergegeven in drie cijfers: de *hit rate*, de *miss rate* en de *false alarm ratio*. Deze zijn uitgaande van de volgende contingentietabel als volgt gedefinieerd:

AUTMETAR	METAR	
	Cb	No Cb
Cb	hit (a)	false alarm (b)
No Cb	miss (c)	correct rejection (d)

$$\text{Hit Rate (HR)} = a/(a+c)$$

$$\text{Miss Rate (MR)} = c/(a+c)$$

$$\text{False Alarm Ratio (FAR)} = b/(a+b)$$

De *hit rate* geeft aan welke fractie van de waargenomen Cb's (METAR) correct zijn gedetecteerd. De *miss rate* is de fractie waarbij dit niet is gebeurd. De som van *hit rate* en *miss rate* is 1.

De *false alarm ratio* is de fractie vermeende Cb-detecties (AUTOMETAR) waarbij geen Cb is waargenomen. Bij de analyse van zeldzame gebeurtenissen wordt *d (correct rejection)* buiten beschouwing gelaten.

Om een volledig beeld te krijgen van de kwaliteit van de methodiek en de toegepaste verfijningen zijn de resultaten voor een aantal configuraties doorgerekend. Deze zijn:

- analyse op basis van een enkel beeld (alleen tijdstippen overeenkomstig METAR), gemiddeld per kwartaal en over het gehele jaar en per locatie;
- invloed van Safir op de kwaliteit van de analyse (jaargemiddelde);
- TCu-detectie.

De resultaten worden gepresenteerd in een contingentietabel met daarnaast de berekende HR, MR en FAR.

## Resultaten

### Algemene instructies voor het interpreteren van de tabellen

De linker tabel geeft een frequentieverdeling overeenkomstig de contingentietabel hierboven weer, voor de signaalniveau's uit tabel 1. Het aantal hits neemt altijd af bij verhoging van het niveau; het aantal missers neemt toe bij verhoging van het niveau; het aantal false alarm neemt af bij verhoging van het niveau. Deze regels gelden altijd. Indien de *false alarm ratio* groter is dan 0,5 betekent dit dat van het aantal detecties er meer false alarms zijn dan hits. De *hit rate* kan worden verhoogd door de drempel te verlagen, maar dit zal ten koste gaan van de *false alarm ratio*. De cijfers voor Schiphol en Rotterdam zijn gedifferentieerd tussen de dagperiode (d.w.z. tussen zonsopkomst en zonsondergang) en 's nachts. Dit onderscheid komt voor Beek en Groningen te vervallen, omdat er 's nachts geen visuele waarnemingen beschikbaar zijn. De verificatie beperkt zich dan ook alleen tot de dagperiode voor zover beschikbaar, dat wil zeggen tussen 06.00 en 23.00 uur lokale tijd.

## EHAM

### Analyse op basis van een enkel beeld overdag (vergelijking tijdstippen METAR)

#### Jan/Feb/Dec 2005

level 1		level 2		level 3	
41	81	22	64	2	9
69	–	88	–	108	–

	lev 1	lev 2	lev 3
miss rate	0.627	0.800	0.982
false alarm ratio	0.664	0.744	0.818
hit rate	0.373	0.200	0.018

#### Maa/Apr/Mei 2005

level 1		level 2		level 3	
34	114	28	93	2	24
49	–	55	–	81	–

	lev 1	lev 2	lev 3
miss rate	0.590	0.663	0.976
false alarm ratio	0.770	0.769	0.923
hit rate	0.410	0.337	0.024

#### Jun/Jul/Aug 2005

level 1		level 2		level 3	
131	120	123	97	67	52
71	–	79	–	135	–

	lev 1	lev 2	lev 3
miss rate	0.351	0.391	0.668
false alarm ratio	0.478	0.441	0.437
hit rate	0.649	0.609	0.332

#### Sep/Okt/Nov 2005

level 1		level 2		level 3	
159	113	148	84	50	28
55	–	66	–	164	–

	lev 1	lev 2	lev 3
miss rate	0.257	0.308	0.766
false alarm ratio	0.415	0.362	0.359
hit rate	0.743	0.692	0.234

#### 2005

level 1		level 2		level 3	
365	428	321	338	121	113
244	–	288	–	488	–

	lev 1	lev 2	lev 3
miss rate	0.401	0.473	0.801
false alarm ratio	0.540	0.513	0.483
hit rate	0.599	0.527	0.199

**Analyse op basis van een enkel beeld 's nachts (vergelijking tijdstippen METAR)**

**Jan/Feb/Dec 2005**

level 1		level 2		level 3	
43	186	33	131	6	20
91	-	101	-	128	-

	lev 1	lev 2	lev 3
miss rate	0.679	0.754	0.955
false alarm ratio	0.812	0.799	0.769
hit rate	0.321	0.246	0.045

**Maa/Apr/Mei 2005**

level 1		level 2		level 3	
33	128	29	100	13	43
29	-	33	-	49	-

	lev 1	lev 2	lev 3
miss rate	0.468	0.532	0.79
false alarm ratio	0.795	0.775	0.768
hit rate	0.532	0.468	0.21

**Jun/Jul/Aug 2005**

level 1		level 2		level 3	
52	82	45	65	25	32
22	-	29	-	49	-

	lev 1	lev 2	lev 3
miss rate	0.297	0.392	0.662
false alarm ratio	0.612	0.591	0.561
hit rate	0.703	0.608	0.338

**Sep/Okt/Nov 2005**

level 1		level 2		level 3	
109	158	97	123	39	47
28	-	40	-	98	-

	lev 1	lev 2	lev 3
miss rate	0.204	0.292	0.715
false alarm ratio	0.592	0.559	0.547
hit rate	0.796	0.708	0.285

**2005**

level 1		level 2		level 3	
237	554	204	419	83	142
170	-	203	-	324	-

	lev 1	lev 2	lev 3
miss rate	0.418	0.499	0.796
false alarm ratio	0.700	0.673	0.631
hit rate	0.582	0.501	0.204

**Totaaloverzicht (etmaalgemiddelde)**

**Totaaloverzicht 2005**

level 1		level 2		level 3	
602	949	525	724	204	213
414	-	491	-	812	-

	lev 1	lev 2	lev 3
miss rate	0.407	0.483	0.799
false alarm ratio	0.612	0.580	0.511
hit rate	0.593	0.517	0.201

**Analyse in combinatie met Safir (etmaalgemiddelde)**

**Totaaloverzicht 2005**

level 1		level 2		level 3	
602	103	525	812	204	313
414	6	491	-	812	-

	lev 1	lev 2	lev 3
miss rate	0.407	0.483	0.799
false alarm ratio	0.632	0.607	0.605
hit rate	0.593	0.517	0.201

**Analyse TCu (etmaalgemiddelde)**

**Totaaloverzicht 2005**

level 1		level 2		level 3	
14	157	13	126	10	
40	0	41	9	44	-
	-		-		

	lev 1	lev 2	lev 3
miss rate	0.741	0.759	0.815
false alarm ratio	0.991	0.990	0.978
hit rate	0.259	0.241	0.185

**EHRD**

**Analyse op basis van een enkel beeld overdag (vergelijking tijdstippen METAR)**

**Jan/Feb/Dec 2005**

level 1		level 2		level 3	
25	77	19	48	2	5
74	-	80	-	97	-

	lev 1	lev 2	lev 3
miss rate	0.747	0.808	0.980
false alarm ratio	0.755	0.716	0.714
hit rate	0.253	0.192	0.02

**Maa/Apr/Mei 2005**

level 1		level 2		level 3	
34	129	27	97	5	24
39	-	46	-	68	-

	lev 1	lev 2	lev 3
miss rate	0.534	0.630	0.932
false alarm ratio	0.791	0.782	0.828
hit rate	0.466	0.370	0.068

**Jun/Jul/Aug 2005**

level 1		level 2		level 3	
103	179	91	152	49	74
43	-	55	-	97	-

	lev 1	lev 2	lev 3
miss rate	0.295	0.377	0.664
false alarm ratio	0.635	0.626	0.602
hit rate	0.705	0.623	0.336

**Sep/Okt/Nov 2005**

level 1		level 2		level 3	
158	111	140	81	46	29
70	-	88	-	182	-

	lev 1	lev 2	lev 3
miss rate	0.307	0.386	0.798
false alarm ratio	0.413	0.367	0.387
hit rate	0.693	0.614	0.202

**2005**

level 1		level 2		level 3	
320	496	277	378	102	132
226	-	269	-	444	-

	lev 1	lev 2	lev 3
miss rate	0.414	0.493	0.813
false alarm ratio	0.608	0.577	0.564
hit rate	0.586	0.507	0.187

**Analyse op basis van een enkel beeld 's nachts (vergelijking tijdstippen METAR)**

**Jan/Feb/Dec 2005**

level 1		level 2		level 3	
42	178	28	123	4	17
76	-	90	-	114	-

	lev 1	lev 2	lev 3
miss rate	0.644	0.763	0.966
false alarm ratio	0.809	0.815	0.810
hit rate	0.356	0.237	0.034

**Maa/Apr/Mei 2005**

level 1		level 2		level 3	
16	127	14	93	5	27
24	-	26	-	35	-

	lev 1	lev 2	lev 3
miss rate	0.600	0.650	0.875
false alarm ratio	0.888	0.869	0.844
hit rate	0.400	0.350	0.125

**Jun/Jul/Aug 2005**

level 1		level 2		level 3	
42	104	37	89	20	43
22	-	27	-	44	-

	lev 1	lev 2	lev 3
miss rate	0.344	0.422	0.688
false alarm ratio	0.712	0.706	0.683
hit rate	0.656	0.578	0.312

**Sep/Okt/Nov 2005**

level 1		level 2		level 3	
114	141	105	115	28	28
28	-	37	-	114	-

	lev 1	lev 2	lev 3
miss rate	0.197	0.261	0.803
false alarm ratio	0.553	0.523	0.500
hit rate	0.803	0.739	0.197

**2005**

level 1		level 2		level 3	
214	550	184	420	57	115
150	-	180	-	307	-

	lev 1	lev 2	lev 3
miss rate	0.412	0.495	0.843
false alarm ratio	0.720	0.695	0.669
hit rate	0.588	0.505	0.157

### Totaaloverzicht (etmaalgemiddelde)

#### Totaaloverzicht 2005

level 1		level 2		level 3	
534	102	461	778	159	227
	9				
376	–	449	–	751	–

	lev 1	lev 2	lev 3
miss rate	0.413	0.493	0.825
false alarm ratio	0.658	0.628	0.588
hit rate	0.587	0.507	0.175

### Analyse in combinatie met Safir (etmaalgemiddelde)

#### Totaaloverzicht 2005

level 1		level 2		level 3	
534	1064	461	816	159	751
376	–	449	–	272	–

	lev 1	lev 2	lev 3
miss rate	0.413	0.493	0.825
false alarm ratio	0.666	0.639	0.631
hit rate	0.587	0.507	0.175

### Analyse TCu (etmaalgemiddelde)

#### Totaaloverzicht 2005

level 1		level 2		level 3	
36	154	31	122	9	397
	4		8		
58	–	63	–	85	–

	lev 1	lev 2	lev 3
miss rate	0.617	0.670	0.904
false alarm ratio	0.977	0.975	0.978
hit rate	0.383	0.330	0.096

**EHGG****Analyse op basis van een enkel beeld (vergelijking tijdstippen METAR overdag)****Jan/Feb/Dec 2005**

level 1		level 2		level 3	
20	90	14	71	2	12
58	–	64	–	76	–

	lev 1	lev 2	lev 3
miss rate	0.744	0.821	0.974
false alarm ratio	0.818	0.835	0.857
hit rate	0.256	0.179	0.026

**Maa/Apr/Mei 2005**

level 1		level 2		level 3	
71	120	60	94	26	28
28	–	39	–	73	–

	lev 1	lev 2	lev 3
miss rate	0.283	0.394	0.737
false alarm ratio	0.628	0.610	0.519
hit rate	0.717	0.606	0.263

**Jun/Jul/Aug 2005**

level 1		level 2		level 3	
114	256	109	217	73	86
31	–	36	–	72	–

	lev 1	lev 2	lev 3
miss rate	0.214	0.248	0.497
false alarm ratio	0.692	0.666	0.541
hit rate	0.786	0.752	0.503

**Sep/Okt/Nov 2005**

level 1		level 2		level 3	
67	112	61	85	21	28
39	–	45	–	85	–

	lev 1	lev 2	lev 3
miss rate	0.368	0.425	0.802
false alarm ratio	0.626	0.582	0.571
hit rate	0.632	0.575	0.198

**2005**

level 1		level 2		level 3	
272	578	244	467	122	154
156	–	184	–	306	–

	lev 1	lev 2	lev 3
miss rate	0.364	0.430	0.715
false alarm ratio	0.680	0.657	0.558
hit rate	0.636	0.570	0.285

**Analyse in combinatie met Safir (overdag)****Totaaloverzicht 2005**

level 1		level 2		level 3	
275	575	247	464	124	152
161	–	189	–	312	–

	lev 1	lev 2	lev 3
miss rate	0.369	0.433	0.716
false alarm ratio	0.676	0.653	0.551
hit rate	0.631	0.567	0.284

**Analyse TCu (overdag)****Totaaloverzicht 2005**

level 1		level 2		level 3	
32	818	27	684	12	264
74	–	79	–	94	–

	lev 1	lev 2	lev 3
miss rate	0.698	0.745	0.887
false alarm ratio	0.962	0.962	0.957
hit rate	0.302	0.255	0.113

**EHBK****Analyse op basis van een enkel beeld (vergelijking tijdstippen METAR overdag)****Jan/Feb/Dec 2005**

level 1		level 2		level 3	
8	33	6	29	2	19
20	–	22	–	26	–

	lev 1	lev 2	lev 3
miss rate	0.714	0.786	0.929
false alarm ratio	0.805	0.829	0.905
hit rate	0.286	0.214	0.071

**Maa/Apr/Mei 2005**

level 1		level 2		level 3	
50	101	45	76	13	26
40	–	45	–	77	–

	lev 1	lev 2	lev 3
miss rate	0.444	0.500	0.856
false alarm ratio	0.669	0.628	0.667
hit rate	0.556	0.500	0.144

**Jun/Jul/Aug 2005**

level 1		level 2		level 3	
114	125	107	101	67	45
41	–	48	–	88	–

	lev 1	lev 2	lev 3
miss rate	0.265	0.310	0.568
false alarm ratio	0.523	0.486	0.402
hit rate	0.735	0.690	0.432

**Sep/Okt/Nov 2005**

level 1		level 2		level 3	
9	90	8	75	5	23
18	–	19	–	22	–

	lev 1	lev 2	lev 3
miss rate	0.667	0.704	0.815
false alarm ratio	0.909	0.904	0.821
hit rate	0.333	0.296	0.185

**2005**

level 1		level 2		level 3	
181	349	166	281	87	113
119	–	134	–	213	–

	lev 1	lev 2	lev 3
miss rate	0.397	0.447	0.710
false alarm ratio	0.658	0.629	0.565
hit rate	0.603	0.553	0.290

**Analyse in combinatie met Safir (overdag)****Totaaloverzicht 2005**

level 1		level 2		level 3	
185	345	168	279	88	112
145	–	162	–	242	–

	lev 1	lev 2	lev 3
miss rate	0.439	0.491	0.733
false alarm ratio	0.651	0.624	0.560
hit rate	0.561	0.509	0.267

**Analyse TCu (overdag)****Totaaloverzicht 2005**

level 1		level 2		level 3	
12	518	11	436	6	194
48	–	49	–	54	–

	lev 1	lev 2	lev 3
miss rate	0.800	0.817	0.900
false alarm ratio	0.977	0.975	0.970
hit rate	0.200	0.183	0.100

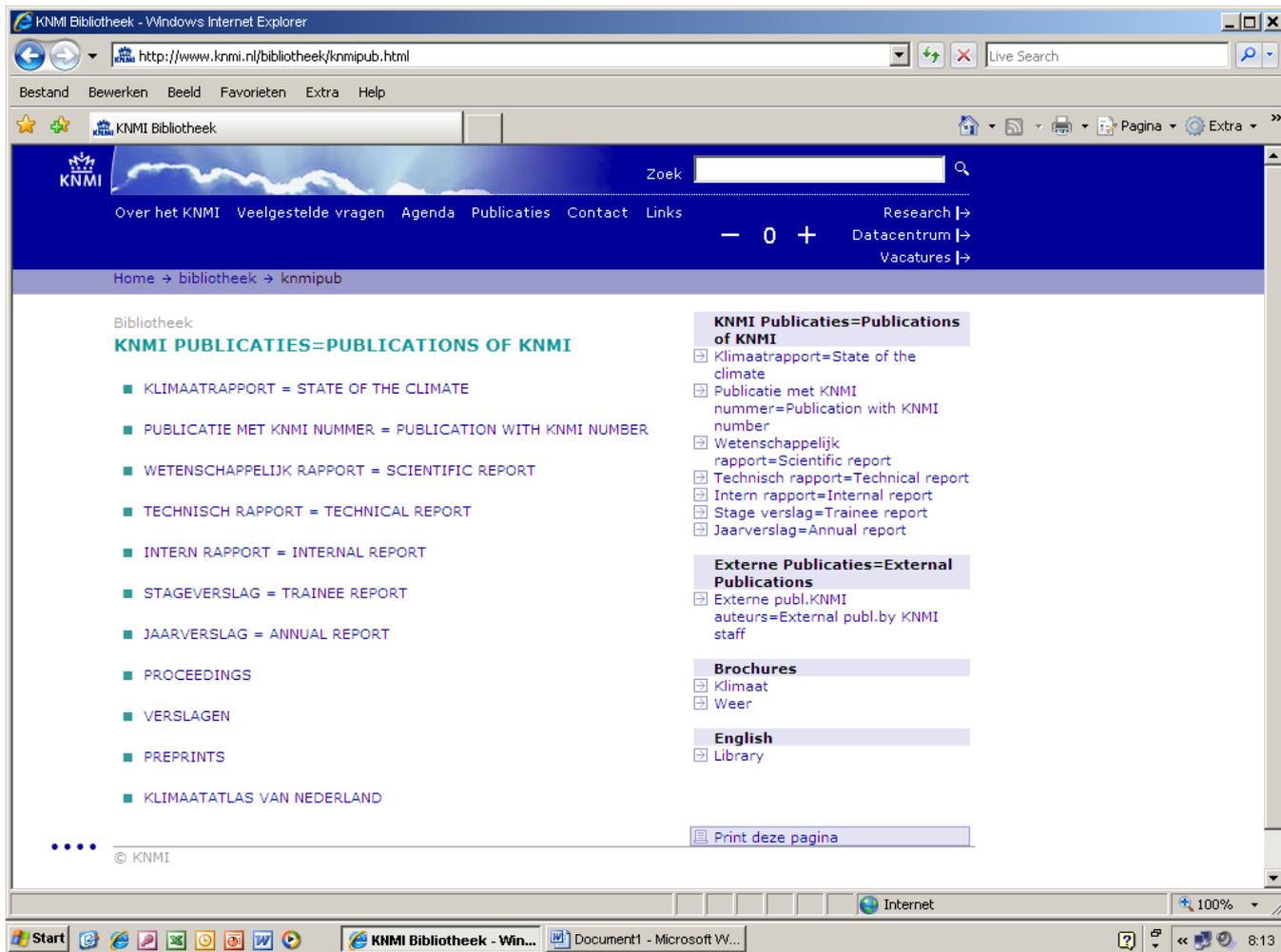
## Discussie en aanbevelingen

- De resultaten zijn vergelijkbaar met die van het Franse onderzoek<sup>2</sup>.
- Het systeem werkt gedurende de zomermaanden en het najaar beter dan in het andere halfjaar. In vrijwel alle situaties blijft de *false alarm ratio* echter aan de hoge kant.
- In de wintermaanden is het systeem niet betrouwbaar en is het gebruik ervan niet aangeraden.
- Bij EHAM en EHRD zijn de *false alarm ratio* overdag structureel gunstiger dan 's nachts. De overige twee indicatoren (*miss rate* en *hit rate*) zijn vergelijkbaar. Dit wijst in ieder geval op een verschillende wijze van waarnemen overdag en 's nachts. Of dit verschil samenhangt met het feit dat er 's nachts visueel minder Cb's worden geregistreerd dan overdag of dat de Cb's 's nacht andere karakteristieken hebben, valt niet a priori te zeggen. Bepalend voor een juiste conclusie is de verhouding tussen het aantal Cb's dat in beide perioden wordt waargenomen. Deze bedraagt voor de visuele waarnemingen 3:2 (overdag vs. 's nachts) en voor de radardetectie ongeveer 1:1. Deze verhoudingen gelden voor beide locaties.
- Voor alle vier de locaties zijn de scores (overdag) vergelijkbaar en kan volstaan worden met een enkele set van drempelwaarden. Het zij opgemerkt dat de winterperiode voor EHBK weinig data bevat hetgeen mogelijkerwijze de kwaliteit van de statistiek heeft kunnen beïnvloeden.
- Onder alle omstandigheden is het gebruik van de hoogste drempelwaarde (41 dBz) te hoog en leidt dit tot een veel te lage detectiegraad (*hit rate*).
- Het gebruik van Safir voor Cb-detectie leidt niet tot een verbetering. Dit hangt samen met de aanwezigheid van valse waarnemingen (zie tabel 4) waardoor de *false alarm ratio* omhoog gaat. Daarentegen wordt het aantal hits niet verhoogd omdat de ondergrens van de radar-echo's waarbij ontladingen worden waargenomen<sup>3</sup> min of meer samen blijkt te vallen met level 1 (tabel 1).
- Het gebruik van drie niveau's in combinatie met de beslistabellen 1 en 2 is niet wezenlijk en kan vervallen. Het is voldoende om één drempelwaarde te kiezen.
- Het is onmogelijk om met behulp van het systeem TCu waar te nemen omdat dit type bewolking niet gecorrigeerd kan worden met een bepaalde neerslagintensiteit (*false alarm ratio* tegen de 100%). Gebruik van radar om TCu te detecteren wordt afgeraden.
- *Wezenlijke verbetering van het systeem is waarschijnlijk alleen te verwachten indien deze methodiek wordt gecombineerd met andere vormen van detectie. Te denken valt aan wolkeoppenecho's (eveneens een radarproduct), satellietinformatie of een stabiliteitsparameter voor de atmosfeer die kan helpen bij de discriminatie convectieve en frontale neerslag. Deze stabiliteitsparameter zou gehaald kunnen worden uit de radiosondewaarnemingen en de korte-termijn modelverwachtingen, en zou met name de false alarm ratio kunnen verbeteren.*

---

<sup>2</sup> Anonymous, ..., Meteo France. Tabel 26 e.v.: 50% miss rate; 70% false alarm ratio

<sup>3</sup> Saskia Noteboom, *Processing, validatie en analyse van bliksemdata uit het SAFIR/FLITS-systeem*, KNMI intern rapport: 2006-01



All titles of KNMI-publications (and a full text PDF for the most recent ones) can be found on

<http://www.knmi.nl/bibliotheek/knmipub.html>

*If you have any questions, please contact us: [bibliotheek@knmi.nl](mailto:bibliotheek@knmi.nl)*

