

Technical report ; TR-300

# **Probabilistic forecasts of winter thunderstorms around Schiphol Airport using Model Output Statistics**

*Aimée Slangen*

De Bilt, 2008

**KNMI technical report = technisch rapport; TR-300**

De Bilt, 2008

PO Box 201

3730 AE De Bilt

Wilhelminalaan 10

De Bilt

The Netherlands

<http://www.knmi.nl>

Telephone +31(0)30-220 69 11

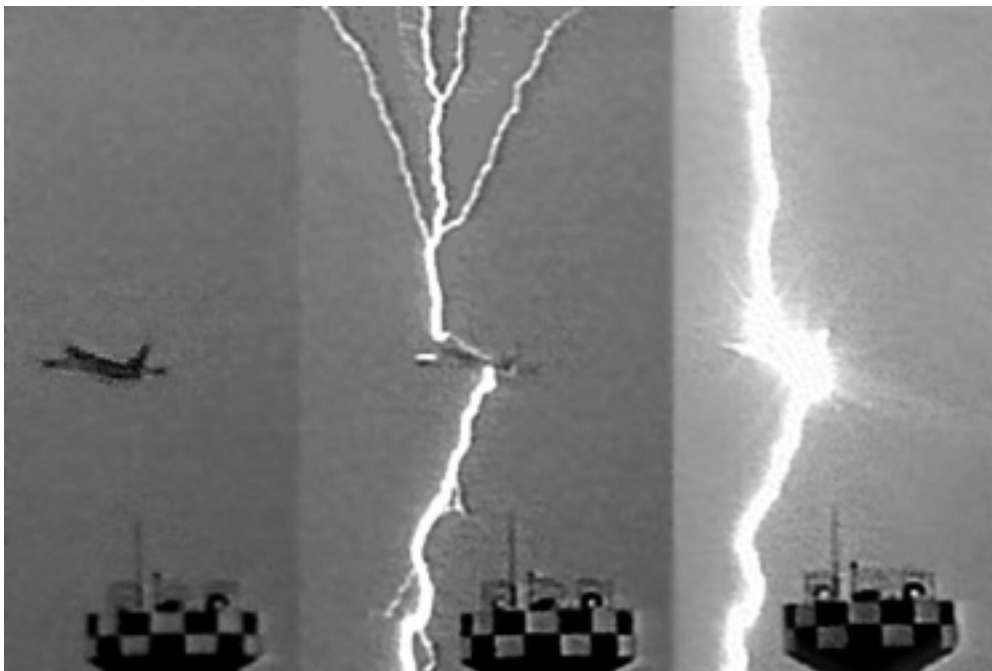
Telefax +31(0)30-221 04 07

Author: Slangen, A.



---

Probabilistic Forecasts of Winter  
Thunderstorms around Schiphol Airport using  
Model Output Statistics



Aimée Slangen

# Probabilistic Forecasts of Winter Thunderstorms around Schiphol Airport using Model Output Statistics

Aimée Slangen (Wageningen UR)

Supervisors:

dr. M.J. Schmeits (Weather Research KNMI)

prof. dr. A.A.M. Holtslag (Meteorology and Air Quality Group Wageningen UR)

This study is part of the Aircraft Induced Lightning project at the Royal Netherlands Meteorological Institute (KNMI), and is also known as WinterKOUW: Winter Kansverwachtingen Onweer t.b.v. Uitgifte Waarschuwingen.

Master Thesis for Meteorology and Air Quality, Wageningen University.

De Bilt, September 2007 - May 2008

# Table of Contents

|   |           |
|---|-----------|
| <b>Abstract</b>   | <b>6</b>  |
| <b>1. Introduction</b>  | <b>7</b>  |
| <b>2. (Synoptical) background of winter thunderstorms and lightning</b>     | <b>10</b> |
| 2.1 Thunderstorms and lightning   | 10        |
| 2.2 Winter thunderstorms  | 12        |
| 2.3 Typical synoptic situations for winter thunderstorms in the Netherlands | 13        |
| 2.4 Aircraft induced lightning (AIL)  | 17        |
| <b>3. Methodology</b>   | <b>19</b> |
| 3.1 Statistical method  | 19        |
| 3.2 Data  | 21        |
| 3.2.1 Predictand  | 21        |
| 3.2.2 Potential predictors  | 24        |
| <b>4. Results</b>   | <b>28</b> |
| 4.1 Selected predictors   | 28        |
| 4.2 Example cases   | 32        |
| 4.2.1 October 30, 2007  | 33        |
| 4.2.2 January 31, 2008  | 34        |
| 4.2.3 November 9, 2007  | 34        |
| 4.2.4 Clutter situations  | 36        |
| 4.3 Testing the equations   | 38        |
| 4.3.1 Brier skill score   | 38        |
| 4.3.2 Verification for 2007-2008  | 43        |
| <b>5. Summary and conclusions</b>   | <b>48</b> |
| <b>References</b>   | <b>51</b> |
| <b>Appendix</b>   | <b>53</b> |
| Appendix A: Forecast equations with coefficients                            | 53        |
| Appendix B: Acronyms  | 55        |

## Abstract

The development and verification of a probabilistic forecast system for winter thunderstorms in the Netherlands is presented in this report. We have used Model Output Statistics (MOS) to develop the probabilistic forecast equations. The MOS system consists of 32 logistic regression equations for two forecast periods (0-6 hours and 6-12 hours), four 90x80 km<sup>2</sup> regions around Schiphol Airport, and four 6-h time periods. For the predictand validated Surveillance et d'Alerte Foudre par Interférométrie Radioélectrique (SAFIR) total lightning data was used. The potential predictors were calculated from postprocessed output of two numerical weather prediction (NWP) models – i.e. the High-Resolution Limited-Area Model (HIRLAM) and the European Centre for Medium-Range Weather Forecasts (ECMWF) model - and from an ensemble of advected lightning and radar data (0-6 h projections only). The predictors that are selected most often are the HIRLAM Boyden Index, the ECMWF 6-h convective precipitation sum, the HIRLAM CAPE and two radar advection predictors. A verification was done using independent data from the winter season 2007-2008, from which it can be concluded that the MOS system is skilful. The forecast system runs at the Royal Netherlands Meteorological Institute (KNMI) on an experimental basis, with the primary objective to warn aircraft pilots for potential aircraft induced lightning (AIL) risk during winter.

# Chapter 1

## Introduction

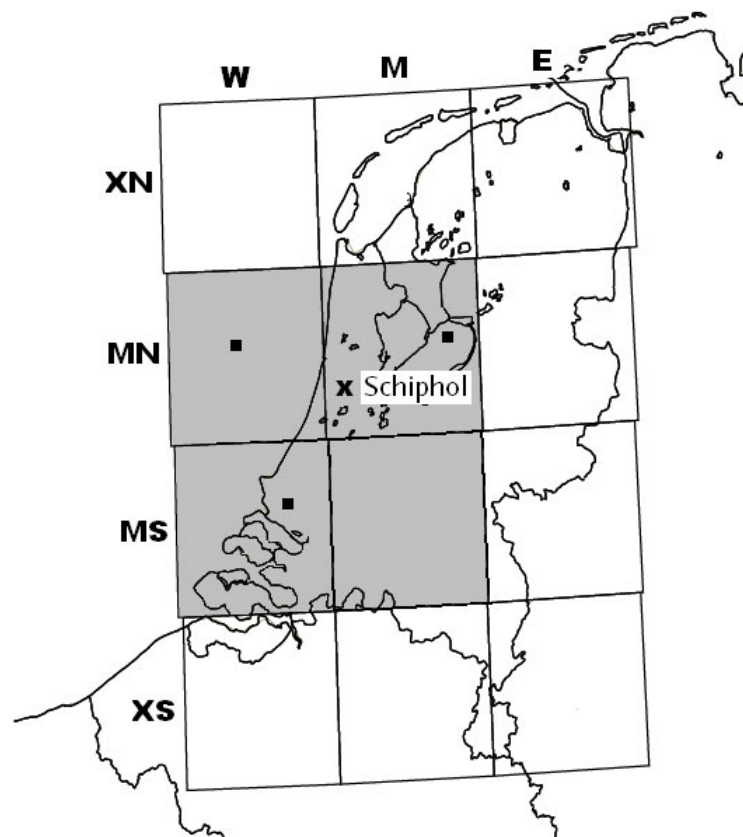
Winter thunderstorms are quite rare in the Netherlands, but despite their rare occurrence they can cause problems to aviation. Pilots have difficulties in estimating the risk of flying through cumulonimbus clouds, because the clouds they see are less deep than in summer and can seem harmless, but they could unexpectedly produce lightning. Aircrafts are mostly not just struck by lightning: 90% of the discharges that contact aircrafts are initiated by the aircraft itself (Rakov and Uman, 2003). This is called aircraft induced lightning. When planes are struck by lightning, this can cause damage to the plane which results in costly reparations or delays. That is why airports want to have a warning system that gives a forecast for the occurrence of lightning. There are several thunderstorm indices, like the Boyden Index (Boyden, 1963) or CAPE (Moncrieff and Miller, 1976), which are generally good thunderstorm predictors when regarded individually. In this study, however, thunderstorm indices will be combined to create a probabilistic forecast system for winter thunderstorms.

Winter thunderstorms have been studied less often than summer thunderstorms. They mainly form when cold polar air flows over the relatively warm North Sea, resulting in an unstable stratification. As a result of this instability, convective clouds can form, in which thunderstorms can occur. Little research has been carried out on this yet in the Netherlands; a larger amount of research about winter thunderstorms has been carried out in Japan. Kitagawa and Michimoto (1994) for instance describe the meteorological conditions and characteristics of Japanese winter thunderstorms. The most important feature in winter is that thunderstorm clouds are less deep than in summer, so the convective activity takes place at lower heights.

In the United States, automated probabilistic forecasts for (severe) thunderstorms from 6 to 72 hours in advance have been provided by the National Weather Service for a long period of time (Hughes, 2001). To our knowledge, in Europe a probabilistic forecast system like this only exists in the Netherlands. It forecasts (severe) thunderstorm probabilities in summer from 0 to 48 hours in advance (Schmeits et al., 2005; Schmeits et al., 2008). In this study a probabilistic forecast system for winter thunderstorms for a short forecast time (0 to 12 hours) will be developed for a few regions in the Netherlands around Schiphol Airport.

The Netherlands were divided into 12 regions of approximately 80x90 km<sup>2</sup> by Schmeits et al. (2005), as can be seen in Figure 1. The regions included in this research are WMN, WMS, MMN and MMS (grey shaded in Figure 1), because these are the regions that surround Schiphol Airport (indicated with **x** in Figure 1). The regions also include the holding beacons where the aircrafts have to wait if not all the runways are available due to bad weather. The

planes have to circle above these beacons until they are granted permission to start the landing process (J. Hemink, personal communication).



**Figure 1: The Netherlands, divided in 12 regions of each 80 km x 90 km (W: west, M: middle, E: east, X: extreme, N: north, S: south). A region name is created by naming it after the column name (West, Middle or East) followed by the row name (eXtreme North, Middle North, Middle South or eXtreme South). The regions included in this research are grey-shaded, the location of Schiphol Airport is indicated with x and the holding beacons are indicated with ■ (adapted from Schmeits et al., 2005).**

To create the forecast system, statistical relationships have been determined between the occurrence of lightning discharges and a set of potential predictors that were calculated from postprocessed output of the numerical weather prediction (NWP) models HIRLAM and ECMWF, and an ensemble of advected lightning and radar data. New radar advection predictors have been added to the predictors already used by Schmeits et al. (2008). Also, the radar advection algorithm has been improved by the addition of a clutter filter that discriminates between rain and false radar echoes.

To determine the occurrence of lightning, a validated lightning data set from the *Surveillance et d'Alerte Foudre par Interférométrie Radioélectrique* (SAFIR) lightning detection network has been used, in which both horizontal and vertical lightning discharges have been included. Model output statistics (Glahn and Lowry, 1972) and logistic regression (Wilks, 2006) with the



so-called forward stepwise method have been used for the selection of the predictors. The run frequency of the system is 4 times per day and the time periods taken are each 6 hours long. An event is called a thunderstorm event if at least 1 lightning discharge is detected in a region in a 6 h period. The time periods are 21-03 UTC, 03-09 UTC, 09-15 UTC and 15-21 UTC. They will be referred to as 00 UTC, 06 UTC, 12 UTC and 18 UTC, respectively. The forecast periods are 0 to 6 hours (referred to as +6 h) and 6 to 12 hours (referred to as +12 h). The equations are valid for the winter half-year only, which runs from October 16<sup>th</sup> to April 15<sup>th</sup>.

The main questions that will be answered in this report are the following:

1. What are typical synoptic situations in which winter thunderstorms occur in the Netherlands?
2. What is the climatological probability of thunderstorms in the winter season?
3. How to create a probabilistic forecast system that predicts winter thunderstorms?
4. Which predictors can be used best to predict winter thunderstorms and what is the physical background of these predictors?
5. Which forecast equations can be made to predict winter thunderstorms and how skilful are these equations?

The structure of this report is as follows. In chapter 2 some theory is provided about (winter) thunderstorms (sections 2.1 and 2.2) and two typical synoptic situations are described in which winter thunderstorms can occur (section 2.3). Section 2.4 gives some background on aircraft induced lightning. In chapter 3 the methodology is treated. This chapter is divided into a part on the statistical method (section 3.1) and a part on the used predictand and potential predictors (section 3.2). In chapter 4 we present the results. In section 4.1 the obtained equations are shown and in section 4.2 some example situations are presented. In section 4.3 the verification results are described. Chapter 5 contains the summary and conclusions of this study.

## Chapter 2

### (Synoptical) background of winter thunderstorms and lightning

Also for research of a statistical nature like this, it is important to know the theory behind the phenomena. Hence this chapter is included. It provides some theoretical background on thunderstorms (section 2.1) and especially winter thunderstorms. Both the physical (section 2.2) and the synoptical (section 2.3) background are treated. Additionally, aircraft induced lightning is treated in section 2.4. For more information about lightning, the reader is referred to Rakov and Uman (2003).

#### 2.1 Thunderstorms and lightning

Single cell thunderstorm development can generally be divided into 3 stages: the cumulus stage, the mature stage and the dissipating stage (Noteboom, 2006). This development is shown in Figure 2.

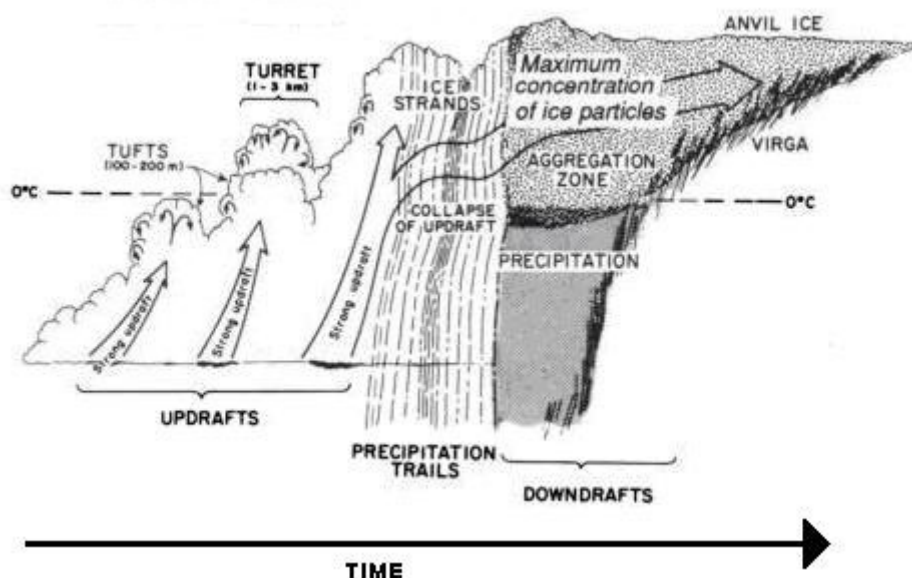


Figure 2: The development of a single cell thunderstorm in time (adapted from Houze, 1993).

Before a thunderstorm can develop, a certain instability has to exist. This can be caused by the temperature difference between a warm (sea) surface and the cold air above. In the cumulus stage this instability causes strong updrafts that make the cloud grow by adding warm, moist air from below to the cloud. The cloud grows until it eventually hits the tropopause. Inside the cloud, there is a large temperature difference: while there are liquid

particles at the bottom of the cloud, supercooled and frozen particles can be found at the top. With continued upward motions, more water condenses and eventually falls down as precipitation. The falling precipitation particles drag the air down and evaporative cooling also causes downdrafts to develop, which means the storm has reached its mature stage. In this stage, updrafts and downdrafts co-exist. The precipitation particles can grow more by collision and coalescence processes, and eventually fall out of the cloud when they are too heavy to be carried upwards again. The downdrafts transport cold air towards the surface, and eventually cut off the updrafts from their source, the warm (sea) surface. This brings the storm to its dissipating stage. In this stage, the strength of the storm is quickly reducing because the updrafts decay, so there is no source for moisture anymore. Eventually, all precipitation falls out and the storm dies out. The life span of such a single cell storm is about half an hour to an hour.

Because of vertical wind shear, the updraft and downdraft may be in different places, in which case the cloud can develop an internal circulation and might exist for a longer time. This is called a supercell thunderstorm. Supercells can be accompanied by severe weather like hail showers and tornadoes. Another possibility for thunderstorms to have a longer lifetime is if one single cell initiates the next single cell, so that multiple cells in different stages exist next to each other: a multicell thunderstorm. Vertical wind shear is needed to form a multicell thunderstorm.

In thunderstorms, lightning is the consequence of a charge-division inside the cloud. Lightning occurs when positively and negatively charged regions connect to transfer the charge. The charge transfer heats the air instantaneously to 30000 K, which causes the pressure to increase 1-2 orders of magnitude and this creates a sound wave (thunder) and a supersonic shockwave. It is still not entirely clear to which process the division of charge can be attributed, and whether it is caused by one mechanism or by several mechanisms working together. Convection is thought to be a part of it, but another theory is the graupel-ice mechanism (Rakov and Uman, 2003). Graupel is the precipitation that forms when supercooled droplets of water condense on a snowflake. Above a critical temperature, which is generally thought to be a value between  $-10^{\circ}\text{C}$  and  $-20^{\circ}\text{C}$ , the falling graupel particles get a positive charge when they collide with ice crystals, while for lower temperatures the charge becomes negative. This results typically in a vertical division of charge over the cloud with a large positive area at the top, a large negative region in the middle and a positive region at the bottom of the cloud. It is still unknown how and why this lower positive region develops (Rakov and Uman, 2003).

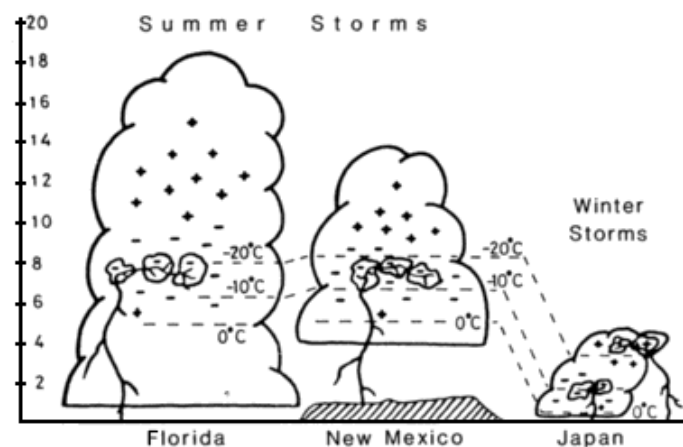
Lightning can occur in 3 ways: cloud-to-ground, cloud-to-cloud and inside the cloud. In this research, no distinction will be made between the three types. The emphasis lies on the

question whether it is possible that lightning will occur and not which kind of lightning or how many strikes are to be expected.

## 2.2 Winter thunderstorms

The main condition for winter thunderstorms to form is that cold (polar) air has to flow over the relatively warm sea surface. This causes the instability that can initiate the formation of convective clouds in which thunderstorms can develop. Winter thunderstorms like this occur frequently in Japan, where a greater amount of research has been carried out on this subject. A few aspects of winter thunderstorms in Japan, taken from Kitagawa and Michimoto (1994) explain the main differences between summer and winter thunderstorms.

1. The duration of winter thunderstorms is generally shorter and has a much lower lightning frequency compared to summer thunderstorms.
2. The vertical wind shear through the troposphere is very strong, comparable to the shear needed for severe summer thunderstorms.
3. The tropopause is much lower in winter, which limits the vertical extent of the convection.
4. The  $-10^{\circ}\text{C}$ -level is closer to the surface in winter (see Figure 3). The altitude of this level determines whether lightning is possible. If the level is above 1.8 km lightning activity is strong, if it is between 1.4 and 1.8 km the activity is weak and below 1.4 km no lightning activity is present.
5. The charge distribution is assumed to be as follows: the upper positive charge can be found above the  $-20^{\circ}\text{C}$ -level on the ice crystals, the main negative charge between  $-10^{\circ}\text{C}$  and  $-20^{\circ}\text{C}$  and the lower positive charge below the  $-10^{\circ}\text{C}$ -level on graupel particles.
6. The updraft velocity is much lower in winter than in summer thunderstorms.



**Figure 3. The charge division and a few temperature levels inside summer and winter thunderclouds (adapted from Geophysics Study Committee, 1986).**

Some of the winter thunderstorm characteristics are shown in Figure 3. The most important difference between summer and winter thunderstorms is clear: winter thunderclouds have a much smaller vertical extent than summer thunderstorms, but the temperature levels associated with the charge division are the same; they occur only at different heights. Summarizing all the different aspects, it can be said that winter thunderstorms have much smaller space and time scales than summer thunderstorms.

### ***2.3 Typical synoptic situations for winter thunderstorms in the Netherlands***

In the previous sections thunderstorms and winter thunderstorms were treated. In this section, we look at two examples of winter thunderstorms in the Netherlands. Since winter thunderstorms develop mainly over a relatively warm sea surface, the wind direction for a winter thunderstorm over the Netherlands can be southwest to northwest. Both the Canal and the North Sea can trigger enough convection to develop a thunderstorm. The examples shown in this section display both situations: example 1 has a north-westerly flow and example 2 has a flow from the southwest. In this research a system has been developed that forecasts these kinds of situations.

The first example, where the wind flows over the North Sea towards the Netherlands, is taken on October 30<sup>th</sup> 2007 and shown in Figure 4. The isobars at the surface are given in Figure 4a, which indicates the wind direction at the surface: northwest. In this figure, a line is drawn over the coast of the Netherlands, indicating a disturbance. The air behind the line is more unstable and is more convective. This convection is characterized by convective cloud cells, visible in Figure 4b and c. In Figure 4b the intensity of the showers is shown, measured by the Dutch weather radars. Figure 4c is a satellite picture, showing the height of the clouds: brighter white means higher clouds. As clouds with higher cloud tops can contain more ice, in which thunderstorms can develop, it is important to know the height of the clouds. In Figure 4d the lightning discharges that were detected by the SAFIR-network are given. The lightning discharges can mainly be found in the province Zuid-Holland. There were not many discharges, as is usual for winter thunderstorms. Most discharges are detected close to the coast, since the thunderstorms decay quickly over land as the warm water source is no longer present. This situation is a typical example of a winter thunderstorm over the Netherlands, with the wind from the northwest, the formation of convective cells and a few lightning discharges. In these conditions, aircraft induced lightning is often observed. In fact, on October 30<sup>th</sup>, 4 planes reported to be struck by lightning in the period 00-06 UTC during the approach to or take-off from Schiphol Airport (J. Hemink, personal communication).

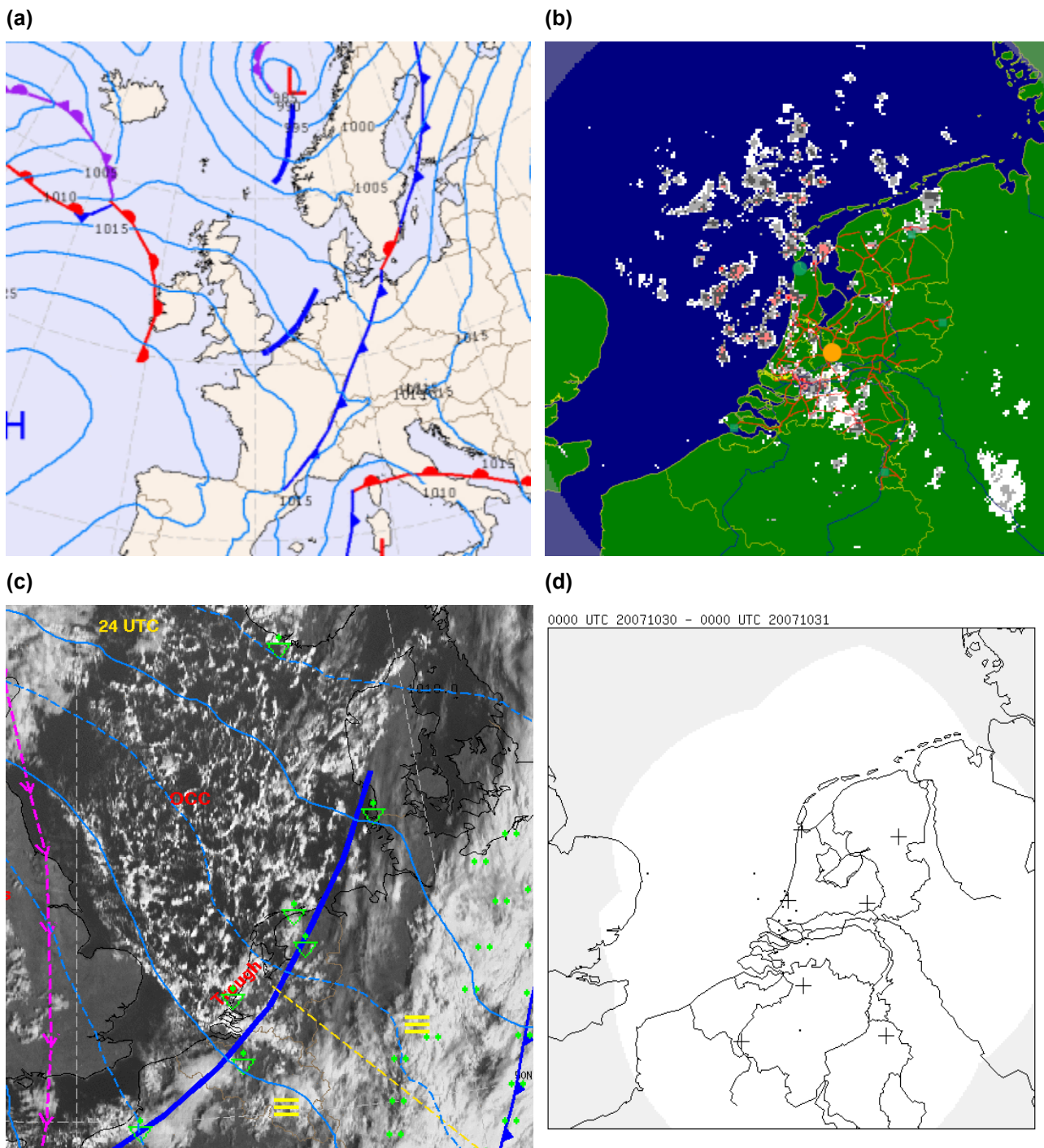


Figure 4: The synoptical situation in the morning of October 30<sup>th</sup>, 2007: (a) 06 UTC surface map analysis (KNMI), (b) radar image at 06 UTC (KNMI), (c) SATREP of the HRVIS-channel at 09 UTC (KNMI), (d) lightning detected by SAFIR/FLITS, 24-h accumulation (KNMI).

Another typical situation is when the wind flows over the Canal and UK and approaches the Dutch coast from the southwest. This situation, taken on January 31<sup>st</sup> 2008, is given in Figure 5. The surface chart is given in Figure 5a. The low pressure system can be found over the North Sea between UK and Norway, which causes the wind to blow from the southwest towards the Netherlands. At 00 UTC a frontal system has just passed the Netherlands and in Figure 5b it can be seen that on this system a thunderstorm occurred. The picture shows the radar precipitation and lightning discharges at 22.25 UTC. The pink pixels indicate a precipitation intensity of 3-10 mm/h, while the dark grey pixels indicate a lower intensity: 1-3 mm/h. Lightning discharges were mainly found over regions with higher precipitation intensity. The cloud tops in Figure 5c are high and cold where the precipitation is the heaviest. The lighter the colour, the colder and higher the cloud is. Figure 5d shows all the discharges detected on January 31<sup>st</sup> 2008. This image gives a good impression of how the thunderstorm moved over the Netherlands: the discharges form a line from southwest to northeast. The situation described here is another typical situation for winter thunderstorms because of the wind direction.

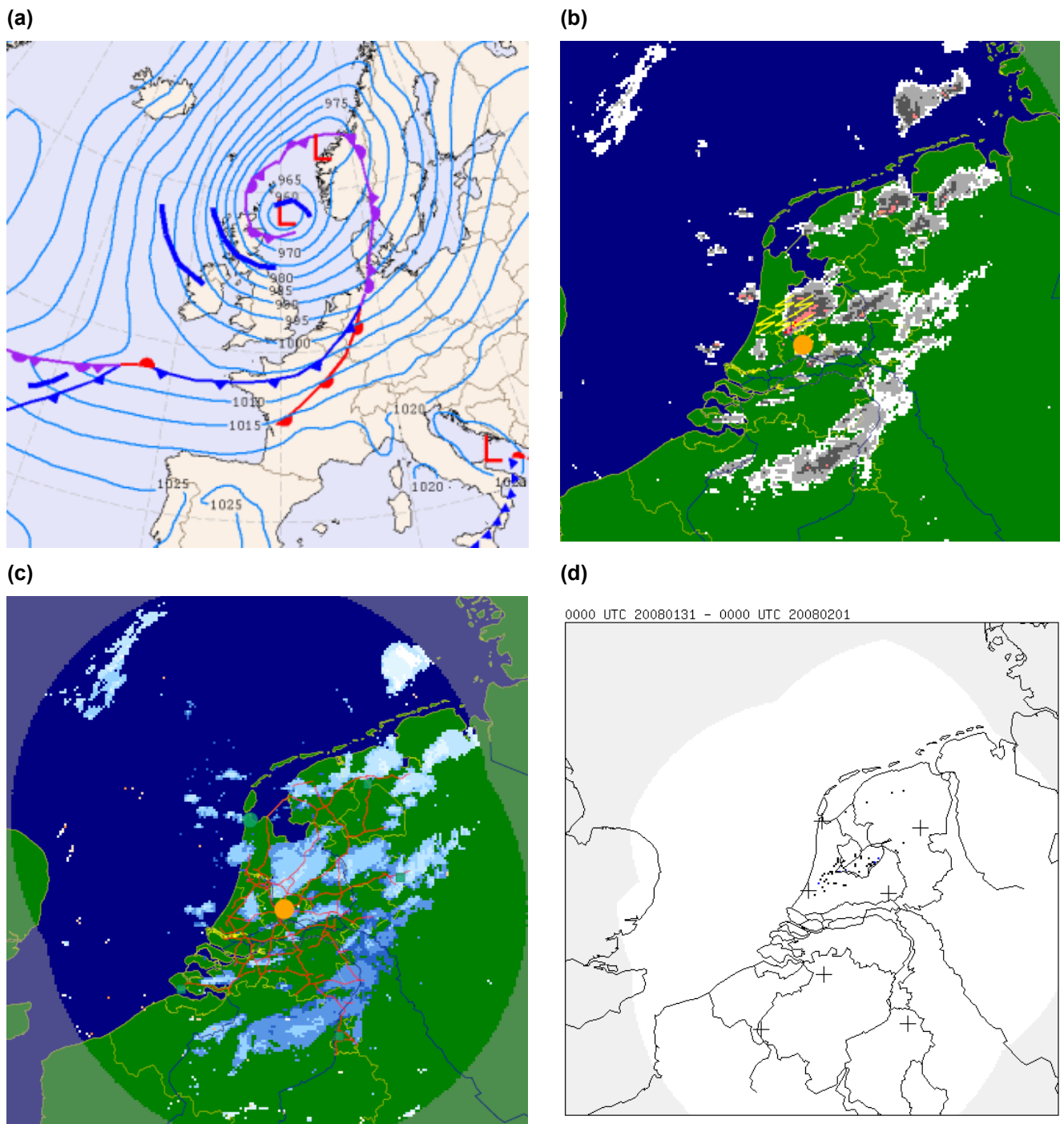
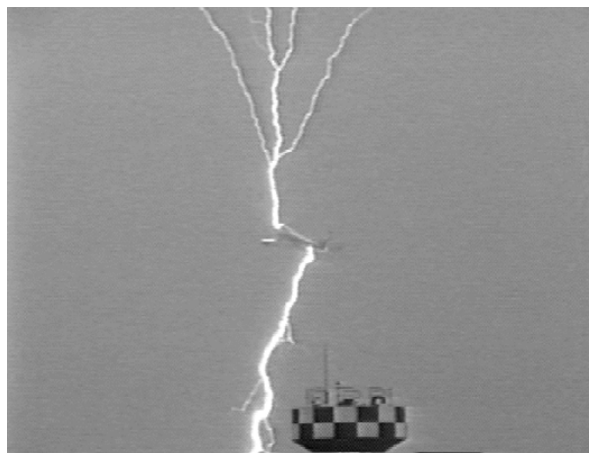


Figure 5: The synoptical situation in the evening of January 31<sup>st</sup>, 2008: (a) 00 UTC surface map analysis of February 1<sup>st</sup> (KNMI), (b) radar image at 22.25 UTC (KNMI), (c) cloud top height at 22.30 UTC (KNMI), (d) lightning detected by SAFIR/FLITS, 24-h accumulation (KNMI).



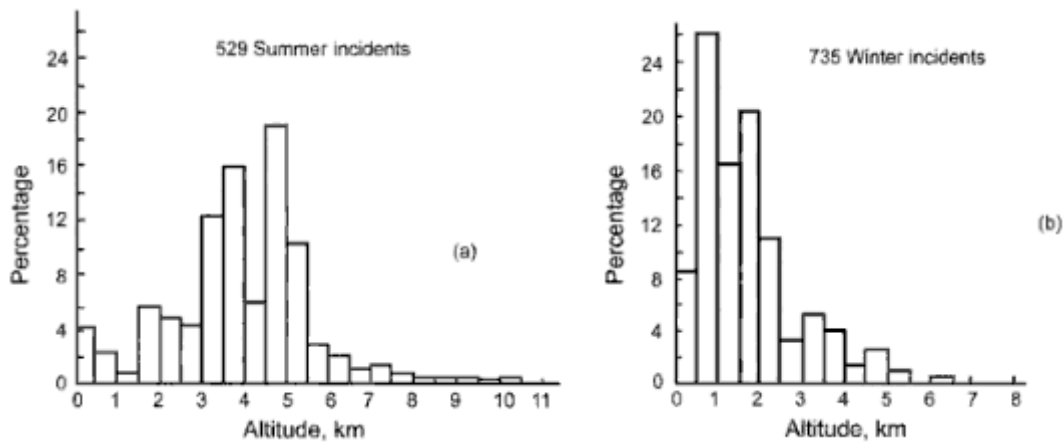
## **2.4 Aircraft induced lightning (AIL)**

In summer, thunderstorms can be dangerous because a lot of discharges may occur in a short time period. In winter this is not the case: discharge rates are much lower than in summer. The problem with winter thunderstorms is that they are inconvenient to aircrafts. The vast majority of time aircrafts are not just struck by lightning: 90% of the discharges that contact aircrafts are initiated by the aircraft itself (Rakov and Uman, 2003). This is called aircraft induced lightning. Evidence for this can be found using pictures, like Figure 6, where the branches above and below the aircraft have different directions and both lead away from the aircraft. The explanation of this two way branching can be found in the bidirectional leader theory (Kasemir, 1950 in Rakov and Uman, 2003) which will not be further described here.



**Figure 6: Lightning strike caused by an airplane during takeoff (Courtesy Z.I. Kawasaki).**

There are no data available on the number of strikes on aircrafts over the Netherlands, but Murooka (1992) investigated at which height and temperature level lightning strikes on planes occurred in Japan. Especially interesting for this research is the change of height per season at which planes were struck by lightning, as can be seen in Figure 7. In summer, the majority of the strikes occurs around 3 – 6 km while in winter most events occur around 0 – 3 km. This means that in winter planes are more often struck during landing and take-off than during the summer. This is exactly what happens around Schiphol airport and is the reason that the research focuses on the regions around Schiphol Airport to make a forecast system for lightning occurrence in winter.



**Figure 7: Aircraft lightning incidents (%) over Japan versus altitude (km), (a) in summer, (b) in winter, for the period 1980-1991 (Murooka, 1992).**

When lightning strikes on an aircraft, different kinds of damage may occur, but fortunately this is mostly non life-threatening. Lightning damage can be divided into direct damage and indirect damage (Uman and Rakov, 2003). Direct damage occurs at the points of entrance and departure of the strike, like holes in the skin of the plane, or a broken radome (i.e. the dome that protects the radar of an aircraft). Indirect damage is caused by the high voltages of the lightning, like magnetic disturbances or damage to electronic systems. These damages are minor, but airplanes have to be taken in for repair after being struck, which costs money and causes delays. The forecast system that has been developed in this research can help reducing the number of aircrafts struck by lightning.

## Chapter 3

### Methodology

This chapter treats the methodology used in this research. It is divided into two main sections: statistical method and data. In the statistical method section the method that was used to develop the forecast equations is explained. The data section is divided into two subsections: predictand and potential predictors.

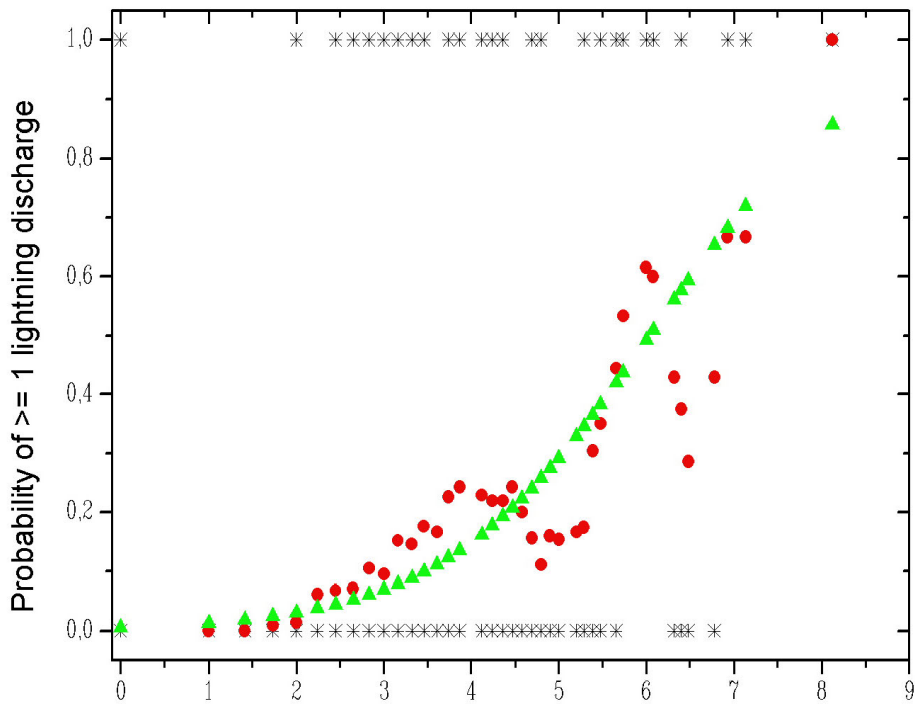
#### 3.1 Statistical method

Model Output Statistics (MOS) (Glahn and Lowry, 1972) is a technique that uses numerical weather prediction (NWP) variables to create a statistical weather forecast. MOS uses an archived record of NWP forecasts as a predictor source to develop forecast equations, contrary to the perfect prog method, which uses historical observations as a predictor source to create the equations. Both methods use observations as the predictand source. The advantage of MOS over perfect prog is that it anticipates to forecast errors of the model due to the use of model forecasts in the development. The disadvantage is that equations will have to be made for each forecast period and that they can only be used in combination with the model that was used for the development. In this research, MOS is applied using logistic regression (Wilks, 2006). With logistic regression the predicted probability of an event can be calculated using the non-linear equation:

$$p_i = \frac{1}{1 + \exp(-b_0 - b_1x_1 - \dots - b_Kx_K)}, \quad (1)$$

where  $p_i$  is the predicted value, also called the predictand (section 3.2),  $b_0 \dots b_K$  are regression coefficients and  $x_1 \dots x_K$  are predictors (section 3.3). The predictors  $x_1, x_2 \dots x_K$  can be selected using the forward stepwise method. In this method, a pool of potential predictors is taken, and for every step the best predictor is chosen to be part of the equation. In SPSS the potential predictors are given a score and the highest score in combination with a sufficiently high significance indicates the best predictor. In this way predictor  $x_1$  is selected first, next predictor  $x_2$ , etc. This is done until the new candidate predictor does not sufficiently change or improve the equation anymore. In this study the maximum number of predictors is three, since the selection of more predictors could result in an overfit of the equation for the dependent data, which could decrease the skill of the equations for independent data. The regression parameters  $b_0 \dots b_K$  are determined using the maximum likelihood method (Wilks, 2006), which assumes that residuals are Bernoulli variables instead of constants.

An example of a selected predictor is given in Figure 8. In this figure, the occurrence of lightning is shown (lightning is indicated by 1 and no lightning by 0) with the corresponding model value for convective precipitation, using crosses. The running conditional mean of the same variable is plotted in red circles, in which the running conditional mean is taken over a bin with the size of one tenth of the total predictor value range. The running conditional mean reveals a relationship between the square root of the precipitation sum and the occurrence of lightning. SPSS can fit a logistic curve to the data (i.e. the crosses in Figure 8). This curve is also plotted (green triangles) and it can be used as the forecast equation.



Square root of convective precipitation sum from 09-15UTC [sqrt(0.1mm)]

**Figure 8: Crosses: the occurrence of lightning in the time period 09-15 UTC for the region WMS as a function of the ECMWF square root of the 6-h convective precipitation sum of that period, using the +27 h forecast of the 12 UTC ECMWF-run. Circles: running conditional mean of the lightning occurrence, where the probability at each point is an averaged value for 1/10<sup>th</sup> of the range of x-axis values. Triangles: logistic curve, fit to the data (i.e. the crosses) determined by SPSS (after Schmeits et al., 2005).**

For the development of the equations a dataset containing lightning data and potential predictors of 3 winter half-years was used: from October 16<sup>th</sup> 2004 until April 15<sup>th</sup> 2007. To select the predictors, the dataset was randomly split into a dependent part  $\left(\frac{2}{3}\right)$  and an independent part  $\left(\frac{1}{3}\right)$ . First the logistic regression was performed on the dependent set, then the equation was tested on the independent part. This has been carried out multiple times, so

the predictors have been selected using different compositions of the same data set, where the random choice should prevent dependency between consecutive days. After the predictors were selected, the coefficients of the equations had to be calculated. To find the most representative coefficients, this has been done using the entire dataset.

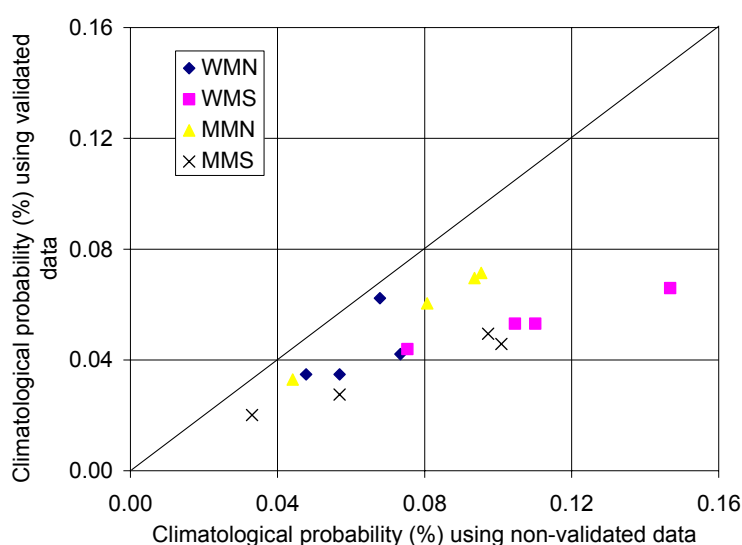
Equations were derived for both forecast times (0-6 h and 6-12 h), all four time periods (00 UTC, 06 UTC, 12 UTC and 18 UTC) and all four regions (WMN, WMS, MMN and MMS). This resulted in a total of 32 forecast equations.

### 3.2 Data

#### 3.2.1 Predictand

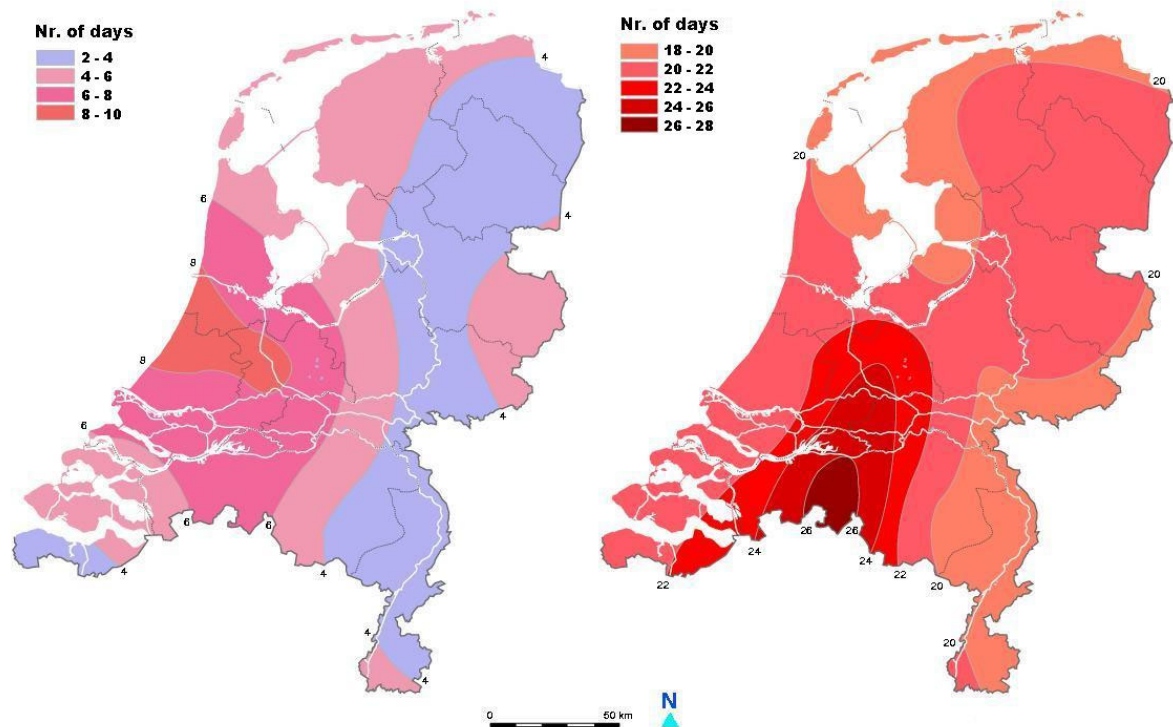
To define the predictand, validated total lightning data were taken from the *Surveillance et d'Alerte Foudre par Interférométrie Radioélectrique* (SAFIR) lightning detection network (Wessels, 1998). The predictand is defined as the probability of a thunderstorm. The predictand is 0 if no lightning is detected in a 6-h time period in a region (see Figure 1, Introduction), and 1 if at least 1 lightning discharge is detected in a 6-h time period in a region.

The validated lightning dataset contains not just the raw data that is detected by SAFIR, it is a dataset in which all observations were compared to radar images. If a lightning discharge was detected by SAFIR and no rain was detected by the radar the discharge was rejected. This means that the validated lightning dataset contains less but probably better lightning data. The difference between the climatological probabilities is given in Figure 9. It can be seen that almost one third and sometimes even half of the data is removed from the raw dataset. In this research the validated lightning dataset has been used.



**Figure 9: Climatological probabilities (%) of validated (y-axis) and non-validated (x-axis) lightning data compared. The diagonal line indicates the 1:1 line.**

To learn more about the predictand, the climatology of thunderstorms in the Netherlands can be used. The climatology for the period 1971-2000 is given in Figure 10. There are a few differences between winter (left) and summer (right). Firstly, the number of thunderstorm days in winter is lower than in summer because deep convection occurs less often in winter. Secondly, the place where the maximum occurs is different, because in winter thunderstorms develop mainly over sea, while in summer they develop mainly over land.



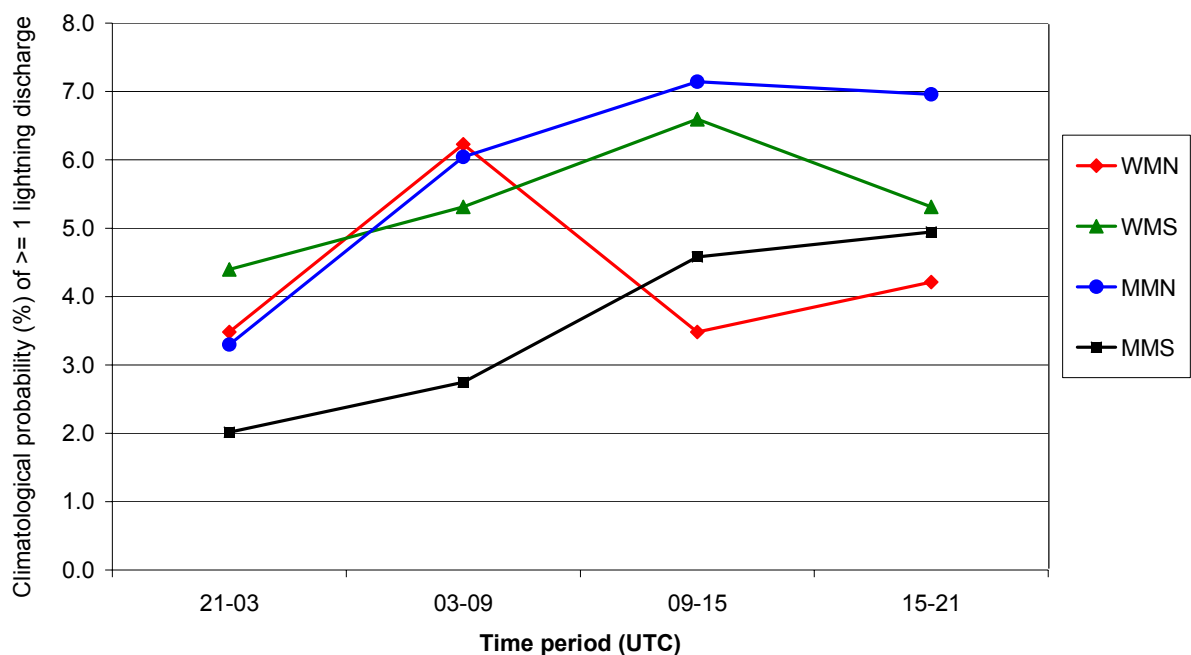
**Figure 10: Climatology for the number of thunderstorm days over the period 1971-2000, for November to April (left) and May to October (right) (adapted from Heijboer and Nellestijn, 2002).**

Using the validated lightning detection data of our own study, another climatology of winter thunderstorms can be made. The advantage is that the lightning occurrence can be split up for the different time periods used in this study and that the data are more recent than the data shown in Figure 10. A disadvantage is that the data contain only three winter half-years, which is a very short period for a climatology. This climatology was created as follows: every time lightning occurs in a region and time period, a value of 1 is given and if no lightning is detected the value is 0. Averaging these values over the total amount of days gives the climatological probability of lightning in a certain region for a certain period of the day. The results are shown in Figure 11. One can see how rare winter thunderstorms are, especially during the night.

MMS is the region with the lowest lightning occurrence, with a minimum of 2% during the night and a maximum of 5% in the evening. MMS is the only land region, which could explain

why the lightning occurrence is low: winter thunderstorms mainly form above sea and they generally have a short lifetime. WMS and MMN both have their maximum during the day (both 7%), while WMN has its maximum in the morning (6%). WMS and MMN are similar in the sense that both are coastal regions, where the thunderstorms are advected from the sea. WMN is a region that contains mainly sea.

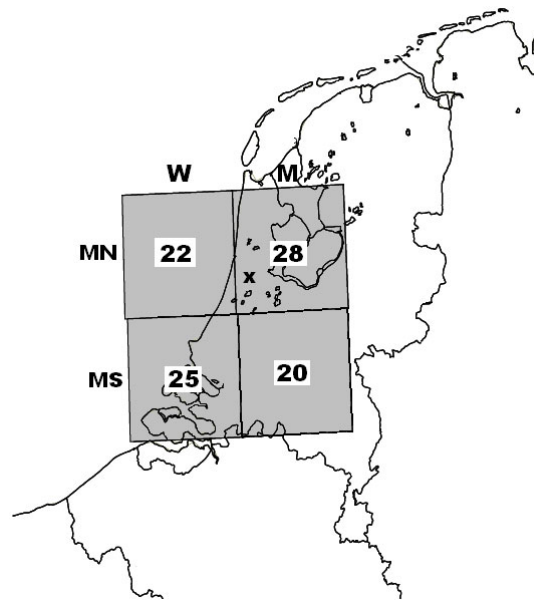
In summer the occurrence of thunderstorms has a strong diurnal cycle with a maximum occurrence from 12 UTC to 21 UTC (Schmeits et al., 2005). Winter thunderstorms also have a diurnal cycle. All regions have their minimum value during the night because of the lack of deep convection. WMN surprisingly has a maximum value in the morning. Possible explanations for this might be the use of the short dataset, so a few occasional events can cause relatively large changes in the climatology, or an earlier occurrence of thunderstorms over sea than over land. The diurnal cycle reaches its maximum in the afternoon and evening. The average occurrence in the afternoon is 5.5% and has a maximum in the regions WMS and MMN, while in the evening the average is 5.4% and MMS attains its maximum value.



**Figure 11: Climatological probability of  $\geq 1$  lightning discharge sorted by region and time period, based on the period October 2004 to April 2007 (winter half-years only).**

In Figure 12 the average number of winter thunderstorm days per year is given for the period used in this research. In this figure, a day is a thunderstorm day when at least 1 lightning discharge was detected by the SAFIR lightning detection network, while in Figure 10 a day is a thunderstorm day when thunder was heard by an observer. Another difference is that in Figure 12 the lightning strikes above sea are taken into account. This results in a strong deviation of thunderstorm days: the numbers in Figure 12 are about three times the values in

Figure 10. The cause of this is that observers have a much smaller observation range than the lightning detection network, which covers the entire country. Also, an exceptional year with a lot of winter thunderstorms has more influence in a shorter climatological period and may cause the climatology to overestimate the number of thunderstorm days in the case of Figure 12.



**Figure 12: Average number of days per year in winter with  $\geq 1$  lightning discharge sorted by region, based on the period October 2004 to April 2007 (winter half-years only).**

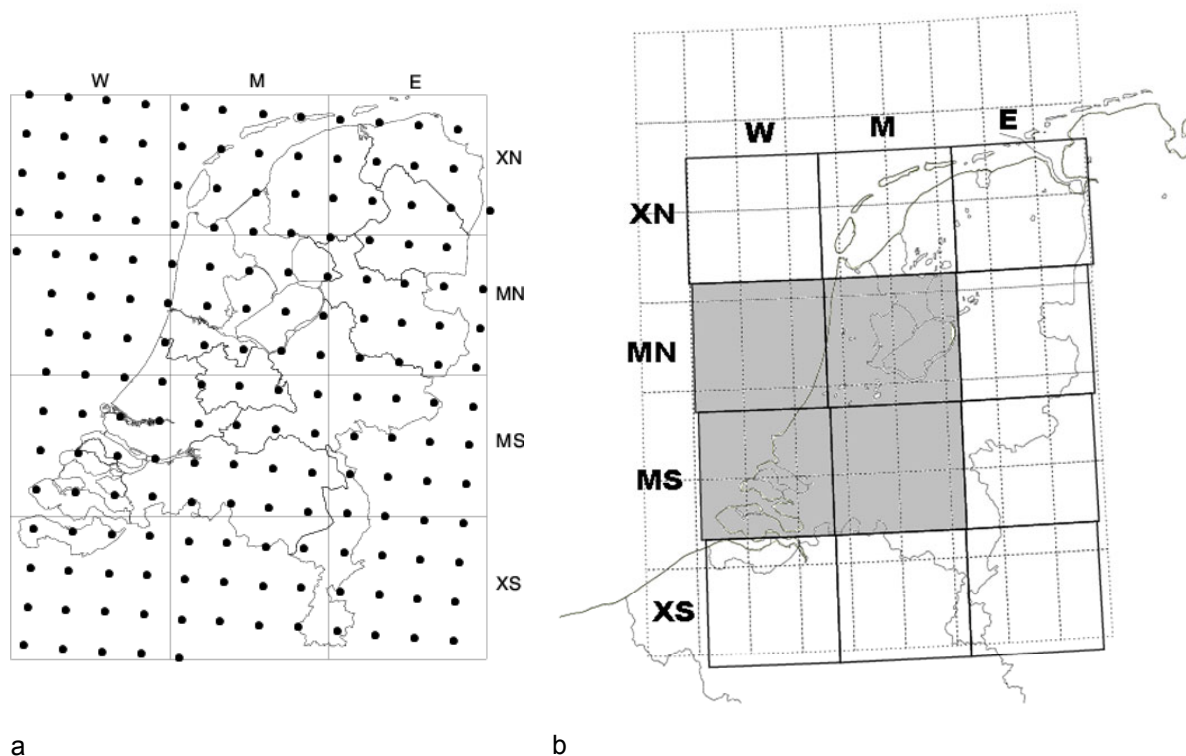
It can be concluded from the previous graph and figures that the occurrence of winter thunderstorms is quite low. Despite this low occurrence a system has been developed that gives probabilistic forecasts for the occurrence of lightning in winter. Unfortunately this can only be done for the 4 regions WMN, WMS, MMN and MMS. Outside these regions the climatology is too low to make skilful forecast equations, as can be deduced from Figure 10. Stations in the chosen regions show a thunderstorm climatology of 6 to 10 days per winter half-year according to Figure 10, while those in the surrounding regions have only 4 to 6 lightning days per winter half-year. To create the forecast equations, not only a predictand is needed but also a set of potential predictors. This will be discussed in the next section.

### 3.2.2 Potential predictors

In Schmeits et al. (2008) several thunderstorm indicators from the NWP models HIRLAM and ECMWF, and an ensemble of advected radar and lightning observations were used as potential predictors. The set included not only the traditional indices, like Jefferson of Boyden, but also other possible indicators of thunderstorms, like wind speed or temperature advection.



In this study, the 0000, 0600, 1200 and 1800 UTC runs from HIRLAM were used in the potential predictor set for the +6 h and +12 h forecasts. The HIRLAM predictor set consists of several traditional thunderstorm indices, like the Boyden index, the Jefferson index, CAPE and the Total Totals Index. For each of the indices the minimum, average and maximum value was calculated using all grid points in each of the regions, and these values are used as potential predictors. This results in 51 potential predictors. HIRLAM had a horizontal resolution of 22 km (shown in Figure 13a) but in October 2006 it changed to 11 km. The vertical resolution changed from 40 to 60 levels.



**Figure 13: Resolution of NWP models over the Netherlands with the regional grid used in this research. a: HIRLAM (22 km resolution), b: ECMWF (0.5° resolution). The grey shaded areas in b indicate the research regions (adapted from Schmeits et al., 2008).**

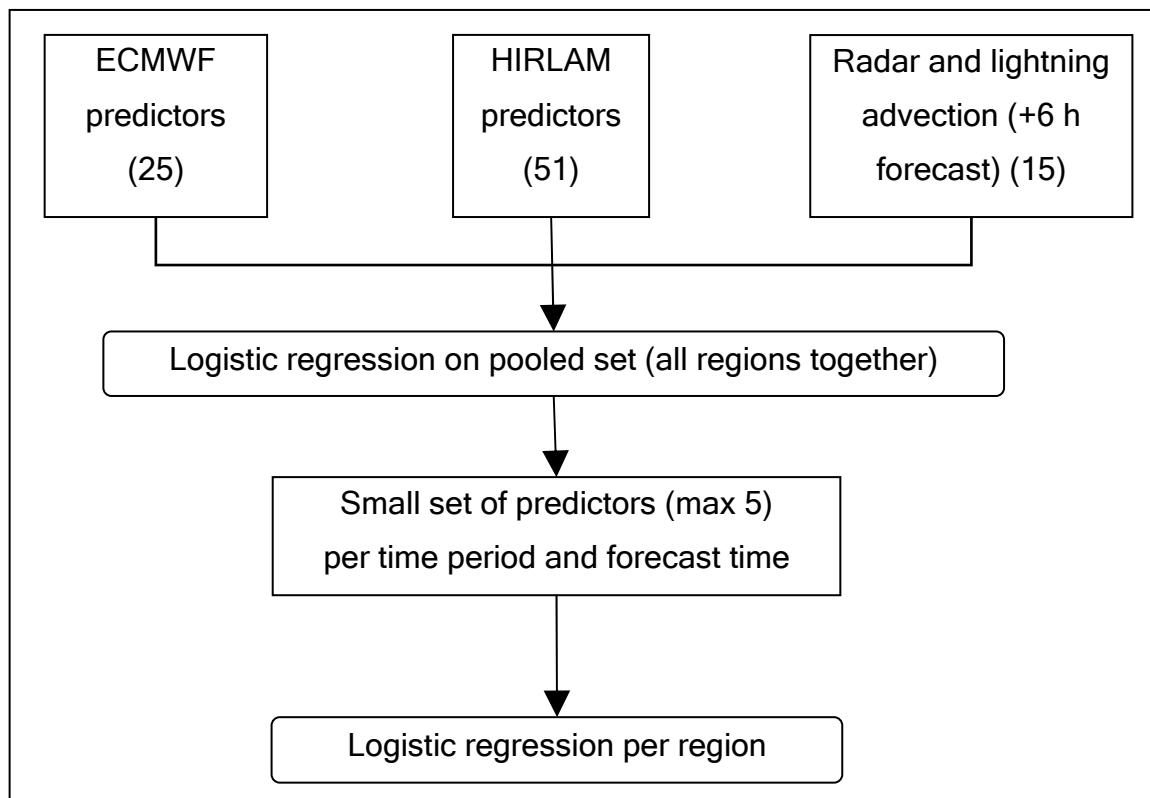
From ECMWF the 12 UTC run was used, with the +12 h, +18h, +24h, +30h and +36 h forecasts. The predictors were used for the +6 h and +12 h forecast period. Various predictors, like vorticity or convective precipitation, are included in this predictor set, which in total consists of 25 potential predictors. ECMWF currently has a horizontal resolution of T799 (0.25°) but until February 2006 the resolution was T511 (0.5°). In the vertical, the number of levels was increased from 60 to 91.

Additional to the NWP model predictors, a few observational predictors were added to the +6 h forecast potential predictor set. As in Schmeits et al. (2008), an 18 member ensemble system of advected radar and lightning data was used. The lightning and radar images of 2.40, 8.40, 14.40 and 20.40 UTC were taken and advected 6 hours ahead, using HIRLAM

700 hPa wind vectors and vectors calculated from previous radar images. The HIRLAM and radar vectors were each varied in length (25% shorter and 25% longer) and direction (+10 and -10 degrees), which resulted in a total ensemble size of 18 members. It was found in Schmeits et al. (2008) that most of the selected predictors were based on statistics of the 9- or 18-member ensemble instead of individual ensemble members. Therefore, these statistics were used to create potential predictors. In this research a few new predictors were developed, such as the square root of the precipitation sum above certain thresholds (3 mm/h, 10 mm/h and 30 mm/h). This predictor is similar to the convective precipitation of the ECMWF model. 15 predictors of the radar and lightning advection ensemble were used as potential predictors.

Besides the development of new predictors, the radar advection program was improved by adding a clutter removal scheme. On clear days and with certain temperature and moisture profiles in the atmosphere, radar images may be contaminated by false radar echoes, called clutter. These false echoes give the impression that the precipitation rates are very high, when actually there is no precipitation at all. Since the radar advection predictors use precipitation rates, clutter may cause predictors to give high values which is undesirable because it decreases the quality of the predictors. Since clutter consists mainly of single pixels without a clear storm structure, the clutter removal's set up is to compare each pixel with its nine surrounding pixels. If a pixel with an intensity above 0 mm/h has less than three neighbours with an intensity above 0 mm/h, its value will be set to 0 mm/h. The number of three was chosen because clutter mostly occurs in single or double pixels and storms often have structures in which more than two pixels are connected. Additionally to this the entire radar image was set to 0 mm/h if one of the pixels gave an intensity of more than 250 mm/h, which is a value that typically is too high for normal precipitation, but can be found in a lot of clutter situations. It appeared that with this clutter filter most clutter was removed and storms were barely reduced in size and intensity, which improved the quality of the radar predictors.

With the complete potential predictor set the process of predictor selection can start. This process is shown in Figure 14. The large set of potential predictors was first tested for each time period and forecast time with all four regions taken together. The best predictors were then used as potential predictors for each region separately. By using this system for the predictor selection it is ensured that neighbouring regions use approximately the same predictors in the equations, which should lead to more coherent forecasts.



**Figure 14: Schematic process of predictor selection.**

Resolution changes can have impact on the behaviour of the predictors, and the dataset used in this research contains resolution changes for both NWP models. This means that the equations that will be used with the current resolution have been developed using a dataset with also lower resolution data. New resolution changes will only require the coefficients to be recalculated, but the implementation of for instance a new convection scheme could alter the statistics of some predictors drastically, which would require totally new equations.

## Chapter 4

### Results

In this chapter the results of this research are presented. In the first section the predictors are given that were selected for the forecast equations. In the second section a few example situations are shown to demonstrate how the forecast system works. These example situations have been taken from the winter season 2007-2008, a season that was not included in the development of the equations. In the third section the skill of the equations is evaluated using the Brier skill score and reliability diagrams are shown for the independent data set of the winter season 2007-2008.

#### **4.1 Selected predictors**

Using forward stepwise selection, predictors have been selected for the forecast equations. Thirteen different predictors have been found for the 32 forecast equations. However, many of these predictors show similarities. In the set of chosen predictors five different 'families' can be distinguished: Boyden, CAPE, RTCP, Advection and Others. These 'families' are shown in Table 1 and discussed here.

- The Boyden Index is a predictor computed from HIRLAM output, that gives information about the instability of the atmosphere. The maximum and average Boyden Index are selected as predictors. The Boyden Index is positively orientated, which means that a higher value for the index gives a higher probability of thunderstorms.
- Convective available potential energy (CAPE) is a measure for the amount of potential energy that is available in unstable air masses. A higher CAPE indicates a higher probability of thunderstorms. CAPE and the level of neutral buoyancy (LNB) are closely related, because LNB is one of the variables used to calculate CAPE. Both are computed from HIRLAM output. Cap100 and cap50 are very similar to CAPE, but CAPE is the lowest level-CAPE, while CAPE100 (CAPE50) always takes the lowest 100 hPa (50 hPa) of the atmosphere and averages the temperature of this layer to acquire the virtual temperature ( $T_v$ ) of the parcel. However, the physics are the same, so they will be counted as one family.
- The radar advection family uses the radar advection ensemble. Tpg03ra takes the member of the radar advection ensemble with the highest value for the square root of the mean 6-h radar precipitation sum, using a minimum intensity threshold of 3 mm/h. It is striking that the lower intensity threshold of 3 mm/h is preferred in the morning, while the 10 mm/h-threshold is preferred in the afternoon. This could be due to an

increase in convection during the day and shows that lightning in the morning is typically accompanied by lower rainfall intensities than in the afternoon.

- RTCP is an ECMWF-based predictor, that takes the square root of the convective precipitation sum, accumulated over different time periods. For instance, if rtcp6ec0 is calculated for 18 UTC verification time, the convective precipitation from 12 UTC to 18 UTC is taken, while rtcp6ec3 takes the 15 UTC to 21 UTC convective precipitation sum. Rtcp6ec is the maximum value of rtcp6ec0 and rtcp6ec3.
- In the family Others, two predictors were placed that did not fit in the previously mentioned families and each of them is chosen only once. Both predictors are calculated from the ECMWF model output, and give respectively the maximum temperature anomaly at 1000 hPa and the relative vorticity-advection at 500 hPa.

**Table 1: Predictors sorted into families, with a formula or an explanation for each predictor. For the predictors, ma(x) takes the maximum value of the predictor in a region, ave takes the average value of the predictor in a region. Here, z is the (geopotential) height [m], T is the temperature [°C], g is the acceleration due to gravity, LNB is the level of neutral buoyancy, LFC is the level of free convection and  $T_v$  is the virtual temperature.**

| Family   | Predictor | Nr. of times used | Formula / explanation  |
|--|-----------|-------------------|--|
| HIRLAM<br>Boyden Index<br>(Boyden, 1963)                                 | boydmax   | 21                | $0.1 \cdot (z_{700} - z_{1000}) - T_{700} - 200$   |
|  | boydave   | 6                 |  |
| HIRLAM<br>CAPE (Moncrieff<br>and Miller, 1976);<br>(Craven et al., 2002) | cap50max  | 1                 | $g \int_{LFC}^{LNB} \frac{T_v(parcel) - T_v(environment)}{T_v(environment)} dz$                                  |
|  | cap100ma  | 6                 |  |
|  | capemax   | 4                 |  |
|  | lnbmax    | 3                 | Level of neutral buoyancy  |
| Radar advection  | tpge03ra  | 5                 | Square root of the mean 6-h radar precipitation sum above 3 or 10 mm/h (maximum of the radar advection ensemble) |
|  | tpge10ra  | 2                 |  |
| ECMWF<br>RTCP  | rtcp3ec   | 4                 | Square root of ECMWF convective precipitation sum over 3 or 6 hours  |
|  | rtcp6ec   | 10                |  |
|  | rtcp6ec3  | 10                |  |
| ECMWF  | mxt1000a  | 1                 | Max. temperature anomaly at 1000 hPa   |
| Others   | rva500    | 1                 | Relative vorticity-advection at 500 hPa  |

In Table 2 the predictors are given separately for each verification time, forecast time and region. The coefficients are not given in this table, they can be found in Appendix A. The predictors are given in order of decreasing importance from left to right. Some predictors were

chosen many times, while others were selected only once or twice. The predictor that is selected most often is Boydmax, the maximum Boyden Index, in 21 of the 32 equations. The second most chosen are rtcp6ec3 and rtcp6ec, the convective precipitation predictors from the ECMWF model. Both can be found in 10 equations.

Different versions of the Boyden Index and the ECMWF convective precipitation predictor can be found in all time periods for both forecast times. Versions of CAPE are chosen in all but one block, the +6 h forecast for 00 UTC. Observational predictors can only be chosen in the +6 h forecast periods; they are chosen in 7 of the 16 equations. Each time period has at least one region with an observational predictor, but for 00 UTC all regions have the observational predictor as the most important predictor. In this time period, the radar advection precipitation predictor has a better skill than the ECMWF convective precipitation predictor.

**Table 2: Selected predictors, sorted by verification time, forecast time and region.**

| Verification time |     | Forecast +6 h |          |          | Forecast +12 h |          |          |
|-------------------|-----|---------------|----------|----------|----------------|----------|----------|
| 00 UTC            | WMN | tpge03ra      | boydmax  |          | capemax        | boydave  |          |
|                   | WMS | tpge03ra      | rtcp6ec3 | boydmax  | capemax        | rtcp3ec  |          |
|                   | MMN | tpge03ra      | boydmax  |          | rtcp3ec        | boydave  |          |
|                   | MMS | tpge03ra      | boydmax  |          | rtcp3ec        | boydave  |          |
| 06 UTC            | WMN | rtcp3ec       | boydmax  | tpge03ra | lnbmax         | rtcp6ec  |          |
|                   | WMS | cap100ma      | rtcp6ec3 |          | rtcp6ec        | boydmax  | mxt1000a |
|                   | MMN | cap100ma      | boydmax  | rtcp6ec3 | lnbmax         | rtcp6ec  | boydmax  |
|                   | MMS | cap100ma      | boydmax  |          | lnbmax         | rtcp6ec  |          |
| 12 UTC            | WMN | rtcp6ec3      | boydmax  |          | cap100ma       | rva500   | boydmax  |
|                   | WMS | rtcp6ec3      | boydmax  |          | rtcp6ec3       | boydmax  |          |
|                   | MMN | capemax       | rtcp6ec3 | boydmax  | cap100ma       | rtcp6ec3 | boydmax  |
|                   | MMS | rtcp6ec3      | tpge10ra | boydmax  | rtcp6ec3       | boydmax  |          |
| 18 UTC            | WMN | rtcp6ec       | boydmax  |          | capemax        | rtcp6ec  |          |
|                   | WMS | rtcp6ec       | boydmax  | tpge10ra | rtcp6ec        | boydave  |          |
|                   | MMN | cap50max      | boydmax  |          | cap100ma       | boydave  |          |
|                   | MMS | rtcp6ec       | boydmax  |          | rtcp6ec        | boydave  |          |

Putting aside the verification time and forecast time, one can count which predictors occur in which regions. It occurs that most predictors are used in all 4 regions: boydave, boydmax, cap100ma, rtcp6ec, rtcp6ec3, rtcp3ec and tpge03ra. There are two predictors that can be

found in only three regions: capemax (never selected in MMS) and lnmax (never selected in WMS). So 7 of the 13 predictors can be found in all regions and 2 of the 13 predictors in three regions. This indicates a certain consistency in the choice of the predictors.

Another sign that the choice of the predictors is not random can be found by looking at the differences in predictor choice between the two forecast times. To clarify this, part of the predictors in Table 2 were put in grey shaded boxes. This is done in the cases for which the same predictor family was selected for the +6 h and the +12 h forecast period. For instance, for verification time 12 UTC and region WMS, RTCP and Boyden are chosen for the +6 h and the +12 h forecast time. There is always at least one predictor from the same family that can be found in both the +6 h and the +12 h forecast period. A few regions even have all their predictors from the same families for both forecast times (06 UTC MMN, 12 UTC MMN and 18 UTC MMN and MMS), but the only region with exactly the same predictors is WMS (12 UTC). These similarities again show that there is a consistency in the predictors that are used for the same region and verification time, even when the forecast times are different.

But why are these predictor families chosen over others? Intuitively, the combination of CAPE, Boyden and RTCP is a good combination to predict lightning occurrence, since (winter) thunderstorms require a certain amount of instability with cold air above and warm air below (which is represented by Boyden), (convective) precipitation is needed since lightning is often accompanied by showers and CAPE says something about how high air parcels can rise and this also gives information on the stability of the atmosphere. So these three predictors together in one equation combine available potential energy, conditional instability, moisture and convection. However, in Table 2 it can be seen that this combination is only chosen 4 times and only in the MMN-region, the region with the highest climatological lightning probability (Figure 11). More often only 2 of these 3 predictors are selected in the equations that are used for the forecasts. The combination of RTCP & Boyden is selected 10 times, CAPE & RTCP 5 times and CAPE & Boyden 4 times. It is striking that Boyden is present in 27 of the 32 equations, although it has never been selected as the first predictor. Apparently Boyden can only be a good predictor for lightning occurrence when it is accompanied by another predictor that adds information about moisture, convection and/or available potential energy. The radar advection predictors, used in the +6 h forecasts, are quite similar to the ECMWF convective precipitation predictors, but replace RTCP only 3 times.

It is striking that no lightning advection predictors have been selected, while in summer this was the most selected predictor in the thunderstorm forecast system for the +6 h forecast (Schmeits et al., 2008). Even with a longer accumulation time for the lightning discharges (20 minutes instead of 5), lightning advection predictors were not good enough to be selected in the winter forecast system. This may be attributed to the shorter life time of winter

thunderstorms: often they form above sea and decay quickly above land. This means that an advection of the lightning activity 6 hours ahead can lead to an overestimation, which decreases the strength of the predictor. Another difference with the summer forecast system is that the Jefferson index (Jefferson, 1963a,b) was selected often in summer and never in winter. This may be explained by looking at the definition of the Jefferson index:

$$\text{Jefferson index} = 1.6 \cdot \theta_{w925} - T_{500} - 11 \quad (2)$$

Here,  $\theta_{w925}$  is the wet-bulb potential temperature at 925 hPa [°C] and  $T_{500}$  the temperature at 500 hPa [°C]. This means that the Jefferson index is calculated using a temperature in the upper troposphere. In summer this is not a problem, because summer thunderstorms easily reach the 500 hPa level. But in winter, when thunderstorms have a much smaller vertical extent, this level is not always reached. This means that the Jefferson Index gives different values, which are apparently not as indicative for a thunderstorm in winter as they are in summer. There are also similarities between the summer and winter system: CAPE, RTCP and Boyden are all frequently used in both systems.

## 4.2 Example cases

In section 2.3 two typical synoptic situations for winter thunderstorms are described. These situations and others will now be used to demonstrate the forecast system developed in this study. All situations are taken from the winter season 2007-2008, which was not included in the development of the forecast system, so these are examples from an independent dataset. Additionally to the examples, a few clutter situations will be shown, where forecasts using the clutter filter will be compared with forecasts without the clutter filter.

The forecast system operates automatically and provides forecasts four times a day. At 03, 09, 15 and 21 UTC, two probabilistic forecasts are calculated: a 0-6 h forecast and a 6-12 h forecast. At 03 UTC for instance, a forecast for the period 03-09 UTC and for 09-15 UTC is given. These forecasts give the probability of thunderstorms in the four different regions of Figure 1. As a reference, the climatology of the different regions and time periods can be taken, as was given in Figure 11. When the forecast probability is higher than the climatological probability, there is an increased probability of lightning. In practice, warnings for thunderstorms will probably only be issued for probabilities that are much higher than climatology. A fixed percentage, above which warnings should be issued, depends on the user, and this can only be determined by doing a user-specified cost-loss analysis (Katz and Murphy, 1997).

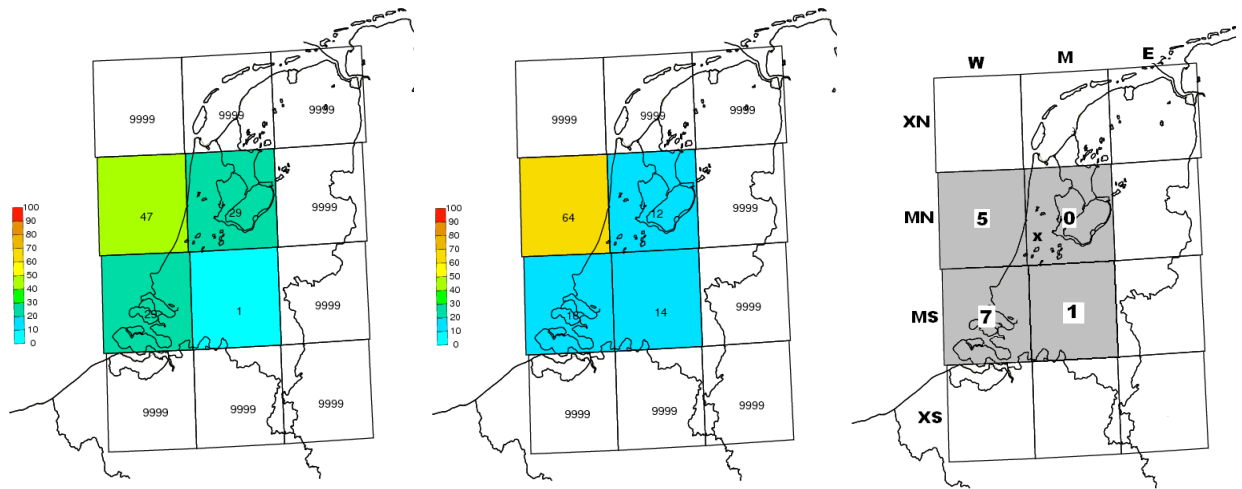
One should keep in mind that the situations shown here are just examples. They are used to illustrate the forecast system, but they give no information on the overall skill of the system. To learn how skilful the system is, a verification is shown in section 4.3 using Brier skill scores and reliability diagrams.



### 4.2.1 October 30, 2007

The first situation that has been described in section 2.3 is from October 30<sup>th</sup> 2007. This is a typical winter thunderstorm situation, with showers coming in from the northwest. For every time period, a 6-12 h forecast and a 0-6 h forecast is given. In this section, the 03-09 UTC time period is discussed.

In Figure 15 the 6-12 h forecast is given. The range of probabilities is from 1% in the lower right region (MMS) to 47% in the upper left region (WMN). The other two regions have a 29% probability of at least 1 lightning discharge in the 03-09 UTC time period.



**Figure 15: 6-12 h probability forecast (%) of  $\geq 1$  lightning discharge for October 30 2007, 03-09 UTC.** **Figure 16: 0-6 h probability forecast (%) of  $\geq 1$  lightning discharge for October 30 2007, 03-09 UTC.** **Figure 17: Number of lightning discharges per region on October 30 2007, 03-09 UTC.**

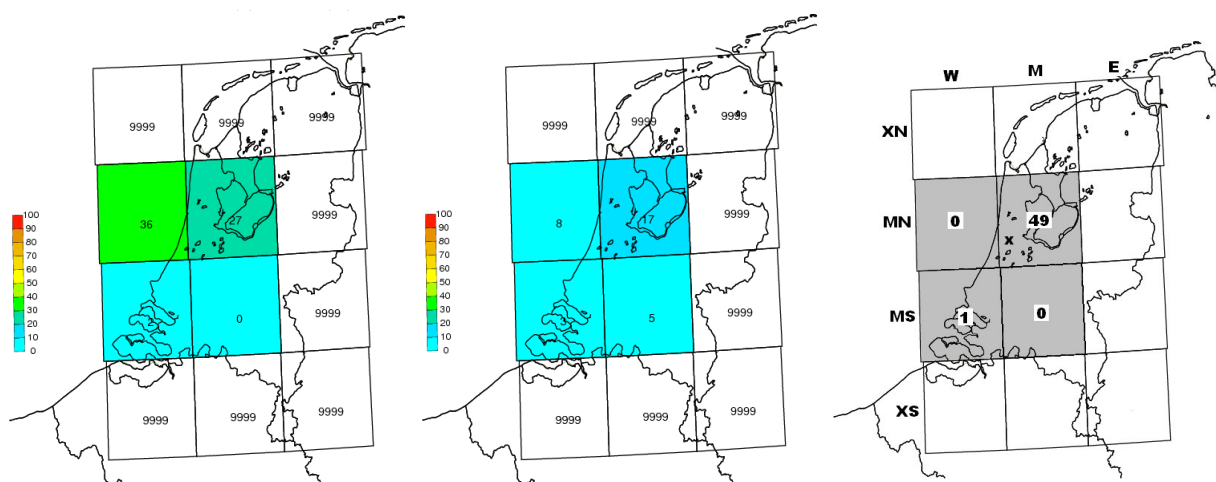
A new forecast is calculated 6 hours later, which is shown in Figure 16. This is the 0-6 h forecast for 03-09 UTC. The probabilities in WMN and MMS increased, while those in WMS and MMN decreased to values still above the climatology, which means there is still an increased probability of lightning occurrence.

The number of lightning discharges that occurred in this time period is given in Figure 17. It can be seen that in all regions, except for MMN, lightning discharges were detected. For 4 of these 13 discharges aircraft induced lightning was reported (J. Hemink, personal communication).

### 4.2.2 January 31, 2008

In chapter 2.3 also the situation in the night of January 31<sup>st</sup> 2008 has been described. This situation will now be used as a second example of the forecast system. The wind direction over the Netherlands was southwest, and the time period discussed is 21-03 UTC, the midnight period.

The 6-12 h forecast for 21-03 UTC is given in Figure 18. Probabilities are above climatology in the north (36% and 27%) and below climatology in the south (2% and 0%). 6 Hours later, in the 0-6 h forecast (Figure 19) the probabilities in the north decreased to 8% and 17%, respectively, while in the south the probabilities increased to 3% and 5%, respectively. Still, the probabilities are above climatology in the north and below in the south. The lightning discharges that occurred in this time period are given in Figure 20, where 49 flashes were found in the upper right region and 1 in the lower left region. In Figure 5d a more precise location can be found, and here one can see that the only lightning discharge in the lower left region is very close to the centre of the 4 regions.



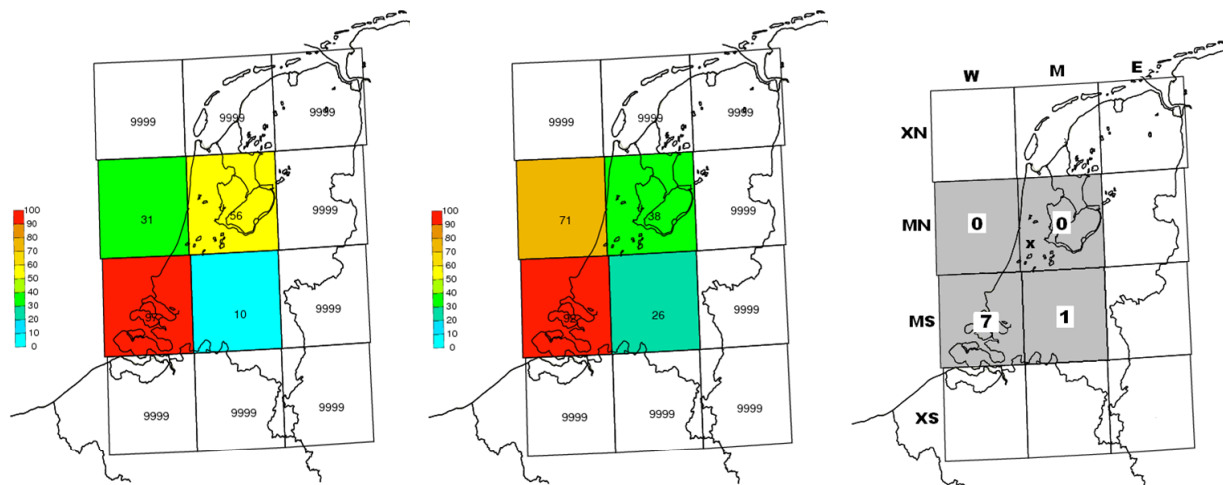
**Figure 18: 6-12 h probability forecast (%) of  $\geq 1$  lightning discharge for January 31 2008, 21-03 UTC.** **Figure 19: 0-6 h probability forecast (%) of  $\geq 1$  lightning discharge for January 31 2008, 21-03 UTC.** **Figure 20: Number of lightning discharges per region on January 31 2008, 21-03 UTC.**

### 4.2.3 November 9, 2007

The 9<sup>th</sup> of November is a day that has not been discussed in section 2.3, but it is an interesting winter thunderstorm day. On this morning, small but intensive showers came towards the Netherlands from the Northwest. Figure 21 shows why this situation is interesting: the system is for region WMS (lower left) already in the 6-12 h forecast for 03-09 UTC very confident that lightning discharges will occur; the forecast gives a probability of 97%. The

other regions give less high probabilities, but still values above the climatological values. In Figure 22 the 0-6 h forecast is given. The probability in WMS decreased slightly to 92%. In WMN and MMS the probabilities increase, while the probability decreases in MMN.

The observed number of lightning discharges is given in Figure 23. Lightning discharges were detected in WMS and MMS, while no discharges were detected in WMN and MMN. This example shows that even for a few lightning discharges the probability can be very high, since it is the probability of at least 1 lightning discharge. MMN still gave a high +6 h forecast probability, but it decreased compared to the +12 h forecast. WMN however showed an increasing probability, while no lightning was detected.



**Figure 21: 6-12 h probability forecast (%) of  $\geq 1$  lightning discharge for November 9 2007, 03-09 UTC.** **Figure 22: 0-6 h probability forecast (%) of  $\geq 1$  lightning discharge for November 9 2007, 03-09 UTC.** **Figure 23: Number of lightning discharges per region on November 9 2007, 03-09 UTC.**

The above situation illustrates that it is important to keep in mind that the forecast is probabilistic. This means that in case of a 71% probability the event should occur 71% of the time. This case turned out to be one of the other 29% of the cases. Although the probabilities are objective, their interpretation can be subjective. If a probability increases from 10% to 30%, a forecaster could decide to issue a warning, while a decrease from 50% to 30% might have the opposite effect where no warnings are sent out. In this way, two probabilities that have the same value can have very different consequences. To prevent this, a fixed percentage could be determined above which warnings should be issued. The height of this percentage can be determined by doing a cost-loss analysis in which, for instance, the costs of flight delays are compared against losses in case of a lightning strike on an airplane. This will result in a percentage that is adjusted to the requirements of the user.

#### 4.2.4 Clutter situations


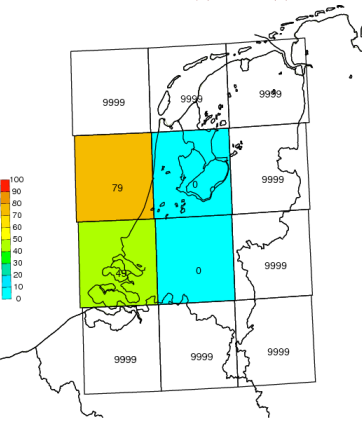
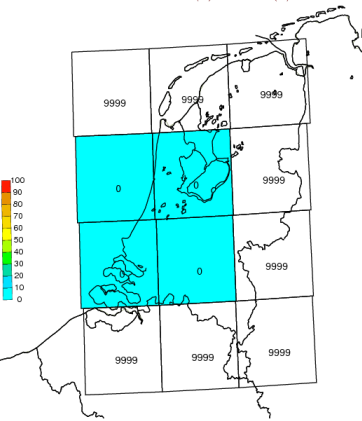
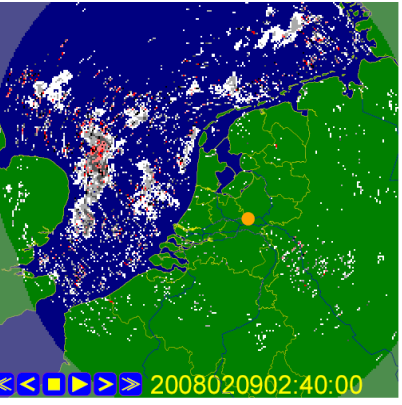
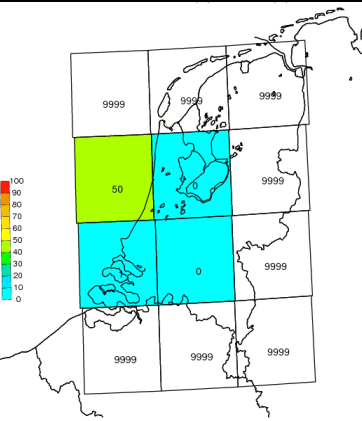
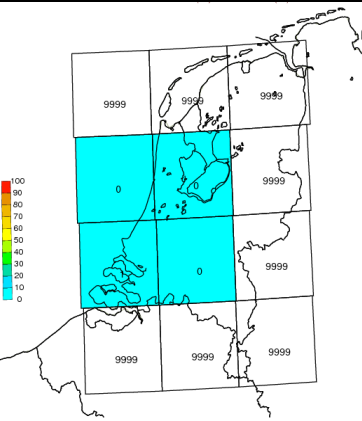
As said before, a clutter filter was added to the radar advection program. This was done because some probabilities of lightning depend on radar advection predictors, so the forecast can give high probabilities in clutter situations where no lightning is possible. This means that the forecast system is less skilful, because the wrong predictand value is used in these situations and this diminishes the quality of the forecast equations. In this section, a few radar images used in the radar advection program will be shown, to give an impression what clutter looks like. The system without clutter filter is compared to the system with clutter filter.

As can be seen in the radar images in Table 3, clutter can be mainly found above sea. It consists of precipitation structures that do not resemble actual showers, with a lot of single pixels with high precipitation intensities. Sometimes the clutter does have the structure of a storm, but when it shows no movement in subsequent radar images, it can be identified as clutter. The situations shown in Table 3 are two forecasts in a period with high pressure above the Netherlands. These days were characterized by cloudless skies and low wind speeds. The forecast system for winter thunderstorms however, gave very high probabilities of lightning discharges. The radar images used in the radar advection program explained why the probabilities were that high: due to clutter, the radar advection predictors had too large values. To prevent this, a clutter filter was added to the radar advection program (section 3.2.2). This filter checks for single pixels and for pixels with unrealistically high intensities (more than 250 mm/h) and removes them from the radar image. In this way, the radar advection should improve and with this the thunderstorm forecast system.

The results can be seen in Table 3. In this table, the old forecast system without clutter filter is compared to the new forecast system with clutter filter. For both situations the radar image used in the radar advection program is shown in the first column. In the second column the old forecast is shown and in the third column the new forecast (all 0-6 h forecasts).

The first situation, February 8, gave very high probabilities for the coastal regions when the old forecast system was used: 79% and 49%. As can be seen in the radar image, the clutter was mainly found over sea, so only the coastal regions were influenced. Below the forecast image, the regions that use radar advection predictors are given. The two coastal regions (WMS and WMN) use them, but also MMN (upper right). Because no clutter is advected to this region, the probability is low. In the new forecast system all probabilities are 0%, due to the absence of a high value for the radar advection predictors, and no indication for a convective situation by any of the other predictors.

**Table 3: Comparison of radar advection with and without clutter filter, for the days February 8, 2008 (21-03 UTC) and February 9 2008 (03-09 UTC). In the first column the radar image (KNMI) used in the advection program is shown, in the second column the 0-6 h forecast without clutter filter (probabilities in %) and in the third column the 0-6 h forecast with clutter filter (probabilities in %). The regions given below the figures are the regions in which radar advection predictors are used in the forecast equations.**

| Radar   | Without clutter filter  | With clutter filter  |
|---|---|--|
|   |  <p style="text-align: center;"><b>WMN &amp; WMS &amp; MMN</b></p> |  <p style="text-align: center;"><b>WMN &amp; WMS &amp; MMN &amp; MMS</b></p> |
|  |  <p style="text-align: center;"><b>WMN &amp; MMS</b></p>          |  <p style="text-align: center;"><b>WMN</b></p>                              |

The second situation, 6 hours later, still shows a radar image with clutter above sea. Some of the pixels are more organised and resemble storm structures, but they remain on the same location in subsequent radar images. In the old system, the upper left (WMN) and the lower right (MMS) region used radar advection predictors, but only WMN had a high probability: 50%. The other regions had very low probabilities. In the new system, only WMN is vulnerable to clutter, but with the use of the clutter filter the probability is set to 0%.

### 4.3 Testing the equations

#### 4.3.1 Brier skill score

In section 4.1 the selected predictors were presented. In this section, an objective verification is done for both forecast periods, by using the Brier skill score (BSS). The BSS is computed using the Brier score (BS), which is the squared error of the probability forecasts, averaged over the number of forecast-event pairs (Wilks, 2006):

$$BS = \frac{1}{n} \sum_{k=1}^n (y_k - o_k)^2, \quad (3)$$

in which  $k$  is the numbering of the  $n$  forecast-event pairs,  $y$  is the forecast and  $o$  the observation with value 1 (if lightning occurred) or 0 (if lightning did not occur). The BS can also be expressed as follows (Wilks, 2006):

$$BS = \text{reliability} - \text{resolution} + \text{uncertainty}$$

in which the reliability should be as small as possible and the resolution as large as possible. Reliability and resolution can be optimized by improving the forecast equations, while uncertainty depends on the variability of the observations only, so it cannot be changed.

The BSS can be calculated using the BS (Wilks, 2006):

$$BSS = 1 - \frac{BS}{BS_{ref}}, \quad (4)$$

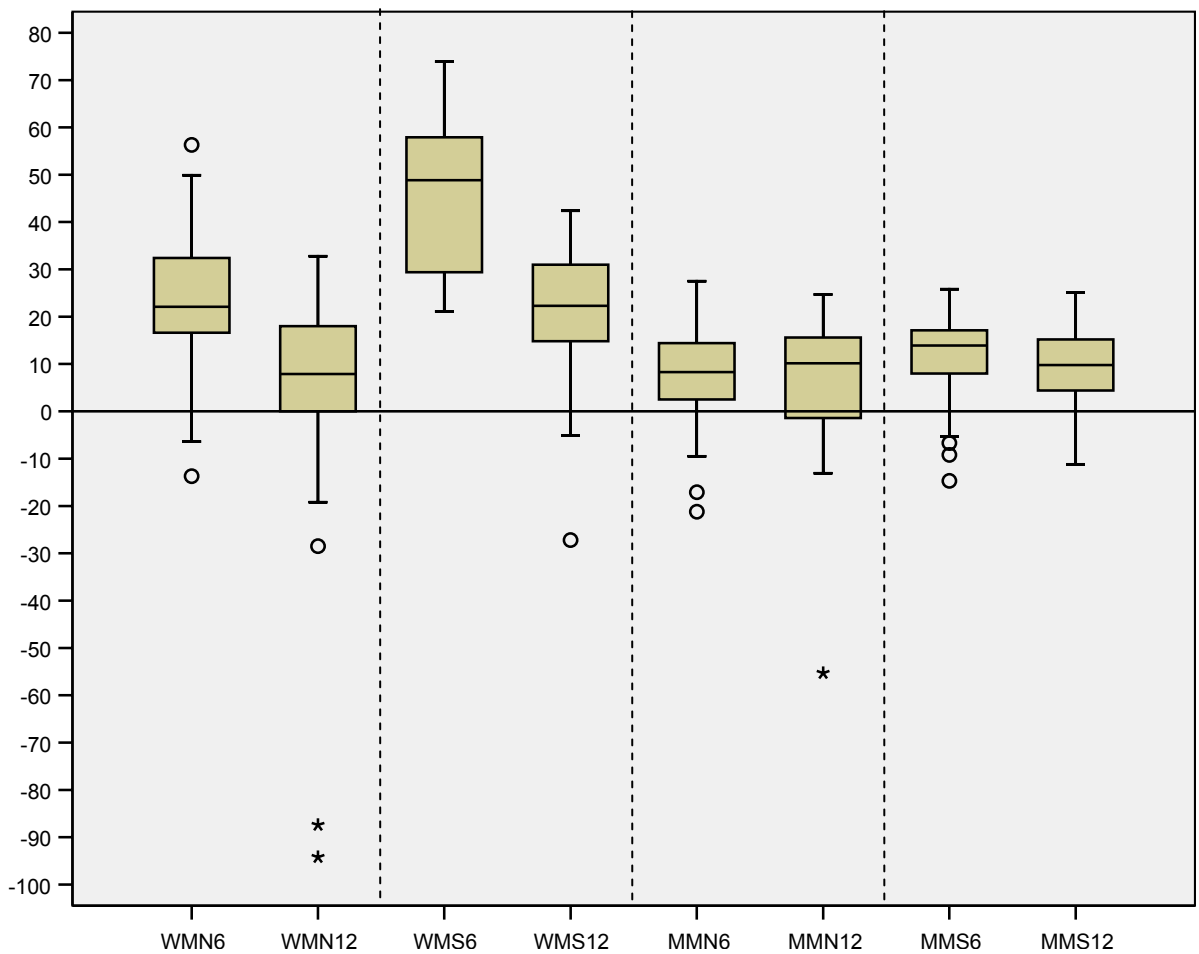
using the climatology to compute  $BS_{ref}$ . The BSS is only a scalar so it is not a definitive measure of the skill, but it gives at least an indication of how well the equation represents the data. The higher the score (with a maximum of 100%) the better. When the BSS is zero, it means that the equation (according to this skill score) is no better or worse than the climatology. When the BSS is negative it means that, according to BSS, the equation is worse than climatology.

In this section the BSS of independent sets is computed. This is done in the following way. First, the data set is randomly divided into a dependent set (2/3 of the data) and an independent set (1/3 of the data). Since the division is random, the amount of cases with lightning occurrence may differ per set, and is not necessarily 2:1. As a lower limit for the number of cases with lightning occurrence in the independent set the amount of 4 cases is taken, to assure a certain level of statistical significance. With the statistical programme SPSS the coefficients are calculated using the previously selected predictors. Because the sets are randomly chosen, this means that the coefficients corresponding to the predictors will differ. This is no problem; it is in fact a good way to test the sensitivity of the dataset to (small) changes in the coefficients of the equations.

The next step is to apply the equation to the independent data set and calculate the BSS. For every region and every time period, the BSS was calculated 30 times for both the dependent set and the independent set. The Brier skill scores of the independent sets are shown in the boxplots below and give an indication whether the chosen predictors can be used to forecast the occurrence of lightning. These Tukey boxplots should be read as follows: the lower limit of the box gives the 0.25 quartile ( $q_{0.25}$ ) and the upper limit gives the 0.75 quartile ( $q_{0.75}$ ) of the data. The line inside the box is the median ( $q_{0.50}$ ). The whiskers indicate the spread of the data; they give the minimum and maximum value, apart from outliers and extremes. Outliers are represented by an open circle (o), when they are 1.5-3 boxlengths away from  $q_{0.25}$  or  $q_{0.75}$ . A value is labelled extreme and marked with an asterisk (\*) when it is more than 3 boxlengths away from the end of the box.

In Figures 24 to 27 the boxplots are shown for the verification times 00 UTC, 06 UTC, 12 UTC and 18 UTC, respectively, for all regions and both forecast times. The figures will be discussed per verification time. In each figure the boxes are ordered by region (in the order WMN, WMS, MMN, MMS), with the +6 h and +12 h forecast time next to each other.

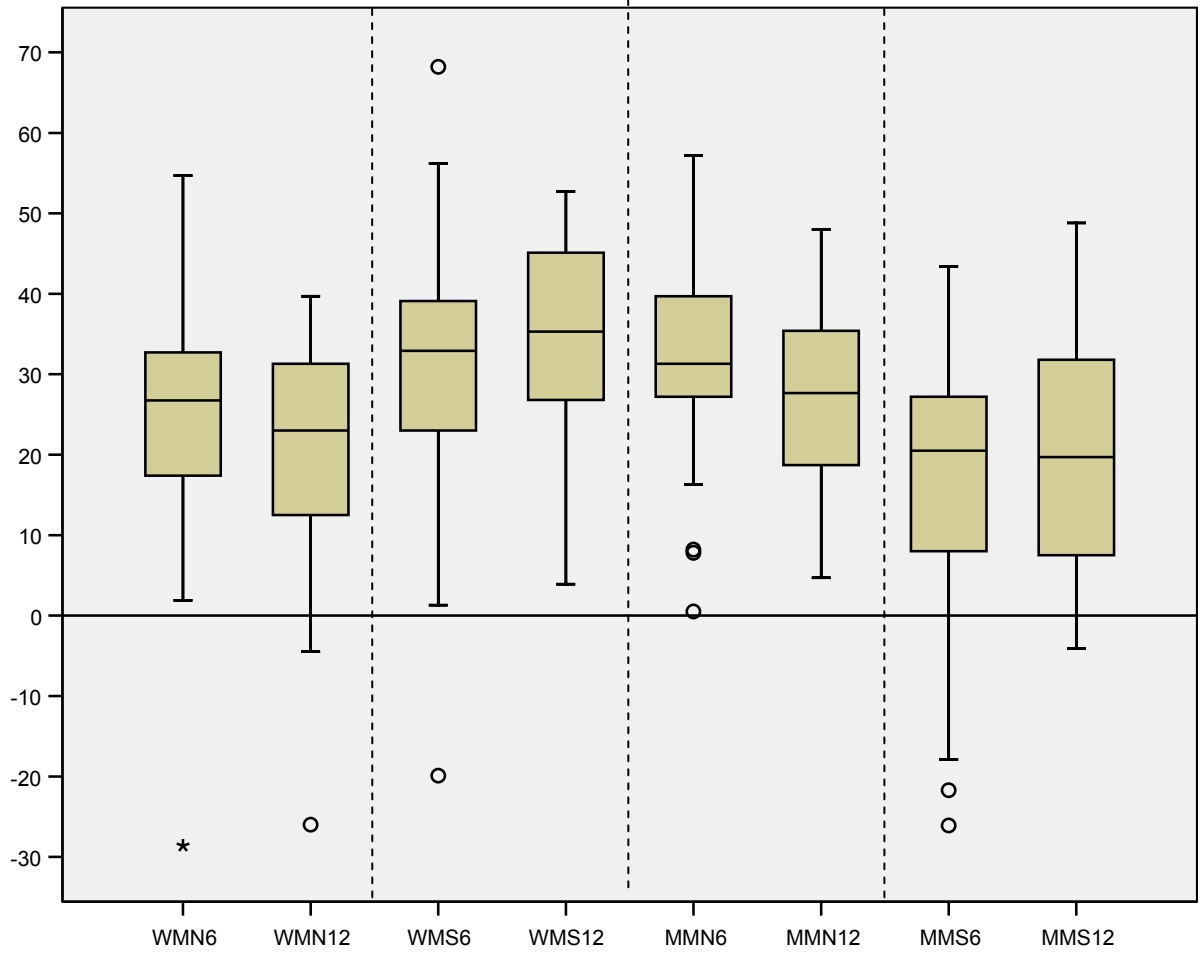
00 UTC is the most difficult time to predict for, as can be deduced from Figures 24 to 27. This is the only time period where some boxes touch (WMN; +12 h) or cross (MMN; +12 h) the 0-skill line (Figure 24). This is also the period where the most extreme value was found: -94% for WMN (+12 h). The BSS is roughly between 0% and 20% with the exception of WMN (+6 h) and WMS (+6 h and +12 h), which have higher values. All regions have an radar advection predictor for the +6 h forecast period. In the case of WMN and WMS, this results in remarkably better skill scores for the +6 h forecasts, but for MMN and MMS this is not the case: they show only a very small improvement for the shorter forecast time.



**Figure 24: Boxplots showing BSS (%) for time period 21-03 UTC for all regions (WMN, WMS, MMN, and MMS) and forecast times (+6 h left, +12 h right).**

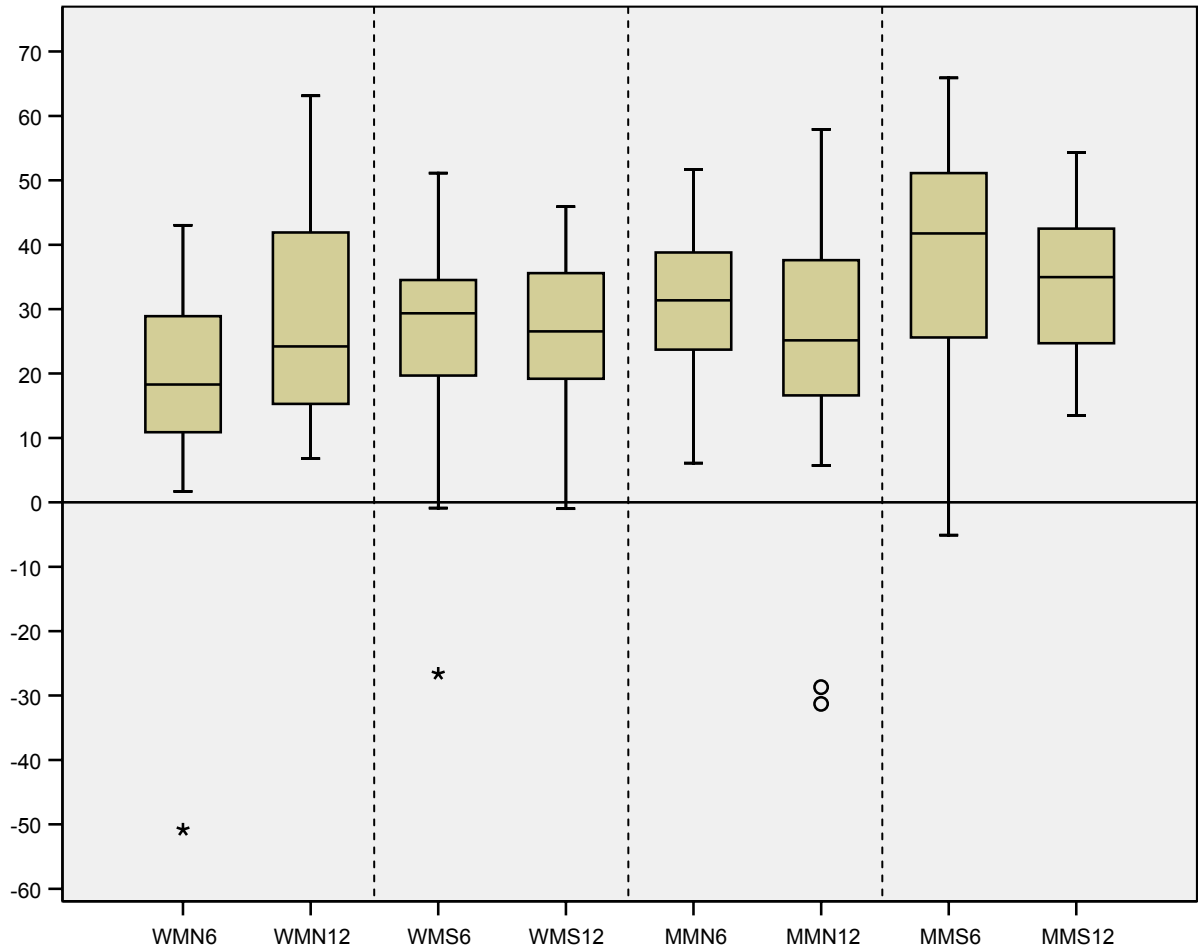


In Figure 25 the scores of time period 03-09 UTC are given. The situation is different from 21-03 UTC period: now all the boxes and most of the whiskers are above the 0-skill line. In most regions the +6 h forecast performs better than or is approximately equal to the +12 h forecast. The BSS is roughly between 20% and 40%, and the lowest outlier has a BSS around -30%.



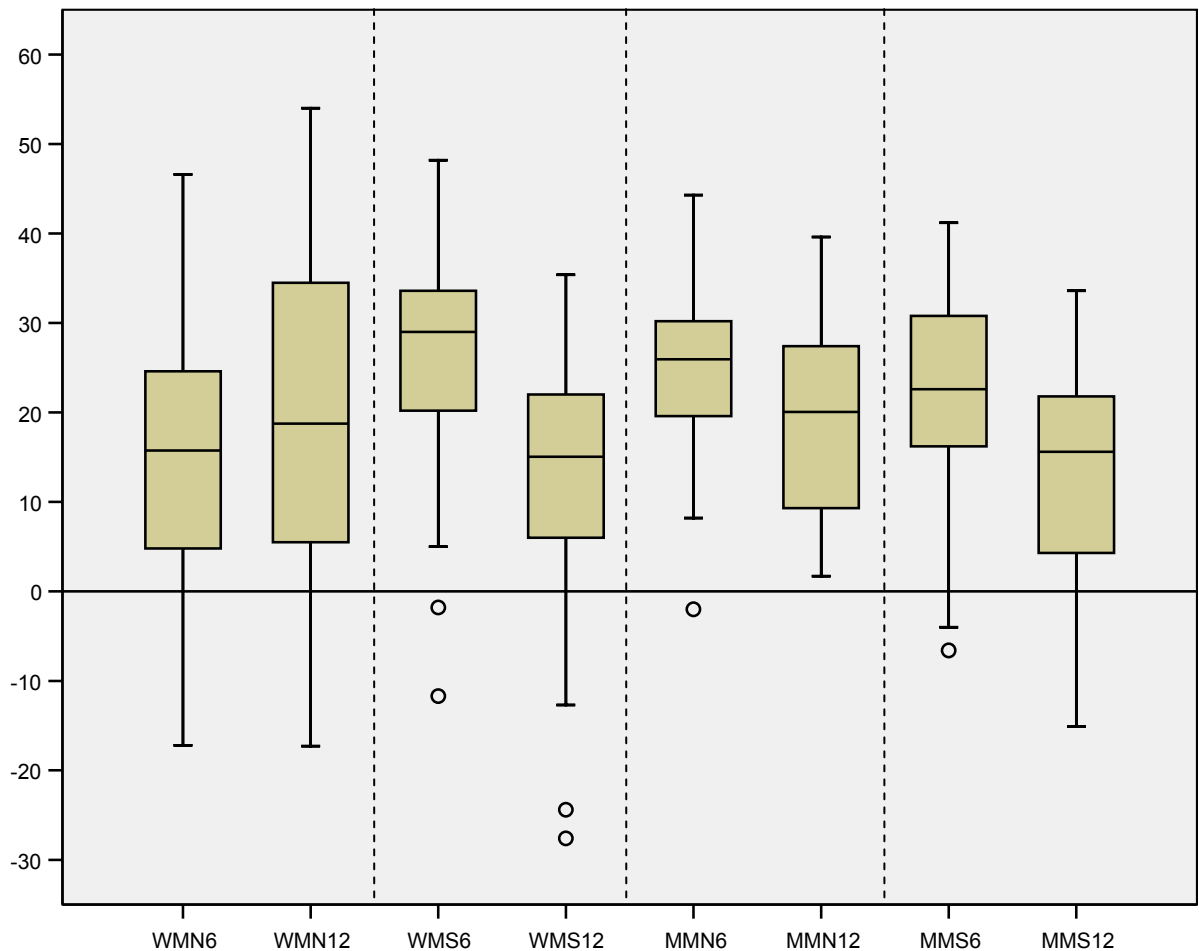
**Figure 25: Boxplots showing BSS (%) for time period 03-09 UTC for all regions (WMN, WMS, MMN, and MMS) and forecast times (+6 left h, +12 h right).**

In Figure 26, the 12 UTC verification time shows even better Brier skill scores than 06 UTC. Again all the boxes are above the no skill line and only three whiskers cross it. Most boxes show BSS-values between 20% and 40%. Only for WMN the +6 h forecast skill is slightly less than the +12 h forecast skill. The other regions show a better skill for the +6 h forecasts, as expected.



**Figure 26: Boxplots showing BSS (%) for time period 09-15 UTC for all regions (WMN, WMS, MMN, and MMS) and forecast times (+6 h left, +12 h right).**

The 18 UTC verification time (Figure 27) shows worse skill scores than 06 UTC and 12 UTC, but better scores than 00 UTC. Most boxes show BSS-values between 5% and 30% BSS. No outliers lower than -30% were found.



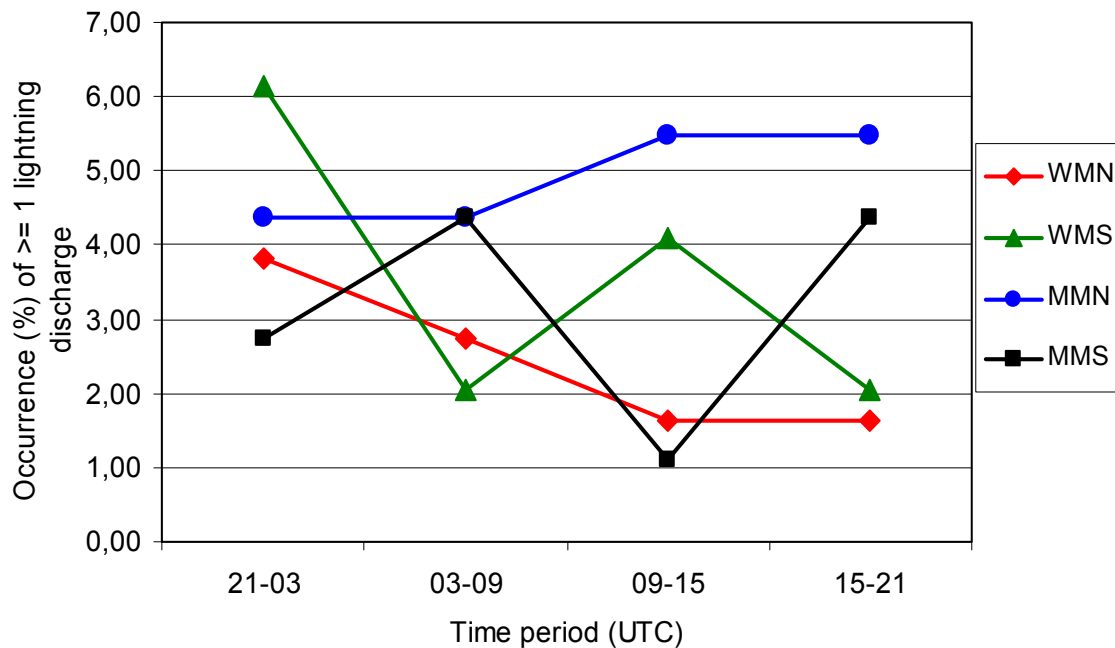
**Figure 27: Boxplots showing BSS (%) for time period 15-21 UTC for all regions (WMN, WMS, MMN, and MMS) and forecast times (+6 h left, +12 h right).**

All the equations obtained with logistic regression, shown above, generally have scores above the 0-skill line. This gives confidence that the forecast system developed here is skilful compared to climatology. The highest skills were found for the midday, the lowest skills for midnight. Another way to examine how skilful the equations are, is by testing them on a completely independent data set. This will be done in the next section.

### 4.3.2 Verification for 2007-2008

In this section the winter thunderstorm forecast system is verified using an independent data set, the winter season of 2007-2008. The question is: how good are the forecast equations in forecasting lightning occurrence, using data that was not used during the development of the system?

The climatology of the winter season of 2007-2008 is given in Figure 28. When this figure is compared to Figure 11, it can be seen that in the verification season slightly less lightning occurred: while the development seasons have a climatology between 2% and 7%, the verification season has a climatology between 1% and 6%. However, in Figure 11 a quite smooth diurnal cycle can be seen, while in Figure 28 the diurnal cycle looks very irregular. This might be caused by the fact that Figure 11 is averaged over three seasons, while Figure 28 contains only one season.



**Figure 28: Occurrence of  $\geq 1$  lightning discharge sorted by region and time period, in the period October 16<sup>th</sup> 2007 to April 15<sup>th</sup> 2008.**

The verification is done using reliability diagrams. The reliability diagram (i.e. Figure 29) is a tool that shows the full joint distribution of observations and forecasts; it shows the strengths and weaknesses of a forecast equation (Wilks, 2006). While the box plots of the Brier skill scores are a good way to give an indication of the skill of forecast equations (see section 4.3.1), the BSS is still only a scalar. Therefore, it cannot give a complete image of the quality of the equations. It gives for instance no impression whether the observed frequency is higher for high probabilities than for lower probabilities.

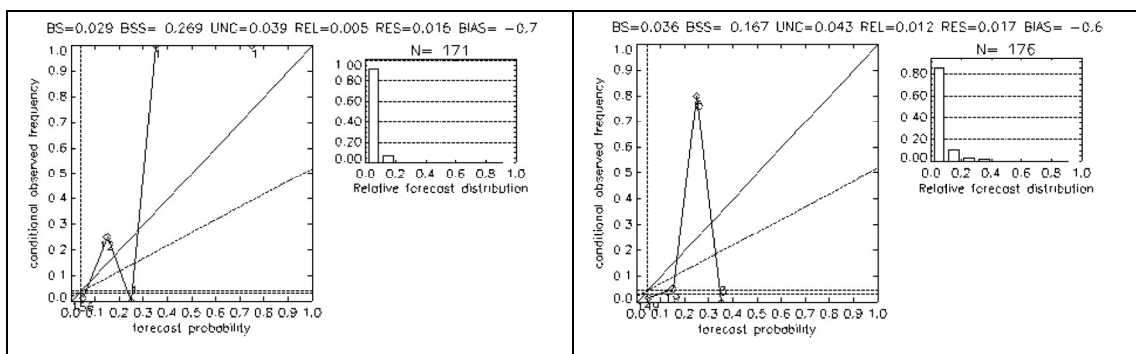
This can be deduced from reliability diagrams. With a reliable forecast equation, a probability of e.g. 60% leads to the occurrence of the event in 60% of the cases and a non-occurrence in 40% of the cases that this probability was forecast. A reliability diagram visualises this. It consists of the observed relative frequency on the y-axis and the forecast probability on the x-axis. A diagonal line, where the observed relative frequency is equal to the forecast probability, is always given as a reference in the graph. A forecast with perfect reliability gives points on this reference line. Furthermore, a horizontal line is drawn to indicate where the

equation has no resolution and gives forecasts with the corresponding observed frequency equal to the climatological probability.

To read and interpret the reliability diagram correctly, a frequency diagram of the forecast probabilities is shown next to the reliability diagram. In the frequency diagram it is indicated how often a certain probability was forecast. In the case of extreme events, for instance winter thunderstorms, a forecast probability of 0 occurs a lot and 1 occurs rarely.

The reliability diagrams of the MMN region (right upper region) for all time periods are shown in the Figures 29 to 32. MMN is chosen because all time periods have a relatively high lightning occurrence in the verification season compared to the other regions, namely more than 4%. For each time period, the two forecast times are shown with the +6 h forecast on the left and the +12 h forecast on the right.

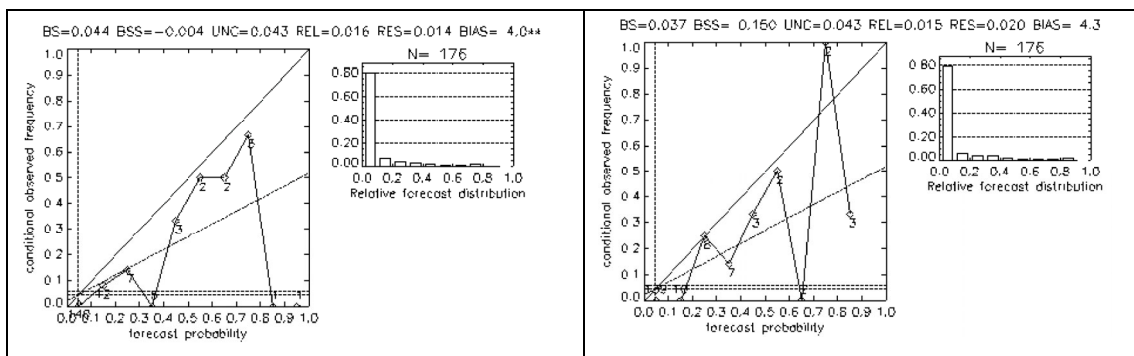
The first time period, 21-03 UTC, is shown in Figure 29. It can be seen in the frequency diagrams, displayed next to the reliability diagrams, that low probabilities were forecast very often, while higher probabilities barely occurred. As can be expected for a rare event, a probability of 0 is often forecast. This will also be the case in all the other reliability diagrams shown in Figures 30 to 32. Almost no probabilities higher than 0.3 were forecast. For a reliable equation, all the points should be on the solid diagonal line in the diagram. In this case, the few points that are available can be found around this line. The BSS is quite high: 26.9% for the +6 h forecast and 16.7% for the +12 h forecast. Both biases are very small, -0.7% and -0.6%, respectively, which means there is no substantial underforecasting.



**Figure 29: Reliability diagrams for region MMN and time period 21-03 UTC. Left: +6 h forecast; right: +12 h forecast.**

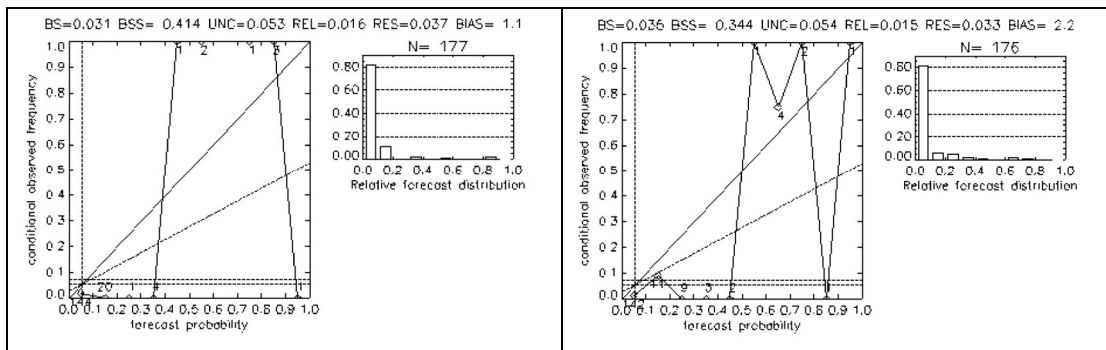
More points are available in the next time period, 03-09 UTC, shown in Figure 30. In this case also some higher forecast probabilities are found. Both reliability diagrams show the entire range of the forecast probabilities, because both high and low probabilities were forecast. In Figure 30 both forecast periods show some overforecasting: most forecast probabilities are higher than the actual observed frequencies. This is also indicated by the positive bias, which is 4.0% for the +6 h forecast and 4.3% for the +12 h forecast. The BSS is quite low for the left

figure, -0.4% and 15% for the right figure. Especially the negative score for the +6 h forecast is remarkable because the positive-sloping reliability curve clearly indicates that the forecast equation has more forecast value than climatological forecasts (Mason, 2004). This case shows one of the flaws of the BSS. Sometimes the BSS is near zero, while the reliability diagram might indicate forecast value (i.e. Figure 30, left). In other cases, the BSS is high, but the reliability diagram might not look that good because not many different forecast probabilities are issued (i.e. Figure 29, left). A BSS is only one number, but the reliability diagram shows the total distribution of the observations and forecasts, so it is a better way to determine how skilful a forecast equation is.



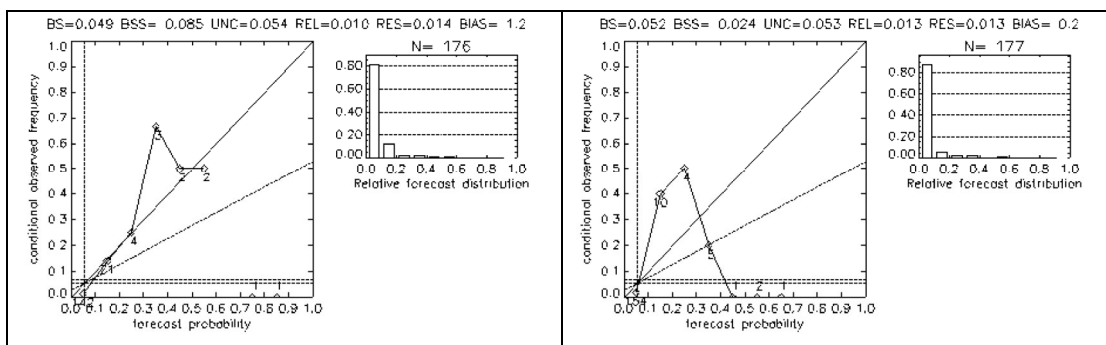
**Figure 30: Reliability diagrams for region MMN and time period 03-09 UTC. Left: +6 h forecast; right: +12 h forecast.**

In Figure 31 (09-15 UTC) it can be seen that there is an overforecasting bias of the low forecast probabilities and an underforecasting bias for the higher forecast probabilities. For lower forecast probabilities the conditional observed frequency is 0, while for the higher forecast probabilities the conditional observed frequency often is 1. The conditional observed frequency is an average of all the observations in a forecast probability bin. If there is only 1 case in the forecast probability bin and it is forecasted wrongly according to the observations, it is much more clearly visible in the reliability diagram than if 1 case was forecasted wrongly out of 100 cases. This is the reason why the frequency diagram is given: If the conditional observed frequency is exactly 0 or 1, one can easily check how many cases were taken into account. The BSS is high for both graphs: 41.4% and 34.4%, and the biases are slightly positive, 1.1% and 2.2%, so there is some small overforecasting, especially of the low probabilities, as discussed above.



**Figure 31: Reliability diagrams for region MMN and time period 09-15 UTC. Left: +6 h forecast; right: +12 h forecast.**

The last time period, 15-21 UTC (Figure 32), shows a reliable curve for the lower +6 h forecast probabilities, but a less reliable curve for the +12 h forecast. However, both diagrams lack information about the higher forecast probabilities. The BSS is not very high: 8.5% for the +6 h forecast and 2.4% for the +12 h forecast. The +6 h forecast in fact shows an almost perfectly reliable pattern, apart from the two points between 0.7 and 0.9 forecast probability. These points each are the result of only 1 forecast, which makes them not significant: more forecasts should be available to determine the quality of forecasts with such high probabilities. For the low probabilities however, most points are around or very close to the perfect reliability line. In the +12 h forecast the lower probabilities are not too bad. There are again a few higher probabilities (which are based on only 1 or 2 points) that show overforecasting.



**Figure 32: Reliability diagrams for region MMN and time period 15-21 UTC. Left: +6 h forecast; right: +12 h forecast.**

The diagrams above are shown for only one region, but they give an average image of how skilful the forecast equations are. For all the other regions and time periods with a reasonable amount of lightning occurrence, the results are similar (not shown). For the regions where the occurrence was below 3% the reliability diagrams are not very useful, because not enough points are available to create a decent reliability diagram. In these cases it is hard to determine the skill of the forecast equations. However, from the regions and time periods that had reasonable lightning occurrence in the verification season, it can be concluded that the forecast equations are skilful.

## Chapter 5

### Summary and conclusions

In this report the development of a probabilistic forecast system for winter thunderstorms in the Netherlands is described. Winter thunderstorms are rare events, which occur approximately 4 to 10 days a year, depending on the location (Heijboer and Nellestijn, 2002). They can develop when cold polar air flows over the relatively warm North Sea surface towards the Netherlands, mainly in small convective clouds or squall lines. Since winter thunderstorms have a smaller vertical extent than summer thunderstorms, aircraft pilots can have difficulties in estimating the lightning risk of cumulonimbus clouds. Also, lightning occurs closer to the surface in winter than in summer, which makes descending and ascending airplanes more vulnerable to lightning. The forecast system developed here can be used to warn aircraft pilots, if high probabilities of winter thunderstorms are forecast.

MOS was used to develop the probabilistic forecast equations, by using logistic regression and the so-called forward-stepwise selection method. The predictand is the probability of at least 1 lightning discharge in a region in a 6-h time period. For the predictand validated SAFIR total lightning data was used. The potential predictors were taken from the NWP models HIRLAM and ECMWF, and from lightning and radar advection (only +6 h forecast). The datasets used in the development of the forecast system contain data from October 16<sup>th</sup> 2004 until April 15<sup>th</sup> 2007, but only the cold half-years (Oct. 16<sup>th</sup> – Apr. 15<sup>th</sup>).

This resulted in 32 forecast equations for the forecast periods 0-6 hours and 6-12 hours, for four regions around Schiphol Airport (each region approximately 80x90km<sup>2</sup>, see Figure 1), and four time periods (21-03, 03-09, 09-15 and 15-21 UTC). These four regions were chosen because they include Schiphol Airport and its holding beacons.

The predictor ‘families’ (section 4.1) that were selected in the MOS system are:

1. HIRLAM Boyden Index (in 27 of the 32 equations). The Boyden index accounts for the conditional instability of the atmosphere, and was selected most often, but never as first predictor. It needs one of the other predictors to provide information about moisture.
2. ECMWF square root of the convective precipitation (in 24 equations). This predictor is selected because lightning is often accompanied by convective precipitation.
3. HIRLAM CAPE (in 14 equations). Unlike the Boyden Index, CAPE is always selected as first predictor. CAPE is the convective available potential energy.
4. Radar advection square root of the precipitation over 6 hours (in 7 of the 16 equations). This newly created predictor turned out to be a good predictor in the short



term forecast equations. For the night, it is a better predictor than the ECMWF convective precipitation. The radar advection was improved by the addition of a clutter filter, because it appeared that the predictor quality decreased without the filter.

5. ECMWF Max. temperature anomaly at 1000 hPa (in 1 equation) and relative vorticity-advection at 500 hPa (in 1 equation). Both predictors are only selected once, in addition to two other predictors.

In the summer thunderstorm forecast system (Schmeits et al., 2008), apart from CAPE, Boyden Index and convective precipitation, the Jefferson Index and lightning advection predictors were often chosen. The Jefferson index is probably not selected in winter because it uses  $T_{500}$ , and winter thunderstorms generally do not reach the 500 hPa-level. It is a good thunderstorm index in summer because summer thunderstorms can reach very high into the troposphere. In future research, this index could be modified by taking another height, for instance 700 hPa, and be checked if it is selected then. Another predictor that was not selected in winter is a lightning advection predictor, probably because winter thunderstorms generally have shorter lifetimes and less lightning activity than summer thunderstorms. This decreases the quality of a predictor that advects lightning activity, since it would probably overestimate lightning occurrence.

A few example situations were described to demonstrate how the forecast system works. However, this does not give a clear indication how skilful the system is. To determine this, the Brier skill score was examined on independent data sets. The best skills can be found for the day, less good but still generally positive skill scores for the night. A BSS boxplot was made using 30 cases. This turned out to be a sufficient number of samples, after finding only a small sampling error when comparing two boxplots of each 30 cases of the same equation. The skill scores indicate that the equations generally have skill compared to climatology.

The best way to test the forecast system is doing a verification on a completely independent dataset. A verification was done on the winter season 2007-2008. Using reliability diagrams from regions and time periods that had a reasonable amount of lightning occurrence in the verification season, it can be concluded that the forecast equations are skilful.

Even with a skilful system, one should never look at the probabilistic forecasts without additional information, as we learned from the clutter situations. It will be a matter of time for the forecasters to learn how the forecast system behaves and at which probabilities action should be taken. This also depends on the demands of the users.

The forecast system developed and described in this report runs at the Royal Netherlands Meteorological Institute (KNMI) on an experimental basis. Forecasters can use the forecast

system to issue warnings to pilots, which could potentially reduce the number of aircrafts struck by lightning. An AIL-specific warning system is currently being developed that gives pilots an indication for the AIL risk, in terms of low, medium or high risk. The forecast system developed here can be used to estimate the AIL risk.

Possible future research could comprehend an extension of the MOS system towards forecast times of +24 h or +48 h. Besides, the quality of the forecast system might be improved by taking the predictors over an area that is larger than the region for which the forecast is done (Kok et al., 2008). Finally, new potential predictors might be investigated, for instance from other observational sources like MSG or, in the future, from a short-term ensemble prediction system and from the non-hydrostatic NWP model HARMONIE (Lynch, 2008).

### *Acknowledgements*

I would like to thank Maurice Schmeits, my supervisor at KNMI, for guiding me through this research. You always made time to talk with me about ideas and problems, which I really appreciated. I would also like to thank Daan Voegelezang, Rudolf van Westrhenen and John Nellestijn from KNMI, who provided me with data. Also, there were a lot of people at KNMI who gave me advice and showed their enthusiasm about my project; thanks for that! I have had a very good time working with you!

## References

- Boyden, C.J., 1963. A simple instability index for use as a synoptic parameter. *Meteor. Mag.* **92**, 198-210.
- Craven, J.P., R.E. Jewell, and H.E. Brooks, 2002. Comparison between observed convective cloud-base heights and lifting condensation level for two different lifted parcels. *Wea. Forecasting*. **17**, 885-890.
- Geophysics Study Committee, 1986. *Studies in Geophysics: The earth's electrical environment*. National Academy Press, 263 pp.
- Glahn, H.R., and D.A. Lowry, 1972. The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.* **11**, 1203-1211.
- Heijboer, D., and J. Nellestijn, 2002. *Klimaatatlas van Nederland, de normaalperiode 1971-2000*. Uitgeverij Elmar, 182 pp. (In Dutch)
- Houze, R.A., 1993. *Cloud dynamics*. Academic Press, 573 pp.
- Hughes K. K., 2001: Development of MOS thunderstorm and severe thunderstorm forecast equations with multiple data sources. Preprints, *18th Conf. on Weather Analysis and Forecasting*, Fort Lauderdale, Amer. Meteor. Soc., 191-195.
- Jefferson, G.J., 1963a. A further development of the instability index. *Meteor. Mag.* **92**, 313-316.
- Jefferson, G.J., 1963b. A modified instability index. *Meteor. Mag.* **92**, 92-96.
- Kasemir, H.W., 1950. Qualitative Uebersicht ueber Potential-, Feld-, und Ladungsverhältnisse bei einer Blitzentladung in der Gewitterwolke. In *Das Gewitter*, ed. H. Israel, Akad. Verlags. Ges. Geist and Portig K.-G, Leipzig, Germany.
- Katz, R.W., and Murphy, A.H., 1997. *Economic Value of Weather and Climate Forecast*. Cambridge University Press, 238 pp.
- Kok, K., B. Wichers Schreurs and D. Vogelesang, 2008. Valuing information from mesoscale forecasts. *Met. App.* **15**, 103-111.
- Kitigawa, N., and K. Michimoto, 1994. Meteorological and electrical aspects of winter thunderclouds. *J. Geophys. Res.* **99**, 10713-10721.
- Lynch, P., 2008. The origins of computer weather prediction and climate modeling. *J. Comp.Phys.* **227**, 3434-3444.
- Mason, S.J., 2004. On Using "Climatology" as a Reference Strategy in the Brier and Ranked Probability Skill Scores. *Mon. Wea. Rev.* **132**, 1891-1895.
- Moncrieff, M.W., and M.J. Miller, 1976. The dynamics and simulation of tropical cumulonimbus and squall lines. *Quart. J. Roy. Meteor. Soc.* **102**, 373-394.
- Murooka, Y., 1992. A survey of lightning interaction with aircraft in Japan. *Res. Lett. Atmos. Electr.* **12**, 101-106.
- Noteboom, S., 2006. Processing, validatie, en analyse van bliksemdata uit het SAFIR/FLITS system. *Internal report KNMI, IR-2006-01*. (Available from KNMI, in Dutch)

- Rakov, V.A., and M.A. Uman, 2003. *Lightning, Physics and Effects*. Cambridge University Press, 687 pp.
- Schmeits, M.J., C.J. Kok, and D.H.P. Vogelesang, 2005. Probabilistic forecasting of (severe) thunderstorms in the Netherlands using Model Output Statistics. *Wea. Forecasting*. **20**, 134-148.
- Schmeits, M.J., C.J. Kok, D.H.P. Vogelesang and R.M. van Westrhenen, 2008. Probabilistic forecasts of (severe) thunderstorms for the purpose of issuing a weather alarm in the Netherlands. *Wea. Forecasting, in press*.
- Uman, M.A., and V.A. Rakov, 2003. The interaction of lightning with airborne vehicles. *Progress in Aerospace Sciences*. **39**, 64-81.
- Wessels, H.R.A., 1998. Evaluation of a radio interferometry lightning positioning system. *Scientific report KNMI, WR-98-04*. De Bilt, The Netherlands, 26 pp.
- Wilks, D.S., 2006. *Statistical methods in the atmospheric sciences, second edition*. Academic Press, 627 pp.

## Appendix

### **Appendix A: Forecast equations with coefficients**

On the next page a table is given that contains all 32 forecast equations with the coefficients. The equations are grouped by verification time and forecast time and in each block the 4 regions can be found. The probability of a thunderstorm can be calculated using formula (1):

$$p_i = \frac{1}{1 + \exp(-b_0 - b_1 x_1 - \dots - b_K x_K)}$$

For verification time 00 UTC, forecast time +6 h and region WMN, this results in:

$$p_i = \frac{1}{1 + \text{EXP}[-(-121.3333) - (0.0400 \cdot \text{TPGE03RA}) - (0.1234 \cdot \text{BOYDMAX})]}$$

|               |            | forecast +6 h |          |          |          | forecast +12 h |          |          |          |
|---------------|------------|---------------|----------|----------|----------|----------------|----------|----------|----------|
| <b>00 UTC</b> | <b>WMN</b> | constant      | tpge03ra | boydmax  |          | constant       | capemax  | boydave  |          |
|               |            | -121.3333     | 0.0400   | 0.1234   |          | -106.4677      | 0.0006   | 0.1087   |          |
|               | <b>WMS</b> | constant      | tpge03ra | rtcp6ec3 | boydmax  | constant       | capemax  | rtcp3ec  |          |
|               |            | -85.3534      | 0.0472   | 0.0054   | 0.0840   | -5.7112        | 0.0014   | 0.0061   |          |
|               | <b>MMN</b> | constant      | tpge03ra | boydmax  |          | constant       | rtcp3ec  | boydave  |          |
|               |            | -95.8229      | 0.0210   | 0.0971   |          | -106.7093      | 0.0042   | 0.1089   |          |
|               | <b>MMS</b> | constant      | tpge03ra | boydmax  |          | constant       | rtcp3ec  | boydave  |          |
|               |            | -122.9340     | 0.0286   | 0.1250   |          | -164.4444      | 0.0059   | 0.1684   |          |
| <b>06 UTC</b> | <b>WMN</b> | constant      | rtcp3ec  | boydmax  | tpge03ra | constant       | lnbmax   | rtcp6ec  |          |
|               |            | -96.3630      | 0.0072   | 0.0962   | 0.0260   | -6.0234        | 0.00005  | 0.0064   |          |
|               | <b>WMS</b> | constant      | cap100ma | rtcp6ec3 |          | constant       | rtcp6ec  | boydmax  | mxt1000a |
|               |            | -5.3786       | 0.0011   | 0.0065   |          | -85.1120       | 0.0059   | 0.0853   | -0.0311  |
|               | <b>MMN</b> | constant      | cap100ma | boydmax  | rtcp6ec3 | constant       | lnbmax   | rtcp6ec  | boydmax  |
|               |            | -112.3131     | 0.0007   | 0.1143   | 0.0059   | -70.7969       | 0.00005  | 0.0045   | 0.0700   |
|               | <b>MMS</b> | constant      | cap100ma | boydmax  |          | constant       | lnbmax   | rtcp6ec  |          |
|               |            | -83.5258      | 0.0010   | 0.0842   |          | -5.9941        | 0.0001   | 0.0056   |          |
| <b>12 UTC</b> | <b>WMN</b> | constant      | rtcp6ec3 | boydmax  |          | constant       | cap100ma | rva500   | boydmax  |
|               |            | -132.2693     | 0.0059   | 0.1336   |          | -129.5581      | 0.0014   | 0.0317   | 0.1315   |
|               | <b>WMS</b> | constant      | rtcp6ec3 | boydmax  |          | constant       | rtcp6ec3 | boydmax  |          |
|               |            | -119.4271     | 0.0055   | 0.1213   |          | -125.3190      | 0.0052   | 0.1276   |          |
|               | <b>MMN</b> | constant      | capemax  | rtcp6ec3 | boydmax  | constant       | cap100ma | rtcp6ec3 | boydmax  |
|               |            | -82.4575      | 0.0011   | 0.0050   | 0.0823   | -107.3187      | 0.0009   | 0.0054   | 0.1086   |
|               | <b>MMS</b> | constant      | rtcp6ec3 | tpge10ra | boydmax  | constant       | rtcp6ec3 | boydmax  |          |
|               |            | -187.9453     | 0.0113   | 0.0599   | 0.1901   | -181.6986      | 0.0093   | 0.1848   |          |
| <b>18 UTC</b> | <b>WMN</b> | constant      | rtcp6ec  | boydmax  |          | constant       | capemax  | rtcp6ec  |          |
|               |            | -117.7633     | 0.0063   | 0.1185   |          | -5.8610        | 0.0009   | 0.0060   |          |
|               | <b>WMS</b> | constant      | rtcp6ec  | boydmax  | tpge10ra | constant       | rtcp6ec  | boydave  |          |
|               |            | -150.1830     | 0.0035   | 0.1536   | 0.0382   | -90.4804       | 0.0046   | 0.0912   |          |
|               | <b>MMN</b> | constant      | cap50max | boydmax  |          | constant       | cap100ma | boydave  |          |
|               |            | -70.6658      | 0.0011   | 0.0716   |          | -52.8963       | 0.0014   | 0.0528   |          |
|               | <b>MMS</b> | constant      | rtcp6ec  | boydmax  |          | constant       | rtcp6ec  | boydave  |          |
|               |            | -137.2569     | 0.0063   | 0.1396   |          | -76.6737       | 0.0067   | 0.0762   |          |

## **Appendix B: Acronyms**

|        |   |
|--------|---|
| AIL    | Aircraft induced lightning  |
| BSS    | Brier skill score   |
| CAPE   | Convective available potential energy                               |
| ECMWF  | European centre for medium-range weather forecasts                  |
| HIRLAM | High-resolution limited-area model                                  |
| HRVIS  | High resolution visible channel                                     |
| LFC    | Level of free convection  |
| LNB    | Level of neutral buoyancy   |
| MOS    | Model output statistics   |
| NWP    | Numerical weather prediction  |
| RTCP   | Square root of convective precipitation                             |
| SAFIR  | Surveillance et d'alerte foudre par interférométrie radioélectrique |
| SATREP | Satellite and report  |
| UTC    | Universal time coordinated  |