Technical report; TR-289

Cloud masking using MSG and HIRLAM

J.P.J.M.M. de Valk

De Bilt, 2006

KNMI Technical report = technisch rapport; TR-289

De Bilt, 2006

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: Valk, J.P.J.M.M. de

Cloud Masking Using MSG and HIRLAM

J.P.J.M.M. de Valk

KNMI

Report of NIVR Demonstration project 1.1 DE 01

NIVR project number : 52310 KN

KNMI Technical Report No :289

November 2005

NIVR
Nederlands Instituut voor Vliegtuigontwikkeling en Ruimtevaart
Netherlands Agency for Aerospace Programmes

P.O. Box 35
2600 AA Delft
The Netherlands
Kluyverweg 1
2629 HS Delft
The Netherlands
Tel. +31 (0)15 278 80 25
Fax +31 (0)15 262 30 96
Email info@nivr.nl
www.nivr.nl

NATIONAL USER SUPPORT PROGRAMME (NUSP)

2001-2005

http://www.ao-go.nivr.nl

The National User Support Programme 2001-2005 (NUSP) is executed by the Netherlands Agency for Aerospace Programmes (NIVR) and the SRON Netherlands Institute for Space Research. The NUSP is financed from the national space budget. The NUSP subsidy arrangement contributes to the development of new applications and policy-supporting research, institutional use and use by private companies.

The objectives of the NUSP are:

- To support those in the Netherlands, who are users of information from existing and future European and non-European earth observation systems in the development of new applications for scientific research, industrial and policy research and operational use;

- To stimulate the (inter)national service market based on space-based derived operational geo-information products by means of strengthening the position of the Dutch private service sector;

- To assist in the development of a national Geo-spatial data and information infrastructure, in association with European and non-European infrastructures, based on Dutch user needs;

- To supply information to the general public on national and international space-based geo-information applications, new developments and scientific research results.

The Royal Netherlands Meteorological Institute
In the Netherlands, KNMI (Royal Netherlands Meteorological Institute) is known mainly for its weather forecasts and warnings, but is does a lot more in its capacity as a national data and knowledge centre for weather, climate research and seismology.
KNMI was founded by Royal Decree on 31 January 1854. C.H.D. Buys Ballot (1817-1890), its first general director, chose to establish KNMI at the Sonnenborgh observatory in Utrecht. In 1897, the Institute moved to De Bilt, where its headquarters is still located today. In all, KNMI has some 500 employees.

KNMI's mission
KNMI is the national institute for weather, climate research and seismology. It disseminates weather information to the public at large, the government, aviation and the shipping industry in the interest of safety, the economy and a sustainable environment. To gain insight into long-term developments, KNMI conducts research on climate change. Making the knowledge, data and information on hand at KNMI accessible is one core activity. KNMI is an agency of the Ministry of Transport, Public Works and Water Management (Ministerie van Verkeer & Waterstaat). KNMI's duties are set forth in KNMI Act (Wet op het KNMI).

Working globally - 24 hours a day
KNMI operates twenty-four hours a day. The combination of infrastructure, technology, sciences and service is unique in the fields of weather, climate research and seismology. On an international level, KNMI works closely with other institutes and research establishments, such as the European Centre for Medium-Range Weather Forecasts (ECMWF) in England. KNMI also represents the Netherlands in numerous other international organisations, such as the World Meteorological Organisation (WMO), the Intergovernmental Panel on Climate Change (IPCC) and the European meteorological satellite network EUMETSAT.

Operational meteorology
At the weather centre in De Bilt, KNMI prepares weather forecasts and warnings for the public at large and the aviation and shipping industries.

Meteorological research
KNMI conducts both strategic and applied research. The strategic research focuses on understanding the most important processes that determine developments in the atmosphere, the oceans and on land, while the applied research focuses on answering questions posed by society and on developing the observation and modelling systems required to generate the necessary data. By housing operational meteorology and meteorological research under one roof, KNMI harnesses the power of synergy.

Climate research
Climate research at KNMI focuses on observing, understanding and predicting changes in climate systems. Our selection of research topics is based on the state of international and Dutch climate research and on questions posed by the government and the public

Seismology
In addition, KNMI also conducts research in the field of seismology, makes observations that form the basis for this research and disseminates information to a wide audience on earthquakes and related phenomena.

Contents

Executive Summary

Clouds occurrence is the direct consequence of atmospheric processes relevant to the (scientific) community. Their presence is a significant indication about the atmospheric state. Knowledge about cloud patterns is therefore important for forecasters.
Remote sensing of clouds from space has become an irreplaceable information source for meteorological and climatological applications. Lacking of the satellite information would lead to a significant degradation of the forecast skill of meteorologists.

Since January 2004 observations from a new type of satellite instrument the Spinning Enhanced Visible and Infra Red Imager (SEVIRI) became available. SEVIRI is located on a new geostationary platform, referred to as METEOSAT 8, operated by the inter-governmental organisation EUMETSAT, located at Darmstadt, Germany. The SEVIRI instrument offers observations in eleven spectral channels, with improved spatial (in the infrared nadir 3x 3 km²) and temporal resolution (15 minute repetition cycle) , where the predecessor instrument, the Meteosat Visible and Infra Red Imager MVIRI offered three spectral channels, a spatial resolution at nadir in the infrared of 5x5km², and a repetition cycle of 30 minutes. To evaluate the wealth of information provided by SEVIRI the forecaster can no longer rely on visual inspection of all the spectral channels separately. New synergetic interpretation methods using all of the spectral channels are required to profit optimal from the possibilities SEVIRI offers.

This project aims to help forecasters, climatologists, and numerical weather prediction modellers by producing an automated interpretation of the satellite observation. First a cloud mask is produced, differentiating cloud contaminated pixels from cloud free pixels. The cloud mask is the first important part of automated interpretation. The cloud masking is then followed by cloud typing.

To achieve the goal an algorithm is developed. It uses the SEVIRI observations, and the surface temperature from the KNMI numerical weather prediction model (HIRLAM). The method is based on knowledge attained in a BCRS USP funded project, named MetClock (METeosat CLOud Characterisation KNMI). The output is a cloud mask. The results are evaluated by comparison to surface observations (called synops) done at meteorological stations.

During the period of the project approval and the actual project execution EUMETSAT released through the Satellite Application Facility on Now casting and short range forecasting (NWC-SAF) a software package. The software package also provides also a cloud mask. Within the project it was decided to evaluate this cloud mask simultaneous with the KNMI derived cloud mask. The chosen approach allows for a comparison between the NWC-SAF products and the KNMI algorithm.
The NWC-SAF software package supplies, next to the cloud mask, a cloud type product and a cloud top height product

Conclusion

The performances of the two algorithm's (KNMI and NWC-SAF) are comparable. The probability of detection (POD) scores are generally over 80 percent. The False alarm rates for cloudy (FAR8) and clear skies (FAR0) are acceptable.
During daytime the thresholding tests for the Reflection channels contributed most to the FAR0. A better tuning of these tests will improve the model performance.
The high POD scores prove that the use of HiRLAM surface temperatures has a significant positive impact on the successful detection of cloud contaminated pixels. It was one of the aims of the project to prove this concept.

The NWC-SAF algorithm is more elaborate compare to the KNMI algorithm. It produces not only a cloud mask but also a cloud type and cloud top heights. These additional products were accepted as added value by the end users, the forecasters.

At the end of the project a performance dependence on latitude and twilight conditions became clear. Due to time constraints this dependence could not be fully explored in this project.

The end users can be subdivided into two groups.
The forecasters require easy interpretable products having a clear added value to the data they already have. They accept the higher level products like cloud types and cloud top heights of the NWC-SAF package as added value products.
The second group consists of researchers who incorporate products in automated process chains. They use a basic product like the cloud mask in their software.

Both groups are eager to use the products as can be derived from the new generation geostationary satellites. It is planned to serve the users with the NWC-SAF products, based on the following arguments:

- the NWC-SAF algorithm provides higher level products, such as cloud type, cloud top height. These products are not available in the KNMI algorithm.
- The performance of both packages show a comparable quality. This does not provide an argument to choose for one of the packages. So the first argument, given above, prevails.

1 Introduction

Figure 1.1 The sunset over the Adriatic Sea. The scene highlights the beauty of clouds.

Clouds are important and often beautiful (Figure 1.1) phenomena in the atmosphere. Clouds consist out of water or ice particles. Clouds are formed through condensation of water vapour in cooling air. The cooling is caused by ascending or mixing of warm moist air with cooler air. The earth atmosphere is unique in the solar system that it has a temperature range in which water can exist in solid, (ice), fluid (water) and gaseous (vapour) form. During the cloud formation process the cloud affects the local radiance balance, reflecting solar radiation to space and reflecting infrared earth radiation back to earth. This will affect the local temperature regime.
The occurrence of clouds thus is a subtle balance between various underlying atmospheric physical processes which are of interest for the (scientific) community.
The role of clouds in the earth climate system is at present not fully understood. On the one hand their reflectivity decreases the incoming solar radiation. On the other hand their coverage limits the infra red emission of the earth. The average global cloud cover is well over sixty percent. Therefore clouds dominate the reflection of the direct sun light.
In long term cloud coverage studies it is still not clear if there is an upward or downward trend of total coverage.
Cloud occurrence has a direct impact on the society. There is a direct link between outdoor activities like recreation and cloud coverage. Cloud coverage mitigates the

possible UV dose on the skin in the European summer. Solar energy production is significantly reduced due to cloud coverage. Hence the infrastructure of solar power stations is determined by cloud climatology.
Associated with clouds is the occurrence of extreme precipitation, potentially related to climate change. In the summer of 2005 precipitation created floods in Europe causing loss of lives and drawing large media attention. Also several deaths were reported due to lightning. Lighting is another phenomena related to convection.
Hence forecasters, hydrologists, climate researchers, and the general public are interested in reliable cloud detection. Satellite imagery is the only means providing information with significant coverage about the (horizontal) spatial and temporal distribution of clouds.

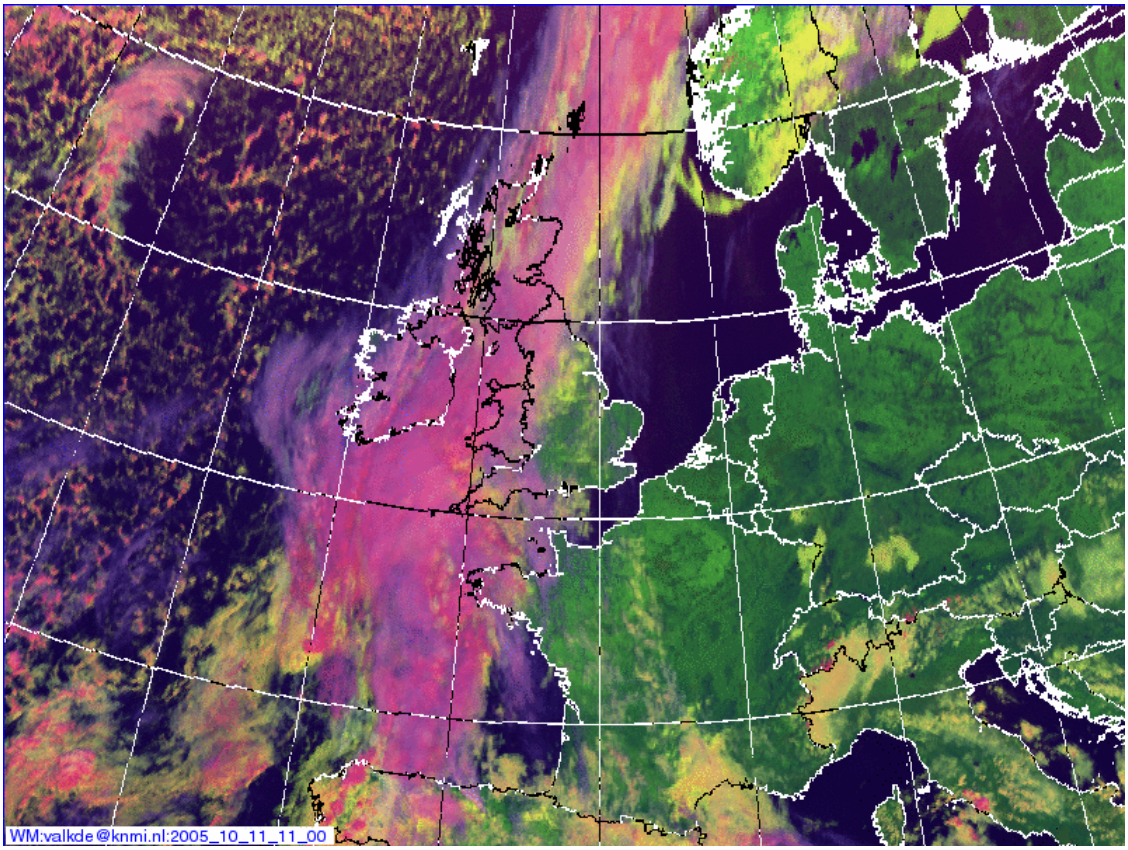

Figure 1-2 A colour composite of the METEOSAT 8 channels. Red 0.6 µm, Green 1.6 µm, Blue 10.8 µm. The chosen combination allows for discrimination between sea (dark bleu) land (green) water clouds (yellow) and (semi transparent) ice clouds (pink, purple), snow (lurid pink).

The successfully launched satellite Meteosat Second Generation (MSG), in 2004 renamed METEOSAT 8, observes the earth in eleven spectral bands with the main payload of the MSG satellite, i.e. the Spinning Enhanced Visible and Infra Red Imager (SEVIRI). The Meteosat Visible and Infra Red Imager (MVIRI) on board of the predecessor satellite METEOSAT 7 observes in three spectral bands. The increased spectral information of METEOSAT 8 can no longer be evaluated in all its possible exploitations by a forecaster behind a screen studying all the spectral bands. This procedure was still common for the interpretation of METEOSAT 7 imagery. In addition to this interpretation complexity it is already proven there is also value added information

in the subtraction of various channels. In Figure 1-2 the added value of the extra spectral channels is demonstrated.

During the former USP a number of projects have been successfully completed through which automated cloud detection acquired a significant role in the analysis of satellite imagery. Automated cloud detection using Meteosat imagery is now the solid base for short-time-term-cloud-forecasting models. It is also used to correct the forecast of the UV- dose for cloud coverage.
The potential of the automated cloud detection to monitor the performance of the ground based present weather sensors with regard to cloud coverage is not yet exploited at KNMI.

When the project of this report was proposed it was anticipated that there was no tested and validated software available for interpretation at the start of the operational data dissemination. A gap was foreseen which could hamper production continuity of automated cloud masks.
To avoid this gap for a number of operational products at KNMI, the project proposal was prepared and applied for in the GO-NIVR programme.

During the period of the project approval and the actual project execution the European Organisation for the exploitation of Meteorological Satellites (EUMETSAT) provided through the Satellite Application Facility on Now casting and short range forecasting (NWC-SAF) a software package. The software package determines amongst other products a cloud mask. Within the project it was decided to evaluate this cloud mask simultaneous with the KNMI derived cloud mask. The chosen approach allows for a comparison between the NWC-SAF products and the KNMI algorithm.

1.1 Project objective

The project objective is to demonstrate that a cloud mask based on SEVIRI imagery is superior to the cloud mask derived from MVIRI imagery. The extraction of a cloud mask based on METEOSAT 8 is required as METEOSAT 7 will stop its service in 2006.
At present the METEOSAT 7 cloud mask is used in an operational mode for the forecasters. The METEOSAT 7 mask serves as an initialization for short range (less than 12 hours) cloud forecast model. In 2005 the METEOSAT 8 cloud mask will replace the METEOSAT 7 cloud mask.
Future applications will demand cloud information with a higher quality and a higher update frequency. In the future NWP models are foreseen that assimilate cloud products.
As the NWC-SAF software became available throughout the project duration two goals are added:

- The comparison between the KNMI algorithm and the NWC-SAF cloud masks.
- The performance of the NWC-SAF validated against the same validation environment for the KNMI detection algorithm.

The project is proposed as a demonstration project as part of the research is already done in the former USP MetClock project, Valk et al, 1998. Within this project a number of aspects are addressed:

- The re-evaluation of the thresholding tests applied for cloud masking. The METEOSAT 8 spectral channels are different from the ones evaluated for the MetClock algorithm. Here the heritage of MetClock is used. This implies the

use of the 10.8 channel versus the HiRLAM surface temperature, and the use of the 0.6 and 0.8 channels versus a derived surface reflectance.
- Research on how to use the added channels of METEOSAT 8 relative to the METEOSAT 7 channels in order to derive a higher quality cloud mask.
- Collocation of the HiRLAM NWP fields versus the METEOSAT 8 pixels
- Collocation between the chosen area of study and the NWC-SAF chosen area.
- Collocation of both NWC-SAF and KNMI algorithm cloud mask with the validation environment based on synops stations.

Most of the project effort is dedicated to the (fine) tuning of the tests within the KNMI algorithm to obtain a good balance between correct and erroneous detection of cloudy and clear skies.

1.2 Relevant expertise at KNMI

The project builds on the expertise already gathered in previous USP and NOP projects.
KNMI has a proven ability to come from research projects via demonstration to implementation projects e.g. MetClock. The developers of MetClock are still working in this field of research at KNMI. This minimizes the acquaintance time with the problem.
Within the same research group the NWC-SAF software is acquired and implemented. The gained expertise here will reflect upon the project results.
Within KNMI there is an environment where the cooperation between researchers and end-users, (NWP researchers and forecasters) is well embedded. The short communication lines enable fast, good and thorough feedback.

1.3 (International) relevance of the project

Reliable cloud masks are beneficial to the whole meteorological community consisting of forecasters, modellers and climate researchers.
The comparison with NWC-SAF cloud mask and evaluation against the KNMI derived cloud mask will be of interest to the NWC-SAF community and the users of this software package. The NWC-SAF products are used by the Climate-SAF, the Ocean and Sea Ice SAF, and the Land Surface Applications SAF. This report includes an independent evaluation of the NWC-SAF cloud mask product performed at KNMI.

Climate researchers are interested in the best available data. Given the significant role of clouds within the climate there is interest in good quality cloud mask and when possible an estimate about the uncertainties. For the climate researchers the present study provides an opportunity to compare this cloud mask to other sources of cloud information.

There is a growing tendency to automate synoptical observations. Cloud masks from satellite provide a tool to monitor the performance of the automated stations.
A new area of interest has become a focal point of the KNMI, the chemical weather forecast. Part of the chemical reactions leading to air pollution depend on (sun) light. The occurrence of clouds can therefore hamper chemical reactions leading to smog.
There is also a direct link to the UV dose on the skin and cloud occurrence. The cloud mask is essential to determine the forecasted UV dose.

Next to the identified KNMI end-users, the radar group emerged as a potential end-user during the project. Radar echo's can be contaminated by false echoes caused by insects or sharp temperature inversions within the atmosphere or ground reflections. False radar echo's can be filtered with a good cloud mask, as radar reflections related to precipitation are not to be expected in a clear sky.

For the general public the weather presentation via television presentations can be improved. Commercial Providers can filter their cloud loops. Here still a relative simple threshold mask is used. In 2005 it was frequently noticed that low level clouds were not represented in the animation loops on television. The weather presenter had to explain this omission. With a higher quality cloud mask these omissions could be minimized, thus contributing to the weather presentation.

1.4 Outline

The structure of the report comprehends after this introduction, the context of automated cloud masking. This is followed by description of the required data, and the development of the algorithm, including software portability issues. The method of assessing the algorithm results is treated in chapter five. Obtained results and evaluation by end users are given separately in chapters six and seven and precede the conclusions given in chapter eight.

2 Cloud masking in context.

Since the first images of the earth from space it was realized what valuable information content these images have. It was then that meteorologist could actually see the cloud patterns they were describing and forecasting like spirals around a depression and other distinct cloud patterns. The images of the first geostationary satellite ATS 1 available since January 5, 1967 added the sensation of movement represented in time loops
The need for automated detection of clouds from satellite imagery as an aid to the forecasters and model developers has stimulated research. Cloud characteristics make them relative easily detectable in imagery by the human eye. From 0.4 μm up to 2.5 μm clouds reflect sunlight mostly more effective than the earth surface. At infrared wavelengths from 6 μm to longer wavelengths they are generally colder then the under laying surface. There are however numerous exceptions with clouds being warmer than the surface. Also clouds over snow covered areas are a challenge to see, even for the human eye. Humans can also easily spot patterns which indicate that a cloud is there, especially when the imagery is presented in an animation loop.
All these subtle characteristics of clouds make the automated detection of clouds a challenge.

2.1 A literature review

A review on cloud detection in satellite imagery is given by Rossow (1989). The coarse spatial resolution of METEOSAT forms a serious constraint to obtain quantitative cloud parameters. Most common researched methods have been concentrated on the detection of cloud contaminated pixels in imagery using the deviating characteristics of clouds opposite to the Earth surface: clouds are colder, reflect more sunlight and have a higher spatial variability. The concepts of these approaches are well described: asymmetric Gaussian histogram analysis (Simmer et al., 1982), dynamic clustering method (Desbois et al., 1982), hybrid bispectral threshold analysis (Minnis and Harrison, 1984a,b,c), spatial coherence method (Coakley and Bretherton, 1982), histogram clustering method (Seze and Desbois, 1987), and segmented spatial coherence method (Shin et al, 1996). The methods have in common that they analyse a large area in which there are cloudy and cloud-free areas. They differ in the retrieval of the Earth surface characteristics. The operational use of the mentioned methods is constrained by the fact that they provide averaged and not accurately determined cloud surface properties. Furthermore the results of the methods depend on the region and atmospheric conditions (Rossow et al, 1985; Seze and Rossow, 1991a,b). In the framework of the International Satellite Cloud Climatology Project (ISCCP), part of the World Climate Program, the ISCCP cloud detection method is developed (WCP, 1988; Rossow et al., 1993a,b,c). For the first time the high temporal resolution of the METEOSAT was used in this algorithm to detect clouds based on the high variability of clouds (in relation to the variability of Earth surface).
The basic assumptions in the ISCCP-method are:

- if the apparent brightness temperature of an observed area has dropped more than can be expected from the Earth surface a cloud must have emerged,
- if the area reflects significantly more sunlight than at another instant this must be caused by a cloud.

Wielicki and Parker (1992) found that the ISCCP-method is the most effective cloud detection method. However, the quality of the algorithm depends on the accuracy of the

estimated change of surface radiative properties for example due to diurnal variations of air mass change. In the ISCCP-algorithm the detection threshold is determined statistically for sea, land, mountains and coastal areas as these surface types have different properties. The surface temperature and reflectance are retrieved from the cloud-free pixels for a 200x200 km² area for a 5-10 day period. This approach introduces detection errors on a pixel-by-pixel basis due to spatial and temporal in homogeneity of the surface radiative properties (vegetation, orography, changes in air masses, wet surfaces). Still it is very suitable for the purpose of ISCCP: a global cloud climatology.

Most work on cloud characterisation, reported in literature, has been focused on application of methods for the purpose of climate research. Only a limited number of methods are developed in Europe that use cloud masks as input for a forecast model (van der Veen and Feijt, 1996 and Zelenka et al, 1998, Veen, 2002).

2.2 The MetClock system

The NRSP funded project MetClock (METeosat CLOud Characterisation KNMI) established a cloud masking algorithm based on METEOSAT 6 and 7 imagery. This cloud mask was operationally implemented in an Implementation project of the NRSP-2/GO 1996-2000.
METEOSAT 7 is the predecessor of METEOSAT 8. METEOSAT 7 has three spectral channels, where the METEOSAT 8 has eleven channels. The METEOSAT 7 has therefore a limited capacity to detect clouds in comparison to METEOSAT 8. Cloud detection in MetClock is based on a thresholding technique. Observed brightness temperatures are compared to calculated surface temperatures from HiRLAM, the KNMI NWP model. MetClock also derives 'cloud cleared' surface masks for the visual [reflection] channel, 0.5-0.9 µm, by averaging clear pixel values over the previous fortnight. These cloud cleared reflection maps are used to determine a cloud mask by thresholding the pixel values of the visual channel in comparison to the cloud cleared composite.
There is a physical difference between the observed brightness temperature of the surface from satellite and the calculated surface temperature of a model. The former is the radiance of the true surface skin affected by the atmospheric contribution, amongst which the radiance from aerosol, water vapour. The latter is a temperature representation of the first finite layer of the surface represented by the model, balancing the latent and sensible heat fluxes within the model.
A significant improvement in cloud detection in MetClock is caused by the introduction of a correction term for HiRLAM surface temperatures for the above described difference between modelled temperatures and occurring skin temperatures.
The correction term is derived from clear situations, reported by the synops. For each synops station reporting a clear sky the brightness temperature is subtracted from the HiRLAM surface temperature. These values are gathered for the previous fortnight. The spatial and time difference weighted averaging of these point values determine a correction field for the temperature difference between HiRLAM surface temperature and observed brightness temperature. (Feijt and de Valk, 1998). The correction term showed a daily cycle with an amplitude (maximum minus minimum value) up to ten Kelvin.
The introduction of the term as a correction on the HiRLAM surface temperature allowed for lower threshold values increasing the detection efficiencies and obtaining low false alarm rates.
The thus obtained MetClock results, de Valk et al, 1997, 1999, Feijt and de Valk, 2001, are used in the short range cloud forecasting MetCast system, Veen et al 1996. An

example of a resulting cloud mask is given in Figure 2.1. Artificial colouring adds information about which threshold test resulted in a cloud masking.

The acquired knowledge in MetClock served as a base for the present project.

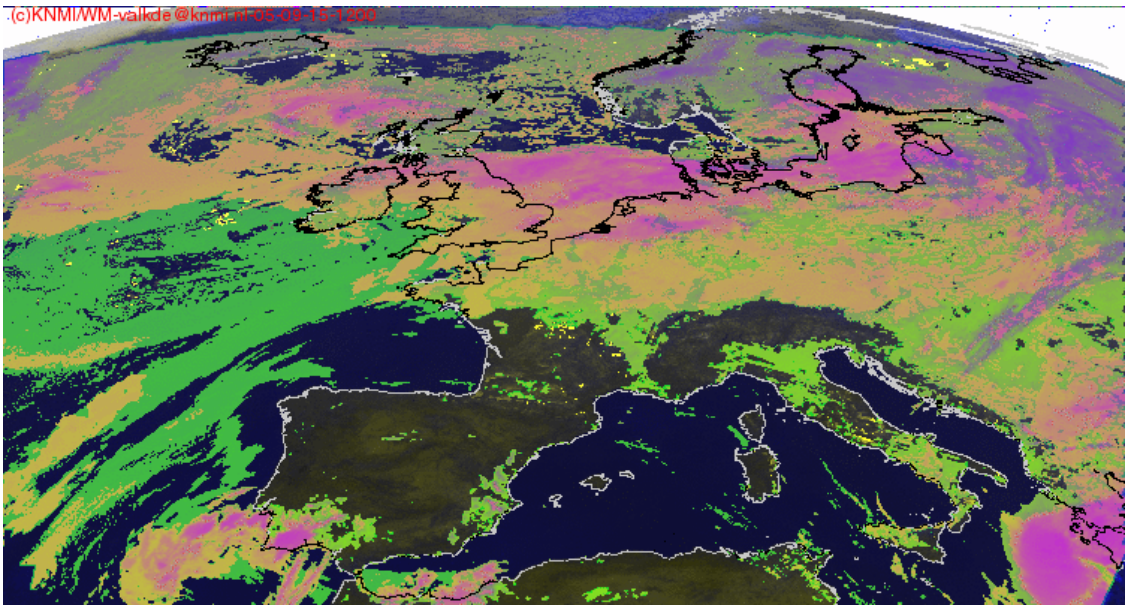

Figure 2-1 The MetClock cloud mask based on METEOSAT 7 imagery. Here the colours indicate the succesful tests, resulsting in a differentiating between low clouds (green, yellow) and highclouds (pink, purple) .

2.3 NWC-SAF

In an early stage of the “new generation satellite” development programme EUMETSAT has recognised the need for a central research and development of high level METEOSAT 8 products. Therefore, EUMETSAT has initiated the Satellite Application Facility (SAF) concept. The SAF’s are formed by consortia of mostly national weather institutes that have specific expertise. Among other SAF’s, the Now casting SAF (NWC-SAF) has been formed by the national meteorological institutes of Spain, France, Sweden and Austria. Over a period of 5 years this NWC-SAF has developed a software package that processes in real time METEOSAT 8/SEVIRI data towards now casting products to be used in the operational tasks of weather services. More detailed information about the NWC-SAF project and products can be found at: www.eumetsat.de/en/area4/saf/internet/main_safs/nwc/main_nwcsaf.html

Within the SAF in support to now casting and short range forecasting, Meteo France (MF) is responsible for the algorithm detecting clouds. MF has a long heritage in satellite interpretation and software development. MF has developed software to extract cloud parameters from the METEOSAT 8 imagery, LeGleau and Derrien, 2000 and 2003. The prototype software was developed on the AVHRR, GOES and HIRS satellite products. The software produces cloud mask, cloud type, and cloud top height and temperature. Except for the cloud mask all other parameters require NWP information.

To obtain the cloud mask the software applies a number of thresholding tests using the 0.6, 0.8, 1.6, 3.9, 8.7, 10.8 and 12.0 channels. It also considers the local spatial texture within the 0.6, 0.8, 3.9 and 10.8 channels.

The latest version of the software, version 1.2, is accepted by KNMI to be in a mature phase. It became available in May 2005. Just in time to evaluated the results together with the results of the KNMI algorithm.
At KNMI the software is installed and NWP fields of ECMWF are used in a semi-operational setting.

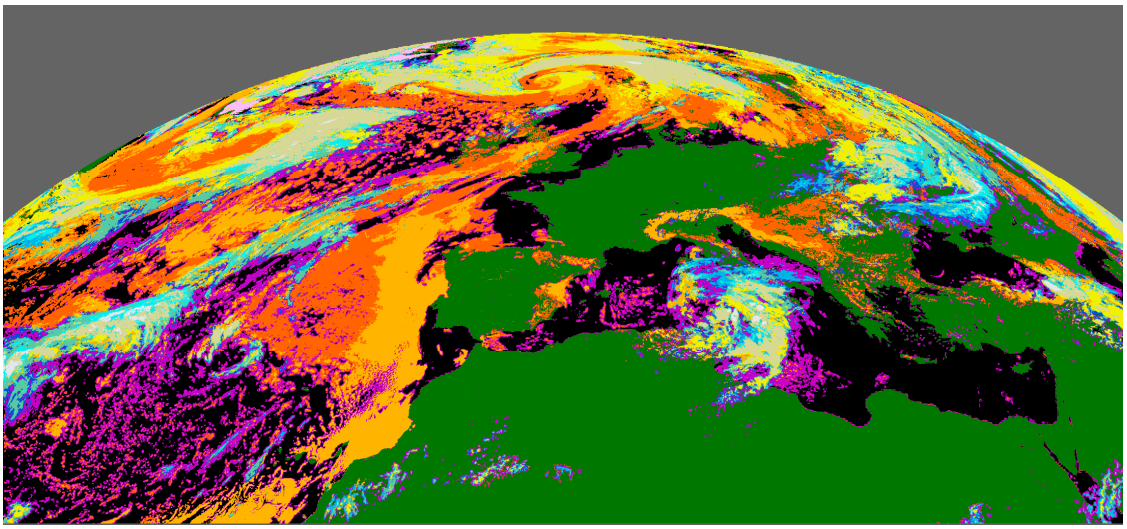

Figure 2-2. The NWC-SAF cloud type product. Each colour represents a different cloud type ranging from very low clouds (orange) to high clouds (purple, cyan)

In Figure 2.2 an example of a cloud type product in colour is given. Each colour represents a different classification type. The discrimination between the various levels of clouds is based on NWP information. The height classes for clouds are very low, low, medium, high and very high. Additionally semitransparent, semi transparent over lower clouds, and fractional clouds are distinguished. The surface types comprise sea, land, snow, and ice.

3 Description of used data within the project.

The principal input data for the algorithm are METEOSAT 8 channels. For cloud detection the algorithm requires an estimation of the surface skin temperature based on NWP model output. HIRLAM NWP model data is used as input here.
For validation of the results synops reports are used.
Here the relevant characteristics of the data types used in this project are described.

3.1 METEOSAT 8 data

The Spinning Enhanced Visible and Infra Red Imager data (SEVIRI, onboard the Meteosat Second Generation geostationary satellite, recently renamed METEOSAT-8) offers a wealth of information in comparison to its predecessor METEOSAT 7.
It observes the earth at medium resolution (3 x 3 km²) with a high frequency (every 15 minutes) and 11 spectral channels. The satellite is operated by the inter-governmental organisation EUMETSAT financially supported by seventeen European countries. METEOSAT's primary mission is to observe the evolution of weather systems for the purpose of operational meteorology.
The raw images are centrally processed (calibrated and navigated) at EUMETSAT in Darmstadt and after that transmitted to a large community of users through a communication satellite. The end users can display the images roughly 10 minutes after the observation of the European continent. The rapid delivery enables a monitoring of the atmosphere in near real time. No other observing system offers this wealth of information with this spatial coverage and rapid update cycle.
METEOSAT 8 also provides a high resolution visual channel, referred to as HRVIS. This channel has a sub satellite resolution of 1 x 1 km². As the collocation with the other channels was an issue, this channel is not incorporated in the cloud masking algorithm, but in a later stage it is an option.
The data is received at KNMI in the original HRIT format, and then processed to the local KNMI hdf (hierarchical data format). The hdf product is distributed to the various end-users.

3.2 HiRLAM data

The HIgh Resolution Limited Area Model (HIRLAM; Gustafsson, 1993) is a numerical weather prediction (NWP) model developed at the national weather services of Denmark, Finland, Iceland, Ireland, Sweden, Norway, Spain and the Netherlands. Each member state runs its own version dedicated to the local region of interest.
HIRLAM as operated at KNMI covers Europe and the North Atlantic. Its' grid is based on a shifted longitude latitude grid. This shift causes the grid cells to be rectangular in the main area of interest, Western Europe. The horizontal grid point distance is 0.2 degrees corresponding to 22 km. The results of HiRLAM used in this project are the products from the three hourly runs.
In preparation of the forecasting work HIRLAM generates analysis fields of observations (e.g. synops, radio sondes) taking into account surface type and orography. HIRLAM surface temperature analysis fields, representative for the parameterised soil top layer, are used in this project.

The HiRLAM output is also stored in a long term archive. Like every NWP model it has frequent updates as new insights are implemented in order to improve the model performance. The latter implies that a long term continuity of parameters is not guaranteed. A new parameterisation may cause a jump in e.g. the surface temperature. One should be aware of these updates when performing a long term comparison study.

3.3 SYNOPS data

Synops data are values of meteorological parameters observed or measured from the ground. To derive standardisation of measurements the type, definition and quality of the parameters, time of observation and requirements to the location of the station are defined by the World Meteorological Organisation (WMO). Each national weather service is responsible for a network of stations from where observations and measurements are done regularly at fixed times (hourly, 3-hourly or 6- hourly). In the Netherlands 15 synops stations at land are operated by KNMI. In other European countries the synops networks are mostly less dense. At sea the situation is even less favourable. From a limited number of platforms in the North Sea synops are observed. Also synops from “selected ships” sailing the seas and oceans all over the world are available. The latter reports are randomly distributed and not continuous in space. Synops are in near real time available for the weather forecasters via the Global Telecommunication System (GTS) of WMO. KNMI archives all synops observed around the world.
For validation of results synops observations of cloud cover are most valuable. Until recently cloud cover is mostly observed by qualified human observers (both at land and at sea). An observer views the total sky and estimates the percentage of the sky which is covered by clouds. This percentage is converted to a measure called okta. 2 okta means 2/8 part (25 percent) of the sky visible for the observer is covered with clouds. 0 okta means the sky is completely cloud-free. 8 okta means the sky is completely cloud covered. The 1 and 7 okta classifications are different. The observer has to report 1 okta when a cloud is visible even when it comprises only 1 percent of the total sky. The same applies for 7 okta but then for a hole in the cloud covered sky.
The area for which the observation is valid is related to the atmospheric conditions (visibility), the amount of obstacles around the observer (mountains, trees, buildings, etc.) and the height of the clouds (the higher the larger the area). At an average it can be assumed that the area viewed by an observer is around 30 x 30 km². At sea the area is larger (50 x 50 km²) as view obstructing features do not occur.
More and more automated stations are used within Europe. On these stations the upward looking lidar system, called ceilometer observes one point above the station for cloud reflections continuously. By rationing the period with a cloud signal to the total observing period cloud coverage is derived in okta’s.
The automated system has problems with high semi-transparent clouds. Also in situations with broken clouds and low wind speeds the derived okta’s can differ from a human observer report. It is difficult to asses the volume of air observed by the system as this is partly determined by the wind speed at cloud level. But in any case is the volume of air observed by the automated system far less than the volume observed by a human observer.
Although the WMO has set very strict quality demands to the synops measurements in practice the quality of cloud cover synops varies considerably. The estimation of cloud cover in the range of 2 until 6 okta appears to be a very difficult task and the result is very much related to the experience and education of the observer. This problem with variable quality of synops is partly solved here within the validation work by not using

synops from stations or ships which are on a “grey list” of bad quality synops stations. This grey list is generated at KNMI based on experiences with the use of synops during many years.
Approximately 50 synops stations at sea and 1000 at land within the studied area are used for the validation of the results.

3.4 NWC-SAF data

As described in section 2.3 EUMETSAT initiated so-called Satellite Application Facilities. The SAF relevant for this project, the SAF on Now casting and short range forecasting produces a Cloud mask. Le Gleau and Derrien, 2003.

When the project was applied it was foreseen that the NWC-SAF products would take more than a year to become available. Due to the postponement of the decision of funding the NWC-SAF products became available during the development phase of the cloud masking algorithm.
It was decided to compare the results of the developed cloud mask to the NWC-SAF products, and it was beneficial to incorporate the knowledge of the NWC-SAF community.

The cloud masks generated by the NWC-SAF are stored in archive since version 1.2 became available in May 2005.

3.5 Data storage

The archiving of all the data is a significant effort. Especially data continuity proved to be a problem.
At KNMI the METEOSAT 8 data is stored in a locally defined hierarchical data format.
http://www.knmi.nl/onderzk/imageformat/bik_web/bik_hdf5/hdf5_intro.htm
Throughout the project the stored area is redefined in accordance with the wishes of KNMI forecasters. This causes changes in the size of the stored area in the archive.
The software has to account for these area changes.

The analyses of HiRLAM used in the project are calculated for an area compassing the studied area, see Figure 3.1. Throughout the project it appeared that the stored area in the long term archive is considerably smaller (20 percent of the area, Figure 3.2). This was unforeseen and limited the possibilities to perform runs on archived data.

The storage of the synops information is consistent and nearly continuous. There are minor gaps due to system failures. Synops information is filtered for clear outliers.

Throughout the project it was realized that the NWC-SAF cloud mask is a useful reference mask. The storage of this cloud mask only started towards the end of the project. This limited the comparison period.

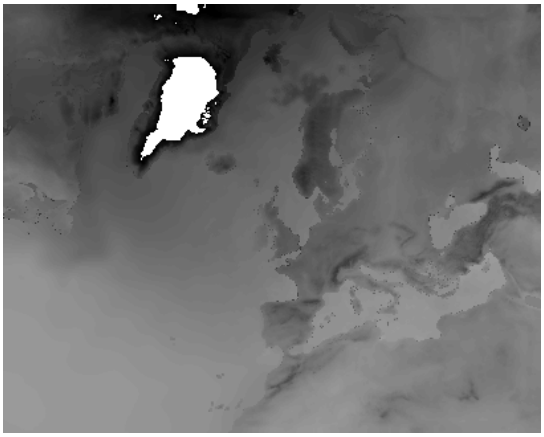

Figure 3-1 Arial coverage of the HIRLAM model. Shown is the surface temperature.

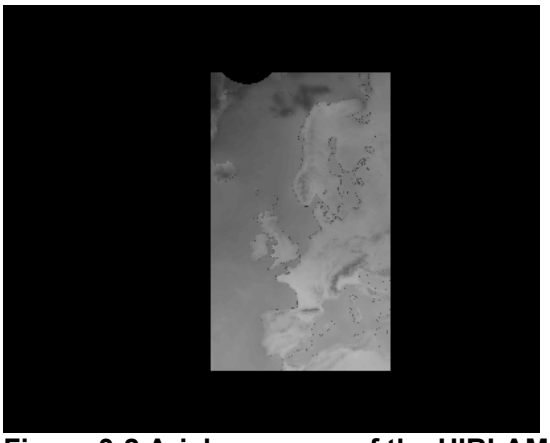

Figure 3-2 Arial coverage of the HIRLAM model in archive

4 Development of the algorithm.

The primary objective of the project is to produce a cloud mask with a high level of confidence. This cloud mask serves as a base for further products. One should however realize that the intended use of the cloud mask affects the detection method and results. Users requiring a high probability of cloud detection and accepting erroneous clear situations have interest in another cloud mask than users requiring a high probability of clear situations and accepting erroneous cloud coverage. The present goal of the algorithm is to detect clear and cloudy situations both with high detection efficiencies.

4.1 The area coverage

The European weather predominantly originates from the Northern Atlantic Ocean. To support the use by forecasters and modellers the algorithm should have a good coverage in this area. For now casting of present weather a good coverage of the European continent is necessary.
The most effective cloud detection method applied here is based on the comparison between the HiRLAM surface temperatures to satellite observed brightness temperatures.
The area for cloud masking in this study is based on the HiRLAM area coverage, the considerations given above and the limitations in viewing geometry of METEOSAT 8. The latter argument refers to the limited view of the METEOSAT 8 of the Polar Regions. Simple geometric calculations show that the METEOSAT 8 can not observe the earth above 81 degree latitude North or below 81 degree latitude South.

Figure 4.1 shows the present study area. The Figure shows a cloud cleared composite of surface reflection.

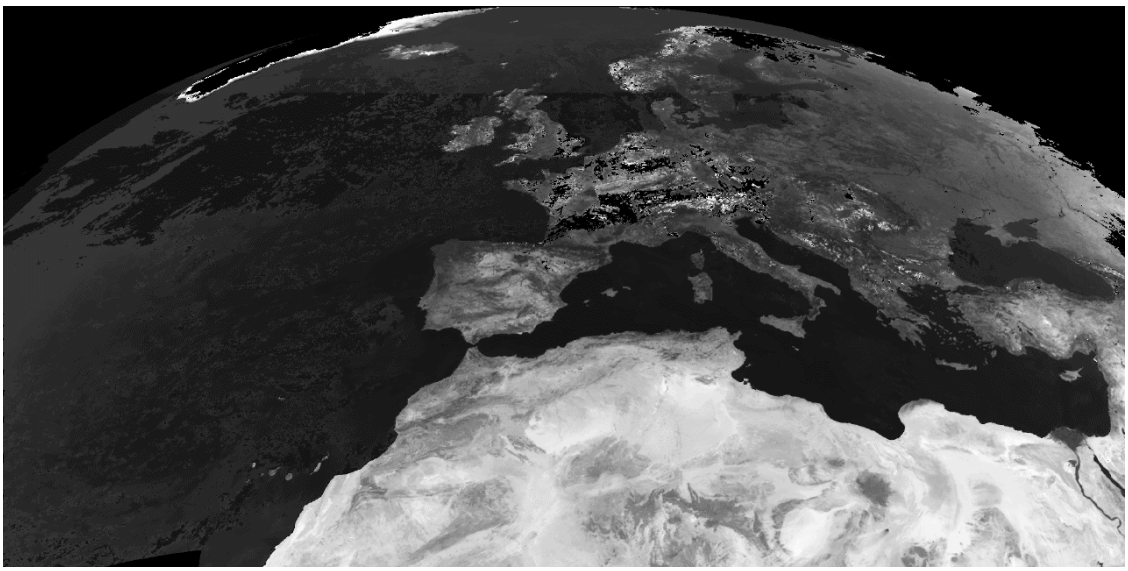

Figure 4-1 Cloud cleared composite of surface reflectance in August 2005. Some dark spots indicate persistent cloudiness during this period. In this cloudy area no surface reflectance could be determined.

Within the area four different surface types are distinguished, sea, coast, land and mountains. Coast areas are defined as those satellite pixels containing both sea and land. Also the direct adjoining pixels are classified as coast. Mountains areas are defined here as areas with an average height within a pixel above 800 meter. For the definition of each surface type the Gtopo30 data base (gtopo 30) was used. Each type has its' specific characterization in reflection and daily temperature cycle. For each type different thresholds are determined here.

4.2 The thresholding tests

The algorithm comprises a number of methods to classify a pixel as cloud contaminated or not. The most efficient test is a comparison between observed brightness temperatures and NWP forecasted (or analysed) surface temperature. During sun illuminated conditions comparisons between the visual channels and cloud cleared surface reflection maps are performed. A number of multi-spectral thresholding techniques are applied additionally. In the following sections each test is described together with the associated theory.

4.2.1 T10.8 versus HiRLAM surface temperature

The 10.8 micron channel referred to as T10.8 is compared to the surface temperature of HiRLAM. Similar to the MetClock system a temperature difference field is determined as described in section 2.2. This temperature difference field, referred to as Tdif, partly accounts for the difference between model surface temperature and observed brightness temperature. The Tdif also comprises atmospheric radiation contributions.

The test flags a pixel as cloud contaminated as:
T10.8 < Thir + Tdif – Threshold (surface, satellite zenith angle, Tdif)

The threshold is a function of the surface type, satellite zenith angle (SZA), and the availability of Tdif. Water vapour can affect the brightness temperature. The contribution to the signal of water vapour is therefore related to the atmospheric path length which is a function of SZA. There are parts in the area under study, e.g. on the ocean, with no synops stations, where no Tdif can be determined. For these areas higher thresholds values are used to minimize erroneous detections.

For a period of a week in August 2005 the T10.8 of correctly determined cloud coverage and clear situations are represented in a number of Figures as function of SZA and for two surface types, sea and land. The other surface type's coast and mountain do not give additional information relative to the shown results. They are therefore not shown here.
Two different time's 3:00 hr and 12:00 hr are shown to illustrate the occurring difference between day and night time situations.

In Figure 4.2 the T10.8 is shown for night time (03:00 hr) sea pixels as a function of cos (SZA). At the nadir point of the satellite the cos (SZA) equals one, towards the poles and the East and West sides the satellite zenith angle increases hence cos (SZA) decreases. The southern boundary of area chosen in this study (Figure 4.1) limits the cos (SZA) to values below 0.8.

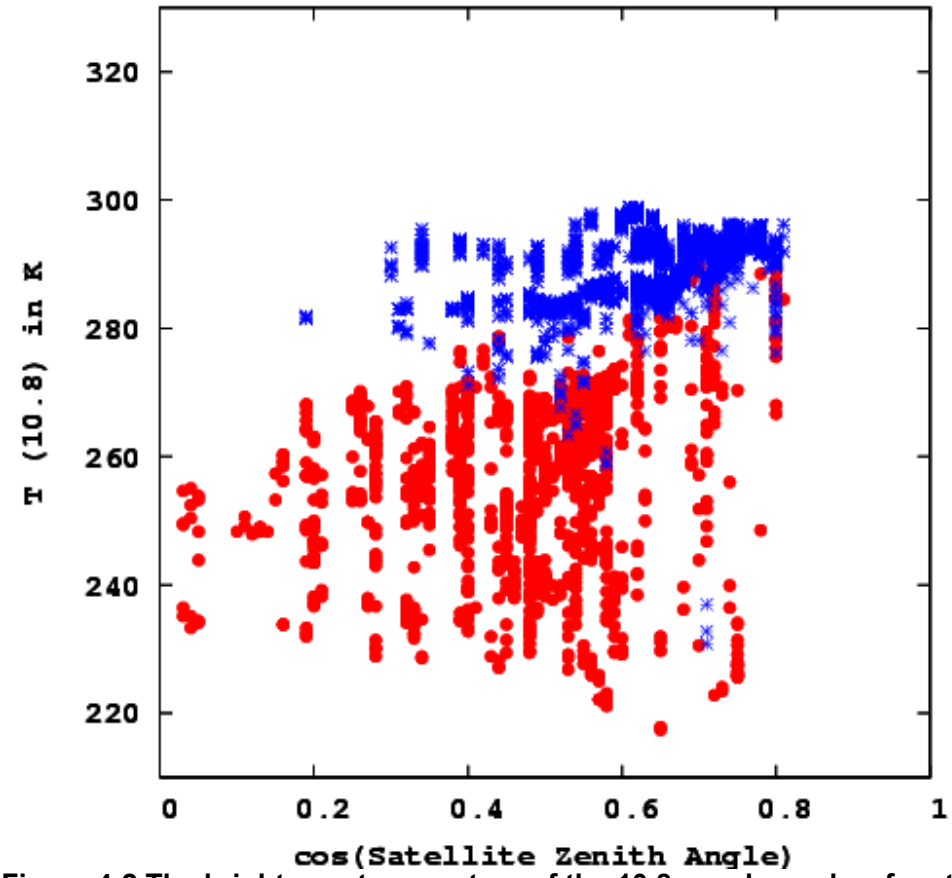

Figure 4-2 The brightness temperature of the 10.8 µm channel as function of the cosine of the satellite zenith angle. Observations are made from 8 to 16 August 2005 at 3:00 hr GMT over sea surfaces only. Blue stars indicate clear skies and red bullets indicate cloud coverage.

Two pools of pixel data are shown, for clear and cloudy cases, both correctly identified and observed cases. The two pools are selected from the total of synops observations of this selected week. For the explanation and examples of the thresholding technique applied here for cloud detection we limit the pools to the cases with high confidence in cloud and clear detection. Chapter 5 describes the method of pixel selection and the validation applied in this study in more detail.
As expected the cloudy pixels have in general a lower temperature compared to the clear pixels. There are however some observed clear pixels with a low temperature. This could be due to poor synops observations.
The pixel values also appear to be clustered around certain cos (SZA) values. This phenomenon is related to the non random distribution of the synops stations within the field of satellite view. The selected week of synops data also seem to have no clear cases with a high (SZA) value. This is an artefact due to the selection of sea cases as a first result. Sea cases have generally a lower number of synops observations. In Figure 4.5 for land cases with a higher number of observations the artefact does not occur.

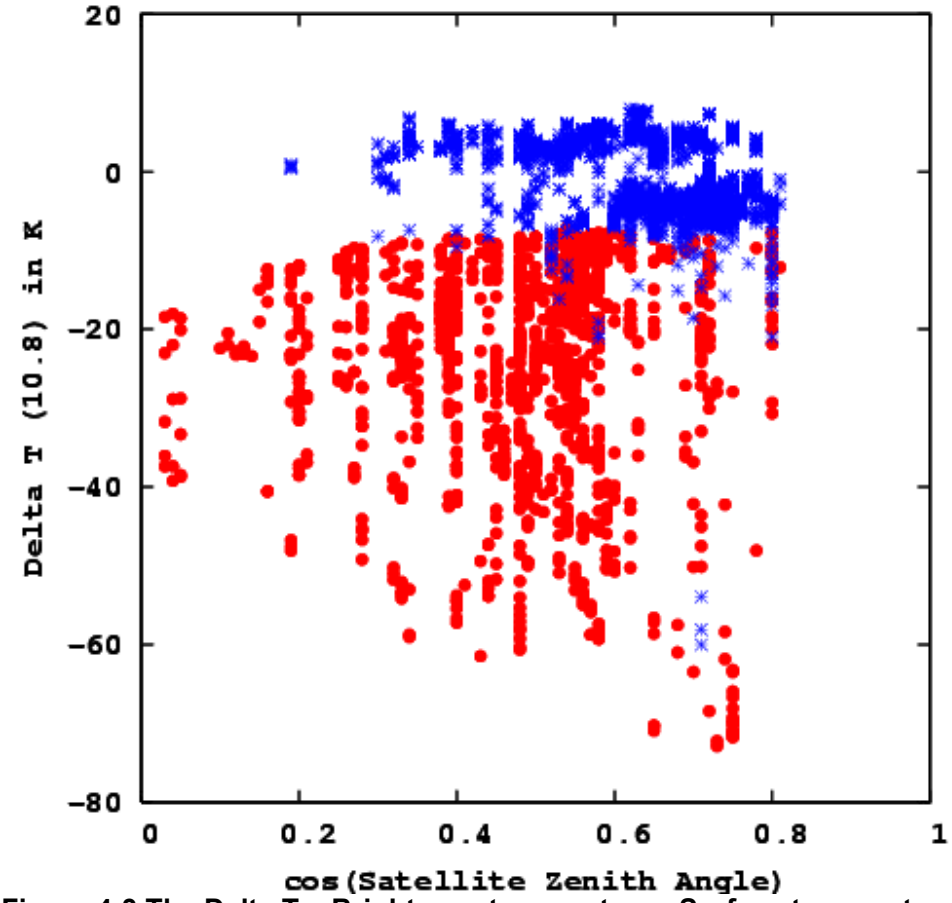

Figure 4-3 The Delta T = Brightness temperature - Surface temperatue from NWP + Tdif for the same cases as in Figure 4.2. The symbols represent the same as in Figure 4.2

In Figure 4.3 the difference delta T which equals T10.8 - Thir + Tdif is shown. The value of delta T for clear cases is around zero, but with a certain bandwidth. This bandwidth is an indication for the threshold to be applied here for cloud detection. Also the exceptional low temperature values for (erroneous) clear cases appear also here.
The data shown are the selected cases from the two data pools. In Figure 4.4 all the synops observations are merged into the Figure given in grey underneath the selected cases. The grey points illustrate the complexity of cloud detection by thresholding. The appearing grey symbols represent the pixel values which are not correctly detected as cloud or cloud free or have a synops of cloud coverage between 2 and 6. The numbers of occurrence of the two data pools versus the grey points are denoted in the figure caption. The numbers indicate that the defined clear and overcast cases represent more than forty percent of the occurring cases. The clear sky pool is dominant for this night case. For the day time cases the overcast sky has higher occurrence likelihood, see the numbers in the Captions of Figure 4.4, 4.5, 4.6 and 4.7. The shown period is a summer period with a lower chance of cloud occurrence.

In Figure 4.5 The delta T of pixels for land observations are shown as a function of cos (SZA).

The observation network of stations on land has a higher density. Hence the data has a higher density in comparison to sea synops. Apart from the higher density the conclusions which can be drawn from the Figure are similar to the sea case. Also here there are pixels with unrealistic low temperatures in clear conditions, possibly related to erroneous observations. A merging with all the other data is not shown because the higher density of data results in a messy Figure while the conclusion is the same as for Figure 4.4.

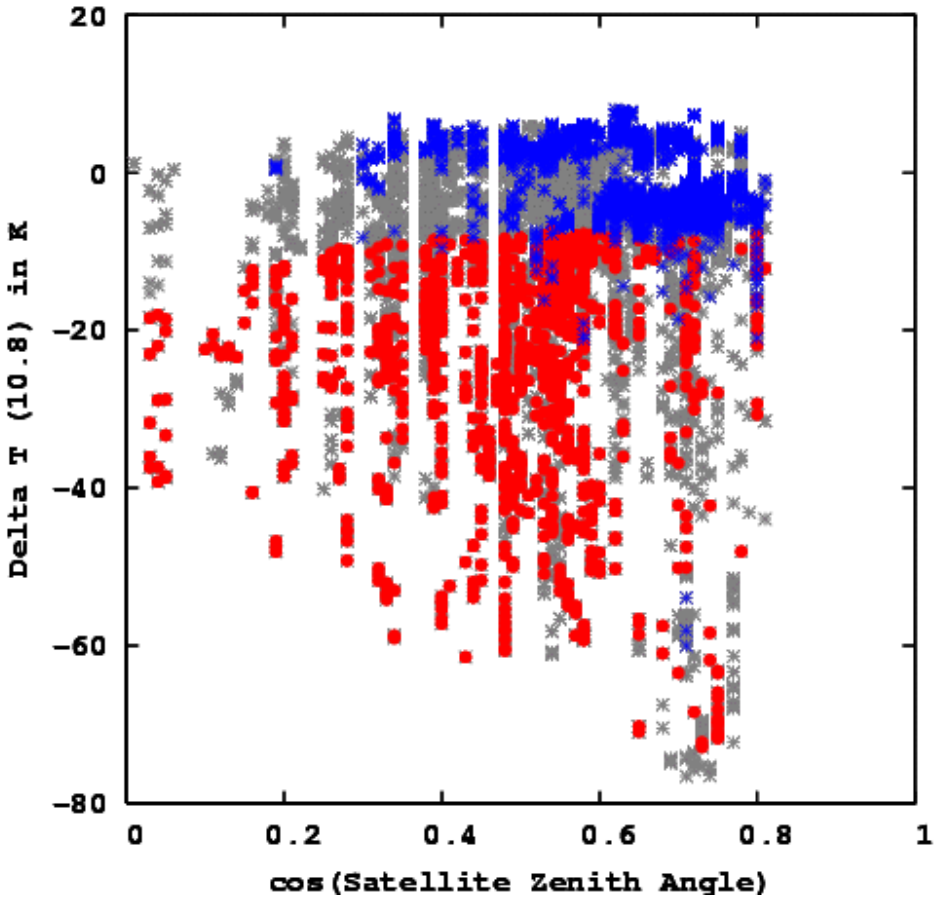

Figure 4-4 Caption is as for Figure 4.3 but added are all synops observations give in grey(10537), blue(3088) red (1326).

The night cases are relative simple in comparison to daytime cases. During day time the extra information available in the solar reflection channels allows for a better cloud detection.
At the same time the heating due to sunlight results in a wider spread in surface temperatures. This hampers the discrimination between clear and cloud pixels, forcing the thresholding technique to apply larger threshold values to minimize erroneous classifications. The increase in threshold values will give a decrease in successful detection of overcast and simultaneously increase the (successful) detection of clear sky. The wider spread of surface temperatures renders day time cloud detection more complex.

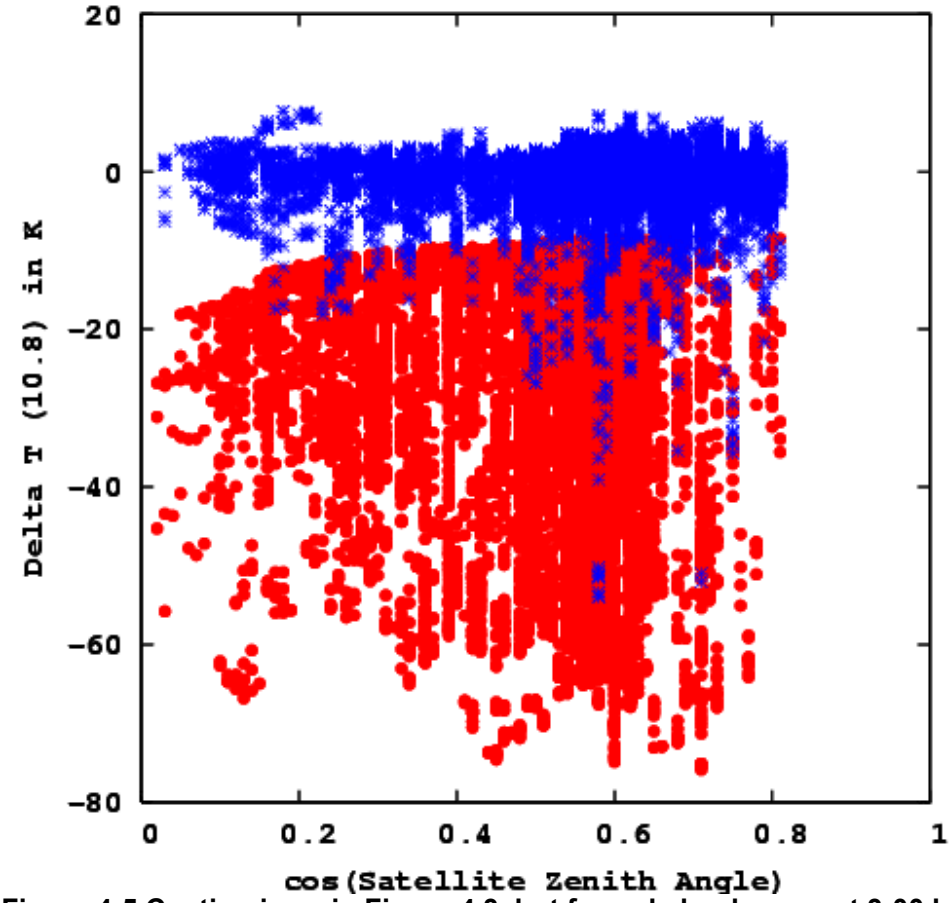

Figure 4-5 Caption is as in Figure 4.3, but for only land cases at 3:00 hr GMT. Clear (24624), overcast (11309) total (77262)

In Figures 4.6 and 4.7 the delta T at noon for sea and land respectively are shown. Note the delta T bandwidth for clear cases over land which becomes lower for lower cos (SZA) values. This is simply caused by the heating by the sun which is higher at lower latitudes, leading to wider spread of delta T values.
The day time synops observations appear to be of a higher quality. There are less extremely low temperatures for the clear data pool, indicating that the synops did not miss a cloud.
The distinction in delta T between clear and cloudy cases is less pronounced compared to the night time cases. The detection using the solar reflection channels contributes to the overall (using all tests) detection and classifies pixels which can have a less pronounced delta T value.
The bimodal distribution in the day time over sea for clear cases is probably caused by the frequently missing of Tdif. Tdif is depending on synops reporting clear situations. At sea with a lower density observation network there is a higher chance on data gap occurrences leading to missing Tdif and hence jumps in delta T. The detection using the solar reflection channels still classifies the pixel correctly, but when incorporated in Figure 4.6 the lacking of Tdif manifests itself in a bi-modal distribution.

In Figures 4.8 and 4.9 two additional data pools are shown, again to illustrate the complexity of cloud detection. For land observations the pool of cloudy situations not

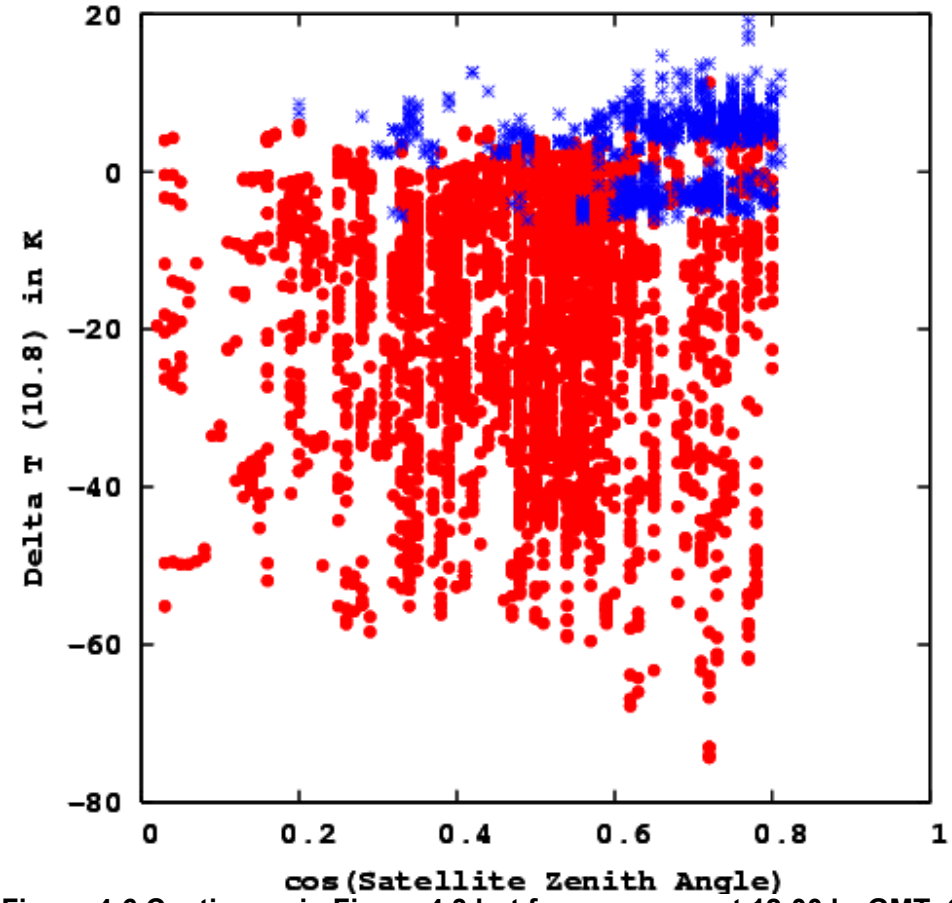

Figure 4-6 Caption as in Figure 4.3 but for sea cases at 12:00 hr GMT. Clear (1306), overcast (4562) total (11936)

detected by the algorithm is shown in 4.8. In 4.9 the clear situations classified as cloudy by the algorithm. In Figure 4.8 all the erroneous classified pixels have a higher delta T than the prescribed threshold value. In Figure 4.9 the apparent erroneous classified pixels are more randomly distributed. In the latter data pool erroneous synops observations could be included, with exceptional low delta T values.

The threshold values for different surface types are determined empirical from the bandwidth of the delta T of clear skies. They are given in table 4.1

Table 4-1 The threshold values in Kelvin for the T10.8 µm versus NWP surface temperature test.

Threshold values for T10.8	Sea	Land	Coast	Mountain
day	-5-1/cos (SZA)	-10-1/cos (SZA)	-10-1/cos (SZA)	-5
Night	-6-1/cos (SZA)	-7-1/cos (SZA)	-7-1/cos (SZA)	-5

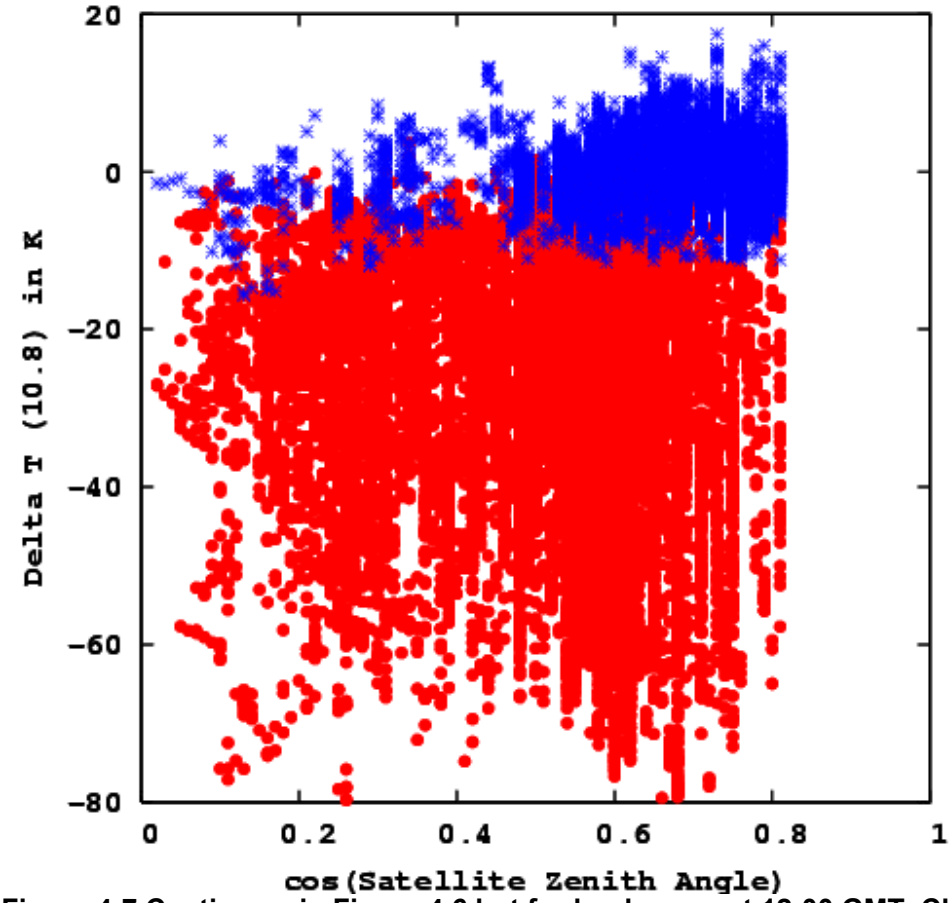

Figure 4-7 Caption as in Figure 4.6 but for land cases at 12:00 GMT. Clear (7322), overcast (19767) total (59477)

During the evaluation process it was realized that there is also a latitudinal dependence of the threshold values. This dependence will be studied in more detail in the near future. It is not included in this project as the project time was not sufficient.

4.2.2 Refl006 Refl 008 versus surface reflection composites

The threshold determination for the channels 0.6 µm and 0.8 µm are treated here simultaneously. The techniques are very similar. The only difference is that for sea surfaces the 0.6 µm channel is used and for land surfaces the 0.8 µm. The ratio behind this is the higher contrast of clouds above sea in the 0.6 µm band and the better contrast of clouds above land in 0.8 µm.
Of course the reflections test are only applied in sun illuminated areas.

Similar as described for the MetClock system in section 2.2 a cloud cleared reflection map is derived using a week of images. Cloud contaminated pixels are filtered; the remaining pixels are averaged, resulting in a surface reflectance map. This method is applied to both channels 0.6 and 0.8 µm. An example is already shown in Figure 4.1.
It is possible that in wintertime no cloud free pixels can be found. Over a sea surface extrapolation is then possible from clear pixels to overcast pixels. Over land this is not

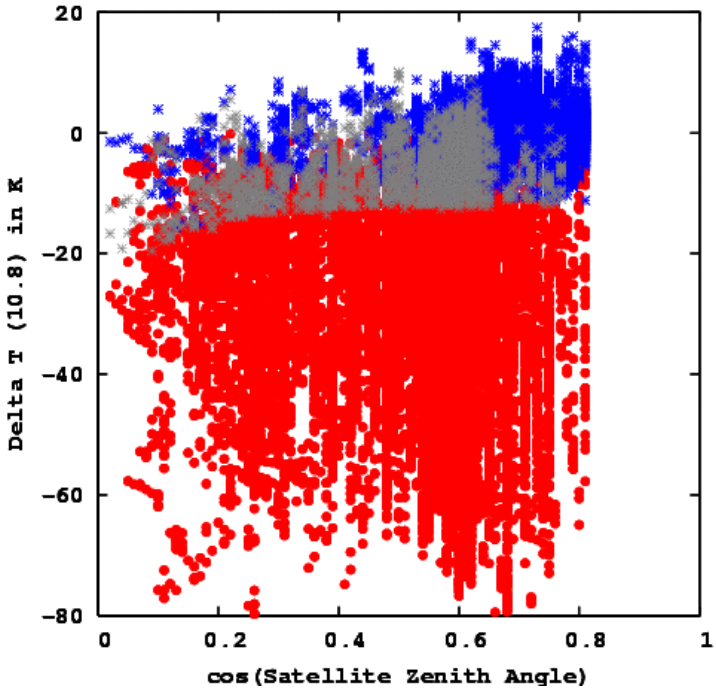

Figure 4-8 Caption as in Figure 4.7. Added in grey are the observations of clouds given by the synops and missed by the algorithm.

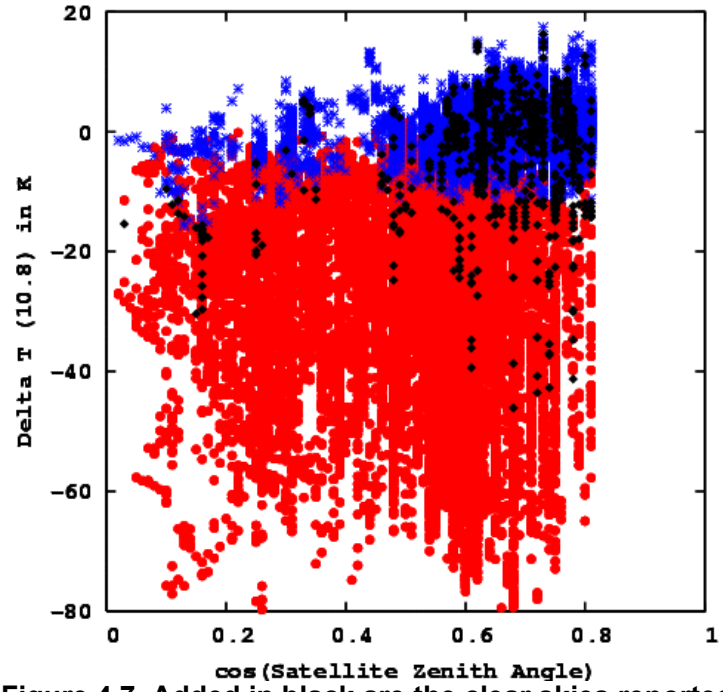

Figure 4-9 As Figure 4.7. Added in black are the clear skies reported by the synops and classified as cloud by the algorithm.

possible. The inhomogeneous land surface will lead to erroneous surface reflection maps. This can be solved by introducing a longer period for the gathering of cloud cleared pixels.

The test flags a pixel as cloud contaminated as:
R0.6 > R0.6_ surface + Threshold (SolZen, satellite zenith angle, R0.6_ surface)
R0.8 > R0.8_ surface + Threshold (SolZen, satellite zenith angle, R0.6_ surface)

Where the R0.6 test is only evaluated above sea and the R0.8 test is evaluated above land. Like in the T10.8 test the thresholds are a function of surface type, SZA, the Solar zenith angle (SolZen), and the surface reflection map. Where no surface reflection pixel values can be derived the threshold is higher.

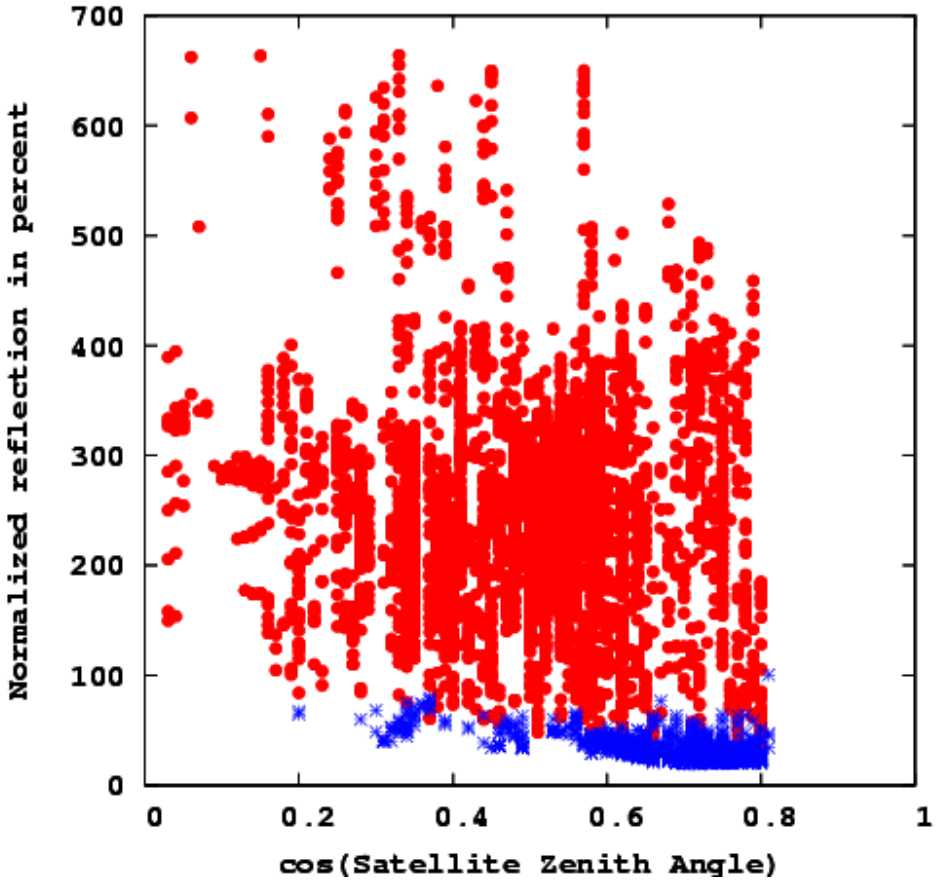

Figure 4-10 Normalized reflection (is reflection divided by cosine of solar zenith angle), as a function of the cosine of the satellite zenith angle for sea surfaces at 12:00 GMT. Bleu and red symbols are as in Figure 4.9

In Figure 4.10 the normalized reflection for 0.6 µm is given as function of the cosine of the SZA. The reflection is normalized by dividing the reflection by the cosine of the solar elevation angle, causing values to rise to values higher than hundred percent. The same data pools, clear and cloudy, as in section 4.2.1 are used.

The clear pixels have low reflection values; the cloudy pixel values have higher reflection values. They are easily distinguishable in this Figure.
In Figure 4.11 the comparison is made between the reflection values from the surface, derived from a week of cloud cleared images and the observed reflection values are subtracted from each other. This difference reflection is given as a function of the cosine of SZA. Due to the low reflection values of the sea surface the detection of clouds is very much feasible, using the 0.6 µm channel.

For the 0.8 µm channel cloud detection is more complex. In Figure 4.12 the Normalized reflection difference is represented for the two previous defined data pools as function of the SZA. Due to the higher reflectance of the land surface, the difference between clear and cloudy situations is less pronounced. In addition the surface reflectance composite can have gaps. This will cause high reflection difference values for clear cases. The threshold technique has to account for this deficit.
The low reflectance values for cloudy cases may look un-natural. However they can occur when high clouds cast shadows over the surface or over lower clouds.

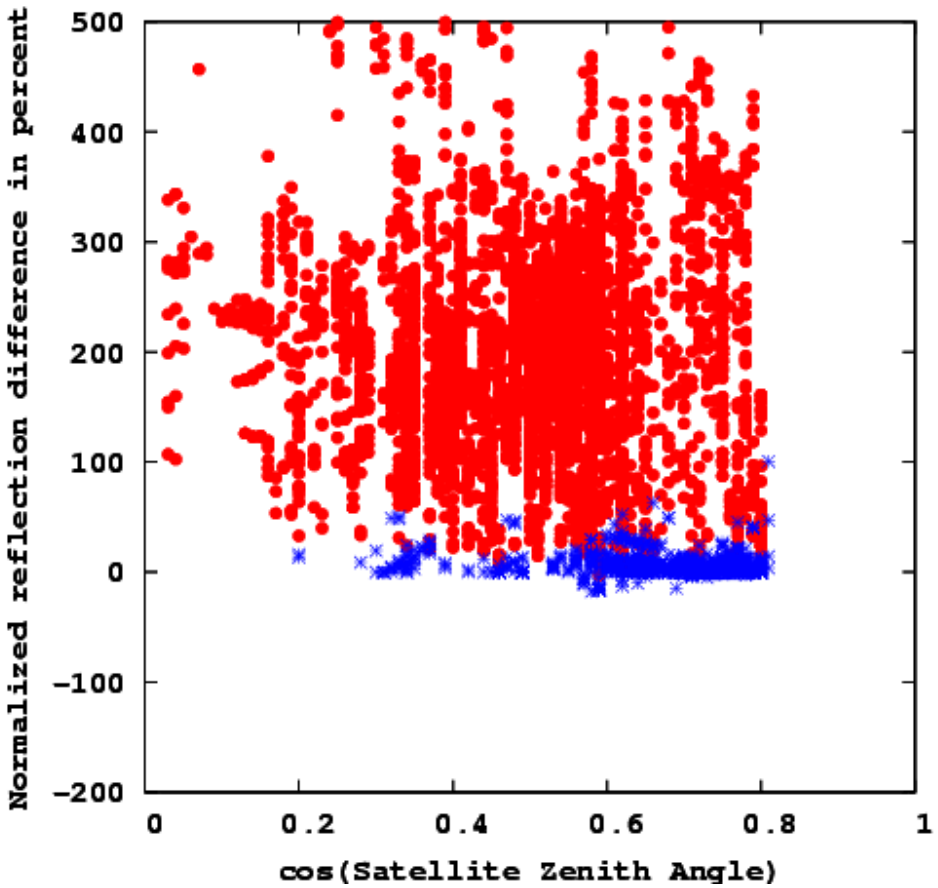

Figure 4-11 As Figure 4.11 but for the reflection difference (R0.6 µm - surface reflectance)

As cloud detection appears to be more complex in the 0.6 and 0.8 µm channels. The T10.8 is chosen as the primary test.

The threshold values are determined empirically from figures like Figure 4.11 and 4.12.

4.3 Additional tests.

The above described two tests are very efficient. Together they cover more than seventy five percent of the probability of detection as will be shown later in the validation section.

The following tests aim to contribute to the cloud detection but simultaneously to contribute minimal to the number of erroneous detections. This means a different strategy. Whereas the first two tests are intended to capture the clear and cloudy situations with a high confidence, the strategy for the following tests is only to fine tune the detection, and concentrate on those clouds missed by the tests T10.8 – Thir and Refl006/Refl 008. Simultaneously the additional test should have a minimum contribution to the error.

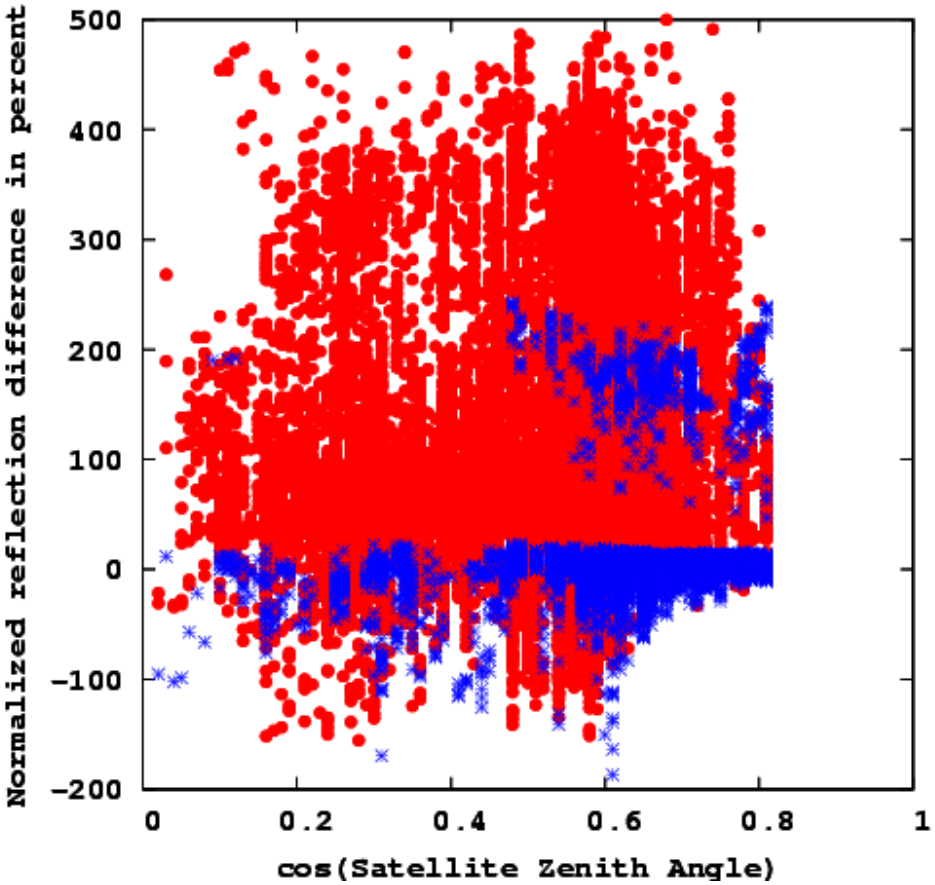

Figure 4-12 As Figure 4.11 but for land surfaces and R0.8

The additional tests described below will therefore be applied when the T10.8 test was NOT successful. The clouds targeted here are optical thin clouds and clouds close to the surface having a relative high top temperature. These clouds show little contrast in the infra red channels with the surface temperature.

During night time the METEOSAT 8 offers more channels than the METEOSAT 7 has, thus enabling possibilities to extract extra information about clouds and detecting them.

4.3.1 T10.8 T12.0 thin cirrus over sea

Thin cirrus clouds have a significant signal in the difference between T10.8 and T12.0. However the signal can be contaminated by all kind of other contributions, especially above land.
In the evaluation and fine tuning period this test appeared only to contribute positively to the results during day time over sea surface. Over land the signal disturbances due to various causes contributed too much to erroneous classifications,

So for sea pixels during daytime the test flags a pixel as cloud contaminated as:
T10.8 – T12.0 > 2 K and T10.8-T12.0 < 4 K

4.3.2 T108-T039

The low clouds emit a significant signal at night in the T03.9 channel. This is used here to detect the lower clouds.
The T03.9 is a unique channel. During night time it behaves like an infrared channel. However when there is sunlight the clouds also reflect in the 3.9 µm channel. The 3.9 µm observed radiance is then a combination of reflection and emission. To evaluate the signal the radiance should be split into an emission and a reflection part. This splitting involves radiance calculations with radiance transfer models. This extra modelling is not suitable for the straight forward threshold technique applied here. In addition to this all radiation models have their specific uncertainties, rendering a straightforward approach even more difficult.
This interpretation complexity is avoided by not using solar illuminated T039 images in the algorithm.

The test flags a pixel as cloud contaminated as:
T10.8-T03.9 > Threshold (SZA, T10.8, T03.9)

The threshold is depending on the SZA, There are additional conditions here to assure that the test contributes to correct low cloud detections. The conditions are that T10.8 should be warmer than Thir + Tdif -8K and the T03.9 should be colder than Thir + Tdif – 7 K.

4.3.3 SD 108 SD 039

The spatial coherence of the observed field is used to detect small clouds or cloud edges. This test is well known and well documented, Kriebel et all (1989). A significant difference in the standard deviation (SD) within an area of 3 x 3 pixels can indicate cloud contamination. Over a homogeneous surface temperature field, e.g. sea a larger SD can indicate a cloud. Over an expected inhomogeneous scene, e.g. a coastal area a small SD can indicate a homogeneous layer of stratus.

The heterogeneity of the surface, especially over land, complicates the cloud detection. Combining a test on both the SD of the T10.8 and T3.9 enables a more robust testing of the SD. Here again conditions are set for the T10.8 and T3.9 to focus on the low clouds lacking a clear temperature difference with the under-laying surface. The use of the T03.9 µm limits the use of the test to night time conditions.

The test flags a pixel over sea or land as cloud contaminated as:
SD10.8 > Threshold (SZA, T10.8, Surface type) and
SD (10.8-3.9)> Threshold (SZA, T3.9, Surface type)
Under the condition that
T10.8 > Thir + Tdif -8 and T3.9 < Thir + Tdif – 7

During the development phase no positive impact of this test on the results could be found. A more thorough investigation is required here. We do believe in the potential of this test and will pursue this subject further. But for the present study this test was not applied.

For coastal regions during night time the SD was evaluated in a slightly different way:
The test flags a pixel as cloud contaminated as:
SD10.8 < Threshold (SZA, T10.8, Surface type) and
SD (10.8-3.9) <Threshold (SZA, T3.9, Surface type)
Under the condition that
T10.8 > Thir + Tdif -8 and T3.9 < Thir + Tdif – 7

For coastal regions this test contributed positively to the detection of low shallow clouds.

4.4 System requirements and development

The requirements for the system for the algorithm were higher compared to the systems requirements for MetClock. The input files and output files are larger. More tests are introduced demanding more computer resources. Still the system is relatively modest and can be implemented on standard pc.

4.4.1 Development

The algorithm development could rely on experience gained in the MetClock project. There were, however, some issues to solve. The tuning of the thresholds was more complex. The chosen strategy where two tests are the dominant tests and the other tests are only intended to have a marginal contribution to the error required a careful recalibration of the whole system as soon as one of the dominant tests was changed. The collocation of pixels relative to the HIRLAM grid had to be re-calculated as METEOSAT 8 has a smaller pixel size compared to METEOSAT 7. This also applied to the collocation with the synops stations positions.

Next to the program errors found in software in the development phase all the above described issues required a substantial amount of effort.
During the project the NWC-SAF software became available. As this package provides cloud type and cloud top height in a thorough and elaborate way it appeared better to

use these products of the NWC-SAF package and not to double the efforts of the NWC-SAF.

4.4.2 Portability of the software.

The software was developed and implemented on two different platforms, a UNIX and a Linux environment. This simultaneous development introduced a number of complexities, solved throughout the project duration.
The algorithm is developed in the program language 'C'. The MetClock algorithm was also in 'C', so it is a logical continuation. The scripts to prepare input and run the MetClock program for extended periods were written in C-shell. As C-shell appeared to behave differently on different platforms and as the Linux environment does not work with the C-shell, the Korn shell was chosen as the scripting language.
The use of scripts enabled automated invocation of the involved programs by the 'CRON daemon'. Automated invocation is required as the system should produce results continuously, including weekends, and nights. The programs have to be scheduled carefully. Some parts can run every 15 minutes when a new METEOSAT 8 image is available. Other parts like the storage of the synops data or the calculation of the Tdif and surface reflection maps can be run once per day. For computer demanding calculations off-office time are preferred.

The algorithm produces a number of intermediate files, surface masks, temperature mapping files from NWP fields to satellite pixels, temperature correction files, etc. used to obtain the final results. These intermediate files are sensible for Big and Little Endean use, a difference existing between Unix and Linux platforms. A special routine was introduced for performing byte swaps required to assure platform portability.

The compilation of the algorithm on two different platforms was a beneficial exercise. It naturally involved two different compilers with different performing characteristics.
One of the two compilers allowed for passing past array boundaries. This error was found only after compilation of the software on the second platform. The software became more robust and portable due to the development on two platforms.

For portability it is relevant that commands exist in both the UNIX and Linux operating systems. Not all commands are available in both environments. E.g. the Unix 'compress' command does not exist in Linux. It was replaced by 'gzip' which is available in both operating systems.

5 Validation Method

To evaluate the performance of the algorithm and to assess improvements a validation of the product is required. For the comparison between the cloud mask and the NWC-SAF products it is essential to have a common validation tool.
Within the NWC-SAF a validation method is described. This method resembles the method applied in the MetClock validation. It was decided to use the NWC-SAF method, Le Gleau and Derrien, 2000, as it facilitates communication with other research groups about efficiencies of the algorithm. The method is described below.

For ground truth in the validation work the synops is used. As mentioned in the data part there are deficiencies associated with the synops reports. However, it is the still best available and accessible data base to validate cloud masking. It provides a large coverage and has continuity. Again it is stressed that the validation is done using a ground based observation with a very different viewing geometry, including the difference between the automated and human observations, versus a satellite based observation. One should be aware of the shortcomings of the synops when interpreting the validation results.
The synops reports are available every hour but not every station reports hourly or even three hourly. To minimize the effect of the spatial incoherence caused by hourly vs. three hourly reporting stations it is decided to evaluate the algorithm results on a three hourly basis.

For a 3 by 3 pixel area around the pixel with the reporting station in the centre the number of pixels flagged as cloud contaminated are counted.

The definitions of clear and cloudy cases are based on synops observations:

- Observed cloudy when the total coverage is larger or equal to six okta. The corresponding Cloud Mask (CM) is identified as erroneous when the CM classifies clouds in less than four pixels.
- Observed clear when the total coverage is below or equals to two okta The corresponding CM is identified as erroneous when the CM classifies clouds in more than five pixels.
- Detected cloudy when CM classifies clouds in more than five pixels.
- Detected clear when CM classifies clouds in less than four pixels.

This means that the stations reporting three to five okta are neglected in the validation. Also the synops stations where four or five of the collocated pixels are cloudy (or clear) are not reflected in the validation results

A contingency table convention is created defining the probability of detection (POD) the False Alarm Rate for clear skies (FAR0) and the False Alarm Rate for cloudy skies (FAR8) according to table 1:

Table 5-5-1 Contingency table convention to define POD, FAR0, and FAR8

	Detected cloudy	Detected clear
Observed cloudy	Na	Nb
Observed clear	Nc	Nd

POD is 100 x (Na + Nd) / (Na + Nb + Nc + Nd) correct performance of the algorithm
FAR8 is 100 x (Nb) / (Na + Nb) underestimation of clouds by the algorithm
FAR0 is 100 x (Nc) / (Nc + Nd) overestimation of clear skies by the algorithm

The POD, FAR0, and FAR8 are calculated in percent for the different surface types.

6 Validation results and conclusions

The results of cloud masking are evaluated for a number of periods throughout the year. As indicated an evaluation run for a longer period was hampered by the limited area of surface temperature archiving, see Figure 3.2 in section 3.5. To discuss the performance of the software relative to the NWC-SAF results the comparison period is limited to the moment the latest version 1.2 became available until the end of the project.
In the first section of this paragraph the results of the various cloud masking tests are shown and discussed. In the second part the comparison to the NWC-SAF results are discussed.

6.1 Validation for separate tests

To illustrate the results of the algorithm a cloud mask (Figure 6.1) is shown together with a red-green-blue (Figure 6.2) composite of the original METEOSAT 8 observation for August 31, at 12:00 GMT.

Figure 6.1 shows a colour composite of the METEOSAT 8 1.6, µm, 0.8 µm and 0.6 µm in the red, green and blue combination. The combination discriminates ice (blue-cyan) from water (white-yellow) clouds. Land appears as green and sea and lakes as dark blue. Desert appears as red. Note the green Nile valley in the lower right corner just visible.

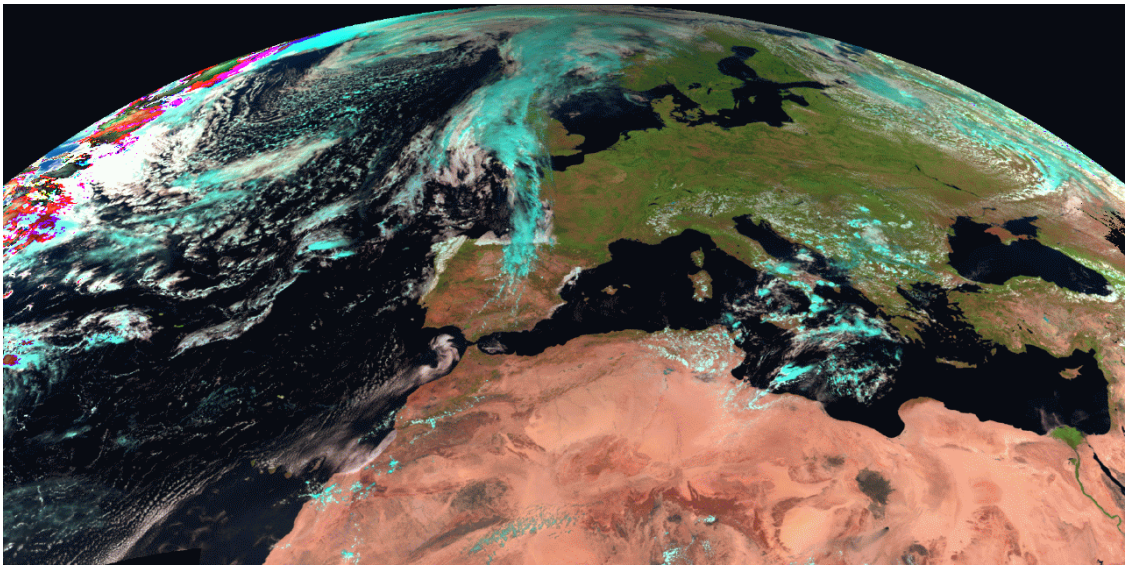

Figure 6-1 Red (1.6 µm) Green (0.8 µm) Blue (0.6 µm) Composite of the METOSAT imagery of August 31, 2005, 12:00 GMT

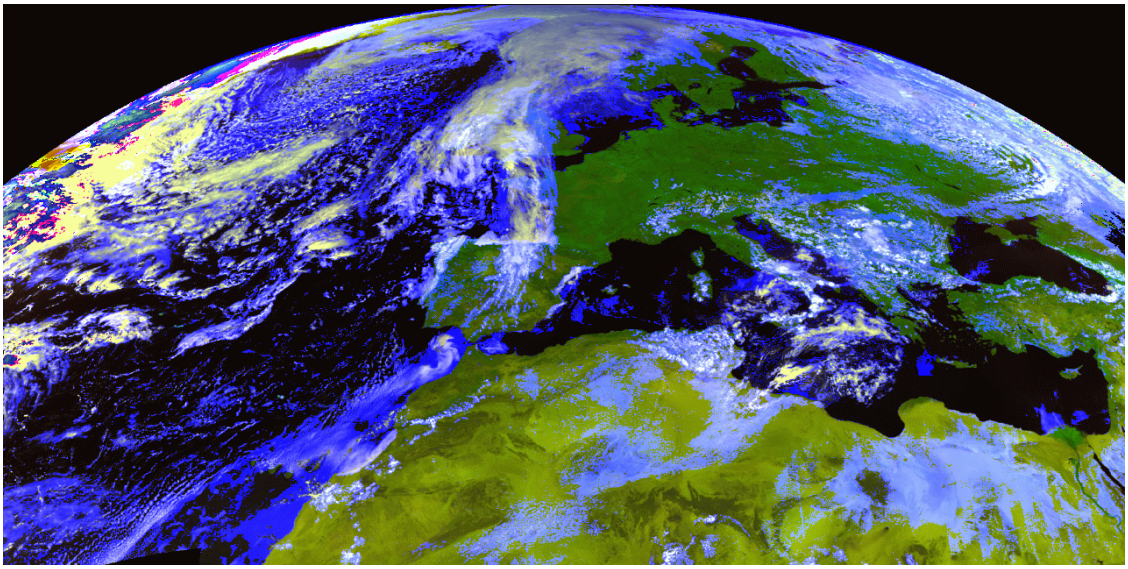

Figure 6-2 Red (0.6 µm) Green (0.8 µm) Blue (Cloud MASK) Composite of the METOSAT imagery of August 31, 2005, 12:00 GMT

In Figure 6.2. the 0.6 (red) and 0.8 µm (green) are blended with the cloud mask in blue. The pixels classified as cloud contaminated are represented in high blue values, in order to make them visible in the image.
From a visual inspection of the Figures it can be deduced that a majority of the cloudy pixels are classified correctly. Also visible is that for a large number of pixels classified as cloudy over Northern Africa and the Mediterranean e.g. for the coast of Egypt and near Crete it is not apparent from Figure 6.1 that the classification is correct. Still the chosen representation does not highlight thin cirrus clouds so the cloud mask could have added value here, especially over the Mediterranean. For Northern Africa dust particles could be the cause for the erroneous classification.
At the edges of the earth within the image the satellite zenith angle is too high. The chosen colour representation leads there to wild colours but they are meaningless for this study. In the validation algorithm the areas with a too high satellite zenith angle are not considered.
Other conclusions drawn from this Figures are in accordance with the conclusions to be drawn from the tables below, and are discussed below.

In the following tables the results of detection are presented. The numbers are derived based on the methods described in section 5. The results are differentiated on base of test types and day and night conditions. It is attempted to limit the amount of data shown and simultaneously still give an illustrative impression of the performance of the algorithm.

Next to the relative scores also the absolute numbers are shown. These numbers show the relative density of the various observations. The number of observations at sea and in mountain areas is relative low. Over land the number is high and coastal areas have a moderate number of observations. The low number of observations at sea and in mountains can affect the conclusions, especially as some synops observations have lesser reliability

The T10.8-Thir test is presented first in table 6.1. The test is tuned to be defensive in cloud detection as a significant detection contribution is expected from the Refl 0.6 and

Refl 0.8 tests. The defensive tuning is reflected in a relative high number in the missed clouds category (FAR8). A part of the high number in FAR0 for sea can be related to erroneous observations over sea. Still a closer study would be fruitful here, considering also the latitudinal dependence of the results.

Table 6-1 T10.8-Thir test results for various surfaces

August 19-31, 12:00 GMT	Na	Nb	POD in percent	FAR8 false cloud detection	FAR0 false clear detection
	Nc	Nd			
Sea	519	240	72	31	19
	82	349			
Land	3798	894	85	19	8
	320	3590			
Coast	810	320	83	28	5
	64	1080			
Mountain	313	116	83	27	4
	16	364			

The incorporation of the Refl0.6 and Refl0.8 thresholding test led to a higher POD but also to higher FAR0. Especially above sea this caused a problem. Above land and coast the POD, FAR8 and FAR0 are in a good balance.

Table 6-2 T10.8-Thir and Refl0.6 and Refl 0.8 test for various surfaces

August 19-31, 12:00 GMT	Na	Nb	POD in percent	FAR8 false cloud detection	FAR0 false clear detection
	Nc	Nd			
Sea	715	41	83	5	35
	151	275			
Land	4357	342	89	7	15
	587	3200			
Coast	1081	94	88	8	16
	176	899			
Mountain	399	41	87	9	15
	56	297			

Table 6-3 T10.8-Thir and Refl0.6 and Refl 0.8 test and T10.8-T12.0 test for various surfaces

August 19-31, 12:00 GMT	Na	Nb	POD in percent	FAR8 false cloud detection	FAR0 false clear detection
	Nc	Nd			
Sea	721	33	82	4	40
	171	252			
Land	4357	342	89	7	15
	587	3200			
Coast	1085	90	87	7	17
	184	869			
Mountain	399	41	87	9	15
	56	297			

Table 6-4 T10.8-Thir for various surfaces

August 19-31, 3:00 GMT	Na	Nb	POD in percent	FAR8 false cloud detection	FAR0 false clear detection
	Nc	Nd			
Sea	178	55	79	23	15
	31	163			
Land	2084	707	85	27	6
	244	3778			
Coast	606	174	84	22	10
	104	898			
Mountain	113	64	73	36	10
	10	89			

Table 6-5 T10.8-Thir and T3.9 test for various surfaces

August 19-31, 3:00 GMT	Na	Nb	POD in percent	FAR8 false cloud detection	FAR0 false clear detection
	Nc	Nd			
Sea	195	38	83	16	16
	31	163			
Land	2377	472	88	16	8
	324	3633			
Coast	680	103	86	13	13
	133	843			
Mountain	125	54	76	30	10
	10	89			

Inclusion of the T10.8-T12.0 did not improve the detection results over sea. One can consider omitting this test in an operational environment.

In table 6.4 the night time results are shown. Note the lower number of observations. At night there are generally less observations. Again the T10.8-Thir test is chosen to be defensive implicitly assuming that other tests will compensate the underestimation of cloud occurrence. The combination of T10.8-Thir test and T3.9 test results in table 6.5 do show a better cloud detection but also affect the FAR0 negatively and FAR8 positively.
The deviating lower score in the mountains can be partially attributed to the low density of observations.
For night time the inclusion of the SD3.9 test results only changed the Na, Nb, Nc and Nd numbers slightly and are not shown here. The impact on the POD and other scores were not significant in table 6.5.

From the results shown in the table's one can conclude that the detection is quite efficient. Generally more than eighty percent of the clouds are detected by the algorithm. Unfortunately the False alarm rates are high, especially FAR0. For operational purposes these numbers have to improve.
The question about the relative quality of the cloud masking can be partly answered by comparison to the NWC-SAF results shown in the next chapter.

6.2 Comparison to NWC-SAF results

The results of the NWC-SAF are incorporated, offering a unique possibility to compare the performance of the KNMI algorithm to results of a third party. The comparison is not completely independent as the methods of cloud detection are similar.
When interpreting the results one should be aware that the NWC-SAF software is tuned using synops of land only.

For a visual impression of the performance a NWC-SAF cloud mask is shown in Figure 6.3 for the same moment as Figures 6.1 and 6.2. Note that the NWC-SAF does not classify clouds over the desert, and the abundance of clouds around Sicily in comparison with the KNMI cloud mask result.

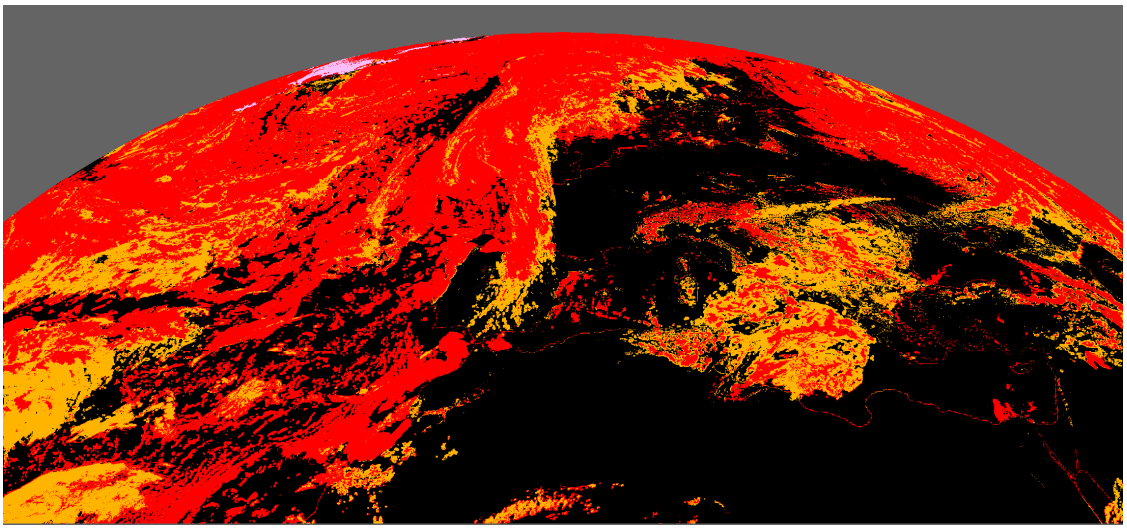

Figure 6-3 NWC-SAF cloud mask for August 31, 2005 12:00 GMT. Red indicates a cloud filled pixel, orange indicate a partially cloud pixel, black is cloud free.

The same method for calculation for POD, FAR8 and FAR0 are applied to the NWC-SAF results. The NWC-SAF software uses similar tests as the KNMI algorithm. It includes the test using T10.8 versus climate surface values and a test using Refl0.6 and Refl0.8. The NWC-SAF software package, however, applies more tests including the use of more spectral channels than the KNMI algorithm.

To enable a straightforward interpretation the NWC-SAF T10. 8 test is compared to the T10.8-Thir of the KNMI algorithm. The NWC-SAF results obtained with T10. 8 and Refl0.6 and Refl0.8 is compared tot the same combination of the KNMI algorithm.
The comparisons of the cloud masking results of both software packages using all tests finalize the comparison. All the comparisons are made for the same period and for same day and night time cases. In the tables the corresponding KNMI scores are incorporated to facilitate the comparison.

Table 6-6 NWC-SAF T10.8 test for various surfaces with the corresponding KNMI results in bold

August 19-31, 12:00 GMT	Na Nc	Nb Nd	POD in percent	FAR8 false cloud detection	FAR0 false clear detection
Sea	512 27	231 406	78 **72**	31 **31**	6 **19**
Land	3338 79	1309 3872	83 **85**	28 **19**	1 **8**
Coast	900 50	282 1138	85 **83**	23 **28**	4 **5**
Mountain	283 3	157 379	80 **83**	35 **27**	0 **4**

Table 6-7 NWC-SAF T10.8 and Refl test with the corresponding KNMI results in bold

August 19-31, 12:00 GMT	Na Nc	Nb Nd	POD in percent	FAR8 false cloud detection	FAR0 false clear detection
Sea	693 74	71 346	87 **83**	7 **5**	17 **35**
Land	4387 352	402 3511	91 **89**	8 **7**	9 **15**
Coast	1175 250	41 772	86 **88**	3 **8**	24 **16**
Mountain	432 54	25 320	90 **87**	5 **9**	14 **15**

Table 6-8 NWC-SAF All tests with the corresponding KNMI results in bold

August 19-31, 12:00 GMT	Na Nc	Nb Nd	POD in percent	FAR8 false cloud detection	FAR0 false clear detection
Sea	729 137	34 277	85 **83**	4 **5**	33 **35**
Land	4578 508	207 3273	91 **89**	4 **7**	13 **15**
Coast	1158 188	51 858	89 **88**	4 **8**	17 **18**
Mountain	450 96	9 253	87 **87**	1 **9**	27 **15**

Comparing the results of the shown table's one can conclude that:

- the NWC-SAF software package has a similar approach as the KNMI algorithm starting with a defensive T10.8 test, and thus allowing for a high FAR8 value in this specific test.
- the NWC-SAF software package has a better performance in the detection of clear cases resulting in lower FAR0 values.

- The NWC-SAF results obtained by using only the T10.8 and Reflection tests shows the best performance, see table 6.7. The results appear to degrade by the inclusion of the additional tests, see table 6.8. Please note however that the period under study is a summer period. The additional tests may be beneficial in other seasons.
- The KNMI algorithm overestimates the FAR0, which can be attributed for the day time cases to the Reflection tests classifying too many pixels as cloudy.

Comparison of the performance of all the tests for day and night cases the two algorithms results are very similar except for the mountain regions. During day time the NWC-SAF scores are generally better with high POD and low FAR0 and FAR8 values. At night the POD of the NWC-SAF algorithm is better. The FAR0 of the KNMI package is better than the NWC-SAF one. In all cases the FAR8 of the NWC-SAF algorithm is better than the KNMI scores.

Table 6-9 NWC-SAF T10.8 results with the corresponding KNMI results in bold

August 19-31, 3:00 GMT	Na	Nb	POD in percent	FAR8 false cloud detection	FAR0 false clear detection
	Nc	Nd			
Sea	185	50	83	21	10
	21	174	**79**	**23**	**15**
Land	2320	553	86	19	9
	398	3611	**85**	**27**	**6**
Coast	655	138	87	17	9
	95	916	**84**	**22**	**10**
Mountain	130	49	77	27	10
	12	85	**73**	**36**	**10**

Table 6-10 NWC-SAF T10.8 and T3.9 tests results, KNMI results in bold

August 19-31, 3:00 GMT	Na	Nb	POD in percent	FAR8 false cloud detection	FAR0 false clear detection
	Nc	Nd			
Sea	231	10	85	4	28
	54	133	**83**	**16**	**16**
Land	2697	185	88	6	14
	569	3360	**88**	**16**	**8**
Coast	746	53	88	6	15
	149	838	**86**	**13**	**13**
Mountain	153	21	86	12	14
	14	81	**76**	**30**	**10**

In this comparison the latitudinal dependence of the results is not evaluated due to time restrictions. Also the dependence of the KNMI results on twilight conditions is not addressed.
From this reference point the KNMI algorithm has produced very comparable results to the NWC-SAF algorithm. The NWC-SAF algorithm, however, provides next to a cloud mask more added value products, like cloud type, cloud top height.

7 Evaluation Forecaster experience, Metcast experience

For the evaluation of the products the end-users forecasters and researchers were approached. To facilitate evaluation of the products it was required to import the products into the working environment of the forecasters. The research people cloud be directly served with the output of the algorithm.

The products of the cloud mask where distributed to the meteorological work stations and the forecasters were prompted for reactions. The cloud mask itself was not considered to have added value for forecasting purposes by the reacting forecasters.
Their arguments were that the cloudiness of a pixel is in most cases easily recognized by the trained eye of the forecaster and the cloud mask here did not give any contribution to their interpretation.

To understand the arguments of the forecasters one has to realize that at present forecasters are confronted with a large data quantity. This data has to be digested and the forecaster has to choose the relevant information to make the best forecast. The schedule to make the forecast is tight, leaving little space to study and appreciate innovative products. Researchers offer them new products on a regular basis. These new products can not replace old products before the merits of the new product are proven and accepted by the forecasters. So this means that a new product increases the workload of the forecaster, at least for a certain period, until there is consensus over the removal of an old redundant product. Therefore forecasters are a bit reluctant in using new products unless the added value is straightforward and clear.
The acceptance by the forecasters of the water vapour channel of the ‘old’ Meteosat satellite is a striking illustration of the occurring time paths of implementation of new information. It took nearly two decades after the launch of the first operational European satellite in 1977 before the information content of the water vapour channel was eminent. The patterns visible in the imagery were shown to relate to potential vorticity in the upper atmosphere. Once this relation was proven the forecaster environment used the water vapour channel in a more efficient way. The water vapour images enabled an extra check on the performance of the NWP model output (Mansfield, 1997).
In case of the cloud mask, be it from the KNMI algorithm or the NWC-SAF algorithm, the forecasters do not see an added value in comparison to their visual interpretation. Hence the response on the request to evaluate the cloud mask was low.

The forecasters, however, did appreciate the cloud type product of the NWC-SAF software. The forecasters did see added value in this product, especially when animated. At present a number of forecasters use this product to facilitate their interpretation of the cloud types. The cloud type product is highly appreciated by General aviation forecasters, as it gives a clear presentation of the lower clouds and fog. The visual interpretation of low clouds in night time situations is difficult due to the relative low contrast between cloud and clear skies. Here the automated interpretation by the NWC-SAF software classifies low clouds directly and displays them in an easy understandable colour coding. It facilitates the interpretation of the forecasters, and releases them from the task to peer at the screen to identify low clouds with little contrast in comparison to the underlying surface. The added value of the cloud type product compared to the cloud mask is the ability to distinguish various cloud types from each other. This enables a crude 3 Dimensional interpretation by the forecaster enabling him or her to distinguish relevant cloud occurrences from irrelevant ones.
Also the other cloud mask related NWC-SAF products, cloud top temperature and cloud top height were considered to have potential for the forecasters. The evaluation

response on these products was, however, also low as no direct added value was clearly recognizable.

The actual use of a cloud mask is more relevant in an environment of automated processing. In an automated process chain there is no human (eye) quality control of the resulting cloud mask. Therefore a good quality cloud mask is essential for the usefulness of the end products of the automated processing chain.

It was intended within the project to perform tests in the summer of 2005 on such an automated process called MetCast. Unfortunately this appeared to be impossible due to time constraints and other obligations. In October 2005 there was time to do a limited number of experiments described in the next section.

7.1 MetCast experiments

At KNMI an operational cloud forecast model is developed, called MetCast (Veen et al 1996). The model uses NWP (HIRLAM) information for the prognostic variables of wind.

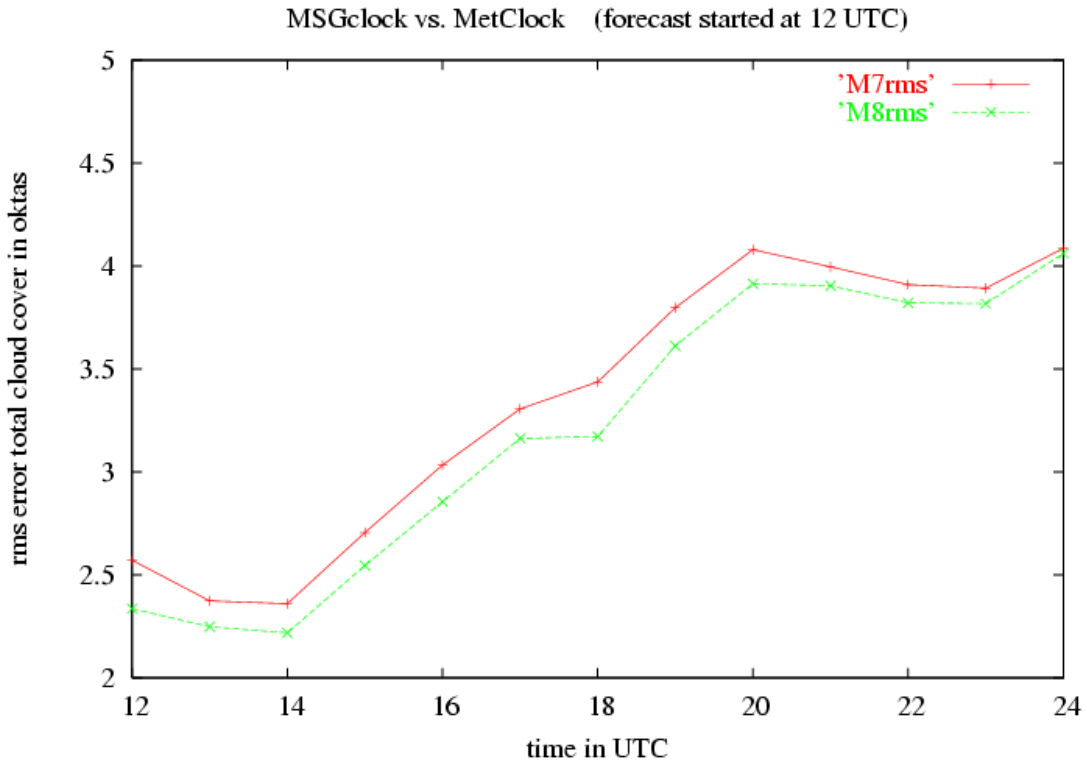

Figure 7-1 The RMS difference of the forecast cloud coverage versus the synops observed coverage as a function of forecast period. The coverage is expressed in okta's. Red indicates the initialization with the cloud mask derived from the METEOSAT 7 observations, green denotes the same but for the METEOSAT 8 observations.

The temperature and humidity fields are initialized based on NWP model information. The initialization of the cloud field of the model is derived from satellite observation and

synops information. The cloud mask provides information about the cloud extension, coverage and cloud top height, the synops provides information about the cloud base height. Together they enable a 3 Dimensional cloud initialization. After initialization a forecast is calculated for a period of 12 hours.

For a comparison done for this study ten forecasts are produced for the period of September 30 till October 9 2005 all starting at 12:00 GMT, for the Western Europe area. For this study the coverage results of the cloud mask based on METEOSAT 7 and METEOSAT 8 are compared during the forecast period.

The Forecast results are compared to synops observations for a number of stations in Western Europe. This enables a determination of the standard deviation (RMS) and the bias as function of the forecast time, shown in Figures 7-1 and 7-2.

The standard deviation of the METEOSAT 8 based cloud mask remains lower than the METEOSAT 7 based cloud mask, throughout the whole forecasting period. This indicates that initialization with the METEOSAT 8 based cloud mask has a positive impact which remains throughout the whole forecasting period.

The bias shown in Figure 7-2 is significantly lower compared to the METEOSAT 7 based cloud initialization. The METEOSAT 8 cloud mask shows better results in comparison to the METEOSAT 7 cloud mask.

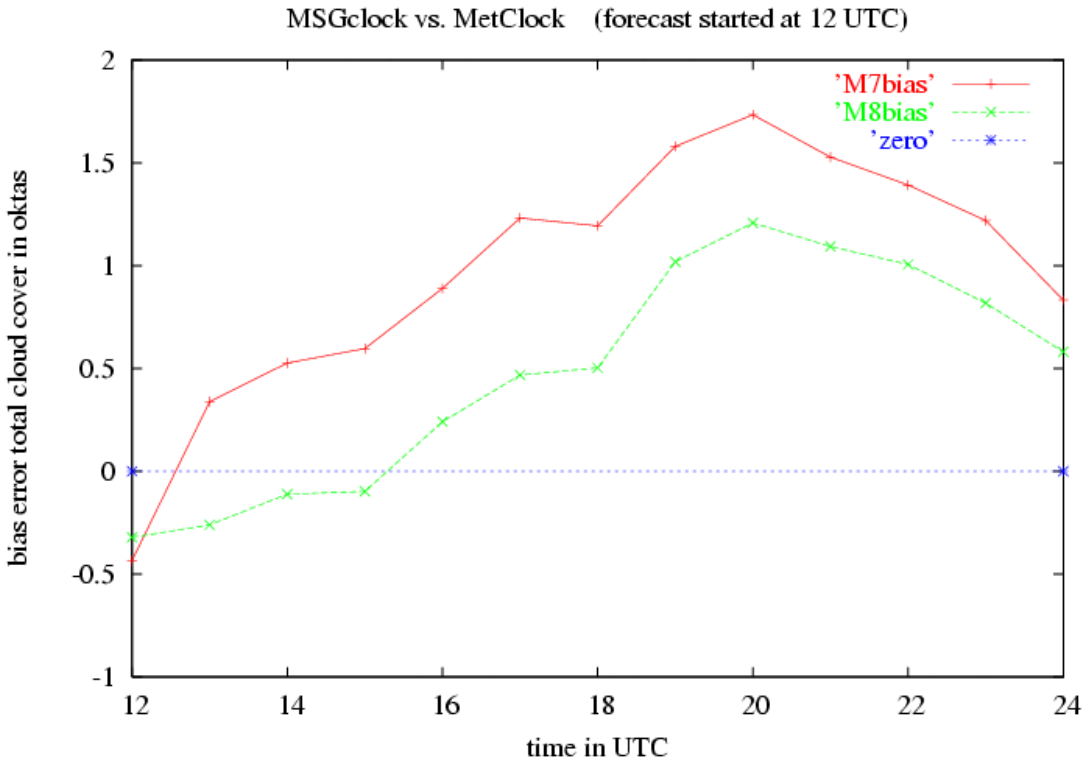

Figure 7-2 The caption is the same as in Figure 7-1 but here the bias difference between synops and forecast is shown as function of the forecasting period. For interpretation purposes the zero axes is added in blue.

A similar comparison using the NWC-SAF cloud mask could have been useful. This comparison, however, was not performed as the MetCast software was not ready to handle the NWC-SAF cloud masks within the project time frame.

8 Conclusions

A cloud detection algorithm is developed at KNMI. Its' purpose is the classification of pixels within a satellite observation as cloud contaminated or clear. This image classification aims to fulfil the needs of a number of applications ranging from forecasting to assimilation of cloud information into numerical weather prediction models.

For classification a thresholding technique is developed. In order to discriminate cloud contaminated pixels from clear pixels information is required about the signal to be expected from a clear pixel. The use of surface temperatures from a NWP model can supply this information for infrared spectral channels. For channels which reflect sun light cloud cleared composites supply the required cloud free information. This technique was already successfully applied in the MetClock algorithm. The MetClock algorithm was developed in a project of the BCRS USP 2 program. It was implemented at KNMI to classify the METEOSAT 7 imagery.

The launch of METEOSAT 8 which became operational in 2004 required a significant change in the MetClock algorithm. METEOSAT 8 offers an improved spectral resolution observing in eleven spectral channels, versus three channels on METEOSAT 7, a higher temporal resolution, 15 minutes versus the former 30 minutes update cycle, and a higher spatial resolution, 3 x 3 km² versus the former 5 x 5 km² resolution for the infrared channels.

A number of steps were required to update the MetClock algorithm to an algorithm which can perform adequately on the METEOSAT 8 imagery. The steps involved a collocation with other sources of information used in the algorithm, e.g. NWP fields, a choice of aerial coverage, redesigning the software, adding additional threshold tests, tuning of the algorithm.
The tuning of the KNMI algorithm by changing the thresholds to obtain an optimum result for both detection of clouds and clear skies required most of the time.
Furthermore the algorithm was implemented and tested in two different system environments, Unix and Linux. The benefit of the multi-platform approach was a minimization of program errors and software portability.

Simultaneous with the development of the new algorithm the NWC-SAF algorithm became available. This NWC-SAF algorithm also provides a cloud detection method.
It was foreseen that this software would become available well after the end of this demonstration project. Due to uncontrollable circumstances the start of this project was significantly delayed.
Since the NWC-SAF products were available they were used as a reference of the results obtained with the KNMI algorithm.

The performances of the two algorithms are comparable. The probabilities of detection (POD) scores are generally over 80 percent. The False alarm rates for cloudy (FAR8) and clear skies (FAR0) are acceptable.
During daytime the thresholding tests for the Reflection channels contributed most to the FAR0. A better tuning of these tests will improve the model performance.
The high POD scores prove that the use of HiRLAM surface temperatures has a significant positive impact on the successful detection of cloud contaminated pixels. It was one of the aims of the project to prove this concept.

The NWC-SAF algorithm has been developed over a period of five years. The KNMI algorithm is developed in less than one year. It is based on the experience of the MetClock algorithm. The KNMI algorithm results presented here show that there are a number of possibilities to improve the results. This also applies of course to the NWC-SAF algorithm.

Next to the KNMI algorithm the NWC-SAF algorithm is more elaborate. It produces not only a cloud mask but also a cloud type and cloud top heights. The latter results are obtained with additional NWP information. These extra products are more advanced in comparison to the KNMI algorithm.

At the end of the project the performance dependence on latitude and twilight conditions became clear. Due to time constraints this dependency could not be fully explored in this project.
The results of the research will be shared within the NWC-SAF group. This will lead to a better understanding of relevant processes and hopefully to better scores.

The end users can be subdivided into two groups. One group are the forecasters having a large interest in higher level products like cloud types and cloud top heights. The second group are the researchers who benefit more from the basic products like a cloud mask.

Both groups are eager to use the products as can be derived from the new generation geostationary satellites. It is planned to serve the users with the NWC-SAF products, based on the following arguments:

- the NWC-SAF algorithm provides higher level products, such as cloud type, cloud top height. These products are not available in the KNMI algorithm.
- The performance of both packages show a comparable quality. This does not provide an argument to choose for one of the packages. So the first argument, given above, prevails.

9 Literature

Desbois, M., Seze, G. and Szejwach, G., 1982, Automatic classification of clouds on METEOSAT imagery: Application to high-level clouds, Journal of Applied Meteorology, Vol. 21, pp. 401-412.

Feijt, A.J., and J.P.J.M.M. de Valk, 1998, Quantifying the difference between NWP surface temperatures and cloud free satellite apparent brightness temperatures, Contr. Atmos Physics, pp 455-460

Feijt, A.J., and J.P.J.M.M. de Valk, 2001, The use of NWP model surface temperatures in cloud detection from satellite, International Journal of Remote Sensing, 22, 2571-2584.

Gtopo30, 1998, Description of world topography data set, Internet URL-address: http://edcwww.cr.usgs.gov/landdaac/gtopo30/README.html.

Gustafsson, N., 1993, HIRLAM 2 Final Report, SMHI Technical Report 9, 126 pp.

Kriebel, K.T., Saunders, R.W., and Gesell, G., 1989, Optical properties of clouds derived from fully cloudy AVHRR Pixels, Beitr. Phys. Atmosph., Vol. 62, no. 3, pp. 165-171.

Le Gléau H. and M. Derrien , 2003, NWC-SAF/MSG SEVIRI cloud products, *Proc. Of the 2003 EUMETSAT Meteorological satellite conference*, Weimar, Germany, 29 sept.-3 oct., 191-198.

Le Gléau H. and M. Derrien, 2000, Prototype Scientific Description for Meteo-France/CMS, Documentation for the Now casting and very short range forecasting SAF, SAF/NWC/MFCMS/MTR/PSD, Issue 1, Rev. 0

Mansfield, D.A., 1997, The ue of potential vorticity and water vapour imagery to validate numerical models, Meteorol. Appl., 4, pp. 305-309.

Minnis, P. and Harrison, E.F., 1984a, Diurnal variability of regional cloud and clear sky radiative parameters derived from GOES data. Part I: Analysis method, Journal of Climate and Applied Meteorology, Vol. 23, pp. 993-1011.

Minnis, P. and Harrison, E.F., 1984b, Diurnal variability of regional cloud and clear sky radiative parameters derived from GOES data. Part II: November 1978 cloud distributions, Journal of Climate and Applied Meteorology, Vol. 23, pp. 1012-1031.

Minnis, P. and Harrison, E.F., 1984c, Diurnal variability of regional cloud and clear sky radiative parameters derived from GOES data. Part III: November 1978 radiative parameters, Journal of Climate and Applied Meteorology, Vol. 23, pp. 1032-1051.

Rossow, W.B., Mosher, F., Kinsella, E., Arking, A., Desbois, M., Harrison, E., Minnis, P., Ruprecht, E., Seze, G., Simmer, C. and Smith, E., 1985, The ISCCP cloud algorithm intercomparison, Journal of Climate Applied Meteorology, Vol. 24, pp. 877-903.

Rossow, W.B., 1989, Measuring cloud properties from space: A review, Journal of Climate, Vol. 2, pp. 201-213.

Rossow, W.B. and Garder, L.C., 1993a, Cloud detection using satellite measurements of infrared and visible radiances for ISCCP, Journal of Climate, Vol. 6, pp. 2341-2369.

Rossow, W.B. and Garder, L.C., 1993b, Validation of ISCCP cloud detection, Journal of Climate, Vol. 6, pp.2370-2393.

Rossow, W.B., Walker, A.W. and Garder, L.C., 1993c, Comparison of ISCCP and other cloud amounts, Journal of Climate, Vol. 6, pp. 2394-2418.

Saunders, R.W. and Kriebel, K.T., 1988, An improved method for detecting clear sky and cloudy radiances from AVHRR data, International Journal of Remote Sensing, Vol. 9, No. 1, pp 123-150.

Seze, G. and Rossow, W.B., 1991a, Time-cumulated visible infrared histograms used as description of surface and cloud variations, International Journal of Remote Sensing, Vol. 12, pp. 877-920.

Seze G. and Rossow, W.B., 1991b, Effects of satellite data resolution on measuring the space/time variations of surfaces and clouds, International Journal of Remote Sensing, Vol. 12, pp. 921-952.

Seze G. and Desbois M., 1987, Cloud cover analysis from satellite imagery using spatial and temporal characteristics of the data, Journal of Climate and Applied Meteorology, Vol. 26, pp. 287-303.

Shin, D., Pollard, J.K., and Muller, J.P., 1996, Cloud detection from thermal infrared images using a segmentation technique, International Journal of Remote Sensing, Vol. 17, pp. 2845-2856.

Simmer, C., Raschke, E. and Ruprecht, E., 1982, A method for determination of cloud properties from two-dimensional histograms, Annual Meteorology, Vol. 18, pp. 130-132.

Valk, J.P.J.M.M. de, Feijt, A.J. and Roozekrans, J.N., 1997, The application of NWP data and METEOSAT imagery in the retrieval of cloud parameters, Proceedings of the 1997 Meteorological Satellite Data Users' Conference, EUMETSAT publication EUM P 21, pp 139-146.

Valk, J. de, A. Feijt, H. Roozekrans, and S. van der Veen., Cloud field characterization using METEOSAT imagery and NWP Model data, Proceedings of the 1999 EUMETSAT Meteorological Satellite Data Users' Conference, Copenhagen, Denmark, September 6-10, 1999a.

Valk, J. de, A. Feijt, H. Roozekrans, R. Roebeling, A. Rosema, 1998, Operationalisation of an algorithm for the detection and characterisation of clouds n METEOSAT imagery. BCRS report USP-2-98-09, ISBN 90-5411-246-8

Veen, S.H. van der and Feijt, A.J., 1996, Liquid Water initialisation in a cloud prediction model using METEOSAT imagery, Proceedings of the 1996 Meteorological Satellite Data Users' Conference, EUMETSAT publication EUM P 19, pp. 257-264.

Van der Veen, S. H.,2002, 'MetClock cloud initialisation in a high resolution limited area model', Proceedings of 'The 2002 Eumetsat Meteorological Satellite Conference', Dublin, Ireland, 2 - 6 September 2002, pp 200-206.

Wielicki, B.A. and Parker, L., 1992, On the determination of cloud cover from satellite sensors: The effect of sensor spatial resolution, Journal of Geophysical Resources, Vol. 97, D12, pp. 12799-12823.

World Climate Program, 1988, International Satellite Cloud Climatology Project (ISCCP): documentation of cloud data, WMO/TD-No. 266, 76 pp.

Zelenka, A., Binder, P. and Schubiger, F., 1997, Total Cloud Cover derived from the METEOSAT VIS channel for monitoring the Swiss NWP Models' performance, Proceedings of the 1997 Meteorological Satellite Data Users' Conference, EUMETSAT publication EUM P 21, pp. 155-161.

10 Acronyms

BCRS	Beleids commissie Remote Sensing
CM	Cloud Mask
FAR0	False alarm rate for clear skies
FAR8	False alarm rate for overcast (cloudy) skies
HDF	Hierarchical Data Format
MF	Meteo France in Lannion
NWC-SAF	SAF on Now casting and short range forecasting
POD	Probability of Detection
Refl0.6	Observed reflection from the 0.6 µm channel
Refl0.8	Observed reflection from the 0.8 µm channel
SAF	Satellite Application Facility
SD	Standard Deviation
T10.8, (T039)	Observed brightness temperatures from the 10.8 (3.9) µm channel.
Tdif	Calculated difference between NWP surface temperature and observed Brightness temperature for clear situations.
USP	User support program