CLOUD HEIGHT DETERMINATION AND COMPARISON WITH OBSERVED RAINFALL BY USING METEOSAT SECOND GENERATION (MSG) IMAGERIES

Peter S. Masika
Kenya Meteorological Department,
P.O Box 30259 00100,
Nairobi, KENYA.
Email: pmasika@meteo.go.ke

ABSTRACT:

To obtain accurate estimates of surface and cloud parameters from satellite data an algorithm has to be developed which identifies cloud-free and cloud-contaminated pixels. Data from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) on board Meteosat Second Generation (MSG) satellites have been available since February 2004. The data is accessible to National Meteorological and Hydrological Services (NMHSs).

This study attempts to utilize available MSG data for developing simple cloud mask and height algorithms and thereafter compare and determine the relationship between cloud height and observed rainfall on a ground station. A multispectral threshold technique has been used: the test sequence depends on solar illumination conditions and geographical location whereas most thresholds used here were empirically determined and applied to each individual pixel to determine whether that pixel is cloud-free or cloud-contaminated. The study starts from the premise of an acceptable trade-off between calculation speed and accuracy in the output data. For this reason, only three infrared channels of MSG satellite were used alongside climatological data provided by National Oceanographic and Atmospheric Administration (NOAA) and also land surface climatological data available from the WorldClim website.

The accurate measurement of spatial and temporal variation of tropical rainfall around the globe remains one of the critical unresolved problems in the field of meteorology. This study attempted to compare computed cloud height and observed rainfall on ground station (CGIS-Butare, Rwanda) and derived cloud height-total rainfall relationship from storms over the same station.

Results from the simple cloud mask algorithm were validated using EUMETSAT cloud mask products for a tropical region (≈ 11°N - 14°S and ≈ 6° - 51°E) over Africa. Overall accuracy of the simple cloud mask developed here was found to be 87% for four scenes which were during day- and night-time as well as twilight time as defined by sun elevation angles. Analysis of recorded rainfall at CGIS and comparison of the same with computed cloud height showed that rainfall mainly occurred when cloud heights were greater than 3000m. Further, deriving a relationship between the observed rainfall and the cloud height was found to follow a Gaussian model in which clouds at approximate heights between 4000m and 5000m produced higher amounts of rainfall. Below and above this height range, rainfall amounts were found to be generally low. The derived cloud height-total rainfall relationship was applied to other storms over this station. Initial results show low correlation between estimated and observed rainfall. More synoptic observations have to be used to evaluate the derived relationship. Next to this a better procedure to differentiate *nimbostratus* and *cumulonimbus* has to be incorporated. Different relations between height and observed rainfall for the two types of clouds may be derived which may improve the overall results.

***Key Words*:** MSG-SEVIRI, cloud mask, cloud height, rainfall comparison.

1.0 Introduction

For more than 40 years, meteorological satellites have been the best way to observe the changing weather on a large scale (EUMETSAT, 2006). Typically, operational meteorology utilizes two types of satellites, namely; polar orbiting and geostationary satellites, to provide the required information. Polar orbiting satellites fly at relatively low altitudes of approximately 800km above the earth surface and can provide information based on a high spatial resolution. Geostationary satellites, on the other hand, are in the equatorial plane and at high altitudes of about 36000km above the earth surface. Their revolution time is the same as that of the earth itself and therefore the satellites are always viewing the same area on the earth. They have low spatial resolution due to their altitudes.

Meteosat Second Generation (MSG) satellites with the first one (MSG-1 now Meteosat-8) was launched on 28th August 2002 and became operational in early 2004. The second of this series (MSG-2 now Meteosat-9) was launched on 20th December 2005. These two satellites are located at 0°N and 0°E.

MSG satellites are spin-stabilized and capable of greatly enhanced earth observations (EUMETSAT, 2006) . The Spinning Enhanced Visible and Infrared Imager (SEVIRI) sensor on board MSG has a high temporal resolution of 15 minutes and spatial resolution of 3 km (sub-satellite) for all channels except 1km for high resolution visible (HRV) channel. The major improvement for this series of satellites is the enhanced spectral resolution of 12 channels. The presence of the 3.9μm channel in the current sensor has allowed analyses of cloud cover especially at night-time.

The primary mission of MSG satellites is the continuous observation of earth's full disk with a multi-spectral imager. The repeat cycle of 15 minutes for full-disk imaging provides multi-spectral observations of rapidly changing phenomena such as deep convection. They also provide better retrieval of wind fields which are obtained from the tracking of clouds, water vapour and ozone features. In this study, main attention is given to cloud properties, such as cloud height, that may be associated with rainfall amounts observed a ground station.

Presence and characteristics of clouds gives information about the state of the atmosphere. For many cloudy situations, the reflected visible radiation and the emitted thermal radiation are not simple to interpret because the cloud is not the only reflecting/radiating source (Dlhopolsky and Feijt, 2001). Of importance is to determine cloud properties by first distinguishing cloud-free pixels from cloud-contaminated pixels. Quantitative data sets obtained from the cloud-contaminated pixels have several potential applications one of which is for water resources and environmental management.

In general, effective integrated water resources management requires timely, accurate and comprehensive meteorological, hydrological and other related information. Use of satellites in observing variables such as rainfall, evaporation and soil moisture has enhanced provision of these data in a timely and effective manner for the water resources management sector. These meteorological variables needs to be monitored effectively and since they are associated with atmospheric moisture hence clouds, there is need to identify the clouds first through masking all cloud-contaminated areas in satellite images.

Firstly, cloud mask is done by using multispectral threshold technique and explore the performance of the algorithm by comparing with the EUMETSAT cloud mask products. Then cloud height is developed for those pixels that are isolated as cloud-contaminated. The cloud height products over Centre for Geographical Information System (CGIS)-Butare, Rwanda station are then compared with the total rainfall observed over the same station in order to determine the relationship.

2.0 Cloud Masking and Cloud Height

Cloud masking allows identifying cloud-free areas where other products such as land or sea surface temperatures may be computed. It also allows identifying cloudy areas where other products (e.g. cloud types and cloud top temperature/height) may be derived. Cloud type on the other hand provides a detailed cloud analysis. It may be used as input

to an objective meso-scale analysis which in turn may be used in a simple nowcasting scheme (Météo-France, 2005b). Cloud type product is essential for generation of cloud top temperature and height products and for identification of precipitating clouds which in turn may be used to estimate rainfall intensity/amount.

Cloud detection from remote sensing data is required for many applications. Some of these are such as determination of cloud cover, identification of cloudy pixels for the retrieval of cloud-related parameters, or exclusion of pixels with even minor cloud contamination if further processing would be affected by the presence of clouds (Schröder *et al.*, 2002). Several methods can be used to perform cloud detection. Some of these methods are such as multispectral thresholding techniques that can be applied to individual pixels (Saunders and Kriebel, 1988) , (Derrien *et al.*, 1993), (Stowe *et al.*, 1999). Dynamic cloud cluster analysis relying on histogram analysis was suggested by (Desbois *et al.*, 1982) whereas (Bankert, 1994) indicated use of artificial neural networks which needs manual training. Another approach was suggested by (Ebert, 1987) which involve pattern recognition techniques based on large scale texture analysis.

Cloud mask and type software modules have been developed by the Centre de Météologie Spatiale of Météo-France and are embedded in the Satellite Application Facility for supporting NoWCasting and very short range forecasting (SAFNWC)/MSG software package that is distributed by EUMETSAT (Derrien and Le Gléau, 2005). These cloud mask and type algorithms uses transfer functions derived from atmospheric models which are not published. Most of National Hydrological Services, especially in Africa, have no access to these transfer functions and even then may not be in a position to derive, on their own, the transfer functions. Besides, due to financial limitations for most of these National Hydrological Services, shareware or freeware (such as Integrated Land and Water Information System (ILWIS)) can be used for masking clouds and determining the cloud type through semi-automated processing.

Thus there is a need for masking out clouds and determining their basic properties in order to improve forecasted rainfall estimates from MSG. It is envisaged that improving rainfall estimation will assist most of National Hydrological Services to provide information on the status of water resources within their area of jurisdiction. It is also expected that it would further improve timely decision making for, areas prone to disasters related to weather such as floods, landslides or areas frequently affected by droughts. This therefore calls for a need to develop simple cloud mask (SCM) and cloud height (SCH) algorithms which may be embedded in readily available shareware such as ILWIS.

3.0 Materials and Methods

3.1 Data

The study attempted to use various data whose source was either straightforward to obtain or needed pre-processing. The data used here was MSG Satellite data and climatological data (Land Surface Temperature (LST) and Sea Surface Temperature (SST)) for the calculation of the cloud height. Field data from the CGIS-Butare, Rwanda was used for the comparison purpose with the calculated cloud height(s).

3.1.1 MSG Satellite Data

The Spinning Enhanced Visible and Infrared Imager (SEVIRI) sensor onboard MSG satellite provides data to EUMETSAT at Darmstadt (Germany) which is processed and then uplinked to HOTBIRD-6 in wavelet compressed format (Gieske *et al.*, 2004). The images are received and archived at the International Institute for Geo-Information Science and Earth Observation (ITC) in compressed form on external drives which are linked to the ITC network and hence accessed through ordinary personal computers. The image geocoding and radiometric calibration coefficients are supplied in so called EPI and PRO files. The data is not atmospherically corrected. Therefore direct ground observation(s) can only be related to the satellite observation(s) (at the required resolution) after atmospheric correction of the images. In this study this step is not necessary since the focus is on clouds which are the main atmospheric parameters aimed at removing from satellite images for direct ground observation(s) relations.

The retrieval of MSG data is straightforward using import utilities developed. External batch files were created using the MSG data Retriever software available at ITC. For more details about the software refer to (Maathuis *et al.*, 2005).

3.1.2 Climatological Data

Climatological data required in this study were minimum, maximum, and mean land surface temperature as well as sea surface temperature. Minimum, maximum, and mean land surface temperature were obtained from 'WorldClim' database available for download from http://www.worldclim.org (Hijmans *et al.*, 2005b). This dataset contains global climate grids with a spatial resolution of a square kilometre and can be used for mapping and spatial modelling in GIS (Hijmans *et al.*, 2005a). Sea surface temperatures (SST) were derived from climatological data using the NOAA National Oceanographic Data Centre (NODC) and the University of Miami Rosenstiel School of Marine and Atmospheric Science (RSMAS) AVHRR Version 5.0 Pathfinder SST dataset available at ftp://data.nodc.noaa.gov/pub/data.nodc/pathfinder/Version5.0_Climatologies/ (NOAA-NODC, 2006) for the period 1985 to 2001. This averaged data was already resolved to 4km and in 5-day, 7-day, 8-day, monthly, seasonal, and annual periods and each period provided daytime-only, night-time-only, and day-night combined.

Here day-night combined monthly mean sea surface temperature HDF file dataset was imported to ERDAS software and then into ILWIS software. The dataset provided needs to be rescaled and transformed to represent SST in degree Kelvin. The scale and offset provided are 0.075 and -3°K respectively, so that the expression for calculation of SST appears as given in equation 3.1.

$$SST_K = (SST_{orig} * 0.075 - 3) + 273.15 \qquad (3.1)$$

where: SST_K and SST_{orig} are corrected SST (in °K) and original SST (in °C), respectively.

At this stage the land surface minimum, maximum, and mean temperatures given in monthly were merged with the generated mean monthly SST. The final images of climatological monthly day-time, night-time, and mean temperature of the entire globe were generated. These were used in computing the dew point temperature as well as performing cloud mask.

3.1.3 Dew Point Temperature

Dew point temperature is an important geophysical parameter that indicates the state of moisture content in the air under given conditions (Hubbard *et al.*, 2003). In their study, Hubbard *et al.*, (2003) presented a temperature-based (daily maximum, minimum, and mean) daily dew point temperature estimation method for historical studies in the Northern Great Plains in USA. They developed four regression-based methods incorporating daily maximum, minimum, and mean temperatures and also daily precipitation for different locations in the plains. After statistical analysis of the results obtained from the four methods, they concluded that the model that performed satisfactorily was in the form:

$$T_d = \alpha(T_m) + \beta(T_n) + \gamma(T_x - T_n) + \lambda \qquad (3.2)$$

where: T_d is the daily dew point temperature in °C
T_m is the daily mean air temperature in °C
T_n is the daily minimum air temperature in °C
T_x is the daily maximum air temperature in °C
α, β, γ, and λ are coefficients of the regression equation.

The method was further supported by the fact that the associated data set required are easily available in most typical meteorological weather stations. The model as Hubbard *et al.*, (2003) pointed out can estimate dew point temperature with sufficient accuracy under varied climatic conditions. Moreover, they also indicated that the climatic conditions observed within the Northern Great Plains are representative of many other regions in the world.

Based on the above statements, dew point temperature was therefore computed by use of the model as given in equation 3.2 and consequently equation 3.3 with all the associated coefficients was adopted. However, here use of the climatological monthly mean, maximum, and minimum temperatures was made instead of the

daily temperatures. Thus monthly climatological dew point temperature was obtained as follows:

$$T_d = -0.0360(T_m) + 0.9679(T_n) + 0.0072(T_x - T_n) + 1.0119 \quad (3.3)$$

where: T_d is the calculated monthly climatological dew point temperature in °C

T_m is the mean monthly temperature in °C

T_n is the minimum monthly temperature in °C

T_x is the maximum monthly temperature in °C

Minimum, maximum and mean monthly temperatures used here were those obtained from the centres mentioned in section 3.1.2 above. However, it is expected that some slight differences may occur in the final dew point temperature values obtained since there was no recalibration of the model with local (African region) data which would otherwise provide more suitable coefficients and subsequently more accurate estimation of dew point temperature. In addition to this, differences due to use of monthly instead of daily temperatures are expected since the regression is based on daily temperatures.

From dew point concept and with the earth's surface, the cloud height can be extracted. Dry adiabatic lapse rate of 1°C per 100m and saturated/moist adiabatic lapse rate of 0.6°C per 100m were used here as suggested by Strahler, (1965) and widely accepted in many studies. Cloud height would be calculated by use of the following expression:

$$H = ((T_x - T_d) * 1 * 100) + ((T_d - T_b) * 0.6 * 100) \quad (3.4)$$

where: H is the cloud height in meters

T_x is the maximum monthly climatological temperature in °K

T_d is the dew point temperature (in °K) as calculated from equation 3.3

T_b is the brightness temperature (in °K) at the top of the cloud

3.2 Cloud Masking Method

Cloud mask method chosen in this study was based on multispectral thresholding technique. This included creating ILWIS scripts in order to generate the necessary images required for processing the cloud mask image.

In this method, a number of tests that allowed identification of pixels contaminated by clouds were performed. The main characteristic of these tests, applied to sea or land pixels, depended on the solar illumination conditions and on the satellite viewing angle.

Here use of non-linear algorithm, as developed by (Météo-France, 2005a), was made in order to compute sea surface temperature using climatological SST. Split window approach was used with IR10.8 and IR12.0 (brightness temperature of bands 9 and 10) averaged and applied in the algorithm which is in the form given below.

$$T_s = 0.98826 T_{10.8} + (1.18116(\sec\theta - 1) + 0.07293 SST) * (T_{10.8} - T_{12.0}) + 1.10718 + 0.2 \quad (3.5)$$

where: T_s is the calculated sea surface temperature (SST) in °C

$T_{10.8}$ and $T_{12.0}$ are brightness temperature (in °C) of bands 9 and 10 respectively

SST is the climatological sea surface temperature (°C)

Sec θ is the inverse of cosine of the satellite zenith angle.

Here climatological night-time temperature was used as SST. A number of tests were performed in an attempt to extract cloud-contaminated pixels. The final cloud mask obtained was then ready for further cloud classification.

Cloud mask results were also validated using available cloud mask products from EUMETSAT, thanks to the ITC Geodata software development section that developed the GRIB2 decoder. An example of cloud mask from EUMETSAT is as given in figure 3-1 showing cloud as white, clear land as green, and clear sea as blue.

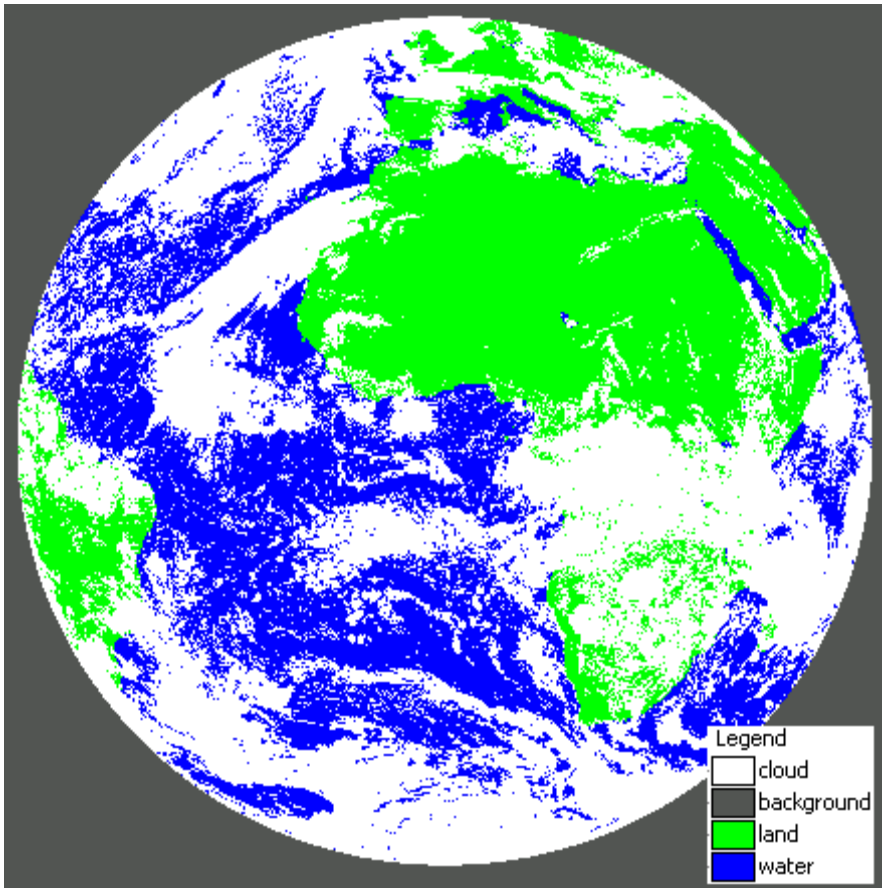

Figure 3-1: MSG cloud mask for 25th December 2006 at 12:00 UTC *(EUMETSAT, 2006)*

3.3 Rainfall Comparison Method

Rainfall is one of the most difficult atmospheric parameter to measure due to its variability in space and time. One-dimensional cloud model-based technique was used in this study to find the basic relationship between cloud height and observed rainfall at a ground station. This was based on relating cloud top height to rainfall amount. Regression-based model as given in equation (3.3) was used to obtain cloud height images. Cloud top height was processed from the cloud top temperature as recorded by MSG satellite. Here brightness temperatures for bands 9 and 10 were averaged.

In this study, rainfall estimation was based on comparison between point observation and satellite estimation using the one-dimensional cloud-model based technique as mentioned above. In this case therefore, and as Maathuis *et al.*, (2006) pointed out; there is a need to incorporate an averaging procedure in order to account for the collocation problems such as spatial and timing offsets. Here spatial average was carried over 5x5 pixels and temporal average was carried for an hour (four MSG images in an hour). Retrieved temperatures of infrared bands 4, 9, and 10 were averaged and used in the simple cloud mask algorithm developed.

The method used in this study follows the idea that clouds produce different amount of rainfall or have different rainfall intensity at different stages of development. This is best represented by (Hong *et al.*, 2004b) in their study on cloud patch-based rainfall estimation using satellite image classification approach. It should be noted here that due to different climatological regimes, empirical relations (either temperature versus rainfall intensity or cloud height versus rainfall intensity) derived may vary significantly. (Adler and Mack, 1984) studied the impact of the regime-to-regime on various empirical rain

estimation schemes based on satellite-observed cloud height or cloud temperature information in which curves representing coastal and inland regimes were strikingly different. They pointed out that these differences had obvious implications for the application of an empirical satellite rain estimation derived in one location and applied in other climatological regimes even with a simple local adjustment. Varying synoptic situations may also cause these types of differences.

Rainfall data from CGIS station in Rwanda was investigated to identify storms which produced rainfall over a given period (recorded after every 30 minutes). The observations were recorded at e.g. 1000hrs, 1030hrs, 1100hrs, 1130hrs, etc. The data was in local time and was converted to Universal Time Convention (UTC). For the case of Rwanda, 2 hours are subtracted from local time to change to UTC.

Twelve storms of different days and time from CGIS station were used to develop a regression function between the height and observed rainfall. The function was consequently used to estimate rainfall amount from other storms over the same station to validate the performance of the method developed and thresholds selected during the simple cloud mask algorithm development.

4.0 Results and Discussions

Generating MSG satellite and sun angles was done by creating a batch file which could be adapted for any date and time in case new angles were required. This particular applet, which can be executed into an active directory, works in a java environment which must be installed in the system. Generated angles were imported into ILWIS for further processing. Figure 4-1 shows the flow chart for generating the satellite and solar angles. The processing was done for mainly MSG field of view covering Africa (≈ 39˚N - 38˚S and ≈ 34˚W - 53˚E).

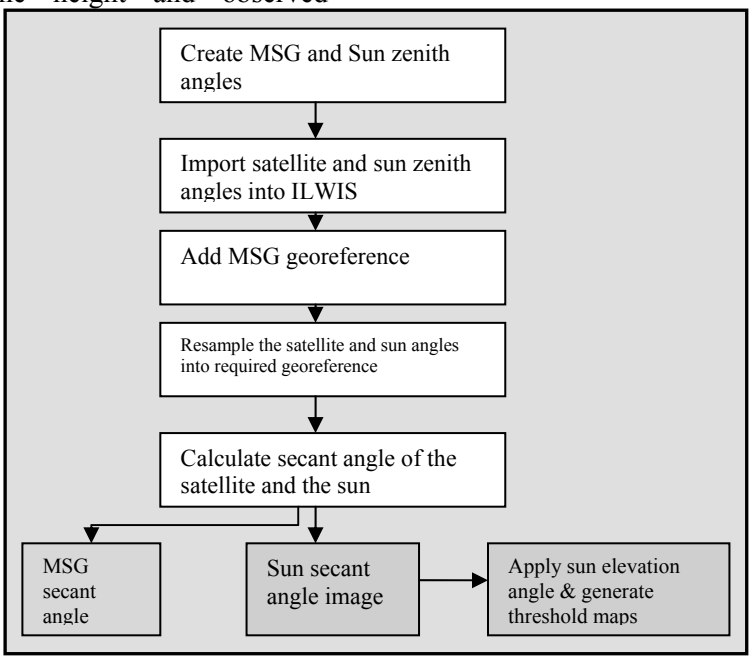

Figure 4-1: Flow chart for generating MSG satellite and Sun angles

After calculating sun elevation angle, solar illumination conditions were generated. This was based on various sun positions in which the condition is day-time when the sun elevation angle is greater than 10° and night-time when the sun elevation angle is less than -3°. The condition is twilight, that is, either before night-time or before day-time when the sun elevation angle is between -3° and 10°. An example of such an

image of 7th March 2006 at 15:30 UTC is provided in figure 4-2.

As earlier pointed out, the algorithm is based on a multispectral threshold technique applied to each pixel of the image. A number of tests for each solar illumination condition in which an example of cloud mask is given in figure 4-3 were applied. The tests applied in the algorithm attempts to address both the land and sea surfaces based on their characteristics.

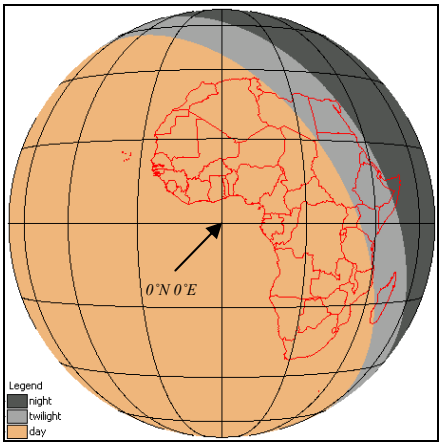

Figure 4-2: Solar illumination conditions on 26th December 2006 at 15:00 UTC

4.1 Day-time Cloud Mask

Day-time cloud mask was obtained by following a number of test steps. Over the land surface, during the day cloud contaminated pixels were identified by using standard deviation from the climatological surface temperatures (mean, minimum, and maximum) which should be greater than 1K. Minimum surface temperature was taken as the monthly climatological night-time temperature as processed from the 'WorldClim' database whereas maximum surface temperature was taken as day-time temperature and the mean surface temperature was average of the day-time and night-time surface temperatures.

To remove false cloud assignment to pixels over desert areas, brightness temperature of band 9 (IR10.8) less than 293.15K was applied otherwise the pixels with a higher temperature were considered cloud free. Further, all pixels already defined as cloudy were subjected to tests in order to avoid cool areas or higher elevated areas. These involved using monthly climatological temperature standard deviation (amplitude). Cloudy pixels with brightness temperature (IR10.8) less than maximum (T_{max}) day-time monthly climatological temperature less half the monthly standard deviation are assigned cloudy else not cloudy. This does not affect the ocean areas since the climatological standard deviation is very small. This test allows us to reduce misclassifications to the minimum except in high elevated areas and desert areas.

Over the sea surface, cloud-contaminated pixels were identified by using standard deviation from the climatological surface temperatures (mean, minimum, and maximum) which should be less than 1K. Further, small difference of -1K (and above) between the local sea surface temperature (as calculated using equation 3.5), here referred to as SST_{cal}, and minimum monthly climatological sea surface temperature (here referred to T_{smin}) was also used to mask cloudy pixels over the sea surfaces.

As earlier explained, Météo-France have developed cloud mask in which this study adopted some of the basic ideas to develop some of the thresholds used. Estimating SST by using IR10.8 and IR12.0 brightness temperature together with minimum monthly climatological SST was used by Météo-France. Here the same two bands are used, as top of atmosphere (in °K) together with minimum monthly SST (here taken as the night time temperature). Météo-France took a small difference of 4K between estimated SST (by using IR10.8 and IR12.0 brightness temperatures) and the monthly climatological minimum SST. Monthly climatological minimum SSTs are derived from a global Pathfinder night-time bulk SST climatology. The bulk night-time SST, as (Derrien and Le Gléau, 2005) pointed out, does not account for the thermal heating at midday observed in infrared satellite

measurements. In this study this difference was set at -1K over the sea surface as stated above.

Brightness temperature of band 9 (IR10.8) was applied by Météo-France as well as by (Kidder *et al.*, 2005) in which the idea was to estimate the temperature that would be observed if there was no water vapour in the atmosphere. Météo-France computed threshold from surface temperatures forecast by NWP model. In this study threshold of 293.15K was set as the maximum temperature for any pixel to be flagged cloud contaminated. Météo-France again used IR10.8 and IR12.0 difference to detect thin cirrus clouds and cloud edges characterized by higher IR10.8-IR12.0 values than cloud-free surfaces. Here use of IR10.8 less than the maximum climatological surface temperature (with half amplitude of the climatological monthly minimum, maximum, and mean temperatures) was to extract thin cirrus clouds as well as to avoid confusion of moist, warm, cloud free areas with clouds. With these few tests day-time cloud mask was obtained of which an example is as given in figure 4-3 (a).

Notable features of this cloud mask are such as sharp boundary between the land and the sea that appears along some coastal areas, especially in this particular case to the North West of the continent. This depicts cloudy conditions over the ocean and non-cloudy conditions over the land, which may not be always the case. The sharp boundary is due to the land-sea temperature effects and increases as we move from equatorial regions to higher latitudes where temperatures are generally low over the sea such as the case in the north-western part of the continent (over the Atlantic ocean). This is more pronounced especially when desert (usually with high temperatures) areas lie next to water body.

Cloud mask image shows presence of clouds over the northern part of Africa whereas from the false colour composite of the visible and near infrared bands does not show the same situation. Over central Africa and Atlantic Ocean (the specific region of interest in this study) most of the cloudy pixels (as can be seen from the false colour composite image) have been masked out.

Also as can be seen from the false colour composite image in figure 4-4 (b), there appears no thick clouds in the northern part of the continent. However, the algorithm has classified the region to be under low level clouds which are semi-transparent in the visible and near infrared bands.

4.2 Night-time Cloud Mask

During night-time, the standard deviation of the monthly climatological temperature was set at a minimum of 1K and the mean brightness temperature (T_{mean}) of IR10.8 and IR12.0 was taken less than 283.15K over the land surface for any pixel to be flagged as cloudy. The difference of the monthly minimum climatological temperature and the brightness temperature of IR10.8 is used is set to be greater than 9K for any pixel to be assigned cloudy. This ensured avoiding cooler areas at night which would otherwise be assigned cloud contaminated. Use of mean brightness temperature for IR10.8 and IR12.0 followed Météo-France developed cloud mask idea in which the difference between the two is used to detect thin cirrus clouds and cloud edges characterised by higher difference (IR10.8-IR12.0) values than cloud-free surfaces. However, in this study the mean of the two was expected to simply avoid the confusion of very moist, warm, cloud free areas with clouds.

Over the sea, the standard deviation of the monthly climatological temperature is less than 1K. The difference between the local calculated sea surface temperatures (SST_{cal}) and the monthly mean climatological temperature was taken to be greater than 10K. Low clouds over the sea were screened by use of IR03.9 to scale down aggregated temperatures of IR10.8 and IR12.0 (i.e. IR10.8*IR12.0). The difference between their mean temperatures and the scaled temperature is set at a minimum threshold value of 2K for the cloudy pixels. This test is based on the fact that the water cloud emissivity is lower at IR03.9 than in IR10.8 or IR12.0. The test allows detecting low clouds at night time. The approach is the same as that of Météo-France using the

difference between IR03.9 and IR10.8. An example of night-time cloud mask is given in figure 4-3 (c).

4.3 Twilight Cloud Mask

At twilight time the difference between climatological minimum temperature and the brightness temperature of band 9 (IR10.8) was set at a threshold of 9K such that any pixel with greater difference than this value and with mean brightness temperature (IR10.8 and IR12.0) less than 283.15K were cloud contaminated. This ensured screening cloudy pixels over the land surfaces where also standard deviation of the monthly climatological temperatures was set at a minimum of 1K.

Over the sea, the difference of mean monthly climatological SST and the calculated SST was taken to be greater than 5K for the cloudy pixels. Here Météo-France used IR10.8 and IR12.0 brightness temperatures to estimate SST by using a nonlinear split window algorithm. A pixel is flagged cloud contaminated if its estimated SST value is lower than a minimum monthly climatological SST value by 4K. However, Météo-France does not apply this test where climatological SST is lower than 270.15K. In this study low clouds were extracted by use of IR3.90 to scale down brightness temperature of bands 9 and 10 (IR10.8 and IR12.0, respectively). Here maximum threshold value of 2K as the difference between the scaled temperature and the mean brightness temperature of IR10.8 and IR 12.0 was used. Threshold for the difference between estimated SST and the climatological SST from Météo-France gives a threshold of 4K which is comparable to the set value in this study.

Météo-France uses IR03.9 and IR10.8 difference to extract low clouds for both day-time and twilight time basing the fact that solar reflection at IR03.9 (approximated by the IR03.9-IR10.8 brightness temperature difference) may be rather high for clouds (especially low clouds), which is not the case for cloud free areas . An example of twilight cloud mask from this study is as given in figure 4-3 (b).

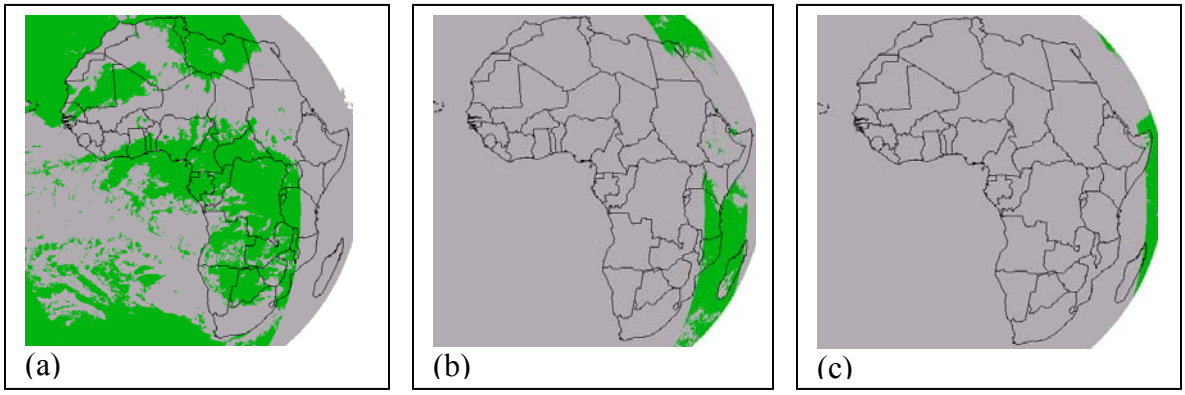

Figure 4-3: Cloud masks for (a) day-time, (b) twilight time, and (c) night-time; for MSG-1 image of 7th March 2006 at 15:30 UTC.

Cloudy pixels are represented as green whereas grey represents non-cloudy pixels. Merging the three images resulted in final cloud mask as given in figure 4-4 (a). RGB Colour composite image of the same day and time is as in figure 4-4 (b).

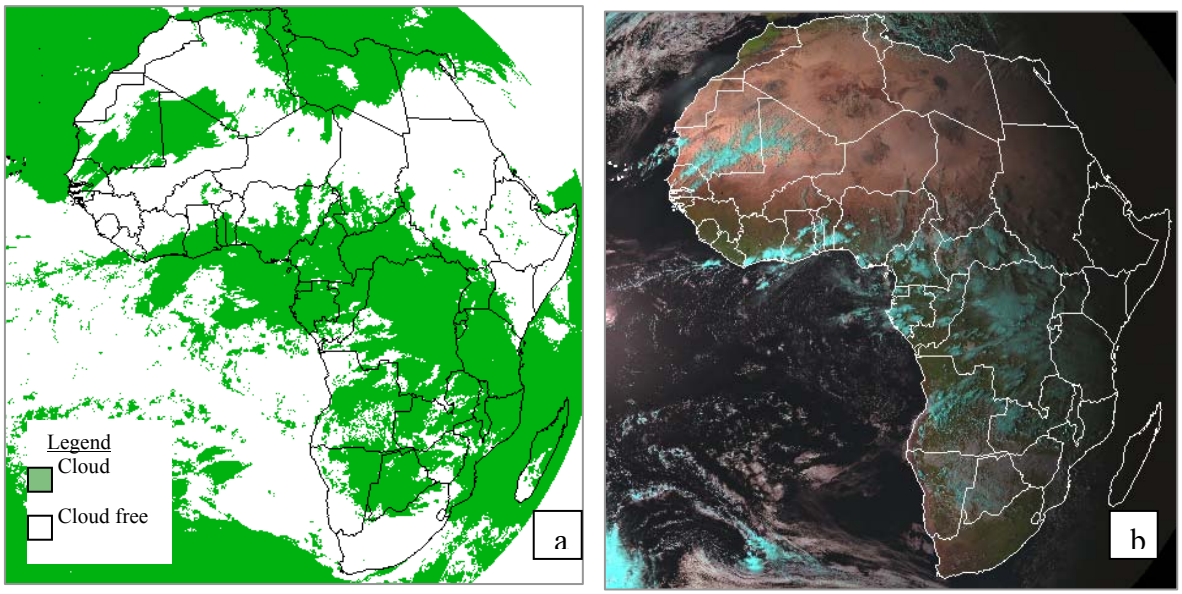

Figure 4-4: Cloud mask (a) and RGB colour composite (b) for MSG image of 07/03/2006 at 15:30 UTC.

Sharp boundary between the land and the sea can be seen to appear along some coastal areas, especially in this particular case to the North West. This depicts cloudy conditions over the ocean and non-cloudy conditions over the land, which may not be always the case. The sharp boundary is due to the land-sea temperature effects. This is more pronounced especially over desert (usually with high temperatures) areas next to water body. Cloud mask image shows presence of clouds over the northern part of Africa whereas from a visual check using the RGB colour composite of the visible and near infrared bands does not show the same scenario. Over central Africa and Atlantic Ocean most of the cloudy pixels (as can be seen from the false colour composite image) have been masked out.

The next step was to process heights for the extracted clouds based on the formula for estimating dew point temperature as given in equation 3.4 in which an example is given in figure 4-5.

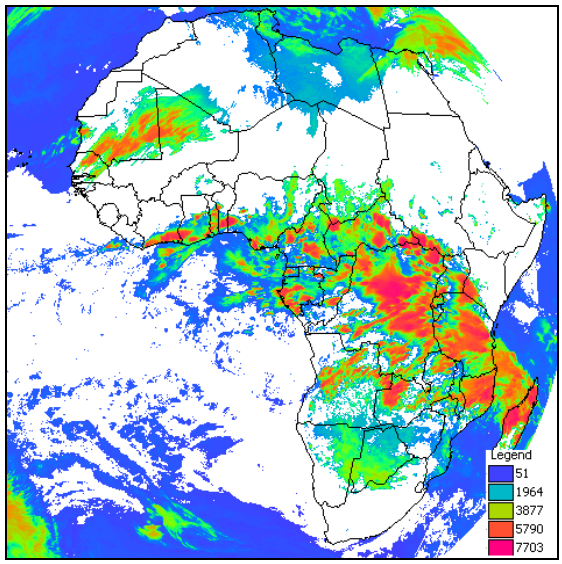

Figure 4-5: Cloud height (in Meters) image (MSG image of 07/03/2006 at 15:30 UTC)

In order to check whether all cloudy pixels have been extracted properly for the area of study, segmentation of the classified cloud image was done and overlaid on to RGB colour composite image of visible and near-infrared bands of the same day and time. Figure 4-6 shows an example of a small window (eastern part of Africa) of such an overlay for MSG satellite image of 23rd November 2005 at 13:30 UTC. This area was under twilight condition on 7th March 2006 at 15:30 UTC and thus such an overlay is not provided here. Segmentation was performed on the classified cloud height image of 23rd November 2005 at 13:30 UTC and the segments overlaid on the RGB colour composite image of the same date and time.

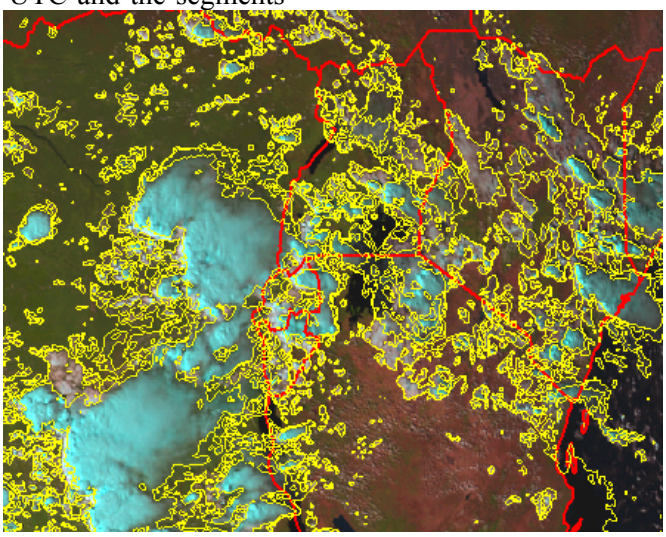

Figure 4-6: Segments (yellow lines) of cloud mask of 23/11/2005 at 13:30 UTC overlaid on RGB colour composite (NIR016, VIS008, and VIS006)

Clouds appear as cyan in colour in the RGB colour composite image. As can be observed visually from figure 4-6, most of the cloudy pixels have been identified. This is more visible over areas where deep cyan colour (mostly deep convective clouds) appears. Some semi-transparent clouds have not been masked out. However, this is not of serious concern in this current study since most of these semi-transparent clouds do not contribute to precipitation, and if they do, very little rainfall is expected from them.

The clouds were selected and a small window (≈ 11°N - 14°S and ≈ 6° - 51°E), covering a part of the tropics over the African continent considered in this study in developing the SCM, was extracted. To evaluate the accuracy of the SCM confusion (contingency) matrix was created. This method compares all pixels, within a selected window, to find out whether the pixels are assigned as cloudy or non-cloudy in both masks.

Here the two cloud masks images were crossed to built a contingency table that indicates the number of pixels in each category. This will show the ability of the SCM to detect cloudy and non-cloudy events based on EUMETSAT CLM. In order to get better results of accuracy of the SCM, there is a need to use a number of images. Since the SCM algorithm was developed based on different solar illumination conditions, it was appropriate to choose MSG images based on these three conditions. Here four days images were used for validation and their contingency matrices were created.

The first MSG image was taken for the 25th December 2006 at 12:00 UTC whose SCM accuracy is found to be 89.1%. The second image was that of the 26th December 15:00 UTC whose SCM accuracy is found to be 88.9% while the third image was that of 4th January 2007 at 22:00 UTC whose SCM accuracy is 88.0% and the fourth image was that of 10th January 2007 at 17:00 UTC of which SCM accuracy is found to be 83.3%.

4.4 Direct comparison of Cloud Height and Total Rainfall

Diurnal trend of observed total rainfall and processed cloud heights, over the weather station, were investigated. Two days were selected for this purpose and figure 4-7 shows how cloud height and total rainfall varied during the selected days. MSG images of 30 minutes interval were processed to obtain cloud heights whereas total rainfall observed at the station at the same time of MSG image acquisition was used.

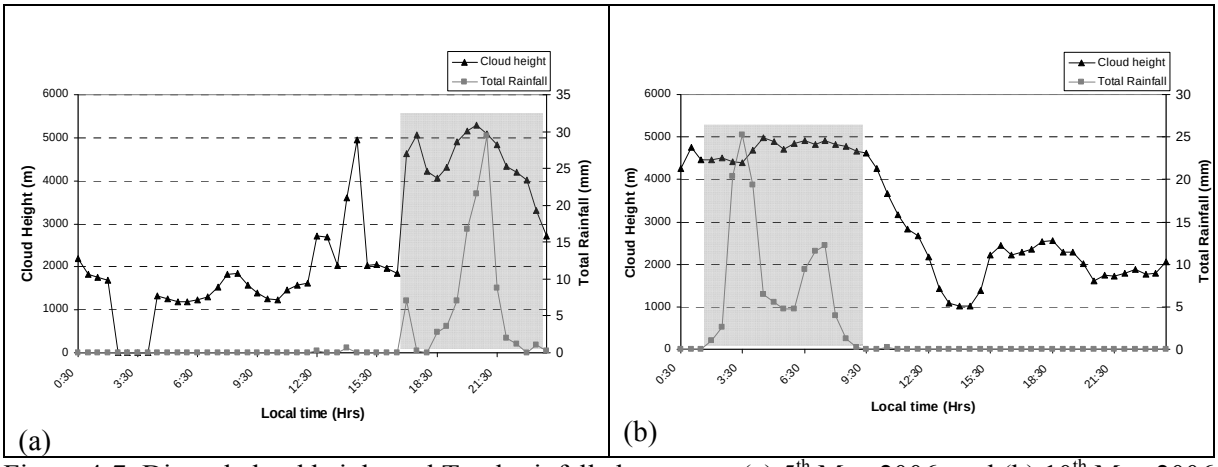

Figure 4-7: Diurnal cloud height and Total rainfall changes on (a) 5th May 2006, and (b) 10th May 2006

It can be observed that rainfall was recorded at the station when the processed cloud height was at high levels (above 3000m). It is clear then that high clouds over this station are the main rain producing rainfall clouds. This indicated that there is a relationship between cloud height and total rainfall produced by the cloud at certain height. The general idea followed in this comparison is that the more the cloud is sustained at a certain height while producing rainfall, the more the rainfall is observed at a ground station.

In order to check whether the results of comparing diurnal change of total rainfall and storm height were mere coincidence, rainfall data from and independent station were investigated. Data from Ministry of Water and Irrigation, Naivasha were used. Two days, one with long rainfall records and the other without rainfall were chosen. Here the day with rainfall records was 1st March 2006 and the day without rainfall was 28th October 2006. Simple cloud height (SCH) algorithm was applied to compute the cloud heights for the two days. Results to this are presented graphically in figure 4-8.

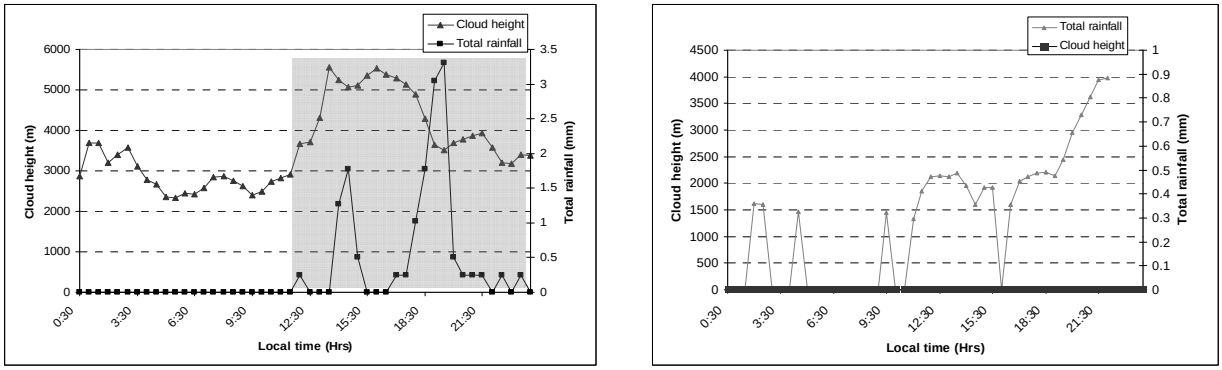

Figure 4-8: Diurnal height and Total rainfall changes on 1st March 2006 (left) and 28th October 2006 (right) over Naivasha station

As can be observed, rainfall occurred generally when cloud height was slightly higher than 3000m especially in the afternoon (the shaded part in the left graph). During this time the likely clouds over the station are convective type of clouds which predominantly occur in the afternoon over this region. Early in the morning, no rainfall was recorded even though the cloud height was slightly more than 3000m. These are likely to be cirrus clouds which mainly occur after dissipation of convective clouds. Thus the convective clouds that produced rainfall in the afternoon and in the night must have dissipated and cirrus clouds appeared in early morning.

The right graph of 28th October 2006 shows that cloud heights above 3000m occurred in the night and there was no rainfall recorded on this day. This may be attributed to the fact that the rainfall was measured at a point and that it may have rained away from the rain gauge.

The situation over CGIS station is slightly different as can be observed that on 5th May 2006 rainfall occurred in the afternoon whereas on 10th May 2006 it occurred in the morning. This means that convective clouds (the likely clouds producing this rainfall) over this region may be sustained at various times during the 24 hours period. Nevertheless, the results of Naivasha station and those of CGIS station are similar in that rainfall is observed when cloud height was more than 3000m.

Determination of the model fit showed Gaussian fit as the best for CGIS by using the 12 storms that appeared over the station selected for this analysis. The above storms were used to determine a regression function between the two variables. The best fit obtained was again a Gaussian model (y=a*exp ((-(x-b)^2)/(2*c^2)); where: a = 60.6, b = 4405.3, and c = 583.0 with correlation coefficient of 0.96 and standard error of 6.56mm. This is presented graphically in figure 4-18. The model agrees with the fact that very low clouds (e.g. stratocumulus, cumulus, and stratus) and very high clouds (e.g. cirrus, cirrocumulus, and cirrostratus) produces very low rainfall.

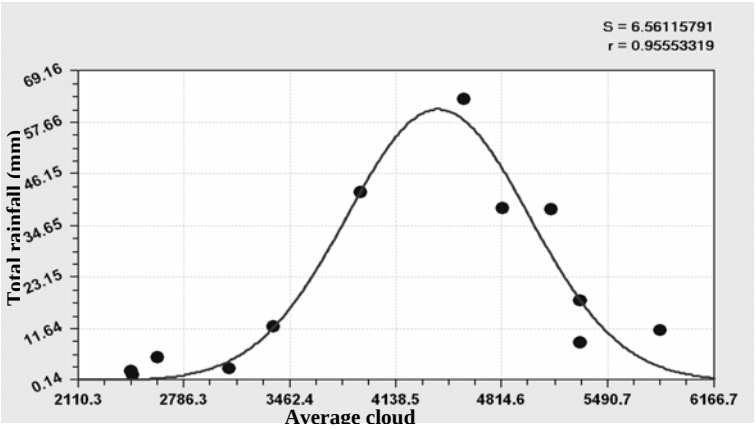

Figure **Error! No text of specified style in document.**-3: Gaussian model fit, X= Average storm height (m), Y= Total rainfall (mm)

Using the function obtained here to estimate total rainfall expected from some storms that appear over the CGIS, Butare Rwanda showed very low correlation with what was observed at the station.

5.0 Conclusions

The main objective is to develop simple cloud mask and height algorithms that can be used for further studies. Availability of satellite images based on thermal infrared bands is essential and is the first focal point in this study. In addition to this is the importance of geospatial data on meteorological parameters that are associated with formation of clouds. Various sources were explored in order to obtain the long term climatological meteorological data, specifically temperature (minimum, maximum, and mean).

Firstly, climatological data from the identified sources were used to process input data for the simple cloud mask algorithm. Secondly, ground rainfall data was used for comparison with the processed cloud heights on various satellite images. During simple cloud mask algorithm development, various thresholds (multi-spectral threshold technique) were explored in order to optimize on extracting all clouds present on any particular day and time.

Based on the developed simple cloud mask algorithm, it was found that setting thresholds for screening all cloudy pixels in satellite images is the most difficult part in threshold techniques. The main problem is that the thresholds are functions of many variables such as; surface type (land, ocean, ice), surface conditions (vegetation, soil moisture), recent weather (which changes surface temperature and reflectance significantly), atmospheric conditions (temperature inversions, haze, foggy), season, time of day and even satellite-earth-sun geometry (hence bidirectional reflectance and sun glint) (Kidder and Haar, 1995).

An automatic simple cloud mask algorithm has been presented ready for use in other applications among them those interested in identification of cloudy pixels for the retrieval of cloud-related parameters (e.g. cloud heights) especially those for clouds which contribute to rainfall (e.g. *cumulonimbus* and *nimbostratus*). The simplicity in the algorithm and significant accuracies based on EUMETSAT data thirsty cloud mask algorithm, and the possibility of automation into shareware or freeware such as ILWIS, may greatly improve cloud detection for specifically weather forecasting in most of the African National Hydrometeorological Services (NHMS).

Thus it was envisioned at the developmental phase that this algorithm would be simple and physically sound and that the MSG satellite imageries and the necessary processing tools (software e.g. ILWIS) would be available in these NHMSs in Africa.

Further to the above, the simple cloud height (SCH) algorithm was developed and consequently comparison with the observed rain gauge data was done. It was then found out that the dew point temperature concept can be used to estimate cloud height which can thus be used to infer rainfall observed on the earth surface. Despite empirical formulation in obtaining geospatial dew point temperature and replication from a different region (USA Northern Great Plains), high correlations when comparing rain gauge observations and processed cloud heights have been obtained.

It can be concluded that there is always need for spatial and temporal averaging of satellite data in order to get better results while comparing point observations on the earth surface.

Nevertheless, further research is needed in order to improve on thresholds tests based on different cloud microphysical processes on formation of cloud particles as well as on variables such as surface type, surface conditions, recent and prevailing weather conditions, and atmospheric conditions. Recalibration or deriving a relationship between dew point temperature and readily available meteorological data e.g. minimum, maximum, and mean temperatures, as suggested by Hubbard *et al.*, (2003) for any region under consideration is recommended for better results.

The SCM and SCH algorithms seem to work well, but they will benefit a lot from a more thorough validation method. Synoptic (Meteorological) data could be used for this purpose.

References

Adler, R.F. and Mack, R.A., 1984. Thunderstorm Cloud height - Rainfall Rate Relations for use with Satellite Rainfall Estimation Techniques. *Journal of Climate and Applied Meteorology*, 23: 280-296.

Bankert, R.L., 1994. Cloud Classification of Avhrr Imagery in Maritime Regions Using a Probabilistic Neural-Network. *Journal of Applied Meteorology*, 33(8): 909-918.

Derrien, M., Farki, B., Harang, L., Legleau, H., Noyalet, A., Pochic, D. and Sairouni, A., 1993. Automatic Cloud Detection Applied to Noaa-11/Avhrr Imagery. *Remote Sensing of Environment*, 46(3): 246-267.

Derrien, M. and Le Gléau, H., 2005. MSG/SEVIRI cloud mask and type from SAFNWC. *International Journal of Remote Sensing*, 26(21): 4707-4732.

Desbois, M., Seze, G. and Szejwach, G., 1982. Automatic Classification of Clouds on Meteosat Imagery: Application to High-level Clouds. *Journal of Applied Meteorology*, 21: 401 - 412.

Dlhopolsky, R. and Feijt, A.J., 2001. Cloud products retrieval for MSG: Final Report. BCRS Report 2000 : USP-2 Report 2000. Beleidscommissie Remote Sensing (BCRS), Delft, 88 pp.90-5411-337-5

Ebert, E., 1987. A Pattern Recognition Technique for Distinguishing Surface and Cloud Types in the Polar Regions. *Journal of Climate and Applied Meteorology*, 26: 1412 - 1427.

EUMETSAT,2006.MSG CHANNELS Interpretation Guide: Weather, Surface conditions and Atmospheric constituents, http://oiswww.eumetsat.org, January 2007

Gieske, A.S.M., Maathuis, B.H.P., Hendrikse, J.H.M., Retsios, V., Leuwen, B.,

Romaguera, M., Sobrino, J.A., Timmermans, W.J. and Su, Z., 2004. Processing of MSG-1 Seviri data in the Thermal Infrared- Algorithm Development with the use of the SPARC2004 data set, *ESA Proceedings WPP-250: SPARC Final workshop, 4-5 July, 2005*. ESA, 2005, Enschede, the Netherlands, pp. 8.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. and Jarvis, A., 2005a. Very High Resolution Interpolated Climate Surfaces for Global Land Areas. *International Journal of Climatology*, 25: 1965-1978.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. and Jarvis, A., 2005b.WORLDCLIM, http://www.worldclim.org. , June 2006

Hong, Y., Sorooshian, S. and Hsu, K., 2004b. Cloud Patch-based Rainfall Estimation using a Satellite Image Classification Approach, 2nd Workshop of the International Precipitation Working Group, Monterey, 25-28 October 2004.

Hubbard, K.G., Mahmood, R. and Carlson, C., 2003. Estimating Daily Dew Point Temperature for the Northern Great Plains Using Maximum and Minimum Temperature. *Agron J*, 95(2): 323-328.

Kidder, S.Q. and Haar, T.H.V., 1995. Satellite Meteorology: An Introduction. Academic Press, 466 pp.0124064302

Kidder, S.Q., Kankiewicz, J.A. and Eis, K.E., 2005. Meteosat Second Generation Cloud Algorithms for use at AFWA, *BACIMO 2005, October 12 - 14, 2005*, Monterey, CA, USA.

Maathuis, B.H.P., Gieske, A.S.M., Retsios, B., Hendrikse, J.H.M. and Leeuwen, B., 2005. MSG Data Retriever: Tool for converting raw MSG SEVIRI L1.5 files into Raster-GIS or Raster image file format.

Maathuis, B.H.P., Gieske, A.S.M., Retsios, V., van Leeuwen, B. and Hendrikse, J.H.M., 2006. Meteosat-8: from temperature to rainfall. *ISPRS 2006 : ISPRS mid-term symposium 2006 remote sensing : from pixels to processes, 8-11 May 2006, Enschede, the Netherlands*

Météo-France, 2005a. O & SI SAF Project team: Atlantic Sea Surface Temperature; Product Manual, version 1.5; Nov 2005: SAF/OSI/M-F/TEC/MA/121.

Météo-France, 2005b. User Manual for the PGE01-02-03 of the SAFNWC/MSG: Scientific part SAF/NWC/IOP/MFL/SCI/SUM/01. (Issue 1, Rev.2. 65).

NOAA-NODC, 2006.Satellite Oceanography at NODC, ftp://data.nodc.noaa.gov/pub/data.nodc/pathfinder/Version5.0_Climatologies, June 2006

NOAA-NWS-CPC, 2005.Monitoring and Data, ftp://ftp.cpc.ncep.noaa.gov/wd51we/wgrib2/Windows_XP/. February 2007

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*, 9(1): 123-150.

Schröder, M., Bennartz, R., Schüller, L., Preusker, R., Albert, P. and Fischer, J., 2002. Generating cloudmasks in spatial high-resolution observations of clouds using texture and radiance information. *International Journal of Remote Sensing*, 23(20): 4247-4261.

Stowe, L.L., Davis, P.A. and McClain, E.P., 1999. Scientific basis and initial evaluation of the CLAVR-1 global clear cloud classification algorithm for the advanced very high resolution radiometer. *Journal of Atmospheric and Oceanic Technology*, 16(6): 656-681.

Strahler, A.N., 1965. Introduction to Physical Geography. John Wiley & Sons, Inc., New York.London.Sydney.65-12699