Field trial of an automated ground-based infrared cloud classification system

Emal Rumi[a], David Kerr*[b], Andrew Sandford[a], Jeremy Coupland[b], and Mike Brettle[a]

[a]Campbell Scientific Ltd, Shepshed, UK

[b]Wolfson School of Mechanical and Manufacturing Engineering, Loughborough University, UK.

Abstract: Automated classification of cloud types using a ground-based infrared imager can provide invaluable high resolution and localised information for Air Traffic Controllers. Observations can be made consistently, continuously in real time and accurately during both day and night operation. Details of a field trial of an automated, ground-based infrared cloud classification system are presented. The system was designed at Campbell Scientific ltd in collaboration with Loughborough University, UK. The main objective of the trial was to assess the performance of an automated infrared camera system with a lightning detector in classifying several types of clouds, specifically Cumulonimbus and Towering Cumulus, during continuous day and night operation. Results from the classification system were compared with those obtained from Meteorological Aerodrome Reports (METAR) and with data generated by the UK Meteorological Office from their radar and sferics automated cloud reports system. In comparisons with METAR data, a Probability of Detection of up to 82% was achieved, together with a minimum Probability of False Detection of 18%.

Keywords: cloud classification, infra-red, pattern recognition, image processing, texture analysis, convective cloud, remote sensing, cumulonimbus cloud

**Corresponding author's address: Wolfson School of Mechanical & Manufacturing Engineering, Loughborough University, Ashby Road, Loughborough, Leicestershire LE11 3TU, d.kerr@lboro.ac.uk*

1. Introduction and background

Convective clouds such as Towering Cumulus (TCU) and Cumulonimbus (CB) are associated with thunderstorms, turbulence, and atmospheric instability. Low-based clouds like CB and TCU can be dangerous to aircraft at landing and take-off, therefore meteorological monitoring and reporting is vital for air traffic and civil aviation control and safety. The Meteorological Aerodrome Report (METAR) provides data on current weather conditions in the vicinity of an airport or airfield. It contains data about cloud amount, cloud base height, thunderstorms, and other weather information. METAR reports, which don't include cloud types apart from CB and TCU when they are present, are traditionally issued by human observers every half an hour during day and night. This requires a human presence for continuous visual inspection of the sky. This visual observation is expensive, time limited, subjective and not reliable at night. Some instruments have been introduced to automate cloud base height (Costa-Surós*, et al.* (2013)) and cloud cover (Cazorla*, et al.* (2008)) estimation throughout the last few decades. However, automatic cloud type recognition was first tested in France and the Netherlands in 2006 at a few airports. Automated cloud type recognition was introduced operationally in 2011 by the Royal Netherlands Meteorological Institute (KNMI), based on detecting TCU and CB presence from lightning and precipitation radar networks.

Cloud properties are extremely variable in time and space, and there are six main features to be considered in describing the visual appearance of clouds. These features are: brightness, texture, size, shape, organisation, and shadow effects. Cloud spectral properties may change but their texture properties are unique to a given cloud type, according to Lamei*, et al.* (1994) and Pankiewicz (1995). Texture analysis of visible wavelength images of clouds obtained from ground-based observations has been demonstrated by different authors. Singh*, et al.* (2005) trained a classifier system to recognise cumulus, towering cumulus and cumulonimbus clouds using five different feature extraction methods: autocorrelation, co-occurrence matrices, edge frequency, Law's filters and primitive length. Their k-nearest neighbour (KNN) and neural network classifiers were trained using segmented visible cloud images to identify cloud types. The correct recognition rate of the combined features in classifying the three types was 72.4%. Heinle*, et al.* (2010) used visible wavelength whole-sky images to classify cloud types based on statistical properties of both grey scale texture and colour components of the observable sky. They used a KNN type classifier to distinguish between seven different sky conditions, including cumulonimbus and nimbostratus clouds. On a test run of 275 random images, independently classified by human observers, their classifier reached an average Probability of Detection (POD) of 74.6%, and for the CB/Nimbostratus class alone they achieved 85.7% POD.

A reasonable number of airports nowadays are continuously operational and some, such as East Midlands Airport (EMA) in the UK, have substantial night-time activity involving cargo aircraft. Human observation or visible wavelength cameras cannot provide the required data during the hours of darkness, so an automated, ground-based infrared (IR) cloud classification system offers the advantage of real time uninterrupted cloud monitoring to compensate for the lack of human vision at night (Rumi*, et al.* (2013)). Liu*, et al.* (2013) studied cloud cover and type based on measurement from a whole-sky IR cloud measuring system (WSIRCMS) and ceilometer. CB classification was excluded from their analysis, due to the system's inability to identify CB clouds without atmospheric electric field measurement. To the authors' knowledge there is no ground-based IR system yet available that classifies CB and TCU clouds in real time.

An experimental ground-based IR system using texture analysis was recently presented by the authors (Rumi*, et al.* (2013)). Results from the supervised classification method gave a POD of up to 90% with a Probability of False Alarm (POFA) as low as 16%. This paper presents the extended field trial results using the same system. The system was designed at Campbell Scientific Ltd (CSL) as a collaborative project with Loughborough University. The trial, located close to EMA, ran for one year from May 2013 to May 2014. The objective of the trial was to assess the viability and performance of an automated IR camera system with a lightning detector in classifying different types of clouds, specifically CB and TCU, during both day and night.

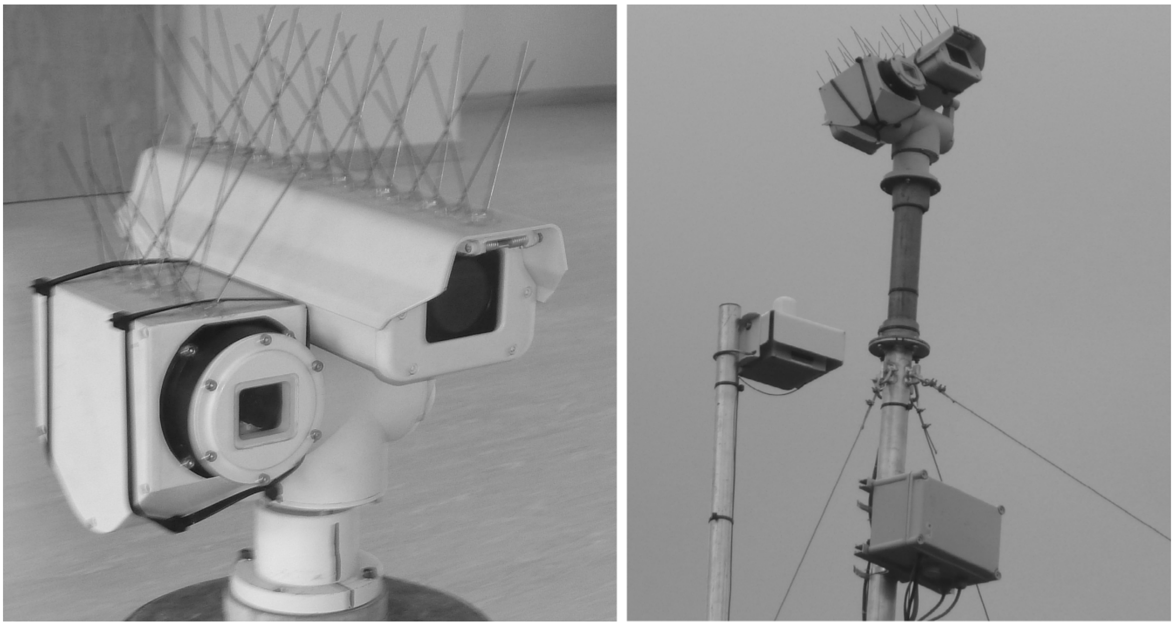

Figure 1. Visible and infrared cameras (a), P&T unit with cameras on 15m mast, with lightning detector on second mast (b).

2. Overview of trial apparatus

The trial system was set up at a location 5 km south of EMA at co-ordinates 52.79231 ° N, -1.2718007 ° W. The system consisted of both IR and visible wavelength cameras housed in a Pan and Tilt (P&T) unit on a 15 m height mast as shown in Figure 1. A NEC Thermo Tracer TS9230 IR camera was used for the main imaging task. It had an uncooled micro-bolometer providing a quarter-VGA (320x240 pixels) resolution calibrated thermal image. The standard lens provided a field of view (FOV) of 21.7 ° horizontally by 16.4 ° vertically. In addition, a F610A CCD IP ZAVIO camera was used to provide visible wavelength images with a 704x576 pixel resolution. CB clouds are identified visually based on their shape, height, accompanying lightning, and sometimes acoustically, when thunder is heard. By convention a cloud is reported as a CB if accompanied by lightning or thunder, therefore a Strike Guard lightning detector was mounted on a second mast about 2 m away from the camera mast, as shown in Figure 1 (right). The detector was fully automated and monitored cloud-to-cloud and cloud-to-ground lightning within an 8, 16 and 32 km radius. Information from the lightning detector was used to confirm the existence of CB clouds when they were embedded within other clouds or at night.

The system setup block diagram is shown in Figure 2. Camera signal and power connections as well as pan and tilt power and control were provided via an external connection box. Both IR and visible cameras are IP devices and were connected to a LAN network switch via Ethernet. The Strike Guard sensor data were communicated via a lightning proof fibre optic cable. The bi-directional fibre-optic link data were converted to RS232 serial and then to USB using the lightning-proof fibre-optic converter, the sensor being connected to the computer via USB. A secure, metal container located in a nearby building housed a five-way network switch, a computer, a MOXA Nport 5110 serial device server and a fibre to RS232 converter. An uninterruptable power supply (UPS) was added to insure continuity in case of a power cut. The system was located about 2 km from the main CSL site, therefore remote communication was established using a cellular 3G mobile network. A remote access 3G router was used to re-establish communication in the event that data calls drop out. An antenna was attached to the router and placed outside the metal container. The router was connected to the computer via the Ethernet switch. The computer was connected to the network switch and provided communication to the system locally and remotely.

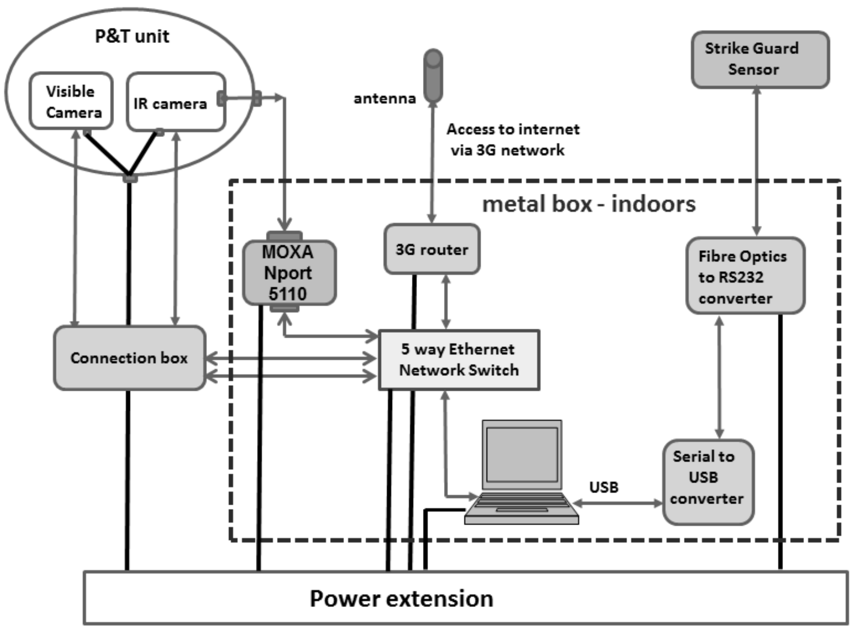

Figure 2. Trial system block diagram.

3. Details of trial procedure

The trial system was fully automated, scanning the sky three times around approximately 360º every half an hour. The method used for capturing images was similar to the way a human observer inspects the sky, taking into account the limitations in the FOV of the IR camera and the P&T operation. Normally, a manual observation would be made every half an hour, repeated in several directions to complete a 360º scan of the prevailing sky conditions. A human observer would watch for lightning and record any sounds of thunder. To increase the FOV without reducing spatial resolution, a series of nine 320x240 images with overlap were taken and then stitched together in the form of a 3x3 rectangle, as shown in Figure 3. Adjacent edges were re-aligned and merged to form a single panoramic sky image. This method of operation and image capture was applied to both the IR and visible camera as they moved and captured images simultaneously. Visible images were used for verification and only IR images were processed for cloud classification. Following the method described in Rumi*, et al.* (2013), one image was captured every 13 seconds, and a total of 15 stitched panoramic images were generated and classified every half an hour. Thus real time data capture and classification was achieved day and night.

The objective of the system was to classify CB and TCU clouds as they appeared in real time. A supervised classification method was selected and therefore a large amount of training data needed to be captured, inspected by an experienced observer and separated into 8 classes (CB, CL1, CLS, HC, MX, OTH, OVC, and TCU) as defined in Table1. Ground truth data were generated from inspection of the visible camera images and accompanying local METAR reports. A total of 420 images were used to create the training data set, which was used at the start of the trial in May 2013.

Table 1 Cloud classes

Class number and name	Symbol	Definition
1 Cumulonimbus	CB	Tall vertical extent of cloud with formation of ice crystals at the top. Can form an anvil shape
2 Small Cumulus	CL1	Represents small clouds known as cumulus humilis, cumulus fractus, or cumulus mediocris
3 Clear Sky	CLS	Clear sky with no cloud
4 High Cloud	HC	Includes cirrus, cirrocumulus, cirrostratus and contrails
5 Mixed Clouds	MX	Mixed layers and types of clouds, with no CB or TCU present
6 Others	OTH	Any other unidentifiable cloud type, not covered by the other 7 classes
7 Overcast	OVC	Cloud cover complete due to one or more layers with no gaps. No CB or TCU present
8 Towering Cumulous	TCU	Cumulus cloud of strong vertical development, also known as cumulus congestus

A further training data set was created in June 2013 after manual analysis of the initial trial data. This training data set contained a total of 455 observations of the 8 cloud types. This data set was used from 21 June 2013 onwards. Subsequent tests on classification results showed some increase in false alarms during OVC conditions using this set. A decision was therefore made to increase the number of OVC samples to create a better balance with the number of CBi. On 17 October 2013, a new training data set was introduced with 500 observations, and this was used until the end of the trial in April 2014.

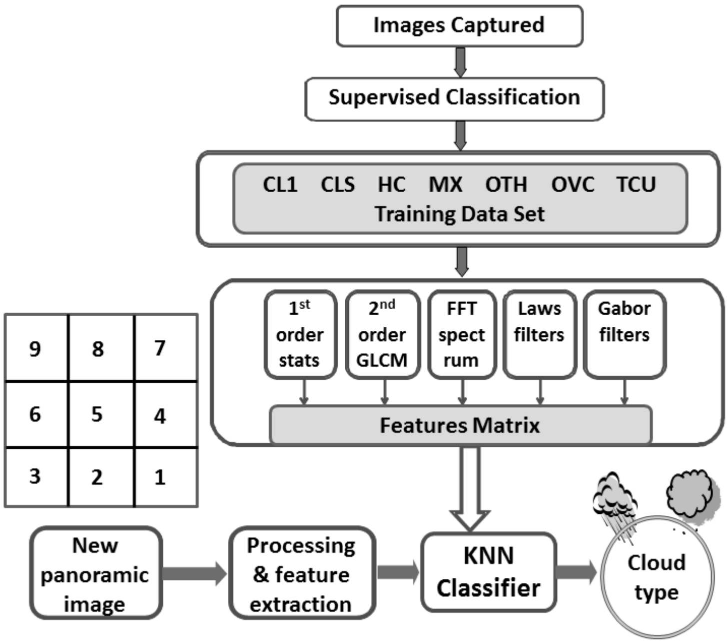

Figure 3. Classification methodology.

The classification process is summarised in Figure 3, where a total of 45 image features were extracted for each observation in the training data set. The features used were a combination of measures relating to image texture, including grey-level co-occurrence matrix (GLCM) statistics, Fourier Transform (FFT) power spectrum properties and energy outputs from Laws and Gabor digital filter banks. Features were normalised by dividing by their sample standard deviations and were then grouped within a training set according to cloud class, determined independently by trained observers. In order to improve discrimination between CB cloud and all other 7 cloud types, the number of features used by the classifier was reduced to 27 by selecting those with the highest Fisher's Discriminant Ratio (FDR). A weighted k = 20 nearest neighbours (KNN) method was used to classify subsequent test images, weights being applied to each nearest neighbour "vote" in the training set space to take account of unequal numbers of observations of each of the 8 classes. Detailed descriptions of the features used and their extraction methods can be found in Rumi*, et al.* (2013).

During the course of the trial, automated METAR data was available every half an hour. In order to compare the IR trial data with the local METAR data from EMA, capture times had to be synchronised as far as possible. The IR system was designed to generate cloud types automatically at every 20 minutes and 50 minutes past the hour reflecting the timing used to generate the METAR data.

The system scanned the sky three times in every half hourly interval to generate five panoramic image sets during each scan. The complete classification process is shown in Figure 4. Each panoramic image was classified into one of the eight cloud types immediately after it was generated. Images containing the sun or moon were treated with a special algorithm to remove any localized thermal influence, as described in Rumi*, et al.* (2013). It took about two minutes to capture nine images, stitch, enhance, and classify the resulting panoramic image. Each scan took about 10 minutes. At the end of every 10 minutes the lightning detector data was read. Using the lightning detector data together with the cloud classifications from the five panoramic images, a decision was produced following the heuristic rules 1-5 defined below.

Heuristic rule set applied to each scan (every 10 minutes):

1. If any strike recorded by lightning detector within the last 10 minutes, then report cloud type as CB.
2. If one or more CBi classified in any one scan, then report cloud type as CB.
3. If one or more CBi and one or more TCUs classified in any one scan, then report cloud type as CB.
4. If two or more TCUs classified in any one scan, then report cloud type as TCU.
5. Otherwise report (Non-Significant Cloud) NSC.

This process was repeated three times and the three decisions were combined to generate one class type for the specific half an hour, for comparison with the equivalent METAR report from EMA. The three scan decisions have one of three possible outcomes: CB, TCU or NSC. A score of 3, 6, and 9 was applied to the three scans respectively, so that the most recent image had the highest weighting. The final decision for the half-hour period was obtained from these by majority vote. In the event of a tie, priority was given to a CB and then to a TCU condition.

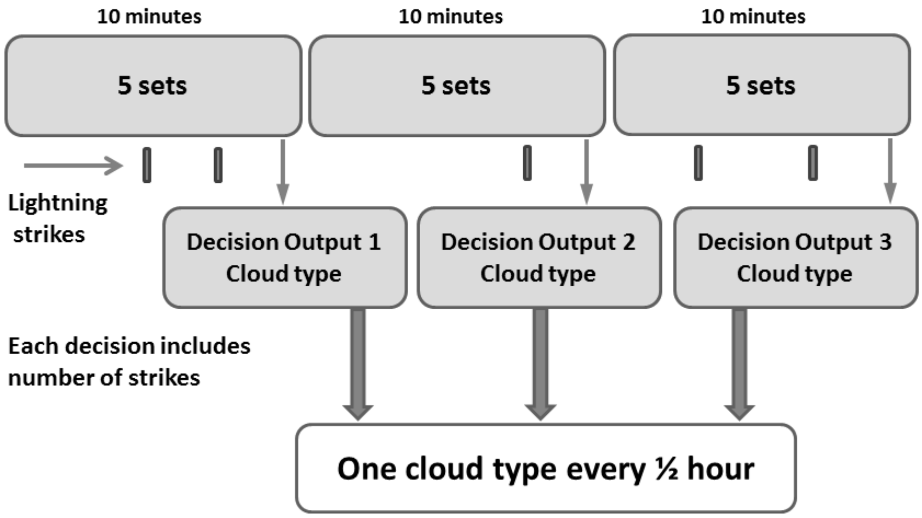

Figure 4. Automated cloud classification process.

4. Results from IR system

The field trial ran for one year, and the system was automated to generate a cloud type for every half an hour, potentially providing 17,520 observations in total. Due to problems with mechanical limitations of the pan and tilt unit and software limitations in handling the large number of data files, in some cases the system got out of sync with the METAR data from EMA. It was important for comparison purposes to synchronize the cloud classification results exactly with the METAR reports, therefore some of the IR system data had to be rejected. This left us with 13,197 usable classifications samples, or just over 75% of the total annual data. Figure 5 shows the total number of samples available per month. Data from June, July, and October were limited due to the technical issues mentioned above. The total number of CB/TCU events reported by EMA during that period was 217, out of which there were 26 TCUs. The Strike Guard system detected lightning and/or thunder alarms, most of which coincided in time with CB reports from EMA. However, 19 alarms appeared at a time when no CB/TCU events were reported by human observation at EMA. This represented about 10% of the total reported CB by EMA for the whole year. These additional CBi were added to the METAR reports to create a ground truth of observed convective clouds and as a result, the total number of CB/TCU events during the trial period was considered to be 236, as shown in Figure 6. CB events were reported in every month apart from June, because the IR system was unfortunately not operational during the few times in that month when CBi were present. More CBi than TCUs were reported overall, TCUs being reported in 6 out of the 12 months, and CBi being reported in every month of the trial period. Most convective clouds appeared during the summer months as expected, based on previous observation records at EMA for the East Midlands region.

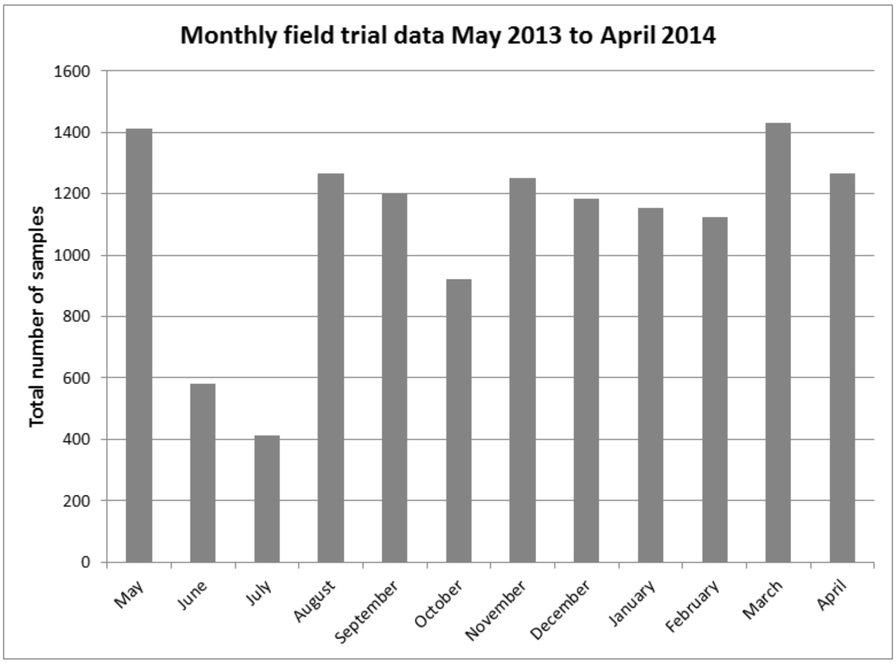

Figure 5. Total number of samples *per* month.

A typical meteorological practice for evaluating cloud classification calculates the Probability of Detection (POD), and the Probability of False Detection (POFD) for the classifier outcomes, following for example Barnes*, et al.* (2009) , where:

POD = true positives (TP) / (true positives (TP) + false negatives (FN)), and

POFD = false positives (FP) / (true negatives (TN) + false positives (FP))

Since the main interest was in the presence of CB and TCU cloud types, POD and POFD were assessed for two cases in this study; one case was for evaluating CB cloud classification and the other was for evaluating the combined classification of CB and TCU clouds. The contingency construction for the field trial result for the whole year was based on using EMA and Strike Guard reports as the ground truth; the results are shown in Table 2. The contingency table for CB/TCU and CB for every month is shown in Tables 3a and 3b.

5. Data analysis and comparison with METAR and Met Office data

From Tables 3a and 3b, the overall yearly POD rate achieved was 75% for CB/TCU and 73% for the CB category. The overall POFD was 23% for CB/TCU and 20% for CB category. In general, the automated IR system recorded more CB/TCU than human observers, which caused a relatively high number of false positives. There could be several reasons for this:

- On some occasions, human observers reported a thunderstorm (TS or VCTS) but failed to include CB in the METAR reports.
- During the trial period, there were occasions where Strike Guard reported lightning but there were no CB, TS or VCTS reports in the METAR. These observations were recorded during daylight, and it is more difficult to see lightning during the day. CBi may at these times have been embedded in other cloud and observers may not have heard any thunder from the inside of the traffic control tower. It should be noted that Strike Guard is an extremely reliable commercial instrument. To prevent false alarms, the device requires an optical signal to coincide with a changing magnetic field signal before reporting lightning.
- Some apparent false positives were reported at night time by the IR system, but EMA reported heavy rain and no convective clouds. It is difficult for human observers to detect CB clouds at night, especially if there are no thunderstorms or lightning present.
- The TS will only be reported as present weather in METAR if a thunderstorm occurred during the 10 minute period up to the METAR time, in line with recommendations. The automated IR system however, reported a CB whenever a thunderstorm alarm was captured during the 30 minutes of the report.
- METARS report TS/VCTS and therefore a CB, for up to 8 km from the Aerodrome Reference Point (ARP); however the automated IR system would report TS at 8 , 16 and 32 km distance from the location of the Strike Guard sensor.
- On a few occasions, hail and thunder were observed at the location of the trial and the automated IR system reported a CB, however there were no convective clouds reported in the METAR data.

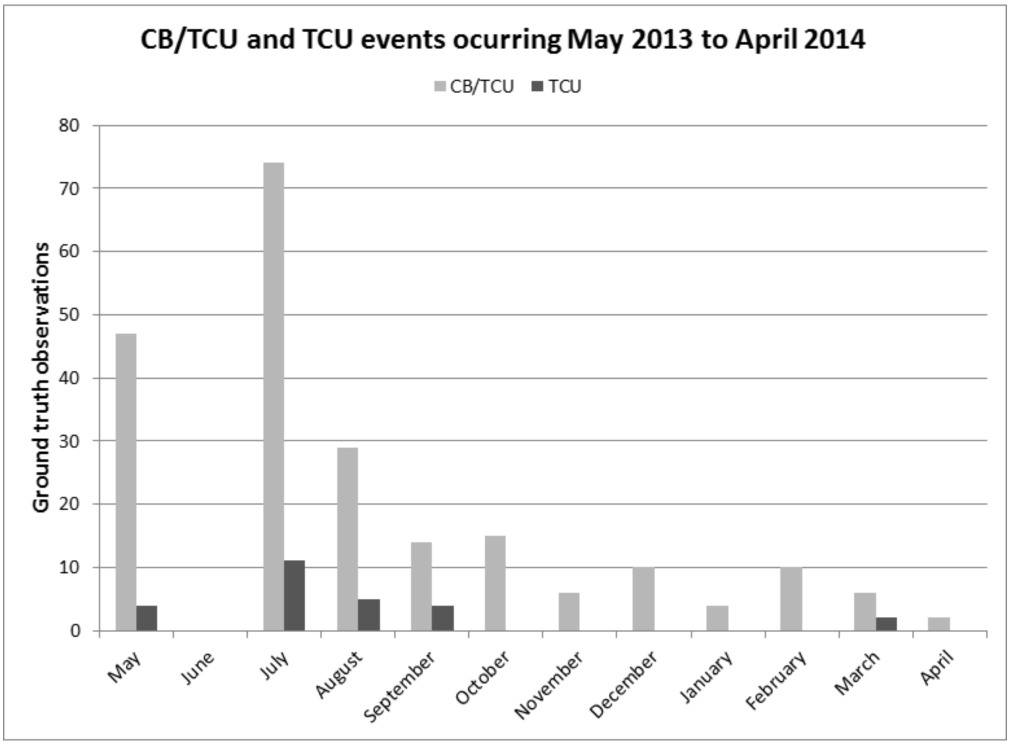

Figure 6. Total number of observed cumulonimbus (CB) and towering cumulus (TCU) by the East Midlands Airport (EMA) and Strike Guard *per* month.

Table 2 Contingency table for CB and CB/TCU field trial results for one year

Classified as	Event Observed			Classified as	Event Observed		
	CB	No CB	Marginal total		CB/TCU	No CB/TCU	Marginal Total
CB	153	2555	2708	CB/TCU	178	3008	3186
No CB	57	10,432	10,489	No CB/TCU	58	9953	10,011
Marginal total	210	12,987	13,197	Marginal total	236	12,961	13,197

According to Bluestein (1993), CB clouds most often develop from large cumulus clouds, and TCU clouds are generally at an intermediate stage between cumulus of strong vertical development and a CB. Although TCU clouds are short lived compared to CB clouds, it was still surprising to see METAR reports of TCU as low as 26 compared to 191 CB reports during the trial period. It was evident, during both the proof of concept of the IR system and during the trial, that human reporting of TCU was generally at a lower rate than automated reporting, and that this contributed to the high number of false positives encountered in the trial. There were 462 false positives due to TCU reports in the CB/TCU category as can be seen in Table 3a. Reporting TCU clouds is more uncertain and subjective, and it is in general not possible to identify a TCU cloud that is obscured by another cloud mass. It was also confirmed that at some airports including EMA, if a TCU has a base higher than 5000 feet (1524 m) it will not be reported in METARs.

Table 3a Contingency table for CB/TCU field trial results per month

Month	TP	TN	FN	FP	Monthly Observation Total	POD (%)	POFD (%)
May	41	1118	10	241	1410	80	18
June	0	486	0	95	581	N/A	16
July	73	209	9	123	414	89	37
August	26	685	8	548	1267	76	44
September	11	927	3	259	1200	79	22
October	7	678	8	230	923	47	25
November	3	1001	3	242	1249	50	19
December	4	894	6	278	1182	40	24
January	4	952	1	196	1153	80	17
February	2	898	8	214	1122	20	19
March	5	1093	2	329	1429	71	23
April	2	1012	0	253	1267	100	20
Entire year of trial	178	9953	58	3008	13197	**75**	**23**

The number of false negatives for CB/TCU and CB-only categories were 58 and 57 respectively. This indicates that almost all false negatives were CBi. After inspection of the visible images, it was clear that in some situations, CB clouds were far-off and captured near the horizon, and were visible only in the lower part of the image. The high resolution and relatively small FOV of the IR panoramic image didn’t contain the CB cloud, and therefore it was not reported. Increasing the FOV of the camera would allow the capture of far-off convective clouds as well as allowing capture of a complete extent of TCU clouds, thus reducing false negatives rates.

Table 4b Contingency table for CB field trial results per month

Month	TP	TN	FN	FP	Monthly Observation Total	POD (%)	POFD (%)
May	35	1173	12	190	1410	74	14
June	0	521	0	60	581	N/A	10
July	64	270	7	73	414	90	21
August	21	822	8	416	1267	72	34
September	8	962	2	228	1200	80	19
October	7	707	8	201	923	47	22
November	3	1012	3	231	1249	50	19
December	4	895	6	277	1182	40	24
January	4	965	1	183	1153	80	16
February	2	919	8	193	1122	20	17
March	3	1129	2	295	1429	60	21
April	2	1057	0	208	1267	100	16
Entire year of trial	153	10432	57	2555	13197	**73**	**20**

To better assess the performance of the CB and TCU classification it was necessary to compare the results with other ground-based classifications systems that are used for AUTO METAR generation. It is difficult to find one system that operates in the same way i.e. is a ground-based IR system. At KNMI, the automated observation system of present weather and CB/TCU information has been derived from lightning and precipitation radar reflectivity data since 2006, and recently improved by using METEOSAT satellite data. A POD of 65.2% and a POFA of 35.4 % were reported for CB/TCU cloud detection by (Carbajal*, et al.* (2009)). Although not yet achieved to the authors' knowledge, it has been suggested by KNMI, that for air traffic control, the desired POD is greater than 80% and POFA is less than 20%. To evaluate our results, the receiver operating characteristic (ROC) was compared with these KNMI requirements. The ROC describes the relationship between the False Positive Rate (FPR) and the True Positive Rate (TPR), Metz (1978). POD is equivalent to Sensitivity and is the same as TPR, however FPR is equal to (1 – Specificity), where:

$$\text{Specificity} = TN / (TN + FP), \text{ and}$$

$$\text{Sensitivity} = TP / (TP + FN)$$

A graphical representation of the yearly field trial result was compared with KNMI requirements and is shown for both CB/TCU and CB cases in Figure 7. The upper-left area denoted by the dotted lines represents the ideal operating points as defined by KNMI. The CB and CB/TCU results from the IR trial fall just outside this area, with the CB class operating point lying on the POFD 20% boundary.

A similar system to the KNMI system was trialled for six months in the UK between June and November 2010 by the Met Office and the National Air Traffic Services (NATS). NATS data were based on a test at 24 airports, where 58,918 samples were gathered and analysed. Trial results showed a POD of 57.8% for CB/TCU and only 36% for CB, however a small false positive rate of 9% for CB/TCU cases and 3% for the CB category was reported by Hord (2011). An improved version of this system became operational in the UK before the time of the current trial. This gave the opportunity to compare the IR system with another automated system at the exact location of the current trial; however this system cannot be considered as a ground truth, based on the system UK trial results mentioned above. Therefore data from IR automated and radar automated systems had to be evaluated using EMA METARs as the ground truth. Data was available from the Met Office for the trial site from July 22nd until August 22nd 2013. From the data in Table 3a, the total number of TP +

FN of convective clouds reported in July and August was 116, of which103 occurred during the period under consideration. This made the selected sample very useful and representative for CB/TCU classification analysis. After removing the missing data from METARs, Met Office and IR trial for this period, the total number of data samples available to be analysed was 1,417 containing 414 samples from July and 1003 samples from August.

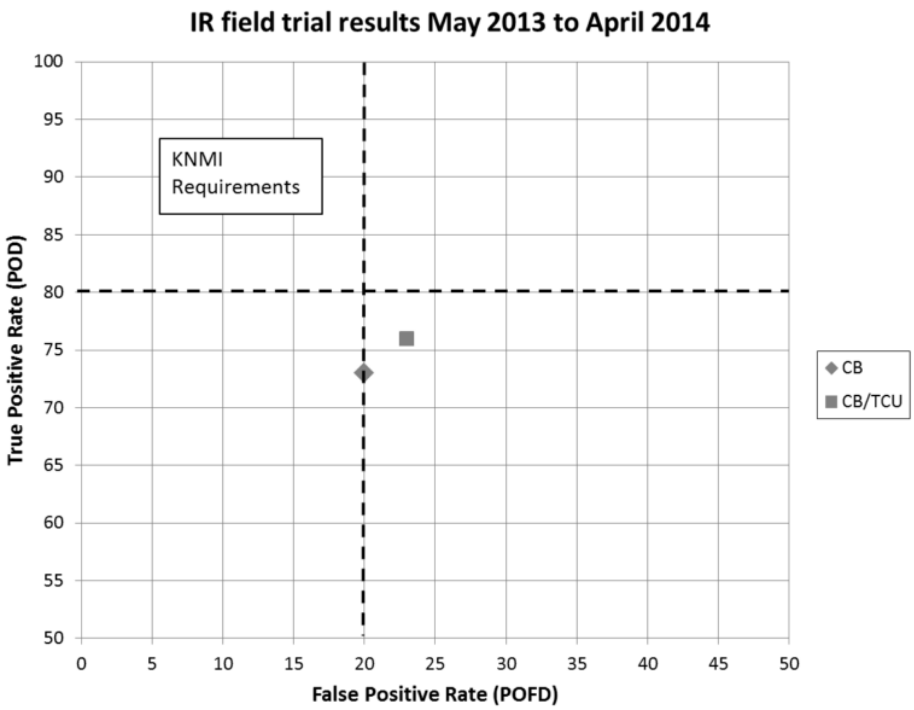

Figure 7. Receiver operating characteristic (ROC) curve for infrared (IR) trial and Met Office trial compared.

Data from the Met Office automated cloud reports and the automated IR system were compared with METAR data, assuming the human observer METARs to be the ground truth, and results are shown in Table 4.

Table 5 MET Office data and IR system data comparison results using METAR as ground truth and with old training data

July 22 – August 22 CB/TCU	TP	TN	FN	FP	Totals	POD (%)	POFD (%)
July	71	272	11	60	414	87	18
August	24	699	10	270	1003	71	28
IR Trial	95	971	21	330	1417	**82**	**25**
MET Radar	80	1126	23	188	1417	**78**	**14**
July 22 – August 22 CB only	**TP**	**TN**	**FN**	**FP**	**Totals**	**POD (%)**	**POFD (%)**
July	63	299	8	44	414	89	13
August	15	775	14	199	1003	52	20
IR Trial	78	1074	22	243	1417	**78**	**18**
MET Radar	69	1218	22	108	1417	**76**	**8**

A POD of 82% was achieved with the automated IR system and a POD of 78% with the Met Office automated system. The IR trial system demonstrated a much better true positive rate and an equivalent false negative rate to the Met Office automated system. The automated IR system showed a high value of false positive rate with a POFD of 25% for the CB/TCU case and a more acceptable 18% for the CB case. The graphical representation of these results is shown on the ROC plot for both CB/TCU and CB in Figure 8, where it can be seen that both systems have performed very closely to the KNMI criteria.

There were more than double the number of CBi and TCUs reported by both automated systems compared to the number reported by human observation at EMA, which is in line with the IR yearly results reported here. Both systems reported TS at times when there were no TS or CB reports in METARs.

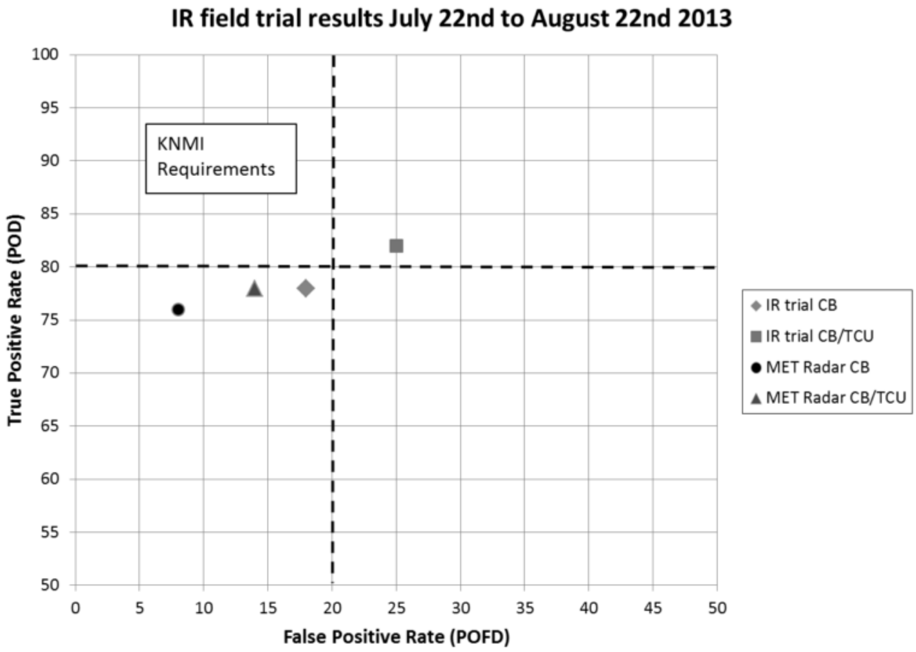

Figure 8. Receiver operating characteristic (ROC) curve representing a comparison of Met Office Radar–automated system and automated IR trial system for 1month of data from 22 July to 22 August 2013.

6. Conclusions

A field trial of an experimental automatic ground based IR cloud classification system has been conducted and presented. The results demonstrate the feasibility of classifying CB and TCU clouds along with other sky conditions using high resolution IR images. A CB/TCU POD of 75% was achieved for the whole year, with a maximum of 82% over one summer month. Results were based on using human observations as the ground truth, but human observation is often unreliable and cannot be considered as 100% accurate. By using lightning detection in both IR and Met Office automated systems, it was clear that at least 10% of CBi accompanied by lightning and/or thunder were missed by human observation. It is very possible therefore, that many more convective clouds were either embedded or not seen during the night time, and were not reported. This likely explains the high number of false positives reported by both automated systems. In general, comparison of any automated cloud recognition system with that of human observers may never show PODs much higher than those achieved in the current trial, due to lack of consistency and variations in experience and objectivity of the observer.

Both automated systems worked more reliably at night than human observers, however the Met Office system relies on a large network of radars and lightning detectors, that is expensive to install, operate and maintain. It can cover a very large area on a national scale, but it is sensitive to precipitation and it does not perform well during snow and hailstones according to Hord (2011). In contrast, the IR system worked well during snow, rain and hail, and was easy to install, run and maintain. However it covers a much smaller area than the Met Office automated system, although it does provide local results out to a range that is more in tune with the requirements of airport operators. In all the analyses, the automated IR system performed at a similar level to that of the Met Office automated system with regard to true positives, and slightly worse for false positives. Results showed comparable performance with Met Office automated reports but there is still room for improvement in order to get the system to operate at a more advantageous point in the ROC space.

Further development work could thus include reducing the number of texture features, as a precise selection of the most influential measures would enhance the classifier performance still further. Increasing the IR camera FOV by replacing its lens would help to provide the larger panoramic images needed to gather more complete CB cloud data. Calculation of elevation temperature would also help in separating high clouds from low clouds, and measuring cloud base height would enable the introduction of accurate automated cloud cover estimation from the IR images. Finally more reliable hardware and software would help to improve the performance of the system overall.

Acknowledgments
The initial part of this work was accomplished under a Knowledge Transfer Partnership (KTP) program with Loughborough University. We wish to acknowledge the UK Technology Strategy Board (TSB) for funding and support of the project that initiated this work. We would like to thank the Met Office team who provided Campbell Scientific Ltd with the automated METAR data.

References

Barnes LR, Schultz DM, Gruntfest EC, Hayden MH ,Benight CC. 2009. Corrigendum: False Alarm Rate or False Alarm Ratio? *Weather and Forecasting.* 24: 1452-1454.

Bluestein HB. 1993. Synoptic dynamics meteorology in mid-latitudes. Volume II: Observations and theory of weather systems. Oxford University Press New York

Carbajal HC, Roelofs G ,Roebeling R. 2009. 'Detection of Cb and TCu clouds using MSG-SEVIRI cloud physical properties and weather radar observations.' Thesis Institute for Marine & Atmospheric Research Utrecht and Royal Netherlands Meteorological Institute

Cazorla A, Olmo FJ ,Alados-Arboledas L. 2008. Development of a sky imager for cloud cover assessment. *Journal of the Optical Society of America A.* 25: 29-39.

Costa-Surós M, Calbó J, González JA ,Martin-Vide J. 2013. Behavior of cloud base height from ceilometer measurements. *Atmospheric research.* 127: 64-76.

Heinle A, Macke A ,Srivastav A. 2010. Automatic cloud classification of whole sky images. *Atmospheric Measurement Techniques.* 3: 557-567.

Hord C. 2011. 'Report on the use of remote sensing capabilities to further automate aeronautical meteorological reports in the United Kingdom.' In Aerodrome meteorological observation and forecast study group (AMOFSG), Montreal, 26-30 September 2011 2011. AMOFSG.9.IP.003.5.en.doc

Lamei N, Hutchison KD, Crawford MM ,Khazenie N. 1994. Cloud-type discrimination via multispectral textural analysis. *Optical Engineering.* 33: 1303-1313.

Liu L, Sun X-j, Gao T-c ,Zhao S-j. 2013. Comparison of Cloud Properties from Ground-Based Infrared Cloud Measurement and Visual Observations. *Journal of Atmospheric and Oceanic Technology.* 30: 1171–1179.

Metz CE. 1978. Basic Principles of ROC Analysis. *Seminars in Nuclear Medicine.* 8: 283-289.

Pankiewicz GS. 1995. Pattern recognition techniques for the identification of cloud and cloud systems. *Meteorological Applications.* 2: 257-271.

Rumi E, Kerr D, Coupland JM, Sandford AP ,Brettle MJ. 2013. 'Automated cloud classification using a ground based infra-red camera and texture analysis techniques.' In Remote Sensing of Clouds and the Atmosphere XVIII and Optics in Atmospheric Propagation and Adaptive Systems XVI, Dresden, Germany September 23rd, 2013 2013. SPIE

Singh M ,Glennen M. 2005. Automated ground-based cloud recognition. *Pattern Analysis & Applications.* 8: 258-271.