Clouds over the summertime Sahara: An evaluation of Met Office retrievals from Meteosat Second Generation using airborne remote sensing

John C. Kealy[1†], Franco Marenco[1], John H. Marsham[2,3], Luis Garcia-Carreras[2], Pete N. Francis[1], Michael C. Cooke[1], and James Hocking[1]

[1]Met Office, Exeter, United Kingdom
[2]School of Earth and Environment, University of Leeds, Leeds, United Kingdom
[3]National Centre for Atmospheric Science
[†]Now at College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, United Kingdom

Correspondence to: John Kealy (J.Kealy@exeter.ac.uk)

Abstract.

Novel methods of cloud detection are applied to airborne remote sensing observations from the unique Fennec aircraft dataset, to evaluate the Met Office derived products on cloud properties over the Sahara based on the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) on-board the Meteosat Second Generation (MSG) satellite. Two cloud mask configurations are considered, as well as the retrievals of cloud-top height, and these products are compared to airborne cloud remote sensing products acquired during the Fennec campaign in June 2011 and June 2012. Most detected clouds (67 % of the total) have a horizontal extent which is smaller than a SEVIRI pixel (3 km x 3 km). We show that, when partially cloud-contaminated pixels are included, a match between the SEVIRI and aircraft datasets is found in 80 ± 8 % of the pixels. Moreover, under clear skies the datasets are shown to agree for more than 90 % of the pixels. The mean cloud field, derived from the satellite cloud mask acquired during the Fennec flights, shows that areas of high surface albedo and orography are preferred sites for Saharan cloud cover, consistent with published theories. Cloud-top height retrievals however show large discrepancies over the region, which are ascribed to limiting factors such as the cloud horizontal extent, the derived effective cloud amount, and the absorption by mineral dust. The results of the CTH analysis presented here may also have further-reaching implications for the techniques employed by other satellite applications facilities across the World.

1 Introduction

The challenge of observing atmospheric processes over the Sahara desert remains a fundamental obstacle in our understanding of the climate system of Africa, a key region for diagnosing global radiation budgets (Allan et al., 2007), aerosol transport (Schepanski et al., 2009), Atlantic hurricane activity (Dunion and Velden, 2004), and climate change (Giorgi, 2006). Much of this difficulty arises from a lack of in situ observational networks across vast expanses of uninhabited, inhospitable desert, leading to difficulties in the investigation of important climatological features such as the Saharan heat low (Lavaysse et al.,

2009). This observational deficit also has a large impact on the accuracy of numerical weather and climate models (Agustí-Panareda et al., 2010; Garcia-Carreras et al., 2013).

The Saharan Atmospheric Boundary Layer (SABL) is one of the deepest on Earth, often reaching 5–6 km by afternoon in the summertime (Garcia-Carreras et al., 2015). Despite the extremely low relative humidities near the surface, the exceptional

5   depth of the SABL can often lead to cloud formation at its top (Parker et al., 2005; Cuesta et al., 2009), with implications for radiative balance (Ramanathan et al., 1989) and dust transport (Birch et al., 2012). These clouds are typically the result of thermal updrafts in the well-mixed SABL, and exhibit limited depth due to being capped by a strong inversion that marks the transition into the free troposphere (Cuesta et al., 2009). Observational case-studies show that albedo anomalies drive boundary-layer circulations in the Sahara and it has been hypothesised that these drive cloud formation (Marsham et al., 2008;

10   Cuesta et al., 2009), but direct observations of this are limited (Ryder et al., 2015). Modelling studies by Huang et al. (2010) and Birch et al. (2012) support this hypothesis, but the exact locations most favourable for cloud formation across the Sahara remain unclear.

Since ground-based observations in this sparsely populated part of the World are rare, satellite products are invaluable for investigating the Saharan atmosphere. The detection and transport of mineral dust in particular has been widely investigated

15   using both geostationary satellites and polar-orbiters (Schepanski et al., 2007; Banks and Brindley, 2013; Banks et al., 2013). Geostationary data, such as that provided by the Meteosat Second Generation (MSG) satellite, have proven very useful in these studies because of the high temporal resolution it offers for a given area, allowing the tracking of features as they cross the desert or pass into other geographical zones (Schepanski et al., 2012). Thus-far however, satellite-based studies specific to cloud over the Sahara have been limited to polar-orbiter data (Stein et al., 2011).

20   A variety of full-disk geostationary cloud products are readily available from the Satellite Applications Facility to support Nowcasting and Very Short Range Forecasting (NWC-SAF, Derrien and Le Gléau, 2005; Derrien et al., 2013), the National Oceanic and Atmospheric Administration (NOAA, Schmit et al., 2002; Chang et al., 2010), the Met Office (Saunders et al., 2006; Hocking et al., 2011), and many other satellite applications facilities across the World. Recently these products have been the subject of cross-collaborations, under the framework of the International Clouds Working Group (ICWG). Before

25   ICWG[1] was formed in 2014, a series of annual workshops known as the cloud retrieval evaluation workshops (CREWs) were held, in order to discuss and intercompare the results of the products (Hamann et al., 2014; Roebeling et al., 2014). Such products include cloud mask, cloud-top height, cloud-top temperature, cloud optical thickness, and cloud effective radius, and the task is to evaluate them using observations from other sources; this is a necessary step to authenticate the robustness of the products, or reveal any weaknesses they may have (Derrien et al., 2005). This is particularly true in an environment such as

30   the Sahara where a bright land surface and large atmospheric dust loads may make retrievals of cloud quantities challenging. During CREW-4, the Sahara was identified as one of three areas in the SEVIRI full-disk image with the largest discrepancy between the members in an intercomparison of 12 different cloud mask products (Roebeling et al., 2014).

As part of the Fennec programme (Washington et al., 2012), the Facility for Airborne Atmospheric Measurements (FAAM) BAe-146 research aircraft was based in Fuerteventura (Canary Islands) in June 2011 and 2012 and performed 30 research flights

---

[1]http://www.icare.univ-lille1.fr/crew/

(Fig. 1, Ryder et al., 2015). Fennec led to new quantification of the boundary-layer processes leading to clouds in the Sahara (Garcia-Carreras et al., 2015; Ryder et al., 2015), demonstrated the role played by downdrafts generated from evaporation of precipitation falling from clouds over the Sahara (Marsham et al., 2013; Trzeciak et al., 2016), and showed that the day-to-day variability in top-of-atmosphere radiative heating is largely a function of total column water vapour and not dust, suggesting that clouds may be key to the radiative balance of the Saharan Heat Low (SHL, Marsham et al., 2016).

The aim of the present paper is to analyse and evaluate the Met Office cloud products based on MSG over the Sahara, by comparing the satellite dataset to airborne cloud retrievals. A description of both the airborne and satellite datasets is provided in Sect. 2, along with a description of the methods used. Two Met Office cloud products for MSG are then investigated individually, the cloud mask product (Sect. 3.2) and the cloud-top height product (Sect. 3.3). We also consider the ideas proposed by Marsham et al. (2008) and Cuesta et al. (2009) regarding the role of surface albedo and topography in cloud formation, within the context of these datasets (Sect. 3.4). Sect. 4 discusses the results and Sect. 5 summarises our conclusions.

2   Datasets

2.1   Satellite derived products

The cloud comparison described in this paper is based on data from the Meteosat Second Generation (MSG) Spinning Enhanced Visible and InfraRed Imager (SEVIRI) and derived cloud products produced by the Met Office operational systems (Saunders et al., 2006). SEVIRI generates an image every 15 minutes using twelve spectral channels (Schmetz et al., 2002) and within the latitude band of interest in this study (approximately $15°$ N – $25°$ N) the spatial resolution is close to 3 x 3 km^2. A constant geo-referencing offset of 1.5 km to the North and West in SEVIRI level 1.5 data has been identified by EUMETSAT, and this has been corrected for in Lat/Lon grids used in this study.

In order to derive information from SEVIRI, the first step is to identify which pixels contain cloud. The cloud mask developed at the Met Office by Hocking et al. (2011) is chosen here as the basis for comparison with the aircraft cloud data. This mask is widely used at the Met Office for the derivation and assimilation of a series of products, and is derived by applying a variety of threshold tests which use the observed brightness temperatures (T_B), reflectances, and simulated clear-sky radiances to infer whether cloud is present in each SEVIRI pixel. These tests are summarised in Table II of Hocking et al. (2011) along with a description of the scheme and a validation against both the Derrien and Le Gléau (2005) cloud mask for SEVIRI and the MODIS cloud mask.

Here we use two separate configurations of the cloud mask. The first detects a cloud for a given pixel if one or more of any of the tests in Hocking et al. (2011) returns "true". It is intended primarily for applications that require a cloud-free pixel, where even the effects of a small amount of cloud-contamination is undesirable, and we therefore refer to this configuration hereafter as the "*AllCloud*" mask. The second configuration of the cloud mask, hereafter referred to as the "*CloudRetrieval*" mask, is intended as a detection tool for those pixels where it is feasible to carry out cloud property retrievals. Pixels which correspond to low fractional coverage or low cloud optical thickness are less radiatively significant in the infrared, and an attempt to retrieve a cloud parameter (e.g. cloud-top height or cloud-top temperature) for these pixels can produce misleading

results. A subset of the tests described by Hocking et al. (2011) have a general tendency to identify such pixels. These consist of three types of spatial coherence test using the 0.8 μm, 10.8 μm and HRV channels, and a test known as the 15-minute temporal differencing test. The spatial coherence tests work on the principle that cloud edges tend to exhibit large standard deviations in their brightness temperature and reflectance values, while the 15-minute temporal differencing test identifies a pixel as cloudy

5 when its brightness temperature decreases by more than a predetermined threshold value since the previous observation time. The radiatively less-significant cloudy pixels (i.e. those which are flagged only by one or more of these four tests and not by any other cloud test) are excluded from the *CloudRetrieval* mask (whereas they have been included in the *AllCloud* mask along with all other cloud-flagged pixels, as noted above). For this reason the *AllCloud* mask will always contain more cloudy pixels than the *CloudRetrieval* mask.

10 The second product investigated here is cloud-top height (CTH). In order to calculate satellite-derived CTH, the algorithm incorporates numerical weather prediction (NWP) model profiles and performs radiative transfer calculations, which in this study are based on the Met Office Unified Model (MetUM). Over the African continent, the Global Model (GM) is currently the chosen configuration of the MetUM used in the cloud retrievals. By inputting vertical atmospheric profiles from the GM into a fast radiative transfer model, radiance information can be simulated at discrete levels throughout the depth of the at-

15 mosphere. The radiative transfer model used at the Met Office is known as RTTOV (v11, Saunders et al., 1999). Land-based surface emissivity values are incorporated into this model using the 'UWiremis' IR emissivity atlas (Borbas and Ruston, 2010), averaged over each model grid box.

In determining the cloud-top height (CTH), the Met Office algorithm uses three spectral channels in the infrared, 10.8 μm, 12.0 μm and 13.4 μm. Firstly, a slightly updated version of the minimum residual technique described in Eyre and Menzel

20 (1989) is employed. This method uses a two-parameter state vector, the active variables being the cloud-top pressure p and the effective cloud amount (N), with the latter being the product of the true cloud fraction and the cloud emissivity (and which is assumed independent of wavelength). For each RTTOV model level, we use an error-weighted version of Eyre and Menzel's Eq. (4) to derive a profile of best-fit effective cloud amount for each cloud-top pressure p:

$$N(p) = \frac{\sum_j (R_j^m - R_j^c)[R_j^o(p) - R_j^c]/\sigma_j^2}{\sum_j [R_j^o(p) - R_j^c]^2/\sigma_j^2} \qquad (1)$$

25 where R_j^m is the measured radiance in channel j, R_j^c is the RTTOV calculated clear sky radiance in channel j, $R_j^o(p)$ is the RTTOV calculated blackbody radiance at pressure level p for channel j, and σ_j is the variance for channel j. The meaning of 'error-weighting' in this context is that each SEVIRI channel is given a different weight depending on the assumed accuracy of the observation, represented by dividing by σ_j^2 for each channel in Eq. 1. The variances are fixed, and equal to 1.23 K, 1.25 K, and 0.57 K for each of the channels used, which are centred at the 10.8 μm, 12.0 μm, and 13.4 μm channels respectively.

30 These R-matrix (observation error) values are a combination of two sources: (a) the measurement "noise" (i.e. instrumental error, from EUMETSAT-published radiometric error data), and (b) the error in the simulated brightness temperatures, derived

from off-line O-B monitoring statistics, which represent any errors in the background NWP profiles and the radiative transfer. Using these values of N, a profile of minimum residual cost (J_{mr}) is constructed for each p value:

$$J_{mr}(p) = \sum_j (R_j^m - [R_j^c(1 - N(p)) + N(p)R_j^o(p)])^2/\sigma_j^2 \tag{2}$$

A stability cost is then added to the minimum residual cost based on the temperature lapse rate calculated using the MetUM, and the value of p corresponding to the minimum of the resulting cost profile is selected as the cloud-top pressure. Generally the minimum residual scheme is intended to find suitable solutions for mid- to high-level clouds.

If a solution is not found by the minimum residual scheme (for example, if the resulting solution variance is greater than a pre-determined threshold, or if the resulting cloud-top is put at a relatively low height and hence considered to be unreliable), a method known as the "stable layers scheme" is attempted, the second of three component schemes within the algorithm. This scheme employs only the 10.8 μm channel, and attempts to match the observed radiance from SEVIRI ($R_{10.8}^m$) with each vertical level in RTTOV ($R_j^o(p)$) . Similar to the minimum residual scheme, an additional cost is applied to reduce the likelihood of solutions placing the cloud-top at the bottom of an unstable layer, since this could potentially cause unrealistic convection if the output were to be subsequently assimilated into an NWP model. This background instability cost is given a much higher weighting for the stable layers retrieval compared with the previous minimum residual retrieval.

If the stable layers scheme also fails (for example, if a suitable cloud-top which satisfies the stability constraint and is also consistent with the measured radiance cannot be found, or if the resulting cloud-top is put at a relatively high altitude and hence considered to be unreliable), then a final method is applied which makes a direct comparison between the observed ($R_{10.8}^m$) and the simulated value, assuming an opaque cloud layer, and with no account taken of atmospheric stability or cloud transparency. This is known as the "profile matching scheme", and is always applied if both the stable layer and minimum residual schemes fail to find a solution (which is typically about 4 % of all SEVIRI pixels in the full disk). A value for the effective cloud amount (N) has been assigned to each pixel within the profile matching scheme (this value is derived from the minimum residual scheme). Since cloud is assumed to be opaque in this scheme, this N-value does not have the same quantitative meaning as the equivalent values found within the minimum residual scheme. However, it is nonetheless retained here for its qualitative value, because useful information about cloud properties within a pixel can still be inferred from the N-value originally assigned to it.

Finally, we note that in general the cloud-top height product will sample multi-layered cloud for a significant portion (~20–40 %) of the SEVIRI full-disk. This can sometimes have the effect of cloud-top heights in multi-layer scenes appearing at an intermediate level for pixels retrieved with the minimum residual scheme, if the higher layer is optically thin. However, for the present dataset we are confident that the assumption of a single cloud layer is reasonable because: (1) the boundary layer is very dry as discussed in section 3.3 so that the formation of a lower cloud is very unlikely (Cuesta et al., 2009); (2) the inspection of the lidar curtains (not shown here) does not show any evidence of lower cloud layers, and (3) cases with clouds higher than the aircraft are filtered out, as discussed in section 2.3 below.

2.2 Airborne cloud dataset (cloud under aircraft)

During the Fennec campaign, 24 out of the 30 FAAM BAe146 aircraft flights were carried out into the remote Sahara of Mauritania and the western part of Mali (Fig. 1 and Table 1), while the other six flew mainly over the ocean. Much of the flying time was spent at cruising altitude (~8,000 m above sea level) when the aircraft was transiting from the operating airport to the target area. During this time, vertical profiles of the atmosphere below the aircraft were sampled with the on-board lidar, and measurements of upwelling radiance were obtained with the Heimann radiometer. These two datasets are principally investigated here, as they provide remotely-sensed cloud information. The meteorological conditions during Fennec, details of each individual flight's location, and principal objectives are described in Ryder et al. (2015).

A dataset of aircraft cloud mask and CTH retrievals for each FAAM flight using the Leosphere ALS450 backscatter lidar and Heimann KT19 radiometer has been derived. The lidar has an operational wavelength of 355 nm with daytime capability, and a nadir viewing geometry (Marenco et al., 2011; Chazette et al., 2012). The full overlap between the emitted light and the receiver field of view (overlap range) is estimated at 300 m. Lidar signals were acquired with a vertical resolution of 1.5 m and an integration of 2 s; to improve the signal-to-noise ratio, the signals have been further smoothed to a vertical resolution of 45 m. The lidar field of view is 4 mrad, meaning that the footprint at sea surface from 8,000 m altitude is ~32 m and at cloud-level (assuming a cloud ~2,000 m below aircraft) is ~8 m. The lidar qualitative plots show in general a significant dust load in the first ~6,000 m of the atmosphere from the surface, consistent with Cuesta et al. (2009), although no attempt has been made to invert the data to extinction coefficient, as this requires a time consuming, non-automated procedure.

Configured in the infra-red spectral range of 8 to 14 μm, the Heimann radiometer measures broadband upwelling radiation at a temporal frequency of 1 Hz, providing a high temporal resolution dataset of brightness temperature across a 22° field of view around nadir. This implies a footprint at sea level of ~3 km when flying at ~8,000 m altitude, with a footprint of of ~800 m at cloud level (again assuming a cloud level ~2,000 m below the aircraft). This instrument is calibrated in the laboratory using known targets. The dataset of the Heimann radiometer has been integrated to yield a 2 s resolution, time-matched with the lidar dataset, thus allowing the two datasets to be considered together. At typical aircraft speeds, this integration time is equivalent to a horizontal resolution of approximately 300 m (along-track).

A cloud detection and cloud-top retrieval algorithm has been applied to each lidar vertical profile, making use of the uncalibrated range-corrected signal profile P(R), using the thresholds given in Allen et al. (2014). The basis for this algorithm lies in the detection of intense peaks with large gradients in the backscattered range-corrected signal. A cloud is identified if (a) P(R) > 4000; (b) P(R) > 3 × P(R - 200 m); and (c) no other cloud-top is found in the same profile between R - 500 m and R, where R is the range below the aircraft. Once a cloud is detected, the cloud-top range R_c is determined as the first lidar measurement point, starting at R and moving inwards to (R - 200 m), where $P(R_c)$ < 1.5 P(R - 200 m). The cloud-top range R_c is then converted to CTH, which we estimate to be accurate to within ±150 m. As an additional quality control, cloud-top heights under 4500 m have been discarded, because the lidar signal tended to be noisy at low altitude due to attenuation by dust, and the noisy signals have shown to lead to a number of false detections. The assumption that no clouds occur under 4500 m is reasonable over the Sahara, where the summer atmosphere is hot and dry in daytime and clouds are found at the top

of a very deep boundary layer (Stein et al., 2011; Marsham et al., 2013). In all locations where a CTH has been determined, a lidar cloud flag has been set, indicating detection. Note that this algorithm detects the cloud-top spike, and would lead to a missed detection when the aircraft is flying in the cloud or when the cloud-top is within the overlap range; additional rare missed detections were encountered where the cloud cover was partial within the integration time used. We estimate that these

5   missed detections represent ~15 % of all cloudy pixels for days exceeding 10 % cloud coverage.

To address the missed detections, a second cloud detection algorithm has been applied using the Heimann radiometer brightness temperature (T_B) dataset. First of all, a clear-sky baseline T_B was derived, using the lidar cloud flag to select cloud-free data, and smoothing to remove the effect of outliers. The baseline T_B is determined as follows: (1) pixels flagged as cloudy by the lidar and pixels measuring a T_B < 290 K are discarded; (2) data points have been grouped in 10 minute intervals; (3) points

10   outside 1 standard deviation of the local group mean have been considered outliers and also discarded; (4) the remaining data in each group have been averaged; and (5) the cloud-free T_B calculated in this way for each group has been interpolated back to the dataset resolution by fitting a cubic spline with a 4-point neighbourhood. Cloud detection has then been based on the difference between the baseline T_B and the measured T_B. A difference greater than 3 K has been considered an indication of cloud presence, and has been used to set a Heimann cloud flag. Sensitivity testing on this threshold shows that a change of ± 1

15   K affects the cloud flag for < 5 % of the pixels in the Fennec dataset. Note that this cloud flag, being sensitive to T_B, permits to resolve most of the missed detections by the lidar, including those when flying in cloud or very near its top. As an additional quality control, clouds detected when flying lower than 4500 m have been omitted for the same reason that they were omitted from the lidar dataset. Data acquired during aircraft turns involving a bank angle larger than 10° have also been omitted, as the Heimann T_B would be tilted from the vertical and affected by the different viewing geometry.

20   The two cloud flags are then combined into an aircraft cloud mask, with a cloud being flagged as detected if either the lidar or the Heimann indicates that a cloud is present. The aircraft cloud mask has been reviewed by looking at the lidar qualitative plots and the Heimann T_B plots, and the thresholds have been adjusted until we have been satisfied that we have a robust product. Note however that the Heimann method may well be dependent on the season and the geographical region (the Sahara being hot with an extremely deep boundary layer in the summer enabling a significant thermal contrast between the clouds and

25   the underlying surface); therefore we do not expect it to work as well outside the specific Fennec dataset.

At the end of this procedure, we end up with the following data on a ~300 m horizontal resolution (along-track), based on the airborne data:

(a) aircraft cloud mask (also articulated into lidar cloud flag and Heimann cloud flag): indicates the presence of a cloud under the aircraft;

30   (b) lidar CTH, only for those points that indicate cloud detection from the lidar;

2.3 Airborne cloud dataset (cloud above aircraft)

Downwelling short wave irradiance measured by a modified Eppley pyranometer (a Broad-Band Radiometer, BBR, calibrated for the spectral range 0.3 μm - 2.8 μm) is used to detect clouds above the aircraft. Using a ground-based approach, it has been shown that pyranometer data is independently capable of differentiating clear-sky conditions from cloud (Long and Ackerman,

2000; Xia et al., 2007). The BBR produces a noisy field with rapid, large-amplitude variations in signal when the incident solar radiation passes through cloud. By differentiating the radiometer's time series, and thresholding at 7 W m^{-2} s^{-1}, a cloud filter was generated. This threshold was fine-tuned through a trial-error-process. Neighbouring points collected within approximately one minute (~10 km) of those flagged by the BBR method were also flagged, to fill inherent gaps in the filter. A sensitivity test
5   reveals that a change of this threshold by ± 1 W m^{-2} s^{-1} alters the assignment of cloudy datapoints by ~3 %. Since data-points for which cloud lay between the aircraft and the satellite field of view are not useful for a quantitative comparison with MSG, this filter was used to remove these points from both the SEVIRI and aircraft datasets entirely (Table 1).

As with the Heimann cloud mask, the BBR-based cloud detection algorithm is sensitive to roll and pitch angles. Therefore, aircraft points which deviated from a nadir-zenith alignment by more than 10° have been removed from the comparison.

10   ## 2.4   Additional data

Throughout the 30 BAe146 flights of Fennec, a total of 121 Vaisala RD94 dropsondes were deployed from the aircraft. These are used here as supplementary information on the atmospheric thermodynamic profile for specific times during the flights. Also used are images taken from the standard rearward facing digital camera of the BAe146. Topography was provided using data from the GLOBE digital elevation model (Payne et al., 1999). A map of surface albedo, created by a compilation of
15   MODIS satellite data (Gao et al., 2005), was also employed.

2.5   Spatial and temporal matching of datasets

To allow for a direct comparison of the satellite and airborne datasets, a mean value was derived for both the aircraft cloud mask and aircraft CTH within each SEVIRI pixel (the number of aircraft points falling within each satellite pixel varied between 2 and 17, with an average of ~8). This mean value of aircraft cloud mask within each SEVIRI pixel is expressed as an aircraft-
20   derived estimate of cloud fraction, with the value of this "aircraft cloud fraction" represented by a number between 0 and 1.

Since the acquisition time for each SEVIRI frame is 15 min, we evaluate the inherent measurement uncertainty arising from a changing cloud field within this time-step. Such changes may be the result of cloud formation, dissipation, or movement. The actual time difference is assumed to be half of the 15-min uncertainty, because each aircraft point has been matched with the
25   satellite frame nearest to its time-stamp (within either 7½ min before or 7½ min after the middle of SEVIRI's full-disk scan). We have translated the uncertainty from temporal to spatial coordinates by calculating the percentage of pixels in the cloud mask which change from cloudy to non-cloudy or vice versa during a timestep. A single value of uncertainty was then derived for each flight by averaging these percentages across each pair of successive frames. The domains used in this method were chosen by defining a rectangle bounded by the maximum and minimum latitude/longitude points along the BAe146 flightpath
30   for each individual flight, and using acquisition times in accordance with the flight's duration. This was done to aid consistency in the comparison, however, we emphasise that the uncertainty itself was derived from SEVIRI data only.

Changes between subsequent cloud mask frames tend mainly to occur near to cloud boundaries, as one would expect, and so it is noted that scenes with more broken/cellular cloud structures will inherently carry a greater temporal matching uncertainty.

These uncertainty values are included in the analysis for each flight (Tables 2 and 3). In addition, a combined overall uncertainty across all the flights has been derived. This overall value for the entire dataset excludes flights for which the cloud coverage was less than 10 %, since the uncertainty value can be biased by cloud-free (or near cloud-free) conditions.

3   Results

3.1   Cloud scale analysis

The aircraft along-track resolution of approximately 300 m permits us to estimate the horizontal extent (calculated by sampling the number of adjacent pixels along the aircraft track) of clouds in the direction of travel; results of this analysis are shown in Fig. 2. A large fraction of the clouds encountered by the aircraft (67 %) are actually smaller in scale than the 3 km-wide SEVIRI pixels, and are therefore filling these pixels partially. In fact, Fig. 2 shows that the smallest cloud size identifiable by the BAe146 (300 m) is also the size most frequently observed (25 %), implying that one quarter of the entire dataset is made up of clouds with a horizontal extent less than (or equal to) the aircraft resolution. Therefore, radiation reaching SEVIRI will often have a contribution from both the altitude of the cloud-top and from the desert surface below, all within the same pixel. This can have varying implications, which will be investigated throughout the following sections.

3.2   Cloud mask comparison

A direct time-series comparison for flights B602, B608, B614 and B706 is presented in Fig. 3. These flights have been chosen as examples of scenes that have sufficient cloud amounts to make a qualitative comparison. In general, the visual comparison of cloudy SEVIRI pixels against the aircraft cloud mask is encouraging. For flight B602 (Fig. 3a), 16 % of the SEVIRI pixels are cloudy in both configurations of the cloud mask, compared to 14 % cloudiness in the aircraft points. Similarly, both B608 and B614 (Fig. 3b and 3c), show a difference of less than 3 % between the aircraft and *AllCloud* masks. The *CloudRetrieval* mask exhibits a lower percentage of cloud coverage compared to the aircraft for these two flights, but the difference here is still no more than 6 %.

In the case of flight B706 (Fig. 3d), a cloudier scene appears from the aircraft dataset than with SEVIRI, with nearly twice as much cloud detected (26 %) compared with the *CloudRetrieval* mask (14 %). The digital camera footage (Fig. 4a) shows that the cloud during this flight was generally stratiform in its structure, but also exhibited numerous gaps which allowed a contribution of upwelling radiation from the desert floor. Overall however, such an underestimation of the cloud field in SEVIRI was only observed in a small portion of flights. The full comparison for each flight using this method can be found in Table 1.

Figure 5 shows how a pixel by pixel comparison of the SEVIRI cloud mask with the aircraft cloud mask is distributed as a function of aircraft cloud fraction for flights B608 and B706. These particular flights have been selected for individual analysis because of the consistent altitude of cloud-tops apparent in the aircraft data, the absence of cloud detected above the aircraft, and for B608, the larger dataset (due to the BAe146 remaining at cruising altitude throughout the whole flight). Since

AllCloud allows for partially cloud contaminated pixels, the ideal *AllCloud* histogram of Fig. 5 would show 100 % of the cloud-free pixels corresponding to an aircraft cloud fraction less than 0.1, and all of the cloudy pixels spread about the aircraft cloud fraction range 0.1–1.0. Conversely, an ideal version of Fig. 5 for *CloudRetrieval* would show 100 % of the cloudy pixels matching with an aircraft cloud fraction greater than 0.9.

5   The cloud structure of B608 is illustrated by the digital camera image at 1616 UTC (Fig. 4b). As Sect. 3.1 highlights, the broken, small scale nature of these clouds means that care must be taken when interpreting a comparison with SEVIRI cloud flags, each one of which represents an area of ~3 x 3 km^2. In the case of the *AllCloud* mask for B608 (Fig. 5a), 87 % of clear-sky SEVIRI pixels have been matched with an aircraft cloud fraction of less than 0.1. Since *AllCloud* has been designed with clear-sky pixels in mind, this result implies a good agreement between the aircraft and MSG datasets. In terms of cloud,
10   63 % of cloudy SEVIRI pixels have been matched with an aircraft cloud fraction exceeding 0.9, to which can be added the 19 % of cloudy flags matched within the aircraft cloud mask range 0.1–0.9. The remaining 18 % of cloudy pixels along the B608 flight path have been matched with an aircraft cloud fraction of less than 0.1.

It is not expected that SEVIRI should flag cloud with an aircraft cloud fraction below 0.1. Part of this discrepancy can be attributed to the temporal uncertainty induced by a changing cloud field during SEVIRI's scan time (calculated as 16 % for
15   B608, see Sect. 2.5). It is also likely that the small horizontal extent of the clouds in this scene may have an effect on SEVIRI's cloud detection scheme. In addition, the one dimensional nature of the aircraft flight path means that although the datasets should agree for clear sky pixels, the aircraft may miss cloud observed by SEVIRI which lies parallel to, but not along, its track. As this effect is random in its nature, it is assumed that overall, on a wide dataset such as the one discussed here, errors can be reasonably assumed to cancel out. The results of this method of matching SEVIRI *AllCloud* cloud mask flags with
20   aircraft cloud fractions for each individual BAe146 flight is shown in Table 2.

The *CloudRetrieval* cloud mask is more conservative in its approach to flagging cloud, and partly as a result of this, Fig. 5b shows a higher percentage of cloudy SEVIRI pixels which are matched to aircraft cloud fractions exceeding 0.9. However, unlike *AllCloud*, aircraft cloud fractions in the range 0.1 – 0.9 are not intended to be matched with a cloudy flag in the SEVIRI data. This means that for B608, the number of SEVIRI pixels flagged as cloudy by *CloudRetrieval* match with 75 % of the
25   aircraft dataset, in contrast to 82 % from *AllCloud*. The number of clear-sky flagged pixels for *CloudRetrieval* that matched with an aircraft cloud fraction below 0.9 was 96 %.

B706 observed a more uniform cloud layer than B608 during the flight (Fig. 4a), which mostly lay in the region of the Mali-Mauritania border. 85 % of SEVIRI cloud flags detected by *AllCloud* were matched with aircraft cloud fractions above 0.9, with this figure rising to 91 % for *CloudRetrieval*. As was implied in the qualitative analysis shown in Fig. 3d, SEVIRI
30   tends to detect less cloud than the aircraft, with 9 % of the *AllCloud* and 12 % of the *CloudRetrieval* cloud-free flags showing aircraft cloud fractions above 0.9. Nonetheless, these values are within the temporal uncertainty (15 % for *AllCloud* and 9 % for *CloudRetrieval*), and we therefore consider the datasets for B706 to be in excellent agreement.

Across all 24 flights and using every valid data-point, we find that 91 % ± 8 % of pixels where SEVIRI did not detect a cloud in the *AllCloud* mask matched with an aircraft cloud fraction of less than 0.1, showing excellent agreement between
35   the datasets for clear skies (Fig. 6a). For the pixels flagged as cloudy, 59 % of pixels with a SEVIRI *AllCloud* cloud detected

corresponded with an aircraft cloud fraction greater than 0.9. Combining this with all other aircraft cloud fractions which were greater than 0.1, we find an overall value of 80 % ± 8 % for the *AllCloud* mask. As for the *CloudRetrieval* configuration of the cloud mask (Table 3), 68 % ± 5 % of cloudy SEVIRI pixels matched with aircraft cloud fractions of greater than 0.9. Since the *CloudRetrieval* mask is not expected to account for cloud fractions of 0.1–0.9, this seems to flag a 32 % overestimation of cloudy pixels. As for the identification of clear-skies by *CloudRetrieval,* 97 % ± 5 % of all pixels flagged to be non-cloudy are matched with an aircraft cloud fraction below 0.9, again implying very good agreement (since non-cloudy pixels in *CloudRetrieval* include partially cloudy pixels).

Overall, these results suggest that although the datasets match well in general, the SEVIRI detection scheme seems to overestimate cloudiness, as evidenced by the 20 % of cloudy pixels from *AllCloud* with aircraft cloud fractions of less than 0.1, and the 32 % of cloudy pixels from *CloudRetrieval* with aircraft cloud fractions less than 0.9.

3.3 Comparison of cloud-top height

Fig. 7 illustrates how the aircraft and MSG cloud-top heights correlate. The broad range of cloud-top height values retrieved by SEVIRI, spanning from below 500 m to above 14,000 m, is in sharp contrast with the narrow range in lidar retrieved CTHs, which is limited to 6,000–8,000 m, with the lidar heights consistent with the typical summertime SABL depth. Note that cloud lying above the aircraft flight path has been filtered out using the BBR, and the dryness of the lower SABL makes the lower clouds very unlikely. The extremes are therefore assumed to be incorrect retrievals from the same clouds observed by the lidar at the top of the SABL. The result shows therefore that over the Sahara, CTH retrievals from MSG may differ substantially from reality.

To understand the CTH errors, we split the MSG retrievals into the component schemes from which they are calculated, plotted as a function of the aircraft cloud fraction (Fig. 8a) and effective cloud amount, N (Fig. 8b). Fig. 8a shows us that there is no dependence of retrieved CTH on actual cloud fraction, whereas Fig. 8b shows that there is a strong dependence on both the retrieval method and the retrieved effective cloud amount, and this highlights a retrieval problem under certain circumstances. By presenting the data as in Fig. 8b, it becomes apparent that a relationship exists between the highest/lowest retrieved SEVIRI CTH's and low values of N. This implies that the errors in CTH are intimately connected to the calculation of N. The scheme that shows the best agreement is the "stable layers", in which the CTH is retrieved independently from N and the latter is only computed successively. The "minimum residual" scheme, instead, derives N and the CTH simultaneously, with the results shown in Fig. 8b. Finally, in the profile matching scheme N is not used, and the value reported is the one derived from the "minimum residual"; the fact that there still remains a dependence tells us therefore that there are issues common to both schemes taking place at the same time.

From Fig. 8b, a strong contrast is apparent between the schemes used above and below 6,000 m. The minimum residual scheme dominates for clouds above 6,000 m, as would be expected, while the profile matching scheme dominates below 6,000 m. The stable layers scheme is used only a small fraction of the time, however, we note that all of the points which do use this scheme lie at an altitude broadly consistent with the aircraft data (~6,000 m).

It is also apparent from Fig. 8b that in this dataset the profile matching scheme returns CTHs between 1,000 m and 6,000 m. Given the extremely low relative humidity within the SABL, particularly near the surface, low CTH's are however not expected in this region. In addition, clouds are detected in the same location but at a higher altitude with the lidar. To illustrate these considerations, Figure 8c shows dropsonde and MetUM profiles from flight B608 at 1611 UTC, a time when some of these low CTHs from SEVIRI appear; the profile exhibits the characteristic deep and well-mixed dry adiabatic profile within the boundary layer depth. The good agreement of this model profile with the dropsonde gives us confidence that the meteorological input to the RTTOV radiative transfer model is not the primary factor driving the profile matching scheme towards unrealistic CTH results. Both the dropsonde and model profiles suggest an approximate CTH range of ~5500–6500 m (450-500 hPa), in line with lidar observations (nearest lidar derived CTH to the dropsonde location was 5927 m). In contrast, the closest SEVIRI retrieval to this location returns a CTH value of 1820 m (with retrieved effective cloud amount N=0.5). It should be noted that more often than not, the NWP profiles do not match so well with reality. However as this particular case shows, CTH errors of this magnitude can still arise even when the NWP profiles are a close match to the true atmospheric state.

In general, the profile matching scheme accounts for ~4 % of all the pixels in the SEVIRI full-disk. However, over the Sahara we find that this number rises for many of the flight days, in one case exceeding 20 % of the domain shown in Fig. 1 (during flight B608). For the profile matching scheme, which is shown here to frequently underestimate CTH, it is likely that the assumption of opacity in cloud is not valid for many of the pixels. If contributions from below the cloud base are affecting the retrieval, then the matched radiance flux may be overestimated, and will hence retrieve a warmer (and therefore lower) cloud-top. This can be linked to a more general issue in cloud retrieval over the Sahara; a large portion of the cloud (67 % in this study) has a horizontal extent smaller than the SEVIRI resolution. Therefore a contribution from the desert surface within each cloudy pixel will be commonplace, and the expected effect of introducing contributions of the warm desert surface would indeed be an observed decrease in altitude of the retrieved cloud-top.

Points in Fig. 8a have been coloured by the retrieved effective cloud amount (N) value for each SEVIRI pixel. The majority of clouds appearing above 10,000 m have retrieved N-values below 0.5 (cyan and yellow). This is also evident in Fig. 8b. Likewise, all clouds with SEVIRI retrieved tops below an altitude of 3,000 m have an N-value of less than 0.5. In the case of the low clouds retrieved using the profile matching scheme, the N-value should be used with caution, because this value is not explicitly calculated (due to the assumption of cloud opacity in this scheme), and is instead taken from the minimum residual scheme's algorithm. Nonetheless, Fig. 8b implies that as the value of N increases, the SEVIRI-retrieved CTH also increases, tending towards 6,000 m (the approximate top of the SABL) as the N-value tends to 1. Similarly for the higher clouds, Fig. 8b implies that as N increases for points using the minimum residual scheme, the SEVIRI-retrieved CTH tends to decrease.

As a specific example of this effect, Fig. 8d shows the CTH retrievals for flight B608. Points with an N-value above 0.7 are highlighted in red and green. These points show a good fit with the lidar data (mean difference lidar CTH – SEVIRI CTH for B608 is 606 m), bearing in mind that the number of pixels used with N-values above 0.7 is limited (45 pixels). For pixels with N-values below 0.7, a large spread in cloud-top height (mean difference for B608 is 4402 m) is present in SEVIRI, with the highest retrieved cloud exceeding 13,500 m (corresponding to a nearby lidar CTH of 6650 m).

3.4 Cloud distribution over the Sahara

Despite a mild tendency to over-flag cloud at times, the results of this study suggest that the Met Office cloud mask is relatively robust. A potential application could be its use for a cloud climatology of the Sahara desert from SEVIRI data. Because MSG lies in a geostationary orbit, such a climatology could expand on previous studies (Stein et al., 2011) to include cloud evolution, diurnal variation, and the location of areas most prone to cloud development.

A first look at the most typical locations of cloud over the western Sahara has been compiled from the SEVIRI data used in this study (Fig. 9a). This data is however limited to daytime retrievals simultaneous to FAAM aircraft flights during the month of June across two years, and so is not sourced from a large enough dataset to be considered a climatology. It is shown here mainly to illustrate the potential of the cloud mask and to suggest certain locations across the Sahara over which clouds most frequently appear to form. Four main areas of cloud formation are identified using Fig. 9a: A) the area surrounding the Western Saharan town of Smara; B) the El Eglab Massif in western Algeria (~300 m); C) the Adrar plateau in Mauritania (340 m); and D) parts of the remote low-lying desert region in northern Mali.

Marsham et al. (2008), Huang et al. (2010), Cuesta et al. (2009) and Birch et al. (2012) suggest two possible triggers for cloud formation: 1) surface albedo anomalies and 2) topography. Figure 9b shows in each of the areas A to D, there is an indication that a dark albedo anomaly does indeed exist, attributed in general to local geological features. These will contribute to higher surface temperature, convergence, a deeper SABL and reduced entrainment (Garcia-Carreras et al., 2011), which in turn may be increasing the convective available potential energy, all favouring cloud formation.

The terrain map in Fig. 9c suggests that three of the cloudy areas (A, B and C), are also associated with high ground (relative to the surrounding area). These mountainous regions also have low albedo values, due to their rocky surfaces, making the distinction between the effects of albedo and those of elevation difficult. However, the area with the highest cloud frequency of the four shown in Fig. 9a is the low lying desert area in Mali (labelled D), which suggests that elevated terrain is not always a critical factor in cloud development.

Overall, these findings broadly show consistency with previous studies (Marsham et al., 2008; Cuesta et al., 2009), although the dataset is not yet comprehensive enough to draw robust conclusions about cloud formation mechanisms. A detailed investigation into these mechanisms are beyond the scope of this paper, but further study is advocated by the authors into the relationship between albedo, terrain and cloud development in the SABL, by using the SEVIRI cloud mask datasets discussed here.

4 Discussion of the possible role of dust in SEVIRI errors

Of the cloud mask threshold tests described in Hocking et al. (2011), we speculate that the *gross* test (usually responsible for ~80 % of detected cloudy pixels) is the test which may be retrieving more cloudy pixels than the aircraft suggests to be present, since if the atmospheric dust loading is high, T_B can be diminished by dust absorption (Haywood et al., 2005). The principle of the gross test is quite simple; cloud tops have a much lower brightness temperature than that of the underlying desert surface

(which is simulated by the MetUM and RTTOV). Therefore dust reducing T_B may lead to an instance of incorrectly flagged cloud.

Figure 10 demonstrates this effect using clear-sky pixels as they cross a boundary between adjacent areas of high and low dust loading, and shows a unique dust event which occured during Fennec (Sodemann et al., 2015). In the eastern half of Fig. 10a, the clear-sky T_B value is shown to differ from the MetUM desert skin temperature by ~5 K. At ~9° W, the SEVIRI T_B value sharply decreases, in line with the location of the boundary into the dustier area in Fig. 10b, which is identifiable by the pink/magenta colors (Brindley et al., 2012) in the RGB imagery. By 13° W, the difference between T_B and the background model skin temperature is in excess of 20 K. This value of 20 K occurs in clear-sky conditions due to dust alone, but this difference will increase further for pixels contaminated by cloud. It is therefore feasible that some of the cloud flagging seen in this study for partially cloud-contaminated (and even cloud-free) pixels may be occurring because of the presence of high dust loading; in other words, SEVIRI's observed radiance (R_j^m) may be modified by dust sufficiently such that the term $R_j^m - R_j^c$ in Eq. 1 (which takes no account of dust in the longwave end of the spectrum) could be large enough that the *gross* test's threshold for identifying a cloud is met.

This effect may explain the lower-than-expected overall value of 68 % ± 5 % obtained by *CloudRetrieval* for aircraft cloud fractions greater than 0.9. If the threshold in the Hocking et al. (2011) *gross* test is met due to a combination of partial cloud contamination and dust, this would explain why the scheme sometimes flags a cloud when aircraft cloud fractions are low. Since the *gross* test will flag a cloud in *AllCloud* whenever a cloud is flagged in *CloudRetrieval*, this would also explain why the general shape of the distribution in each histogram (Figs. 5 and 6) always appears to be similar for the range 0.1–0.9 . The aircraft cloud fraction range 0.1–0.9 has been included in Table 3 to show each discrete contribution.

In an intercomparison of 12 cloud mask products under the framework of the fourth-annual cloud retrieval evaluation workshop (CREW-4), Roebeling et al. (2014) identified the Sahara as one of three areas in the SEVIRI full-disk image which showed the largest discrepancy between the cloud masks (for a specific SEVIRI frame taken in June 2008). The cause of this disagreement was identified as differing thresholds for the detection of thin cirrus. However, in light of the tendency found in this study for SEVIRI to overestimate cloudiness, we propose that perhaps at least some of this disagreement might be related to the effect of dust absorption on whether the different cloud mask algorithms will trigger a cloud flag. Dust was also a factor in another key region in the CREW-4 study, the Arabian Peninsula, with a dust storm in the area being the suggested cause of the mismatch between the 12 cloud masks. The incorrect flagging of pixels in SEVIRI images by dust has also been suggested by Banks and Brindley (2013).

Dust, combined with low cloud fractions, may also be a key cause for the poorly matched cloud-top height estimates identified in this study. When a cloud is observed by SEVIRI, two of the three Met Office retrieval schemes seem to dominate over the Sahara; the minimum residual scheme and the profile matching scheme. One hypothesis is that the minimum residual scheme could be incorporating an underestimate of the surface T_B contribution in partially cloudy pixels, due to dust absorption between the surface and the cloud base. Using the dropsonde profile of Fig. 8c, the 15-20 K difference shown in the western half of Fig. 10a would translate to a height error of ~1,500–2,000 m. When cloud contamination is added, this height error would undoubtedly increase.

Hamann et al. (2014) showed that the Met Office CTH retrievals performed comparably to those developed by EUMETSAT, NOAA, Meteo-France, and NASA. Although not the only methods of cloud retrieval, the "radiance fitting" and "radiance ratioing" methods (characterised here by the profile matching and minimum residual schemes respectively) are widely employed by other centres (See Table 4 of Hamann et al. (2014)). Aerosols are not accounted for in the longwave radiative transfer calculations by any of the centres, which implies the possibility that over/underestimates of CTH might be occurring over the Sahara in many of these products. In summary, with no ground truth to verify CTH in intercomparison studies, it is well possible that at least some of the cirrus observed by satellites over the Sahara is, in actual fact, cumuliform cloud atop the SABL.

4.1 Considerations regarding model skin temperatures

Aside from dust, another consideration in locating the source of cloud product errors over the Sahara is the accuracy of the MetUM model skin temperature. Although Fig. 8c shows that large CTH errors can arise despite an NWP profile which is well-matched with observational data, it cannot necessarily be assumed that the model skin temperature is also an accurate reflection of reality. Given the rapidly changing temperatures of the desert surface, especially in the summer, the value of T_B at any given instant can be problematic for the model, leading to another source of potential error. Although Fig. 10 shows that dust can have an effect on the *relative* difference between model skin temperature and the observed SEVIRI T_B at 10.8 μm, errors in the *absolute* value of model skin temperature also have the potential to introduce unknown effects. Any error in the assumed surface emissivity in RTTOV would also have a similar repercussion.

5 Conclusions

Derived products used to identify and characterise cloud over the Sahara desert based on the SEVIRI instrument aboard the MSG geostationary satellite have been compared to airborne cloud remote sensing products from the FAAM BAe146 aircraft during the Fennec Campaign. This provides the first evaluation of satellite cloud data using detailed observations in the extremely data-sparse region of the Sahara, which has an important impact on the global radiation budget. A novel method for identifying cloud above the aircraft using pyranometer data is described, as well as the methodology for cloud identification below the aircraft using a combination of the lidar and Heimann radiometer.

Two configurations of the Met Office cloud mask product, *AllCloud* and *CloudRetrieval*, are used in the comparison. *AllCloud*, a configuration which is intended to include partially cloud-contaminated pixels, is matched to an aircraft cloud fraction greater than 0.1 in 80 % ± 8 % of the SEVIRI pixels found along the combined BAe146 flight paths. In the case of the *CloudRetrieval* mask, which is not intended to flag partially cloudy SEVIRI pixels, we find that 68 % ± 5 % of the cloudy pixels are matched with aircraft cloud fractions greater than 0.9. The cloud mask products achieve higher accuracy for non-cloudy pixels, with 91 % ± 8 % and 97 % ± 5 % correctly identified pixels in *AllCloud* and *CloudRetrieval* respectively. Some partially-cloudy pixels (aircraft cloud fraction 0.1-0.9) are masked as cloudy in *CloudRetrieval* (19 %), suggesting care must be taken when using *CloudRetrieval* for applications which require completely cloud-free pixels. These results are complemented

by a qualitative visualisation of how the satellite and aircraft datasets compare. This is especially important for scenes in which the cloud patterns are broken / cellular in nature, and show that visually speaking, the datasets agree very well.

There is a tendency for SEVIRI to flag cloud more frequently than the aircraft. We speculate that this may be due predominantly to the influence of two key factors: 1) high dust loadings, which reach cloud base (~5–6 km) in the great majority of cases and act to decrease the 'background' brightness temperature and 2) small cloud horizontal extents, as 67 % of clouds were found to be smaller than the SEVIRI resolution (~3 x 3 km^2). However, despite these limitations, we believe that both of the cloud mask configurations discussed here are capable of providing very useful datasets with many potential applications, including NWP model assimilation, operational forecasting, meteorological research, and climate studies.

We present a first estimate of the typical daytime geographical distribution of cloud over the Sahara in June. We identify four areas which seem prone to cloud formation, and show that these areas are consistent with existing theories that surface albedo anomalies and topography may be the key drivers of Saharan cloud formation (Marsham et al., 2008; Cuesta et al., 2009; Huang et al., 2010). The dataset used is too limited to represent a true climatology, but we advocate the use of the cloud masks to create a Saharan cloud climatology in a future publication.

Finally, estimates of cloud-top height from the Met Office's suite of cloud products have been compared with lidar-derived CTH estimates from the BAe146. We find in general that the retrievals differ substantially from the lidar-derived CTH over the Sahara. We have identified a relationship between the effective cloud amount (the product of cloud fraction and cloud emissivity) and errors in the CTH values, and suggest that both cloud horizontal extent and dust loading may again be the key factors which are negatively affecting the results. A SEVIRI-derived CTH, theoretically speaking, might still be useful if the data were to be restricted to using high values of effective cloud amount. However, it is more likely that in order to create a sufficiently reliable dataset, further investigation will be required into the effects of mineral dust, model skin temperatures, and selection criteria for component retrieval schemes in this particular part of the World.

The results of the present research will be used to develop a newer version of the CTH product. The "stable layers" scheme, which mainly makes use of the model thermodynamic profile to identify CTH, is the one giving better results. This seems to be an indication that, for the conditions encountered in this research (i.e. deep and well-defined boundary layer in the daytime summertime Sahara) the model is actually a more reliable source of information for determining CTH than satellite radiances. On the other hand, only a few retrievals make use of this scheme; this seems to indicate that in a future revised version of the product this balance will probably have to be revisited. Since the the methods used by the Met Office in the determination of CTH are also applied in other forms by other satellite applications facilities globally, the relevance of the results presented here are not limited to Met Office products, but may also have implications for other cloud-retrieval algorithms which employ similar techniques.

Acknowledgments

Airborne data were obtained using the BAe-146-301 Atmospheric Research Aircraft (ARA) flown by Directflight Ltd and managed by the Facility for Airborne Atmospheric Measurements (FAAM), which is a joint entity of the Natural Environment Research Council (NERC) and the Met Office. Fennec was funded by NERC (grant NE/G017166/1), and all those involved in

making the campaign a success are warmly thanked. John Kealy's time for this project has been awarded within the framework of the Met Office Chief Scientist internal secondment scheme. John Marsham was funded by the NERC project SWAMMA (NE/L005352/1). Franco Marenco has developed the aircraft cloud mask at the University of Leeds, within the framework of the Met Office Academic Partnership scheme.

References

Agustí-Panareda, A., Beljaars, A., Cardinali, C., Genkova, I., and Thorncroft, C.: Impacts of assimilating AMMA soundings on ECMWF analyses and forecasts, Weather and Forecasting, 25, 1142–1160, 2010.

Allan, R. P., Slingo, A., Milton, S. F., and Brooks, M. E.: Evaluation of the Met Office global forecast model using Geostationary Earth Radiation Budget (GERB) data, Quarterly Journal of the Royal Meteorological Society, 133, 1993–2010, 2007.

Allen, G., Illingworth, S., O'Shea, S., Newman, S., Vance, A., Bauguitte, S.-B., Marenco, F., Kent, J., Bower, K., Gallagher, M., et al.: Atmospheric composition and thermodynamic retrievals from the ARIES airborne TIR-FTS system-Part 2: Validation and results from aircraft campaigns, Atmospheric measurement techniques, 2014.

Banks, J. and Brindley, H.: Evaluation of MSG-SEVIRI mineral dust retrieval products over North Africa and the Middle East, Remote Sensing of Environment, 128, 58–73, 2013.

Banks, J., Brindley, H., Flamant, C., Garay, M., Hsu, N., Kalashnikova, O., Klüser, L., and Sayer, A.: Intercomparison of satellite dust retrieval products over the west African Sahara during the Fennec campaign in June 2011, Remote Sensing of Environment, 136, 99–116, 2013.

Birch, C. E., Parker, D. J., Marsham, J. H., and Devine, G. M.: The effect of orography and surface albedo on stratification in the summertime Saharan boundary layer: Dynamics and implications for dust transport, Journal of Geophysical Research: Atmospheres, 117, 2012.

Borbas, E. E. and Ruston, B. C.: The RTTOV UWiremis IR land surface emissivity module, http://nwpsaf.eu/vs_reports/nwpsaf-mo-vs-042.pdf, EUMETSAT, 2010.

Brindley, H., Knippertz, P., Ryder, C., and Ashpole, I.: A critical evaluation of the ability of the Spinning Enhanced Visible and Infrared Imager (SEVIRI) thermal infrared red-green-blue rendering to identify dust events: Theoretical analysis, Journal of Geophysical Research: Atmospheres, 117, 2012.

Chang, F.-L., Minnis, P., Lin, B., Khaiyer, M. M., Palikonda, R., and Spangenberg, D. A.: A modified method for inferring upper troposphere cloud top height using the GOES 12 imager 10.7 and 13.3 μm data, Journal of Geophysical Research: Atmospheres, 115, 2010.

Chazette, P., Dabas, A., Sanak, J., Lardier, M., and Royer, P.: French airborne lidar measurements for Eyjafjallajökull ash plume survey, Atmospheric Chemistry and Physics, 12, 7059–7072, 2012.

Cuesta, J., Marsham, J. H., Parker, D. J., and Flamant, C.: Dynamical mechanisms controlling the vertical redistribution of dust and the thermodynamic structure of the West Saharan atmospheric boundary layer during summer, Atmospheric Science Letters, 10, 34–42, 2009.

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

Derrien, M., Gléau, H., Daloze, J.-F., and Haeffelin, M.: Validation of SAFNWC/MSG cloud products with one year of SEVIRI data, in: 2005 EUMETSAT Meteorological Satellite Conference, pp. 95–103, 2005.

Derrien, M., Gléau, H., and Fernandez, P.: Algorithm theoretical basis document for cloud products(CMa-PGE01 v3. 2, CT-PGE02 v2. 2 & CTTHPGE03 v2. 2), Tech. rep., NWC SAF Tech. Rep. SAF/NWC/CDOP2/MFL/SCI/ATBD/01, 2013.

Dunion, J. P. and Velden, C. S.: The impact of the Saharan air layer on Atlantic tropical cyclone activity, Bulletin of the American Meteorological Society, 85, 353–365, 2004.

Eyre, J. R. and Menzel, W. P.: Retrieval of cloud parameters from satellite sounder data: A simulation study, Journal of Applied Meteorology, 28, 267–275, 1989.

Gao, F., Schaaf, C. B., Strahler, A. H., Roesch, A., Lucht, W., and Dickinson, R.: MODIS bidirectional reflectance distribution function and albedo Climate Modeling Grid products and the variability of albedo for major global vegetation types, Journal of Geophysical Research: Atmospheres, 110, 2005.

Garcia-Carreras, L., Parker, D. J., and Marsham, J. H.: What is the mechanism for the modification of convective cloud distributions by land

5   surface-induced flows?, Journal of the Atmospheric Sciences, 68, 619–634, 2011.

Garcia-Carreras, L., Marsham, J., Parker, D., Bain, C., Milton, S., Saci, A., Salah-Ferroudj, M., Ouchene, B., and Washington, R.: The impact of convective cold pool outflows on model biases in the Sahara, Geophysical Research Letters, 40, 1647–1652, 2013.

Garcia-Carreras, L., Parker, D., Marsham, J., Rosenberg, P., Brooks, I., Lock, A., Marenco, F., McQuaid, J., and Hobby, M.: The turbulent structure and diurnal growth of the Saharan atmospheric boundary layer, Journal of the Atmospheric Sciences, 72, 693–713, 2015.

10   Giorgi, F.: Climate change hot-spots, Geophysical research letters, 33, 2006.

Hamann, U., Walther, A., Baum, B., Bennartz, R., Bugliaro, L., Derrien, M., Francis, P., Heidinger, A., Joro, S., Kniffka, A., et al.: Remote sensing of cloud top pressure/height from SEVIRI: analysis of ten current retrieval algorithms, Atmospheric Measurement Techniques, 7, 2839–2867, 2014.

Haywood, J. M., Allan, R. P., Culverwell, I., Slingo, T., Milton, S., Edwards, J., and Clerbaux, N.: Can desert dust explain the outgoing

15   longwave radiation anomaly over the Sahara during July 2003?, Journal of Geophysical Research: Atmospheres, 110, 2005.

Hocking, J., Francis, P. N., and Saunders, R.: Cloud detection in Meteosat second generation imagery at the Met Office, Meteorological Applications, 18, 307–323, 2011.

Huang, Q., Marsham, J. H., Parker, D. J., Tian, W., and Grams, C. M.: Simulations of the effects of surface heat flux anomalies on stratification, convective growth, and vertical transport within the Saharan boundary layer, Journal of Geophysical Research: Atmospheres, 115,

20   2010.

Lavaysse, C., Flamant, C., Janicot, S., Parker, D., Lafore, J.-P., Sultan, B., and Pelon, J.: Seasonal evolution of the West African heat low: a climatological perspective, Climate Dynamics, 33, 313–330, 2009.

Long, C. N. and Ackerman, T. P.: Identification of clear skies from broadband pyranometer measurements and calculation of downwelling shortwave cloud effects, Journal of Geophysical Research: Atmospheres, 105, 15 609–15 626, 2000.

25   Marenco, F., Johnson, B., Turnbull, K., Newman, S., Haywood, J., Webster, H., and Ricketts, H.: Airborne lidar observations of the 2010 Eyjafjallajökull volcanic ash plume, Journal of Geophysical Research: Atmospheres, 116, 2011.

Marsham, J., Parker, D., Grams, C., Johnson, B., Grey, W., and Ross, A.: Observations of mesoscale and boundary-layer scale circulations affecting dust transport and uplift over the Sahara, Atmospheric Chemistry and Physics, 8, 6979–6993, 2008.

Marsham, J., Parker, D., Todd, M., Banks, J., Brindley, H., Garcia-Carreras, L., Roberts, A., and Ryder, C.: The contrasting roles of water

30   and dust in controlling daily variations in radiative heating of the summertime Saharan Heat Low, Atmospheric Chemistry and Physics Discussions, 16, 3563–3575, 2016.

Marsham, J. H., Hobby, M., Allen, C., Banks, J., Bart, M., Brooks, B., Cavazos-Guerra, C., Engelstaedter, S., Gascoyne, M., Lima, A., et al.: Meteorology and dust in the central Sahara: Observations from Fennec supersite-1 during the June 2011 Intensive Observation Period, Journal of Geophysical Research: Atmospheres, 118, 4069–4089, 2013.

35   Parker, D. J., Thorncroft, C. D., Burton, R. R., and Diongue-Niang, A.: Analysis of the African easterly jet, using aircraft observations from the JET2000 experiment, Quarterly Journal of the Royal Meteorological Society, 131, 1461–1482, 2005.

Payne, J., Bryant, N., Logan, T., Muller, J., Schreier, G., and MacDonald, J.: The Global Land One-kilometer Base Elevation (GLOBE) Digital Elevation Model, version 1.0. National Oceanic and Atmospheric Administration, National Geophysical Data Center, Boulder, 1999.

Ramanathan, V., Cess, R., Harrison, E., Minnis, P., Barkstrom, B., Ahmad, E., and Hartmann, D.: Cloud-radiative forcing and climate: Results from the Earth Radiation Budget Experiment, Science, 243, 57–63, 1989.

Roebeling, R., Baum, B., Bennartz, R., Hamann, U., Heidinger, A., Fokke Meirink, J., Stengel, M., Walther, A., Watts, P., and Thoss, A.: Evaluating and improving cloud parameter retrievals: Summary report 2014 on the inter-comparison and validation studies of CREW (Cloud Retrieval Evaluation Workshop), www.icare.univ-lille1.fr/crew/, 2014.

Ryder, C., Mcquaid, J., Flamant, C., Rosenberg, P., Washington, R., Brindley, H., Highwood, E., Marsham, J., Parker, D., Todd, M., et al.: Advances in understanding mineral dust and boundary layer processes over the Sahara from Fennec aircraft observations, Atmospheric Chemistry and Physics, 15, 8479–8520, 2015.

Saunders, R., Matricardi, M., and Brunel, P.: An improved fast radiative transfer model for assimilation of satellite radiance observations, Quarterly Journal of the Royal Meteorological Society, 125, 1407–1425, 1999.

Saunders, R., Francis, R., Francis, P., Crawford, J., Smith, A., Brown, I., Taylor, R., Forsythe, M., Doutriaux-Boucher, M., and Millington, S.: The exploitation of Meteosat Second Generation data in the Met Office, in: Proceedings of the 2006 EUMETSAT Meteorological Satellite Conference, Helsinki, Finland, Citeseer, 2006.

Schepanski, K., Tegen, I., Laurent, B., Heinold, B., and Macke, A.: A new Saharan dust source activation frequency map derived from MSG-SEVIRI IR-channels, Geophysical Research Letters, 34, 2007.

Schepanski, K., Tegen, I., and Macke, A.: Saharan dust transport and deposition towards the tropical northern Atlantic, Atmos. Chem. Phys, 9, 1173–1189, 2009.

Schepanski, K., Tegen, I., and Macke, A.: Comparison of satellite based observations of Saharan dust source areas, Remote Sensing of Environment, 123, 90–97, 2012.

Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., and Ratier, A.: An introduction to Meteosat second generation (MSG), Bulletin of the American Meteorological Society, 83, 977–992, 2002.

Schmit, T. J., Prins, E. M., Schreiner, A. J., and Gurka, J. J.: Introducing the GOES-M imager, National Weather Digest, 25, 28–37, 2002.

Sodemann, H., Lai, T., Marenco, F., Ryder, C., Flamant, C., Knippertz, P., Rosenberg, P., Bart, M., and McQuaid, J.: Lagrangian dust model simulations for a case of moist convective dust emission and transport in the western Sahara region during Fennec/LADUNEX, Journal of Geophysical Research: Atmospheres, 120, 6117–6144, 2015.

Stein, T. H., Parker, D., Delanoë, J., Dixon, N., Hogan, R. J., Knippertz, P., Maidment, R., and Marsham, J.: The vertical cloud structure of the West African monsoon: A 4 year climatology using CloudSat and CALIPSO, Journal of Geophysical Research: Atmospheres, 116, 2011.

Trzeciak, T., Garcia Carreras, L., and Marsham, J. H.: Cross Saharan transport of water vapour via recycled cold-pool outflows from moist convection, Geophysical Research Letters, 2016.

Washington, R., Flamant, C., Parker, D., Marsham, J., McQuaid, J., Brindley, H., Todd, M., Highwood, E., Ryder, C., Chaboureau, J., Kocha, C., Bechir, M., and Saci, A.: Fennec - The Saharan Climate System, CLIVAR Exchanges, No. 60, Vol. 17, No. 3, p31-325 [Available online at http://www.clivar.org/sites/default/files/documents/Exchanges60.pdf], 2012.

Xia, X., Chen, H., Li, Z., Wang, P., and Wang, J.: Significant reduction of surface solar irradiance induced by aerosols in a suburban region in northeastern China, Journal of Geophysical Research: Atmospheres, 112, 2007.
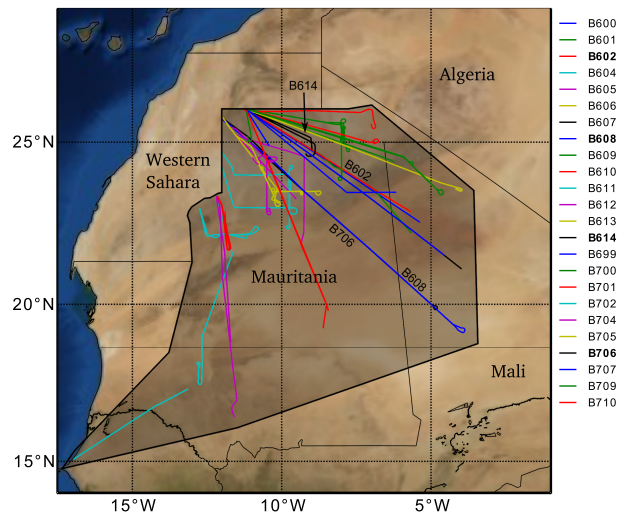
Figure 1. Domain of interest in the Sahara desert, with flight tracks of the BAe146 during Fennec. The shaded polygon surrounds the sampling area. Important flights which are analysed in detail throughout the text are highlighted.
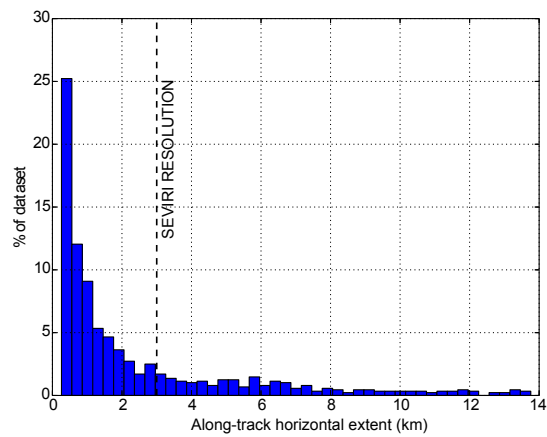
Figure 2. Distribution of cloud sizes from the full aircraft dataset. 67 % of the clouds detected by the aircraft exhibit a horizontal extent smaller in size than SEVIRI's resolution.
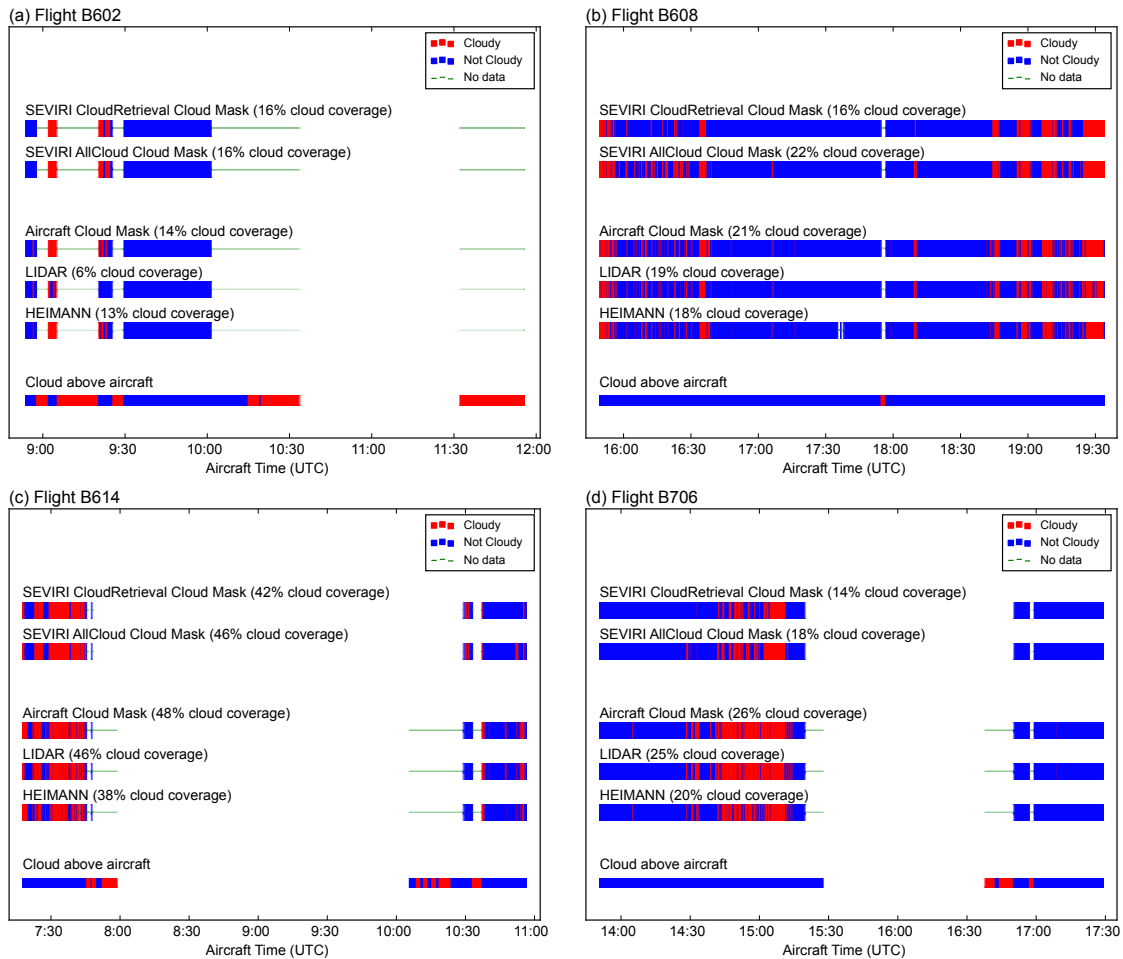
Figure 3. Comparison of *CloudRetrieval* (SEVIRI pixels fully cloud contaminated) and *AllCloud* (partially cloud contaminated SEVIRI pixels also included) cloud masks against the aircraft cloud masks for flights **(a)** B602; **(b)** B608; **(c)** B614; and **(d)** B706. Percentages show the amount of cloudy points (red) as a fraction of the total (red+blue) points for which a cloud flag has been assigned. Gaps in the data (green) can be attributed to either the filtering of cloud detected above the aircraft, aircraft altitudes of < 4500 m, or missing points in the raw data. The BBR-derived filter for cloud above the aircraft is also included.
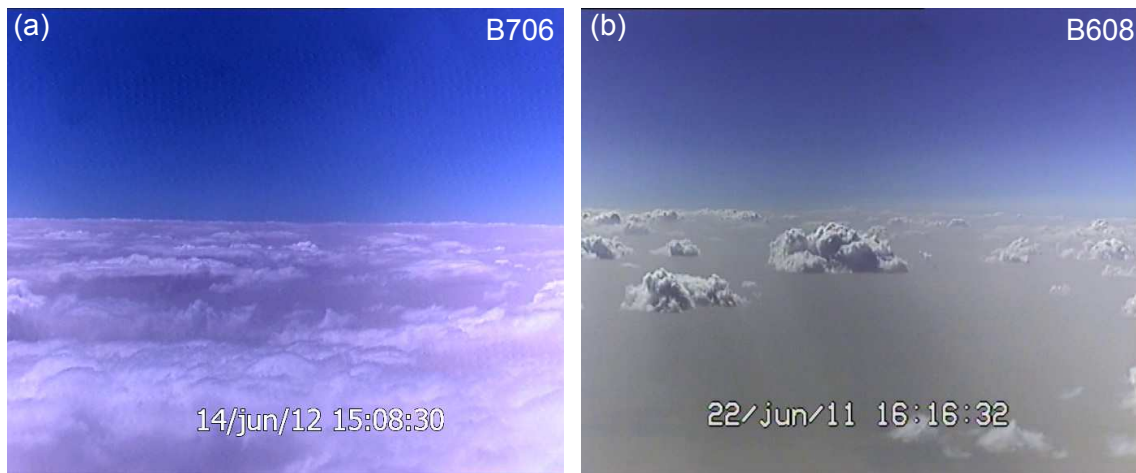
Figure 4. Snapshot from the BAe146 rearward-facing digital camera at **(a)** 1508 UTC during B706 and **(b)** 1616 UTC during B608, showing the cloud horizontal extent and structure of the cloud in each scene. Note the strong contrast in (b) between the dusty SABL and the free-atmosphere above, with the clouds lying at this boundary.
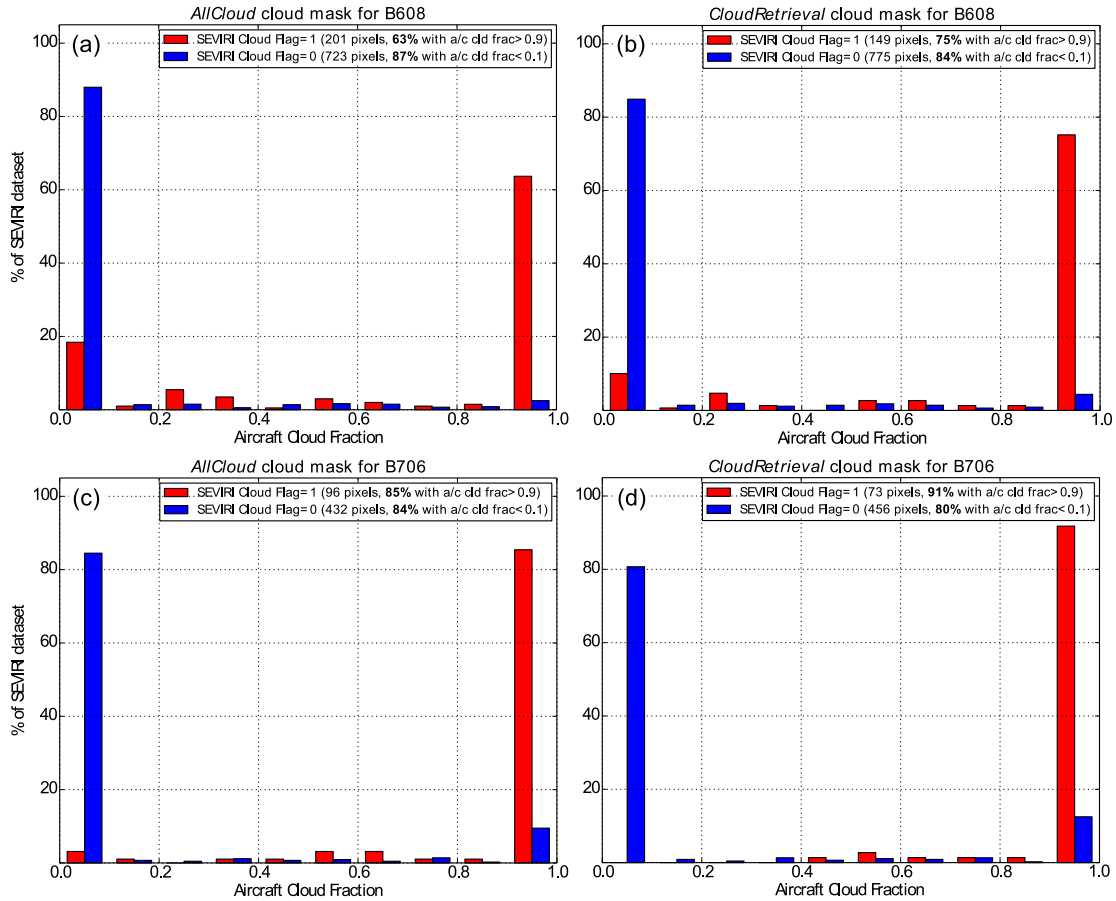
Figure 5. Histogram comparison of aircraft cloud fractions as matched to SEVIRI cloud flags showing pixels flagged as cloudy (red) and non-cloudy (blue) in SEVIRI. Percentage values in the legend pertain to aircraft cloud fractions greater than 0.9 for cloudy SEVIRI pixels, and less than 0.1 for non-cloudy SEVIRI pixels, to indicate the proportion of the dataset deemed to match. **Top:** Flight B608 **(a)** *AllCloud* and **(b)** *CloudRetrieval* cloud masks. **Bottom:** Flight B706 **(c)** *AllCloud* and **(d)** *CloudRetrieval* cloud masks.
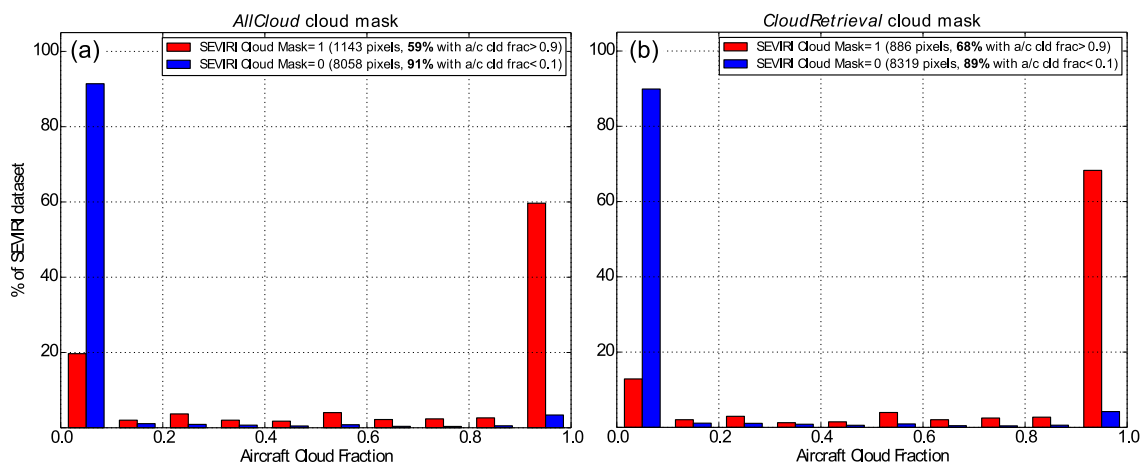
Figure 6. Histogram containing the full Fennec dataset to show the statistical agreement between the SEVIRI and aircraft cloud masks, for (a) *AllCloud* and (b) *CloudRetrieval*. Histograms are as described in Fig. 5.
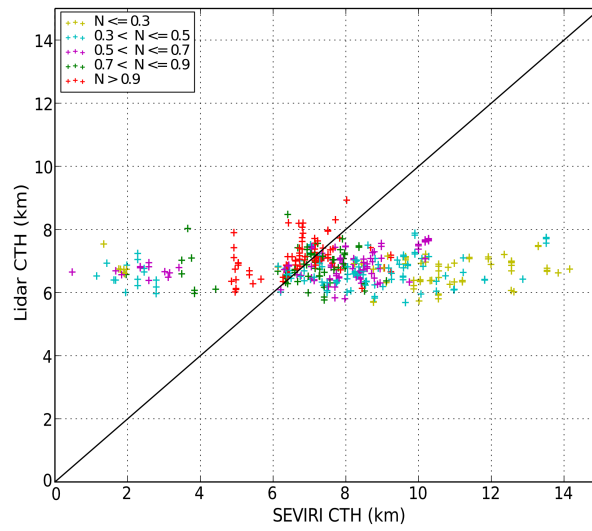
Figure 7. Scatterplot of CTH between SEVIRI and the aircraft datasets for all flights and all retrieval schemes. The altitude range of SEVIRI CTHs is much greater than the range of aircraft CTHs.
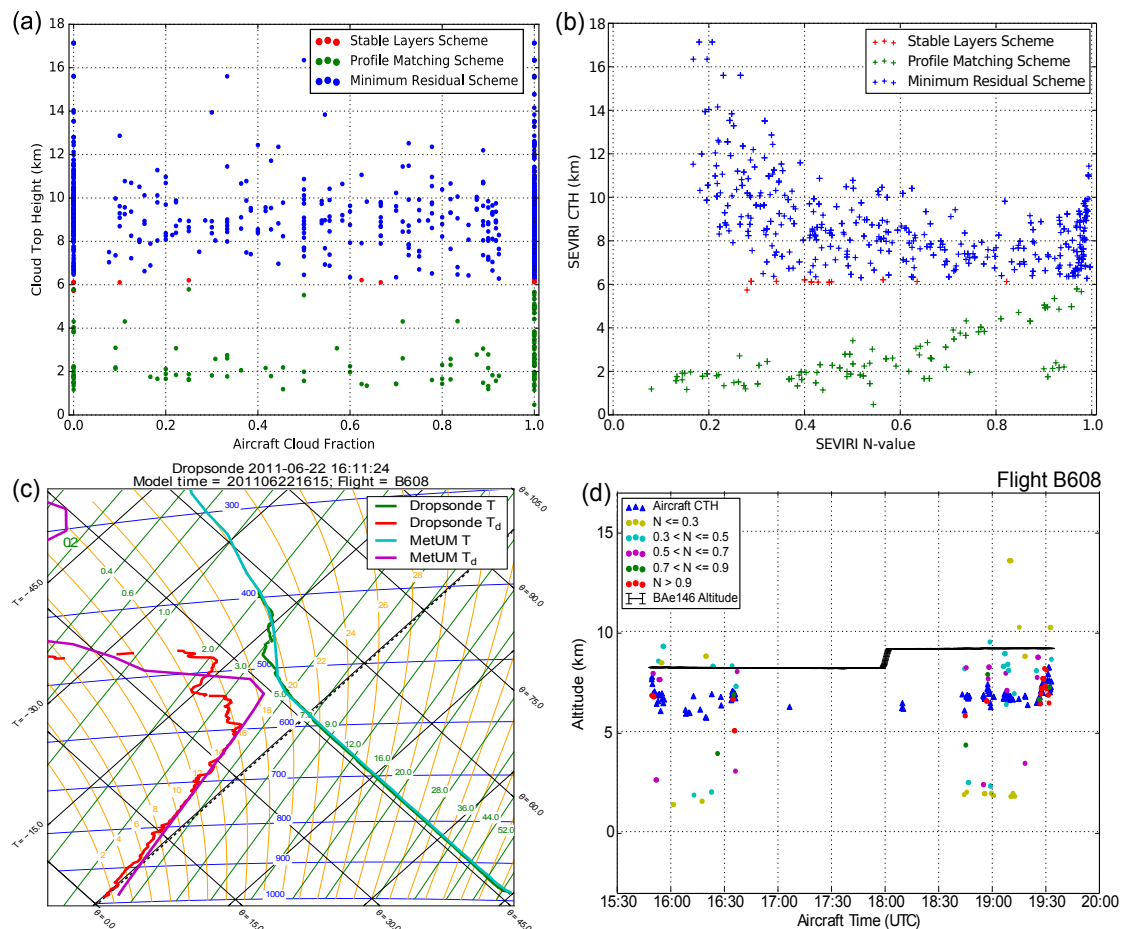
Figure 8. (a) Scatterplot of SEVIRI CTH plotted against aircraft cloud fraction (all flights) **(b)** SEVIRI CTH retrievals for every pixel associated with a BAe146 data-point, plotted against the SEVIRI N-value, and streamed by the type of retrieval scheme used (all flights) **(c)** Thermodynamic diagram of temperature and dewpoint from a dropsonde profile of the SABL from flight B608, overlaid with the nearest MetUM temperature profile **(d)** Time-series of lidar CTH (blue triangles) and SEVIRI CTH retrievals for B608. Pixels where cloud above the aircraft was detected have been omitted in each panel.
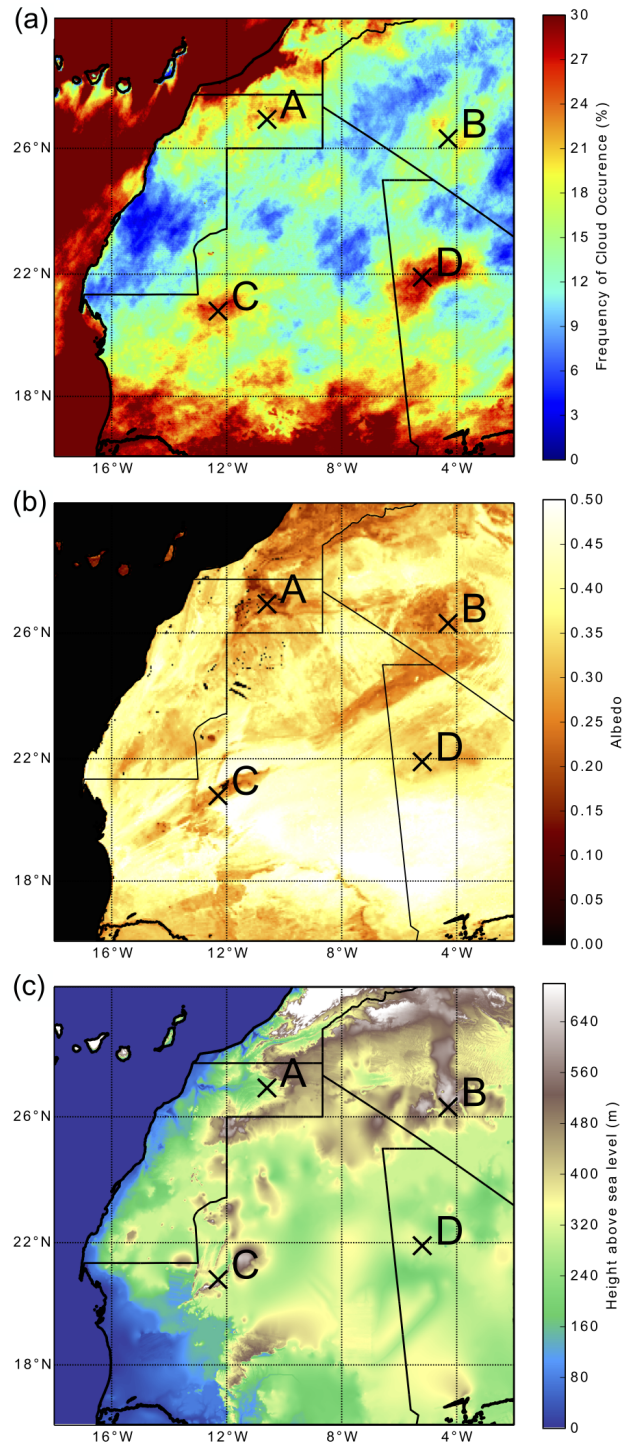
Figure 9. (a) Compilation of all cloudy SEVIRI pixels (*CloudRetrieval* mask) across all the Fennec BAe146 flight dates. Lettering (A–D) indicates the cloudiest land-based areas in the region of interest. Image compiled from 670 SEVIRI *CloudRetrieval* cloud mask frames in the daytime for June 2011 and 2012 (simultaneous to research flight times only) **(b)** Surface albedo map of the region. **(c)** Terrain map of the region.
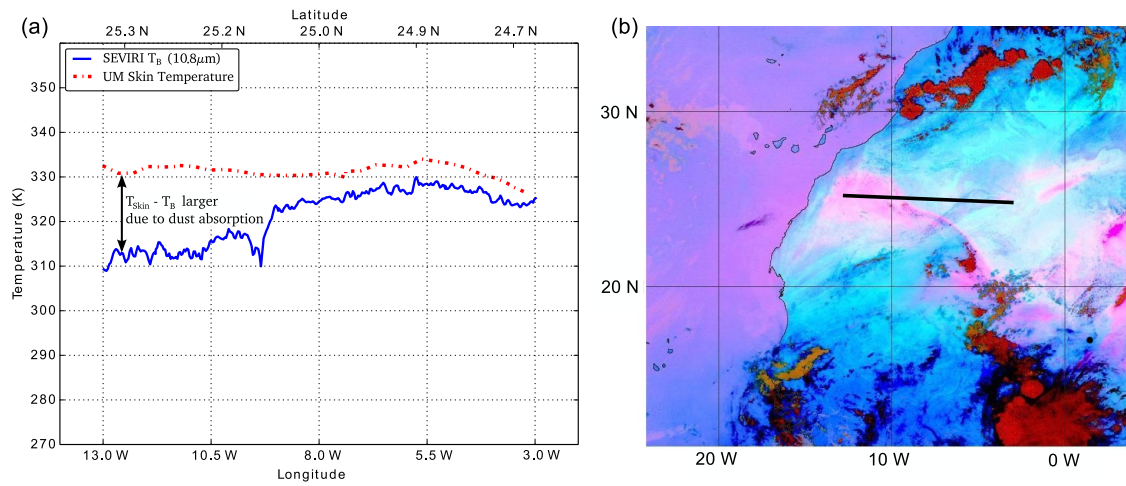
Figure 10. (a) Cross section of the Sahara from 20 June 2011 at 1430 UTC, showing how MetUM model skin temperatures (used as an input to the CTH calculations) compare with SEVIRI brightness temperature in the IR. **(b)** Dust RGB showing the location of the cross section in (a) as a black line. Areas of high dust loading show up in bright pink/magenta. The difference between SEVIRI T_B and model T_{skin} is larger in the dusty area to the west, likely due to absorption by the aerosols.

Date / Time	Flight Number	Percentage of cloud in BAe146 data	Percentage of cloud in SEVIRI data		Cloud detected above aircraft? (y/n)
			AllCloud	*Cloud-Retrieval*	
2011					
17 June 07:48-12:41	B600	2 %	8 %	5 %	y
17 June 14:43–19:37	B601	32 %	43 %	39 %	y
18 June 08:10–12:40	B602	14 %	16 %	16 %	y
20 June 12:47–17:51	B604	8 %	7 %	3 %	y
21 June 08:10–11:58	B605	7 %	4 %	3 %	n
21 June 14:04–19:20	B606	3 %	3 %	2 %	n
22 June 08:04–12:37	B607	7 %	2 %	2 %	n
22 June 15:10–20:16	B608	21 %	22 %	16 %	n
24 June 11:29–16:45	B609	42 %	47 %	40 %	y
25 June 07:31–12:17	B610	0 %	1 %	0 %	n
25 June 14:14–19:16	B611	3 %	7 %	1 %	y
26 June 07:29–12:22	B612	1 %	0 %	0 %	n
26 June 13:55–18:59	B613	3 %	2 %	1 %	n
27 June 06:34–11:39	B614	48 %	46 %	42 %	n
2012					
6 June 12:01–16:50	B699	2 %	5 %	5 %	y
8 June 07:56–12:57	B700	3 %	0 %	0 %	n
9 June 07:55–13:08	B701	0 %	0 %	0 %	n
10 June 08:04–12:41	B702	12 %	3 %	3 %	y
11 June 12:14–17:19	B704	50 %	62 %	60 %	y
12 June 11:27–17:07	B705	1 %	8 %	3 %	y
14 June 13:07–18:13	B706	26 %	18 %	14 %	n
15 June 09:13–14:33	B707	39 %	27 %	14 %	y
17 June 12:14–17:24	B709	6 %	4 %	3 %	n
18 June 07:51–13:11	B710	0 %	0 %	0 %	n

Table 1. Dates/Times for each BAe146 flight, with the percentage of the dataset which is flagged as cloudy shown for the aircraft, the *AllCloud* SEVIRI mask, and the *CloudRetrieval* SEVIRI mask. Also shown is whether or not cloud above the aircraft was observed by the BBR.

Flight Number	Cloudy SEVIRI pixels with ACF > 0.1	Clear-sky SEVIRI pixels with ACF < 0.1	Ratio of Cloudy SEVIRI pixels to total sample size	Uncertainty
	AllCloud	*AllCloud*	*AllCloud*	*AllCloud*
B600	57 %	100 %	14 / 183	± 14 %
B601	84 %	91 %	106 / 246	± 17 %
B602	79 %	95 %	29 / 177	± 6 %
B604	60 %	89 %	25 / 366	± 8 %
B605	80 %	93 %	15 / 358	± 1 %
B606	63 %	96 %	16 / 459	± 8 %
B607	100 %	91 %	11 / 689	± 1 %
B608	82 %	88 %	201 / 924	± 16 %
B609	84 %	77 %	154 / 325	± 20 %
B610	33 %	99 %	3 / 370	± 1 %
B611	60 %	95 %	15 / 211	± 12 %
B612	0 %	98 %	0 / 441	± 0 %
B613	80 %	96 %	5 / 296	± 7 %
B614	93 %	65 %	95 / 207	± 14 %
B699	50 %	98 %	8 / 156	± 15 %
B700	0 %	96 %	0 / 252	± 1 %
B701	0 %	100 %	0 / 572	± 0.2 %
B702	92 %	88 %	12 / 430	± 6 %
B704	81 %	73 %	152 / 247	± 15 %
B705	0 %	98 %	36 / 451	± 12 %
B706	97 %	84 %	96 / 528	± 15 %
B707	80 %	63 %	128 / 474	± 23 %
B709	95 %	93 %	22 / 570	± 4 %
B710	0 %	99 %	0 / 274	± 0 %
All Flights	80 %	91 %	1143 / 9205	±8 %

Table 2. Cloud mask matching statistics for the *AllCloud* SEVIRI cloud mask, sample size for the number of SEVIRI pixels which contain aircraft points, and the temporal measurement uncertainty as described in Sect. 2.5. Aircraft cloud fraction is denoted by ACF.

Flight Number	Cloudy SEVIRI pixels with ACF > 0.9	Cloudy SEVIRI pixels with 0.1 < ACF < 0.9	Clear-sky SEVIRI pixels with ACF < 0.9	Ratio of Cloudy SEVIRI pixels to total	Uncertainty
	CloudRetrieval	*CloudRetrieval*	*CloudRetrieval*	*CloudRetrieval*	*CloudRetrieval*
B600	78 %	11 %	100 %	9 / 183	± 10 %
B601	68 %	24 %	99 %	97 / 246	± 13 %
B602	59 %	21 %	99 %	29 / 177	± 4 %
B604	58 %	17 %	97 %	12 / 366	± 5 %
B605	78 %	22 %	96 %	9 / 358	± 0.8 %
B606	20 %	60 %	99 %	10 / 459	± 3 %
B607	100 %	0 %	96 %	10 / 689	± 1 %
B608	75 %	15 %	96 %	149 / 924	± 10 %
B609	69 %	19 %	92 %	131 / 326	± 12 %
B610	0 %	0 %	100 %	0 / 371	± 0.5 %
B611	0 %	67 %	100 %	3 / 212	± 3 %
B612	0 %	0 %	99 %	0 / 441	± 0 %
B613	0 %	100 %	99 %	1 / 296	± 2 %
B614	75 %	20 %	98 %	87 / 207	± 10 %
B699	13 %	38 %	99 %	8 / 156	± 4 %
B700	0 %	0 %	96 %	0 / 252	± 0.4 %
B701	0 %	0 %	100 %	0 / 572	± 0 %
B702	91 %	0 %	91 %	11 / 430	± 4 %
B704	59 %	21 %	89 %	148 / 246	± 11 %
B705	0 %	0 %	100 %	15 / 451	± 6 %
B706	92 %	8 %	88 %	73 / 529	± 9 %
B707	65 %	21 %	77 %	66 / 475	± 12 %
B709	67 %	33 %	98 %	18 / 570	± 3 %
B710	0 %	0 %	100 %	0 / 274	± 0 %
All Flights	68 %	19 %	97 %	886 / 9205	± 5 %

Table 3. As Table 2 but for the *CloudRetrieval* SEVIRI cloud mask. The ACF range 0.1–0.9 is also shown in this table, as discussed in Sect. 4.