ACTRIS WP 5. NA-5 Clouds and aerosol quality controlled-observations.
Deliverable 5.5 Target classification algorithms

Introduction

The core instrumentation for clouds and aerosols quality controlled-observations at a CloudNet station comprises:

i) Lidar/ceilometer providing profiles of
   a. attenuated backscatter coefficient from aerosol and cloud
   b. linear depolarisation ratio

ii) Doppler cloud radar operating at 35 or 94GHz, providing profiles of
   a. reflectivity, Z,
   b. mean Doppler velocity, v,
   c. Doppler spectral width, σ_v,
   d. standard deviation of the mean Doppler width, σ_{vbar}
   e. linear depolarisation ratio, LDR.

iii) Microwave radiometer providing
   a. brightness temperatures at two or more wavelengths in the range 20-30 GHz
   b. derived water vapour and liquid water path

iv) Additional observations
   a. surface rain rate from raingauge
   b. radiosonde profiles, or NWP model output (temperature, pressure, humidity, winds)

Measurements in blue are not required, but are utilised within the Cloudnet processing scheme, if available.

The profile of attenuated backscatter coefficient, from a lidar/ceilometer, exhibits returns from aerosol, ice and liquid. The molecular return may also be present at visible lidar wavelengths. Attenuation by liquid is severe, and the lidar signal does not penetrate liquid layers that are deeper than about 300 m, thus only reliably detecting the cloud base. Penetration through ice clouds can reach 3 km or more. The lidar depolarisation ratio potentially provides information on the shape and type of particles responsible for the lidar return.
The cloud radar is sensitive to precipitation, ice, snow, insects and liquid droplets.The cloud radar is much less affected by attenuation and provides profiles of the first three moments of the Doppler spectra; reflectivity, Z; Doppler velocity; and the spread of Doppler velocities. Additionally, the variation of the mean Doppler velocity can also be calculated; this can be used for estimating the turbulent properties of the atmospheric particulate, or highlighting the random motion displayed by an insect. The linear depolarisation ratio, if available, is useful for distinguishing insects from clouds and precipitation within the boundary layer, and for identifying the location of melting ice particles.

Facilitating cloud radar and lidar algorithms:

The Cloudnet Instrument Synergy/Target Categorization product

Original authors: Robin J. Hogan and Ewan J. O'Connor

Original document: August 17, 2004

1 Introduction

This document describes an intermediate data product designed to facilitate the application of multi-sensor algorithms by performing most of the typical preprocessing that such algorithms require and providing the results in a common format for all sites. It was developed for the *Cloudnet* project[1] involving four European remote sensing sites, but is also applicable to data from the similar Atmospheric Radiation Measurement (ARM) sites[2] and has been extended for ACTRIS sites with suitable instrumentation. In addition, the procedures have been updated to include some new parameters that are now commonly available. The key procedures that are carried out are:

Standardization of conventions: The output data are provided in a common format for all sites, with common units and with the same conventions such as height being above mean sea level, Doppler velocity being positive upwards.

Ingestion of model data: Many algorithms require temperature and horizontal wind speed, and unless there are regular 6-hourly radiosonde ascents from close to the site, these are best obtained from a model analysis or forecast.

Regridding: The model data are interpolated on to the same time-height grid as the radar. Likewise, the rain rate and liquid water path are interpolated onto the time axis of the radar. The lidar data is placed on the same time-height grid as the radar using nearest-neighbour for the time grid.

Target categorization: Each pixel is categorized in terms of the presence of liquid droplets, ice, insects, aerosol etc., thereby allowing algorithms specific to one type of target to be applied.

Gaseous attenuation correction: Model temperature, pressure and humidity are used to correct radar reflectivity for gaseous attenuation.

Liquid attenuation correction: The radar can be significantly attenuated by the presence of liquid water, but liquid water path from the microwave radiometer, in combination with the location of the liquid water in the profile from lidar and radar, allows this effect to be corrected.

[3]**Instrument errors:** For radar and lidar, variables representing both random error (due to the finite number of samples averaged and the accuracy of any correction for radar attenuation) and systematic error (an indication of the accuracy of the calibration) are added, allowing subsequent retrieval algorithms to make realistic estimates of their associated error.

Instrument sensitivity: Knowledge of the minimum detectable signal of the radar allows one to take account of clouds that may not be detected in comparisons with a model.

Data quality flags: These inform the user when signals are contaminated by ground clutter, unknown radar attenuation in rain and other effects. In section 2 the essential and optional sources of input data are described, and in section 3 the details of the algorithm are provided. The format of the data is outlined in the appendix.

2 Input datasets

The essential instruments that must be present are vertically pointing cloud radar and backscatter lidar, together with either regular radiosonde profiles, or hourly model forecast data available over the site. Recommended but non-essential instruments are microwave radiometer (for providing liquid water path) and rain gauge.
Processing is done one day at a time. Ideally the instruments operate continuously, but if at any time in the day radar, lidar or model data are missing then there will be a gap in the output product. The absence of data from the non-essential

[1] http://www.met.rdg.ac.uk/radar/cloudnet/

[2] http://www.arm.gov/

[3] ACTRIS Deliverable 5.6: Establish a common set of definitions for instrument errors

instruments may mean that correction for radar attenuation is not possible, but data quality flags will indicate that this is the case.

The decision regarding which instruments and auxiliary data to include in this product was made by considering those that are most used in retrieval algorithms and those that are necessary to correct for radar attenuation and flag bad data. Radar and lidar are complementary due to their very different dependence on particle size; this means that the combination of the two offers the most accurate estimates of cloud occurrence and cloud fraction (Mace et al. 1998, Hogan et al. 2001), and can also be used to retrieve particle size (Donovan et al. 2001, O'Connor et al. 2004). However, it is not the intention that all instruments at a site that could conceivably be used in a retrieval algorithm (or that might be useful for model comparison, such as broadband fluxes) should be combined into this product; instead, routines for placing other instrument data on the same time-height grid are provided for this task.

The algorithm is implemented in Matlab[4], although future implementation in C and/or python is possible. The input datasets are read into the processing algorithm using separate functions that may make modifications depending on the particular instrument and site.

2.1 Cloud radar

The vertically pointing radar would typically operate continuously at 35 or 94 GHz and provides profile with a time resolution of around 10-30 seconds and a height resolution of 30-60 m (better than 100 m). The radar resolution is used as a master grid on to which all other datasets are placed. The temporal resolution of 30 s is long enough that the datasets are conveniently small to use, but short enough that cloud fraction and other parameters to be used to evaluate models are of sufficiently high precision. Essential parameters that the radar must provide are radar reflectivity factor Z and Doppler velocity v[5]. Radar reflectivity factor should have had the following processing applied to it:

- Linear averaging to the resolution required.
- Noise subtraction: the reported Z should be that of the atmospheric targets without the contribution from instrument noise and thermal emission that is present in the raw measurements.
- Clear sky clearing: areas with no detectable atmospheric signal should be flagged (or SNR should be available). Any remaining speckle noise should be removed using masking algorithms.
- Elimination of artefacts, such as the near field effect for instruments with large antennas, overlap effects for some bistatic systems and spurious instrument-specific echos.

Note: Although definitions and processing may be performed with Z in linear units, it is conventional to report the values in logarithmic dBZ units defined thus:

$$Z\,[\mathrm{dBZ}] = 10 \log_{10} Z\,[\mathrm{mm}^6\,\mathrm{m}^3]$$

The quality control algorithm attempts to flag likely ground clutter, but since the clutter behaviour can be very different for different radars it is better if it can be removed prior to being read in. Likewise, it is beneficial but not necessary for insects to have been identified; if available, the LDR parameter is an excellent discriminator. The radar should report values as low as possible above the ground. No attempt should be made to correct for attenuation. The radar should obviously be calibrated as well as possible and an indication of the likely accuracy of the calibration should be available. If necessary the appropriate factor will be applied to Z in order to conform to the following calibration convention, defined for a distribution of Rayleigh-scattering liquid water droplets in the absence of attenuation:

in the absence of attenuation, a cloud at 0 C containing one million 100 μm (i.e. Rayleigh scattering) droplets per cubic metre will have a reflectivity of 0 dBZ at all frequencies.

If several reflectivity channels are available from a particular site, such as the four specialized modes used by the MMCR radars on the ARM sites (Clothiaux et al. 1999), then the one most free from artefacts will be taken, even if this is not the most sensitive. This is because the Instrument Synergy/Target Categorization product is intended to be used in automated algorithms that operate on large volumes of data without the need for significant human quality control. The Doppler velocity should have had clear sky pixels removed in the same way as Z. If necessary, the values will be inverted to ensure that the convention of positive velocities upwards is adhered to. The velocities need not be unfolded, but the

[4] http://www.mathworks.com

[5] There are some sites with current and/or legacy data without Doppler velocity; the algorithm may be modified in future to cope without Doppler velocity.

folding velocity v_{fold} should be known to the algorithm. It is possible that the algorithm for locating the melting layer (see section 3.4.1) will perform poorly if the folding velocity is too small.
Other radar parameters, such as spectral width (σ_v), will be transferred into the output dataset. If available, the 30-s standard deviation of the 1-s mean velocities, $\sigma_{\bar{v}}$ (used to estimate turbulence levels; Bouniol et al. 2003), will be used to assist in the diagnosis of ground clutter.

2.2 Cloud lidar

The cloud lidar is principally used to identify the base of liquid water clouds. Most commonly this instrument will be a near-infrared lidar ceilometer reporting only attenuated backscatter coefficient, β. The instrument should be operated pointing between 2 and 5 degrees from zenith to avoid specular reflection from horizontally aligned pristine crystals, which could be mistaken by the algorithm for the presence of supercooled liquid water (Hogan et al. 2003b). The lidar should be calibrated as well as possible (e.g. using the technique of O'Connor et al. 2004). Units are in m^{-1} sr^{-1}. The algorithm assumes that all signals from the lidar are due to atmospheric particulates such as cloud and aerosol, rather than Rayleigh scattering from air molecules. This is generally valid in the near infrared (e.g. 905 nm used by the Vaisala CT25K and CT75K instruments) but for lidars at visible wavelengths (e.g. 532 nm micropulse lidars used by ARM), an additional processing step is necessary to identify and flag the molecular return. Background and solar noise should also be identified and flagged in the pre-processing step. Note that lidars with Raman or depolarisation channels are far superior in their discrimination between the various atmospheric particulates; the depolarisation ratio and the extinction are extremely useful parameters for liquid water diagnosis and cloud retrievals, and these are provided in the Instrument Synergy/Target Categorization for instruments in quasi-continual operation.

2.3 Model parameters

At least four radiosonde profiles per day would be required from close to the site to provide adequate dynamic and thermodynamic data above the instruments. Because most sites do not have this information, we use hourly profiles from short-range model forecasts. These models assimilate the data from the radiosonde network so the temperature (T), pressure (p), humidity (q) and horizontal wind (u and v) will usually be accurate enough for our purposes. We do not make use of the cloud variables as these are to be evaluated, and will generally be much less accurate.
Before it can be used the model output must be converted into the *Cloudnet* single-site model format, which includes calculation of a number of radar propagation and scattering parameters at 35 GHz and 94 GHz that depend on thermodynamic state. The variables T, p, u and v are then interpolated on to the time grid of the radar and provided as part of the Instrument Synergy/Target Categorization product. To conserve space, they are not interpolated in height; subsequent processing algorithms will need to do this. The following additional parameters are ingested and used by the algorithm: specific gas attenuation κ_g (dB km^{-1}) (predicted from model q, T and p using the line-by-line model of Liebe 1985), specific gas attenuation for saturation with respect to liquid water κ_{gs} (dB km^{-1}) and specific liquid water attenuation κ_l (dB km^{-1} [g m^3]$^{-1}$), using the formulation given by Liebe et al. (1989). For details of their use, see sections 3.4.1 and 3.5.

In *Cloudnet* we use the model over the site which has the highest horizontal resolution, and for which forecasts closest to t+0 are available.

2.4 Microwave radiometer

A recommended but optional input dataset is liquid water path (LWP) derived from dual- or multi-frequency microwave radiometers. LWP is used in many retrieval algorithms for liquid water clouds. Ideally LWP will not be derived using climatologically tuned coefficients, as these can result in errors exceeding 50 g m^{-1} and retrieved values that are negative. One method is described by Gaussiat (2004) which makes use of other sources of data to provide more accurate coefficients; lidar is used to locate the height (and therefore the temperature) of liquid cloud, and the model to provide a more appropriate humidity distribution than climatology. Also, profiles identified by the lidar to be free of liquid water cloud are used to estimate the zero offset and by interpolation across periods of cloud, to provide considerably more accurate LWP in thin cloud. This approach also has the advantage that it can produce accurate LWP even if with poorly calibrated radiometer channels.
LWP is important for the correction of radar attenuation in liquid water clouds; at 94 GHz it can exceed 5 dB and so if correction is not performed then reliable retrievals in ice clouds above liquid clouds based on the value of Z are impossible

2.5 Rain gauge

The presence of rain on the dish or radome of a radar can result in a large and variable attenuation (Hogan et al. 2003a), of order 10 dB, making it impossible to use the absolute value of Z with any reliability. Rain itself also extinguishes the signal and can be difficult to correct for; microwave radiometers do not provide accurate LWP estimates in rain (Crewell and Loehnert, 2003). Radar attenuation due to the melting level is also uncertain. To diagnose the presence of rain we use either a rain gauge or the radar itself. Ideally the rain gauge should have a fairly high sensitivity; tipping-bucket gauges in particular can be very slow to register the start of light rain events. The radar parameters Z, v, and σ_v in the lowest few gates can also be used to diagnose probable rain on the ground. This approach tends to be more sensitive than a rain gauge, so the rain gauge is not treated as an essential input dataset.

3 The algorithm

3.1 Standardization of conventions

The data ingested consist of the individual fields and the associated meta-data (e.g. from the NetCDF attributes). First, the vertical coordinate of each input dataset is converted to height above mean sea level in metres. This correction is necessary as most instruments report range from the instrument in kilometres, and they may have been mounted at different heights above the ground. As the lidar often points several degrees from zenith, the range reported by this instrument is multiplied by the sine of the zenith angle to obtain height. Variable and attribute names are standardized, and variables are typically "expanded" into floating-point representation.

3.2 Regridding

A "universal" grid is defined which consists essentially of the radar time and height grid. However, if there are any temporal gaps in the model (radiosonde) or lidar data, or if the model or lidar height grids start higher or end lower than the radar height grid, then the universal grid will be reduced so that there is always a model and lidar pixel in the vicinity of a radar pixel. The lidar and model variables (β, T, p, q, κ_g, κ_{gs}, κ_l) are then interpolated on to the universal grid. Model data is interpolated linearly; for the lidar we wish to conserve the integrated backscatter as it is useful for calibration (O'Connor et al. 2004) and estimating optical depth in certain situations (Hogan et al. 2003b). This is done by using nearest-neighbour interpolation in time followed by integration in height, interpolation of the integral on to the universal grid, and differentiation. The one-dimensional fields, rain rate and liquid water path, are interpolated linearly on to the radar grid. Missing values are assigned as NaN. In the case of missing rain rates the radar is used to indicate the presence of rain at the ground and hence when the reflectivity values may be unreliable (see section 3.3.1), while in the case of missing liquid water path the attenuation flags are used to indicate pixels when liquid water below them is likely to have caused attenuation that has not been corrected (see sections 3.3.3 and 3.5.2). The result is a set of 2D fields that share the same grid, and a few 1D fields that share the time grid of the 2D fields.

3.3 Data quality flags

Before the target categorization algorithm can be applied, quality control is performed, including diagnosis of the likelihood of rain at the ground and identifying radar pixels affected by ground clutter. The resulting data quality information is presented as the time-height bit field quality_bits, where the bits have the following interpretation:
Bit 0: An echo is detected by the radar.
Bit 1: An echo is detected by the lidar.
Bit 2: The apparent echo detected by the radar is ground clutter or some other non-atmospheric artifact.
Bit 3: The echo detected by the lidar is due to clear-air molecular scattering.
Bit 4: Liquid water cloud or rainfall below this pixel will have caused radar and lidar attenuation; if bit 5 is set then a correction for the radar attenuation has been performed; otherwise do not trust the absolute values of reflectivity factor. No correction is performed for lidar attenuation.
Bit 5: Radar reflectivity has been corrected for liquid-water attenuation using the microwave radiometer measurements of liquid water path and the lidar estimation of the location of liquid water cloud; be aware that errors in reflectivity may result.
Bit 6: The echo detected by the lidar is due to noise.

We now describe the criteria that are used to determine the quality of each pixel.

3.3.1 1-D Rain bit

Rain at the ground can wet the radar dish or radome, causing strong (and unknown) attenuation and therefore rendering the reflectivity values above unreliable. Additionally, the raindrops themselves can cause attenuation that is not well estimated using the technique for liquid clouds described in section 3.5.2, as the attenuating particles will be outside the Rayleigh size regime and attenuation may no longer be proportional to liquid water content. Furthermore, it is difficult to partition the measured liquid water path (which may itself be affected by water on the instrument) with height. To facilitate the formulation of the 2-D attenuation flags which indicate regions of corrected and uncorrected attenuation, a 1-D bit is first generated to indicate that the reflectivity values in the profile may be unreliable due to rain at the ground. If rain gauge data are available then the bit is set to unity for all times when the rain gauge indicates non-zero rain rate. If

rain gauge data are not available then the rain rate variable is instead created from the radar information, and consists of zeros when the radar indicates no rain at the ground (when Z is less than 0 dBZ in the third (or lowest-clutter free) range gate above the ground) and NaN when the radar indicates the likelihood of rain at the ground but is unable to accurately estimate the rate (when Z is greater than 0 dBZ, corresponding to a rain rate of around 0.05 mm h-1). The 1-D rain bit is then set to unity whenever rain rate is NaN.
Finally, all pixels within 2 minutes of rain are also deemed to be raining in the 1-D rain bit, and additionally any clear spells between rain bits that are less than 2 minutes long are also set to raining.

3.3.2 Ground clutter

The characteristics of the ground clutter are strongly instrument dependent, and an algorithm is available to identify ground clutter. Note that most radar manufacturers now implement their own clutter algorithm, so that this section is not required.
Hereafter Z refers to radar reflectivity with clutter pixels removed, and Z_c refers to reflectivity with clutter untouched. Note that in the final product it is Z_c that is reported, but Bit 2 of the quality_bits bit field indicates the location of probable radar clutter so that it may be removed. The most robust ground clutter algorithm utilises the fact that ground clutter tends to have a Doppler velocity close to zero. The whole day is analysed at a time, but only the first 10 gates are considered for groundclutter removal. If rain is detected at the ground then it is assumed that the return from the lowest gates will be dominated by rain and no attempt is made to identify clutter in these profiles. At each height starting with the lowest, pixels are deemed to be dominated by clutter whenever the rain bit is not set, v lies between 0.05 and -0.05 m s^{-1} and $\sigma_{\bar{v}}$ is less than 0.2 m s^{-1}. If a height is reached where no clutter is found at any time in the day then the gates above are not analysed.

3.3.3 Molecular scattering

Note that most ceilometers do not detect molecular scattering so Bit 3 of quality_bits is not set by them. However, molecular scattering is much stronger at visible, and especially UV, wavelengths; and will be detected by virtually all instruments operating at these wavelengths.

3.3.4 Attenuation bits

Two bits are provided to indicate the radar pixels that have been affected by attenuation and those for which a correction has been made. The first (Bit 4 of quality_bits) indicates that the radar and lidar returns have been attenuated by liquid water cloud, rain and/or melting ice in the intervening pixels, while the second (Bit 5 of quality_bits) indicates that microwave radiometer estimates of liquid water path have been used to correct this attenuation, but that errors in reflectivity may result. These bits are calculated after liquid attenuation correction has been performed, so we defer further discussion until section 3.5.2.

3.4 Target categorization

The type of target(s) present in each pixel is important for the application of subsequent algorithms, and is diagnosed using all available data. As several target types may be present in a given pixel, this information is also presented in the form of a bit field, with each bit typically representing a different type of target. Of course, some types of target can never be identified when others are present (such as aerosol when liquid water is present), but this format is intended to allow the maximum flexibility for when different types can be identified simultaneously. There are separate bits diagnosing the presence of liquid cloud droplets, melting ice particles, aerosols and insects. Two further bits are used to define liquid precipitation and ice: falling_bit, indicating particles that have appreciable terminal velocity, and cold_bit, indicating whether such falling particles are likely to be composed of ice or liquid water. In this simple scheme we are accepting that there is no distinct difference between ice cloud and ice precipitation (at least, not one that can be discerned from the observations; see Hogan et al. 2001), and also that we are unable to distinguish supercooled drizzle from ice, although this may be possible in future. We also do not distinguish between rain originating from melting ice, and drizzle originating from the warm rain process, although it would be a simple matter to do that from the final target categorization bit field. The format of the resulting category_bits variable is as follows:
Bit 0: Small liquid droplets are present (droplet_bit).
Bit 1: Falling hydrometeors are present; if Bit 2 is set then these are most likely to be ice particles otherwise they are drizzle or rain drops (falling_bit).
Bit 2: Wet-bulb temperature is less than 0 C, implying the phase of Bit-1 particles (cold_bit).
Bit 3: Melting ice particles are present (melting_bit).
Bit 4: Aerosol particles are present and visible to the lidar (aerosol_bit).
Bit 5: Insects are present and visible to the radar (insect_bit).

One would use category_bits to diagnose the presence of cloud as when Bit 0 is set (indicating the presence of liquid water droplets), or both Bits 1 and 2 were present (indicating the presence of ice particles). For convenience, *Cloudnet* also provides a simpler "classification" product at Level 2, consisting of numbers from 0 to 10 indicating the main combinations possible in the more complex "target categorization" bit field described above. It is important to stress that the intention with these two products is not to provide automated cloud classification in the classical sense (e.g. to distinguish cirrocumulus from cirrostratus), but to use criteria that are objectively defined from the measurements, useful for subsequent retrieval algorithms and which match some of the distinctions that are made in numerical forecast models.

We now describe the categorization algorithm in more detail, considering each bit in turn. The following sections describe the separate functions (some of which calculate more than one bit) and are headed by the input variables that they make use of. It should be noted that many of the procedures contain seemingly arbitrary parameters; these have been chosen to produce the best agreement with a subjective analysis of real cases. The order in which the bits are described matches the order in which they are calculated in the program.

3.4.1 Cold bit, melting bit

Input fields: Model wet bulb temperature T_w, Doppler velocity v, folding velocity v_{fold}.

The purpose of cold_bit is to indicate where "falling" particles are likely to be composed of ice rather than liquid. It is initially defined to be where the wet-bulb temperature T_w (calculated from model temperature, pressure and humidity) is less than 0C (note that falling ice melts when T_w, rather than T, becomes positive). To cope with isothermal layers, this is actually implemented such that all pixels below the highest 0C isotherm in the profile are deemed to be "warm", since melted ice precipitation is unlikely to refreeze. This field is then refined using the radar to locate the melting layer more precisely in stratiform precipitation. The radar reflectivity profile usually provides a distinct step at 94 GHz where ice particles melt (see Mittermaier and Illingworth 2003) and a bright band at lower frequencies. However, the Doppler profile provides a more distinct signal, with a large and sharp increase in fall speed at the point of melting. Additionally, for those instruments with LDR capability, the melting layer displays much higher values of LDR (> -15 dB) than in ice or rain (< -20 dB). The algorithm is also designed to deal with low folding velocities (less than 4 m s-1).

3.4.2 Droplet bit

Input fields: Attenuated lidar backscatter coefficient β, radar reflectivity factor Z, cold bit (defined above), temperature T.

Each lidar ray is examined in turn, and searched for one or more liquid layers. Lidars with depolarisation provide direct detection of liquid water layers (at cloud base), but not all sites are equipped with such instruments. In these situations, we utilise the fact that to lidar the base of liquid clouds appears as a strong echo that is confined over only a few hundred metres. Note that Hogan et al. (2003b) used the integrated backscatter through the layers to assist in their diagnosis; while this enabled the optical thickness necessary to trigger their algorithm to be defined, it only allowed one liquid layer to be identified in any given profile. Here we are interested in identifying pixels that, on the balance of probabilities, contain liquid droplets, so need not be so restrictive in our criteria. The first liquid layer is found by locating the lowest pixel in the ray where both $\beta > 2 \times 10^{-5}$ m^{-1} sr^{-1}, and the β' value 250 m higher up is a factor of 10 lower; this is denoted the "pivot" value from now on. The maximum gate-to-gate increase in β' in the 100 m below the pivot, Δβ, is calculated. Liquid cloud base is defined as the lowest pixel in this 100 m range for which the difference in β' between it and the pixel above exceeds Δβ.
Lidar cloud top is defined as follows. If the lidar return falls to zero (below instrument sensitivity) within 300 m above the pivot then the top is defined to be the last non-zero pixel just below this point. Otherwise the procedure is similar to cloud base; Δβ is calculated as the maximum decrease in the gate-to-gate β' in the 300 m above the pivot, and cloud top is declared to be the highest pixel in this range where the decrease in β' from the pixel below exceeds Δβ/4.
Then the radar profile is analysed to determine cloud top in the case that the lidar has been extinguished while the radar still has a signal. If we are in a sub-zero region (as determined by cold_bit), the radar is only searched a further 300 m above the lidar-diagnosed top; otherwise we search up to the last pixel where cold_bit is zero. If there are any radar pixels in this region in which no signal is detected, then cloud top is changed to be the pixel immediately beneath the first pixel where no radar signal is detected. If on the other hand there is a radar signal throughout this region, then it is regarded as ice or drizzle falling from further up in the profile. The use of rather arbitrary search distances sometimes results in erroneous liquid water profiles. This simply represents the difficulty in locating the tops of clouds when the radar is dominated by larger particles falling through them and the lidar signal has been extinguished. The droplet_bit is then set to unity in all pixels between cloud base and top. The next layer is diagnosed by repeating this process, with the lidar profile searched above the pivot for the next pixel with $\beta > 2 \times 10^{-5}$ m^{-1} sr^{-1} to act as a new pivot. Note that if several contiguous pixels have β' exceeding this value then the pixels in the vicinity may be analysed several times to determine

if they qualify for droplet_bit being set. Finally, the droplet bit is set to zero for any pixels with T<-40 C, as liquid water cannot persist at these temperatures.

3.4.3 Falling bit, insect bit

Input fields: Radar reflectivity without clutter Z, radar reflectivity with clutter Z_c, attenuated backscatter coefficient β, cold bit and droplet bit (defined above), 1D rain bit.

The falling_bit incorporates rain, drizzle and all ice particles. Discrimination between ice and liquid is then possible using cold_bit defined in the previous section. The procedure is to first assign all radar echos that have not been identified as clutter as "falling", then to remove drizzle-free liquid clouds and to reassign those due to insects using insect_bit. We first consider profiles containing no liquid water droplets at any height and no radar echo in the lowest sub-zero pixel (the first in the profile with cold_bit =1 (indicating ice about to melt). In these profiles all pixels with a finite radar echo (using Z with clutter removed) that have cold_bit set are assigned to be "falling", while those in the warm region are assigned as "insects". However, if rain is detected at the ground (as indicated by the 1-D rain bit), then all radar echos in the profile are assigned to be falling.
We next consider profiles containing liquid water droplets as indicated by droplet_bit. For the purposes of this algorithm, if a radar echo is received from the lowest pixel in the profile with cold_bit set, then this is also considered to be a (single-pixel) liquid cloud. The reason is that both liquid clouds and ice just on the verge of melting can be considered as sources of liquid precipitation, and the same methodology is used on the pixels below them to distinguish precipitation from insects. Thus the bases and tops of each of the liquid clouds in the profile are determined. A simple method is then used to discriminate insects from drizzle beneath the first cloud base. If radar echos are recorded continuously between the ground and the first cloud base then the minimum reflectivity with clutter, Z_c, is found, and pixels above this are designated "falling" while those below are "insects". If the radar echo is not continuous then falling_bit is set only in the contiguous finite radar echos below cloud base; all finite radar echos below this are assigned as insects.

Next the pixels within the cloud are considered; the aim is to remove from falling_bit those for which the radar echo is predominantly due to liquid droplets. Firstly cloud top is examined. If there is a radar signal in the pixel above it then this implies that precipitation particles are falling into the liquid cloud, so all radar pixels within the cloud are deemed to be falling. If there is no radar signal immediately above cloud top then we must decide whether any radar echo within the cloud is predominantly due to the cloud itself or to growing drizzle or ice. As precipitation-free liquid water clouds tend to have liquid water content increasing with height, the reflectivity of such clouds also increases with height. Conversely, when precipitation is present it tends to grow due to accretion of liquid water, and Z decreases with height. We therefore compare the Z values 20% above cloud base and 20% below cloud top. If from these two pixels Z increases with height then the entire cloud is deemed to be free of precipitation and falling_bit is set to zero (although note that Z might be finite just below cloud base and so falling_bit might still be set here). If Z decreases with height then drizzle or ice is deemed to be present in the profile, and falling_bit is set to unity between cloud base and the highest pixel below cloud top where Z exceeds 30 dBZ. This procedure is repeated for all liquid layers, although the insect/drizzle distinction is only performed for the first layer. This has the unfortunate consequence that any pixel with a finite radar echo above the first cloud base that is not within not within a cloud is declared to be "falling" even though subjectively it might be identified as insects. A major improvement in insect discrimination is possible when LDR is available (LDR > -12 dB indicates insects, clutter, or melting particles).

Some tenuous ice clouds are only detected by the lidar. These are added by declaring any pixel colder than -40 C for which a lidar echo is received to be "falling". Additionally, pixels between 0 and -40 C are also declared as ice, provided that cold_bit is set, droplet_bit is not set, and that there is a gap between the lowest such pixel and aerosol in the boundary layer. Note that it is possible for aerosol pixels to satisfy these criteria. Lidars with depolarisation capability are much better at discriminating between tenuous ice and aerosol layers.

Finally a modification is made to melting_bit, which previously was set for all radar echoes immediately beneath the lowest cold_bit. For those pixels that have been reassigned to insects, melting_bit is unset.

3.4.4 Aerosol bit

Input fields: Attenuated lidar backscatter coefficient β, falling bit, droplet bit, cold bit.

A pixel is deemed to be aerosol (aerosol_bit is set) if a finite lidar signal is present, and it is at or below a specific altitude (site dependent), or if cold_bit is not set, and neither droplet_bit, nor falling_bit have been set. Again, lidars with depolarisation capability are much better at discriminating between the various atmospheric particulates.

3.5 Attenuation correction

The radar reflectivity profile is corrected for the effects of both liquid water and gas (predominantly water vapour and oxygen) attenuation. Ice attenuation in vertical profiles can be considered negligible at frequencies below 100 GHz. Two fields are generated on the same universal grid as used by the radar; the two-way gas attenuation and the two-way liquid attenuation, both in dB. These are simply added to the reflectivity field (in dB*Z*). They are also recorded as part of the product, so that the user may recover the actual measured *Z* field if desired.

3.5.1 Gas attenuation

The *Cloudnet* model data includes κ_g, the specific gas attenuation at the frequency of the radar calculated from the model temperature, pressure and humidity, and also κ_{gs}, the specific gas attenuation for 100% humidity with respect to liquid water. A "most likely" specific gas attenuation field is generated as a combination of the two, using droplet_bit dictate when κ_{gs} should be used rather than κ_g. The result is integrated with height to obtain the total 2-way gaseous attenuation, designated radar_gas_atten in the output NetCDF file.

3.5.2 Liquid water attenuation

Liquid water attenuation is estimated by partitioning the LWP measured by microwave radiometer with height amongst those pixels with droplet_bit set to obtain the "most likely" profile of liquid water content (LWC). This is done by considering each individual liquid water cloud (i.e. each contiguous sequence of pixels with droplet_bit = 1) and calculating the adiabatic profile of LWC using *T* and *p* at each cloud base (specifically LWC is assumed to increase linearly with height from the base of each liquid layer). Then the entire profile is scaled to match the LWP value. The resulting LWC is then multiplied by $2\kappa_l$ and integrated with height to obtain the cumulative 2-way liquid attenuation in that profile, designated radar_liquid_atten in the output NetCDF file. If LWP reported by the microwave radiometers is zero or negative then no liquid water correction is performed; even if droplets are present then it is assumed that the LWC is so small that liquid attenuation is negligible. If no LWP is available, then the liquid attenuation in the lowest pixel containing droplets, and all those above, will be set to NaN. The radar *Z* will not then be corrected for liquid attenuation.
At this point in the algorithm enough information is available to populate the two data quality bits referring to attenuation, discussed briefly in sections 3.3. The first (Bit 4) indicates if any radar pixel has been attenuated by either liquid water or rain; this bit is set if radar_liquid_atten is non-zero (either positive or NaN) or if the 1-D rain bit is set for that profile. The second (Bit 5) indicates if a correction has been performed for liquid attenuation, and is set if radar_liquid_atten is non-zero and non- NaN, *and* the 1-D rain bit is *not* set for that profile. Hence if no LWP measurements are available, pixels within and above liquid cloud would have the first of these bits set and the second unset, indicating that liquid attenuation correction has not been performed and the *Z* values should not be relied upon by subsequent retrieval algorithms.

3.7 Instrument errors

As detailed here[6], all variable are assigned a *bias* and an *error*, where appropriate. In addition to the precision estimate of radar reflectivity factor, Z, there are also uncertainties in the attenuation corrections.
The error in gas attenuation. The humidity profile from the model will not be perfect, resulting in an error due to the correction of the radar reflectivity for gas attenuation. We assume the error in *Z* from this source to be 10% of the 2-way cumulative gas attenuation to that height.
The error in liquid attenuation. The combination of error in LWP and the partitioning of liquid water with height results in an error in *Z* due to correction for liquid attenuation. This is estimated by following the same steps for computing the liquid water attenuation as was carried out in section 3.5.2, but using the *error* in LWP to scale the LWC profile rather than LWP itself. To account for the possibility that LWC does not increase linearly with height but has more of a symmetric profile, this calculation assumes constant LWC with height in each liquid cloud. As liquid water attenuation is much less at 35 GHz than 94 GHz, radars at 94 GHz have a considerably higher error in retrievals from ice clouds above liquid water clouds.

4 Conclusions

A data product has been described that performs much of the preprocessing that is necessary before synergetic radar/lidar algorithms can be applied, and should facilitate the development of cloud retrieval algorithms that can be applied to large volumes of data. The target categorization component should be straightforward to use in simpler studies just concerned with cloud boundaries, such as evaluation of cloud fraction in models and investigations into cloud overlap.

[6] ACTRIS Deliverable 5.6: Establish a common set of definitions for instrument errors

References

Atlas, D., 1964: Advances in radar meteorology. *Advances in Geophys.,* **10,** 318–478.
Bouniol, D., A. J. Illingworth and R. J. Hogan, 2003: Deriving turbulent kinetic energy dissipation rate within clouds using ground based 94 GHz radar. *Proc. 31st AMS Conf. on Radar Meteorology, Seattle,* 193–196.
Clothiaux, E. E., K. P. Moran, B. E. Martner, T. P. Ackerman, G. G. Mace, T. Uttal, J. H. Mather, K. B. Widener, M. A. Miller and D. J. Rodriguez, 1999: The Atmospheric Radiation Measurement program cloud radars: Operational modes. *J. Atmos. Oceanic Technol.,* **16,** 819–827.
Donovan, D. P., A. C. A. P. van Lammeren, H. W. J. Russchenberg, A. Apituley, R. J. Hogan, P. N. Francis, J. Testud, J. Pelon, M. Quante and J. W. F. Goddard, 2001: Cloud effective particle size and water content profile retrievals using combined lidar and radar observations - 2. Comparison with IR radiometer and in situ measurements of ice clouds. *J. Geophys. Res.,* **106,** 27 449–27 464.
Doviak, R. J., and D. S. Zrni´c, 1993: *Doppler radar and weather observations.* 2nd Ed. Academic Press.
Gaussiat, N.,R. J. Hogan and A. J. Illingworth, 2007: Accurate liquid water path retrieval from low-cost microwave radiometers using additional information from lidar and operational forecast models. *J. Atmos. Oceanic Technol.*, **24**, 1562-1575.
Goddard, J. W. F., J. Tan and M. Thurai, 1994: Technique for calibration of meteorological radars using differential phase. *Electronics Letters,* **30,** 166–167.
Hogan, R. J., 1998: *Dual-wavelength radar studies of clouds.* PhD Thesis, University of Reading, UK.
Hogan, R. J., A. J. Illingworth and H. Sauvageot, 2000: Measuring crystal size in cirrus using 35 and 94-GHz radars. *J. Atmos. Oceanic Technol.,* **17,** 27-37.
Hogan, R. J., C. Jakob and A. J. Illingworth, 2001: Comparison of ECMWF cloud fraction with radar-derived values. *J. Appl. Meteorol.,* **40,** 513–525.
Hogan, R. J., D. Bouniol, D. N. Ladd, E. J. O'Connor and A. J. Illingworth, 2003a: Absolute calibration of 94/95-GHz radars using rain. *J. Atmos. Oceanic Technol.,* **20,** 572–580.
Hogan, R. J., A. J. Illingworth, E. J. O'Connor and J. P. V. Poiares Baptista, 2003b: Characteristics of mixed-phase clouds - 2. A climatology from ground-based lidar. *Quart. J. Roy. Meteorol. Soc.,* **129,** 2117– 2134.
Hogan, R. J., N. Gaussiat and A. J. Illingworth, 2004: Stratocumulus liquid water content from dual-wavelength radar. *Submitted to J. Atmos. Oceanic Technol.*
Liebe, H. J., 1985: An updated model for millimeter-wave propagation in moist air. *Radio Science,* **20,** 1069–1089.
Liebe, H. J., T. Manabe and G. A. Hufford, 1989: Millimeter-wave attenuation and delay rates due to fog/cloud conditions. *IEEE AP,* **37,** 1617–1623.
Mace, G. G., C. Jakob and K. P. Moran, 1998: Validation of hydrometeor occurrence predicted by the ECMWF using millimeter wave radar data. *Geophys. Res. Lett.,* **25,** 1645–1648.
Mittermaier, M. P., and A. J. Illingworth, 2003: Comparison of model derived and radar-observed freezing level heights: Implications for vertical reflectivity profile correction schemes. *Quart. J. Roy. Meteorol. Soc.,* **129,** 83–96.
O'Connor, E. J., R. J. Hogan and A. J. Illingworth, 2005: Retrieving stratocumulus drizzle parameters using Doppler radar and lidar. *J. Appl. Meteorol.*, ***44****, 14-27*
O'Connor, E. J., A. J. Illingworth and R. J. Hogan., 2004: A technique for autocalibration of cloud lidar, *J. Atmos. Oceanic Technol*., **21**, 777-786.
Sekelsky, S. M., W. L. Ecklund, J. M. Firda, K. S. Gage and R. E. McIntosh, 1999: Particle size estimation in ice-phase clouds using multifrequency radar reflectivity measurements at 95, 33 and 2.8 GHz. *J. Appl. Meteorol.,* **38,** 5–28.