The Representation of Low Cloud in the Antarctic Mesoscale Prediction System

THESIS

Presented in Partial Fulfillment of the Requirements for the Degree Master of Science in
the Graduate School of The Ohio State University

By
Karen Pon, B.S.
Graduate Program in Atmospheric Science

The Ohio State University
2015

Master’s Examination Committee:
Professor David Bromwich, Advisor
Professor Jay Hobgood
Professor Jialin Lin
Abstract

The accuracy of cloud prediction in Antarctica can have a significant impact on aviation operations. Unforecast low cloud can endanger an aircraft attempting to land, and affect a pilot’s ability to distinguish the horizon and surface features while in flight. Over-forecasting of low cloud results in fewer missions completed. A number of cloud forecast products have been developed over the years however forecasters often prefer to use the low level relative humidity (RH) fields to forecast low cloud. This study investigated the use of the Stoelinga-Warner algorithm to generate the current Antarctic Mesoscale Prediction System (AMPS) cloud base height forecast and whether a RH threshold could be used as a proxy for cloud base height.

The Stoelinga-Warner algorithm was tested using a case study of a mesoscale low in Prydz Bay near Davis station. The algorithm was insensitive to changes in the phase scheme and light extinction threshold used to predict cloud base. Further investigation revealed inadequate quantities of cloud hydrometeors, indicating a problem with the model’s microphysics scheme. Therefore, AMPS combined with the Stoelinga-Warner algorithm does not accurately predict cloud base height.

Cloud base heights derived from radiosonde RH thresholds were compared with synoptic observations for Davis, McMurdo and Halley. Lidar observations were also tested against both synoptic observations and radiosonde-derived cloud base heights at
Halley. The optimal RH threshold for predicting cloud base height was ~70% at Davis and McMurdo, and ~90% at Halley.

AMPS RH data was used to generate cloud base heights at different thresholds, and these were verified against synoptic observations. Results were mixed due to the comparatively large scatter in the model RH field, with the optimal RH threshold changing according to the verification metric used. However there was broad agreement that Davis and McMurdo required a lower RH threshold than Halley. The thresholds found for Davis and McMurdo are consistent with a study by Inoue et al. (2015) which found optimal RH thresholds between 58% and 66% for Davis, Casey and Mawson stations. The reason for the much higher threshold at Halley is unclear, and further studies are required to determine whether a general RH threshold can be applied across the continent to predict cloud base height.
Acknowledgments

Firstly, I would like to express my sincere gratitude to my advisor Professor David Bromwich for his encouragement, direction and support. His guidance helped me in the research and writing of this thesis.

I would like to thank Professor Jay Hobgood and Professor Jialin Lin for their time and effort as part of my committee.

Steve Colwell of the British Antarctic Survey provided valuable assistance with obtaining the Halley data.

Kevin Manning of UCAR patiently answered many queries about AMPS.

My colleagues from the Polar Meteorology Group at the Byrd Polar and Climate Research Center provided copious amounts of technical and, more importantly, moral support throughout my time here.

This research was supported by the National Science Foundation grant GRT00032749 via UCAR.
Vita

1994...............................................................Leeming Senior High School

1999 .............................................................B.Sc. (Hons) Environmental Science /

Mathematics, Murdoch University

2000 .............................................................G.Dip. Meteorology, Bureau of

Meteorology Training Centre

2013.............................................................M.B.A., Chifley Business School

2013 to present .............................................Graduate Research Associate, Byrd Polar

and Climate Research Center, The Ohio

State University

Fields of Study

Major Field: Atmospheric Sciences
Table of Contents

Abstract ........................................................................................................................................ ii

Acknowledgments ......................................................................................................................... iv

Vita.................................................................................................................................................... v

List of Tables ...................................................................................................................................... viii

List of Figures ..................................................................................................................................... ix

Chapter 1: Introduction ...................................................................................................................... 1

Chapter 2: Literature Review ............................................................................................................ 8

Chapter 3: Methods ........................................................................................................................... 18
  3.1 Part I: The Stoelinga-Warner Algorithm .............................................................................. 18
  3.2 Part II: Relative Humidity Thresholds ................................................................................. 19
  3.3. Part III: AMPS Model Cloud Base Height Predictions from Relative Humidity Thresholds ......................................................................................................................... 21

Chapter 4: Results and Discussion ................................................................................................ 26
  4.1. Part I: The Stoelinga-Warner Algorithm .............................................................................. 26
  4.2. Part II: Relative Humidity Thresholds ................................................................................. 28
3.3. Part III: AMPS Model Cloud Base Height Predictions from Relative Humidity Thresholds

Chapter 5: Conclusions

References
List of Tables

Table 1: WMO SYNOP cloud base height code. ................................................................. 44
Table 2: 2x2 contingency table. ....................................................................................... 45
List of Figures

Figure 1: Davis Ski Landing Area (Photo: Karen Pon) ................................................. 46

Figure 2: Davis Ski Landing Area (Photo: Bob Heath) ................................................. 47

Figure 3: Map of Antarctica. Stations used in this study circled. (http://www.mapsofworld.com/antarctica/) .......................................................... 48

Figure 4: Map of Davis area. Courtesy of the Australian Antarctic Division (2011) ..... 49

Figure 5: MODIS visible channel (0.62-0.67 µm) satellite image for the Davis area at 0415 UTC on 22 January 2013. Distinct areas of low cloud are marked ......................... 50

Figure 6: Archived AMPS 27 hour forecast valid at 03 UTC on 22 January 2013 showing cloud base height ......................................................................................... 51

Figure 7: Archived AMPS 27 hour forecast valid at 03 UTC on 22 January 2013 showing 1000 ft relative humidity and horizontal wind fields ........................................... 52

Figure 8: Domain of AMPS Grid 2 (10 km continental grid). ....................................... 53

Figure 9: 27-hour cloud base (ft) forecast for the Davis area valid for 03 UTC 22 January 2013, using AMPS initialized at 00 UTC 21 January 2013, generated by the Stoelinga-Warner algorithm using the (a) Dudhia and (b) Reisner schemes .......................... 54

Figure 10: 27-hour forecast for the Davis area valid for 03 UTC 22 January 2013, using AMPS initialized at 00 UTC 21 January 2013, showing 1000 ft mixing ratio (g/kg) of (a) cloud liquid water and (b) cloud ice. ...................................................... 55
Figure 11: Australia Day swimming in open water at Davis station. (Photo: Bill de Bruyn)

Figure 12: 27-hour forecast for the Davis area valid for 03 UTC 22 January 2013, using AMPS initialized at 00 UTC 21 January 2013, showing 1000 ft mixing ratio of water vapor (g / kg).

Figure 13: 27-hour cloud base (ft) forecast for the Davis area valid for 03 UTC 22 January 2013, using AMPS initialized at 00 UTC 21 January 2013, generated by the Stoelinga-Warner algorithm using the Reisner scheme and light extinction threshold of (a) 2% and (b) 90% remaining light.

Figure 14: Frequency analysis of radiosonde relative humidity threshold of 70% vs synoptic cloud base height observation at Davis for December 2011 and January 2012.

Figure 15: Frequency analysis of radiosonde relative humidity threshold of 70% vs synoptic cloud base height observation at Davis for October to November 2011 and February to March 2012.

Figure 16: Frequency analysis of radiosonde relative humidity threshold of 70% vs synoptic cloud base height observation at McMurdo for December 2011 and January 2012.

Figure 17: Frequency analysis of radiosonde relative humidity threshold of 90% vs synoptic cloud base height observation at Halley for December 2011 and January 2012.

Figure 18: Frequency analysis of radiosonde relative humidity threshold of 90% vs synoptic cloud base height observation at Halley for October to November 2011 and February to March 2012.

Figure 19: Frequency analysis of lidar vs synoptic cloud base height observation at Halley for December 2011 and January 2012.
Figure 20: Frequency analysis of lidar vs synoptic cloud base height observation at Halley for October to November 2011 and February to March 2012

Figure 21: Frequency analysis of lidar vs synoptic cloud base height observation at Halley for December 2011 and January 2012 with sky coverage ≥ 5 oktas

Figure 22: Frequency analysis of lidar vs synoptic cloud base height observation at Halley for October to November 2011 and February to March 2012 with sky coverage ≥ 5 oktas

Figure 23: Frequency analysis of radiosonde relative humidity threshold of 90% vs lidar cloud base height observation at Halley for December 2011 and January 2012

Figure 24: Frequency analysis of radiosonde relative humidity threshold of 90% vs lidar cloud base height observation at Halley for October to November 2011 and February to March 2012

Figure 25: Comparison of RMSE of AMPS cloud base height prediction by RH threshold at each station for the summer period

Figure 26: Comparison of RMSE of AMPS cloud base height prediction by RH threshold at each station for the shoulder period

Figure 27: Comparison of bias of AMPS cloud base height prediction by RH threshold at each station for the summer period

Figure 28: Comparison of bias of AMPS cloud base height prediction by RH threshold at each station for the shoulder period

Figure 29: Gerrity Skill Score for Davis

Figure 30: Gerrity Skill Score for McMurdo

Figure 31: Gerrity Skill Score for Halley
Figure 32: Critical Success Index for the summer period.......................... 77

Figure 33: Critical Success Index for the shoulder period.......................... 78

Figure 34: Proportion of correct forecasts over the summer period. ............... 79

Figure 35: Proportion of correct forecasts over the shoulder period. ............... 80
Chapter 1: Introduction

Accurate weather forecasting is critical for many operations in Antarctica, in particular for logistics and transport. Many programs use aircraft for transport both to and within Antarctica (Bromwich et. al. 2013, Powers et. al. 2012, Turner and Pendlebury 2004). Accurate weather forecasts contribute to the optimal use of aviation resources, minimizing the number of flights required to return to their point of origin due to poor weather, and fewer field trips canceled or delayed during periods of good weather. For example, the introduction of the Antarctic Mesoscale Prediction System in 2001 saw a significant reduction in aborted flights on the intercontinental Christchurch-McMurdo route in subsequent years, representing substantial cost savings (Powers et. al. 2012).

Aviation safety is another major concern in Antarctica. Some inter- and intra-continental flight routes extend beyond the point of safe return (PSR) of the aircraft, so pilots and aviation managers need a high level of confidence in flight forecasts and forecasters need a high level of confidence in their decision aids. There have been numerous weather-related incidents in recent years. A fatal helicopter crash near Dumont D’Urville in 2010 was attributed to deteriorating weather conditions (Bureau d’Enquêtes et d’Analyses 2012, Werfelman 2012). In 2013 a Royal New Zealand Air Force Boeing
757-200 jet was forced to land in below minimum visibility conditions at McMurdo having passed the PSR with a forecast of visual conditions (Croft 2015, Transport Accident Investigation Commission 2015). Also in 2013, a helicopter crashed near Davis in white-out conditions after low cloud moved in earlier than forecast, resulting in serious injuries to those aboard (Australian Transport Safety Bureau 2015). Reduced visibility in cloud may also have contributed to a fatal Twin Otter crash in 2013 at Mount Elizabeth in the Queen Alexandra Range (Transportation Safety Board of Canada 2015).

Cloud is an important meteorological phenomenon for aviation operations, especially in Antarctica. Each of the incidents described above involved reduced visibility for aircraft flying in cloud. The general remoteness of Antarctica and scarcity of alternate airfields or suitable landing areas increases the importance of accurate cloud prediction in the region. Cloud base, which is the height above ground level of a layer of cloud, can significantly impact aviation operations.

Few airfields in the region have the necessary infrastructure for instrument landings, so pilots require adequate visibility of the landing area. Airfields have a minimum cloud base (also known as ceiling) for aircraft landing. Some aircraft types are restricted to flying under visual flight rules, and require ground visibility at all times. Low cloud, usually defined as having a base up to 6500ft or 2000m above the surface (Ahrens 2013), impedes visibility and thus restricts flying. Unpredicted low cloud at an airfield presents a serious safety hazard for aircraft in flight if there is no alternate landing area.
Cloud cover also affects the surface contrast and horizon definition, two parameters specific to polar aviation operations. The surface contrast describes “the ease with which features on a snow–covered surface can be distinguished” (Turner and Pendlebury 2004), which affects the assessment of a landing area. The horizon definition is “the ease with which the boundary between the ground and the sky can be determined” (Turner and Pendlebury 2004), which affects pilots in flight. Both of these parameters are assessed on a four-point scale of “good”, “fair”, “poor” and “nil”. Even modest amounts of cloud can reduce these two parameters significantly.

Figure 1 shows an aerial view of the Davis Ski Landing Area, which is on the Antarctic Plateau near Davis station. There is scattered thin high level cloud, which normally would not be considered a hazard to aviation. On the right side of the image the horizon is clearly delineated, however on the left the horizon definition is degraded to “poor” or “nil” as the distant cloud merges with the snow and ice landscape of the Antarctic Plateau. Although the surface features can still be identified, the pilot assessed the surface contrast as being “poor”. Figure 2 shows a Twin Otter aircraft parked at the Davis Ski Landing Area under a broken layer of thick low cloud. The horizon is barely discernible and no surface features can be distinguished.

The Antarctic Mesoscale Prediction System (AMPS) is a high resolution numerical weather prediction system used for operational weather forecasting in Antarctica (Powers et al. 2012). It uses a version of the Weather Research and Forecasting model with modifications for the polar environment (Polar-WRF) (e.g., Bromwich et al. 2013, Hines et al. 2015). The model uses a polar stereographic
projection centered at the South Pole (AMPS 2015c). There are several nested horizontal
grids, with the outer grid at 30 km resolution (AMPS 2015a, c). There are 61 vertical
levels, using a terrain-following vertical co-ordinate (AMPS 2015a).

Products for operational forecasting are generated by a post-processing routine
based on the Read/Interpolate/Plot (RIP) program, which was designed for the
visualization of model output and commonly used for WRF output (Stoelinga 2009).
These are published on the publicly available AMPS website
(http://www.mmm.ucar.edu/rt/amps). One of the current operational products shows the
lowest cloud base over the forecast area. This is an important forecast and decision aid
for aviation operations.

Feedback from operational forecasters indicates that models currently perform
poorly with cloud prediction over Antarctica, in particular with respect to low cloud. A
study by Fogt and Bromwich (2008) comparing a previous AMPS pseudosatellite product
to satellite imagery also found poor performance of the model. Forecasters report that the
relative humidity (RH) fields from AMPS are often a more useful predictor of low cloud
presence in Antarctica than the current dedicated cloud base product. Model RH
thresholds have in the past been used to generate cloud forecast products, such as the
pseudosatellite image previously produced by the Australian Bureau of Meteorology
(Turner and Pendlebury 2004).

However the RH thresholds used by forecasters vary by location. During my
operational forecasting seasons at Davis (November to March of 2010-11 and 2011-12) I
found that model RH values of 70% provided the most reliable cloud predictions while
forecasters at McMurdo use 90-95% as a rule of thumb (Art Cayette, personal communication, 11 June 2013).

Three locations have been selected for this study – McMurdo, Halley and Davis. All three stations are near the coast, with open water during the height of summer. Figure 3 shows these locations on the Antarctic continent.

The U.S. station of McMurdo on Ross Island in McMurdo Sound is the largest station in Antarctica and the logistics hub of the U.S. Antarctic Program (National Science Foundation 2015). Both rotary and fixed-wing aircraft are based at McMurdo and there are regular intercontinental flights from Christchurch, New Zealand. The main summer operational season typically extends from October to February. The airfields at McMurdo are also used by the neighboring New Zealand station, Scott Base. Weather conditions at McMurdo are generally favorable, largely due to the high pressure ridge over the Antarctic continent extending over McMurdo Sound and Ross Island (Turner and Pendlebury 2004). Low cloud is often associated with an upper-level low pressure system known as the Ross Sea Low.

The British station of Halley is located on the Brunt Ice Shelf on the Weddell Sea. Its main summer operational season typically extends from December to February (British Antarctic Survey 2015). Steve Colwell, of the British Antarctic Survey, provided high-resolution radiosonde and lidar ceilometer data for Halley. The lidar data provides an objective measure of cloud base height directly over the station. Few stations in Antarctica have lidar ceilometers, so these data allows for the opportunity to include comparisons against the manual observations and the radiosonde. Halley’s weather is
dominated by synoptic-scale low pressure systems and low pressure systems forming in the lee of the Antarctic Peninsula (Turner and Pendlebury 2004).

The Australian station of Davis in Princess Elizabeth Land was selected due to my own extensive forecasting experience with the area. Davis is located at 68.6°S 78.0°E in East Antarctica, east of the Amery Ice Shelf and on the east coast of Prydz Bay (Fig. 4). Typically the station hosts both rotary and fixed-wing aircraft during the five month summer operational season from November to March.

During the summer, Davis typically experiences weather conditions favorable to aviation operations (Turner and Pendlebury 2004). From my forecasting experience at Davis, there are two main scenarios for low cloud during the operational season. Synoptic-scale low pressure systems often pass to the north, with associated warm fronts along the Davis coastline. These larger low pressure systems directly impact Davis more frequently during the remainder of the year.

The other common scenario for poor weather conditions in the area is a mesoscale low pressure system in Prydz Bay, locally referred to as a Prydz Bay Low. These typically bring low cloud, precipitation and poor visibility to the Davis area, severely restricting aviation operations. When these situations are poorly predicted there is significant disruption to field activities. Figures 5, 6 and 7 show an example of a Prydz Bay Low on 22 January 2013. Figure 5 is the MODIS visible wavelength satellite image for 0415UTC on that date, obtained from the NASA archive (http://ladsweb.nascom.nasa.gov/), with three areas of low cloud marked. Figures 6 and 7 respectively show the archived AMPS cloud base and low level RH forecasts for
0600UTC on the same date. The RH field is with respect to ice and is at 1000 ft above ground level (AMPS 2015b).

In the satellite image, an area of low cloud (A) can be seen along the east coast of Prydz Bay, extending inland and over the northern Amery Ice Shelf. Other layers of low cloud lie over the main part (B) and on the western side (C) of the Amery Ice Shelf. While much of the low cloud is identified by the model cloud base product there are significant areas which are not. In area C to the west of the Amery Ice Shelf there is low cloud evident in the satellite image that is not identified in the cloud base product. In this case there is also a lack of low level moisture in the RH field, indicating a problem with the raw model data. Area B over the Amery Ice Shelf is identified in the 1000 ft RH field as having RH over 70%, but not by the cloud base product as having cloud at any base below 20,000ft. This indicates a problem with the post-processing of the raw model data into the cloud base. It is this latter problem which this investigation will focus on.

This study will evaluate the current AMPS cloud base product and investigate methods to improve the low cloud forecasting skill of the model using the RH or other parameters.
Chapter 2: Literature Review

Tropospheric clouds in Antarctica are poorly understood (Bromwich et al. 2012, Lachlan-Cope 2010) with limited observational data available. Bromwich et al. (2012) conducted a review into several aspects of Antarctic clouds, including observational datasets, climatology, microphysics and model representation. Earlier, Lachlan-Cope (2010) reviewed the available observations of cloud properties.

Observational cloud data are limited in both space and time across Antarctica. Surface observations come from trained observers at staffed stations; however both the quantity and quality of these are reduced in winter (Bromwich et al. 2012). There are relatively few of these stations and most are located along the coast. Ground-based lidar cloud observations are even sparser.

Space-based observations are generated by both passive and active sensors. Passive sensors include the Advanced Very High Resolution Radiometer (AVHRR) and the more recent Moderate Resolution Imaging Spectroradiometer (MODIS). Cloud detection algorithms from passive sensors do not perform well in distinguishing cloud over ice surfaces (Bromwich et al. 2012). Active sensors such as the Geoscience Laser Altimeter System (GLAS), Cloud Profiling Radar (CPR) and the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) can measure cloud accurately, but their narrow path with much smaller spatial coverage reduces their usefulness (Bromwich et al. 2012).
Verlinden et al. (2011) investigated three years of CloudSat-CALIPSO data to study the seasonal variation in the vertical structure of Antarctic cloud. They found a strong seasonality with cloud at a maximum during winter, dominated by changes in upper-level cloud. A seasonal comparison of CloudSat-CALIPSO data by Adhikari et al. (2012) found that low cloud formed the majority of cloud cover and that it was more prevalent during summer in the vicinity of the circumpolar trough around the Antarctic coast.

A subsequent study of the CloudSat-CALIPSO data by Bromwich et al. (2012) showed a distinct seasonal cycle over the interior of East Antarctica. They found that cloudiness reaches a maximum over the Southern Ocean and decreases poleward due to lower temperatures, less moisture and steep coastal topography, with the decline more rapid across the East Antarctic coast. At the South Pole the cloud cover is typically 50-60% year-round and predominantly ice cloud, while in coastal regions the cloud cover rises to 80-90% and is often mixed phase cloud (Lachlan-Cope 2010).

Microphysical data have been obtained using a variety of methods and instruments on airborne and ground-based platforms, however are still very sparse. The review by Lachlan-Cope (2010) concluded that in general, cloud particle size is slightly smaller in the interior than on the coast. Since the review by Lachlan-Cope (2010) efforts on this topic have been increasing, with a number of dedicated cloud observation experiments in recent years (e.g., Gorodetskaya et al. 2015, Lawson and Gettelman 2014, Scott and Lubin 2014).
During January to February 2009, a cloud particle imager attached to a tethered balloon was deployed at the South Pole (Lawson and Gettelman 2014). Simultaneous observations from a ground-based micropulse lidar showed significant differences in the backscatter patterns of mixed-phase clouds compared with ice-only clouds. The micropulse lidar data were then extrapolated over the whole of 2009 to investigate the frequency of mixed-phase cloud. The authors found that mixed-phase clouds occurred more often than previously thought, and at times with very cold surface air temperatures.

During the 2012-13 operational season a spectroradiometer was deployed at McMurdo (Scott and Lubin 2014). The data were compared with satellite, radiosonde and ceilometer data in case studies of various synoptic situations to investigate cloud macro- and microphysical properties. In particular they found differences in cloud liquid water content according to the cloud origin and type.

A new meteorological observatory for cloud and precipitation was recently established at Belgium’s Princess Elisabeth station in Dronning Maud Land as part of the HYDRANT project (Gorodetskaya et al. 2015). A number of instruments were installed, including a lidar ceilometer, automatic weather station, infrared radiation pyrometer, micro rain radar, webcam and a snow pit.

The data were used to conduct several case studies of different synoptic-scale weather events during February 2012, as well as a blowing snow with clear skies event in April 2013. The authors were also able to compile a cloud cover climatology over a 14-month period and compare cloud properties. The cloud cover distribution was strongly
bimodal, with clear-sky conditions occurring 51% of the time, and overcast conditions 35% of the time. During the overcast conditions, 20% of clouds contained liquid water.

Further measurements and longer-term studies from these and similar sites would enhance the understanding of the microphysics of Antarctic clouds significantly. Scott and Lubin (2014) acknowledge the need for improvements to the modeling of Antarctic clouds in both weather forecasting and climate models.

Cloud microphysics schemes for mesoscale models have improved in recent years with the development of polar-optimized parameterization schemes (Bromwich et al. 2012). However difficulties remain with verification of these schemes due to the paucity of observations. In addition, many prediction schemes do not offer an explicit cloud fraction output, instead producing hydrometeor quantities.

Fogt and Bromwich (2008) investigated cloud prediction in AMPS for both point locations and areas. For point locations they calculated Cloud Fraction by vertically integrating cloud liquid water and cloud ice, comparing this to synoptic observations in the vicinity of McMurdo station and at South Pole station. They found that cloud fraction was significantly under-predicted at these locations. An adjustment to the coefficients in the Cloud Fraction algorithm allowed for bias correction however did not improve the correlations achieved, which remained low.

For the areal evaluation, Fogt and Bromwich (2008) compared satellite images from the Defense Meteorological Satellite Program (DMSP) with the former AMPS cloud product, the pseudosatellite image. This product aimed to represent cloud as it
would appear on a satellite image. It was generated by vertically integrating the cloud liquid water and cloud ice species.

The comparison was conducted by visual inspection over an area centered on the Ross Ice Shelf and extending beyond the South Pole and into the Ross Sea for a short period in January 2006. The 12-, 36- and 60-hour forecasts were used and the area was divided into a number of representative regions (Fogt and Bromwich 2008).

They found that the AMPS pseudosatellite product performed poorly across all regions apart from East Antarctica, which saw few instances of cloud. Low cloud was significantly under-predicted over continental areas when compared with DMSP imagery, and over-predicted over the western Ross Sea (Fogt and Bromwich 2008).

Bromwich et al. (2013) examined the performance of a range of variables for more recent versions of Polar WRF (versions 3.0.1, 3.1.1, 3.2.1 and 3.3.1). Model results were compared with observational data from selected stations, including synoptic observations, surface radiation measurements and radiosonde data.

Like Fogt and Bromwich (2008), they similarly found that low cloud was under-predicted compared with surface observations. This was observed from biases in the surface downwelling short- and longwave radiation as well as calculation of the cloud fraction as per the algorithm of Fogt and Bromwich (2008).

RH above the surface was compared to radiosonde data, with the model having small biases in the lower levels but larger errors and lower correlations compared with other variables assessed.
Inoue et al. (2015) investigated the relationship between RH and low cloud base height (LCBH) at the three Australian Antarctic stations of Davis, Mawson and Casey, all located along the coast of East Antarctica. They compared the low cloud base height reported by human observers to the RH profiles from both balloon-launched radiosondes and numerical weather prediction (NWP) to obtain a RH threshold for each location. They also calculated thresholds for the optical depth (τ), derived from the NWP mixing ratios for cloud liquid water and cloud ice, as an alternative measure.

The NWP data was sourced from the Australian Bureau of Meteorology’s Polar Limited Area Predictive System (Polar-LAPS), which was decommissioned in 2010 (Bureau of Meteorology 2015). The study was restricted to cases with extensive cloud cover, i.e., where a layer was reported to cover greater than or equal to 6 oktas and overcast conditions, i.e., where a layer was reported to cover 8 oktas. RH thresholds were assessed according to bias, with the threshold demonstrating bias nearest to zero selected as optimal.

The results gave an optimal RH threshold of approximately 90% for each of the stations in each of the cloud coverage scenarios when the observer LCBH was compared with the radiosonde data. For the NWP data, the RH thresholds varied between 58% and 66%. This is consistent with the model’s dry bias in the lowest 2000 m that the authors found when comparing the NWP RH data to the radiosonde. The authors found that the τ thresholds obtained were significantly lower than the minimum generally accepted for visible cloud, and suggest that this is also related to the dry bias in the model.
Using these RH and τ thresholds, the authors performed a verification analysis against the manual observations, calculating the bias score, probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI). The results showed that the LCBH derived from the radiosonde RH threshold was the most accurate. The LCBH derived from the NWP RH threshold tended to over-forecast the incidence of low cloud, with a slightly higher POD and much higher FAR, resulting in a poorer CSI. The LCBH derived from the NWP τ threshold was less conservative compared with the NWP RH method, but performed worse in each metric for each station than the radiosonde RH method.

A comparison between stations shows that Davis consistently had the highest bias scores and FAR values, and consequently lowest CSI scores. It is suggested that this may be due to the location of the model grid point used, which is 8 km inland from the Davis meteorological observations site.

The current AMPS cloud base product is generated during the post-processing stage using the Stoelinga-Warner algorithm (Stoelinga and Warner 1999). This algorithm is based on the extinction of light from the surface upward through the atmosphere due to water content. The equation for the extinction of light is given by:

\[
\frac{I(z)}{I_0} = \exp\left[-\int_0^z \beta(z) \, dz\right],
\]

(1)

where \(I_0\) and \(I(z)\) are the luminance at the surface and at height \(z\) respectively, and \(\beta(z)\) is the extinction coefficient.

Using a threshold of 2% of light remaining this becomes
\[-\ln(0.02) = \int_0^{clg} \beta(z) \, dz,\]  

where $clg$ is the height of the cloud base, or ceiling. The algorithm calculates the amount of light remaining at each vertical level of the model until the extinction threshold is reached. The height at that vertical level is determined to be the cloud base height.

$\beta(z)$ is calculated as the sum of the contributions from each of the hydrometeor species (cloud liquid water, rain, cloud ice and snow). The coefficients for each of these are obtained from previous observational studies, all conducted in the North American mid-latitudes.

Three options for this calculation are available in this algorithm. The Reisner scheme (Reisner et al. 1998; Stoelinga and Warner 1999) is mixed-phase and includes all four hydrometeor species. Although the most accurate method, it is computationally expensive. The Dudhia scheme (Dudhia 1989; Stoelinga and Warner 1999), also referred to as the 0°C scheme, is a single phase scheme which uses a temperature threshold of 0°C to determine the phase of water in the cloud and precipitation. This has the advantage of computational simplicity at the expense of accuracy. The Bocchieri scheme uses only a single phase for each of cloud water and precipitation but includes a combination using cloud liquid water and frozen precipitation. It determines the hydrometeor species mix using a temperature threshold of -10°C for the cloud phase and the statistical method used by Bocchieri (1980) to determine the probability of frozen precipitation.
Stoelinga and Warner (1999) found that the mixed phase Reisner scheme was more accurate in the cloud base prediction of an East Coast winter storm compared with the single phase Dudhia scheme, however was also more computationally expensive. Using the Bocchieri assumption gave results close in accuracy to the mixed phase scheme while being computationally more economical.

To summarize, the available literature on Antarctic clouds is relatively limited (Bromwich et al. 2012, Lachlan-Cope 2010) however in recent years there have been significant efforts to remedy this. High resolution satellite data have been used to compile cloud climatologies (e.g., Adhikari et al. 2012, Bromwich et al. 2012, Verlinden et al. 2011). Adhikari et al. (2012) found that low cloud was prevalent in coastal areas during the summer. Since most stations in Antarctica are on the coast, and with the majority of aviation activity taking place during the summer, it is important to be able to represent low cloud accurately in forecasting models. In the past few years there have been a number of short- and long-term field deployments of ground-based instrumentation (e.g., Gorodetskaya et al. 2015, Lawson and Gettelman 2014, Scott and Lubin 2014), which will increase the understanding of Antarctic clouds.

One of the key findings of these recent studies has been the higher than expected incidence of mixed-phase clouds, even with very cold surface air temperatures (Gorodetskaya et al. 2015, Lawson and Gettelman 2014). This may have significant implications for the microphysics schemes used to calculate cloud and moisture quantities in numerical models, and subsequent calculations of cloud. Under the Dudhia scheme of the Stoelinga-Warner algorithm used to generate the AMPS cloud base
product, a simple 0°C temperature threshold is used to discard certain hydrometeors, thereby disallowing mixed-phase cloud. Provided that the model hydrometeor quantities are accurate, this would introduce significant errors into the calculation compared with the Reisner mixed-phase scheme.

Forecasters often find that numerical models perform poorly when predicting cloud. This is supported by the findings of Fogt and Bromwich (2008). Operationally, many Antarctic forecasters use a relative humidity threshold as a proxy for cloud. Inoue et al. (2015) investigated the relationship between relative humidity and low cloud base height for the three Australian stations in East Antarctica using the Australian Bureau of Meteorology’s Polar-LAPS model. The optimal RH thresholds they calculated ranged from 58-66% which is reasonably close to the 70% value commonly used as a rule of thumb at the Australian stations. The verification analysis showed that a model-derived RH threshold could be used as a reasonable predictor of low cloud base height.
Chapter 3: Methods

3.1 Part I: The Stoelinga-Warner Algorithm

The first part of this study investigated the use of the Stoelinga-Warner algorithm (Stoelinga and Warner 1999). The previously described Davis region Prydz Bay Low event from 22 January 2013 (Figs. 5, 6, 7) was selected as a case study. From my own experience, AMPS typically performs well with prediction of the low level winds and RH associated with this feature, however usually performs poorly with prediction of the associated low cloud. The performance of the AMPS cloudbase and RH products in this case are typical for this scenario.

The cloud base product in AMPS is currently generated using the Dudhia scheme in the Stoelinga-Warner algorithm (Kevin Manning, personal communication, 2013). This scheme assumes single phase cloud based on a simple temperature threshold. However studies have shown a significant incidence of mixed-phase cloud in Antarctica (e.g., Gorodetskaya et al. 2015, Lawson and Gettelman 2014). This suggests that the mixed-phase Reisner scheme would produce an improved forecast.

For the case study, the algorithm was reproduced using AMPS raw data from Grid 2 (Fig. 8), which covers continental Antarctica and has a horizontal resolution of 10 km (AMPS 2015c).
Cloud base products using both the Dudhia and Reisner schemes were generated. Resulting images were compared by visual inspection with each other, a MODIS satellite image, and the archived 1000 ft RH field. Since this study was not concerned with computational expense, comparisons were made with the Reisner scheme only and the Bocchieri scheme was not tested. The sensitivity of the Reisner scheme was also tested, using light extinction thresholds set at various levels from the original 2% remaining light up to 90%.

3.2 Part II: Relative Humidity Thresholds

The second part of the study investigated the association between selected relative humidity thresholds from radiosonde ascents and observer estimated cloud base height for Davis, McMurdo and Halley stations. Lidar estimated cloud base heights for Halley were also compared with the radiosonde and observer estimates.

While working at Davis, I found that a relative humidity threshold of 70% provided a useful rule of thumb for cloud base. Inoue et al. (2015) calculated an optimal threshold of 64-66% at Davis, a reasonably close agreement. This part of the study will test that threshold for Davis and the other two stations to see if it is more widely applicable. A robust result would enable model predictions of cloud based on a relative humidity threshold. For this exercise, RH thresholds of 50% to 90% at 10% intervals were tested.

Surface synoptic observations in WMO FM-12 SYNOP format were obtained from the OGIMET website (http://www.ogimet.com/home.phtml.en). This format
describes the lowest cloud base observed according to height range groupings (Table 1). The lowest groupings were combined into a single group of cloud base below 300m (~1000ft) due to the low resolution of the radiosonde data and for simplicity of calculation.

Twice-daily radiosonde data were obtained from the University of Wyoming website (http://weather.uwyo.edu/upperair/sounding.html) in text format for Davis and McMurdo. The vertical resolution of these data was highly variable and the distance between consecutive data points would sometimes encompass more than one SYNOP category, particularly in the lower levels of the atmosphere. Radiosonde data for Halley were provided by Steve Colwell of the British Antarctic Survey and were of significantly higher vertical resolution, but were only available once per day. The radiosonde observations were classified into the same height range groupings according to the height at which they first reached each RH threshold.

A frequency analysis of each pair of observations was conducted for the period December 2011 to January 2012. This is the height of the summer operational season, during which more observations are available. Longer daylight hours also increase the accuracy of the synoptic observations of clouds (Bromwich et al. 2012). The four months either side of this period, October and November 2011 and February and March 2012, were also analyzed for comparison. These months form part of the summer operational season and still see considerable activity at the three stations.
3.3. Part III: AMPS Model Cloud Base Height Predictions from Relative Humidity Thresholds

The third part of the study investigated the ability of the AMPS model RH forecast to predict cloud base height. Cloud base heights at each of the three locations were generated from AMPS model RH data and compared to the synoptic observations. The study was conducted for three seasons (2009-10, 2010-11 and 2011-12) for the peak summer period (December and January) and the shoulder period (October, November, February and March).

AMPS cloud base heights were calculated using RH thresholds at the model grid point closest to each station. The RH thresholds used were from 60% to 90% in 5% increments. The 12 UTC model run for AMPS domain 2 with 10 km resolution was used, with values calculated every three hours for forecast hours 12 to 33, giving three-hourly predictions for the following day from 00 UTC to 21 UTC. AMPS predictions were classified according to the SYNOP height range groupings (Table 1) and synchronized with the available SYNOP observations.

Numerous verification statistics were calculated for each station at each RH threshold. These included the root mean square error (RMSE), bias, and the Gerrity Score (Gerrity 1992), a multi-category forecast skill score. The statistical analysis was repeated for observations filtered for the condition of broken or more cloud cover (≥ 5 oktas). This threshold was selected based on the International Civil Aviation Organization (ICAO) definition of ceiling as being the “height above the ground or water
of the base of the lowest layer of cloud below 6000 metres (20,000 feet) covering more than half the sky ” (International Civil Aviation Organization 2005).

The Gerrity Score (GS) is part of the Gandin-Murphy (Gandin and Murphy 1992) family of weighted equitable skill scores. These are considered more suitable for categorical or event-type forecasts with more than two categories and where the likelihoods of different events are not equally distributed (Gandin and Murphy 1992, Livezey 2011). They use a scoring matrix derived from the distribution of the observations to weight the forecast outcomes, with neighboring categories given a higher weighting than those more distant.

The Gerrity Score for a variable with $K$ categories is defined as:

$$ GS = \sum_{i=1}^{K} \sum_{j=1}^{K} p_{i,j} s_{i,j} $$  \hspace{1cm} (3)

where $p_{i,j}$ are the elements of the $K \times K$ contingency table of forecast vs observed events as a proportion of total events. The scoring matrix elements, $s_{i,j}$, are given by:

$$ s_{i,j} = \frac{1}{K-1} \left[ \sum_{r=1}^{i-1} a_r^{-1} - (j-i) + \sum_{r=j}^{K-1} a_r \right] = s_{j|i} $$  \hspace{1cm} (4)

where $1 \leq i \leq j \leq K$, the summations are 0 when the upper index is less than the lower index, and
\[ a_i = \frac{1 - \sum_{r=1}^{i} p_r}{\sum_{r=1}^{i} p_r} \]  

\( p_r \) is the proportion of observed events to total events and represents the climatological average.

The Gerrity Score ranges from -1 to 1, with 0 indicating no skill and 1 being a perfect score.

Binary event verification (Mason 2011) was also conducted for low cloud events. A low cloud event was defined as cloud below 600 m (~2000 ft), corresponding to the boundary between cloud height categories 4 and 5 (Table 1). This level was selected based on my own experience of critical cloud height thresholds at Davis and more broadly as a forecaster in Australia. Effectively, categories 3 and 4 were merged into a single “low cloud” category, and categories 5 to 9 were merged into a single “not low cloud” category. Generating the binary event results allowed for forecast verification metrics using a 2 x 2 contingency table (Table 2).

For this analysis, a “hit” occurred when low cloud was both observed and forecast. A “false alarm” occurred when low cloud was forecast but not observed. A “miss” occurred when low cloud was observed but not forecast. A “correct negative” occurred when low cloud was neither forecast nor observed.

Using these values, metrics such as the Proportion Correct (PC), Probability of Detection (POD), False Alarm Ratio (FAR) and Critical Success Index (CSI) were calculated.
The Proportion Correct is the number of correct forecasts divided by the total number of forecasts. It ranges from 0 to 1, with 1 being a perfect score.

$$PC = \frac{\text{hits + correct negatives}}{\text{total forecasts}}$$ (6)

The Probability of Detection reflects the likelihood that an observed event is forecast and is calculated by dividing the number of hits into the total number of events observed, i.e. both hits and misses. It ranges from 0 to 1, with 1 being a perfect score.

$$POD = \frac{\text{hits}}{\text{hits + misses}}$$ (7)

The False Alarm Ratio is a measure of the over-forecasting of an event and is calculated by dividing the number of false alarms into the total number of events forecast, i.e. both false alarms and hits. It ranges from 0 to 1, with 0 being a perfect score.

$$FAR = \frac{\text{false alarms}}{\text{false alarms + hits}}$$ (8)

The Critical Success Index, also known as the Threat Score (TS), combines the previous two metrics to consider how well events were forecast. It is calculated by dividing hits into the sum of the hits, misses and false alarms. It ranges from 0 to 1, with 1 being a perfect score.
Inoue et al. (2015) used different criteria in their analysis; defining an observed low cloud event as cloud of 6 or more oktas with a base below 2000 m, and a forecast low cloud event with cloud base height below 4000 m. The authors doubled the height threshold for the forecasts “in order to correctly produce negative bias values where the predicted cloud is close to the 2000-m low-cloud cutoff value” (Inoue et al. 2015). Since the SYNOP categories only reach 2500 m this could not be exactly replicated, however a similar analysis was done using thresholds of 1000 m and 2000 m for the observed and forecast events respectively.

\[
CSI = \frac{\text{hits}}{\text{hits} + \text{misses} + \text{false alarms}}
\]
Chapter 4: Results and Discussion

4.1. Part I: The Stoelinga-Warner Algorithm

Figure 9 shows the forecast cloud base using the (a) Dudhia scheme and (b) Reisner scheme. Comparison of the cloud base forecast generated using the Dudhia scheme (Fig. 9a) with the corresponding archived AMPS operational product (Fig. 6) indicates an accurate reproduction of the post-processing algorithm.

Comparison of the cloud base forecast generated using the Reisner mixed phase scheme (Fig. 9b) with that of the Dudhia scheme (Fig. 9a) shows little difference. Comparison of these figures with the MODIS satellite image (Fig. 5) shows poor identification of cloud over the Amery Ice Shelf. Comparison with the low level relative humidity field (Fig. 7) shows that there is moisture in the model at those levels however it is not translating into cloud during post-processing.

The similarity between the two schemes suggests that the model data might be dominated by one water phase. Figure 10 shows the mixing ratio values for (a) cloud liquid water and (b) cloud ice at 1000 ft. Comparing the two cloud components shows that the model cloud over coastal areas is entirely ice cloud. As these are low clouds in a coastal Antarctic environment during summer, this is unlikely to reflect the physical reality in this environment, where mixed phase clouds have been frequently identified (Lachlan-Cope 2010). Additionally, this event occurred in late January when there is
typically open water along the coastline at Davis out to Prydz Bay. Images from the weekly station newsletter show that this was indeed the case for 2013 (Australian Antarctic Division 2013), with the annual Australia Day swimming event taking place a few days earlier. Figure 11 shows open water extending out beyond Gardner Island, which is 3 km from shore.

Ample open water was available and in a Prydz Bay Low event winds tend to have an onshore component for that section of the coast. This can be seen in the low-level wind barbs in Figure 7. So some moisture transport onto the coast would be expected, however this is not evident in the model cloud liquid water. This implies that there is a problem with the cloud microphysics scheme in its generation of cloud hydrometeor quantities.

Figure 12 shows the mixing ratio values for water vapor at 1000 ft. This shows much larger quantities of water vapor than cloud liquid water or cloud ice. Over the Davis coast the mixing ratio of cloud liquid water is less than 0.01 g/kg and cloud ice is up to 0.08 g/kg, compared with up to 3 g/kg of water vapor. Of the five mixing ratio quantities (including snow and rain) only water vapor is present in the model over the main part of the Amery Ice Shelf. This suggests that there is insufficient conversion of moisture to cloud ice or cloud liquid water compared with water vapor. The Stoelinga-Warner algorithm only includes the solid and liquid hydrometeors, so despite the large quantity of water vapor, it is not factored into the calculation of the cloud base height.

The Stoelinga-Warner algorithm also showed low sensitivity to changes in the visible light extinction threshold and extinction coefficients. Figure 13 shows the
predicted cloud base using the Reisner scheme and extinction threshold of (a) 2% and (b) 90%. It can be seen that there is only a modest expansion of the clouded area, with much of the Amery Ice Shelf remaining cloud-free.

The low sensitivity of the algorithm to the extinction threshold is interesting. The algorithm might be very sensitive to the first layer of moisture it encounters.

4.2. Part II: Relative Humidity Thresholds

Comparison of the radiosonde relative humidity data with synoptic observations shows that Davis demonstrates a reasonably good correspondence between relative humidity and cloud base at 70% RH over the peak summer period of December 2011 – January 2012 (Fig. 14), with most of the radiosonde-derived cloud bases within one category of the synoptic observation. The RMSE was 1.66 and the bias was 0.08 for this threshold. This concurs with local forecaster experience. There were some notable discrepancies, in particular the eight observations of category 5 cloud (base height 600 - 1000 m) for which the radiosonde relative humidity did not meet the 70% threshold below 2500 m.

The best relationship during the shoulder period of October to November 2011 and February to March 2012 was also at a threshold of 70% (Fig. 15), although the correspondence was weaker compared with the peak summer period, with the RMSE at 1.99 and bias of 0.35. Again there were a large number of missed category 5 events.

In contrast, McMurdo shows no strong relationship at any of the relative humidity thresholds tested. Figure 16 shows the frequency of each pair of observations at a
threshold of 70%. This value showed the strongest correspondence with RMSE of 2.13 and bias of 0.15. However there is considerable scatter of the observation pairs, so it does not seem to be a reliable indicator of cloud base. Like Davis, the peak summer period showed a stronger relationship between the relative humidity thresholds and cloud base than the shoulder period.

A significant issue for this analysis was the low vertical resolution of both the synoptic observations and the radiosonde data. While the use of the synoptic height range groupings simplified the analysis, it reduced the accuracy and made it impossible to calculate errors in terms of actual height. The radiosonde data were not consistent in their vertical resolution and on occasion there were large gaps in data. This typically occurred nearer the surface, and may have affected the cloud base height value determined for these instances.

The British station of Halley, on the Brunt Ice Shelf in East Antarctica towards the eastern Weddell Sea, had high resolution radiosonde data available and a similar analysis was performed. At Halley, the optimal RH threshold in this analysis was 90%, although it was still comparatively poor. Figure 17 shows the frequency analysis for Halley for the RH threshold of 90% for the peak summer period, which had RMSE of 2.66 and bias of 1.5. Interestingly the shoulder period (Fig. 18) had better correspondence with RMSE of 2.08 and bias of 0.53.

Lidar ceilometer data were also available for Halley, which provides a useful objective comparison against the synoptic observations. Comparing the two, there looks to be reasonably good agreement for both the peak summer (Fig. 19) and shoulder (Fig.
periods. There are few instances in which the lidar detected a lower cloud base than that reported by the observer. This is not unexpected since the lidar is one of the tools which observers can use to estimate cloud base height. There are more instances of a lower cloud base height reported by the observer than detected by the lidar. This too is not surprising since the observer is required to observe the entire sky while the lidar employs a vertical beam and can only detect cloud directly overhead. However these factors may have had an impact on the statistics, with the summer period RMSE at 2.49 and bias at 1.34 and the shoulder period RMSE at 2.16 and bias at 0.88.

To test this, the analysis was repeated only with the instances where the synoptic observation reported broken cloud ($\geq 5$ oktas) to increase the likelihood that the cloud layer was overhead the lidar. The resulting frequency graphs (Figs. 21, 22) show a large reduction in the instances of a lower cloud base height reported by the observer than detected by the lidar. The RMSE showed large improvements for both periods, with summer value at 1.95 (2.49) and 1.51 (2.16) for the shoulder period. The bias did not change much for the summer, down slightly to 1.32 (1.34), while the shoulder period showed more improvement, down to 0.68 (0.88).

The comparison between the lidar and radiosonde RH thresholds again resulted in an optimal RH of 90%. This showed a significant improvement (Figs. 23, 24) against the radiosonde vs observer results with the summer RMSE at 2.05 (2.66) and bias at 0.19 (1.5) and shoulder RMSE at 1.30 (2.08) and bias at -0.14 (0.53). This is a strong result. If the radiosonde is assumed to have little horizontal movement in the low levels, then both the lidar and radiosonde are sampling the same vertical profile above the station and
would be expected to have higher accuracy than comparisons with the synoptic observation which take into account the much larger area of the visible sky dome.

Once again there is a stronger relationship in the shoulder period than during the summer. It is unclear why this is the case; however the result is robust given its consistency across all comparisons.

The variation across locations and the lack of a strong relationship at McMurdo means that using a relative humidity threshold as a proxy for cloud base may not be useful for areal coverage. Therefore it may not be useful as an alternative to the Stoelinga-Warner algorithm for the production of cloud base forecasts in AMPS. However it may still be useful at some locations as a local forecasting rule of thumb.

It is important to note that the cloud base height reported on the synoptic observation may not be fully independent of the radiosonde data. At Davis and other Australian Antarctic stations the duty observer is responsible for both reporting the synoptic observation and monitoring the radiosonde data as they are transmitted in real time. Thus the observer can be influenced by the radiosonde data when estimating cloud base height. It is not uncommon for observers to use the radiosonde data as a guide to the heights of the various cloud layers.
3.3. Part III: AMPS Model Cloud Base Height Predictions from Relative Humidity Thresholds

Three-hourly cloud base height predictions were generated from AMPS RH data over the October to March period for three seasons (2009-10, 2010-11 and 2011-12) with thresholds from 60% to 90% at 5% intervals. These were verified against synoptic observations.

Figure 25 shows the root mean square error (RMSE) for each station according to RH threshold for the summer period. Each station noticeably improves (decreased RMSE) when restricted to situations with broken cloud cover, which is not surprising. For Davis and McMurdo the values are generally worse than the comparative results from the radiosonde analysis of the previous section, with the best results for each station being 1.97 and 2.21 compared with 1.66 and 2.13 respectively. The lowest value for Halley was 2.43, compared with 2.66 (radiosonde vs observer), 1.95 (lidar vs observer) and 2.05 (radiosonde vs lidar). This suggests a less reliable relationship between the AMPS RH threshold and observed low cloud. As expected, Davis has the lowest RMSE overall, however this occurs at RH thresholds below 70%. Likewise McMurdo performs best at lower RH thresholds. Halley performs best at higher RH thresholds, but even here it is lower than expected based on the results of the previous section.

The RMSE results for the shoulder period can be seen in Figure 26. The values are higher compared to the summer period, by about 0.4 for Davis and Halley, and 0.15 for McMurdo. Davis and McMurdo have the best performance at around 65% RH, which is close to the 70% RH threshold found in the comparison between the radiosonde RH
and the observations. The optimal forecast RH threshold for Halley remains at around 80%, lower than the 90% RH threshold obtained from the radiosonde.

The similar values for the radiosonde and AMPS RH thresholds suggest that the AMPS RH field is reasonably accurate. This is consistent with the near zero bias in the AMPS low level RH field found by Bromwich et al. (2013). They noted a relatively large scatter in the RH field (compared with other variables), which may have contributed to the RMSE values obtained here.

Figures 27 and 28 show the bias results. These results are more in line with the radiosonde-derived thresholds, with optimal (near zero) results for Davis at around 70% RH and Halley at 85% RH for both periods. McMurdo has the best results at around 70% RH in summer and 65% RH in the shoulder period. Interestingly there was little difference between the full set of data and the filtered set for each station. This is surprising since the instances of few or scattered low cloud might be expected to skew the bias values higher.

Figures 29, 30 and 31 show the Gerrity Skill Scores for each station. This is a weighted equitable skill score which provides a measure of forecast skill against the climatology derived from the observations. Scores range from -1 to 1, with 0 being no skill and 1 a perfect score. All of the results here are above 0, indicating that at least some skill is evident.

The highest skill (0.43) is shown at Davis at 80% RH in the summer period, dropping to 0.31 at 75% RH in the shoulder period. Although at a higher RH threshold than the radiosonde results, this is consistent with the general result of higher reliability.
in the relationship between RH and cloud base height at Davis in summer. The scores from McMurdo show higher skill in the shoulder period (0.40 at 65% RH) than the summer (0.35 at 75% RH), again consistent with previous results. At Halley the results are a little worse, scoring 0.28 in both the summer (90% RH) and shoulder (85%) seasons.

The second part of the verification involved testing the prediction of low cloud events, with the threshold for low cloud set at 600 m (~2000 ft). The Critical Success Index (CSI), which combines both the Probability of Detection (POD) and the False Alarm Ratio (FAR), is shown in Figures 32 and 33. All three stations have a CSI score around 0.5 in the summer period, with Davis slightly higher at 0.54. CSI ranges from 0 to a perfect score of 1. CSI falls in the shoulder period for each station to 0.42 at both Davis and Halley, and 0.26 at McMurdo. This fall is largest for McMurdo, which is unexpected given previous results which suggest that the shoulder period shows a stronger relationship between RH threshold and low cloud base height. This is due to a large increase in the FAR from 0.48 in the summer to 0.70 in the shoulder period at McMurdo compared with the other stations, which remained around 0.5 for both periods.

Inoue et al. (2015) performed a similar verification on the thresholds they obtained through their RH analysis, using a low cloud threshold of 2000 m and the PolarLAPS model. The CSI values they obtained averaged 0.47, lower than the 0.5 obtained here for the summer period, but higher than the shoulder period values. Davis, the station in common, performed worst in Inoue et al. (2015) with values slightly below 0.4 for
NWP, compared with 0.54 and 0.42 for the summer and the shoulder season respectively in this analysis.

Comparison of the Proportion Correct (PC) in the summer period (Fig. 34) shows Halley peaking at 90%, Davis at 80% and McMurdo at 75% RH, which is consistent with previous findings. Davis performs best with a peak value around 0.65 while McMurdo and Halley are just above 0.60. The shoulder season (Fig. 35) brings some interesting results with each station having slightly better results for the full dataset (i.e., any amount of reported low cloud) compared with the filtered dataset (more than 50% cloud cover). The PC values for the full (filtered) dataset at Davis, Halley and McMurdo were 0.69 (0.64), 0.68 (0.61) and 0.81 (0.78) respectively. This is mostly due to large increases in the number of correct negatives recorded and may reflect a relatively high frequency of clear days. This is also a major factor in the PC peaking at 90% for every station in the shoulder period. The large increase in PC at McMurdo from summer (0.61) to the shoulder period (0.78) reflects the large decrease in the FAR as previously noted for the CSI.

The verification metrics had mixed results in terms of identifying an optimal RH threshold for each station. Since verification metrics each have their own strengths and weaknesses (Mason 2011, The Centre for Australian Weather and Climate Research 2015) this is not unexpected.

In particular the PC performs poorly for rare events due to a high number of correct negatives (Mason 2011). In this case, although low cloud is not a rare event, the climatological probability is not always near 0.5 and can be different across seasons. For
example, using the filtered set of McMurdo observations, the summer period showed 437 low cloud events in 936 observations giving a probability of 0.47. In contrast there were 337 low cloud events out of 1550 observations giving a probability of 0.22 in the shoulder season.

The CSI removes the influence of correct negatives which is useful for rare events; however it remains an inequitable scoring method as it assigns equal weighting to events and non-events despite the unequal climatological probability (Gandin and Murphy 1992, Mason 2011).

On the other hand, the Gerrity Skill Score is designed to be equitable (Gerrity 1992), using the observation sample distribution as climatology to provide the weightings for the scoring matrix. It is also multi-category which is useful in this instance since there are multiple cloud base thresholds, the significance of which depends on the application.

Broadly, the results from the verification analysis supported those of the previous section, indicating that Davis and McMurdo would use a lower RH threshold than Halley. The choice of RH threshold for each would depend on the operational requirements and the level of acceptable risk (misses vs false alarms).
Chapter 5: Conclusions

Investigation of the Stoelinga-Warner algorithm revealed that the under-prediction of cloud was due to insufficient cloud hydrometeor quantities in the model data. Further work is required on the model physics to investigate and improve this. AMPS currently uses the WSM 5-class scheme (AMPS 2015a). Since this is not generating mixed-phase cloud hydrometeors well, it is suggested that other schemes which are more directed towards mixed-phase clouds be tested.

It is also suggested that the post-processing algorithm be adjusted to use the Reisner mixed-phase scheme, as the literature indicates that mixed-phase clouds are not uncommon, even over the Antarctic plateau.

Tests of cloud base heights derived from radiosonde RH thresholds against synoptic observations showed that the relationship between RH and cloud varied by location. A good relationship was evident at Davis for a threshold around 70% RH, which agrees with the findings of Inoue et al. (2015). McMurdo showed a weaker relationship although it also had an optimal RH threshold around 70%, despite the local forecasters’ use of 90% as a rule of thumb (Art Cayette, personal communication, 11 June 2013). Halley was best at 90%, which was consistent in comparisons with the lidar
observations. This diverges with Inoue et al. (2015), who found little difference between stations in the optimal RH threshold.

The reason for the variation is not clear. All of the stations are located on the coast, with open water nearby during the summer. The other Australian stations of Casey and Mawson are also located on the coast. The results for these stations from Inoue et al. (2015) showed a similar RH threshold to McMurdo, with values ranging from 58% to 66% RH.

Of the five stations studied, only Halley requires a significantly different RH threshold. One possible source of the variation is Halley’s location on an ice shelf, permanently surrounded by snow and ice and with no significant topographical features. In contrast, the other four stations are all on land with snow- and ice-free areas of exposed dark rock and soil during the summer, as well as local topography influencing weather regimes. Further studies of other sites (in particular, on ice shelves and the Antarctic plateau) are needed to test the influence of both the exposed rock and the topography.

Higher resolution synoptic and radiosonde data would enable a more robust analysis. However the results obtained here for Davis compared well with those obtained by Inoue et al. (2015), which were based on high resolution data.

The AMPS model RH was tested against the synoptic observations. Reasonable skill was demonstrated and the results were generally consistent with the radiosonde comparison. This shows that AMPS model RH can be useful as a local aid to forecasting low cloud.
While RH thresholds have been shown to be useful for predicting cloud base height locally, a common threshold was not found. Further studies of additional sites are needed before a generalized cloud base product derived from RH can be developed. An improvement to the model physics to more accurately predict the cloud hydrometeors present is also needed.
References


43
Table 1: WMO SYNOP cloud base height code. From http://weather.unisys.com/wxp/Appendices/Formats/SYNOP.html.

<table>
<thead>
<tr>
<th>SYNOP code</th>
<th>Cloud base of lowest cloud seen (meters above ground)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0 to 50 m</td>
</tr>
<tr>
<td>1</td>
<td>50 to 100 m</td>
</tr>
<tr>
<td>2</td>
<td>100 to 200 m</td>
</tr>
<tr>
<td>3</td>
<td>200 to 300 m</td>
</tr>
<tr>
<td>4</td>
<td>300 to 600 m</td>
</tr>
<tr>
<td>5</td>
<td>600 to 1000 m</td>
</tr>
<tr>
<td>6</td>
<td>1000 to 1500 m</td>
</tr>
<tr>
<td>7</td>
<td>1500 to 2000 m</td>
</tr>
<tr>
<td>8</td>
<td>2000 to 2500 m</td>
</tr>
<tr>
<td>9</td>
<td>above 2500 m</td>
</tr>
</tbody>
</table>
Table 2: 2x2 contingency table.

<table>
<thead>
<tr>
<th>Event forecast</th>
<th>Event observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>A (hits)</td>
</tr>
<tr>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>C (misses)</td>
</tr>
<tr>
<td>Yes</td>
<td>B (false alarms)</td>
</tr>
<tr>
<td>No</td>
<td>D (correct negatives)</td>
</tr>
</tbody>
</table>
Figure 1: Davis Ski Landing Area (Photo: Karen Pon)
Figure 2: Davis Ski Landing Area (Photo: Bob Heath)
Figure 3: Map of Antarctica. Stations used in this study circled. (http://www.mapsofworld.com/antarctica/)
Figure 4: Map of Davis area. Courtesy of the Australian Antarctic Division (2011).
Figure 5: MODIS visible channel (0.62-0.67 μm) satellite image for the Davis area at 0415 UTC on 22 January 2013. Distinct areas of low cloud are marked.
Figure 6: Archived AMPS 27 hour forecast valid at 03 UTC on 22 January 2013 showing cloud base height.
Figure 7: Archived AMPS 27 hour forecast valid at 03 UTC on 22 January 2013 showing 1000 ft relative humidity and horizontal wind fields.
Figure 8: Domain of AMPS Grid 2 (10 km continental grid). The color contours represent terrain with green being low elevation and red being higher elevation. (http://www2.mmm.ucar.edu/rt/amps/information/configuration/maps.html).
Figure 9: 27-hour cloud base (ft) forecast for the Davis area valid for 03 UTC 22 January 2013, using AMPS initialized at 00 UTC 21 January 2013, generated by the Stoelinga-Warner algorithm using the (a) Dudhia and (b) Reisner schemes.
Figure 10: 27-hour forecast for the Davis area valid for 03 UTC 22 January 2013, using AMPS initialized at 00 UTC 21 January 2013, showing 1000 ft mixing ratio (g/kg) of (a) cloud liquid water and (b) cloud ice.
Figure 11: Australia Day swimming in open water at Davis station on 20 January 2013. (Photo: Bill de Bruyn)
Figure 12: 27-hour forecast for the Davis area valid for 03 UTC 22 January 2013, using AMPS initialized at 00 UTC 21 January 2013, showing 1000 ft mixing ratio of water vapor (g/kg).
Figure 13: 27-hour cloud base (ft) forecast for the Davis area valid for 03 UTC 22 January 2013, using AMPS initialized at 00 UTC 21 January 2013, generated by the Stoelinga-Warner algorithm using the Reisner scheme and light extinction threshold of (a) 2% and (b) 90% remaining light.
Figure 14: Frequency analysis of radiosonde relative humidity threshold of 70% vs synoptic cloud base height observation at Davis for December 2011 and January 2012.
Figure 15: Frequency analysis of radiosonde relative humidity threshold of 70% vs synoptic cloud base height observation at Davis for October to November 2011 and February to March 2012.
Figure 16: Frequency analysis of radiosonde relative humidity threshold of 70% vs synoptic cloud base height observation at McMurdo for December 2011 and January 2012.
Figure 17: Frequency analysis of radiosonde relative humidity threshold of 90% vs synoptic cloud base height observation at Halley for December 2011 and January 2012.
Figure 18: Frequency analysis of radiosonde relative humidity threshold of 90% vs synoptic cloud base height observation at Halley for October to November 2011 and February to March 2012.
Figure 19: Frequency analysis of lidar vs synoptic cloud base height observation at Halley for December 2011 and January 2012.
Figure 20: Frequency analysis of lidar vs synoptic cloud base height observation at Halley for October to November 2011 and February to March 2012.
Figure 21: Frequency analysis of lidar vs synoptic cloud base height observation at Halley for December 2011 and January 2012 with sky coverage ≥ 5 oktas.
Figure 22: Frequency analysis of lidar vs synoptic cloud base height observation at Halley for October to November 2011 and February to March 2012 with sky coverage ≥ 5 oktas.
Figure 23: Frequency analysis of radiosonde relative humidity threshold of 90% vs lidar cloud base height observation at Halley for December 2011 and January 2012.
Figure 24: Frequency analysis of radiosonde relative humidity threshold of 90% vs lidar cloud base height observation at Halley for October to November 2011 and February to March 2012.
Figure 25: Comparison of RMSE of AMPS cloud base height prediction by RH threshold at each station for the summer period.
Figure 26: Comparison of RMSE of AMPS cloud base height prediction by RH threshold at each station for the shoulder period.
Figure 27: Comparison of bias of AMPS cloud base height prediction by RH threshold at each station for the summer period.
Figure 28: Comparison of bias of AMPS cloud base height prediction by RH threshold at each station for the shoulder period.
Figure 29: Gerrity Skill Score for Davis.
Figure 30: Gerrity Skill Score for McMurdo.
Figure 31: Gerrity Skill Score for Halley.
Figure 32: Critical Success Index for the summer period.
Figure 33: Critical Success Index for the shoulder period.
Figure 34: Proportion of correct forecasts over the summer period.
Figure 35: Proportion of correct forecasts over the shoulder period.