An Assessment of Numerical Weather Prediction–Derived Low-Cloud-Base Height Forecasts

MANA INOUE

Institute for Marine and Antarctic Studies, and Antarctic Climate & Ecosystems Cooperative Research Centre, University of Tasmania, Hobart, Tasmania, Australia

ALEXANDER D. FRASER

Antarctic Climate & Ecosystems Cooperative Research Centre, University of Tasmania, Hobart, Tasmania, Australia, and Institute of Low Temperature Science, Hokkaido University, Sapporo, Hokkaido, Japan

NEIL ADAMS AND SCOTT CARPENTIER

Australian Bureau of Meteorology, Hobart, Tasmania, Australia

HELEN E. PHILLIPS

Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania, Australia

(Manuscript received 13 May 2014, in final form 28 December 2014)

ABSTRACT

As demand for flight operations in Antarctica grows, accurate weather forecasting of cloud properties such as extent, cloud base, and cloud-top altitude becomes essential. The primary aims of this work are to ascertain relationships between numerical weather prediction (NWP) model output variables and surface-observed cloud properties and to develop low-cloud-base (<2000 m) height prediction algorithms for use across Antarctica to assist in low-cloud forecasting for aircraft operations. NWP output and radiosonde data are assessed against surface observations, and the relationship between the relative humidity RH profile and the height of the observed low-cloud base is investigated. The ability of NWP-derived RH and ice–water cloud optical depth profiles to represent the observed low-cloud conditions around each of the three Australian stations in East Antarctica is assessed. NWP-derived RH is drier than that reported by radiosonde from ground level up to ~2000 m. This trend reverses in the higher troposphere, and the largest positive difference is observed at ~10 000 m. A consequence is very low RH thresholds are needed for low-cloud-base height prediction using NWP RH profiles. RH and optical depth–based threshold techniques all show skill in reproducing the observed cloud-base height at all Australian Antarctic stations, but the radiosonde-derived RH technique is superior in all cases. This comparison of three low-cloud-base height retrieval techniques provides the first documented assessment of the relative efficacy of each technique in Antarctica.

1. Introduction

The demand for and importance of flight operations has expanded rapidly in recent years in Antarctica. Aircraft operations are diverse and include transport between stations, field camps, and ships; a wide range of research/field work; and tourism to the continent. Flight operations require accurate weather forecasts to ensure safety. Precise weather forecasting is also essential to facilitate well-organized and successful field work, while minimizing extra cost due to failure/postponement of survey work, and for ground transportation. However, weather forecasting in Antarctica is less accurate than in other places in the world, and fewer forecasting studies have been conducted in this area because of insufficient in situ observations, and the

Corresponding author address: Mana Inoue, Institute for Marine and Antarctic Studies, University of Tasmania/Antarctic Climate & Ecosystems Cooperative Research Centre, Private Bag 129, Hobart, TAS 7001, Australia.
E-mail: mana.inoue@utas.edu.au

DOI: 10.1175/WAF-D-14-00052.1

harsh, isolated environment (Turner et al. 2000; Adams 2002; Pendlebury et al. 2003).

Flight operations across Antarctica are generally performed under visual flight rules (VFR). Minimum requirements for descent procedures into unknown landing areas are a cloud base above 1500 ft (~450 m), where the cloud cover exceeds 50%, and >8-km horizontal visibility (Australian Bureau of Meteorology 2014, chapter 6, section 2.5.4, p. 57). The Antarctic is one of the most data-sparse areas on the planet for quality meteorological observations (Adams et al. 1999) with visual observations of cloud limited to a small number of staffed stations, predominantly around the Antarctic coast. The lack of widespread cloud cover and low-cloud-base height (hereinafter termed LCBH) observations, coupled with the cloud-base restrictions on flying, ensure that forecasting cloud properties in support of aviation remains a challenge.

This study uses balloon-launched radiosonde data and surface observations from the three Australian East Antarctic coastal stations of Mawson, Davis, and Casey, along with near-coincident numerical weather prediction (NWP) forecasts, in order to

(i) determine the accuracy of the NWP relative humidity RH profile by comparing NWP-derived RH (RH_{NWP}) to radiosonde-derived RH (RH_{sonde}),
(ii) investigate if RH profiles (RH_{sonde} and RH_{NWP}) can be used to determine LCBH, and
(iii) investigate if NWP-modeled cloud water and ice mixing ratios, after conversion to optical depth, can be used to determine LCBH.

The dataset and study period are discussed in section 2. An assessment of the NWP-modeled RH profile is presented in section 3. The efficacy of using RH_{sonde} and RH_{NWP} profiles to determine the LCBH is explored in section 4, with section 5 covering the possibility of using NWP-modeled cloud water and ice mixing ratios in order to determine the LCBH. Section 6 presents a discussion of all the LCBH retrieval techniques detailed here and suggests further research.

2. Datasets and study period

Three-hourly surface observation (provided by the Australian Bureau of Meteorology) data were recorded by trained observers at Casey, Davis, and Mawson. These data include low-cloud amount (oktas) and LCBH (feet; later converted to meters). The cloud-base observations are reported to a precision of 100 ft (~30.5 m) between 50 and 5000 ft (~1525 m), then every 1000 ft (~305 m) above. No effort was made to assess or reduce the bias of any particular observer.

Radiosonde observations were conducted at 0000 and 1200 UTC at Casey and Davis. At Mawson, radiosonde observations were conducted at 0000 UTC between November 2005 and March 2008, and at 1200 UTC between April 2008 and May 2010. The radiosondes were instrumented to measure RH (percent), temperature (kelvins), pressure (hectopascal), and wind speed and direction (meters per second). The radiosonde sampled at 0.5 Hz with enough hydrogen to provide an average ascent rate of 5 m s^{-1}. This gave an approximate sample height spacing of 10 m. Linear interpolation was then used to subsample to a regular 50-m vertical grid. Before the balloon flight, the radiosondes are stored inside, where different temperature and RH may alter the response of the humidity probes. If the radiosonde is not allowed to equilibrate to the outside temperature prior to release, then RH measurements near the surface may be affected. An RH spike of up to ~20% was often observed in the lower 50 m of each RH profile, particularly at Mawson Station. For this reason, the lowest measurement (i.e., the bottom 50 m) was discarded from each radiosonde RH profile.

The predominant radiosonde used throughout the study was the Vaisala RS92, with the RS80 radiosondes phased out at all three stations in early to mid-February 2006. The Vaisala RS80 radiosonde humidity measurements have a time constant of around 37 s at −40°C (Miloshevich et al. 2004) compared with a negligible (less than 2 s) time constant for the RS92 radiosonde in the lower atmosphere (Fig. 1; Miloshevich et al. 2009). The mean ascent rate of the radiosondes is on the order of 5 m s^{-1}, meaning RH_{sonde} measurements may be biased toward surface conditions for the lower 185 m of some RS80 profiles (less than 10 m for RS92 soundings). Miloshevich et al. (2009) provides a comprehensive correction algorithm for the RS92 humidity profiles, taking into account the effects of calibration errors, time lag (slow sensor response at low temperatures), and solar radiation errors. Tomasi et al. (2011) analyzed 4-yr worth of radiosonde data from Dome C on the East Antarctic plateau, and applied the RS80 (Miloshevich et al. 2004) and RS92 (Miloshevich et al. 2009) corrections prior to characterizing the temperature and moisture fields. RS80 corrections in the lower atmosphere were appreciable (around 16% of the measured RH); however, the corrections to the RS92 profiles were less significant. Also of note is the fact that Dome C is at an elevation of 3233 m and well inland of the coast, resulting in dry and cold conditions where the RH correction values are most significant. In the present study, only the lower 2000 m of the RH profiles were used for cloud detection and all three stations are on the coast and close to mean sea level. The Miloshevich et al. (2009) estimates

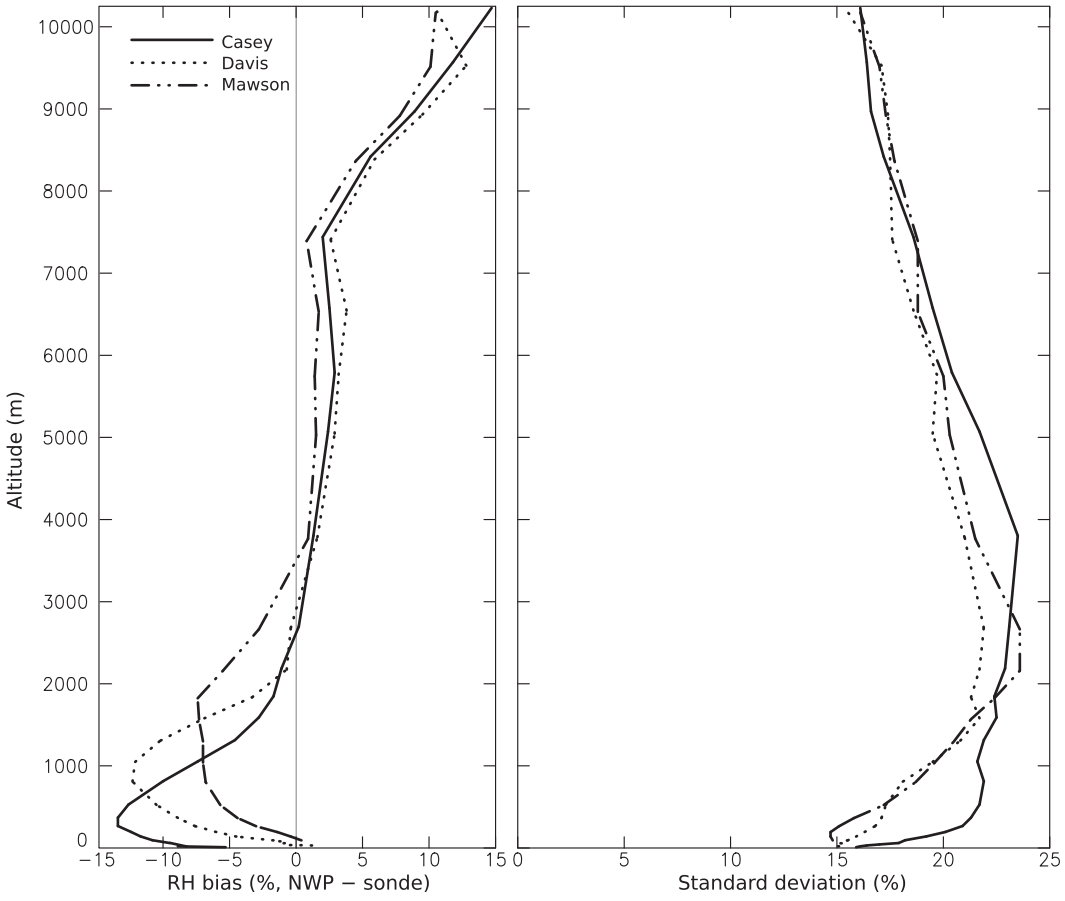

FIG. 1. (left) Bias and (right) std dev of $RH_{NWP} - RH_{sonde}$ during times of ≥6-oktas cloud coverage at Casey, Davis, and Mawson.

of the RS92 nighttime radiosonde dry bias under such conditions were around 2% of the measured RH value and around 5% of the measured values for clear-sky daytime ascents (Miloshevich et al. 2009). Furthermore, all radiosonde ascents considered in the present study were in overcast or near-overcast conditions where solar-radiation-induced errors would be small, so it is expected that the dry bias associated with the RS92 radiosonde would be well under 5% of the measured RH values.

The NWP data used in this study were from the Polar Limited Area Predictive System (polarLAPS; Adams and Powers 2007), which was developed by the Australian Bureau of Meteorology Research Centre (BMRC) based on the Limited Area Predictive System (LAPS) NWP forecasting model (Puri et al. 1998; Adams 2004). A polar stereographic version of LAPS, polarLAPS has a domain covering all of Antarctica and most of the Southern Ocean. The NWP model has a 27.5-km horizontal grid size, and 34 levels of variable vertical grid resolution ranging from sea level pressure (SLP) × 0.9988 (around 10 m above ground level) up to SLP × 0.05 (approximately 20 km). Data assimilation is not performed by polarLAPS, and is nested within the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS; National Weather Service Environmental Modeling Center 2011) model outputs. Provided by polarLAPS is 3-hourly output out to 120 h that is initialized at 0000 and 1200 UTC (Adams and Powers 2007). The 3-hourly forecast files from the +24-h time step through to the +33-h time step from each of the twice-daily forecast runs were composited to generate a contiguous 3-hourly dataset concurrent with the surface observations and radiosonde data. Relevant modeled fields include RH, cloud water mixing ratio (kilograms of water per kilogram of air), and cloud ice mixing ratio (kilograms of water equivalent per kilogram of air). NWP, surface, and radiosonde data are summarized in Table 1.

The study period was determined by overlapping time series of surface observations, radiosondes, and NWP outputs. Thus, the study period is from November 2005 to May 2010 for surface observations versus radiosondes, and from January 1998 to December 2008 for surface observations versus NWP outputs. To retain the most data possible, the study period was not restricted to

TABLE 1. Surface, radiosonde, and NWP datasets used in this study.

Obs type	Data	Unit
Surface	Amount of low cloud	oktas
	Height of low-cloud base	m
Radiosonde	RH	%
NWP	RH	%
	Cloud water mixing ratio	kg_{water}/kg_{air}
	Cloud ice mixing ratio	$kg_{water\ equivalent}/kg_{air}$

the mutual overlap between all three datasets. There is no significant temporal separation between the observations and NWP output. The spatial separation between each station and the nearest NWP grid points are given in Table 2. The NWP grid points are between 118 and 234 m higher than the respective surface station elevations. The consequences of this elevation difference are discussed in the next session.

3. An assessment of the RH_{NWP} profile

The first aim of this study is to determine how well RH_{NWP} matches the observed values obtained from radiosonde measurements. The RH_{sonde} profiles were resampled using averaging to match the NWP levels. Following this, the bias between RH_{sonde} and RH_{NWP} was determined for each radiosonde observation. Profiles of the mean bias and standard deviation over the entire study period are presented in Fig. 1 for each station.

Initially, we performed the analyses presented here using both RH measured over water RH_{water} (as observed directly by the radiosonde instrumentation and modeled by polarLAPS) and RH converted to RH measured over ice RH_{ice} (which is expected to more accurately represent the true RH conditions). This conversion to RH_{ice} frequently produced values well in excess of 100%. It is likely that the coastal sites selected here are not suitable for use with a simple conversion to RH_{ice}, particularly during summertime (i.e., when most flight operations are conducted) when sea ice extent is low. Nevertheless, such a conversion should be investigated if these techniques are to be used farther inland.

Inspection of the mean RH bias profile ($RH_{NWP} - RH_{sonde}$) shows that the lower-troposphere NWP forecast RH values are drier than observed across all three stations, for times of ≥6 oktas (Fig. 1). A dry bias of up to 15% is evident in RH_{NWP} from just above the surface to near 2000 m. The bias is evident at all three stations, although weaker in magnitude but vertically more extensive at Mawson. This dry bias may be partly attributable to the difference between the observed and modeled wind rose climatologies at each station (see Fig. 2). In the Antarctic coastal fringe, the near-surface wind climatology is strongly forced by local orography. The horizontal grid resolution of polarLAPS is 0.25° and it only approximates the steep coastal escarpment observed around Mawson, inland of Davis, and the orography of Law Dome to the southeast of Casey. The modeled wind climatology at Mawson most closely matches the station-based observations. Both the Casey and Davis modeled wind roses show wind from inland directions (i.e., dry air) too frequently. At Casey, station observations show a significant onshore flow from the northeast that is not reproduced in the polarLAPS climatology, and the NWP prevailing wind is east-to-southeasterly and directly off the high drier slopes of Law Dome. Similarly, at Davis the observed surface flow has a more significant flow from the north-northeast than is evident in the NWP climate, which shows a dominant easterly flow. This discrepancy in near-surface wind partitioning may directly account for the observed dry bias.

It is desirable to compare the NWP-modeled wind rose climatology to an observational climatology at various altitudes in the lower troposphere. Unfortunately, production of a quality observational wind rose climatology at altitude from radiosonde data is impossible because of the much lower volume of data collected by radiosondes. Nevertheless, we recognize that the surface wind climatology may not be representative of the climatology above the katabatic layer. Furthermore, the NWP grid point closest to each station is located inland of the station (e.g., the NWP grid point closest to Casey is ~12.3 km inland, at a mean altitude of ~234.0 m; see Table 2). Coastal locations generally experience more moisture from the ocean than inland locations, which are more strongly influenced by dry katabatic winds.

Across all stations, the bias in RH_{NWP} becomes positive (too moist) in the upper levels, with the largest

TABLE 2. Spatial separations between each of the stations and NWP grid points.

Station	Station location	NWP grid point	Separation (km)	Bearing (from station to NWP)
Casey	66.2825°S, 110.5232°E (42.3-m elevation)	66.2447°S, 110.78°E (234.0-m elevation)	12.25	70°
Davis	68.5743°S, 77.9672°E (23.2-m elevation)	68.6412°S, 78.0613°E (141.2-m elevation)	8.34	152°
Mawson	67.6017°S, 62.8753°E (16.0-m elevation)	67.5648°S, 62.9767°E (250.4-m elevation)	5.93	46°

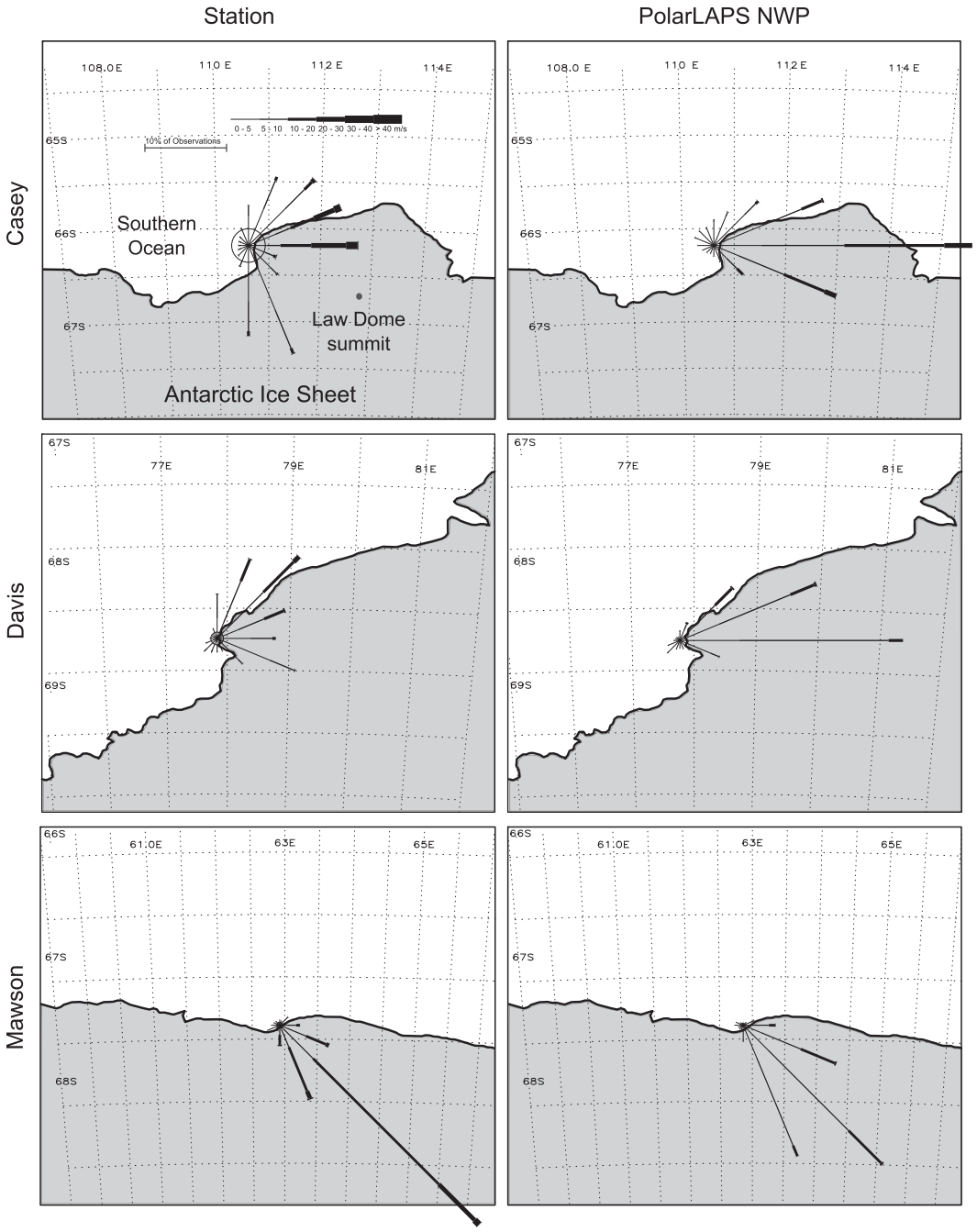

FIG. 2. Wind rose comparison of (left) station 10-m wind observations and (right) coincident near-surface (~9 m) polarLAPS wind forecasts for the period November 2005–May 2010 for (top) Casey, (middle) Davis, and (bottom) Mawson.

differences generally observed at 10 km. This trend is likely associated with generally high levels of moisture seen in the GFS model used to initialize polarLAPS (Luo and Krueger 2004).

The RH_{NWP} minus RH_{sonde} standard deviation reaches a maximum between 2000 and 4000 m, with mid-troposphere biases close to zero in the same layer. Biases are generally larger (negative) near the surface and larger (positive) in the upper troposphere–lower stratosphere but with lower standard deviations. This may be because the upper troposphere exhibits less variability, and near the surface the orographic influences act to constrain the flow and generate highly persistent flows. By contrast, the middle atmosphere has

more variability because of the influence and interaction of synoptic- and mesoscale events, which are far less predictable within NWP systems. Generally, polarLAPS exhibits a 3–6-h lag on large-scale synoptic systems that are associated with moisture (Adams and Powers 2007).

Further discrepancies between the radiosonde- and NWP-reported RH profiles may arise because of the differences in scale between the radiosonde (essentially a spot measurement) and one NWP grid cell (~27.5 km × 27.5 km). Interpretation of each profile must take this scale mismatch into account.

The following section focuses on identifying low-cloud events by thresholding profiles of RH_{sonde} and RH_{NWP}. The dry bias evident in the lower 2000 m of RH_{NWP} indicates that the RH_{NWP} thresholds will need to be relatively low (compared to RH_{sonde}). Nevertheless, it is expected that the NWP RH profiles will show some skill in predicting the LCBH, given that the standard deviation of the differences is smaller in the lower troposphere, compared to the middle troposphere.

4. Using RH_{sonde} and RH_{NWP} to determine LCBH

Because the aim of this study is to improve aviation weather forecasting, it focuses on times when large amounts of low cloud are present. Data analysis is performed for cases of both ≥6 oktas observed low cloud, in which the radiosonde would likely have passed through low cloud, and 8 oktas observed low cloud, where the sky was totally covered by low cloud, ensuring the radiosonde would encounter the observed cloud. The RH profiles at low- to midlatitudes generally show high values (upward of 90%) where clouds are located (Korolev and Isaac 2006). Relative humidity is a useful parameter for determining (in the case of radiosonde observations) or predicting (in the case of NWP) LCBH at these latitudes. Thus, the surface-observed LCBH ($LCBH_{SOb}$) was compared with both RH_{sonde} and RH_{NWP} profiles using thresholds for defining the LCBH, to determine the applicability of this approach in Antarctica. The second objective of this work is to investigate whether RH profiles (both sonde and NWP derived) can be used to determine LCBH. This is achieved by

- observing where the radiosonde RH profile exceeds a threshold value (i.e., finding $LCBH_{RH_{sonde}}$) and comparing this against the $LCBH_{SOb}$ and
- observing where the NWP-derived RH profile exceeds a threshold value (i.e., finding $LCBH_{RH_{NWP}}$) and comparing this against $LCBH_{SOb}$.

The analysis of RH_{sonde} profiles provided in section 3 highlights the limitations in identification of LCBH using RH profiles in coastal regions in Antarctica. Complexities associated with this are shown in Fig. 3, where four examples of RH_{sonde} profiles are shown with the observed cloud-base height marked as a horizontal line. In Fig. 3a the match between the radiosonde-deduced and observed cloud base is excellent, and the observed cloud base occurs at a high RH value. In Fig. 3b the RH_{sonde} profile does not show a cloud base at all in the lowest 2000 m, although a peak in RH at ~1750 m, close to the observed value, is suggestive of a cloud layer. In Fig. 3c no RH peak exists near the altitude of the observed cloud base, and RH values are around 50% near the observed cloud-base height. In Fig. 3d the observed LCBH occurs in very low RH, approximately 30%. A peak to 65% RH was observed at 2300 m and the next at 4300 m.

The good fit displayed in Fig. 3a is typical behavior at low- and midlatitudes, with a peak in low-level RH coincident with the observed cloud base. However, this behavior was not observed consistently in this study and the examples shown in Figs. 3c and 3d highlight the existence of low clouds in quite low RH environments.

To estimate $LCBH_{RH_{sonde}}$ and $LCBH_{RH_{NWP}}$, appropriate threshold values of RH need to be found. The long-term bias between $LCBH_{SOb}$ and $LCBH_{RH_{sonde}}$/$LCBH_{RH_{NWP}}$ can effectively be adjusted by varying the RH threshold. A higher RH threshold means that higher humidity is needed to produce a cloud in the sonde or NWP dataset, making the sonde or NWP minus observed cloud-base height more positive. Here, we use this ability to "tune" the threshold in order to find threshold values that give a long-term near-zero bias in cloud-base height. We vary the RH threshold from 40% to 100% in increments of 1% in order to find the value that gives a near-zero mean difference between $LCBH_{SOb}$ and $LCBH_{RH_{sonde}}$ or $LCBH_{RH_{NWP}}$. These values, and the mean bias obtained using these thresholds, are shown in Table 3. Also shown is the mean standard deviation of the difference between $LCBH_{SOb}$ and $LCBH_{RH_{sonde}}$ or $LCBH_{RH_{NWP}}$, and the percentage of observed clouds that are able to be reproduced (i.e., the RH threshold is exceeded in the lower 4000 m of the model atmosphere) using that threshold. We use an altitude of 4000 m as a search window threshold here in order to correctly produce negative bias values where the predicted cloud is close to the 2000-m low-cloud cutoff value. A sensitivity test (varying the search window threshold from 2000 to 4000 m in 500-m steps) showed that the resulting critical success index (CSI) was largely insensitive to the choice of this threshold; CSI values using a 2000-m threshold were generally 0.02–0.03 lower than when using a 4000-m threshold, for both the $LCBH_{RH_{sonde}}$ and $LCBH_{RH_{NWP}}$ techniques (and also for the $LCBH_{\tau_{NWP}}$ technique discussed in the following section).

For the $LCBH_{RH_{sonde}}$ technique, the near-100% RH threshold values needed to obtain a bias of around zero

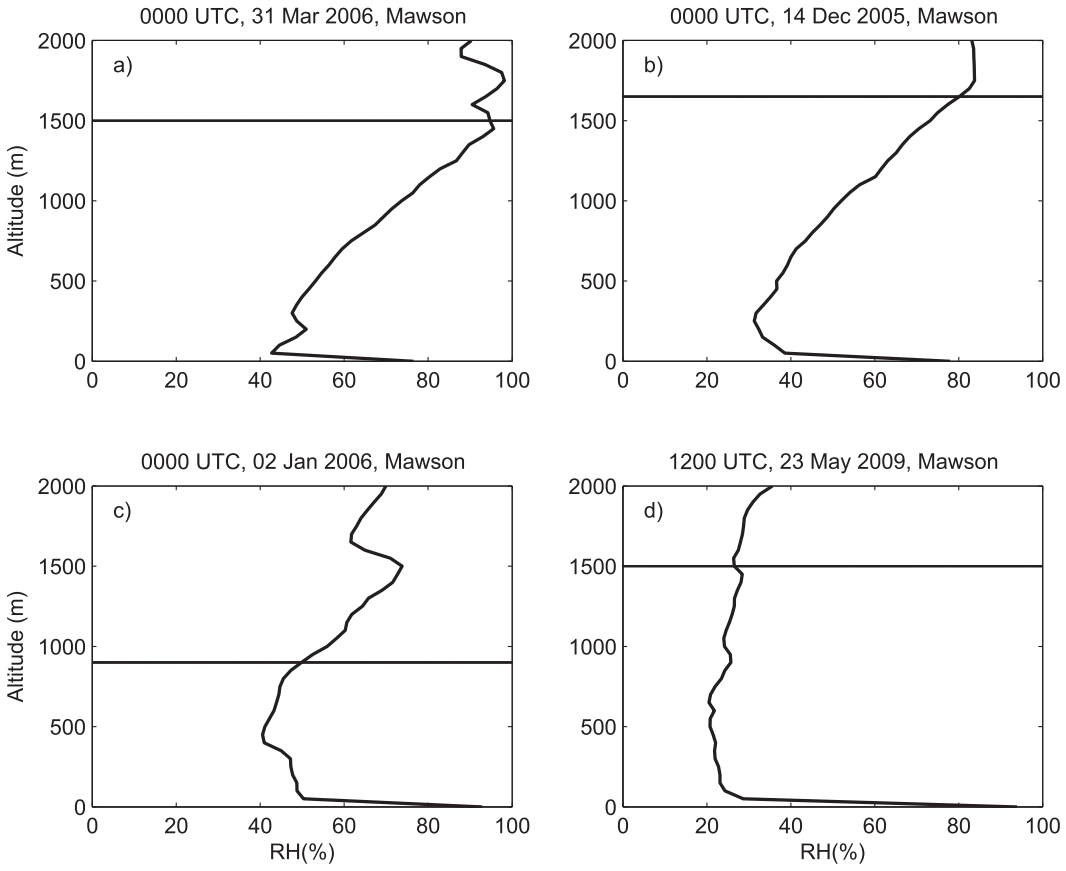

FIG. 3. Example profiles of RH_{sonde} at Mawson Station with $LCBH_{SOb}$ overplotted (horizontal line). (a) At 0000 UTC 31 Mar 2006, the RH_{sonde} profile shows a peak near 100% RH at the observed cloud-base height. (b) At 0000 UTC 14 Dec 2005, the RH profile shows a peak near the LCBH, though the RH is only ~80% at this altitude. (c) At 0000 UTC 2 Jan 2006, the observed LCBH of 900 m does not correspond with an RH peak, and the RH is only ~50% at the observed cloud-base height. (d) At 1200 UTC 23 May 2009, the observed LCBH occurs in very low RH. A peak to 65% RH was observed at 2300 m and the next at 4300 m.

indicate that the radiosonde humidity profile is a useful tool for cloud studies, despite ice clouds predominating over liquid clouds at polar latitudes. Around 89%–93% of low clouds were able to be reproduced using this technique.

For the $LCBH_{RH_{NWP}}$ technique (Table 3), the RH threshold values range between 58% and 66%, depending on location and cloud cover. These values are up to 30% lower than the equivalent thresholds for the radiosonde RH-derived profiles. This is likely a consequence of the dry lower NWP atmosphere, as discussed in the previous section. Nevertheless, this technique is able to reproduce 85%–95% of the observed low clouds. Standard deviation values are on the order of 800 m, compared with ~600 m for the radiosonde technique, likely reflecting the ~20% standard deviation of RH_{NWP} minus RH_{sonde} in the lower troposphere (see Fig. 1), as well as the much coarser vertical resolution in the NWP product.

In addition to these metrics, we also calculated the standard forecast metrics of bias score, probability of detection (POD), false alarm ratio (FAR), and CSI. These four metrics are calculated from total numbers of hits, misses, and false alarms. These terms are defined here:

- Hit—low cloud (i.e., a cloud base below 2000 m) was forecast using NWP outputs or predicted from the sonde RH profile, and low cloud was observed by the observer.
- Miss—no low cloud was forecast or predicted, but a low cloud was observed.
- False alarm—low cloud was forecast or predicted, but no low cloud was observed.

The bias score, POD, FAR, and CSI metrics (Schaefer 1990) are then defined as follow:

- Bias score = (hits + false alarms)/(hits + misses)—this metric answers the question, how does the forecast

TABLE 3. Comparison showing number of observations, RH, or τ threshold used to obtain near-zero mean bias, std dev of difference between predicted and observed cloud-base height, and percentage of clouds reproduced. Variable $LCBH_{RH_{sonde}}$ is defined as the lowest level where the RH_{sonde} profile exceeds the RH threshold. Variable $LCBH_{RH_{NWP}}$ is defined as the lowest level where the RH_{NWP} profile exceeds the RH threshold. Variable $LCBH_{\tau_{NWP}}$ is defined as the lowest level where the optical depth τ_{NWP} profile exceeds the τ threshold. Boldface indicates mean values where relevant (mean excludes values from Mawson when using the $LCBH_{RH_{sonde}}$ technique). Forecast metrics (bias score, POD, FAR, and CSI) are discussed in the text.

		8 oktas				≥6 oktas			
		Mawson	Davis	Casey	Mean	Mawson	Davis	Casey	Mean
$LCBH_{RH_{sonde}}$	No. of observed cloud events	120	109	394		216	401	780	
	RH threshold at min bias (%)	88	91	91	**90**	88	91	92	**90.3**
	Mean std dev (m)	480	600	729	**603**	439	508	698	**548**
	Clouds reproduced (%)	92.5	91.7	89.3	**91.2**	88.9	91.3	81.4	**87.2**
	Bias score	1.00	1.20	1.12	**1.11**	1.00	1.20	1.08	**1.09**
	POD	0.75	0.74	0.72	**0.74**	0.75	0.74	0.70	**0.73**
	FAR	0.24	0.39	0.36	**0.33**	0.24	0.39	0.35	**0.33**
	CSI	0.61	0.50	0.51	**0.54**	0.61	0.50	0.51	**0.54**
$LCBH_{RH_{NWP}}$	No. of observed cloud events	378	271	1221		706	1113	2336	
	RH threshold at min bias (%)	58	64	63	**61.7**	61	66	65	**64**
	Mean std dev (m)	821	794	805	**807**	801	771	865	**812**
	Clouds reproduced (%)	89.2	95.2	94.9	**93.1**	85.3	89.7	91.6	**88.8**
	Bias score	1.38	1.91	1.38	**1.56**	1.29	1.84	1.34	**1.49**
	POD	0.76	0.82	0.83	**0.80**	0.73	0.79	0.80	**0.77**
	FAR	0.45	0.57	0.40	**0.47**	0.44	0.57	0.40	**0.47**
	CSI	0.47	0.39	0.53	**0.47**	0.46	0.39	0.52	**0.46**
$LCBH_{\tau_{NWP}}$	No. of observed cloud events	378	271	1221		706	1113	2336	
	τ threshold at min bias	0.01	0.0125	0.0175	**0.013**	0.005	0.01	0.015	**0.01**
	Mean std dev (m)	829	838	797	**821**	864	804	848	**839**
	Clouds reproduced (%)	80.7	92.3	89.3	**87.4**	73.8	80.4	82.4	**78.9**
	Bias score	0.96	1.29	1.14	**1.13**	1.04	1.32	1.15	**1.17**
	POD	0.59	0.62	0.73	**0.65**	0.62	0.64	0.74	**0.67**
	FAR	0.38	0.52	0.36	**0.42**	0.40	0.52	0.36	**0.43**
	CSI	0.43	0.37	0.52	**0.44**	0.44	0.38	0.52	**0.45**

frequency of low-cloud events compare to the observed frequency? A score of unity indicates no bias.

- POD = hits/(hits + misses)—this metric measures the success of the forecast in correctly predicting the occurrence of events (but is insensitive to false alarms). A score of unity indicates no misses.
- FAR = false alarms/(hits + false alarms)—this metric measures the fraction of forecast events that were nonevents. A score of zero indicates no false alarms, while a score of unity implies all forecasted events were nonevents.
- CSI = hits/(hits + misses + false alarms) = $1/\{(1/\text{POD}) + [1/(1-\text{FAR})] - 1\}$—this metric measures the overall success of the forecast in correctly predicting the occurrence of events. A score of unity indicates a perfect forecast record, with no misses or false alarms.

For the $LCBH_{RH_{sonde}}$ technique (Table 3) these metrics show a large degree of skill. The bias score is close to unity, indicating forecast low cloud occurs at about the right frequency; the POD is reasonably high at ~0.74; the FAR is relatively low at ~0.33; and the CSI has a mean value of 0.54 across all stations.

The bias scores for the $LCBH_{RH_{NWP}}$ technique are higher than those obtained when using the $LCBH_{RH_{sonde}}$ technique. As a corollary, the POD and FAR are both higher when using the $LCBH_{RH_{NWP}}$ technique, and CSI is consequently lower (a mean value of ~0.46). This indicates that the $LCBH_{RH_{NWP}}$ low-cloud forecast is more conservative than that of the $LCBH_{RH_{sonde}}$ technique, at the cost of a higher FAR.

Little difference exists in the statistics of the 8-oktas datasets compared to the equivalent ≥6-oktas datasets, despite the guarantee of the radiosonde passing through the low cloud in the former, and the larger sample size of the latter. This indicates that both techniques perform similarly well under a range of low-cloud-cover conditions.

We conclude that RH profiles from radiosonde data are a reliable indicator of LCBH in Antarctica, despite the predominance of the ice phase in low clouds. Furthermore, despite the need for quite low (~63%) RH thresholds (a probable consequence of the too-dry NWP lower atmosphere), NWP RH profiles also represent LCBH with a comparable but slightly lower degree of accuracy (low cloud is overforecast, at the cost of a higher

FAR). It is worth acknowledging the difficulty in accurately predicting RH in Antarctica: Bromwich et al. (2013) note that the polar-optimized Weather Research and Forecasting (Polar WRF) Model shows little skill in predicting RH.

5. Cloud-base height forecasting from NWP-modeled cloud mixing ratios

Finally, we seek to determine if other NWP outputs can effectively reproduce the observed LCBH. The Antarctic Mesoscale Prediction System (AMPS) Polar WRF team derives cloud by simply drawing the $0.01\,\mathrm{g\,kg^{-1}}$ contour of the sum of the model microphysical fields of cloud ice, cloud water, and snow (K. Manning 2014, personal communication). We expect that the cloud water mixing ratio and cloud ice mixing ratio parameters, explicitly forecast in polarLAPS, will relate to cloud visibility after conversion to optical depth τ. This objective is thus achieved by converting NWP-derived cloud water and ice mixing ratios to optical depth, observing where these profiles exceed a threshold τ (i.e., determining $\mathrm{LCBH}_{\tau_{\mathrm{NWP}}}$), and comparing this against the surface observations.

Four polarLAPS variables explicitly involve water content: water vapor mixing ratio, rainwater mixing ratio, cloud water mixing ratio, and cloud ice mixing ratio. Only cloud water and cloud ice mixing ratios are relevant for Antarctic clouds. Water vapor is irrelevant for this study because it is invisible at visible wavelengths (400–700 nm) and therefore not a hazard to aviation, and rainwater mixing ratio is typically very low in Antarctica compared to cloud water and cloud ice mixing ratios. Thus, the water vapor and rainwater mixing ratios were excluded from this analysis.

Cloud water and cloud ice mixing ratios were converted to liquid water path LWP and ice water path IWP, respectively, for each NWP level. Variables LWP and IWP are defined as the total mass of water ($\mathrm{kg\,m^{-2}}$) (or water equivalent in the case of IWP) in a column of atmosphere.

From simple geometry,

$$\mathrm{LWP} = (\text{cloud water mixing ratio})(\rho_{\mathrm{air}})(\Delta Z) \quad \text{and}$$

$$\mathrm{IWP} = (\text{ice water mixing ratio})(\rho_{\mathrm{air}})(\Delta Z),$$

where ρ_{air} is density of air at the altitude of interest and ΔZ is the thickness of the NWP model vertical level. Following this conversion, LWP and IWP values were divided by the cosine of the solar zenith angle SZA to account for increased cloud visibility during times of low solar elevation. Values of SZA are generally between 45° and 90° during daytime around the coast of East Antarctica. In this study, a constant summertime value of SZA = 70° was used for simplicity.

Values of LWP and IWP were then converted to optical depth, a dimensionless quantity describing the visibility of a cloud, with the conversion to τ from LWP given by

$$\tau_{\mathrm{LWP}} = \left(\frac{9}{5}\right)\left(\frac{\mathrm{LWP}}{\rho_{\mathrm{water}} r_e}\right),$$

from Wood (2013), and for ice water content given by

$$\tau_{\mathrm{IWP}} = 0.068(\mathrm{IWP})^{0.83},$$

from Heymsfield et al. (2003), where ρ_{water} is the density of water and r_e is the effective cloud droplet radius. According to Lachlan-Cope (2010), the effective radius of Antarctic cloud droplets ranges from 4 to 200 μm. In this study, the effective radius was set to a constant value of 20 μm, in line with values observed at Palmer Station (a coastal station on the West Antarctic Peninsula) (Lachlan-Cope 2010).

Figure 4 shows the observable characteristics of several values of optical depth. We define $\tau = 3$ as thin cloud (Lo et al. 2006) and $\tau \simeq 1$ is a suitable value for typical cirrus clouds (Landulfo et al. 2008). A τ value of 0.03 is widely used as a threshold for cloud visibility (Bourassa et al. 2005; McFarquhar et al. 2000; Sassen et al. 1989; Giannakaki et al. 2007).

The τ threshold used here is determined in a similar manner to how the RH threshold is determined in the previous section: The value that gives the lowest mean difference between $\mathrm{LCBH_{SOb}}$ and $\mathrm{LCBH}_{\tau_{\mathrm{NWP}}}$ is taken as the threshold value. For each observation, the optical depth was calculated at each NWP model level from the cloud water and cloud ice mixing ratios. The lowest cloud level above the visible threshold was taken to be the NWP cloud base. This was then compared with $\mathrm{LCBH_{SOb}}$. Values of LCBH are calculated using a range of threshold τ values (from 0.0025 to 0.04 in increments of 0.0025). Results for this $\mathrm{LCBH}_{\tau_{\mathrm{NWP}}}$ technique are given in Table 3.

The threshold τ, standard deviation, and the percentage of observed clouds able to be reproduced, as well as the forecast metrics bias score, POD, FAR, and CSI, are given in Table 3. The τ threshold adjustments were able to effectively tune the mean bias to near zero in all six data subsets. Threshold τ values range from 0.005 to 0.0175 depending on the data subset. This is generally lower than the minimum τ threshold value for visible cloud as reported in the literature (i.e., $\tau = 0.03$; see Fig. 4), though this may be another consequence of a too-dry NWP lower troposphere.

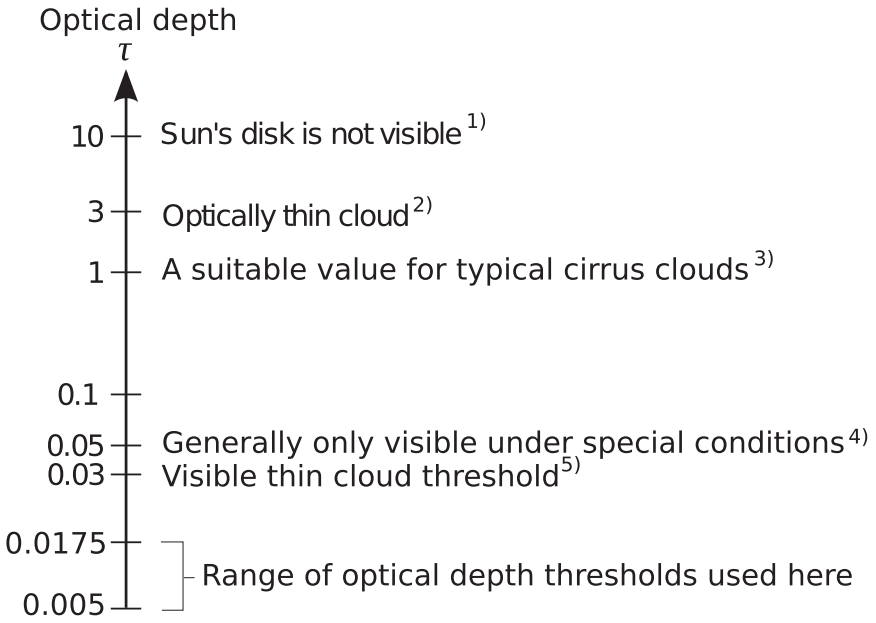

FIG. 4. Visibility characteristics of cloud for a range of values of optical depth: 1) Bohren et al. (1995); 2) Lo et al. (2006); 3) Landulfo et al. (2008); 4) Schmidt et al. (1993); and 5) Bourassa et al. (2005), McFarquhar et al. (2000), Sassen et al. (1989), and Giannakaki et al. (2007).

As a final step, we compare the $LCBH_{\tau_{NWP}}$ technique against $LCBH_{RH_{sonde}}$ and $LCBH_{RH_{NWP}}$ using the statistics in Table 3. The standard deviation of the difference between $LCBH_{SOb}$ and $LCBH_{\tau_{NWP}}$ is very similar to the values obtained with the $LCBH_{RH_{NWP}}$ method, again reflecting the coarse vertical resolution of the NWP outputs. The percentage of clouds able to be accurately reproduced is slightly lower than the $LCBH_{RH_{NWP}}$ method, but the bias score is much closer to unity, and the FAR is consequently much lower. This is reflected in a slightly lower CSI for the $LCBH_{\tau_{NWP}}$ technique. In summary, estimation of LCBH from the cloud mixing ratio conversion to τ is a viable alternative to the NWP-derived RH profile threshold method. By tuning the τ threshold to obtain an approximately zero bias between observed and modeled LCBH, the $LCBH_{\tau_{NWP}}$ method produces a less conservative forecast than the $LCBH_{RH_{NWP}}$ method.

6. Conclusions and further research

The research outlined in this paper has compared low-cloud-base height forecasts from NWP output against radiosonde-based RH threshold methods. The NWP model RH profile was first assessed by comparing RH_{NWP} against RH_{sonde}. We found RH_{NWP} to be drier from ground level to ~2000 m, and moister in the higher troposphere. The upper-level positive moisture bias was most likely due to the diagnosed high levels of moisture in the upper troposphere within the GFS model used to initialize polarLAPS. The low-level moisture deficit was suggested to be a result of inaccurate modeling of low-level prevailing winds, at least partially the result of choosing NWP grid points farther inland than the radiosonde launch sites, as well as elevation differences between observation stations and model grid points. The midtroposphere showed the smallest biases in the NWP–radiosonde comparison but the largest standard deviation at around 22% compared to 15% near the surface and in the upper atmosphere, likely attributable to the more complex synoptic- and mesoscale weather systems influencing the midtropospheric regions (Holton 1992). This moisture deficit resulted in the very low RH and τ thresholds required for LCBH prediction using the NWP product. Despite unrealistically low thresholds, both forecast techniques showed roughly equal skill in predicting LCBH. The RH forecast technique produced a more conservative low-cloud forecast than the τ technique, but both NWP techniques were able to reproduce a similar percentage of observed low clouds, and both produced a similar mean standard deviation. As expected, neither NWP-based technique showed as much skill as the radiosonde RH-based technique.

Across all three LCBH prediction methods, Mawson Station consistently required the lowest RH and τ thresholds for optimum cloud-base height detection. The formation of cloud at lower RH/τ values at Mawson is possibly related to a different cloud microphysics regime along the strongly katabatic-dominated Mawson coast. Relatively high moisture levels in the first 100 m or so from the surface were often measured there by the

sondes within otherwise very dry continental outflow environments, such as in Fig. 3. This moist surface layer was particularly evident during strong wind events and presumably results from the ablation/sublimation of the icy surface and windborne ice crystals. Also worth noting is that blowing snow can remain airborne for days following strong winds (Walden et al. 2003). Many of the instances where low cloud was observed in dry RH sonde environments (such as Figs. 3c,d) occurred in windy continental flow regimes. It would be an interesting extension of this work to investigate whether this moist surface layer can at times be lifted to its condensation level to produce optically visible thin cloud that goes underdetected by the 2-s (10 m) sampling of the sonde. Davis Station consistently reported the lowest CSI score, both in the 8- and ≥6-oktas cases, because of a high FAR score. This may be attributable to the NWP station grid point being situated about 8 km farther inland than the observation and coastal radiosonde release site (see Table 2). This also places the NWP station grid point only 2 km from the Sørsdal Glacier, where drier cold-air drainage is more prone to throughflow, which in turn supports the relatively higher frequency of easterly winds at the NWP gridpoint site (see Fig. 2).

An alternative LCBH determination technique was investigated using NWP-modeled cloud water and cloud ice mixing ratios converted into optical depth as a discriminator of cloud. Both NWP-derived LCBH retrieval techniques ($LCBH_{\tau_{NWP}}$ and $LCBH_{RH_{NWP}}$) were roughly equally effective. Both NWP-based techniques returned similar values for the standard deviation between observed and modeled cloud base and percentage of clouds able to be reproduced. Combining both the RH- and τ-based retrievals for enhanced NWP prediction would be a worthwhile extension of this work. We also suggest for future work that it may be worthwhile to consider grid points that may be more representative of the station (i.e., at the same altitude as the station) rather than the geographically closest one (which is often inland) to see if the low-level moisture forecasts improve.

Although polarLAPS has recently been decommissioned, the intercomparison techniques presented in this study could be used to verify other high-latitude models [e.g., the new Australian Antarctic NWP model, Australian Community Climate and Earth-System Simulator Polar (ACCESS-P)]. This could also incorporate a comparison of different cloud microphysics packages, to assess which models perform best in the Antarctic region.

Acknowledgments. This work was carried out with the support of the Australian Government Cooperative Research Centres Programme through the Australian Bureau of Meteorology and the Antarctic Climate & Ecosystems Cooperative Research Centre (ACE CRC), and was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant 25.03748.

On 23 March 2012 at the age of 51 Neil Adams passed away suddenly. He leaves behind him a tremendous legacy to the Antarctic and meteorological community. Neil has been nationally and internationally recognized for his dedication to, and role in, advancing the science and widespread application of meteorology, particularly in the fields of data presentation, high-latitude meteorology, and service provision. Neil contributed to several national and international polar-related committees/forums. His many authored and coauthored peer-reviewed papers are further testament to his inherent professionalism and expertise. Neil is sadly missed by his many colleagues in the Antarctic community.

Biography provided by P. Reid, Australian Bureau of Meteorology.

REFERENCES

Adams, N., 2002: Advances in forecasting systems at the Antarctic Meteorological Centre, Casey. *Meteor. Appl.,* **9,** 335–343, doi:10.1017/S1350482702003079.

——, 2004: A numerical modeling study of the weather in East Antarctica and the surrounding Southern Ocean. *Wea. Forecasting,* **19,** 653–672, doi:10.1175/1520-0434(2004)019<0653:ANMSOT>2.0.CO;2.

——, and J. Powers, 2007: Advances in Numerical Weather Precipitation and weather forecasting systems in support of the Australian Antarctic program. *Proc. Second Antarctic Meteorological Observation, Modeling, and Forecasting Workshop,* Rome, Italy, International Commission on Polar Meteorology. [Available online at http://amrc.ssec.wisc.edu/presentations_2007/june_27/adams.pdf.]

——, H. Hutchinson, and T. Hart, 1999: An analysis of TOVS data during the FROST special observing periods. *Aust. Meteor. Mag.,* **50,** 12–42.

Australian Bureau of Meteorology, 2014: Aviation observations and reports. *Aeronautical Services Handbook* 2014, Australian Bureau of Meteorology, 57.

Bohren, C. F., J. R. Linskens, and M. E. Churma, 1995: At what optical thickness does a cloud completely obscure the sun? *J. Atmos. Sci.,* **52,** 1257–1259, doi:10.1175/1520-0469(1995)052<1257:AWOTDA>2.0.CO;2.

Bourassa, A. E., D. A. Degenstein, and E. J. Llewellyn, 2005: Climatology of the subvisual cirrus clouds as seen by OSIRIS on Odin. *Adv. Space Res.,* **36,** 807–812, doi:10.1016/j.asr.2005.05.045.

Bromwich, D. H., F. O. Otieno, K. M. Hines, K. W. Manning, and E. Shilo, 2013: Comprehensive evaluation of Polar Weather Research and Forecasting Model performance in the Antarctic. *J. Geophys. Res. Atmos.,* **118,** 274–292, doi:10.1029/2012JD018139.

Giannakaki, E., D. S. Balis, V. Amiridis, and S. Kazadzis, 2007: Optical and geometrical characteristics of cirrus clouds over a southern European lidar station. *Atmos. Chem. Phys.,* **7,** 5519–5530, doi:10.5194/acp-7-5519-2007.

Heymsfield, A. J., S. Matrosov, and B. Baum, 2003: Ice water path–optical depth relationships for cirrus and deep stratiform ice cloud layers. *J. Appl. Meteor.,* **42,** 1369–1390, doi:10.1175/1520-0450(2003)042<1369:IWPDRF>2.0.CO;2.

Holton, J., 1992: *An Introduction to Dynamic Meteorology.* 3rd ed. International Geophysics Series, Vol. 48, Academic Press, 511 pp.

Korolev, A., and G. A. Isaac, 2006: Relative humidity in liquid, mixed-phase, and ice clouds. *J. Atmos. Sci.,* **63,** 2865–2880, doi:10.1175/JAS3784.1.

Lachlan-Cope, T., 2010: Antarctic clouds. *Polar Res.,* **29,** 150–158, doi:10.1111/j.1751-8369.2010.00148.x.

Landulfo, E., E. G. Larroza, F. J. S. Lopes, W. C. de Jesus, M. Bottino, W. M. Nakaema, and J. Steffens, 2008: Sub-visual cirrus LIDAR measurements for satellite masking improvement. *Remote Sensing of Clouds and the Atmosphere XIII,* R. H. Picard et al., Eds., International Society for Optical Engineering (SPIE Proceedings, Vol. 7107), doi:10.1117/12.799872.

Lo, C., J. M. Comstock, and C. Flynn, 2006: An atmospheric radiation measurement value-added product to retrieve optically thin cloud visible optical depth using micropulse lidar. PNNL Tech. Rep. DOE/SC-ARM/TR-077, Pacific Northwest National Laboratory, 11 pp. [Available online at https://www.arm.gov/publications/tech_reports/doe-sc-arm-tr-077.pdf.]

Luo, Y., and S. Krueger, 2004: Cloud types simulated by the NCEP GFS model. *Proc. 14th ARM Science Team Meeting,* Albuquerque, NM, Atmospheric Radiation Measurement (ARM) Climate Research Facility. [Available online at http://www.arm.gov/publications/proceedings/conf14/extended_abs/luo3-y.pdf.]

McFarquhar, G. M., A. J. Heymsfield, J. Spinhirne, and B. Hart, 2000: Thin and subvisual tropopause tropical cirrus: Observations and radiative impacts. *J. Atmos. Sci.,* **57,** 1841–1853, doi:10.1175/1520-0469(2000)057<1841:TASTTC>2.0.CO;2.

Miloshevich, L. M., A. Paukkunen, H. Vömel, and S. J. Oltmans, 2004: Development and validation of a time-lag correction for Vaisala radiosonde humidity measurements. *J. Atmos. Oceanic Technol.,* **21,** 1305–1327, doi:10.1175/1520-0426(2004)021<1305:DAVOAT>2.0.CO;2.

——, H. Vömel, D. N. Whiteman, and T. Leblanc, 2009: Accuracy assessment and correction of Vaisala RS92 radiosonde water vapor measurements. *J. Geophys. Res.,* **114,** D11305, doi:10.1029/2008JD011565.

National Weather Service Environmental Modeling Center, cited 2011: The Global Forecast System (GFS)—Global Spectral Model (GSM). NOAA/NWS/Environmental Modeling Center. [Available online at http://www.emc.ncep.noaa.gov/GFS/doc.php.]

Pendlebury, S. F., N. Adams, T. L. Hart, and J. Turner, 2003: Numerical weather prediction model performance over high southern latitudes. *Mon. Wea. Rev.,* **131,** 335–353, doi:10.1175/1520-0493(2003)131<0335:NWPMPO>2.0.CO;2.

Puri, K., G. Dietachmayer, G. Mills, N. Davidson, R. Bowen, and L. Logan, 1998: The new BMRC Limited Area Prediction System, LAPS. *Aust. Meteor. Mag.,* **47,** 203–223.

Sassen, K., M. K. Griffin, and G. C. Dodd, 1989: Optical scattering and microphysical properties of subvisual cirrus clouds, and climatic implications. *J. Appl. Meteor.,* **28,** 91–98, doi:10.1175/1520-0450(1989)028<0091:OSAMPO>2.0.CO;2.

Schaefer, J. T., 1990: The critical success index as an indicator of warning skill. *Wea. Forecasting,* **5,** 570–575, doi:10.1175/1520-0434(1990)005<0570:TCSIAA>2.0.CO;2.

Schmidt, E. O., J. M. Alvarez, M. A. Vaughan, and D. P. Wylie, 1993: Review of subvisual cirrus morphology. *Passive Infrared Remote Sensing of Clouds and the Atmosphere,* D. K. Lynch, Ed., International Society for Optical Engineering (SPIE Proceedings, Vol. 1934), 230–239, doi:10.1117/12.154907.

Tomasi, C., B. Petkov, E. Benedetti, L. Valenziano, and V. Vitale, 2011: Analysis of a 4 year radiosonde data set at Dome C for characterizing temperature and moisture conditions of the Antarctic atmosphere. *J. Geophys. Res.,* **116,** D15304, doi:10.1029/2011JD015803.

Turner, J., S. Pendlebury, L. Cowled, K. Jacka, M. Jones, and P. Targett, 2000: Report on the First International Symposium on Operational Weather Forecasting in Antarctica. *Bull. Amer. Meteor. Soc.,* **81,** 75–94, doi:10.1175/1520-0477(2000)081<0075:ROTFIS>2.3.CO;2.

Walden, V. P., S. G. Warren, and E. Tuttle, 2003: Atmospheric ice crystals over the Antarctic Plateau in winter. *J. Appl. Meteor.,* **42,** 1391–1405, doi:10.1175/1520-0450(2003)042<1391:AICOTA>2.0.CO;2.

Wood, R., cited 2013: Relationships between optical depth, liquid water path, droplet concentration and effective radius in an adiabatic layer cloud. Dept. of Atmospheric Sciences, University of Washington. [Available online at http://www.atmos.washington.edu/~robwood/papers/chilean_plume/optical_depth_relations.pdf.]