Performance of Weather Forecast Models in the Rescue of Dr. Ronald Shemenski from the South Pole in April 2001*

ANDREW J. MONAGHAN AND DAVID H. BROMWICH

Polar Meteorology Group, Byrd Polar Research Center, and Atmospheric Sciences Program, Department of Geography, The Ohio State University, Columbus, Ohio

HE-LIN WEI

Polar Meteorology Group, Byrd Polar Research Center, The Ohio State University, Columbus, Ohio

ARTHUR M. CAYETTE

SPAWAR Systems Center, Charleston, South Carolina

JORDAN G. POWERS AND YING-HWA KUO

Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research, Boulder, Colorado

MATTHEW A. LAZZARA

Antarctic Meteorological Research Center, Space Science and Engineering Center, University of Wisconsin—Madison, Madison, Wisconsin

(Manuscript received 21 May 2002, in final form 14 November 2002)

ABSTRACT

In late April 2001, an unprecedented late-season flight to Amundsen–Scott South Pole Station was made in the evacuation of Dr. Ronald Shemenski, a medical doctor seriously ill with pancreatitis. This case study analyzes the performance of four of the numerical weather prediction models that aided meteorologists in forecasting weather throughout the operation: 1) the Antarctic Mesoscale Prediction System (AMPS) Polar MM5 (fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model), 2) the National Centers for Environmental Prediction Aviation Model (AVN), 3) the European Centre for Medium-Range Weather Forecasts (ECMWF) global forecast model, and 4) the NCAR Global MM5. To identify specific strengths and weaknesses, key variables for each model are statistically analyzed for all forecasts initialized between 21 and 25 April for several points over West Antarctica at the surface and at 500- and 700-hPa levels. The ECMWF model performs with the highest overall skill, generally having the lowest bias and rms errors and highest correlations for the examined fields. The AMPS Polar MM5 exhibits the next best skill, followed by AVN and Global MM5. For the surface variables, all of the models show high skill in predicting surface pressure but demonstrate modest skill in predicting temperature, wind speed, and wind direction. In the free atmosphere, the models show high skill in forecasting geopotential height, considerable skill in predicting temperature and wind direction, and good skill in predicting wind speed. In general, the models produce very useful forecasts in the free atmosphere, but substantial efforts are still needed to improve the surface prediction. The spatial resolution of each model exerts an important influence on forecast accuracy, especially in the complex topography of the Antarctic coastal regions. The initial and boundary conditions for the AMPS Polar MM5 exert a significant influence on forecasts.

1. Introduction

Over the past decade, numerical weather prediction over the polar regions has made significant progress. In addition to the continual improvement of spatial resolution, forecasters now have access to several global and regional models. The implementation of physical parameterizations that are well suited to polar phenomena has also aided progress. The Polar MM5, for example, is a version of the fifth-generation Pennsylvania State University–National Center for Atmospheric Research (NCAR) Mesoscale Model (MM5) that has been modified for use over extensive ice sheets (Bromwich et al. 2001; Cassano et al. 2001). Recently, the model was tested in forecast mode over Antarctica for a 1-yr period and showed generally high skill (Guo et al. 2003). The Antarctic Mesoscale Prediction System (AMPS; Powers et al. 2003), which em-

* Byrd Polar Research Center Contribution Number 1255.

Corresponding author address: Andrew J. Monaghan, Polar Meteorology Group, Byrd Polar Research Center, The Ohio State University, 1090 Carmack Rd., Columbus, OH 43210.
E-mail: monaghan@polarmet1.mps.ohio-state.edu

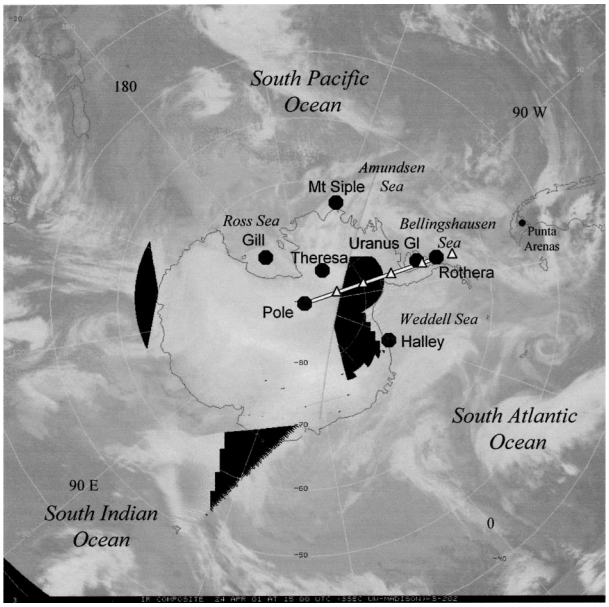

FIG. 1. Approximate flight path from Rothera to Pole (white line), superimposed on a satellite composite for 1500 UTC 24 Apr 2001, showing conditions during the flight. The dots are locations where modeled data are compared with observations. The filled dots correspond to surface stations, while triangles represent study points at 700 hPa. Black areas indicate no imagery available. (Courtesy of the Antarctic Meteorological Research Center.)

ploys the Polar MM5, was implemented in the 2000–01 field season. AMPS resulted from the forecasting needs expressed at the May 2000 Antarctic Weather Forecasting Workshop held at the Byrd Polar Research Center in Columbus, Ohio (Bromwich and Cassano 2001). AMPS is an experimental system employing models run at NCAR and dedicated to real-time numerical weather prediction in Antarctica. For the period covered by this study, AMPS was run for three domains at various resolutions covering most of the Southern Hemisphere from 40°S, with the finest grid (10 km) focusing on the United States Antarctic Program's base of operations at McMurdo Station. At present, AMPS has been upgraded to include an additional 3.3-km-resolution grid over McMurdo, and a 10-km grid over South Pole. AMPS forecasts are available to the public via the Internet (http://www.mmm.ucar.edu/rt/mm5/amps/).

In the Antarctic, the importance of operational forecasting has become more apparent in recent years. One reason is the effort to lengthen the field season, requiring flights in early spring and late autumn when the entire continent is subject to extremely cold temperatures, strong wind episodes, and limited daylight. In addition, there is the occasional need to perform emergency evacuations of injured or ill personnel.

In late April 2001, a DeHaviland Twin Otter aircraft made an unprecedented late-season flight from Rothera to Amundsen–Scott South Pole Station (hereinafter, ''Pole'' or ''South Pole''; Fig. 1) for the evacuation of Dr. Ronald Shemenski, a medical doctor seriously ill with pancreatitis. The operation was complicated by near 24-h darkness over interior Antarctica and extreme cold temperatures at South Pole, below the −55°C operating limits of the commonly used LC-130 military aircraft. Frequent blowing snow conditions and the lack of a lighted runway and a control tower made landing the aircraft at South Pole a hazardous proposition. Throughout the operation, pilots depended on weather forecasters in McMurdo and the Falkland Islands to predict weather conditions accurately for the 10-h flight. Because of the sparse observational network, forecasters relied heavily on numerical weather prediction models to aid in their forecast generation.

Using this high-profile case, the performance of four numerical weather prediction models is analyzed for the purpose of comparison and the identification of strengths and weaknesses. First, the models and data used in the study are discussed in section 2. Next, the synoptic situation over Antarctica from 21 to 28 April 2001 is discussed in section 3. In section 4, we examine model performance at the surface in comparison with observations throughout West Antarctica. Model performance is also assessed at the 500-hPa level in comparison with radiosonde observations at South Pole and the 700-hPa level in comparison with the European Centre for Medium-Range Weather Forecasts (ECMWF) Tropical Ocean and Global Atmosphere (TOGA) operational analysis data (henceforth, ECMWF–TOGA) along the approximate flight path from Rothera to Pole. At the end of section 4 we examine the effects of differing initial and boundary conditions on the AMPS forecast by comparing simulations that employ ECMWF–TOGA data with simulations using the National Centers for Environmental Prediction (NCEP) Aviation (AVN) Model output, which is currently used in the operational version of AMPS. In section 5, some of the important model deficiencies are addressed. Conclusions are presented in section 6.

2. Data and methods

Table 1 gives a summary of each model used in the study. Of the three domains of AMPS (Powers et al. 2002, manuscript submitted to *Bull. Amer. Meteor. Soc.*), the 30-km version used in this study is the mid-range spatial resolution domain, and covers the continent and coastal regions of Antarctica using a polar stereographic projection (the 10-km domain of AMPS did not encompass the South Pole and is, therefore, not used). The NCEP AVN spectral forecast model (Kanamitsu 1989; Kanamitsu et al. 1991) is run globally with approximately 1° latitude × 1° longitude spatial resolution, and is initialized earlier than most models with the main objective of providing timely forecasts in support of aviation. The ECMWF spectral forecast model (Reading, United Kingdom) is run globally with approximately 0.5° latitude × 0.5° longitude spatial resolution. The Global MM5 (Dudhia and Bresch 2002) is

TABLE 1. Summary of models used in the study.

Model	Agency	Alias (for purpose of discussion)	Domain	Approximate resolution with respect to latitude (km)	Projection	Initialization times (UTC)	Forecast duration (h)
AMPS_MM5	NCAR	AMPS	Antarctica	30	Polar stereographic	0000, 1200	48
AVN	NCEP	AVN	Global	110	Equal Lat–lon	0000, 1200	72
ECMWF Forecast	ECMWF	ECM	Global	55	Equal Lat–lon	1200	168
Global_MM5	NCAR	GLO	Global	120	Polar stereographic	0000, 1200	120

Abbreviations: NCAR, National Center for Atmospheric Research; AMPS, Antarctic Mesoscale Prediction System MM5; NCEP, National Centers for Environmental Prediction; AVN, Aviation model; ECMWF, European Centre for Medium-Range Weather Forecasts; ECM, ECMWF forecast model; GLO, Global MM5.

the MM5 run globally employing a polar stereographic projection, and has approximately 120-km spatial resolution in high latitudes. All of the models are initialized twice daily (0000 and 1200 UTC), with the exception of the ECMWF forecast model, which is initialized once daily (1200 UTC). The archived model output for AMPS, AVN, and Global MM5 was obtained from NCAR, while the ECMWF model output was purchased from ECMWF. Henceforth, the alias indicated in Table 1 will be used to refer to each model.

It is noteworthy that there are some models not included in the analysis. We have tried to include the primary models used by the U.S. forecasters. This excludes, for example, the Met Office forecast model (UKMET). An evaluation of the performance of UKMET over Antarctica can be found in Leonard et al. (1997). In addition, archived data could not be obtained for one of the primary models used to aid the U.S. forecasting effort, the Air Force Weather Information Network (AFWIN) standard version of the MM5. Also excluded because of similarities to the AMPS MM5 is a 60-km-resolution version of the Polar MM5 covering Antarctica, which is run at the Byrd Polar Research Center of The Ohio State University (Cassano et al. 2000) and is intended as a backup should technical problems arise when running AMPS.

a. Surface analyses

Model output is compared with observed data at seven sites distributed throughout West Antarctica (indicated by dots in Fig. 1). The model results are interpolated from the four points surrounding the observation site. Data obtained from the University of Wisconsin Antarctic Meteorology Research Center (AMRC) automatic weather stations (AWSs) provide the observations at five of the sites, including South Pole (information online at http://amrc.ssec.wisc.edu/). At the two British sites (Rothera and Halley), observations transmitted via the Global Telecommunications System (GTS) and archived by the British Antarctic Survey (BAS) are used (information online at http://www.antarctica.ac.uk/met/). The cloud fraction observations for South Pole are from manual observations submitted via the GTS and archived at BAS.

The surface variables examined are wind speed, wind direction, pressure, temperature, and cloud fraction. The wind instrumentation is installed at a nominal height of 3 m, and surface temperature at a height of 2 m at the AWS sites (although this varies due to snow accumulation). At the British sites the wind instrumentation is at a height of 10 m (adjusted here to 3 m for consistency), and temperature at 1.5 m. Where necessary, the modeled wind speed and temperature data have been adjusted to the observed heights by applying Monin–Obukhov similarity theory (Stull 1988) to the temperature and wind speed at the lowest model level and the model surface temperature assuming a roughness length of 0.0001 m, which is considered representative of an ice sheet (e.g., Budd et al. 1966). This method accounts for a near-surface temperature inversion in the case that one exists. Modeled temperatures are also adjusted for the difference between modeled and observed elevations assuming a lapse rate of $-dt/dz = 0.01°\text{C m}^{-1}$ (note that this lapse rate need not consider the near-surface inversion as it reflects the lapse rate due to elevation changes along the *surface* of the ice sheet). This value is chosen following Budd et al. (1971) who give annual lapse rates (due to elevation change) along the snow surface in West Antarctica ranging from $0.0025°\text{C m}^{-1}$ (near the coast) to $0.02°\text{C m}^{-1}$ (inland). In most cases this adjustment is small, as the elevation differences are generally within ± 100 m. Most AWS sites at high elevations have a significant amount of uncertainty in the reported elevation because it has been determined using aircraft altimetry. To compensate for this, it is assumed that a given modeled-minus-observed pressure bias is due to elevation uncertainties, and this bias is subtracted from the modeled value. In this manner, a better estimation of the root-mean-square error (rmse) can be ob-

tained, although a true bias cannot be determined. Modeled cloud fraction estimates in AMPS follow that of Kiehl et al. (1996). In AVN, ECM (ECMWF forecast model), and GLO (Global MM5), modeled cloud fraction estimates follow that of Slingo (1987).

To minimize the temporal differences resulting from varying forecast durations, statistics are calculated for all 6-hourly forecasts out 6–48 h from the initial (0 h) time (results for times beyond 48 h for the longer-range models yielded similar results). All available forecasts initialized from 21 to 25 April 2001 are used. The statistical differences induced by comparing once-daily (ECM) and twice-daily (AVN, AMPS, GLO) forecasts were checked and found to be minimal. The statistics computed include bias, rmse, and correlation coefficient. The bias is the difference between the modeled mean for the period and the observed mean for the period. The rmse is the square root of the period-averaged squared difference between the modeled and observed values, and represents the ''typical'' difference between modeled and observed values (Cassano et al. 2001). The correlation coefficient measures the relative agreement between the two series. The strength of the correlation is computed for the 95% confidence interval using a two-tailed *t*-test.

b. 500-hPa analyses (South Pole)

Radiosonde observations (archived at AMRC) are used to examine model performance at the 500-hPa level over South Pole. Variables examined are 500-hPa wind speed, wind direction, geopotential height, and temperature. No adjustments are made to these variables. The statistics are the same as those computed for the surface analyses, and are computed for all 12-hourly forecasts out to 12–48 h from the initial (0 h) time. Twelve-hourly intervals correspond to the availability of radiosonde data.

c. 700-hPa analyses (along flight path)

The flight from Rothera to South Pole flew at approximately the 700-hPa level along the 70°W meridian (white line in Fig. 1). To examine the models at this level, model data are compared with ECMWF–TOGA data at five points (5° latitude intervals from 65° to 85°S, indicated by the white triangles in Fig. 1) along the 70°W meridian. South Pole (90°S, 70°W) is not included because the 700-hPa surface is below ground level. The variables, statistics, and treatment of time are the same as described for the 500-hPa analyses, with the exception that ECMWF–TOGA data are used for validation, as no radiosonde data are available for these points.

The ECMWF–TOGA data were quality checked using radiosonde observations where available. Figure 2 shows the ECMWF–TOGA minus observed values for the 700-hPa wind speed, geopotential height, and temperature means for 15–30 April. Inspection of the wind

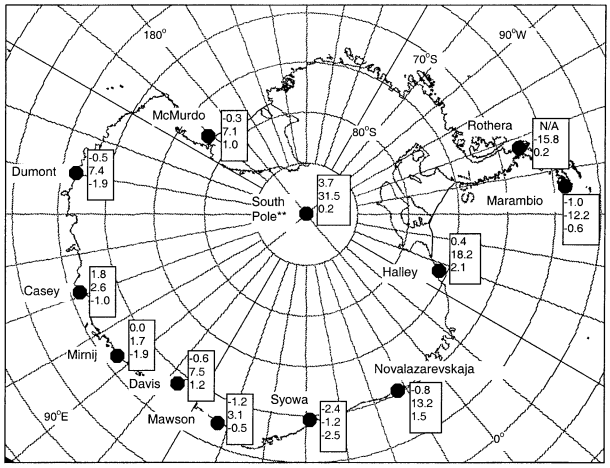

FIG. 2. ECMWF–TOGA minus radiosonde observations at 700 hPa for available sites in Antarctica. The values are averaged from 15 to 30 Apr 2001. The top value is the wind speed difference (m s^{-1}). The middle value is the geopotential height difference (gpm). The bottom value is temperature difference (°C). Note that South Pole data are for the 500-hPa level, because the surface is above the 700-hPa level. Observations not available are denoted by N/A.

speed differences indicates seven coastal sites with biases less than or equal to ± 1 m s^{-1}, and three coastal sites with biases greater than ± 1 m s^{-1}. Small negative biases are most common. No significant regional trends are present. At South Pole, the only site in the continental interior, a large positive bias of 3.7 m s^{-1} is found at 500 hPa (700 hPa is not available at South Pole due to the surface elevation). This bias should be taken into account when considering the statistics along the flight path from Rothera to Pole. Inspection of the 700-hPa geopotential height differences indicates that the ECMWF–TOGA data are negatively biased on the Antarctic Peninsula by approximately 15 gpm (Rothera and Marambio). In contrast, farther east the geopotential height is positively biased by approximately 15 gpm at Halley and Novalazarevskaja. Elsewhere, there are no significant biases except at South Pole. Inspection of the 700-hPa temperature differences does not indicate any regional warm or cold trends in the ECMWF–TOGA data. However, there are several sites with biases greater than $\pm\sim 2$°C. At four of the five sites nearest the flight path (Rothera, South Pole, McMurdo, and Halley), warm biases (0.2°–2.1°C) are observed.

Because most radiosonde stations are on the coastal margin, the skill of ECMWF–TOGA over the continental interior was assessed using other methods. First, the technique of Phillpot (1997) was used to check the ECMWF–TOGA geopotential heights for qualitative accuracy over the continental interior by extrapolating AWS data (using the sites shown in Fig. 1) to the 700-hPa level. This method uses the hypsometric equation to estimate the thickness of the layer between the surface (where the given station is located) and the 700-hPa level for an air column with a mean virtual temperature

(MVT) of 0°C. Second, the ECMWF–TOGA 700-hPa data were compared against analyses obtained from the United Kingdom Met Office (UKMO) for the period 21–28 April. Based on these comparisons, the uncertainty errors in ECMWF–TOGA at 700 hPa (over the coast and the interior) are thought to be on the order of ± 1 m s^{-1} (wind speed), ± 15 gpm (geopotential height), and ± 2°C (temperature). All of the ECMWF–TOGA 700-hPa fields used in statistical computations are in good *qualitative* agreement with the observations. With this in mind, the correlation coefficient may be the best indicator of the model performance at the upper levels, as this value only considers the similarity between two curves, and does not consider biases. The authors feel that the ECMWF–TOGA analysis is the best choice for validating the 700-hPa data in this study. Other studies (e.g., Cullather et al. 1997) support this, showing that ECMWF–TOGA performs better than similar analyses over Antarctica. It is stressed that ECMWF–TOGA is an estimation of the actual atmospheric state based on available observations; this should be considered when making comparisons to the forecast models in data-sparse areas. In addition, the ECM model will be slightly favored by using ECMWF–TOGA, as it is being compared with its own analysis.

3. Synoptic situation over Antarctica, 21–28 April 2001

Figure 3 shows the 0000 UTC ECMWF–TOGA 500-hPa geopotential heights for 21–28 April 2001; the high elevation at South Pole (2835 m) provides some correlation between the 500-hPa height gradient and the surface winds. The relationship between the South Pole surface and 500-hPa levels is shown in Figs. 4a–d, which compare the surface AWS observations and 500-hPa radiosonde data for surface pressure/geopotential height, wind speed, wind direction, and temperature. Figure 4e shows the observed cloud fraction at South Pole. In Fig. 4a, ECMWF–TOGA data are also plotted for surface pressure and 500-hPa geopotential height in order to demonstrate the close correlation to observed fields. To address the fact that every direction is north at South Pole, when discussing the weather here (following the usual convention) it is assumed that Antarctica is overlaid by a grid in which north (''grid north'') is aligned with the 0° meridian. With this notation, grid east would be 90°E, grid south would be 180°, and grid west would be 90°W. Later in the text, when discussing sites other than Pole, directions are with respect to true north.

After reaching Rothera (via Punta Arenas, South America) on 21 April, the flight to Pole was delayed on 22–23 April due to moderate blowing snow conditions at Pole that degraded conditions to below flight minimum requirements (~300 m ceiling and ~5 km visibility). From 21 April through ~0000 UTC 24 April, a trough extending from the Ross Sea across South Pole

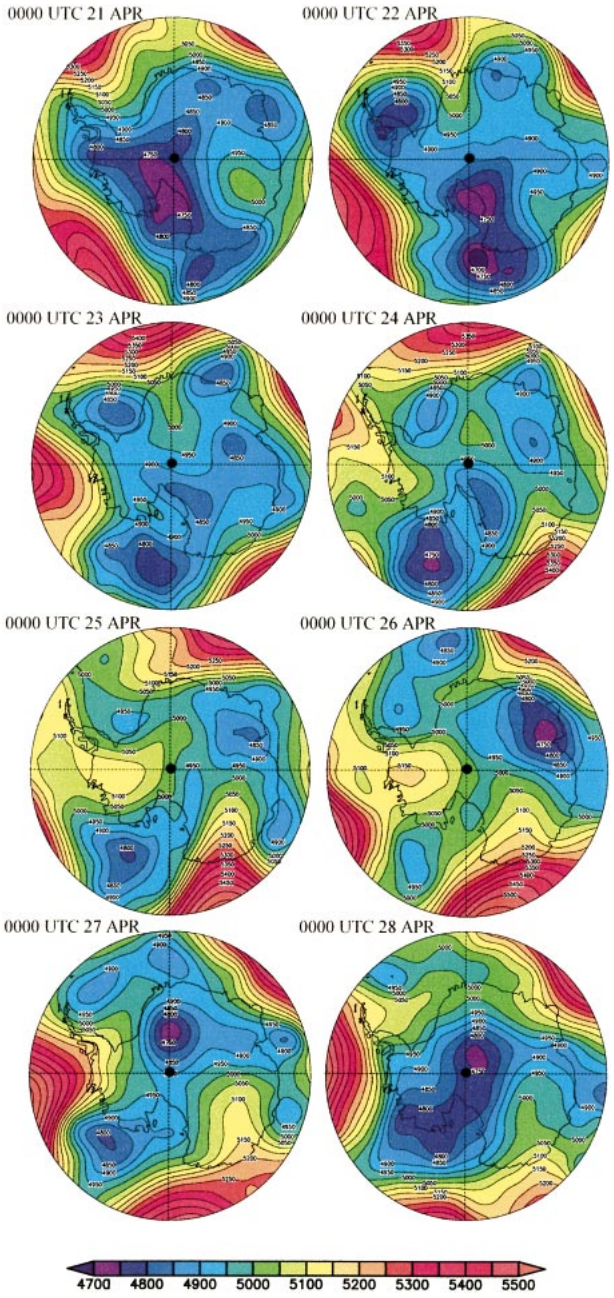

FIG. 3. ECMWF–TOGA 500-hPa geopotential heights (gpm) over Antarctica (60°–90°S) at 0000 UTC each day, 21–28 April 2001. South Pole is indicated by a black dot. The grid north direction (0°) points upward.

to the Weddell and Bellingshausen Seas (Fig. 3) provided support for the observed surface winds in excess of 6 m s^{-1} (the approximate blowing snow threshold), which occurred at Pole until about 24 April (Fig. 4b). On 23 April, a ridge began to extend poleward from the Bellingshausen Sea and continued to strengthen until about 26 April. This was noted at Pole by rising surface pressure and geopotential height (Fig. 4a), and in the

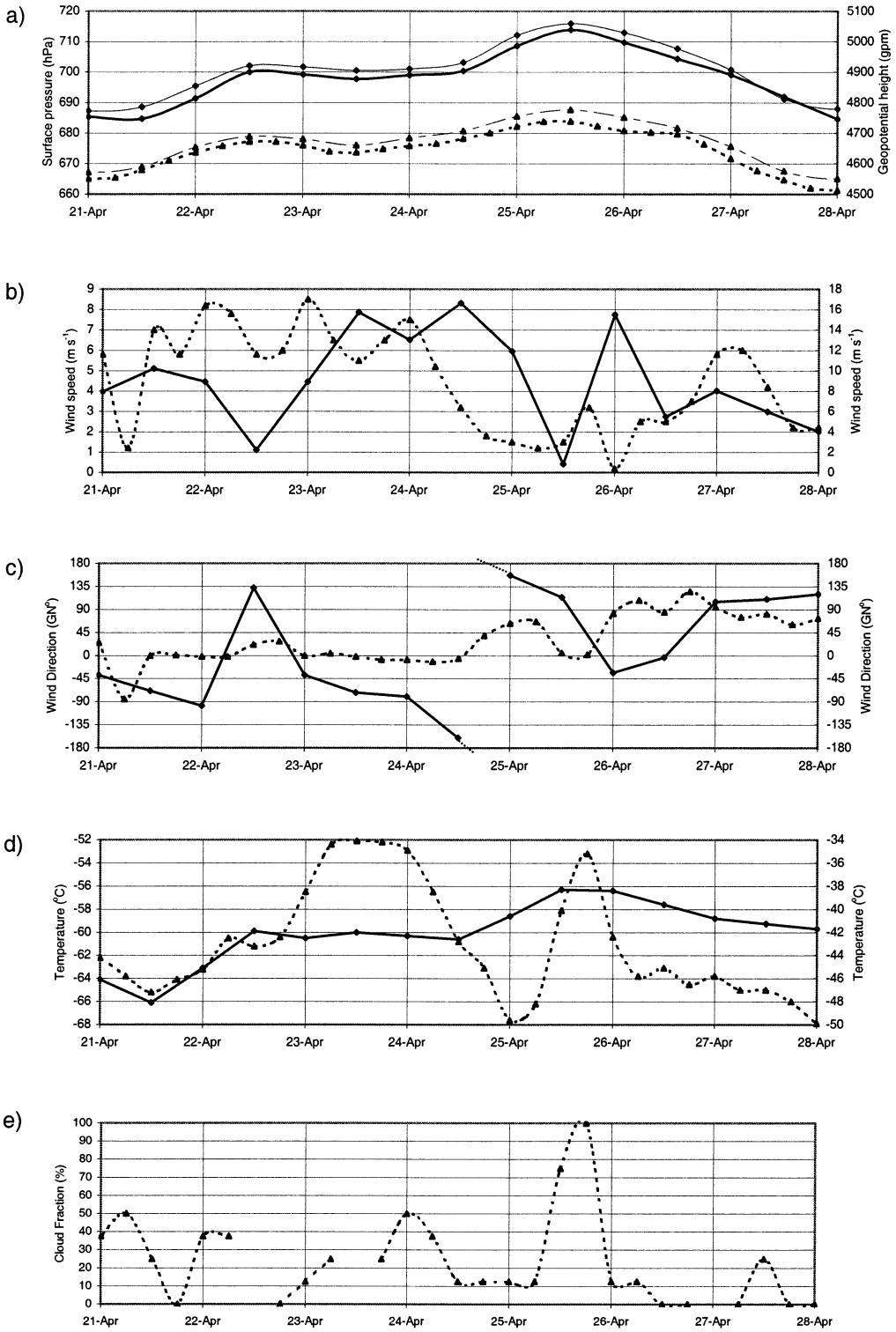

FIG. 4. Surface automatic weather station (thick dashed) and 500-hPa radiosonde (thick solid) data for (a) surface pressure (hPa)/geopotential height (gpm), (b) wind speed (m s^{-1}), (c) wind direction (° from grid north; GN°), (d) temperature (°C), and (e) cloud fraction (%; manually observed) at South Pole, 21–27 Apr 2001. Scales for AWS (radiosonde) variables are on the left- (right-) hand side of each plot. Also shown in (a) are the ECMWF–TOGA 500-hPa geopotential height (gpm; thin solid) and surface pressure (hPa; thin dashed).

radiosonde winds (Figs. 4b and 4c), which intensified and became grid northwest/southwest from approximately 0000 UTC 23 April through 1200 UTC 24 April. The height gradient associated with the ridge weakened, and a decrease in surface winds followed by a decrease in upper-level winds occurred on 24–25 April. This allowed the flight to depart Rothera for Pole at 1434 UTC 24 April. This ridge also opened a pathway for moisture advection from the Amundsen and Bellingshausen Seas toward Pole. Forecasters were concerned the moisture would reach Pole during the flight, bringing increased clouds and reducing the ceiling below the minimum requirement. However, this did not occur until midway through 25 April, and the flight safely landed at Pole at 0002 UTC 25 April. After a brief rest at Pole, the flight crew departed at 1647 UTC 25 April and returned to Rothera without incident at 0052 UTC 26 April. The winds began to increase again on 25–26 April as a low over East Antarctica (~50°E) deepened and moved poleward.

It is noteworthy that the synoptic situation here follows very closely that discussed by Neff (1999); the winds are not forced solely by the pressure/height gradient, as implied by Fig. 3. Neff argues the weather at Pole is strongly influenced by synoptic events, although the associated cross-isobar surface wind regime often leads meteorologists to erroneously assume that a given event is due to slope wind (katabatic) forcing. The author notes a very consistent bimodal wind–temperature regime associated with upper-level (synoptically forced) winds from the grid northwest and southeast directions, which is most prominent in the winter months. When upper-level winds at Pole are from grid northwest (and also from the southwest, to a lesser extent), winds throughout the atmospheric column are strong, and the winds at the surface blow in a cross-isobar direction. A deep layer of (relatively) warm air is advected poleward from the Weddell Sea region, and skies become increasingly cloudy. When upper-level winds at Pole are from grid southeast, lighter surface (and upper level) winds are observed. Associated with this is the advection of a shallow layer of cold air, and a clearing of the skies. Similarly, in this study strong upper-level winds from grid northwest (and briefly from grid southwest) are observed from 0000 UTC 23 April to approximately 1200 UTC 24 April (Figs. 4b and 4c). Associated with this event are strong surface winds (Fig. 4b) and a marked increase in surface temperature (Fig. 4d; although the 500-hPa temperature does not rise, indicating the layer is fairly shallow). Increased cloudiness is also observed (Fig. 4e). After this, upper-level winds become grid southeast and decrease in magnitude from approximately 1200 UTC 24 April through 1200 UTC 25 April. During this period, a drop in surface winds, rapid cooling, and decreased cloudiness are observed.

The extension of the ridge from the Bellingshausen Sea on 25 April was a major concern for the forecasters, who thought that it may extend to South Pole sooner than it actually did and bring inclement weather while the plane was in the air and past the point of safe return. The AFWIN MM5 and AMPS, the two operational Antarctic mesoscale models, predicted somewhat different scenarios. Initially, AFWIN MM5 did not project that this flow would reach Pole, while AMPS showed a stronger flow extending to Pole. According to the forecasters, the flow associated with the AMPS prediction was supported by satellite imagery, which showed extensive cloud streaking in the Executive Committee Range (~77°S, 125°W; Marie Byrd Land). The actual situation was a split between the two model predictions; the flow did move into South Pole on 25 April, but not with the speed and intensity predicted by AMPS.

4. Model performance

a. Surface analyses

1) RESULTS FOR ALL SURFACE SITES

Table 2a presents the means of the statistics for the 6–48-h forecasts versus observed surface wind speed, wind direction, pressure, and temperature for the average of seven sites in West Antarctica (see Fig. 1 for locations).

Observing the wind speed statistics, it is noted that AMPS has the largest positive bias (1.9 m s^{-1}) and rmse (4.0 m s^{-1}). The other models all have biases less than ± 1 m s^{-1} and rmse values less than 3.4 m s^{-1}. Overall, ECM shows the highest correlation (0.59). GLO has the lowest correlation (0.27), which does not meet the 95% confidence interval. It is noteworthy that for Rothera, a coastal site with complex topography that was included in the mean statistics, all of the models performed poorly and were negatively correlated with the observations (not shown), indicating that model topography is very important for accurately resolving surface winds along the Antarctic coast.

The models perform similarly in predicting wind direction, each with an rmse of ~60°, and small negative biases. Correlation coefficients are not calculated in this small sample size, as most of the sites demonstrate variability in excess of 180°, which yields unreliable results. This problem is generally negligible for a larger sample size (e.g., Guo et al. 2003).

The four models show skill in predicting surface pressure, all with correlations above 0.85. ECM has the highest correlation and lowest rmse indicating the greatest skill. AMPS and AVN have similar correlations and rmse values, suggesting the dependence of AMPS upon the AVN initial conditions. Recall, bias statistics are not given due to uncertainty in reported station elevations.

The surface temperature trends indicate that the models all have a slight positive bias for the mean of the seven study sites. However, if the sites are looked at individually (not shown), both negative and positive biases occur for all of the models. For example, at Theresa AWS, the smallest negative bias is −0.8°C (AVN) and

TABLE 2. Forecasts from 6 to 48 h: Statistics for the modeled vs observed surface (3 m) wind speed and wind direction, surface pressure, and surface (2 m) temperature for (a) the mean of seven AWS sites throughout West Antarctica and (b) South Pole. The results were compiled using all model initializations from 21 to 25 Apr 2001 at 6-h intervals for all forecasts 6–48 h from the initialization. Abbreviations: n = number in population, rmse = root-mean-square error; corr = correlation coefficient; TN°, from true north; GN°, from grid north; and N/A, not available. A Y in the 95% column indicates that the strength of the correlation exceeds the 95% confidence interval (N means it does not).

(a)		All sites				(b)		South Pole			
Model	n	Bias	Rmse	Corr	95%	Model	n	Bias	Rmse	Corr	95%
		3-m wind speed ($m\ s^{-1}$)						3-m wind speed ($m\ s^{-1}$)			
AMPS	390	1.9	4.0	0.39	Y	AMPS	65	2.5	3.2	0.62	Y
AVN	480	−0.4	3.2	0.34	Y	AVN	80	−1.3	2.4	0.64	Y
ECM	240	−0.7	2.4	0.59	Y	ECM	40	1.0	1.9	0.79	Y
GLO	342	0.6	3.4	0.27	N	GLO	57	1.5	2.7	0.54	Y
		3-m wind direction (TN°)						3-m wind direction (GN°)			
AMPS	390	−13.8	65.9	N/A	N/A	AMPS	65	1.3	29.2	0.74	Y
AVN	480	−9.3	59.2	N/A	N/A	AVN	80	−19.3	64.3	0.40	Y
ECM	240	−16.2	60.2	N/A	N/A	ECM	40	20.0	34.5	0.74	Y
GLO	342	−23.4	67.3	N/A	N/A	GLO	57	−1.1	41.9	0.47	Y
		Surface pressure (hPa)						Surface pressure (hPa)			
AMPS	455	N/A	3.2	0.90	Y	AMPS	65	N/A	2.2	0.95	Y
AVN	560	N/A	2.9	0.91	Y	AVN	80	N/A	1.4	0.95	Y
ECM	280	N/A	1.8	0.95	Y	ECM	40	N/A	1.4	0.96	Y
GLO	399	N/A	3.4	0.87	Y	GLO	57	N/A	2.9	0.85	Y
		2-m temperature (°C)						2-m temperature (°C)			
AMPS	455	1.5	8.2	0.32	Y	AMPS	65	4.1	6.7	0.49	Y
AVN	560	2.1	8.4	0.27	N	AVN	80	5.6	8.1	0.23	Y
ECM	280	1.2	4.4	0.44	Y	ECM	40	2.0	4.4	0.68	Y
GLO	399	2.1	7.7	0.37	Y	GLO	57	5.2	7.9	0.14	N

the largest is −7.1°C (AMPS). In this case (surface temperature), bias is a poor indicator of true model performance, and the relatively large rmse values observed for all of the models indicate the large amount of variability about the observations. Overall, ECM performs with the greatest skill, having the lowest rmse (4.4°C) and highest correlation (0.44). For all of the models, the low quality of the results implies a need for improved planetary boundary layer parameterizations in support of near-surface temperature prediction. Weaknesses may also be present in the parameterization of radiation and land surface processes.

2) RESULTS FOR SOUTH POLE

It is often instructive to examine a case at one specific point in order to gain insight into model strengths and weaknesses. Here, we choose South Pole, the landing site for the rescue flight. This discussion draws on Table 2b, which presents the model statistics for Pole, and Figs. 5a–e, which show the 6–48-h time series data of forecast surface variables (wind speed, wind direction, pressure, temperature, and cloud fraction, respectively) at South Pole versus Clean Air AWS (90°S, 0°; 2835 m). Cloud fraction is plotted versus manual cloud observations at Pole. The gray-shaded area in the first plot in each figure indicates the flight window to and from South Pole, which coincides with the winds subsiding (and is represented by the synoptic conditions depicted by the satellite composite in Fig. 1). The black bars at the top of the first plot in each figure indicate times when blowing snow was reported.

Perhaps the most important forecasting task for the landing at South Pole was the prediction of blowing snow. The models perform with varying degrees of skill for this endeavor. In predicting the wind speed variability, particularly capturing the decline in wind speed on 24–25 April, AMPS displays the greatest drop (~4 $m\ s^{-1}$), while ECM and AVN predict a more moderate drop (~2 $m\ s^{-1}$). There is a small drop (~1–2 $m\ s^{-1}$) in the GLO wind speed during this period. The lack of the winds subsiding in GLO may be related to a discrepancy at Pole in the surface topographic dataset originally used. This has since been adjusted to reflect the accurate topography. Noting the observed wind speed during the blowing snow event, a blowing snow ''wind speed threshold'' can be approximated at about 6 $m\ s^{-1}$, which is confirmed by the forecasters at McMurdo [Turner and Pendelbury (2000) also give a threshold of 6 $m\ s^{-1}$ at Pole following precipitation; it is conjectured that precipitation may have occurred at or near Pole prior this event, but very few observations are available before 20 April]. Using this

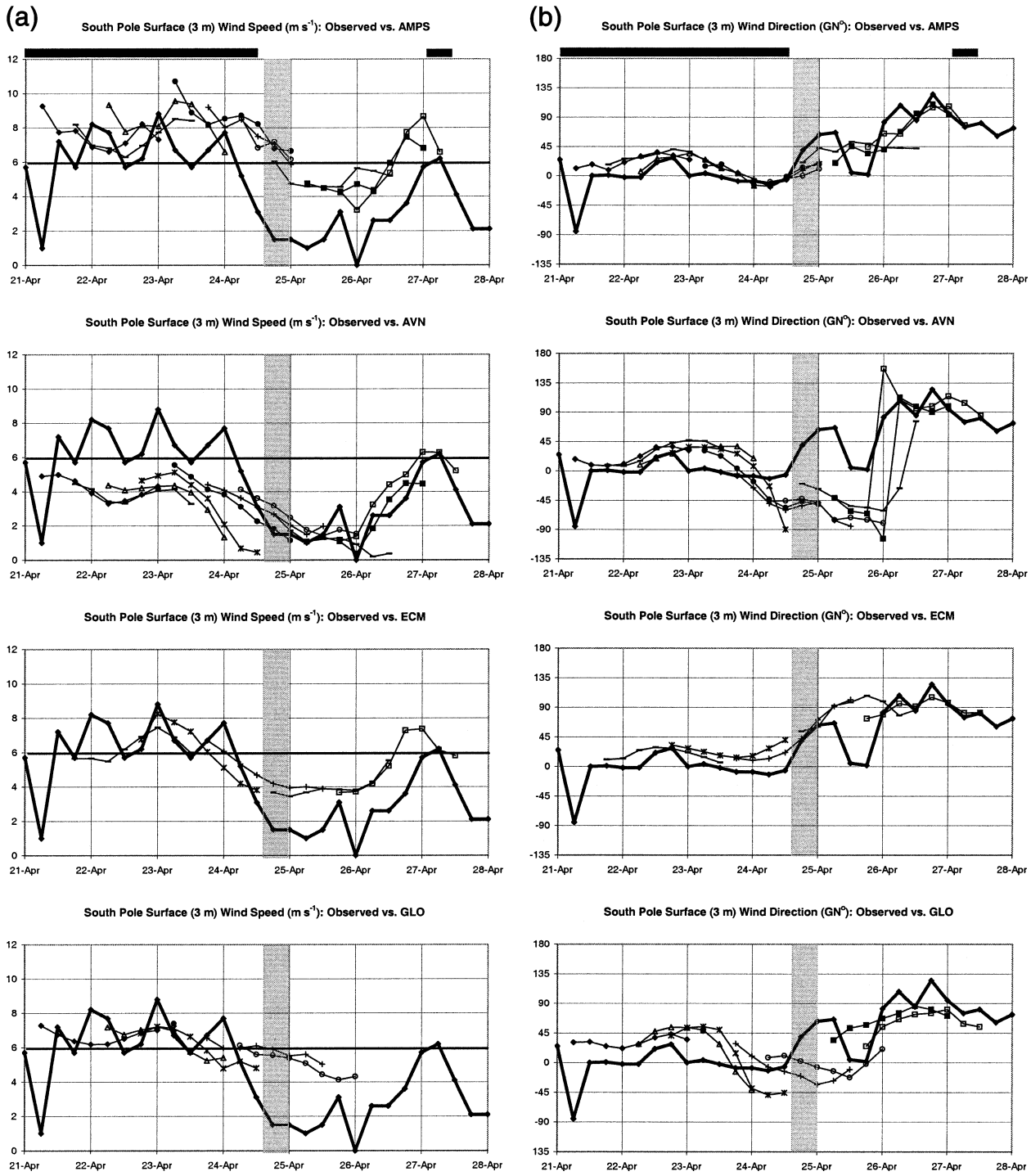

FIG. 5. (a) Observed vs forecast surface (3 m) wind speed (m s^{-1}) at South Pole, 21–27 Apr 2001, for all forecasts out to 6–48 h from the initial time. Observations (thick line) are from Clean Air AWS (90°S, 0°). The gray vertical bar in the first plot indicates the duration of the flight to Pole. The black horizontal bars at the top of the first plot represent times when blowing-snow conditions are reported at Pole. The thick horizontal line at 6 m s^{-1} indicates the approximate blowing-snow threshold. (b) Same as in (a) but for surface (3 m) wind direction (GN°). (c) Same as in (a) but for surface pressure (hPa). (d) Same as in (a) but for surface (2 m) temperature (°C). (e) Same as in (a) but for cloud fraction (%; compared against manually observed cloud fraction at South Pole).

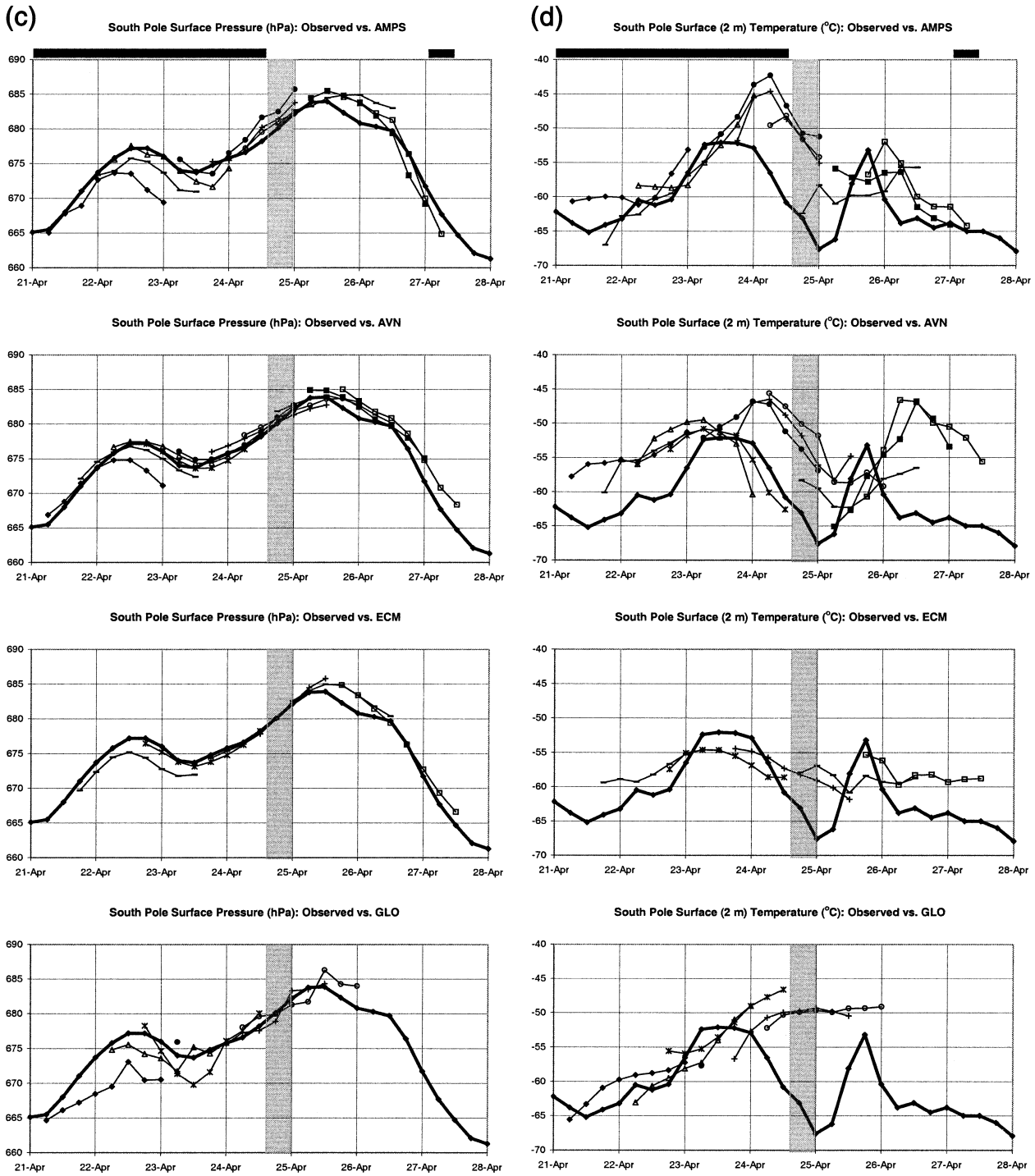

FIG. 5. (*Continued*)

general rule, AMPS and ECM would have accurately predicted the blowing snow event and its subsequent cessation on 24 April (although the AMPS timing is somewhat late). The statistics in Table 2b indicate that ECM has the lowest bias and rmse, and the highest correlation with the observed values. Statistically, AVN performs with the next highest skill, having the second-lowest bias and rmse, and the second-highest correlation with the observations. This is perhaps due to the skill shown from 25 April onward in predicting the magnitude and trend of the observations. However, observation of the plots indicates that AMPS performed with more skill in pre-

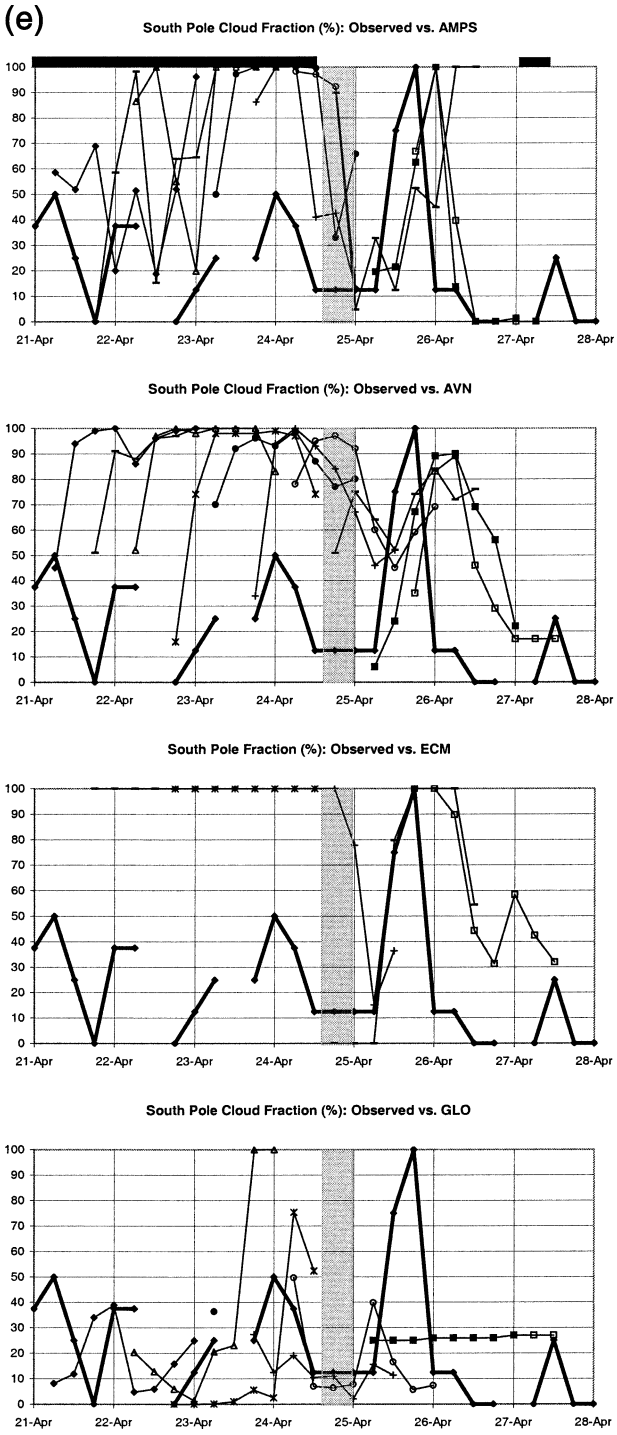

FIG. 5. (*Continued*)

dicting the high winds prior to the drop on 24–25 April, and in capturing the rather large and abrupt drop (but with a delay of ~12 h).

In predicting the wind direction at Pole (Fig. 5b), all of the models demonstrate a reasonable depiction of the large-scale circulation, capturing the general shift of winds from grid north to grid east over the time series. None of the models capture the abrupt (and brief) departure back to grid north on 25 April. The most notable departures are in the AVN forecasts on 25 and 26 April, which make substantial adjustments toward the observed values from what appear to be poor initial conditions. Similarly, Bromwich et al. (1999), in an examination of the AVN/Medium Range Forecast (MRF) model, note that large adjustments from deficient initial conditions are common for many of the poorly represented fields, and the fields generally improve as they are integrated forward in time. Statistically, AMPS showed the highest skill in forecasting wind direction, having a small bias (1.3) and the lowest rmse (29.2) (Table 2b). For this case, the low variability of the South Pole winds allows for the correlation coefficient to be calculated with a high degree of confidence. AMPS and ECM have the highest correlation (0.74); however, ECM has a larger bias and rmse. AVN and GLO have considerably lower correlations, likely reflecting their lower-resolution depiction of the topography.

For the surface pressure at South Pole (Fig. 5c), all of the models capture the variability with skill, correlating at or above 0.85 (Table 2b). AMPS, AVN, and ECM all correlate at or above 0.95. AVN and ECM have the best fit about the observations, and this is reflected in their low rmse values (1.4 hPa).

Observation of surface temperatures (Fig. 5d) indicates that all of the models have a warm bias at South Pole. This is shown in the statistics, with positive biases ranging from 2.0°C (ECM) to 5.6°C (AVN). AMPS and AVN both show anomalous warming on 23–24 April followed by cooling. While capturing variability, they tend to be out of phase with the temperature trend by 6–12 h. It is likely that the similarity of the evolution of features in AMPS to those in AVN arises because AMPS employs AVN for initial and boundary conditions. Bromwich et al. (2003) find that AMPS relies heavily on the AVN initial fields, and this behavior has been noted for other mesoscale models at high latitudes (e.g., Klein et al. 2001). The influence of the initial and boundary conditions on AMPS forecasts is explored further in section 4d. ECM shows the highest skill in predicting the magnitude of the surface temperatures, although it does not capture the period of vigorous cooling on 24 April. GLO shows minimal skill in predicting the surface temperatures (i.e., overpredicting the surface temperature by ~18°C on 25 April), and does not correlate at the 95% confidence interval.

Figure 5e shows the modeled versus observed cloud fraction (no statistics are presented for this field). With the exception of GLO, the models tend to overpredict clouds, a problem that has been reported in other studies over ice sheets and cold continental interior locations (e.g., Hines et al. 1997a,b). Generally, this causes an excess of downwelling longwave radiation and has a warming effect during the polar night. It appears that this may be partially responsible for the warmer than

TABLE 3. Forecasts from 12 to 48 h: Statistics for the modeled vs observed 500-hPa wind speed and wind direction, geopotential height, and temperature for South Pole. Conventions are the same as those used in Table 2.

Model	*n*	South Pole Bias	South Pole Rmse	Corr	95%
		500-hPa wind speed ($m\ s^{-1}$)	($m\ s^{-1}$)		
AMPS	32	0.3	4.3	0.64	Y
AVN	40	0.2	5.3	0.25	N
ECM	20	0.9	5.0	0.38	N
GLO	28	1.3	4.9	0.45	Y
		500-hPa wind direction (GN°)	(GN°)		
AMPS	32	11.6	81.8	N/A	N/A
AVN	40	41.1	97.1	N/A	N/A
ECM	20	−24.8	75.2	N/A	N/A
GLO	28	6.1	78.9	N/A	N/A
		500-hPa geopotential height (gpm)	(gpm)		
AMPS	32	35.1	43.3	0.92	Y
AVN	40	−9.1	22.8	0.95	Y
ECM	20	11.8	16.7	0.98	Y
GLO	28	33.9	40.1	0.95	Y
		500-hPa temp (°C)	(°C)		
AMPS	32	−1.2	2.0	0.69	Y
AVN	40	−1.6	2.1	0.77	Y
ECM	20	−0.9	1.6	0.69	Y
GLO	28	−3.6	3.9	0.77	Y

observed temperatures in AMPS, AVN, and ECM, for example, on 24 April (Fig. 5d), and for the overall warm bias observed in these models (Table 2b), although this cloud–temperature relationship is not apparent in GLO. All three of these models capture the cloud clearing on 24–25 April after the upper-level winds shift to grid southeast [see discussion of Neff (1999) in section 3], and then capture the increase in cloud fraction during the latter part of 25 April (although all are about 6 h late) as the upper-level winds briefly become grid north-northwest, advecting moisture from the Weddell Sea. GLO shows very little correlation with the observations, especially in relation to the cloudiness observed on 25 April, but does not tend to overpredict cloud fraction. It is noteworthy that the modeled clouds are in better agreement with the observed clouds during periods when blowing snow is not reported; it is possible that the blowing snow may be partially obscuring the observer's view during blowing snow events, leading to erroneous cloud estimates.

b. 500-hPa analyses (South Pole)

Extending the analysis at South Pole to the upper-level conditions, the models are examined at 500 hPa. This discussion draws on Table 3, which presents the 500-hPa model statistics at Pole, and Figs. 6a–d, which show the 12–48-h time series data of the 500-hPa forecast variables (wind speed, wind direction, geopotential height, and temperature) versus radiosonde observations.

Examining Fig. 6a, AMPS shows the highest skill in capturing the magnitude and variability of the 500-hPa wind speed, most notably the rise in wind speed associated with a shift in the wind direction to grid northwest on ~23 April, then the drop in winds on ~25 April as the wind direction shifts back to grid southwest, consistent with the synoptic discussion in section 3. ECM and GLO also capture the rise in winds on ~23 April, but do not resolve the decrease in winds on 25 April. Little skill is demonstrated by the AVN in capturing the magnitude or variability of the wind speed. The statistics reflect this performance, with AMPS having the highest correlation (0.64), lowest rmse, and a negligible bias. GLO outperforms ECM, having a lower rmse and higher correlation. ECM and AVN do not correlate within the 95% confidence interval.

Inspection of Fig. 6b shows a wide range of wind directions predicted by the models, and this is reflected in the large rmse values in Table 3, all greater than 70°. Note that the wind directions here are plotted from 0° to 360° for presentation, but are the same as the 500-hPa winds plotted from −180° to 180° in Fig. 2c. All of the models predict (to varying degrees) the gradual shift from grid northwest to grid southeast on 23–25 April, which is important in resolving the bimodal pattern discussed in Neff (1999; see section 3 in this paper). AVN, which did not resolve the decrease in winds associated with this shift in direction to southeast, only shows an ~90° shift in wind direction, from grid northwest to grid southwest. Accordingly, AVN has the largest bias and rmse. ECM, about 45° out of phase, tends to capture this shift with the most skill. The rapid (and brief) departure in wind direction to grid north-northwest on ~26 April is not resolved by any of the models.

Inspection of Fig. 6c reveals that all of the models resolve the geopotential height trends (i.e., large-scale circulation) with reasonable accuracy. Accordingly, all have correlations above 0.90, and ECM is nearly unity (0.98). AMPS and GLO, the two MM5 models, both have positive biases in excess of 30 gpm. The magnitude of the rmse is nearly the same as the bias, which implies that the bias is *systematically* high in all forecasts; this is evident in Fig. 6c for AMPS and GLO.

The forecast temperatures (Fig. 6d) correlate at 0.69 and above for all models, with AVN and GLO having the highest (0.77). In general, the models perform similarly in resolving the variability. All tend to underpredict the warming on 25–26 April, apparently related to the failure of each to capture the brief shift in winds to grid north-northwest discussed above. Inspection indicates all of the models have negative biases. ECM has the smallest bias (−0.9), and GLO the highest (−3.6).

c. 700-hPa analyses (flight path)

Table 4a presents the means of the statistics for the 12–48-h forecast versus observed (ECMWF–TOGA)

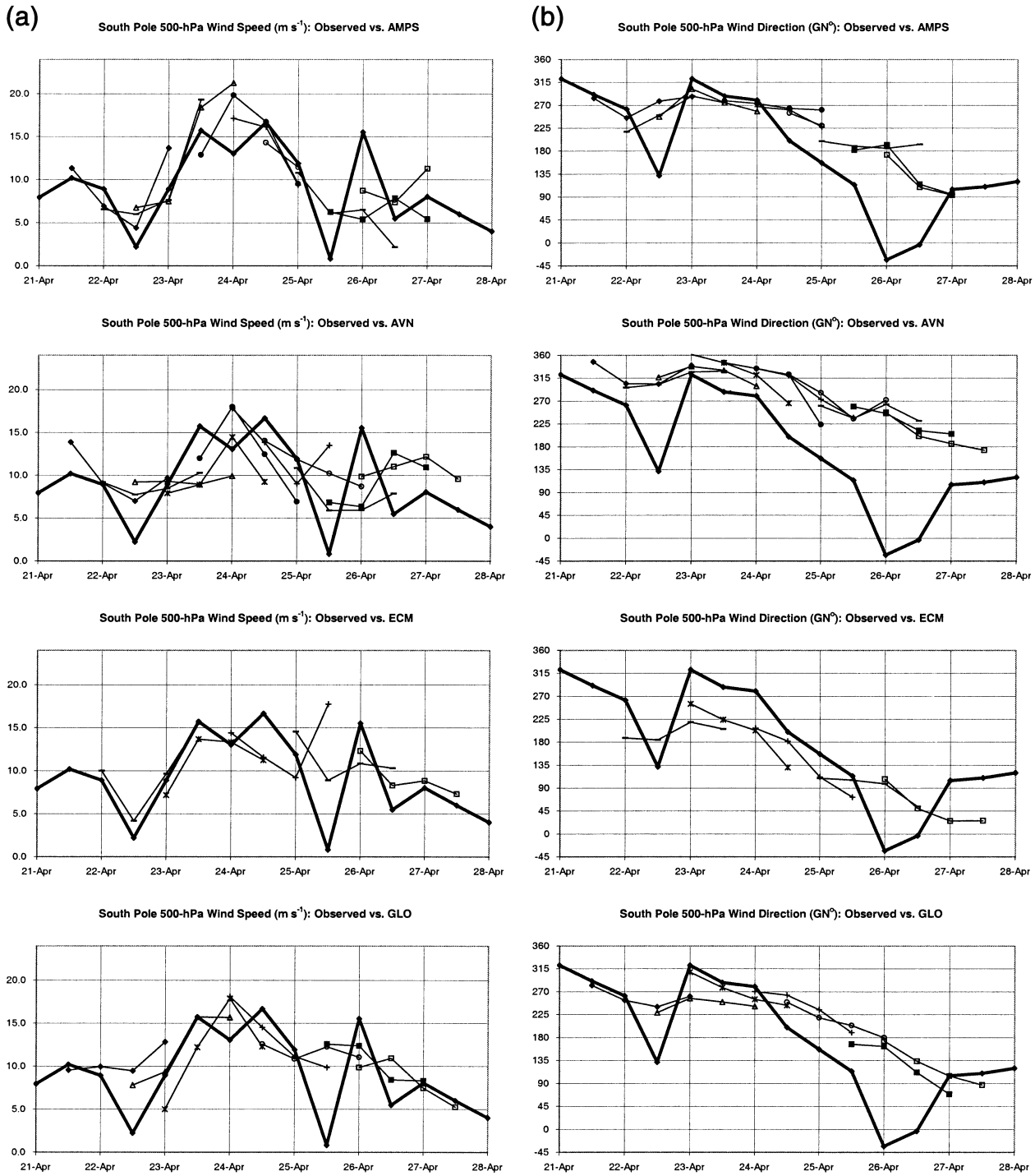

FIG. 6. (a) Observed vs forecast 500-hPa wind speed (m s^{-1}) at South Pole, 21–27 Apr 2001, for all forecasts out to 12–48 h from the initial time. Observations (thick line) are from radiosonde observations at Pole. (b) Same as in (a) but for 500-hPa wind direction (GN°). (c) Same as in (a) but for 500-hPa geopotential height (gpm). (d) Same as in (a) but for 500-hPa temperature (°C).

700-hPa wind speed, wind direction, geopotential height, and temperature for five points along 70°W, the longitude of the flight path (see Fig. 1). Overall, the models show much more skill in depicting variables at 700-hPa than they do for the surface, at least partially due to the absence of the stable boundary layer and the smaller influence of topography.

Inspection of the wind speed statistics in Table 4a indicates that for all of the models the correlations are greater than 0.60. Likewise, the biases are all within

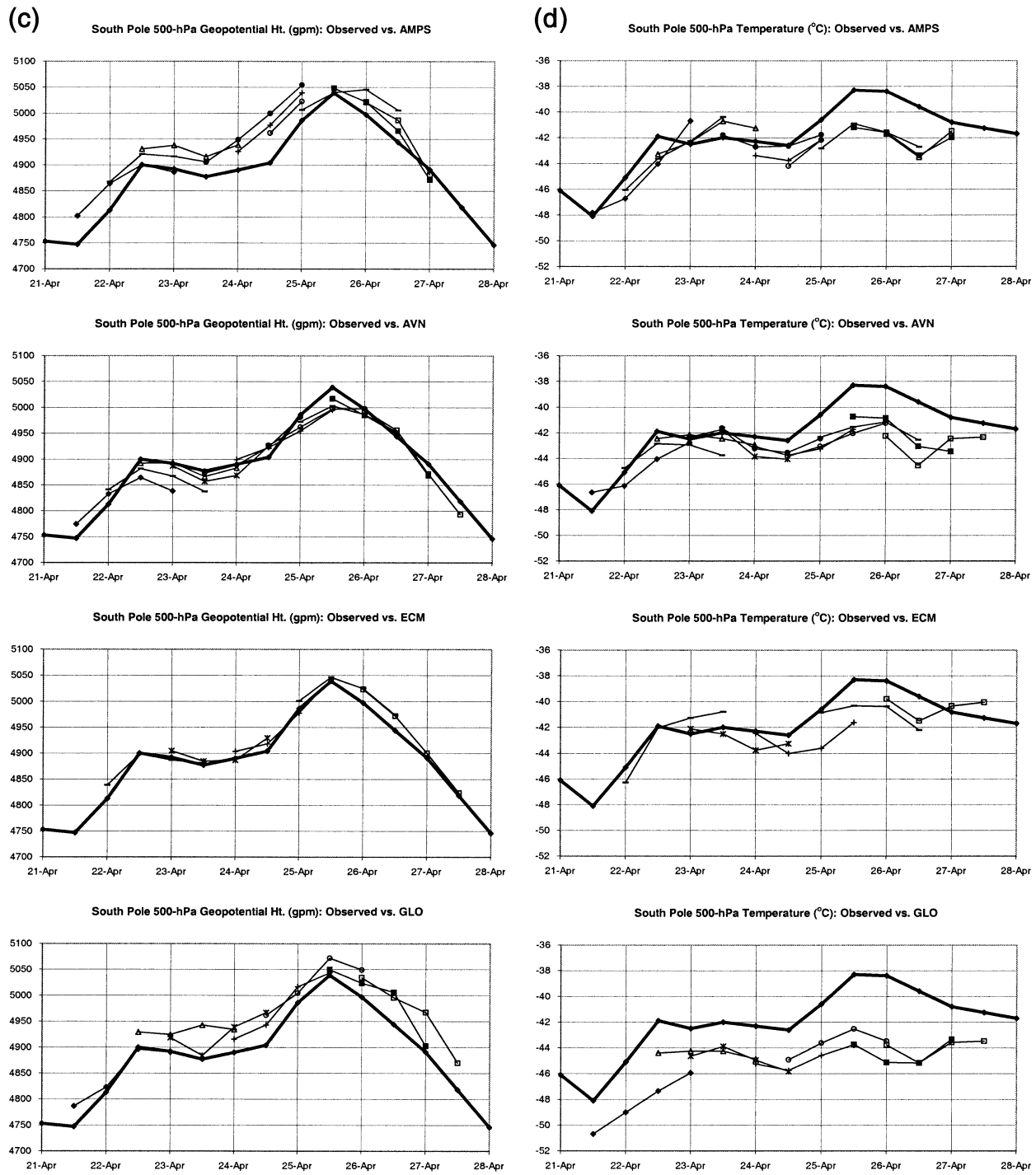

FIG. 6. (*Continued*)

± 1 m s^{-1} of the ECMWF–TOGA wind speed, which is estimated to have an uncertainty of ± 1 m s^{-1} (see section 2). The rmse values for all models are within 3.3 m s^{-1}. ECM has the highest correlation and lowest rmse for all five sites, suggesting the highest skill in predicting the wind speed variability (although it must be considered that ECM is being compared with its own analysis). Surprisingly, the lowest-resolution model, GLO, has the second highest correlation.

The wind direction statistics show all of the models performing with similar skill. The bias and rmse statistics at 700 hPa indicate better model performance than at the surface (Table 2a), likely due to the decreased topographic forcing with distance from the model sur-

TABLE 4. Forecasts from 12 to 48 h: Statistics for the modeled vs observed 700-hPa wind speed and wind direction, geopotential height, and temperature for the mean of five points along the flight path from Rothera to South Pole. Conventions are the same as those used in Table 2.

		All sites			
Model	*n*	Bias	Rmse	Corr	95%
		700-hPa wind speed ($m\ s^{-1}$)	($m\ s^{-1}$)		
AMPS	160	0.1	3.0	0.68	Y
AVN	200	−0.9	3.3	0.64	Y
ECM	100	−0.2	2.6	0.79	Y
GLO	140	0.4	2.8	0.72	Y
		700-hPa wind direction			
		(TN°)	(TN°)		
AMPS	160	−2.1	32.9	N/A	N/A
AVN	200	3.1	36.0	N/A	N/A
ECM	100	8.7	32.3	N/A	N/A
GLO	140	−2.9	40.7	N/A	N/A
		700-hPa geopotential height (gpm)	(gpm)		
AMPS	160	−2.0	22.5	0.97	Y
AVN	200	−5.3	19.0	0.97	Y
ECM	100	5.5	13.3	0.99	Y
GLO	140	24.8	35.8	0.95	Y
		700-hPa temp			
		(°C)	(°C)		
AMPS	160	−1.3	2.1	0.87	Y
AVN	200	−1.3	2.1	0.88	Y
ECM	100	−0.6	1.3	0.94	Y
GLO	140	−2.2	2.8	0.85	Y

face. Correlation statistics are not computed because of the high variability of the wind directions.

For the geopotential height, the correlations for all sites for all models are at or above 0.95. The biases for AMPS, AVN, and ECM are all less than ±6 gpm, while GLO indicates a strong positive bias of about 25 gpm. The rmse values are about 20 gpm for AMPS and AVN, about 13 gpm for ECM, and about 35 gpm for GLO. When considering the ECMWF–TOGA data against which the models are validated is probably within ±15 gpm of the true geopotential height, AMPS, AVN, and ECM all perform well, both qualitatively and quantitatively. GLO, while performing well qualitatively, has a strong positive geopotential height bias, even considering the margin of error of the validating data.

For the temperature, the correlations for all sites for all models are at or above 0.85. Negative biases are observed for all models, ranging from −0.6 to −2.2°C. The rmse values for all models are about 2°C. Considering that the ECMWF–TOGA data may differ from the radiosonde observations on the order of ±2°C [and probably are positively biased along the flight path (Fig. 2)], this suggests that all of the models, when averaged over several sites, are in reasonable quantitative agreement with the observed temperature.

d. Additional runs using ECMWF–TOGA

Three additional AMPS runs were performed for the 0000 and 1200 UTC 23 April and 0000 UTC 24 April initializations using ECMWF–TOGA (analysis) data to depict the initial and boundary conditions (ICs/BCs). These data were used because they are derived from the initial fields of the much higher resolution ECMWF global forecast model (ECM). Simmons and Hollingsworth (2002) show that forecast errors in the ECMWF global forecast model are small after 24 h when validated against its own analyses, and therefore we consider ECMWF–TOGA to be representative of the forecast fields. The ECMWF global forecast data could not be directly used because of the high cost of obtaining all of the necessary fields. In addition, AMPS is already configured to assimilate ECMWF–TOGA, so it is convenient to use these data. It should be noted that the ECMWF–TOGA data have a lower spatial resolution than AVN, which is currently used for the AMPS ICs/BCs, and the temporal resolution is every 12 h as opposed to 6 h for AVN (this influences the lateral boundary conditions, which are imposed throughout the model run). Therefore, while it is expected that the quality of the ECMWF–TOGA data is an improvement over the AVN data, limitations are imposed by the low spatial and temporal resolution, which would not be present if the ECMWF global forecast data were used. It is noteworthy that the ECMWF global forecast model generally performed with the highest skill in this study, and has demonstrated the best performance when compared to similar global models, including AVN (Skinner 1995; Skinner and Hart 2001). Accordingly, the use of this model to derive the ICs/BCs in AMPS may be a potential improvement over AVN.

Figure 7 compares the results of the AMPS runs using ECMWF–TOGA ICs/BCs (Fig. 7a, AMP_NEW) to the original runs, which employ AVN ICs/BCs (Fig. 7b, AMPS). Note that the two latter runs using AVN ICs/BCs (Fig. 7b) are cut short at 0000 UTC 25 April; the cause of this is unknown. Inspection of the surface winds indicates that the runs using the ECMWF–TOGA ICs/BCs have a positive bias similar to that from the runs using AVN ICs/BCs, with a greater bias during the peak of winds at 0000 UTC 24 April. The runs using the AVN ICs/BCs capture the wind direction better than those using the ECMWF–TOGA ICs/BCs. Similar to the results for the other models shown in Fig. 5b, the new runs fail to capture the brief shift in wind direction on 25 April. Inspection of the surface pressure plots indicates that both cases correlate well with the pressure trend, with no significant difference in the runs using ECMWF–TOGA ICs/BCs from the runs using AVN ICs/BCs. Inspection of the surface temperature forecasts indicates improvement in the runs using the ECMWF–TOGA ICs/BCs. The runs using ECMWF–TOGA ICs/

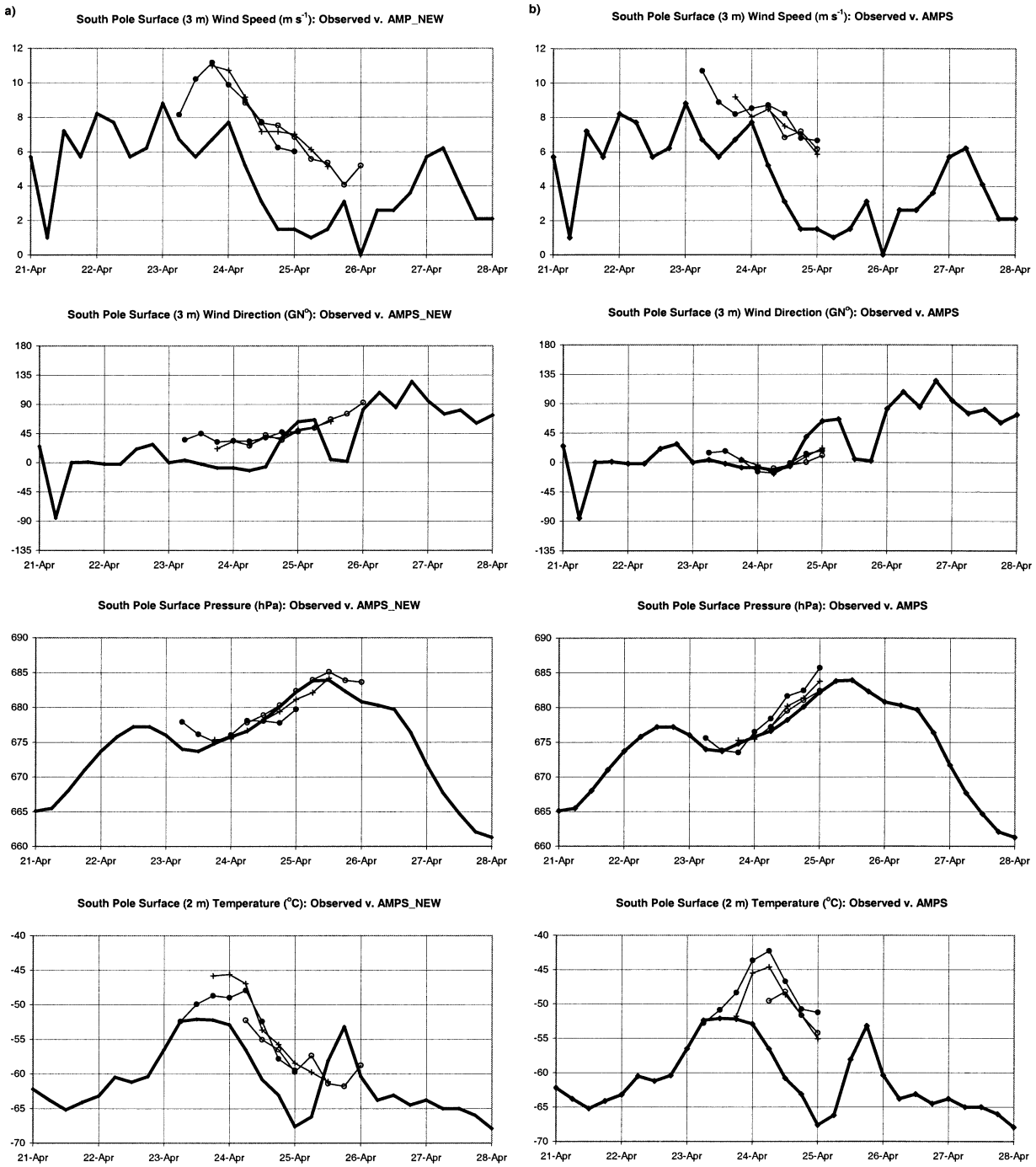

FIG. 7. Surface wind speed (m s^{-1}), wind direction (GN°), temperature (°C), and pressure (hPa) at South Pole for (a) AMPS_NEW (runs using ECMWF–TOGA initial and boundary conditions) and (b) AMPS (original runs using AVN initial and boundary conditions).

BCs correlate more closely to the observed temperatures, and do not reproduce the anomalously high temperatures on 24 April that are observed in the runs using the AVN ICs/BCs.

The new simulations demonstrate that the ICs/BCs exert a significant influence on AMPS forecasts. However, with the exception of the surface temperature prediction, they do not indicate that ECMWF–TOGA ICs/BCs improve simulations compared to AVN ICs/BCs (although it must be considered that the ECMWF–TOGA ICs/BCs have a lower spatial and temporal resolution than the AVN ICs/BCs).

5. Discussion

Addressing some model deficiencies

1) Positive surface temperature bias

All of the models have a positive surface temperature bias. The bias ranges from 1.2°C (ECM) to 2.1°C (AVN, GLO) for the average of all sites, and is much larger for South Pole, ranging from 2.0°C (ECM) to 5.6°C (AVN), suggesting that the bias is related to the continental ice sheet environment at South Pole. AMPS, AVN, and ECM generally overpredict cloud fraction (Fig. 5e), which contributes to surface warming via an increase in downwelling longwave radiation, a problem that is prominent during austral winter over the Antarctic ice sheet for many mesoscale and global models (e.g., Hines et al. 1997a,b). This has been addressed in the Polar MM5 (the model used in AMPS) by replacing the Fletcher (1962) equation for ice nuclei concentration with that of Meyers et al. (1992) in the MM5 Reisner explicit microphysics parameterization (Reisner et al. 1998). Using this parameterization in a 1-yr validation of nonhydrostatic Polar MM5 simulations over Antarctica, Guo et al. (2003) find that the spatial variability of cloud cover is well represented, but there is slightly deficient cloud cover over the composite annual mean of several stations due to low-level dry biases. They also note a cold bias near the interior Antarctic surface that may be related to the underestimated clouds. Assuming the Polar MM5 simulations of Guo et al. (2003) and the AMPS (Polar MM5) simulations are similar, these findings suggest that the case examined here may not be representative of typical wintertime cloud conditions. It is noteworthy that the ICs/BCs for the simulations in Guo et al. (2003) are derived from ECMWF–TOGA, whereas AMPS employs AVN ICs/BCs. The AVN/MRF model was found to have significant adjustment in cloud fraction throughout first few days of each forecast, suggesting poor initial conditions (Bromwich et al. 1999). This may be partially responsible for the higher than observed temperature and cloud cover in these AMPS forecasts.

Another factor that may contribute to the temperature bias is a positive wind speed bias. With respect to the latter, during the period in which the positive temperature bias is greatest (~24–25 April; Fig. 5d), there is a positive wind speed bias in three of the four models (Fig. 5a), which may cause enhanced vertical mixing.

2) Positive surface wind speed bias in AMPS

Table 2 and Fig. 5a reveal that AMPS has a wind speed bias over the mean of all sites, and at Pole. For the composite mean of 28 stations in Antarctica simulated using the Polar MM5, Guo et al. (2003) find a small positive wind bias (~0.60 m s^{-1}) for April 1993. When examining the synoptic variability during winter (July) at a similar interior Antarctic site (Dome C; 75.12°S 123.37°E; 3250 m), the findings indicate a tendency to underpredict strong wind events (>6 m s^{-1}) and overpredict light winds. From Fig. 5a, most of the positive bias at Pole occurs during the period of light winds on 24–25 April, which is consistent with the latter assessment of Guo et al. (2003). The 500-hPa wind speed forecasts at Pole (Table 3, Fig. 6a) represent approximately geostrophic conditions. A negligible bias is present in the AMPS 500-hPa winds, especially during the period of the largest positive surface wind speed bias (light surface winds) on 24–25 April. This indicates that the large-scale synoptic forcing is approximately correct, and is not contributing to the surface wind speed bias in AMPS at Pole. It is possible then that the problem lies in the planetary boundary layer parameterization, and this deserves further evaluation.

6. Conclusions

This study is intended to inform the reader about several of the models commonly used in Antarctic weather prediction and examine their performance in a notable forecasting event. Although it is unwise to draw definitive conclusions based on the short time frame and the few forecasts analyzed, the results are useful in assessing the models' strengths and weaknesses.

While not a unique conclusion, it is noteworthy that the horizontal spatial resolution of each model exerts significant influence on forecast accuracy for these Antarctic predictions. ECM and AVN, because of their projections, increase in resolution with respect to longitude as they near the Pole. For example, at 75°S, the longitudinal resolution of ECM is ~28 km, and ~56 km for AVN, and increases poleward. Considering this, the order of highest to lowest horizontal spatial resolution is also the order of highest to lowest overall skill: 1) ECM, 2) AMPS, 3) AVN, and 4) GLO. However, model performance is not limited to this aspect alone; data assimilation, vertical resolution, and model physics also have important impacts. Other comparisons have demonstrated the high skill of ECM relative to other models (Skinner 1995; Skinner and Hart 2001). Simmons and Hollingsworth (2002) examine some of the major reasons behind the success of this model, which they show has improved significantly throughout the Southern Hemisphere in the past 10 years, now approaching the skill achieved in the Northern Hemisphere. This is in part due to important model enhancements that benefit the data-sparse high latitudes, such as the implementation of four-dimensional variational data assimilation in 1997 (4DVAR; Mahfouf and Rabier 2000) and the continually updated assimilation of satellite infrared and microwave radiances from polar-orbiting satellites.

At the surface, ECM exhibits the best performance, generally having the lowest values for bias and rmse, and the highest correlations. All four models demonstrate high skill in predicting surface pressure (correlations > 0.85), but moderate skill in predicting wind

speed, wind direction, and temperature. AMPS has a systematic positive wind speed bias that is amplified during the period of light winds. The large-scale pressure gradient forcing is examined as a possible cause of this and is shown to be a negligible contributor. All models capture the shifts in surface wind direction at South Pole associated with upper-level forcing, with AMPS and ECM showing the highest skill. Surface temperature is the most poorly simulated variable. The models demonstrate high variability, evident by the large rmse values and low correlations. Additionally, all of the models yield a positive surface temperature bias. This is partially caused by overprediction of cloud fraction over the ice sheet.

All of the models' forecasts have considerably more skill in the free atmosphere, in part due to decreased topographic effects and less dependence on PBL parameterizations; this is demonstrated in the 500-hPa geopotential height and temperature forecasts at South Pole, for which the large-scale circulation is accurately depicted by all models. However, all models exhibit a negative 500-hPa temperature bias compared to radiosonde observations. The highest skill in forecasting the 500-hPa winds, which are shown to have a strong relation to the surface at Pole, is shown by AMPS. Global MM5 and AMPS forecasts exhibit positive geopotential height biases greater than 30 gpm.

ECM exhibits the best performance at 700 hPa, generally having the lowest values for bias and rmse, and the highest correlations; this model is slightly favored by using its analysis (ECMWF–TOGA) to represent the ''observed'' state. All of the models show generally high skill in predicting geopotential height, and good skill in predicting wind speed, wind direction, and temperature within the uncertainties of the validating data. One exception is the geopotential height forecasts for GLO, which exhibit a strong positive bias (~25 gpm) and large rmse (~35 gpm) for the mean of the five points examined.

Similar trends between AMPS and AVN in many of the forecast variables reinforce the findings of Bromwich et al. (2003), who suggest a significant dependence of AMPS on the AVN initial and boundary conditions. This is explored further by performing additional simulations with AMPS using ECMWF–TOGA initial and boundary conditions. The new simulations demonstrate that these exert an important influence on AMPS forecasts. However, with the exception of the surface temperature prediction, the results do not indicate that the simulations using ECMWF–TOGA are a substantial improvement over simulations using AVN, although it must be considered that the ECMWF–TOGA data have a lower spatial and temporal resolution than the AVN output.

Acknowledgments. The National Science Foundation (Grants OPP-9725730, OPP-9905381, ATM-9820037 Amendment No. 41, and UCAR Subcontract S01-22961) and the National Aeronautics and Space Administration (Grant NAG5-9518) sponsored this research. Special gratitude is due to John Cassano at the Cooperative Institute for Research in Environmental Sciences at the University of Colorado for providing software for the manipulation of modeled time series data. Data were obtained from the Antarctic Meteorological Research Center at the University of Wisconsin—Madison, the Arctic and Antarctic Research Center at Scripps Institution of Oceanography, the British Antarctic Survey, the National Center for Atmospheric Research, and the European Centre for Medium-Range Weather Forecasts.

REFERENCES

Bromwich, D. H., and J. J. Cassano, 2001: Meeting summary: Antarctic Weather Forecasting Workshop. *Bull. Amer. Meteor. Soc.*, **82,** 1409–1413.

——, R. I. Cullather, and R. W. Grumbine, 1999: An assessment of the NCEP operational global spectral model forecasts and analyses for Antarctica during FROST. *Wea. Forecasting,* **14,** 835–850.

——, J. J. Cassano, T. Klein, G. Heinemann, K. M. Hines, K. Steffen, and J. E. Box, 2001: Mesoscale modeling of katabatic winds over Greenland with the Polar MM5. *Mon. Wea. Rev.,* **129,** 2290–2309.

——, A. J. Monaghan, J. J. Powers, J. J. Cassano, H. Wei, Y.-H. Kuo, and A. Pellegrini, 2003: Antarctic Mesoscale Prediction System (AMPS): A case study from the 2000–01 field season. *Mon. Wea. Rev.,* **131,** 412–434.

Budd, W. F., W. R. J. Dingle, and U. Radok, 1966: The Byrd snow drift project: Outline and basic results. *Studies in Antarctic Meteorology,* M. J. Rubin, Ed., Antarctic Research Series, Vol. 9, Amer. Geophys. Union, 71–134.

——, D. Jenssen, and U. Radok, 1971: Derived physical characteristics of the Antarctic ice sheet. Meteorology Dept. Publ. 18, University of Melbourne, Melbourne, Australia, 178 pp.

Cassano, J. J., L. Li, and D. H. Bromwich, 2000: BPRC mesoscale numerical weather prediction during the 1999/2000 Antarctic field season. *Antarctic Weather Forecasting Workshop,* Byrd Polar Research Center Misc. Publ. 419, Columbus, OH, 56–59.

——, J. E. Box, D. H. Bromwich, L. Li, and K. Steffen, 2001: Verification of Polar MM5 simulations of Greenland's atmospheric circulation. *J. Geophys. Res.,* **106,** 13 867–13 890.

Cullather, R. I., D. H. Bromwich, and R. W. Grumbine, 1997: Validation of operational numerical analyses in Antarctic latitudes. *J. Geophys. Res.,* **102,** 13 761–13 784.

Dudhia, J., and J. F. Bresch, 2002: A global version of the PSU–NCAR Mesoscale Model. *Mon. Wea. Rev.,* **130,** 2989–3007.

Fletcher, N. H., 1962: *Physics of Rain Clouds.* Cambridge University Press, 386 pp.

Guo, Z., D. H. Bromwich, and J. J. Cassano, 2003: Evaluation of Polar MM5 simulations of Antarctic atmospheric circulation. *Mon. Wea. Rev.,* **131,** 384–411.

Hines, K. M., D. H. Bromwich, and R. I. Cullather, 1997a: Evaluating moist physics for Antarctic mesoscale simulations. *Ann. Glaciol.,* **25,** 282–286.

——, ——, and Z. Liu, 1997b: Combined global climate model and mesoscale model simulations of Antarctic climate. *J. Geophys. Res.,* **102,** 13 747–13 760.

Kanamitsu, M., 1989: Description of the NMC global data assimilation and forecast system. *Wea. Forecasting,* **4,** 334–342.

——, and Coauthors, 1991: Recent changes implemented into the global forecast system at NMC. *Wea. Forecasting,* **6,** 425–435.

Kiehl, J. T., J. J. Hack, G. B. Bonan, B. A. Boville, B. P. Briegleb, D. L. Williamson, and P. J. Rasch, 1996: Description of the

NCAR Community Climate Model (CCM3). NCAR Tech. Note NCAR/TN-420+STR, CO, 152 pp.

Klein, T., G. Heinemann, D. H. Bromwich, J. J. Cassano, and K. M. Hines, 2001: Mesoscale modeling of katabatic winds over Greenland and comparisons with AWS and aircraft data. *Meteor. Atmos. Phys.,* **78,** 115–132.

Leonard, S., J. Turner, and S. Milton, 1997: An assessment of UK Meteorological Office numerical weather prediction analyses and forecasts for the Antarctic. *Ant. Sci.,* **9,** 100–109.

Mahfouf, J.-F., and F. Rabier, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. II: Experimental results with improved physics. *Quart. J. Roy. Meteor. Soc.,* **126,** 2991–3012.

Meyers, M. P., P. J. DeMott, and W. R. Cotton, 1992: New primary ice-nucleation parameterizations in an explicit cloud model. *J. Appl. Meteor.,* **31,** 708–721.

Neff, W. D., 1999: Decadal timescale trends and variability in the tropospheric circulation over the South Pole. *J. Geophys. Res.,* **104,** 27 217–27 252.

Phillpot, H. R., 1997: Some observationally-identified meteorological features of East Antarctica. Meteorological Study 42, Bureau of Meteorology, Melbourne, Australia, 275 pp.

Reisner, J., R. M. Rasmussen, and R. T. Bruintjes, 1998: Explicit forecasting of supercooled liquid water in winter storms using the MM5 mesoscale model. *Quart. J. R. Meteor. Soc.,* **124,** 1071–1107.

Simmons, A., and A. Hollingsworth, 2002: Some aspects of the improvement in skill of numerical weather prediction. *Quart. J. Roy. Meteor. Soc.,* **128,** 647–677.

Skinner, W., 1995: Numerical prediction model performance summary April to June 1995. *Aust. Meteor. Mag.,* **44,** 309–312.

——, and T. Hart, 2001: Numerical prediction model performance summary October to December 2000. *Aust. Meteor. Mag.,* **50,** 149–157.

Slingo, J. M., 1987: The development and verification of a cloud prediction model for the ECMWF model. *Quart. J. Roy. Meteor. Soc.,* **113,** 899–927.

Stull, R. B., 1988: *An Introduction to Boundary Layer Meteorology.* Kluwer Academic, 666 pp.

Turner, J., and S. Pendlebury, Eds., 2000: *The International Antarctic Weather Forecasting Handbook.* British Antarctic Survey, 704 pp.