

**Improved Characterization of Errors within the Antarctic Mesoscale
Prediction System Through Utilization of CONCORDIASI Dropsondes**

James O. Russell

UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

IMPROVED CHARACTERIZATION OF ERRORS WITHIN
THE ANTARCTIC MESOSCALE PREDICTION SYSTEM THROUGH
UTILIZATION OF CONCORDIASI DROPSONDES

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE IN METEOROLOGY

By

JAMES RUSSELL
Norman, Oklahoma
2014

IMPROVED CHARACTERIZATION OF ERRORS WITHIN
THE ANTARCTIC MESOSCALE PREDICTION SYSTEM THROUGH
UTILIZATION OF CONCORDIASI DROPSONDES

A THESIS APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY

Dr. David Parsons, Co-Chair

Dr. Steven Cavallo, Co-Chair

Dr. Lance Leslie

© Copyright by JAMES RUSSELL 2014
All Rights Reserved.

Acknowledgments

I would like to thank my advisors, Dr. David Parsons and Dr. Steven Cavallo for their guidance, patience, and wisdom throughout this research. I would also like to thank Dr. Lance Leslie for his helpful input. I would like to thank the Arctic and Antarctic Atmospheric Research Group for their feedback throughout, especially Andrew Dzambo for his assistance in implementing the radiosonde humidity correction scheme. Further I would like to thank my family and friends for their support, and assistance during this research.

Contents

Acknowledgments	iv
List Of Tables	vii
List Of Figures	viii
Abstract	x
1 Introduction and Background	1
1.1 Motivation	2
1.1.1 Antarctica	2
1.1.2 Numerical Weather Prediction	4
1.1.3 Forecasting Applications	7
1.2 The Antarctic Mesoscale Prediction System (AMPS)	8
1.3 The Concordiasi Project	10
2 Model and Observational Details	13
2.1 AMPS, the Global Forecast System (GFS), and ERA-Interim	13
2.2 Dropsonde data	16
3 Methodology	18
3.1 Interpolation Techniques	18
3.2 Statistics	19
3.3 Lead time	22
3.4 Subsets based on Antarctic Environments	22
4 Observed Atmospheric Characteristics in the Concordiasi Data Set	25
4.1 Observed Characteristics	25
4.2 Discussion	30
5 AMPS Forecast Performance	33
5.1 Continental High Plateau (HP)	33
5.2 Continental Low Elevations (LE)	36
5.3 Permanent Ice Shelves and Sea Ice (IC)	38

6	In Depth Analysis of AMPS Biases	41
6.1	Continental High Plateau (HP)	41
6.2	Continental Low Elevations (LE)	44
6.3	Permanent Ice Shelves and Total Sea Ice Cover (IC)	48
6.4	Inversions	50
6.5	Low-level Wind Maxima	56
6.6	Discussion	60
7	A Comparison between AMPS, GFS and ERA-Interim	66
7.1	Continental High Plateau (HP)	66
7.2	Continental Low Elevations (LE)	72
7.3	Permanent Ice Shelves and Sea Ice (IC)	77
7.4	Discussion	83
8	Conclusions	87
	Reference List	89

List Of Tables

2.1	AMPS Parameterization Schemes during Concordiasi	14
2.2	Vaisalla RS92 sensor accuracy (Vaisalla, 2013).	17
3.1	Interpolation Schemes used in the evaluation of AMPS.	19
3.2	Sample sizes for the various subsets. Forecast subset sizes are those without parentheses and 0 hour subset sizes are in parentheses. . . .	24
4.1	Mean temperature inversion strength and depth, number of inversions present in the subset, and number of inversions that are surface based. Characteristics are presented for day first and then night (e.g. day/night).	28
4.2	Mean LLWM maximum wind speed and height, and number of LLWMs present in the subset. Characteristics are presented for day first and then night (e.g. day/night).	30

List Of Figures

1.1	A terrain map of Antarctica with October 2010 mean sea ice extent shown in white and open water shown in blue. Antarctic radiosonde sites (as of 2010) in pink and Concordiasi dropsondes separated by the surface type over which they dropped are shown. Surface types include continental high plateau (HP) in cyan, continental low elevation (LE) in blue, sea ice or permanent ice shelves (IC) in green, Partial Sea Ice Cover (PI) in brown, and Open Water (OW) in red. Triangles represent profiles during the day and squares represent those at night.	12
2.1	AMPS grid structure during September-December 2010. Source: Mesoscale & Microscale Meteorology Division, University Center for Atmospheric Research (2014).	15
4.1	Vertical profiles of temperature and dew point for (a) HP and day, (b) HP and night, (c) LE and day, (d) LE and night, (e) IC and day, and (f) IC and night. For HP and LE subsets (i.e. a,b,c, and d) only the surface and first 15 model levels are shown to focus on the SBI. For the IC subsets, this is extended to focus on the mid-levels where there are co-located warm and dry biases. Bold black lines represent Concordiasi mean profiles. Colored lines represent corresponding AMPS forecast mean profiles for given lead times. . .	27
4.2	As in Figure 4.1 but for low-level wind speed.	29
5.1	Statistics of (a,b,c) mean bias, (d,e,f) standard deviation bias, (g,h,i) CRMSD, (j,k,l) MSESS for AMPS (a,d,g,j) temperature, (b,e,h,k) dew point, and (c,f,i,l,o) wind speed forecasts. Statistics represent the HP subset for lead times from 6 hours to 4 days.	35
5.2	As in Figure 5.1 but for the LE subset.	37
5.3	As in Figure 5.1 but for the IC subset.	40
6.1	Biases for the HP subset for (a) temperature and day, (b) temperature and night, (c) relative humidity and day, (d) relative humidity and night, (e) wind speed and day, and (f) wind speed and night. Statistically significant biases are hatched.	42

6.2	As in Figure 6.1, except for the LE subset.	45
6.3	As in Figure 6.1, except for the IC subset.	49
6.4	Mean Inversion strength (IS) biases for temperature, (a) day and HP, (b) temperature, night and HP, (c) dew point, day and HP, (d) dew point, night and HP, (e) temperature, day and LE, temperature, (f) night and LE, (g) dew point, day and LE, and (h) dew point, night and LE. AMPS is in red, GFS is in blue, and ERA is in green. Error bars indicate the 95% confidence intervals.	51
6.5	As in Figure 6.4 but for inversion depth (ID).	53
6.6	Fraction of inversions forecast as a total of the whole subset in the IC subset for (a) temperature during day, (b) temperature during night, (c) dew point during day, (d) and dew point during night. Observations are in black, AMPS is in red, GFS is in blue, and ERA is in green.	55
6.7	Mean LLWM maximum wind speed biases for (a) HP and day, (b) HP and night, (c) LE and day, (d) LE and night. AMPS is in red, GFS is in blue, and ERA is in green. Error bars indicate the 95% confidence intervals.	57
6.8	As in Figure 6.7 but for LLWM height.	58
6.9	Fraction of LLWMs forecast as a total of the whole subset for (a) HP and day, (b) HP and night, (c) LE and day, (d) LE and night, (e) IC and day, and (f) IC and night. Observations are in black, AMPS is in red, GFS is in blue, and ERA is in green.	59
7.1	The 95% confidence intervals calculated from 1000 bootstrapped samples for (a,b,c) mean bias, (d,e,f) standard deviation bias, (g,h,i) CRMSD, (j,k,l) Correlation, and (m,n,o) MESS and (a,d,g,j,m) temperature, (b,e,h,k,n) dew point, and (c,f,i,l,o) wind speed. Statistics represent the HP subset for AMPS (red), GFS (blue), ERA-interim (green).	68
7.2	As in Figure 7.1 but for forecasts at 72 hours lead time.	70
7.3	As in Figure 7.1 but for the LE subset.	74
7.4	As in Figure 7.1 but for the LE subset and forecasts at 72 hours lead time.	76
7.5	As in Figure 7.1 but for the IC subset.	79
7.6	As in Figure 7.1 but for the IC subset and forecasts at 72 hours lead time.	81

Abstract

The Antarctic Mesoscale Prediction System (AMPS) is currently the only real-time, limited-area numerical weather prediction (NWP) model with a domain over the Antarctic region. This study presents an overall evaluation of AMPS by comparing its forecasts against dropsonde observations of temperature, dew point temperature, and winds obtained during the Concordiasi field program between September and December of 2010. Concordiasi dropsonde profiles provide a unique spacial and vertical coverage over the region where observations are traditionally provided at only a limited number of stations.

These comparisons reveal that biases in the model forecasts are greatest in the boundary layer over the Antarctic continent, with warm, moist, and slow forecast surface biases. In addition, over the continent, there are weak and shallow temperature and dew point inversions, and slow low-level jets. Over the sea ice, there is a warm and dry bias, at the same location where there is a known bias of low cloud fraction within multiple NWP models. Furthermore, deep temperature and dew point inversions are forecast in AMPS at early lead times, which strongly depend on the initial conditions provided by the NCEP's Global Data Assimilation System.

We also present a comparison of AMPS to the Global Forecast System and the ERA-Interim reanalysis. AMPS forecasts exhibit lower bias in forecasts of low-level atmospheric features over the continent, in particular over the high plateau

region of Antarctica. However, ERA-Interim exhibits lower bias in forecasts of low-level atmospheric features over the sea ice.

Chapter 1

Introduction and Background

In recent decades, advances in computational power and numerical weather prediction (NWP) have allowed for improved representation and future projections of the atmospheric system through higher model resolution, improved data assimilation techniques and more accurate parameterization schemes (Kalnay, 2003; Dee et al., 2011). Despite such advancement, the skill of Antarctic NWP products lag behind those in the lower-latitudes (World Meteorological Organization, 2013; Tilley and Bromwich, 2005; Pendlebury et al., 2003). However, during this period, the importance of accurate representation of Antarctic processes has become more apparent for both weather and climate applications (Mayewski et al., 2009). With the aim of improving the representation of Antarctic processes within NWP models, the Antarctic Mesoscale Prediction System (AMPS) was developed and began producing forecasts in 2000 (Powers et al., 2012). This thesis approaches the issue of poor Antarctic NWP models with the aim of evaluating and recommending areas for which improvement in AMPS is required using a unique new dataset of dropsondes produced during the Concordiasi field project (hereafter referred to as Concordiasi; Rabier et al., 2013).

1.1 Motivation

1.1.1 Antarctica

Since the inception of the satellite-era, the discovery of the ozone hole and other such scientific steps forward, the importance of the Antarctic region in particular, has grown in the scientific community, specifically with regard to global climate. Given that the Antarctic ice sheet holds 90% of the global fresh water, Antarctica holds significant societal importance due to its potential effects on sea level rise (IPCC, 2013). Changes in climatic phenomena govern a large proportion of the Antarctic ice-sheet's mass balance changes. Warming Antarctic temperatures, specifically in West Antarctica (Bromwich et al., 2013a; Turner et al., 2006; Vaughan et al., 2003) where the ice sheet is most vulnerable (Oppenheimer, 1998), contribute to decreases in the ice sheet's mass. Climate change has also led to changes in atmospheric circulation, in particular to the Southern Annular Mode (SAM) as shown by Arblaster and Meehl (2006). As discussed by Russell and McGregor (2010) this contributes to atmospheric temperature changes. In addition, Son et al. (2009) show that projected changes in ozone likely lead to an increase in West Antarctic warming through alteration of the Southern Hemisphere atmospheric circulation. Further, Bengtsson et al. (2006) show that there will likely be a displacement of storm tracks poleward in the future. This leads to changes in precipitation, whilst also affecting ocean currents that result in oceanic temperature changes and melting of the ice sheet from below (Pritchard et al., 2012). However, Steig et al. (2013) and Monaghan et al. (2008) show that warming temperatures have caused increases in precipitation, partially offsetting the ice sheet's decrease in mass. Given the above, Antarctica's climate, and therefore its ice sheet mass balance, are clearly a complicated system that need to be studied in order to estimate future sea level changes.

Other issues concerning Antarctic meteorology stem from a lack of understanding of polar processes, relative to those in lower-latitude regions, such as the dynamical, physical and chemical processes involved with clouds, gravity waves, the boundary layer, katabatic flow, and polar lows. Bromwich et al. (2012) document the current state of knowledge on tropospheric clouds in the Antarctic. In this review study, they also emphasize the lack of knowledge of microphysical and radiative properties in Antarctic clouds, relative to other regions. The scientific communities understanding of orographic gravity waves, forced by the various large orographic features in the Antarctic region, such as the mountain ranges of the Antarctic peninsula and the Transantarctic Mountains, is poor (Plougonven et al., 2013). In addition, their effects on other phenomena such as clouds (Alexander et al., 2013) and the atmospheric circulation (McLandress et al., 2012) have only recently been studied and the literature in these fields is still minimal. The treatment of stable boundary layers (a common occurrence over the predominantly ice-covered surfaces of Antarctica and the surrounding sea ice) lack a framework of knowledge (King and Turner, 2007; Fernando and Weil, 2010; Bourassa et al., 2013; Holtslag et al., 2013; Bromwich et al., 2012) since Monin-Obukhov similarity theory breaks down in environments of strong stability (Ha et al., 2007). Other Antarctic mesoscale phenomena such as katabatic flows and polar lows, and their effects on the overall Antarctic environment are also poorly understood (World Meteorological Organization, 2013). All these phenomena, among others, feed back into the climate system through radiative processes, and changes to atmospheric circulation and surface temperatures (Turner and Marshall, 2011; King and Turner, 2007). Thus a detailed understanding and accurate, high resolution future projections of all the above, are paramount to predicting future climates.

A limiting factor in the understanding of many of the aforementioned Antarctic atmospheric processes is the relative dearth of available observations in the Antarctic (Barker, 2005; Nordeng et al., 2007; Powers, 2007; Bromwich et al., 2012). While some advances have been made in the observation of Antarctic processes in the past 50 years since the inception of meteorological satellites (Bromwich et al., 2007) and automatic weather stations (AWS; Lazzara et al., 2012), few regular in-situ observations are made. For example, upper air observations are key to a wide array of research tasks including the long term measurement of climate processes, short term measurements for analysis and forecasting, verification of remote sensing observations (Wang et al., 2013a) and observations through which to gauge the performance of, and evaluate NWP products. The scarcity of upper air observation sites in Figure 1.1 is evidence for the lack of in-situ observations, with only around 10 sites regularly reporting at a frequency of twice daily. In comparison to the contiguous United States (an area of approximately 10 million km² relative to Antarctica's 14 million km²) with 92 regularly launching upper sites, Antarctica has a very low density of upper air measurements.

1.1.2 Numerical Weather Prediction

In the absence of observations, NWP models can be used to fill the observation voids while also producing forecasts of future weather. However, as earlier discussed, NWP performance in the Antarctic is relatively poor (Pendlebury et al., 2003). This can be attributed to a number of NWP features that are only optimized for use in lower-latitude regions (World Meteorological Organization, 2013).

Firstly, NWP coverage in Antarctica has in most cases been produced by global models, with coarse resolutions and hydrostatic dynamical cores. However, at high-latitudes, an increased coriolis force leads to a decreased Rossby radius of

deformation and smaller synoptic scale disturbances. Further, there are an abundance of mesoscale phenomena such as polar lows, and katabatic flows and orographic gravity waves forced by steep orography. To resolve such phenomena, small grid-spacing and the accompanying non-hydrostatic dynamical cores are required. Agosta et al. (2013) show that high resolution modeling is key to estimating the ice-sheet mass balance with 30% simulated differences between 15- and 60-km grid spacing. Moreover, Bouchard et al. (2010) discuss how improvements to grid structures can lead to improvement in NWP models over Antarctica. They show using the ARPEGE (Action de Recherche Petite Echelle Grande Echelle; Deque et al., 1994) model, that unphysical variations in model orography can be removed by centering the grid on Antarctica. The development of limited area non-hydrostatic NWP models has however, been focused on the tropics and mid-latitudes, since convection is more prominent and the forecasting need for such models is greater.

Second, as discussed by Liu and Xiao (2013), current data-assimilation schemes have not been developed to deal with the lack of observations in Antarctica. Bouchard et al. (2010) recently showed how the addition of new satellite data can improve assimilation in the Antarctic. Also, the implementation of new data assimilation schemes in research has shown the inability of the current schemes to deal with Antarctic conditions. Bouttier (1994) show how the problems with Antarctic NWP, such as this lack of observations, and steep terrain may lead to inhomogeneous background error covariance matrices, since errors are strongly driven by dynamics. Further, Michel and Auligné (2010) show that background error covariances are strongly effected by ocean and land boundaries surrounding Antarctica. Liu and Xiao (2013) show how, for the forecast of a cyclone in the Antarctic, a 4-dimensional ensemble based variational data assimilation scheme can be used to improve on the WRF 3-dimensional variational scheme used in

AMPS. These studies emphasize the lack of skill current data assimilation schemes have to deal with the complex Antarctic environment.

Third, appropriate parameterization schemes that accurately approximate the Antarctic atmosphere, as in lower-latitudes, have not been developed. Harsh, extreme, and remote Antarctic conditions limit the number, location and duration of experiments, aimed at understanding the virtually ubiquitous surface of snow and ice. Development of parameterization schemes based on such limited data and understanding result in schemes that are designed only for certain areas and seasons, that may not be representative of the annual conditions, and across the wider continent. Boundary layers, gravity waves, and clouds, among other Antarctic atmospheric processes and phenomena have all been shown to be poorly parameterized by both operational and reanalysis NWP products.

For example, the Antarctic is often characterized by stable conditions. Holtslag et al. (2013) and Fernando and Weil (2010) provide an overview of the challenges faced in modeling stable boundary layers. These studies show that NWP models produce too much vertical mixing, resulting in deeper than observed boundary layers, too little turning of the wind with height, large downward sensible heat fluxes and weak low-level maximum in wind speed (LLWM). These errors are partially a result of the incorrect specification of surface fluxes (Bourassa et al., 2013), and properties (Kuipers Munneke et al., 2011) of ice surfaces, associated with un-specialized parameterizations for the Antarctic.

Alexander et al. (2013) give a broad overview of research pertaining to gravity wave drag, and its effect on the global circulation. They discuss how gravity wave parameterization is key to obtaining correct estimates within climate models. More specifically, McLandress et al. (2012) show how poorly parameterized orographic gravity waves generated in the Southern Hemisphere effect climate models, with both tropospheric and stratospheric jets weak in future climate projections.

Bodas-Salcedo et al. (2012) found that the Met-Office Unified Model had problems with the under-representation of mid-top and stratocumulus clouds on the cold air side of cyclones, in the Southern Ocean. They found that errors concerning these cloud types could be largely attributed to the boundary layer parameterizations used. Upon this evaluation, Huang et al. (2014) found that the Weather Research and Forecasting model (WRF) had many similar issues regarding the under-representation of clouds and emphasize that this is likely a result of parameterization errors. Further, Fogt and Bromwich (2008) showed that these issues also affect AMPS forecasts, expanding on the other studies by noting that AMPS forecasts upper-level clouds accurately.

With the importance of both the Antarctic cryosphere and atmosphere to climate and sea level change, the discussed NWP deficiencies in the Antarctic are a clear barrier to weather forecasting and climate projections. Thus, the development of specialized non-hydrostatic models, along with the associated data-assimilation and parameterization schemes, is a key component of future research to understanding Antarctica.

1.1.3 Forecasting Applications

Antarctic NWP is utilized to support activities including logistics for science and tourism, which have grown rapidly during the past few decades. Currently, an estimated 5000 scientists operate in and around the Antarctic continent, while tourist visits to Antarctica grew from an estimated 6500 people per year in 1992 to around 35000 people per year in 2009 (International Association for Antarctic Tour Operations, 2014). The cost of transportation to and from the region is highly dependent on weather. For example, the typical cost for a New Zealand to McMurdo flight to turn around at the point of safe return due to expected poor landing conditions is \$US100,000 (World Meteorological Organization, 2013).

Furthermore, the aforementioned lack of observations and the relative isolation of the Antarctic continent increases the reliance upon NWP forecasts and thus they have become a key component of logistical and scientific operations in and around the Antarctic continent.

The need for specialized, non-hydrostatic NWP models to accurately represent phenomena across all scales in the Antarctic is therefore paramount to both forecasting and research, in an otherwise data-sparse region of the globe.

1.2 The Antarctic Mesoscale Prediction System (AMPS)

Until October 2000, NWP capability was provided by global models such as the National Center for Environmental Prediction's (NCEP) Global Forecasting System (GFS) and European Center for Medium Range Weather Forecasting's (ECMWF) Integrated Forecasting System (IFS). After October 2000, in response to the need for higher resolution NWP capability in the Antarctic, the Antarctic Mesoscale Prediction System (AMPS; Powers et al., 2012) was developed and implemented by the National Center for Atmospheric Research (NCAR) in conjunction with the Byrd Polar Research Center (BPRC). AMPS is a NWP model that is designed to address many of the issues discussed in Section 1.1. AMPS, implements a modified version of the polar-optimized Weather Research and Forecasting Model (Polar WRF; Hines and Bromwich, 2008; Bromwich et al., 2009; Hines et al., 2011; Bromwich et al., 2013b). As well as providing forecasting support for the United States Antarctic Program, other international Antarctic scientific efforts and Antarctic tourism, AMPS has been used for a variety of different research studies of Antarctic weather and climate. These include studies of atmospheric

moisture (Fogt and Bromwich, 2008), winds (Nigro et al., 2012; Seefeldt and Casano, 2012; Steinhoff et al., 2012), cyclones (Uotila et al., 2009), local climatologies (Monaghan et al., 2005), and cloud and precipitation regimes (Schlosser et al., 2008).

Such extensive use of AMPS in forecasting and research efforts require that the model be evaluated for areas of poor performance. A recent evaluation of the Polar WRF in the Antarctic by Bromwich et al. (2013b) used AWS sites across the continent to produce an annual time series of surface biases. In addition, the Amundsen Scott South Pole site and five other coastal sites were used to evaluate the upper air performance at standard pressure levels. Bromwich et al. (2013b) showed that seasonal variations in model skill exist. Bromwich et al. (2013b) indicate that there are cold surface temperature biases during the summer with warm surface temperature biases during the winter. Further Bromwich et al. (2013b) showed that there are positive wind speed and pressure biases throughout the year, with larger biases in the winter for wind speed, and in the summer for pressure. These positive surface wind speed biases and an anomalously large downward flux of sensible heat toward the surface produced the warm surface bias in winter. In addition Bromwich et al. (2013b) concluded that model skill is affected more by the analysis than the physical parameterizations or multi-annual differences in the atmospheric circulation. Improving the current network of observations and quality of the analyses was therefore a conclusion drawn for future focus.

Other evaluations of polar NWP models such as Tastula and Vihma (2011), Tastula et al. (2012), and Kumar et al. (2012) have used single sites to verify model performance. Kumar et al. (2012), evaluated Polar WRF simulations over the Maitri region of Antarctica. They showed contrasting results to Bromwich

et al. (2013b) indicating negative surface wind speed biases, although pressure biases were similar to that of Bromwich et al. (2013b). Furthermore, Tastula and Vihma (2011) and Tastula et al. (2012) examined the treatment of the atmospheric boundary layer in Polar WRF over sea ice. They showed many results consistent with Bromwich et al. (2013b) in regard to variables such as temperature and cloud cover although their analysis of surface flux biases within the model was inconsistent with Bromwich et al. (2013b).

These previous evaluations and their inconsistent conclusions highlight the many issues with evaluating NWP models in the Antarctic. Low spatial coverage, a lack of observations, few specialized observations (such as radiative fluxes) and highly transient meteorological conditions make it difficult to identify consistent errors with AMPS and Polar WRF.

1.3 The Concordiasi Project

It is clear that in addition to the issues posed by a lack of observations for operational NWP and forecasting, the dearth of in-situ observations also hinders the ability to improve the model through direct evaluations. In particular an issue that all the previous evaluations encountered was a distinct lack of spatial coverage for observations, specifically for upper air observations. Given this issue the current study attempts to address the discussed problems by evaluating AMPS using dropsonde data produced during the Concordiasi project. Concordiasi was a multidisciplinary, joint France-US initiative designed to study the troposphere, lower stratosphere and land surface of Antarctica. Surface measurements and radiosonde launches from the Concordia, Dumont d'Urville, and Rothera sites were carried out in the austral springs of 2008, 2009, and 2010. From September through December of 2010, 13 driftsondes with instrument packages on tethered gondolas, drifted on isopycnic surfaces within the stratospheric winter polar vortex (Rabier

et al., 2013; Cohn et al., 2013). The driftsondes took in situ measurements of stratospheric properties while also allowing the launch, on command, of around 640 dropsondes. Dropsonde launches were primarily designed to coincide with passes of the Infrared Atmospheric Sounding Interferometer (IASI) with a secondary goal to provide observations in regions with critical gaps in observational coverage. As a result there is a broad spacial coverage of observations in the data set of dropsondes. Given the spacial coverage of observations Concordiasi dropsondes provide an excellent spacial distribution of tropospheric profiles (Figure 1.1) relative to current and previous Antarctic observations.

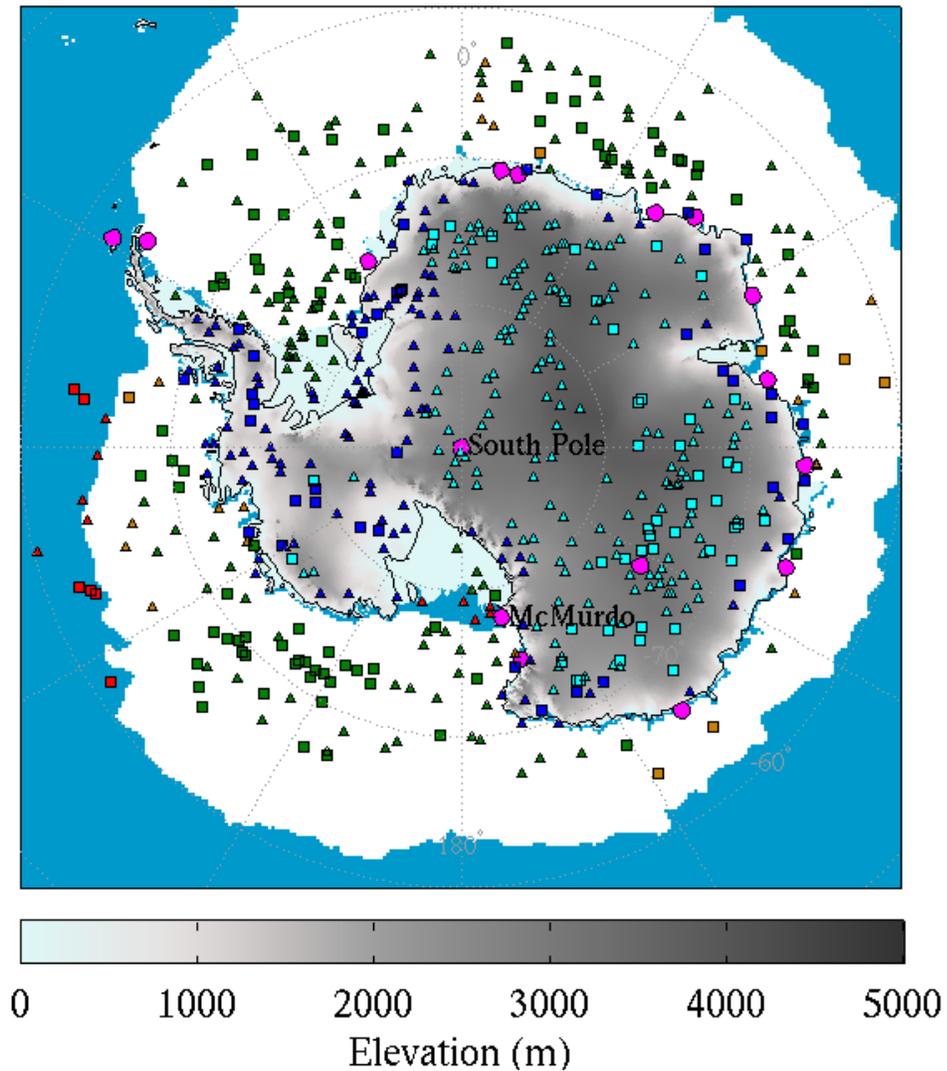


Figure 1.1: A terrain map of Antarctica with October 2010 mean sea ice extent shown in white and open water shown in blue. Antarctic radiosonde sites (as of 2010) in pink and Concordiasi dropsondes separated by the surface type over which they dropped are shown. Surface types include continental high plateau (HP) in cyan, continental low elevation (LE) in blue, sea ice or permanent ice shelves (IC) in green, Partial Sea Ice Cover (PI) in brown, and Open Water (OW) in red. Triangles represent profiles during the day and squares represent those at night.

Chapter 2

Model and Observational Details

In this chapter we discuss the details of the models used in the study. Further, we also discuss the observational errors of Concordiasi dropsondes and measures we have taken to alleviate any problems arising from the errors we discuss.

2.1 AMPS, the Global Forecast System (GFS), and ERA-Interim

We evaluate AMPS forecasts that were performed in real-time during the period September through December of 2010. The forecasts were obtained through the NCAR high performance storage system (Mesoscale & Microscale Meteorology Division, University Center for Atmospheric Research, 2014). During this period, AMPS implemented a modified version of Polar WRF, which in itself is a modified version of the Advanced Research WRF (WRF-ARW) Version 3.0.1. Modifications to improve the performance of WRF in the polar regions are made to the parameterizations. The specific schemes used are documented in Table 2.1. The discussed modifications are primarily to the Noah Land Surface Scheme to account for the different properties of surfaces of snow and ice, and are discussed in Hines

Parameterization	Scheme
Longwave	RRTM
Shortwave	Goddard
Boundary Layer	Mellor-Yamada-Janjic (Eta) TKE scheme
Surface Layer	Monin-Obukhov (Janjic Eta) scheme
Land Surface	Unified Noah LSM
Microphysics	WSM-5 Class Scheme
Cumulus	Kain-Fritsch
Sea Ice	Custom fractional sea-ice implementation

Table 2.1: AMPS Parameterization Schemes during Concordiasi

and Bromwich (2008); Bromwich et al. (2009); Hines et al. (2011); Bromwich et al. (2013b).

AMPS grid spacing included a 45-km outer grid encompassing Antarctica, the southern ocean, New Zealand, and southern parts of South America, Africa and Australia; a 15-km grid encompassing Antarctica and southerly parts of the Southern Ocean; and 4 smaller grids located over the Antarctic peninsula, the South Pole, Ross Sea and Ross Island (Figure 2.1). Given the broad spacial distribution of Concordiasi, the current evaluation focuses on the 45-km grid. The parameterization schemes of these simulations are documented in Table 2.1.

The operational GFS is used for comparison to AMPS forecasts. We use the real-time $0.5^\circ \times 0.5^\circ$ configuration, identical to those used as the initial and boundary conditions for AMPS. This allows a direct comparison to the AMPS initial conditions, with the aim of deducing the source of any analysis errors. Information regarding the GFS setup can be found at NOAA, Environmental Modelling Center (2014).

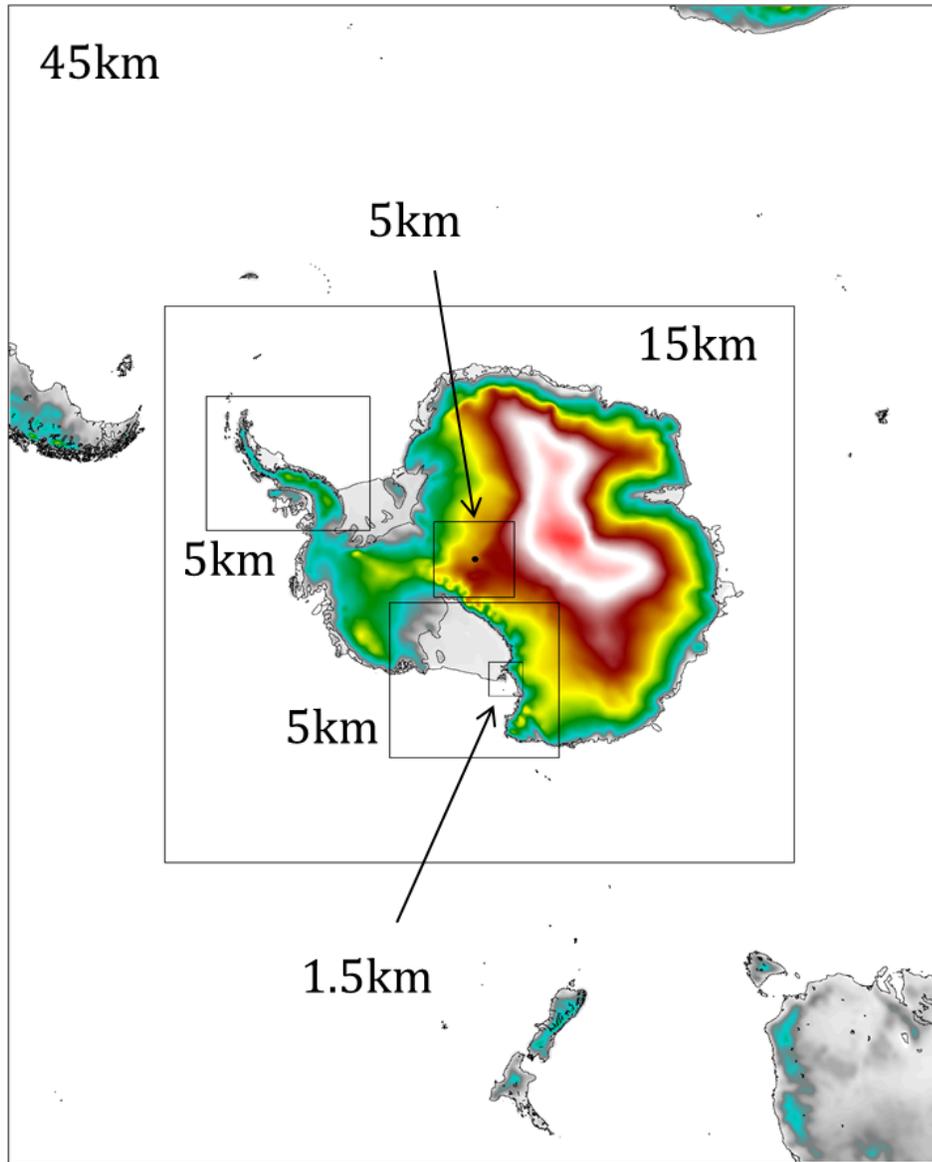


Figure 2.1: AMPS grid structure during September-December 2010. Source: Mesoscale & Microscale Meteorology Division, University Center for Atmospheric Research (2014).

Further, we also use ECMWF’s ERA Interim reanalysis for comparison to AMPS. We use the $0.5^\circ \times 0.5^\circ$ configuration similar to the GFS. Information regarding ERA-Interim can be found at Dee et al. (2011).

2.2 Dropsonde data

The dropsondes, created by the Earth Observing Laboratory at NCAR are miniature in-situ sounding technology (Cohn et al., 2013; National Center for Atmospheric Research, Earth Observing Laboratory, 2011). These use modified Vaisalla RS92 dropsonde sensors and therefore have similar performance characteristics. The measurement of temperature, pressure and winds have no systematic measurement biases reported in the literature and the Vaisalla reported errors are widely accepted. However, biases in measurements of humidity in radiosondes have been documented for some time (Guichard et al., 2000). Vömel et al. (2007) first documented a dry bias in the RS92 radiosonde, caused when direct solar radiation dries the humidity sensor. To counter this bias in the Concordiasi data set, the NCAR radiation bias correction (NRBC; Wang et al., 2013b) is implemented on the Concordiasi humidity data. Other quality control issues with the soundings were treated by EOL (National Center for Atmospheric Research, Earth Observing Laboratory, 2011). These corrections include the removal of anomalous data points due to failed deployment of the sondes, false measurements after dropsondes reach the ground, and equilibration of the sensors after release from the heated driftsonde gondola. Despite these corrections, humidity errors exist in the observations. The time lag of the sensors is one of these errors which is a significant problem for the low humidities found in polar regions, especially at higher elevations. While a correction scheme is available for time lag errors in RS92 radiosondes (Miloshevich et al., 2009), there is no research that extends such schemes to dropsondes and therefore no correction has been applied. The extreme cold and dry conditions

Variable	Total Uncertainty	Response Time
Temperature	0.5C	<1s between 1000- and 100-hPa
Humidity	5%	<20s at -40C, lower at warmer temperatures
Pressure	1hPa	Not provided.

Table 2.2: Vaisalla RS92 sensor accuracy (Vaisalla, 2013).

in Antarctica, especially in the upper atmosphere at pressure lower than 500 hPa limit the accuracy of humidity measurements (Bromwich et al., 2005; King and Turner, 2007). Hence these issues with humidity observations above 500 hPa are taken into account in discussions of statistics in latter sections.

There is little information available that documents the accuracy of the MIST sonde sensors. However since the sensors are primarily those from the RS92 sondes we document the Vaisalla reported errors in the Table 2.2. In the following discussion we take into account the uncertainty given here when discussing the significance of statistics

Chapter 3

Methodology

3.1 Interpolation Techniques

To produce comparable atmospheric profiles through which to evaluate statistics, a number of interpolations are used in both time and space (Table 3.1). This study uses a simple linear interpolation in time between the 2 surrounding 3 hourly model output times to translate the model predictions to the observation time. In horizontal space, a bilinear interpolation to the observation location from the surrounding four model grid point is utilized, This has been used in previous model evaluations (Bromwich et al., 2013b), and in the WRF post-processing system. In the vertical, NWP model evaluations often interpolate to pressure levels. However, due to high and varying orography in the Antarctic region (Figure 1.1), pressure levels do not represent constant environments such as the boundary layer where NWP model errors are often largest. Therefore, this study utilized vertical interpolations to the AMPS eta level given by:

$$\eta = \frac{p - p_t}{p_s - p_t} \quad (3.1)$$

where p is the pressure at a given location, p_t is the pressure at model top (10 hPa) and p_s is the surface pressure. Analogous to the sigma coordinate, η is terrain following, thus interpolations of the observation to this level, produce constant

Dimension	Scheme	From	To
Horizontal Space	Bilinear	Surrounding four model grid points	Observation horizontal location
Vertical Space	Linear	Observation vertical location	Model Level
Time	Linear	3 hourly model output times	Observation time

Table 3.1: Interpolation Schemes used in the evaluation of AMPS.

data sets across constant vertical coordinates, and allow a clearer evaluation of the boundary layer in particular.

3.2 Statistics

Forecast evaluation typically requires a number of characteristics of the model be examined. These characteristics include reliability, typical error, association, variability, and skill (Jolliffe and Stephenson, 2012). To study these characteristics, we use a number of statistical moments of temperature, moisture and winds. Specifically, relative humidity and dew point are the focus for moisture variables, while speed and direction are the focus for winds. Reliability, defined as whether the model represents the observed atmospheric characteristics, or something else, is represented by bias:

$$bias = A_m - A_o \quad (3.2)$$

where A is a given variable (e.g. temperature) with the model variable denoted by the subscript m and the observational variable denoted by the subscript o. Bias therefore represents over or under-forecasting through positive and negative biases respectively. Paired, 2-sided T-tests at the 95% confidence level are performed in

order to note the statistical significance of mean biases in Sections 6.1, 6.2, and 6.3 e.g. the difference from a mean bias of zero (Wilks, 2011).

Error, defined as the accuracy of the model, is represented in most studies (Bromwich et al., 2013b, 2005) by root mean square deviation or error (RMSD or RMSE):

$$RMSD = \sqrt{\frac{\sum (A_m - A_o)^2}{N}} \quad (3.3)$$

However, RMSD accounts for both the bias and the error, thus the centered root mean square deviation (CRMSD: Taylor, 2001) is used in this study:

$$CRMSD = \sqrt{\frac{\sum ((A_m - \bar{A}_m) - (A_o - \bar{A}_o))^2}{N}} \quad (3.4)$$

This accounts for the part of the error that does not include the bias since RMSD can be separated into CRMSD and bias. Variability is represented by the standard deviation of the data set:

$$s = \sqrt{\frac{\sum (A - \bar{A})^2}{N - 1}} \quad (3.5)$$

where N is the number of observations. Differences between the observed s and forecast s (standard deviation bias) are calculated in order to observe whether the model captures the correct range of the observed variable. The standard deviation bias is also used to measure the degree to which the model is correctly capturing atmospheric cycles, such as boundary layer temperature over a 24 hour period. Association, the degree to which the forecasts match the truth, is in many studies represented by correlation coefficient:

$$\rho = \frac{\sum (A_m - \bar{A}_m)(A_o - \bar{A}_o)}{\sqrt{\sum (A_m - \bar{A}_m)^2 \sum (A_o - \bar{A}_o)^2}}. \quad (3.6)$$

However it is shown by McCuen et al. (2006) that ρ is simply the square root of the Mean Square Error Skill Score (MSESS: Jolliffe and Stephenson, 2012):

$$MSESS = 1 - \frac{\sum (A_m - A_o)^2}{\sum (\bar{A}_o - A_o)^2} \propto \rho^2 \quad (3.7)$$

which also represents skill, the degree to which the forecast improves upon some reference forecast, in this case climatology. MSESS directly relates the model forecast to a forecast of climatology while also quantifying association. Since MSESS is related to the square of the bias by:

$$1 - MSESS \propto (A_m - A_o)^2, \quad (3.8)$$

a MSESS of 0.75 represents an absolute error that is 50 % of the magnitude of the bias from a forecast a forecast of climatology. Thus a skill score of 0.75 or greater represents a very good skill score. Using this relation a MSESS of 0.44 gives an absolute error from AMPS that is 75 % of the bias obtained from climatology. Therefore 0.44 represents a mediocre skill score. Anything less represents an absolute error with little improvement upon the climatology and is therefore a poor skill score. By design a score of less than 0 means the prediction is worse than climatology.

These scores combine to comprehensively represent model performance. In order to quantify the uncertainty in these statistics we use 95 % confidence intervals of 5000 re-sampled (with replacement) subsets of the data (Wilks, 2011). These confidence intervals provide an accurate means of noting the statistical significance of all statistics, especially those statistics such as MSESS that only have one value for the entire subset and for which it is therefore not possible to calculate traditional confidence intervals. Re-sampled confidence intervals also provide a non-parametric method through which to test the significance of statistics. However in this case results of re-sampling are comparable to the parametric T-test used to test the statistical significance of biases in Sections 6.1, 6.2, and 6.3 because in most cases the T distribution fits the biases and so a parametric test is appropriate.

3.3 Lead time

To produce a complete time series for the duration of the model, statistics are presented for 10 windows of 12 hours each, 6 hours either side of the presented lead time. These windows vary from 6 to 108 hours (4.5 days; the last 12 hour window possible given the 120 hour forecast of the model). In addition there is a 0 hour window designed to represent the initial conditions. In this window interpolations in time are made between AMPS 0 and 3 hour forecasts. As such, the 0 hour window does not include all dropsonde profiles and is around a quarter of the size of the other windows, since the model was only initialized every 12 hours as shown in Table (3.2). There are 3 windows covering the first 18 hours of the model integration in order to capture the rapid progression of changes in the atmospheric state after the model is initialized and diagnose any biases that may be a result of poorly specified initial conditions.

3.4 Subsets based on Antarctic Environments

During the following analysis, it was noted that low-level model bias is strongly related to the surface. King and Turner (2007) discuss how the atmospheric boundary layer over the Antarctic continent can be separated into three regimes, the high continental interior, the steep coastal katabatic zones, and over the ice shelves. Thus, in order to separate biases by different boundary layer environment we subset the profiles by day and night, as well as by the surface types over which the dropsonde fell. These are (1) over the elevated continental interior and high plateau (HP), (2) over the steep coastal areas and lower elevations surrounding the interior (LE), (3) over the permanent ice shelves and sea ice (IC), (4) over an area of partial sea ice cover (PI), and (5) over open water (OW).

The HP subset is defined as any vertical profiles with a surface pressure less than or equal to 750 hPa, over the continent. This definition is designed to include all profiles over the high plateau region of Antarctica but exclude any profiles over regions with strong orography (e.g. the coasts and mountains) since these represent a different boundary layer environment. The LE subset is defined as any profile with a surface pressure greater than 750 hPa, over the continent. This subset is designed to provide some insight into AMPS performance in the katabatic zone, although in reality, due to the localized nature of strong katabatic flows, it encompasses a mixture of environments found at the lower elevations over Antarctica. The IC subset is defined as any profile over the permanent ice shelves, or over a cell with at least 90% sea ice cover. Sea ice concentration is from the 25-km grid spacing National Snow and Ice Data Center (NSIDC) Climate Data Record of Passive Microwave Sea Ice Concentration, available at National Snow and Ice Data Center (2014). The PI subset is defined as profiles over sea ice with a concentration between 10 and 90 %. The OW subset is defined as any profile over a cell with sea ice concentration of less than 10%, which includes those that fell over the open ocean as well as polynyas. Due to very small sample sizes in the PI and OW subset no statistics are presented here since no statistical significance can be attributed.

Each subset sample size is shown in Table 3.2 while the location and subset of each dropsonde is shown in Figure 1.1. In the Chapter 4 we discuss the atmospheric characteristics of these subsets in further detail.

Table 3.2: Sample sizes for the various subsets. Forecast subset sizes are those without parentheses and 0 hour subset sizes are in parentheses.

Subset	Day	Night	Total
High Plateau (HP)	152 (49)	47 (26)	199 (75)
Low Elevation (LE)	107 (24)	40 (10)	147 (34)
Total Ice cover (IC)	123 (28)	95 (18)	218 (46)
Partial Sea Ice cover (PI)	18 (6)	7 (3)	25 (9)
Open Water (OW)	15 (8)	7 (1)	22 (9)
Total	424 (115)	218 (58)	611 (173)

Chapter 4

Observed Atmospheric Characteristics in the Concordiasi Data Set

This section is designed to provide an overview of the benefits Concordiasi provides in observing atmospheric features across the Antarctic region. Here we discuss the observed characteristics in the Concordiasi data set with a focus on the key low-level features. Features are discussed with respect to the subsets as defined in Section 3.4. We also define some key parameters of the low-level features that will be used in later sections.

4.1 Observed Characteristics

As defined in King and Turner (2007) the High Plateau is characterized by intense radiative cooling at the surface, with accompanying strongly stratified low levels that lead to strong inversions (e.g. Figures 4.1a and 4.1b). In this thesis we define two parameters to characterize inversions; inversion strength and inversion depth. Inversion strength (IS):

$$IS = T_m - T_b, \tag{4.1}$$

is the difference between temperature (or dew point) at the base of the inversion (T_b) and first temperature (or dew point) maximum aloft (T_m). Inversion depth (ID), defined as:

$$ID = \frac{R\bar{T}}{g} \ln\left(\frac{P_{T_m}}{P_b}\right), \quad (4.2)$$

where R is the ideal gas constant for dry air, \bar{T} is the mean temperature of the layer, g is gravitational acceleration, P_b is pressure at the inversion base, and P_{T_m} at the first temperature (or dew point) maximum aloft. These definitions follow those used in Vihma et al. (2012), and Tastula et al. (2012). In the HP subset a majority of temperature inversions are surface based (SBIs), consistent with the forcing of strong radiative cooling at the surface. Inversions are on average stronger during the night with mean IS of 14.3 K during the day and 17.8 K during the night (Table 4.1). Inversions are approximately 500 m deep with inversions approximately 90 m deeper during night (Table 4.1). Dew point inversions are comparable to temperature inversions in both IS and ID.

Despite mild slopes over the High plateau, katabatic effects are still important for wind due to the intense surface cooling, and resulting low buoyancy of surface air parcels relative to those not in contact with the surface. This leads to a strong low-level wind maximum (LLWM) just above the surface (e.g. Figures 4.2a and 4.2b). In this thesis we analyze two variables associated with the LLWMs. The first is the maximum wind speed of the LLWM (V_m). V_m must be 2 ms^{-1} greater than the minimum winds immediately above and below the maximum in order for the LLWM to exist. This condition is based on the formulations used in Andreas et al. (2000) and Tastula et al. (2012) for a LLWM. The second is the height of this LLWM maximum, defined as:

$$z_{V_m} = \frac{R\bar{T}}{g} \ln\left(\frac{P_{V_m}}{P_b}\right). \quad (4.3)$$

Table 4.2 shows that the mean LLWM maximum wind speed is approximately 17 ms^{-1} , the mean height of this wind maximum is approximately 192 m, and

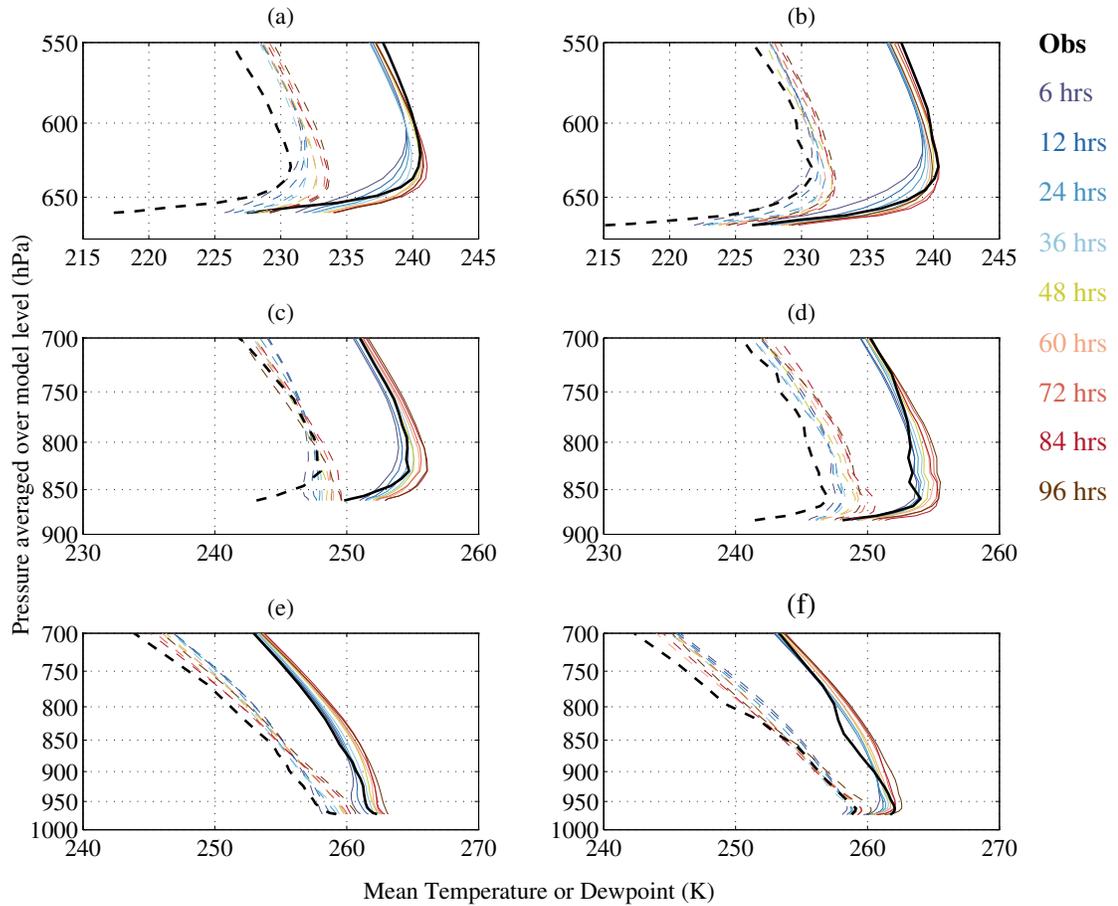


Figure 4.1: Vertical profiles of temperature and dew point for (a) HP and day, (b) HP and night, (c) LE and day, (d) LE and night, (e) IC and day, and (f) IC and night. For HP and LE subsets (i.e. a,b,c, and d) only the surface and first 15 model levels are shown to focus on the SBI. For the IC subsets, this is extended to focus on the mid-levels where there are co-located warm and dry biases. Bold black lines represent Concordiasi mean profiles. Colored lines represent corresponding AMPS forecast mean profiles for given lead times.

Table 4.1: Mean temperature inversion strength and depth, number of inversions present in the subset, and number of inversions that are surface based. Characteristics are presented for day first and then night (e.g. day/night).

Subset	Strength (K)	Depth (m)	Inversions	SBIs
High Plateau (HP)	14.3/17.8	440.7/528.0	149/44	129/44
Low Elevation (LE)	7.2/11.1	419.0/411.3	93/31	71/26
Total Ice Cover (IC)	2.9/6.0	427.1/229.1	88/79	45/43

LLWMs are observed in around 86 % of the profiles. There is little variation in LLWMs between day or night. The SBIs and LLWMs are the main features of the HP environment, and there is a focus on the accuracy of AMPS forecasts for these in following chapters.

The LE subset is characterized by the same features as those that characterize the HP subset but inversions are weaker (Figures 4.1c and 4.1d), approximately 8 K, (Table 4.1), and LLWMs are weaker and higher (Figures 4.2c and 4.2d), approximately 15.5 ms^{-1} and 350 m respectively (Table 4.2). Again observed dew point inversions are comparable to temperature inversions.

In the IC subset, weaker downward surface fluxes and flat terrain result in weak (if any) stratification (Figures 4.1e and 4.1f) leading to weak SBIs, of approximately 4 K (Table 4.1), and weaker and higher LLWMs (Figures 4.2e and 4.2f), of approximately 14.5 ms^{-1} and 500 m respectively (Table 4.2). This environment is comparable to the nocturnal boundary layer at lower latitudes. However, a reduced diurnal cycle results in a steady state rather than the continual diurnal evolution seen at lower latitudes (King and Turner, 2007). Further, LLWMs here are not always forced by the katabatic flow since there is no slope. Close to the coast there is some katabatic effect from winds flowing off the continent but often LLWMs are

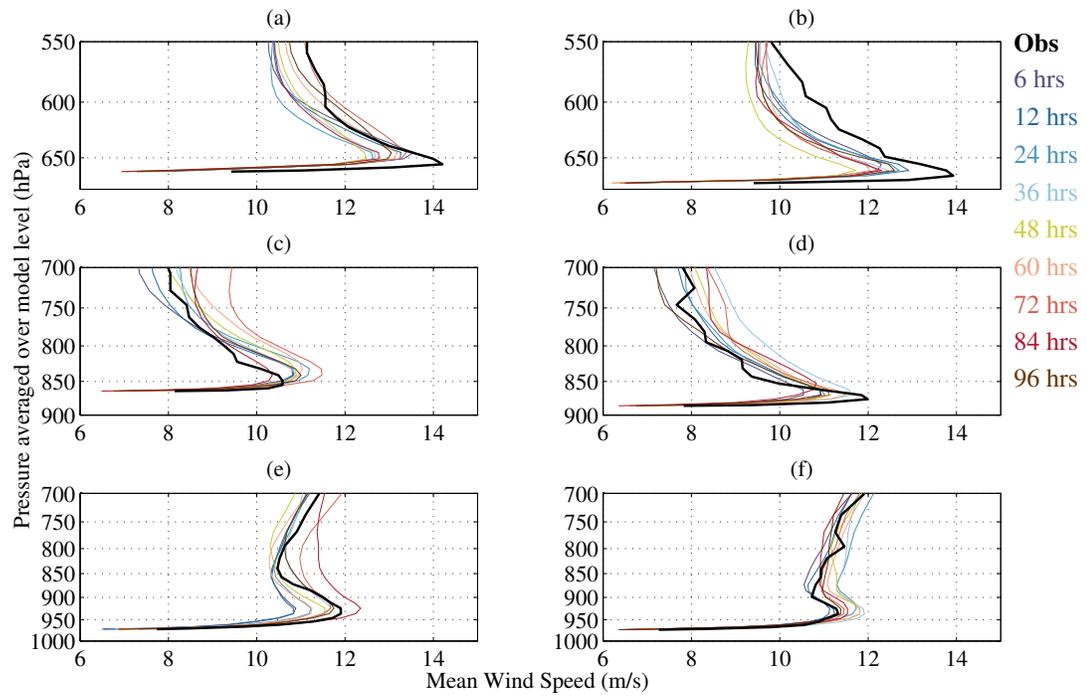


Figure 4.2: As in Figure 4.1 but for low-level wind speed.

Table 4.2: Mean LLWM maximum wind speed and height, and number of LLWMs present in the subset. Characteristics are presented for day first and then night (e.g. day/night).

Subset	Maximum Wind Speed (ms^{-1})	Height (m)	LLJs
High Plateau (HP)	16.4/17.1	192.2/191.9	132/39
Low Elevation (LE)	15.8/15.3	355.7/317.6	80/32
Total Ice Cover (IC)	14.1/15.0	462.8/532.1	84/58

forced by synoptic features such as low pressure systems. Thus LLWMs there are weaker LLWMs in the IC subset relative to the HP and LE subset because LLWMs are less persistent. This is evident in Table 4.2 where the number of LLWMs makes up a smaller fraction of the total subset size than in the HP and LE subsets.

4.2 Discussion

Since the dropsonde profiles provide one of the first data sets with a broad spacial coverage of in-situ measurements across Antarctica, Concordiasi allows a detailed in-situ look at a number of features that are poorly understood in the literature. One of these features is the atmospheric inversion, in particular over the sea ice in Antarctica. Inversion characteristics are consistent with the observations in Pavel-sky et al. (2011) who showed Antarctic inversion characteristic over sea ice using the Atmospheric Infrared Sounder (AIRS). Andreas et al. (2000) also documented the inversion characteristics over sea ice at Ice Station Weddell in Fall and Winter of 1992. They showed that 96% of profiles had inversions while 44% of profiles had SBIs. We show here that 93% of profiles exhibit inversions while 53% of these inversions are surface based, a result consistent with Andreas et al. (2000).

Further, while inversions over the continental interior have been studied in depth at certain sites such as at the Amundsen Scott South Pole Station, or at the Dome C Concordia Research Station (Hudson and Brandt, 2005; Pietroni et al., 2014), there is little research on inversions across the entire continental interior. Here we show that inversions are very strong in the HP subset and that inversions exist in nearly all profiles with a majority of these inversions surface based (100% of inversions are surface based during night). These results are consistent with Hudson and Brandt (2005) and Pietroni et al. (2014) indicating that the results at these sites are representative of the HP region.

In Chapter 4.2 we saw that winds above the surface accelerate and exhibit a LLWM on average. This LLWM is thought to represent the katabatic wind due to its structure and the overall synoptic scale flow in Antarctica (King and Turner, 2007). The katabatic wind is primarily a function of the surface heat fluxes (e.g. cooling of the surface layer by the surface), the slope of that surface, and the surface drag (King and Turner, 2007). While the LE subset was tentatively designed to account for the strong coastal katabatic flows, it actually exhibits a weaker LLWM than in the HP subset. A weaker LLWM occurs, since in reality, the katabatic flows are localized in coastal regions and the subset accounts for all profiles over lower elevations. Thus it is thought here that the continental winds captured by the Concordiasi data set are not representative of the strong katabatic flows in steep coastal regions but more representative of the overall circulation over Antarctica. Thus, since the LE subset is not representative of the steep and varying terrain, the slope and frictional effects have little effect on the katabatic wind and the cooling of air at the surface is the primary driver of katabatic flows. Therefore winds in the LE subset that are cooled less would exhibit weaker katabatic flows than in the HP subset. There are few, if any, studies that have studied the katabatic flow in the interior since most studies of the katabatic winds have focused on the extreme

coastal katabatic flows that are not captured here. Thus this study presents one of the first large scale observational views of the katabatic flow over the high plateau.

While this study was not designed to focus on the observed atmospheric characteristics in the Antarctic region, this simple analysis highlights the benefits that Concordiasi provides in documenting the characteristics of low-level atmospheric features over the Antarctic that have few detailed observations and are poorly understood relative to features in the mid-latitudes.

Chapter 5

AMPS Forecast Performance

In this section we present a general overview of AMPS forecast performance through the statistics discussed in Section 3.2 with a focus on skill represented by MSESS. These sections represent the various surface type subsets, as discussed in Section 3.4. Only the statistics are shown here, with statistical significance provided in more detailed analyses in later sections.

5.1 Continental High Plateau (HP)

Figures 5.1a, 5.1d, 5.1g, and 5.1j show the mean bias, standard deviation bias, CRMSD, and MSESS, respectively, for temperature forecasts in the HP subset. Skill, represented by MSESS is above 0.75 during the early lead times (6 and 12 hours) with the exception of two local minima; at the 5th model level and at the surface. Relatively low skill (MSESS of approximately 0.45) at the 5th model level coincides with a maxima in negative (cold) biases, and a positive maxima in standard deviation biases. Therefore, this minima in skill is a result of locally cold temperatures and relative high variability. The surface relative minima in skill (MSESS of approximately 0.55) coincides with positive (warm) biases, negative standard deviation biases and high errors (CRMSD of over 6 K). Thus, this minima in skill is likely a result of warm temperatures, low relative variability and relative

large error in the forecasts. Aloft (above the 10th model level) skill is high (MSESS of greater than 0.75) with small mean biases, standard deviation biases and errors.

Throughout the first 48 hours of the model integration there are large changes to all statistics for temperature. The cold mean bias and high standard deviation bias at the 5th model level in the early lead times are replaced with small biases by the 2 day lead time. This raises the skill at the 5th model level, such that MSESS is 0.7 at the 5th model level by the 2 day lead time despite an overall 0.1 decrease in MSESS throughout the column. By the later lead times (3 and 4 days) MSESS decreases by 0.1-0.2 such that at the surface skill is poor (although skill remains good aloft), CRMSDs are approximately 8 K and mean bias is 5.5 K.

Dew point MSESS values show that skill is progressively lower with lead time, and is highest in the mid-levels (Figure 5.1k). Above the 20th model level, skill decreases such that above the 22nd model level MSESS is below 0. However, given that there is high uncertainty in observation accuracy above 500 hPa it is not possible to attribute this low skill to AMPS. Skill is also poor at the surface with MSESS at all lead times less than 0.44 and 2, 3, and 4 day lead times less than 0. This low skill coincides with positive (moist) mean bias (Figure 5.1b) of 8-11 K, negative standard deviation biases of 2 K (Figure 5.1e), and high errors of 7-9 K (Figure 5.1h). This implies that low skill is a result of relative moist dew points, low variability, and large errors. Skill is relatively highest between the 5th and 20th model levels where 6 hour MSESS is as high as 0.7. However, there is a decrease in skill after 12 hours, and by 4 days MSESS is approximately, or less than 0 throughout the column. This decrease in skill coincides with increasing errors.

AMPS skill at forecasting wind speed is average to poor (Figure 5.1l). At 6 hours MSESS is above 0.6, except below the 5th model level where MSESS decreases to around 0.2 at the surface. Low skill below the 5th model level coincides

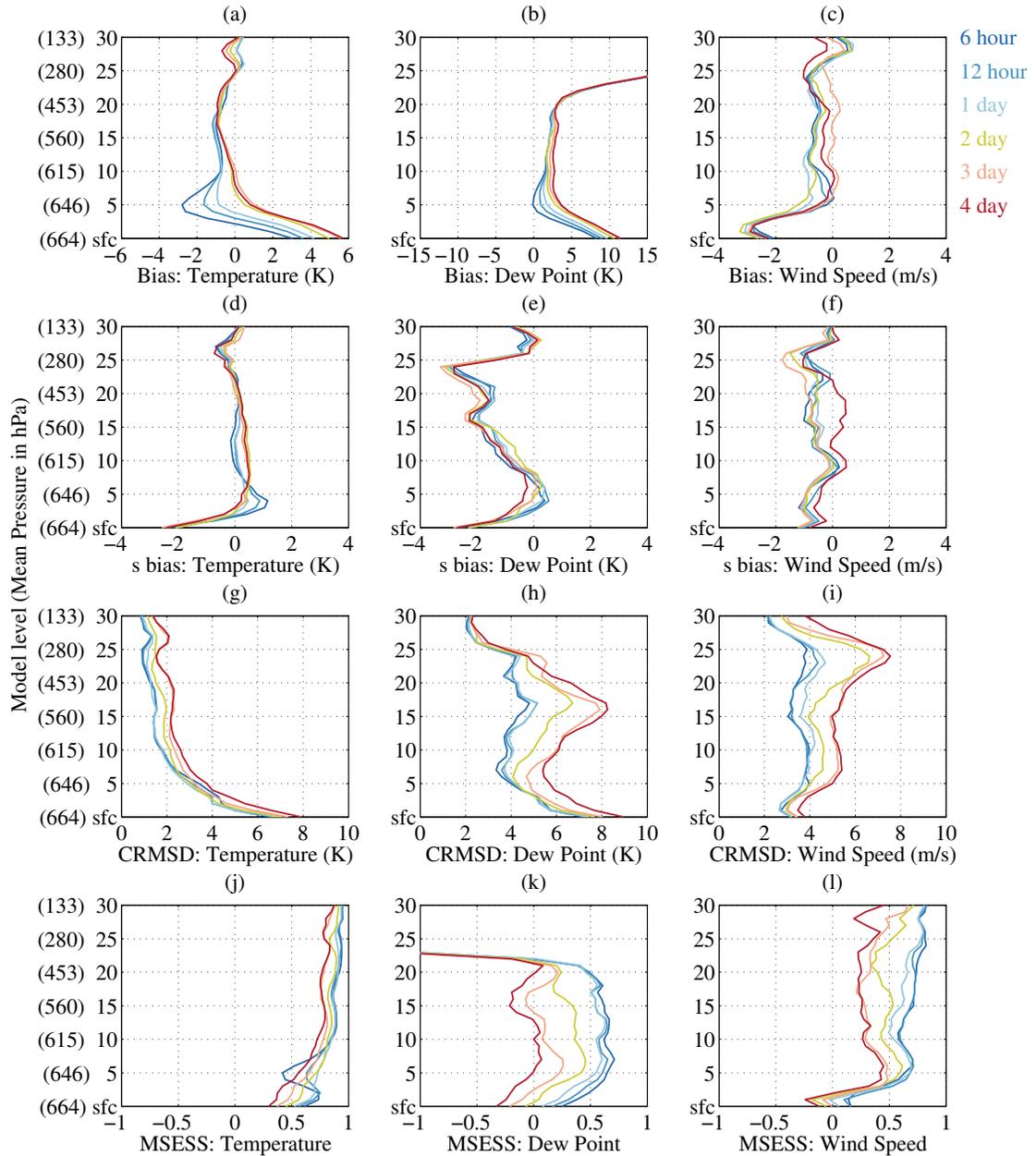


Figure 5.1: Statistics of (a,b,c) mean bias, (d,e,f) standard deviation bias, (g,h,i) CRMSD, (j,k,l) MSESS for AMPS (a,d,g,j) temperature, (b,e,h,k) dew point, and (c,f,i,l,o) wind speed forecasts. Statistics represent the HP subset for lead times from 6 hours to 4 days.

with negative (slow) mean biases up to 3 ms^{-1} implying that the low skill is primarily due to slower predicted winds in AMPS (Figure 5.1c). Above the 5th model level, while skill is good to average at early lead times, after 2 days MSESS drops below 0.44 and skill is poor with large errors of greater than 5 ms^{-1} (Figure 5.1i). Standard deviation biases are primarily negative, between 1.5 and 0.5 ms^{-1} . There are large CRMSEs of up to 7 ms^{-1} at later lead times between the 23rd and 27th model levels. This level coincides with the maximum winds in the polar jet stream. Thus at later lead times high errors are simply an artifact of the incorrect position of the polar jet. Between 1 and 2 days there are large increases in errors. Therefore the position of the polar jet in AMPS is displaced from the observations by 2 days in the model integration.

5.2 Continental Low Elevations (LE)

Skill for forecasts of AMPS temperature is good at early lead times (6 and 12 hours), with MSESS above 0.75 throughout the column except at the surface (Figure 5.2j). Similar to the HP subset there are local minima in MSESS at the surface and at the 6th model level, although the magnitude of the local minima at the 6th model level is smaller, with MSESS above 0.75. The local minima at the surface coincides with warm mean bias (Figure 5.2a), low standard deviation bias (Figure 5.2d), and high error (Figure 5.2g). However, the magnitude of these are smaller than in the HP subset which possibly accounts for improved skill in the LE subset. At the 6th model level cold mean bias and relatively higher errors coincide with the local minimum in skill. Throughout lead time there is an increase in mean bias toward warm biases, and an increase in errors which coincide with a decrease in skill. There are again large changes during the early lead times.

The skill for forecasts of dew point below the 22nd model level is average for most lead times except 4 days (Figure 5.2k). Mean biases are small between the

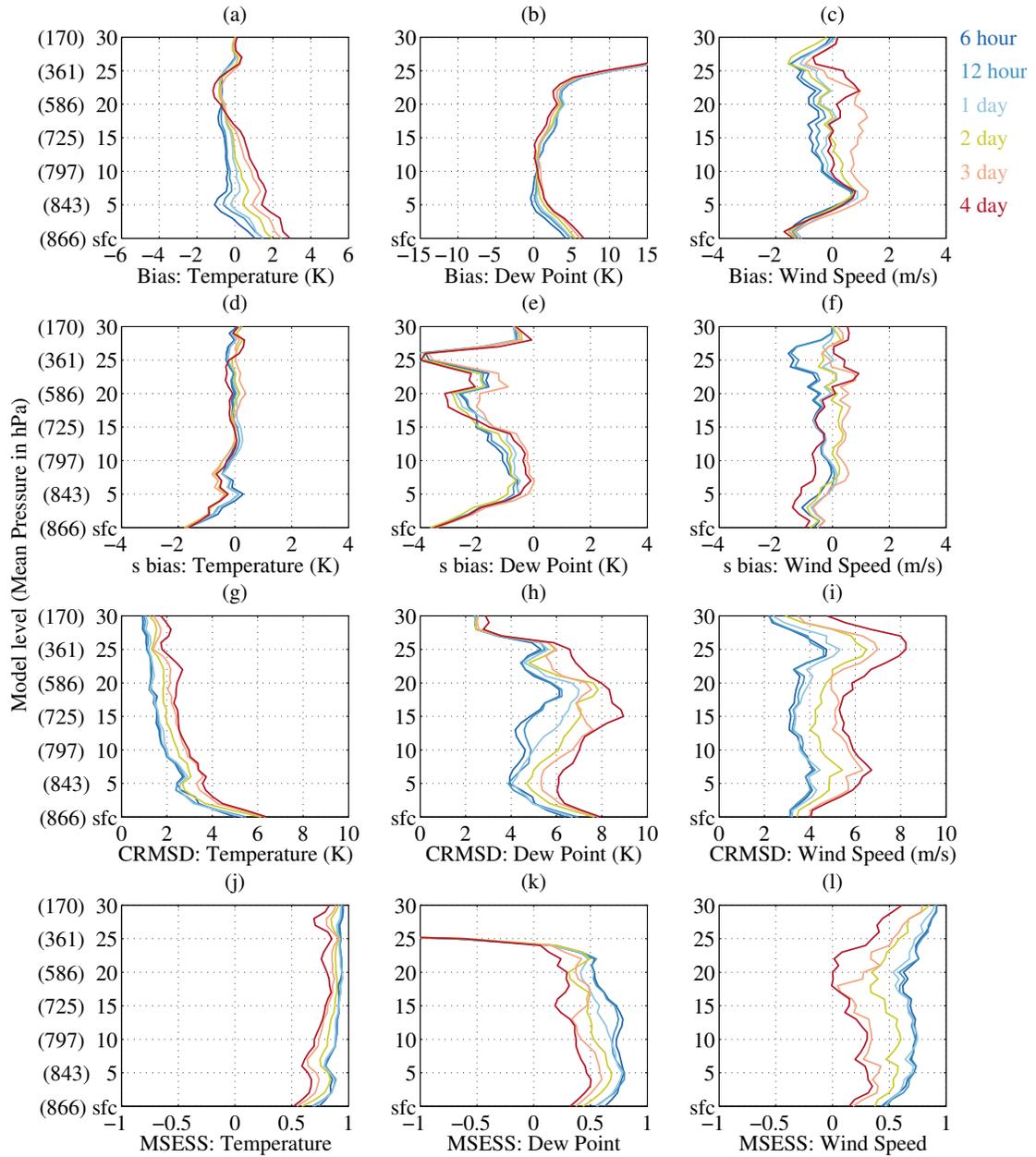


Figure 5.2: As in Figure 5.1 but for the LE subset.

5th and 15th model level (Figure 5.2b). Below the 5th model level there are large standard deviation biases and errors of approximately -3.5 K and 6-8 K respectively (Figures 5.2e and 5.2h respectively). Skill drops from approximately 0.7 at 6 hours to approximately 0.4 at the 4 day lead time. This coincides with a substantial increase in the errors with lead time (at some levels, errors double).

Similar to the HP subset, forecasts of wind speed have the highest skill above the 5th model level (Figure 5.2l). At the surface, during the 6 and 12 hour lead times MSESS is approximately 0.45 indicating marginal skill. This low skill coincides with a negative mean bias of approximately 1.5 ms^{-1} (Figure 5.2c), a small standard deviation bias of $0.5\text{-}1 \text{ ms}^{-1}$ (Figure 5.2f), and low errors (Figure 5.2i). Aloft there are high errors around the 25th model level, possibly associated with the jet-stream maximum, since it coincides with mean pressures of approximately 280 hPa. Further there is a 1 ms^{-1} mean wind speed bias around the 7th model level. Through the model integration there is a 0.2-0.7 decrease in MSESS, and a $1\text{-}3 \text{ ms}^{-1}$ decrease in errors such that by the 4 day lead time skill is low and errors are high.

5.3 Permanent Ice Shelves and Sea Ice (IC)

AMPS forecast skill for temperature in the IC subset is comparable to skill over the continent (Figure 5.3j). MSESS values are greater than 0.75 above the 5th model level at the 6 hour lead time. At the surface MSESS is 0.6 indicating mediocre skill. Relative low skill at the surface coincides with a mean cold bias of 2 K (Figure 5.3a) and relative large CRMSDs of 4 K (Figure 5.3g). Standard deviation biases are small, approximately 0.5 K (Figure 5.3d). This indicates that low surface skill is due to cold biases and high errors. MSESS decreases through the model integration by only 0.1 at the surface. However at the 25th model level, MSESS decreases through the model integration by 0.4 to around 0.45 at the 4

day lead time. This does not coincide with any notable biases or relative large errors thus it is most likely due to a decrease in the correlation between the model and observations. Mean biases at the surface decrease to 0 K by the 4 day lead time. However this warming of temperature with lead time leads to a warm bias of nearly 2 K at the 13th model level.

Dew point forecast skill is good only below the 13th model level at the 6 hour lead time where MSESS is approximately 0.75 (Figure 5.3j). Above here MSESS decreases to approximately 0.45 at the 20th model level, where there are also positive biases of 3 K (Figure 5.3b), low standard deviation biases of 3 K (Figure 5.3e), and high CRMSDs of 6 K (Figure 5.3h). Above the 25th model level (where the observations are not reliable) skill is low and bias is large. Through the model integration, MSESS decreases by 0.2 below the 10th model level such that skill is mediocre to poor by the 4 day lead time. Above the 10th model level skill is poor, or worse than climatology by the 4 day lead time. By the 4 day lead time there is also a decrease in variability with a standard deviation bias of -1 K after 4 days.

AMPS forecast skill for wind speeds increases with height (Figure 5.3l). At the surface, there is a relative low in MSESS of 0.5 at the 6 hour lead time. MSESS increases to 0.9 above the 25th model level. Low skill at the surface coincides with slow mean bias and low standard deviation bias, both of 1 ms^{-1} (Figures 5.3c and 5.3f respectively). After 1 day there are large decreases in skill and increases in errors that lead to MSESS of less than 0.2 and errors of $4\text{-}10 \text{ ms}^{-1}$ at the 4 day lead time. Aloft, above the 25th model there is large positive mean bias, standard deviation bias and errors at the 4 day lead time.

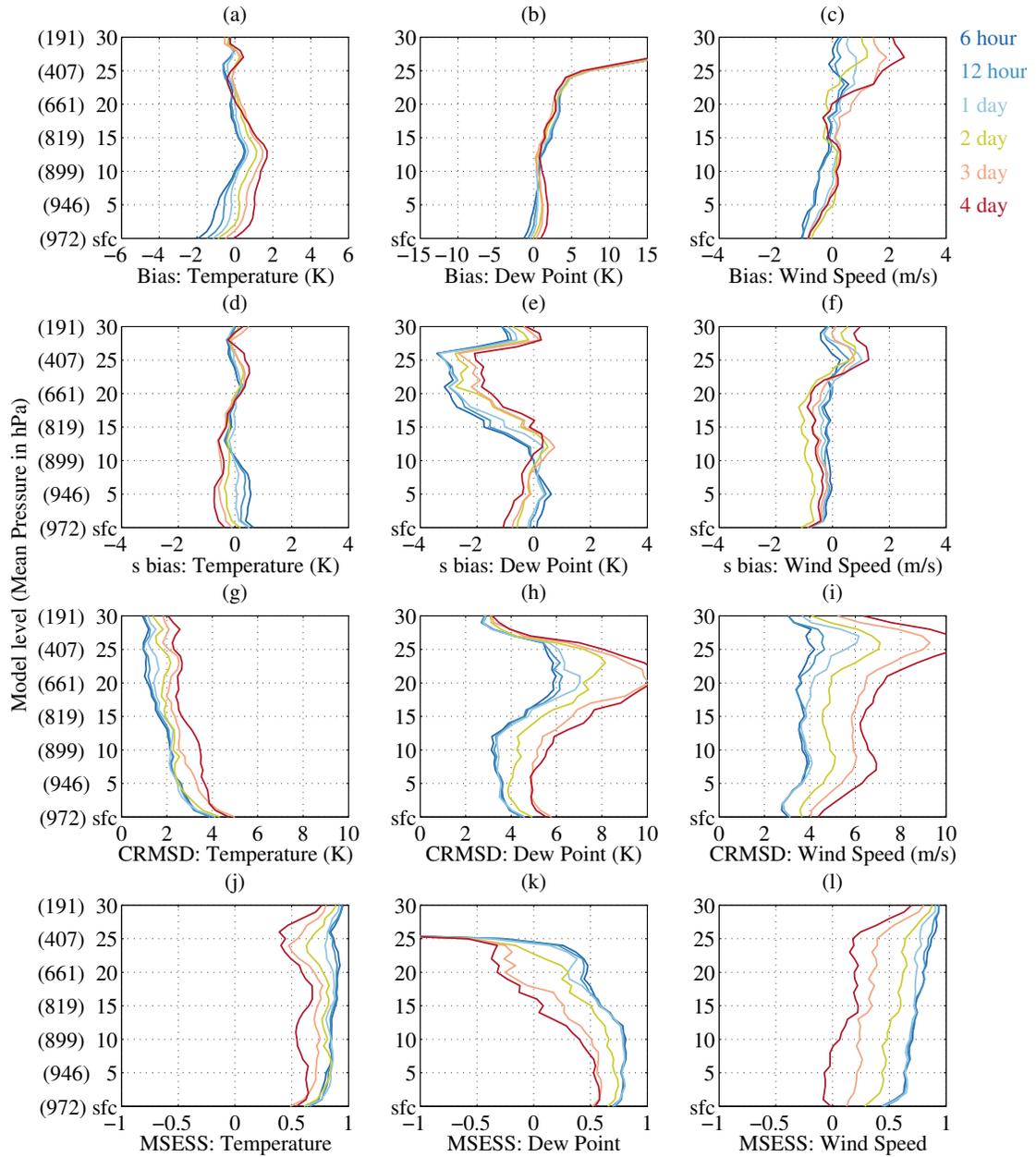


Figure 5.3: As in Figure 5.1 but for the IC subset.

Chapter 6

In Depth Analysis of AMPS Biases

In this Chapter we discuss the biases by their various subsets, as discussed in Section 3.4. Here, statistically significant biases are referred to as significant and all biases presented are mean biases but will be referred to as biases for conciseness.

6.1 Continental High Plateau (HP)

Figures 6.1a and 6.1b show the temperature biases for the HP subset for day and night respectively. There is a similar pattern of biases in both (as in Figure 5.1a), although the magnitudes vary between day and night. During day, before 12 hours there are significant positive (warm) biases of 1-3 K at the surface and 1st model level. At these lead times, between the 3rd and 8th model levels there are significant negative (cold) biases of 2-4 K. The low-level warm biases become stronger and extend through a deeper layer with lead time such that by the 84 hour lead time there are warm biases that extend through the 6th model level reaching a maximum of 6 K. Aloft, between the 14th and 23rd model levels, there are significant cold biases of 0.5-2 K. At lead times of 36 hours or earlier, these small cold biases connect with the larger cold biases at early lead times. During night, the pattern of biases are the same, except that low-level biases are cooler and biases aloft are weaker. This pattern manifests itself with stronger cold biases

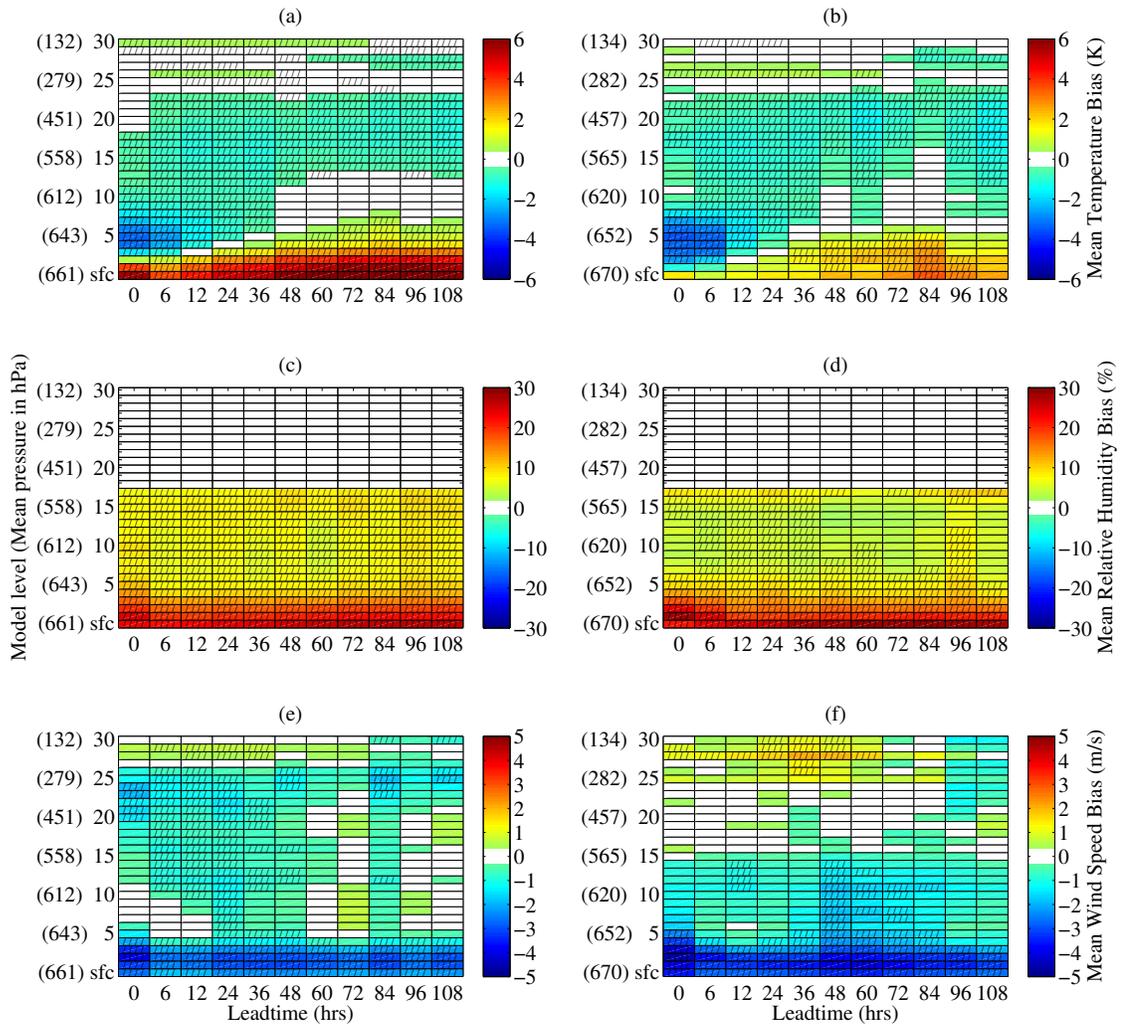


Figure 6.1: Biases for the HP subset for (a) temperature and day, (b) temperature and night, (c) relative humidity and day, (d) relative humidity and night, (e) wind speed and day, and (f) wind speed and night. Statistically significant biases are hatched.

and weaker warm biases in the low levels. As a result, warm biases are no larger than 4 K and are only significant between the surface and 5th model level, at lead times between 36 and 96 hours. Furthermore, low-level cold biases at early lead times are larger than during the day; 3-5 K. Aloft, biases are more sporadically significant and weaker; 0.5-1 K.

Relative humidity biases (Figures 6.1c and 6.1d), are always positive (e.g. the model is too moist). For both day and night, there are larger biases of 20-30 % in the low-levels (between the surface and 3rd model level). As per the discussion in Section 2.2, it is difficult to decipher whether the biases aloft (beyond 500 hPa) are a result of an observational bias or model bias, thus relative humidity biases at these levels have been whited out in the figures. The only difference in relative humidity biases between day and night is the difference in the magnitude of the biases between the 5th and 20th model levels. During day relative humidity biases are in the range 8-15 % but during the night they are in the range 3-10 %. Furthermore during night, many biases are not significant.

Relating these biases to the observed conditions, Figures 4.1a and 4.1b show the mean observed and AMPS forecast low-level vertical profiles of temperature and dew point for day and night respectively. These profiles indicate that at early lead times the couplet of warm and cold biases in the low-levels is due to warm surface temperatures, with weak lapse rates in the SBI and cold temperatures atop the SBI. The temperatures in the SBI then increase with lead time to produce warm biases throughout the SBI. Weaker warm biases during the night are due to improved representation of the surface temperatures in AMPS. Above the SBI there are small cold temperatures, as shown by the biases, and AMPS temperatures slowly change with lead time to be more similar to those of the dropsonde observations.

High dew points throughout the column are consistent with high relative humidity biases. Similar to the temperature SBI, the dew point SBI is poorly represented in AMPS with high surface dew points and weak SBI lapse rates. Also similar to temperature, there is a distinct increase in the moisture with lead time. Relative humidity biases do not increase with lead time since temperature and dew point biases increase in parallel. Above the SBI dew points are all high. Like with temperature, representation of the dew point SBI is improved at night, with lower surface dew points and steeper SBI lapse rates than during day.

There are significant negative (slow) wind speed biases of 3-4 ms^{-1} at the surface and first 3 model levels during both day and night (Figures 6.1e and 6.1f). The only other significant pattern in wind speed biases during day is the 0.5-1.5 ms^{-1} biases between the 6 and 48 hour lead times above the 10th model levels. During night, there are positive (fast) biases at lead times before 60 hours above the 25th model level.

Observed and AMPS forecast wind speed profiles show that the significant slow biases are due to an overly large decrease in wind speed with height below the katabatic maximum (Figure 4.2a and 4.2b). AMPS forecasts of the katabatic maximum are too slow by around 1 ms^{-1} during both day and night. Above this, AMPS forecast wind speeds decrease exponentially with height despite the observed structure of wind speeds having a more complex decrease with height. This results in the weak slow biases shown in Figures 6.1e and 6.1f.

6.2 Continental Low Elevations (LE)

Temperature biases in the LE subset follow a similar structure to those over the HP (Figures 6.2a and 6.2b). Surface biases are positive through all lead times although they are much weaker (relative to the HP subset), 0.5-4 K. During day surface biases are significant at the 6 hour lead time and after, but during night

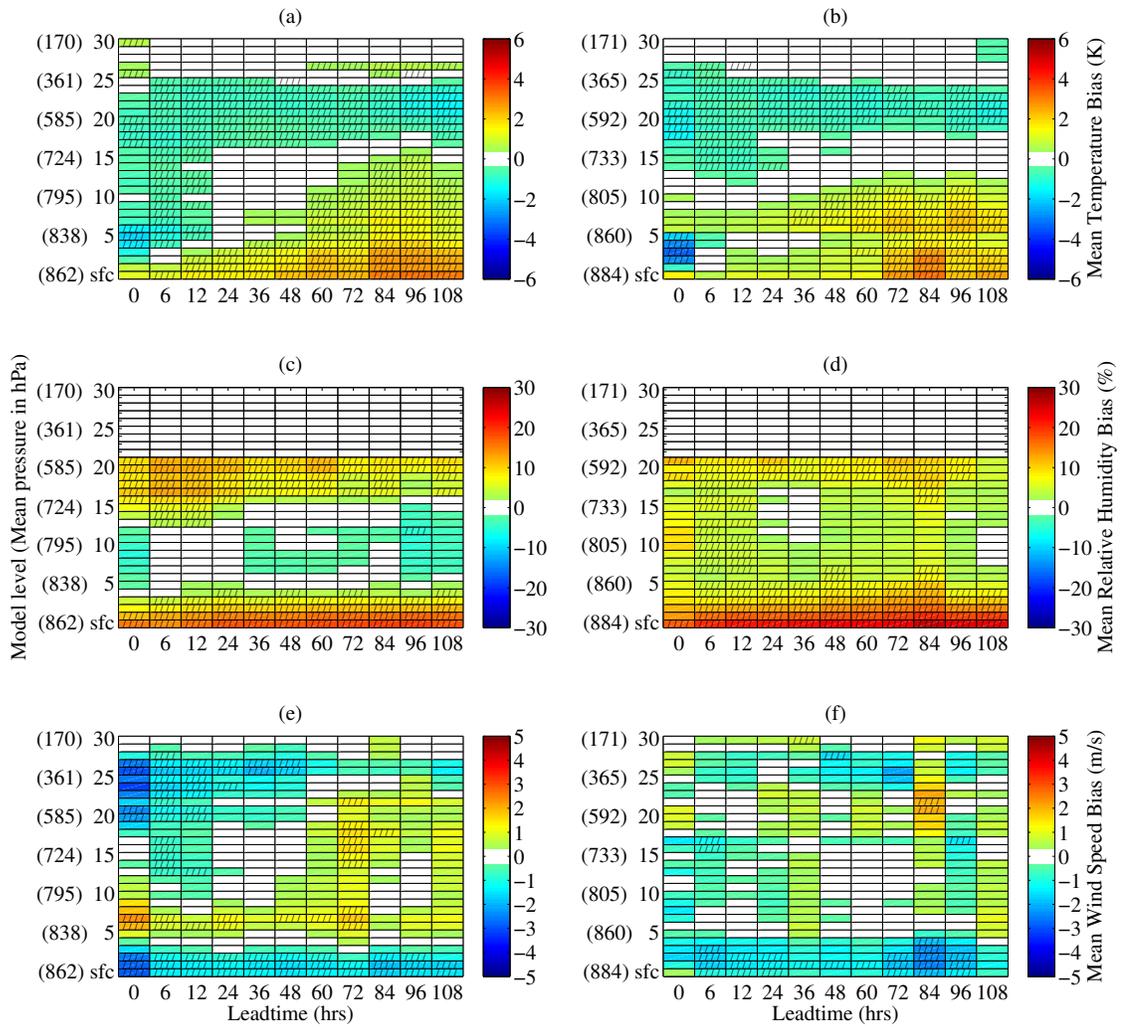


Figure 6.2: As in Figure 6.1, except for the LE subset.

they are only significant at the 72 hour lead time and after. These warm biases also increase in height with lead time more than in the HP subset. During day significant warm biases extend through the 15th model level by the 96 hour lead time, but during night the significant warm biases only reach the 7th model level. At early lead times there are again cold biases between the 3rd and 8th model levels, which are most prominent during the 0 hour lead time. Aloft, between the 15th and 25th model levels there are again significant cold biases of 0.5-1.5 K (at most lead times).

Relative humidity biases during day have a tri-pole pattern with significant moist biases of 10-20 % at the surface through 4th model level, small dry or no biases from the 5th model level through the 15th model level, and some moist biases of up to 12%, above the 15th model level (Figure 6.2c). During night, biases are moist through all levels and lead times (Figure 6.2d). At the surface and first 5 model levels biases are significant, from 15-25 %. From the 6th model level through the 18th model level biases vary from 1-10 % and are only significant during the 6, 12 and 84 hour lead times. Above these levels biases are moist, up to 10%.

Similar to the HP subset, during both day and night, AMPS forecasts a weak lapse rate and high surface temperatures in the mean temperature SBI relative to that observed (Figures 4.1c and 4.1d). This results in the similar structure of temperature biases between the HP and LE subsets. However, since the observed and forecast temperature SBIs are weaker in the LE subset, temperature biases are smaller in magnitude. Temperatures in the SBI exhibit similar warming with lead time as in the HP subset, resulting in the progression of temperature biases with lead time seen.

Dew point biases vary between day and night. During day, the observed SBI is not forecast in AMPS, with isothermal profiles at early lead times progressing to

profiles with negative lapse rates at later lead times (Figure 4.1c). This resulted in the moist surface relative humidity biases and dry relative humidity biases atop the SBI. Above the SBI dew points have small or no bias. During night, the observed dew point SBI is forecast but it is weak (Figure 4.1d). Dew points are too high during night throughout the profile with the smallest biases around a mean pressure of 750 hPa. As with all temperature and dew points already discussed there is an increase in the dew point with lead time.

Wind speed biases are negative (slow) at the surface and first 3 model levels during both day and night (Figures 6.2e and 6.2f). These biases are significant for only the surface, 1st model level, and some lead times at the 2nd model level during the day, and at the 1st model level and in some lead times at the 2nd and 3rd model levels during night. Around the 6th and 7th model levels wind speed biases are positive (fast) although they are only significant at the 7th model level and at other levels for some lead times during the day. Another significant pattern is the slow wind speed biases aloft, between the 20th and 30th model levels during the 6, 12, 24, 36, and 48 hour lead times during day. During night, there are some significant biases but there are no clear patterns.

Figures 4.2c and 4.2d show that the slow surface wind speed biases during both day and night are due to a larger decrease in the winds below the LLWM. In the LLWM biases are different between day and night. During day the observed winds are slower than the forecasts winds during most lead times. During night there is a much larger observed LLWM and forecast wind speeds are slower than the observed as a result.

6.3 Permanent Ice Shelves and Total Sea Ice Cover (IC)

In the IC subset during day, temperature biases are cold during the first 12 hours below the 8th model level (Figure 6.3a). These cold biases are significant for both 0 and 6 hour lead times and between the 4th and 7th model levels at the 12 hour lead time. After the 24 hour lead time there is an area of significant warm biases centered on the 13th model level. These warm biases are largest around the 84 hour lead time with biases of around 1.5 K. After 72 hours these biases extend to the surface but are not significant. During night, there are again cold biases in the early lead times below the 8th model level (Figure 6.3b). In contrast to day, there are significant cold biases at the surface through all lead times. During night there is a significant warm bias between the 10th and 14th model levels through all lead times after 6 hours. At the later lead times (after 60 hours) these biases are strongest, approximately 2 ms^{-1} , and extend through a deeper layer, from the 7th model level through the 17th model level. Aloft there are sporadic warm and cold biases during both day and night with some of these biases significant.

Relative humidity biases during the day are predominantly moist except for between the 10th and 15th model levels after 36 hours (Figure 6.3c). In this area, at these lead times biases are small or slightly negative, but not significant. Below the 10th model level biases are significant and below 10 %. Above the 20th model level biases are significant, up to 15%. During night, (Figure 6.3d), there are distinct significant dry biases, of 5-10 %, between the 9th and 14th model levels, after the 24 hour lead time. This region of biases is co-located with warm temperature biases. Above and below this, biases are the same as during day.

Vertical profiles of low-level temperature are well forecast at the early lead times except for in the lowest levels where forecast profiles have weaker lapse rates than

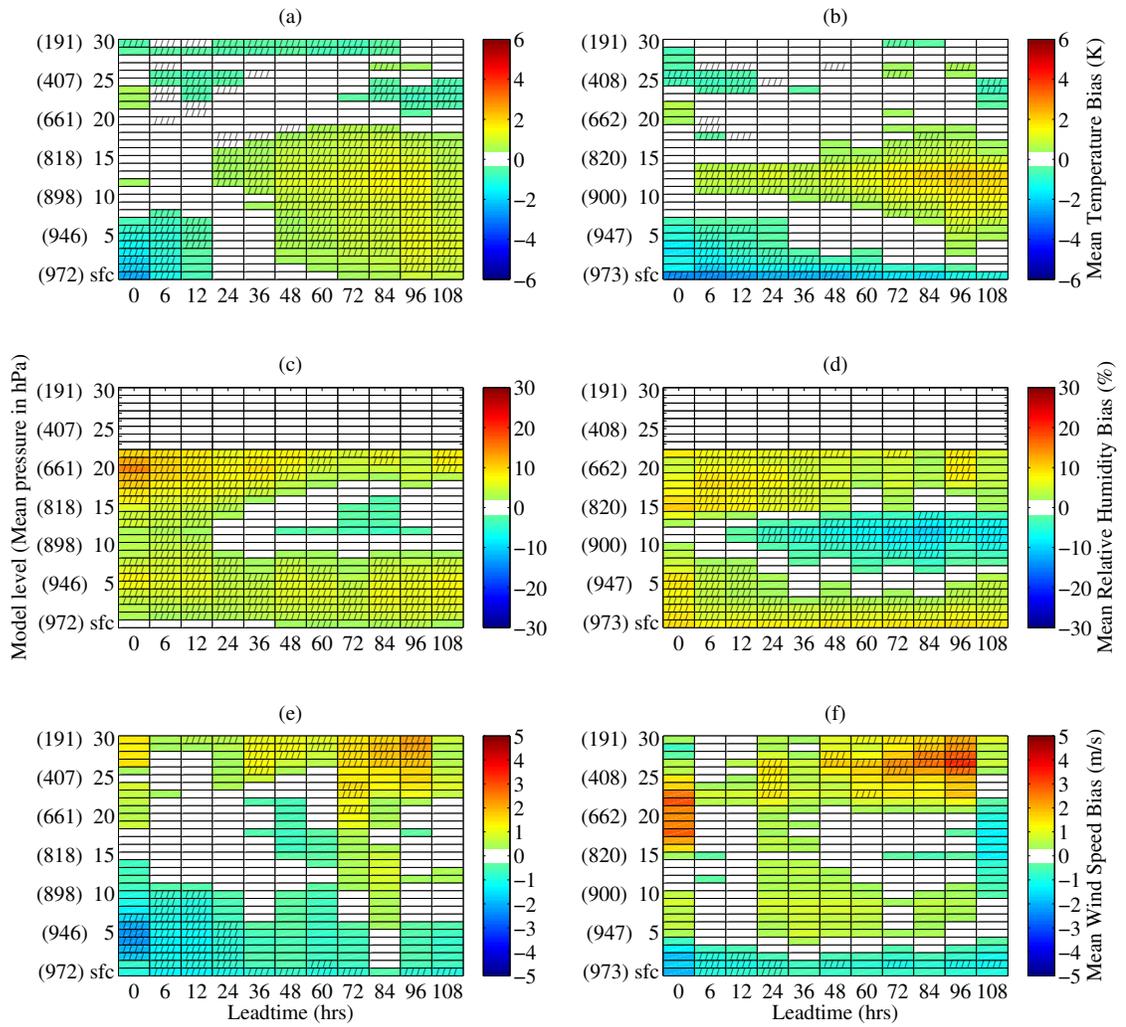


Figure 6.3: As in Figure 6.1, except for the IC subset.

those observed (Figure 4.1e and 4.1f). At later lead times temperatures increase resulting in the warm biases through the low levels at later lead times during day. During the night, lower observed temperatures are not reproduced in AMPS in a layer between 800- and 900-hPa. Since dew point profiles are well-forecast in this layer, this results in dry relative humidity biases. During day, dew point profiles are high through all levels and there is an increase of dew points with lead time.

There are two clear patterns in wind speed biases in the IC subset (Figures 6.3e and 6.3f). The first pattern is a region of low-level slow biases. These are predominantly at the surface and first 3 model levels but there are also biases during the day through the 10th model level before 24 hours lead time. For most lead times the surface biases are significant. The second bias is an area of fast wind speeds aloft, primarily at the later lead times, between 72 and 96 hours.

The lack of any large biases in the wind speeds is portrayed in low-level vertical profiles of wind speed (Figures 4.2e and 4.2f). Similar to the HP and LE subsets, there are slow wind speeds at the surface as a result of too large of a decrease in the winds below the LLWM. The LLWM and winds above this are overall accurate.

6.4 Inversions

Since one of the primary issues for AMPS forecasts of temperature and dew point is the inversion we now focus our analysis on this feature of the Antarctic environment. Biases related to inversions vary strongest in the first 24 hours of the model integration (Figures 6.1a, 6.1b, 6.2a, and 6.2b). Therefore we hypothesize that the initial conditions, are a significant cause of the poorly represented inversions. As such, the same analysis is carried out on the GFS forecasts from which the initial conditions are provided for AMPS.

The 95% confidence intervals for AMPS forecast biases of temperature and dew point inversion strength (in red) during day and night, in both the LE and HP

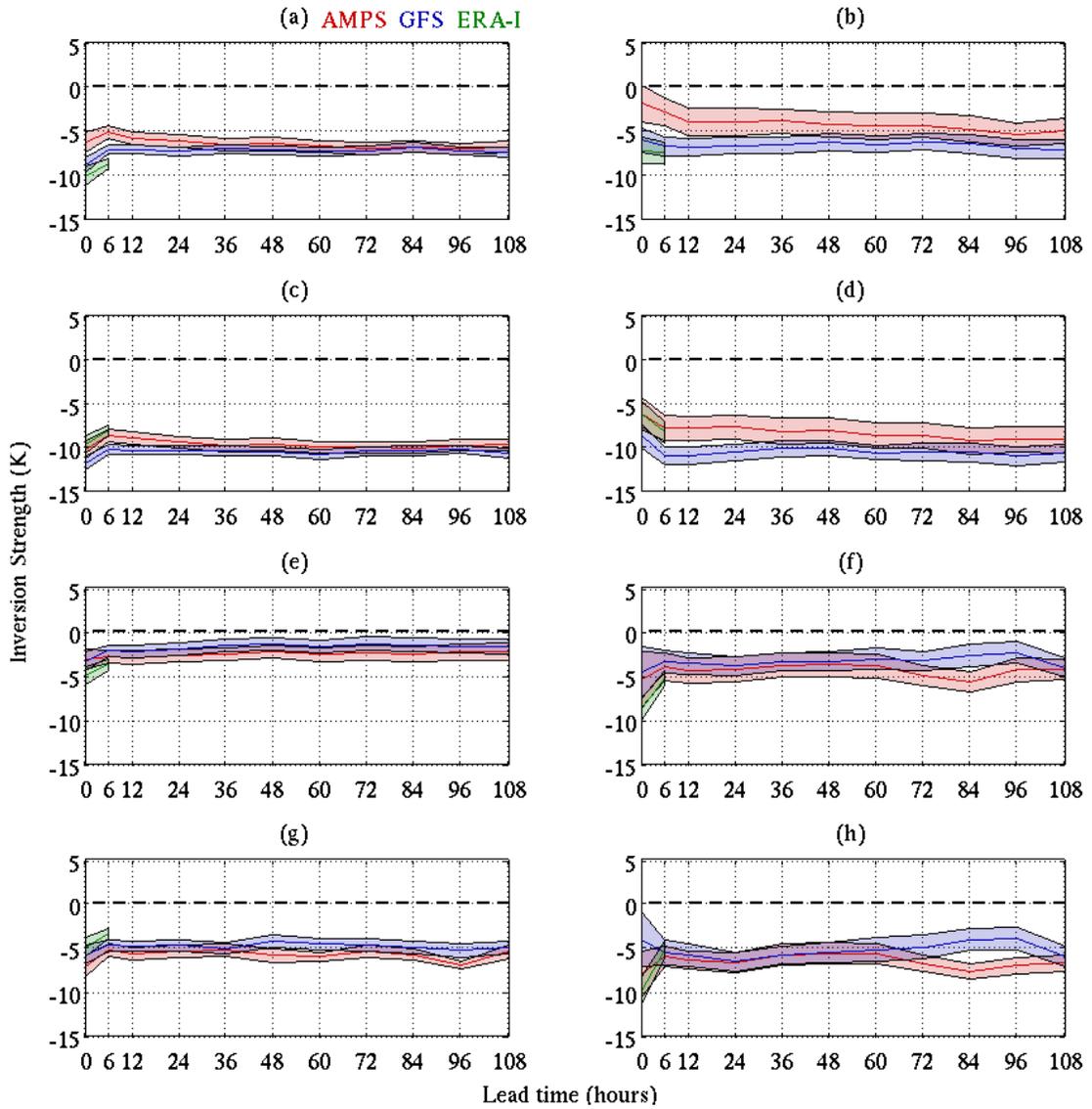


Figure 6.4: Mean Inversion strength (IS) biases for temperature, (a) day and HP, (b) temperature, night and HP, (c) dew point, day and HP, (d) dew point, night and HP, (e) temperature, day and LE, temperature, (f) night and LE, (g) dew point, day and LE, and (h) dew point, night and LE. AMPS is in red, GFS is in blue, and ERA is in green. Error bars indicate the 95% confidence intervals.

subsets are shown in Figure 6.4. With the exception of the 0 hour lead time for temperature during night in both subsets (Figures 6.4b and 6.4f), and the 96 hour lead time for temperature during night in the LE subset (Figure 6.4f), there are negative (weak) inversion strength biases in AMPS over the continent. Overall AMPS forecast inversions are weaker than the observed by 0-10 K.

For both subsets and both variables, AMPS inversions are less variable during the day (relative to night) since there are larger confidence intervals at night. The only significant difference between AMPS biases during day and night is for temperature inversions in the HP subset. AMPS temperature inversion biases range from 4-7 K during day and 0-6 K during night with significant differences at most lead times. Overall, AMPS dew point inversions have weaker inversion biases than temperature inversions with these differences significant at most lead times, both subsets, and during both day and night. Mean inversion biases are 2-5 K weaker in dew point inversions than temperature inversions. With the exception of AMPS temperature inversions during night (Figures 6.4b and 6.4f), the HP subset has weaker inversion biases than the LE subset. Mean inversion biases are 1-5 K weaker in the HP subset than in the LE subset.

The 95% confidence intervals for AMPS forecast biases of temperature and dew point inversion depth are shown in Figure 6.5. During the early lead times (0 and 6 hours) AMPS mean forecast inversion depth biases are positive (i.e. on average inversions are too deep in AMPS relative to the observations). Given the confidence intervals at these early lead times, deep biases are significant in the HP subset, and at night in the LE subset for temperature inversions (Figure 6.5f). AMPS inversion depth biases at the 0 hour lead time range from -100 to 350m, with mean biases of 0 to 200 m (dependent on subset and variable). After the 0 hour lead time, AMPS inversion depth biases transition from the deep biases discussed to either negative (shallow) biases or no bias. Between the 24 and 48 hour lead times (depending on

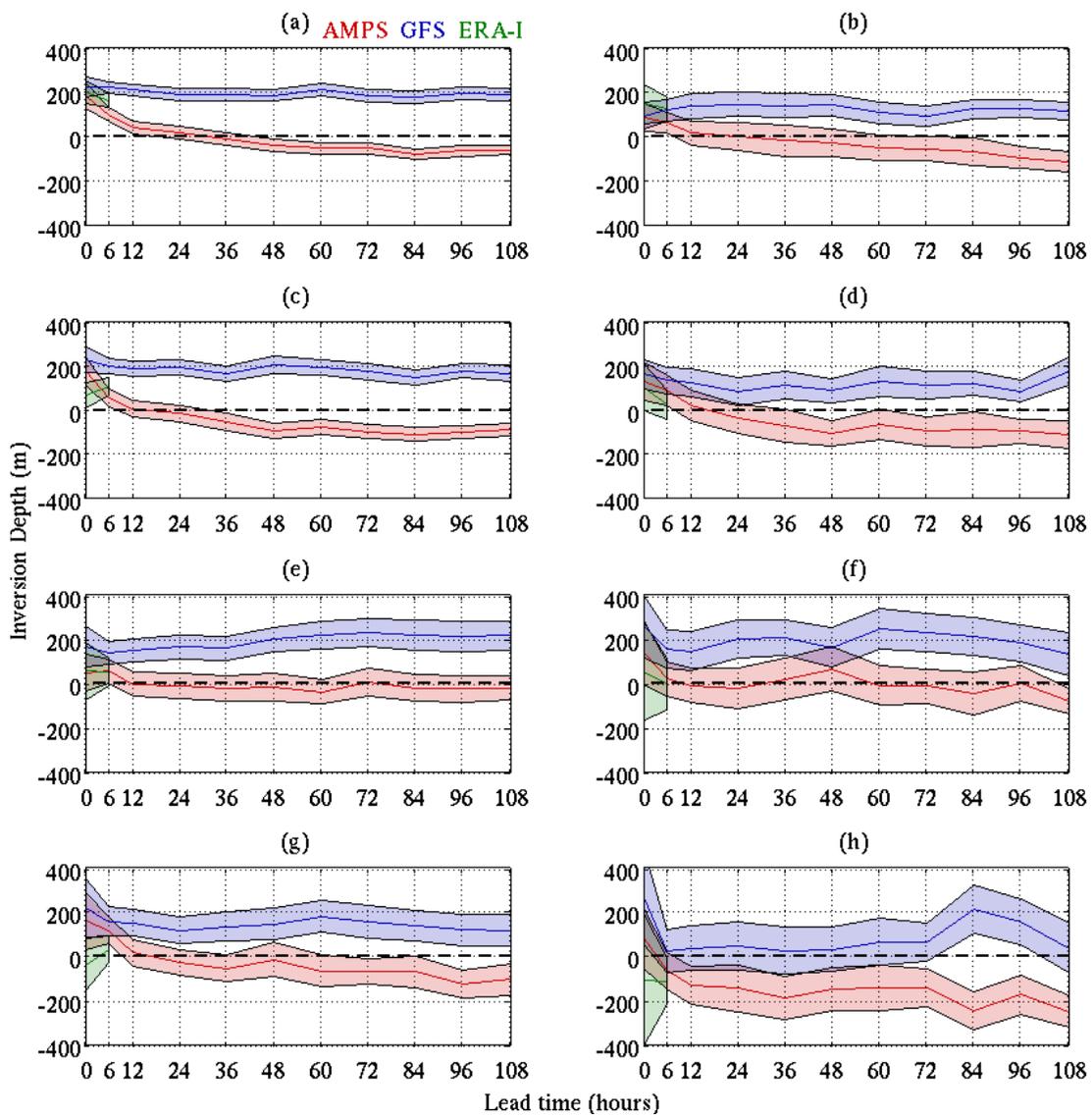


Figure 6.5: As in Figure 6.4 but for inversion depth (ID).

subset and variable) biases stop this transition and remain approximately constant for the rest of the model integration. Inversion depth biases after 48 hours range from -200m to 100m, with mean biases of -25 to -175 m. AMPS inversion depth biases are more variable during night than day as shown by narrower confidence intervals during day. There is little difference between temperature and dew point inversions over the continent, or between inversions in the HP and LE subsets.

When comparing AMPS depth biases at early lead times (0 and 6 hours) to GFS depth biases there are clear similarities between the two models. In all subsets confidence intervals indicate that inversion depth biases are very similar if statistically indistinguishable. Since AMPS initial conditions are provided by the GFS this is expected. It also implies that the data assimilation process in AMPS does not correct the inversion depth from the incorrect GFS initial conditions. While inversion strength biases at early lead times are similar between AMPS and GFS in the LE subset, they are significantly different in the HP subset. Here, AMPS biases are significantly lower suggesting there is correction in the data assimilation process to the inversion strength that improves the poor GFS initial conditions. This in turn appears to result in better forecast inversion strengths in the HP subset throughout the model integration.

In the IC subset, since inversion depth and strength are only on average 328.1 m and 4.5 K respectively (see Table 4.1 for values separated by day/night), there are few significant biases in inversion structure (not shown). However there are biases in the frequency of occurrence for the IC subset. For temperature inversions during day the observed occurrence is just below 0.5 of the subset throughout the model integration (Figure 6.6a). The occurrence of biases in AMPS are similar to this with just over 0.5 forecast during the 0 hour period and approximately the same as AMPS through the model integration. For temperature inversions during night the observed relative occurrence is 0.6 at the 0 hour lead time and 0.4 at later lead

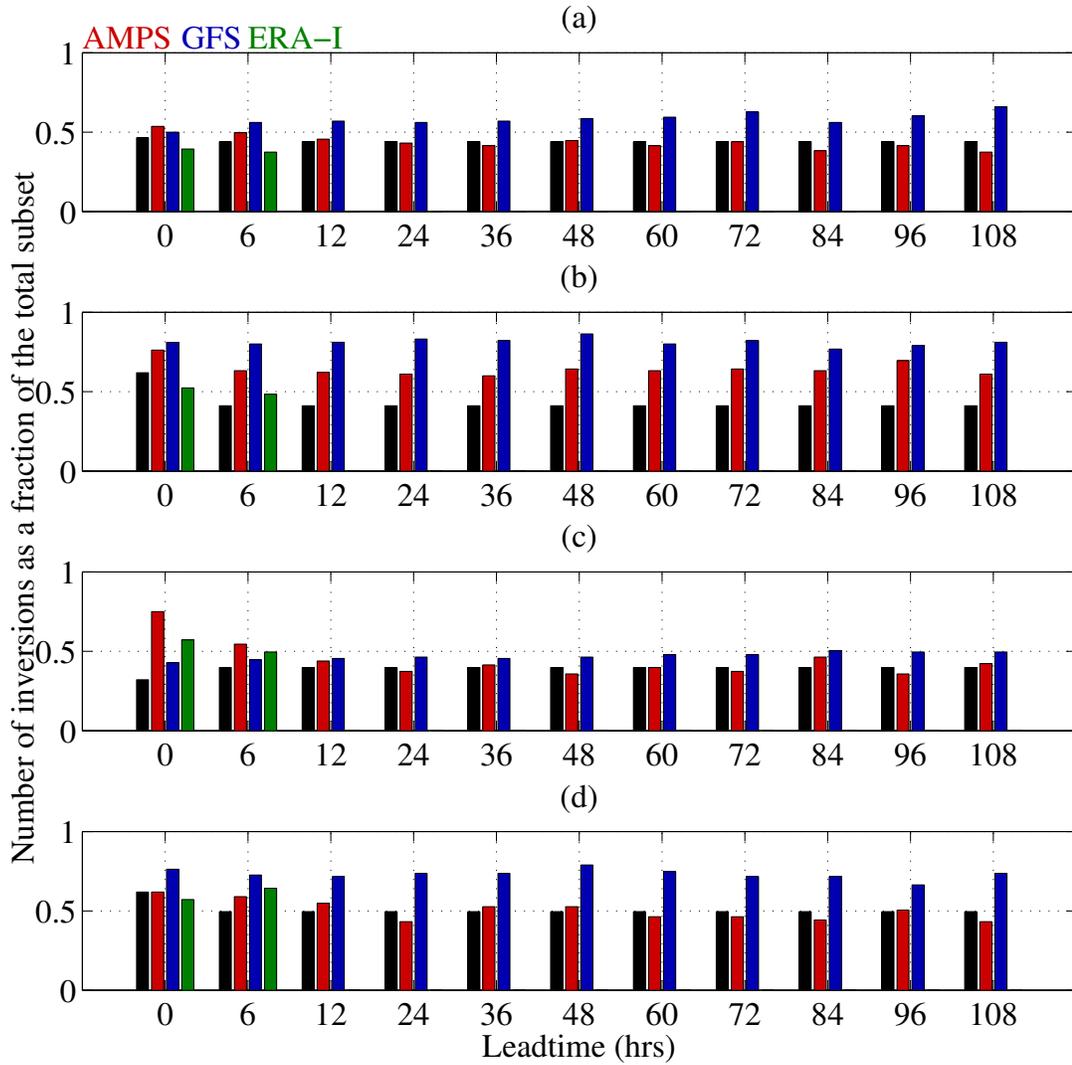


Figure 6.6: Fraction of inversions forecast as a total of the whole subset in the IC subset for (a) temperature during day, (b) temperature during night, (c) dew point during day, (d) and dew point during night. Observations are in black, AMPS is in red, GFS is in blue, and ERA is in green.

times (Figure 6.6a). AMPS forecasts too many inversions throughout the model integration, with a relative occurrence 15-25 % more frequent than the observations throughout.

Figures 6.6c and 6.6d shows the relative occurrence of dew point inversions during day and night respectively. During day the relative occurrence is 0.3 during the 0 hour lead time and 0.4 at later lead times. During 0 and 6 hours lead time the relative occurrence in AMPS is 0.45 and 0.15 larger than the observations respectively. At later lead times AMPS forecasts are approximately the same as the observations. During night dew point inversion relative occurrence is 0.6-0.5 with AMPS relative occurrence similar throughout.

6.5 Low-level Wind Maxima

A second area noted as an issue is the low-level wind maximum (LLWM) wind speed maximum associated with the katabatic flow; a prominent feature in the Antarctic boundary layer. Here we report and discuss the biases in AMPS associated with LLWMs.

Figure 6.7 shows the LLWM maximum wind speed 95% confidence intervals for the HP and LE subsets. With the exception of a few lead times the LLWM maximum wind speeds are significantly slow in the HP subset (Figures 6.7a and 6.7b). On average the LLWM maximum wind speed is approximately $1-2 \text{ ms}^{-1}$ too slow throughout. This represents a maximum wind speed that is 6-12% slower than those observed (Table 4.2). LLWM maximum wind speeds are also slow in the LE subset, although this result is not significant since there is greater uncertainty in the results for the LE subset represented by larger confidence intervals. Here, the biases are again $1-2 \text{ ms}^{-1}$ on average, which represents wind speeds that are 7-13% slower than those observed (Table 4.2). There is little change with lead time throughout the model integration in any of the subsets (Figures 6.7c and 6.7d).

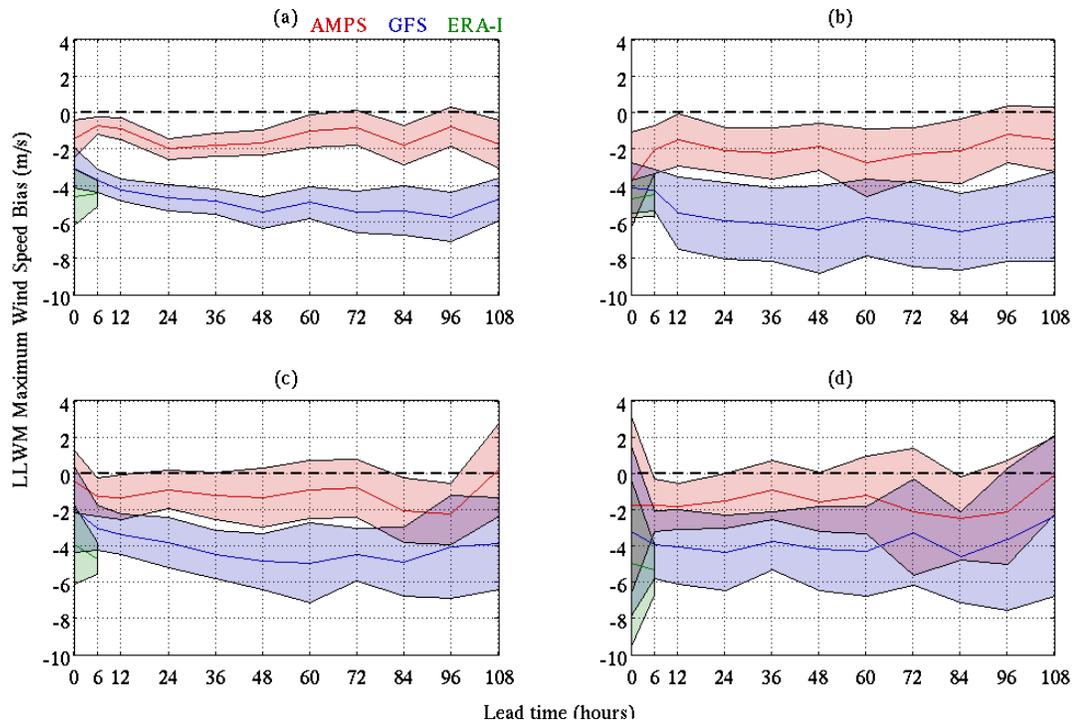


Figure 6.7: Mean LLWM maximum wind speed biases for (a) HP and day, (b) HP and night, (c) LE and day, (d) LE and night. AMPS is in red, GFS is in blue, and ERA is in green. Error bars indicate the 95% confidence intervals.

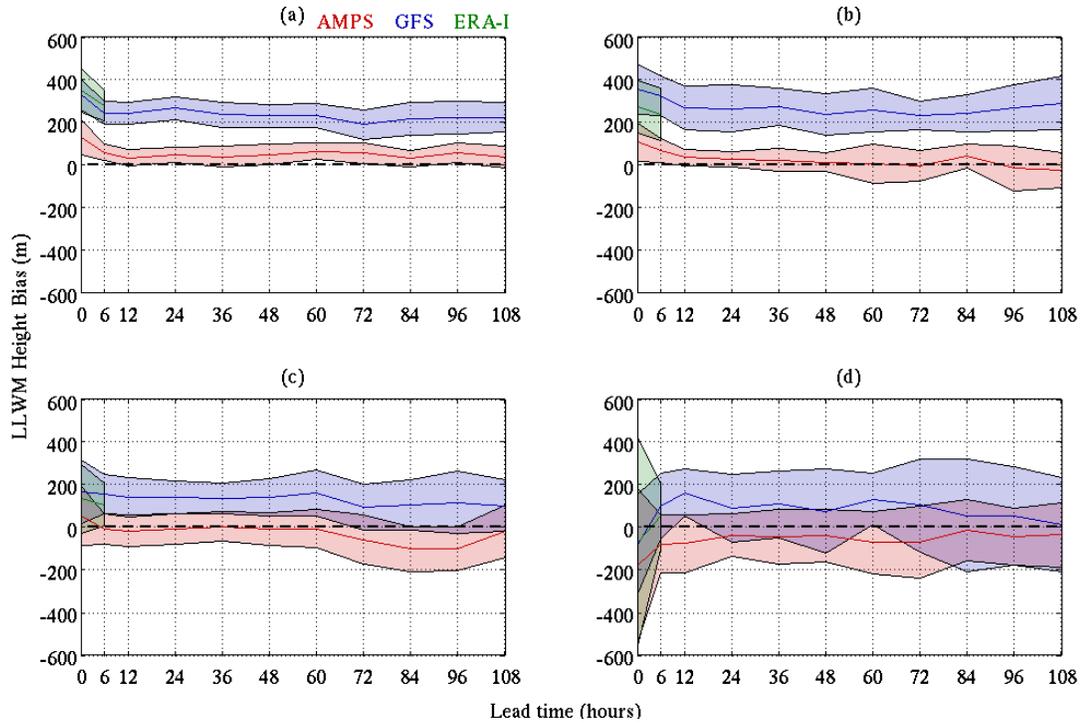


Figure 6.8: As in Figure 6.7 but for LLWM height.

There are no significant biases in the IC subset for LLWM maximum wind speeds (not shown).

LLWM height biases are shown in Figure 6.8. There are few LLWM height biases throughout the model integration with the only significant biases occurring at early lead times (0 and 6 hours) in the HP subset (for both day and night) and also at later lead times in the HP subset during day (Figures 6.8a and 6.8b). At the 0 hour lead time when biases are largest, they are on average 100-150m. There are no significant LLWM height biases in the LE subset with biases that are approximately zero or slightly less on average (Figures 6.8c and 6.8d).

Throughout all subsets LLWMs are predominantly forecast with low occurrence relative to observations (Figure 6.9). AMPS forecasts too few inversions by a relative occurrence of 0-0.3 throughout all subsets. The only exception to this bias is in the LE subset during night at the 0 hour lead time. However there are only

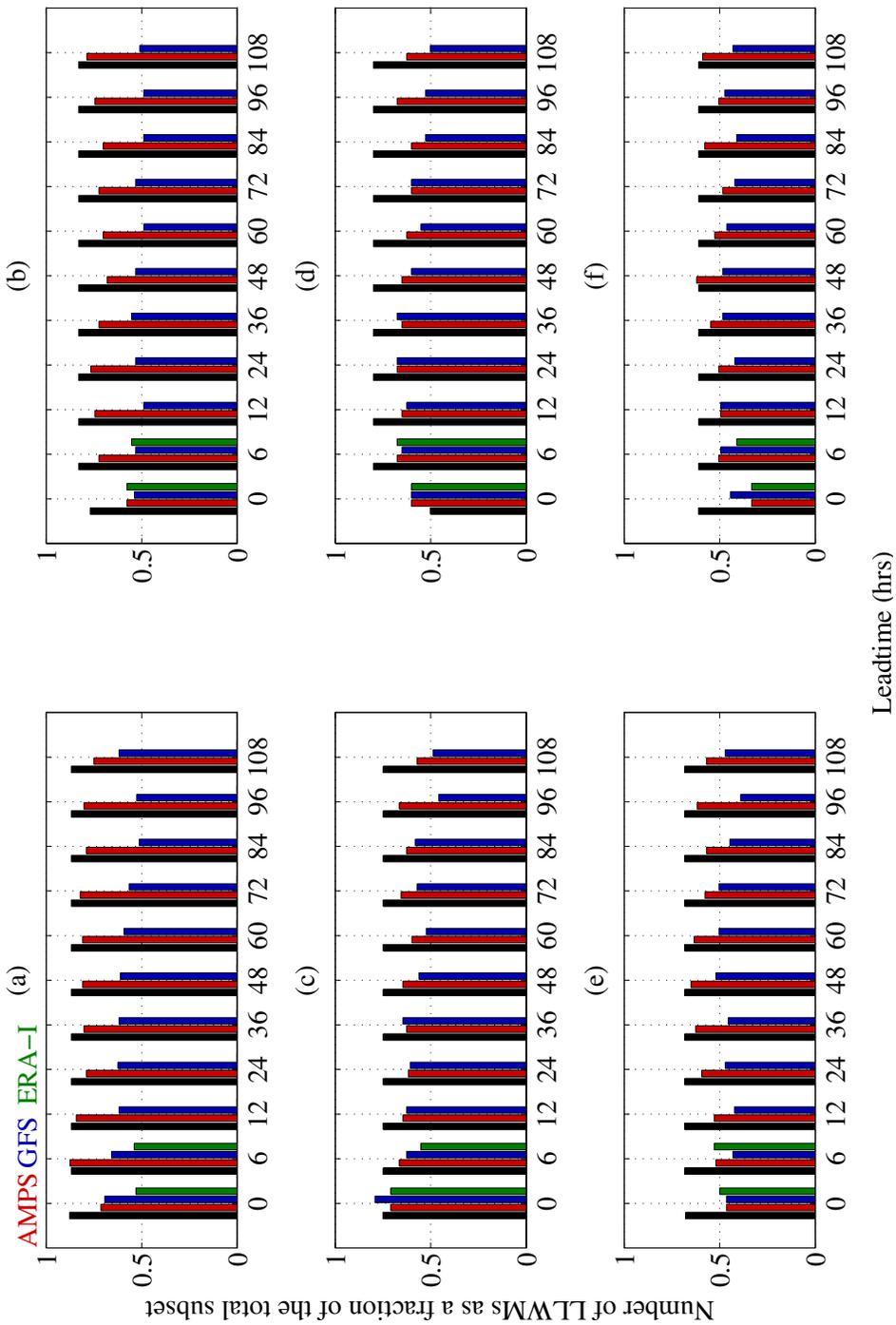


Figure 6.9: Fraction of LLWMs forecast as a total of the whole subset for (a) HP and day, (b) HP and night, (c) LE and day, (d) LE and night, (e) IC and day, and (f) IC and night. Observations are in black, AMPS is in red, GFS is in blue, and ERA is in green.

10 profiles in this subset so the statistical significance of this result is low. The largest relative occurrence biases of -0.3 occur during the 0 hour lead time in the HP and IC subsets. In the LE subset during the 0 hour lead time AMPS forecasts approximately the correct number of inversions. There are no trends in AMPS forecasts of LLWM occurrence with lead time.

6.6 Discussion

Chapter 5 provides a broad overview of the errors associated with AMPS. In particular, during this chapter we noted distinctly high mean biases, standard deviation biases, and errors, and low skill close to the surface in nearly all variables, especially over the continent. These statistics suggest that the area of poorest forecast performance in AMPS occurs in the boundary layer over the continent, with particularly large biases over the high plateau. Also, there were large errors that coincided with the position of the upper level jet stream in all subsets. Large changes between the 1 and 2 day lead times indicate that after 2 days the position of the upper-level jet may be inaccurate in AMPS.

Patterns of biases in Chapter 6 build on the picture provided in Chapter 5. There are some persistent biases in AMPS analyses and forecasts that are collocated with large values of the other statistics shown in Chapter 5. Key biases over the continent include biases associated with low-level inversions and winds, which are strongest over the high plateau. Over the ice surrounding the continent, the clearest biases are the warm and dry biases atop the boundary layer. In this section we will discuss these issues and propose causes.

One of the key points noted in Chapter 5 is the large changes in model forecast conditions at early lead times. In the first 1-2 days of the model integration there are rapid changes in temperature and dew point biases. Further, there are rapid changes to the inversion depth during the first 1-2 days of the model integration.

These changes suggest one point; there are biases associated with the analysis that are altered or corrected later in the forecast and the model spin-up period to remove these biases is around 1-2 days.

Biases at early lead times noted in Chapter 6 include the couplet of warm and cold biases at early lead times over the continent associated with poor estimation of strong observed inversions. Further analysis of the inversions showed that the deep inversion biases at early lead times are associated with the initial conditions provided by GFS. Since biases are similar to the GFS, it is likely that the lack of observations discussed in Section 2.2 is affecting the analysis, through little correction to the GFS initial conditions (NCEP's Global Data Assimilation System: GDAS). Therefore, early lead time errors emphasize the requirement for more in-situ low-level observations over Antarctica to improve data assimilation through more corrections to the initial state provided by GDAS, especially over the high plateau. However this is a very costly solution to the initial condition problem. Therefore other solutions should be sort. Such initial condition problems could be alleviated or removed through more advanced data assimilation systems than the current WRF three dimensional variational system (WRF 3D-Var Barker et al., 2006). Liu and Xiao (2013) show how in replacing the homogeneous and isotropic background error covariance used by the standard WRF 3D-Var, with a flow-dependent background error covariance in an ensemble four-dimensional variational system, analysis and forecast accuracy are both improved. This improvement is partly made through the inclusion of more observations and utilizes the current limited observational network over Antarctica more efficiently.

At later lead times, after the 1-2 day spin-up period, low-level biases dominate the forecast problems with positive temperature, dew point, and relative humidity biases caused by weak inversions over the continent in the HP and LE subsets. Temperature and dew point biases also increase with lead time. Further, there

was also low variability and skill, and high errors in temperatures and dew points close to the surface at early lead times. A ubiquitous surface of ice and snow over the continent leads to high fluxes of heat from the atmosphere to the surface. All the discussed issues imply that there are upward sensible and latent heat fluxes are too strong in AMPS forecasts. Such flux errors therefore imply that land surface temperatures may be incorrect. Slow wind speeds in AMPS could imply low mixing which in turn would lead to the warm and moist surface temperatures. Further, shallow inversions in the forecast imply low vertical mixing since the large surface properties are not advected as high. There are also differences between biases during day and night which imply that radiative properties of the surface are incorrect. Surface biases of temperature and dew point are consistent with the surface biases presented in Bromwich et al. (2013b), the only other study with comparable evaluations of Polar WRF over the continent, at many sites during autumn. These results suggest that parameterization of the strongly stratified boundary layer over the Antarctic interior is inaccurate, with multiple issues.

The only notable issue AMPS encounters in forecasting over sea ice is an over prediction of temperature inversions during night. Since GFS provides the initial conditions and GFS also over-forecasts inversions at early lead times it is possible that the over-forecast in AMPS is a result of poor initial conditions as already discussed. Alternatively the cold surface biases could be the cause of an over-forecasting of inversions. Cold surface biases imply high downward surface fluxes, a known issue in modeling the stable boundary layer as discussed by Holtslag et al. (2013).

In addition, the low-level kinematic properties in AMPS are incorrect, with slow surface wind speed biases through all subsets of the data (but are most prominent over the HP subset). Such surface wind speed biases suggest a high surface drag in the model although this is inconsistent with Bromwich et al. (2013b) who shows

positive wind speed biases. Further, other studies suggest that WRF exhibits low surface drag and positive surface wind speed biases over plains and valleys (Jiménez and Dudhia, 2012). However, it is noted that other studies have not evaluated WRF over the unique environment of Antarctica and many of the observations from Bromwich et al. (2013b) are coastal representing only a very specific area of the Antarctic atmosphere. It is shown by Kumar et al. (2012) that Polar WRF exhibits similar slow wind speed biases over the Maitri region of Antarctica. Given contrasting results, the result presented here should be validated through further evaluation over the Antarctic.

In a summary of AMPS forecasts for LLWMs, biases are largest in the HP subset with the wind speed maximum too weak and the LLWM displaced vertically from that observed during day, significant only in the HP subset. Further, AMPS forecasts too few LLWMs throughout the model integration. Since the katabatic wind is a function of the surface fluxes, slope, and the surface drag, and discussed previously is likely low, this is the likely source of error for the katabatic flow. However, the low level warm biases encountered over the continent could also account for the low wind speeds in the katabatic flow through increased buoyancy. Qualitatively, these two factors likely account for the slow katabatic maxima over the continent that is significant over the HP. There are no model evaluations that focus on the katabatic flow over the continent.

Finally, a warm and dry bias between approximately 750- and 900-hPa over the sea ice during night (and to a much lesser extent in other subsets) is hypothesized to be the result of a forecast low cloud fraction in AMPS. Qualitatively the dry bias is expected to be the result of low water vapor concentration, while the warm bias is expected since there would be less radiative cooling in the model (relative to reality) given a lower cloud fraction. Further, this is in a layer where low cloud is observed in abundance over the sea ice and Antarctica's low elevations

as discussed by Bromwich et al. (2012). A low cloud fraction in NWP models over the southern ocean and Antarctica is a known problem that has been studied relatively heavily in the literature. In AMPS, this problem was first diagnosed in 2008 when Fogt and Bromwich (2008) documented a low cloud fraction over McMurdo. Fogt and Bromwich (2008) attempted to fix this error by developing a new empirical cloud fraction algorithm for AMPS. While this algorithm fixed the cloud fraction bias, it was noted that it also reduced the cloud fraction correlation with observations. Since then, a low cloud fraction in Polar WRF was diagnosed by both Bromwich et al. (2013b) and Huang et al. (2014). This problem was also noted in the Met-Office Unified Model by Bodas-Salcedo et al. (2012) where the issue was assessed to be the result of a low fraction of stratocumulus on the cold air side of cyclones. Bodas-Salcedo et al. (2012) developed a new parameterization that treats shear-dominated boundary layers differently which, to a large extent, removed the bias. Alternatively, the low cloud fraction could be a result of the incorrect initial conditions. In Section 6.3 we see that the bias only appears at lead times later than 36 hours; approximately the model initialization period. Thus, the boundary layer parameterizations may be correct but since they are initialized incorrectly, they act to produce an incorrect result later in the forecast. These theories should be tested in future research through WRF single column experiments.

While the errors have been emphasized in an attempt to note areas for improvement, it is also important to note what was well forecast in AMPS. Above the surface and first few model levels, there are no large biases in temperature and wind speed, and overall profiles of these variables are accurate. In particular, errors over the sea ice area are minimal, especially in forecasts of wind speed, and wind direction (not shown). When forecasting inversions, a key feature of the Antarctic low-levels, in most cases AMPS had either the same or a smaller bias than GFS.

Further, despite small sample sizes there were no notable biases in the open water and partial sea ice subsets (not shown).

Chapter 7

A Comparison between AMPS, GFS and ERA-Interim

In this section we compare Concordiasi profiles to AMPS, GFS and ERA-Interim profiles with the aim to evaluate AMPS with respect to a larger spectrum of analysis estimates. Similar to Section 6 we separate the biases by the surface type (HP, LE, and IC). Here we present statistics for the 6 hour and 72 hour lead times. The aim of the 6 hour lead time is to provide a comparison at the earliest forecast lead time where all 3 models can be compared. The aim of the 72 hour lead time is to provide a comparison at a later lead time when the initial condition impact is further removed.

7.1 Continental High Plateau (HP)

Figure 7.1 shows statistics for temperature, dew point and wind speed in the HP subset. The area contained within each plume represents the area encompassed by the 95 % confidence interval calculated from 1000 bootstrapped samples of each statistic for the 6 hour forecast of each model.

At the surface, the skill of temperature forecasts is best in AMPS close to the surface and skill is only significantly better than ERA (Figure 7.1j). Higher skill in

AMPS at the surface corresponds with smaller magnitude mean bias and standard deviation bias (Figures 7.1a and 7.1d respectively). It is notable that the mean surface temperature biases in GFS and ERA are larger. Such warm mean surface biases are consistent with the inversion biases, where there are significantly weaker inversions in GFS and ERA than in AMPS (Figures 6.4a and 6.4b). GFS has significantly worse skill than AMPS and ERA between the 7th and 10th model levels, above the SBI. These areas of low skill correspond with cold biases that are significantly larger in GFS than AMPS or ERA. Given that there are larger biases at the bottom and top of the mean inversion this also suggests that GFS produces weaker inversions than AMPS. This is consistent with the temperature inversion strength biases in Figures 6.4a and 6.4b where there are significantly stronger biases in GFS at the 6 hour lead time during both day and night. Alternatively it could indicate a deeper inversion, a result also confirmed in Figures 6.5a and 6.5b. ERA temperature inversion depth biases are not significantly different from AMPS which possibly accounts for the lack of significantly different ERA biases atop the inversion (Figures 6.5a and 6.5b). Above the 10th model level there are no significant differences in any statistics. Additionally there are no significant differences in CRMSDs between any of the models.

There are no significant differences between any of the models for MSESS in 72 hour forecasts of dew point (Figure 7.1k). Throughout the column AMPS has significantly smaller magnitude mean biases than both GFS and ERA above the 13th model level. At the surface AMPS also has significantly smaller biases than GFS. At 6 hours ERA and AMPS produce significantly stronger mean inversion strength than GFS (e.g. their inversion strength biases are closer to that observed; Figure 6.4). AMPS also has significantly smaller inversion depth biases than GFS (Figure 6.4). Stronger inversions in AMPS and ERA are consistent with the dew point biases where there are larger surface biases in GFS. Standard deviation

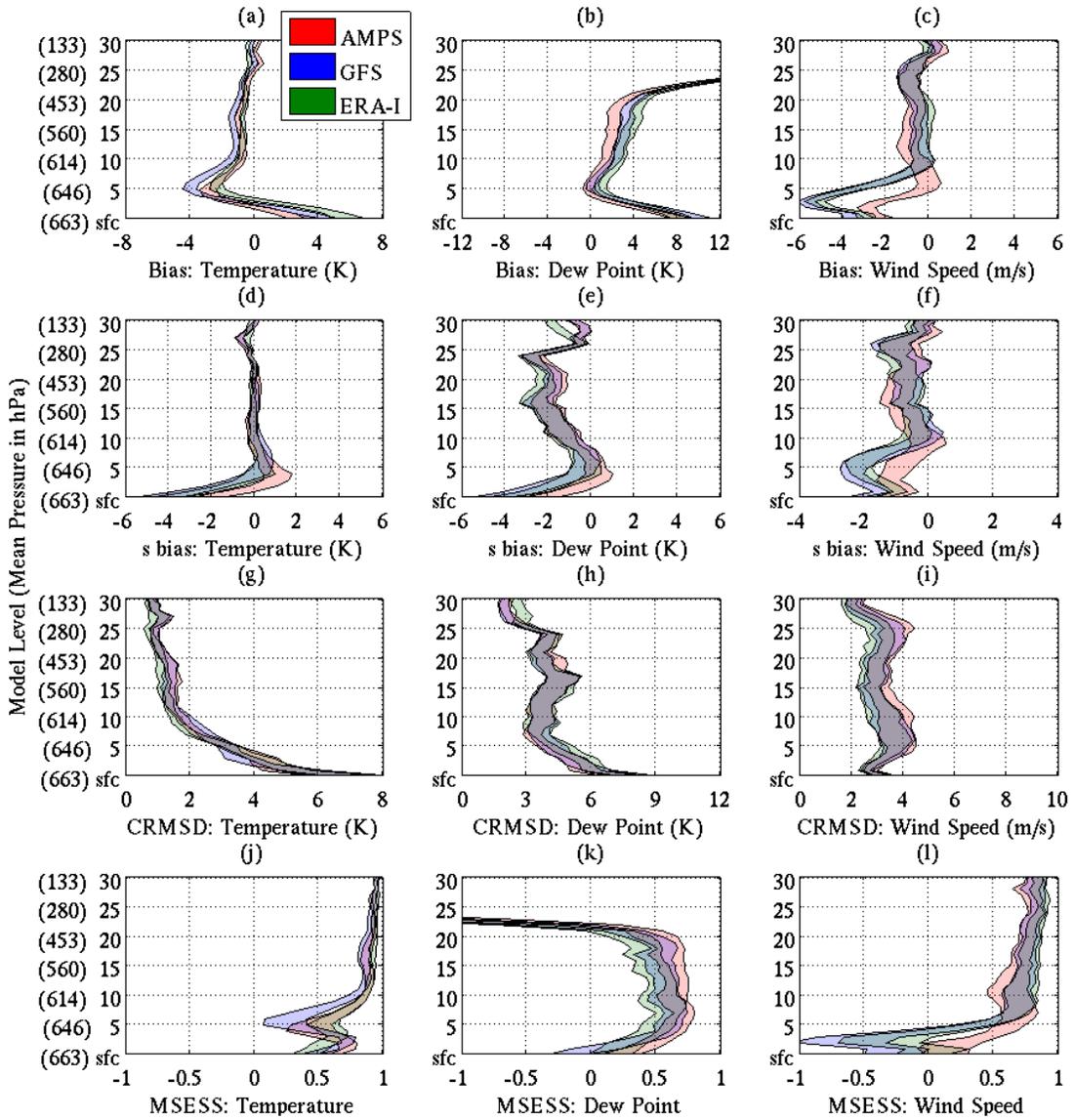


Figure 7.1: The 95% confidence intervals calculated from 1000 bootstrapped samples for (a,b,c) mean bias, (d,e,f) standard deviation bias, (g,h,i) CRMSD, (j,k,l) Correlation, and (m,n,o) MESS and (a,d,g,j,m) temperature, (b,e,h,k,n) dew point, and (c,f,i,l,o) wind speed. Statistics represent the HP subset for AMPS (red), GFS (blue), ERA-interim (green).

biases are not significantly different at any level although AMPS does have smaller standard deviation biases on average (Figures 7.1b and 7.1e respectively). There is little difference in errors between any of the models for forecasts of dew point (Figure 7.1h).

Skill for forecasts of wind speed is similar between all models above the 5th model level (Figure 6.11). However, between the 2nd and 5th model levels AMPS MSESS is significantly larger for wind speed than both GFS and ERA. This coincides with significantly smaller biases in AMPS between the 2nd and 7th model levels relative to GFS and ERA (Figure 7.1c). Further, there are also smaller standard deviation biases in AMPS around the 5th model level (Figure 7.1c). These biases are co-located with the average position of the LLWM (or katabatic flow) discussed in Section 3.4 and throughout Chapter 6. Further, LLWM maximum wind speeds have significantly smaller biases in AMPS than in GFS or ERA at 6 hours during day (Figure 6.7). The height of the LLWM at 6 hours, during both day and night, is also significantly lower (and closer to that observed) in AMPS relative to GFS or ERA (Figure 6.8). Further, at 6 hours during both day and night AMPS produces more accurate relative occurrence of the LLWMs than either GFS or ERA by 0.2-0.4. These results indicate that improved skill in AMPS in the low-levels (relative to GFS or ERA) is due to better representation of the LLWMs. This occurs through faster wind speeds and greater variability in the LLWM, more frequent occurrence of the LLWM, and better representation of the LLWM height for forecasts in AMPS (relative to GFS and ERA). There is little difference between 6 hour forecast errors in any of the models (Figure 7.1i).

Figure 7.2 shows statistics for AMPS and GFS forecasts at 72 hours lead time. MSESS for 72 hour forecasts of temperature between AMPS and GFS is similar above the 10th model level (Figure 7.2j). However, below that level there are significant differences between AMPS and GFS. MSESS in GFS has a local minima

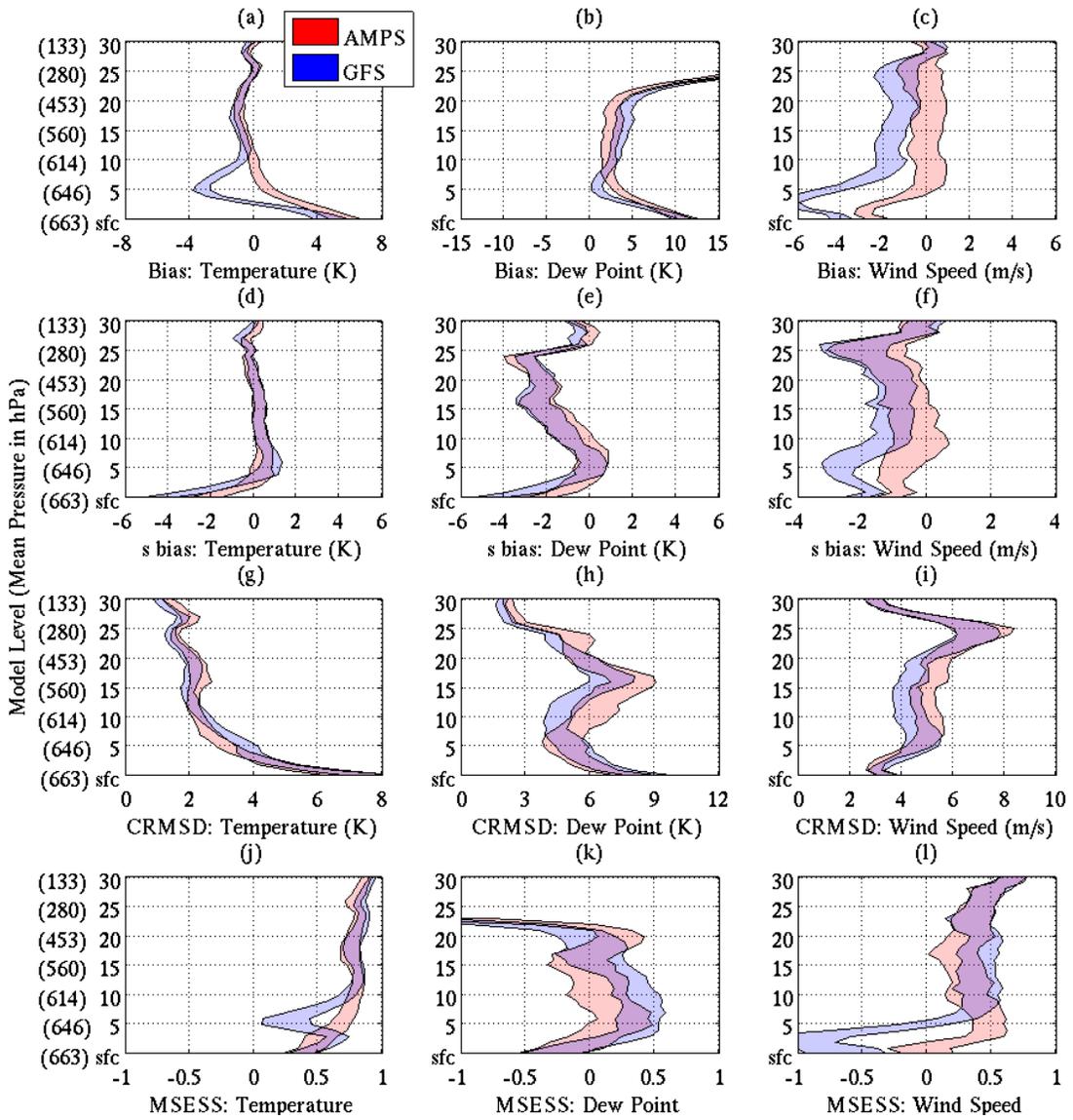


Figure 7.2: As in Figure 7.1 but for forecasts at 72 hours lead time.

at the 6th model level while there is no minima in AMPS. This leads to better skill in AMPS between the 5th and 8th model levels. The local minima in GFS forecast skill coincides with a large GFS cold bias that does not occur in AMPS (Figure 7.2a). Below the 4th model level GFS biases tend toward the same biases as AMPS such that at the surface there is no significant difference in 72 hour temperature biases between the two models. Weaker GFS mean biases below the 4th model level are simply an artifact of the incorrect forecast of inversions. In Figures 6.4a and 6.4b, and 6.5a and 6.5b, we showed that GFS inversions have weak and deep inversions at 72 hours throughout the HP subset. Thus temperature lapse rates in the inversions are much weaker than in AMPS. This leads to the strong cold biases atop the SBI and warm biases at the surface in GFS at 72 hours lead time. Thus any weaker temperature biases in GFS at these levels are simply an artifact of a weak temperature inversion. In addition there are also significantly weaker standard deviation biases at the surface in AMPS indicating AMPS forecasts are consistently less biased (Figure 7.2d). There is little difference between errors in the two models at 72 hours (Figure 7.2g).

There are no significantly different MSESS values between AMPS and GFS for 72 hour forecasts of dew point, and skill throughout both models is poor to average (Figure 7.2k). There are some significantly weaker mean biases in GFS between the 4th and 6th model levels. However, this produces a larger spread of mean biases between the surface and the top of the average inversion indicating weaker inversions in GFS. There are weaker mean inversion biases during night although there are none during day at the 72 hour lead time (Figures 6.4c and 6.4d). Further there are also significantly deep mean inversion biases in GFS at 72 hours during both day and night (Figures 6.5c and 6.5d). Thus significantly weaker and deeper inversions in GFS result in the larger spread of biases in GFS. Further, there are stronger (although not significant) negative standard deviation

biases below the 3rd model level in GFS relative to AMPS indicating low bias is consistent in AMPS (Figure 7.2e). There are no significant differences between the two models in errors although there are on average larger errors in AMPS between the 8th and 16th model levels (Figure 7.2h).

Wind speed MSESS at the 72 hour lead time is significantly different between AMPS and GFS below the 6th model level (Figure 7.2l). Here skill is much worse than climatology in GFS and is average to poor in AMPS. Above there are no significant differences in MSESS. Distinctly higher skill in AMPS between the 2nd and 5th model levels coincides with significantly weaker slow mean biases, and weaker negative standard deviation biases, at levels co-located with the LLWM maximum (Figures 7.2c and 7.2f). Further at the 72 hour lead time GFS has significantly larger LLWM maximum wind speed and height biases. GFS produces LLWMs that are both weaker and at a higher height than the LLWMs in AMPS and those observed. In addition GFS forecasts far fewer inversions than AMPS and than observed. The relative frequency of occurrence for LLWMs in GFS is approximately 0.3 lower than observed and 0.2 lower than AMPS (Figure 6.9). Thus at the 72 hour lead time skill is better in AMPS due to lower biases and improved variability relative to GFS. Further, representation of the LLWM is much better in AMPS. There are few differences in errors between the two models although there are on average higher errors in AMPS between the 10th and 20th model levels (Figure 7.2i).

7.2 Continental Low Elevations (LE)

There are few differences between any of the models for 6 hour forecasts of temperature over the continental low elevations (Figures 7.3a, 7.3d, 7.3g, and 7.3j). The only notable differences are in the mean biases with GFS producing the weakest mean surface biases, significantly better than ERA (but not AMPS). The biases

are consistent with biases of inversion strength where there are no significant differences between forecasts for models at early lead times although there are notably larger biases in ERA than GFS (Figures 6.4e and 6.4f). There are also no significant differences in the biases of inversion depth at early lead times although ERA has notably weaker biases than GFS (Figures 6.5e and 6.5f).

Skill for 6 hour forecasts of dew point is not significantly different anywhere in the column (Figure 7.3k). In the mean biases there are also no significant differences between the models although ERA surface mean biases are on average smaller than AMPS and GFS (Figure 7.3b). There are no significant differences between any of the models for early lead time forecasts of inversion strength or depth (Figures 6.4g and 6.4h, and Figures 6.5g and 6.5h respectively). Further there are no significant differences between the models for standard deviation biases and CRMSDs (Figures 7.3f and 7.3i).

Wind speed MSESS varies little between the models (Figure 7.3l). However there are significant differences in the biases between the 2nd and 5th model levels (Figure 7.3c). Here there are significantly weaker biases in AMPS than in GFS or ERA. Similar to the HP subset this difference in mean bias occurs at levels collocated with the maxima in the LLWM. In addition, there are also weaker standard deviation biases on average at these levels although this difference is not significant (Figure 7.3f). AMPS has significantly lower LLWM maximum wind speed biases than ERA and although AMPS biases relative to GFS are not significantly different, they are smaller by around 2 ms^{-1} . Further there are no LLWM height biases in AMPS during day or night while there are biases of approximately 100-200 m in GFS and ERA. The relative frequency of occurrence during night is the same in all models, approximately 0.1 less than the frequency observed. During day however AMPS has better relative occurrence of LLWMs than both AMPS and

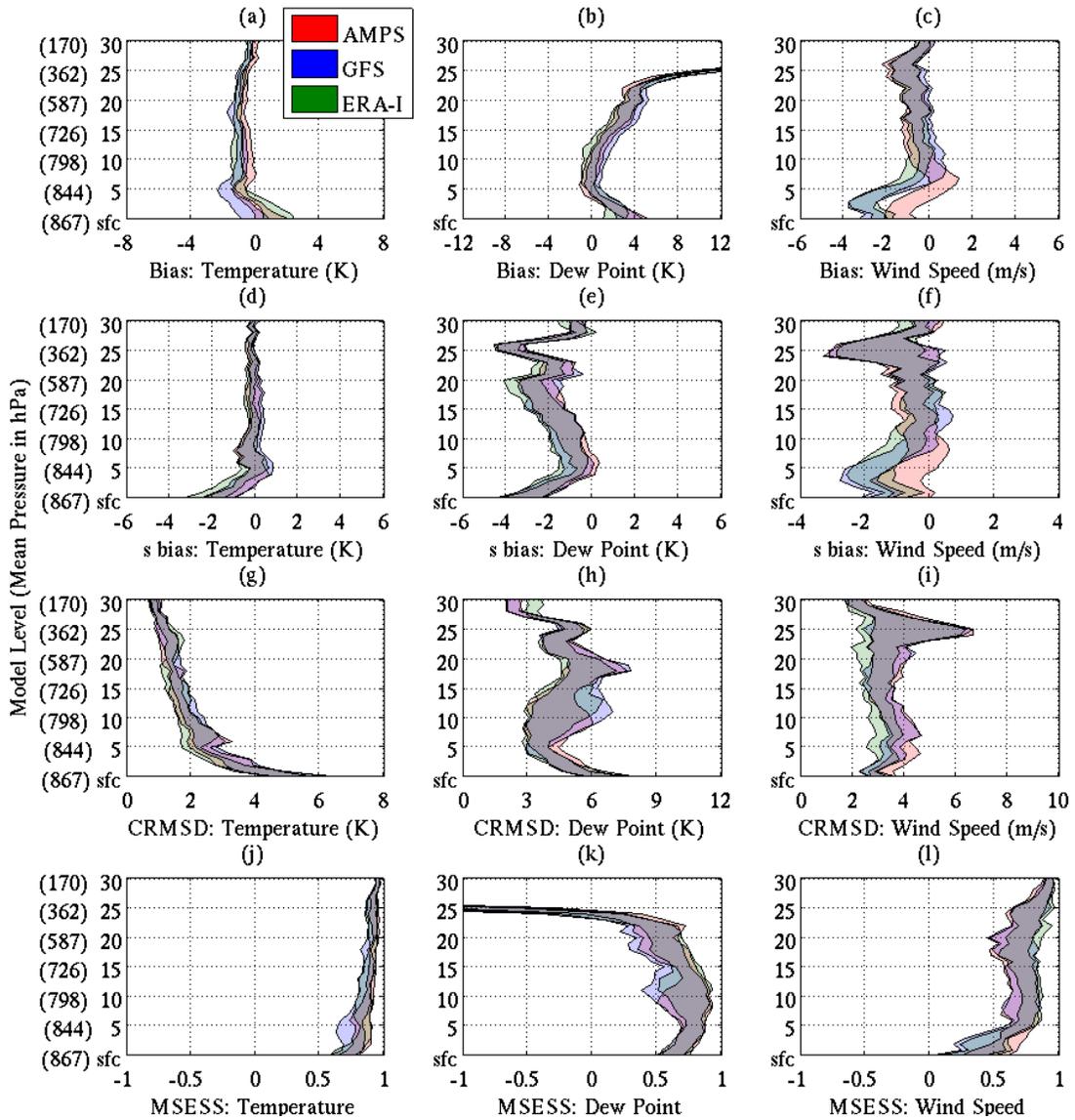


Figure 7.3: As in Figure 7.1 but for the LE subset.

ERA. Again there is evidence to suggest that the AMPS 6 hour forecasts of the LLWM are better than in GFS or ERA.

There are no significant differences between AMPS and GFS in MSESS for 72 hour forecasts of temperature in the LE subset (Figure 7.4j). The only significant differences between AMPS and GFS in statistics for 72 hour forecasts of temperature are in the mean biases where there are cold biases throughout the column in GFS (except at the surface) and warm biases below the 12th model level in AMPS (Figure 7.4a). In the LE subset at later lead times (after 48 hours), temperature inversion strength and depth biases are smaller in magnitude in the LE subset relative to the HP subset (Figures 6.4e and 6.4f, and 6.5g and 6.5h, respectively). This accounts for the improvement of mean temperature biases in GFS, with AMPS and GFS biases similar above the surface while GFS has no biases at the surface. There are better standard deviation biases on average in GFS below the 10th model level indicating more accurate variability in AMPS (Figure 7.4d). There are no major differences between AMPS and GFS 72 hour forecasts with regard to CRMSEs (Figure 7.4g).

MSESS for 72 hour forecasts of dew point are not significantly different (Figure 7.4k). The only significant difference in 72 hour forecasts for dew point statistics are in the mean biases where AMPS has significantly larger warm biases below the 5th model level (Figure 7.4b). At 72 hours there are no significant differences between the models for forecasts of inversion strength although at later lead times there are significant differences at night with greater biases in AMPS (Figure 6.4h). At later lead times, GFS depth biases are not significant while there are shallow AMPS inversion depth biases (Figures 6.5g and 6.5h). There are no significant differences in either standard deviation biases or in CRMSEs, although AMPS has slightly smaller standard deviation biases on average between the 5th and 25th model levels (Figures 7.4e and 7.4h).

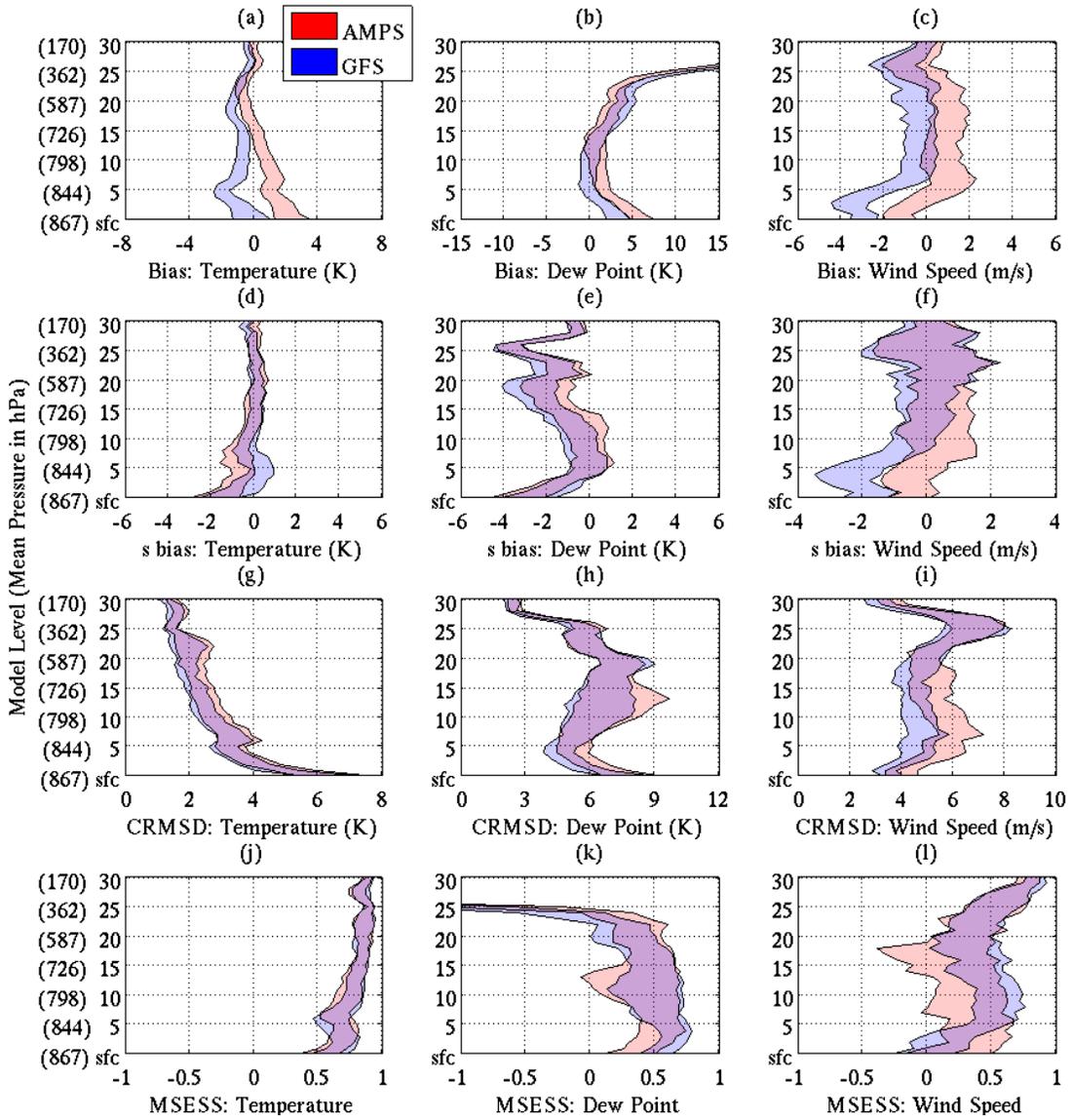


Figure 7.4: As in Figure 7.1 but for the LE subset and forecasts at 72 hours lead time.

While there are no significant differences in MSESS for 72 hour forecasts of wind speed, this is likely a result of the large amount of uncertainty in MSESS with both models producing average to poor skill (Figure 7.4l). However, below the 6th model level, there are significant differences in mean biases and standard deviation biases (Figures 7.4c and 7.4f respectively). Slow mean wind speed biases and negative standard deviation biases below the 6th model level are approximately doubled in GFS relative to AMPS. Similar to the HP subset, these biases are at the same level as the LLWM indicating a much weaker LLWM in GFS than AMPS. Although there are no significant differences in most LLWM biases, there are significantly smaller LLWM maximum wind speed biases in AMPS relative to GFS during the day. Further, although both models have forecast LLWMs at a relative occurrence 0.1-0.3 less than observed AMPS, has slightly higher relative occurrence than GFS through most of the later lead times. There are no significant differences in CRMSDs between the 2 models but AMPS has on average larger errors below the 20th model level (Figure 7.4i).

7.3 Permanent Ice Shelves and Sea Ice (IC)

On average below the 6th model level MSESS in 6 hour forecasts of AMPS temperature in the IC subset are worse than ERA, but better than GFS, although not significant (Figure 7.5j). Improved skill in ERA at these levels also coincides with significantly smaller ERA mean biases and CRMSDs (Figures 7.5a and 7.5g respectively) and weaker standard deviation biases on average although not significant (Figure 7.5d). In addition ERA has a different structure of biases to AMPS and GFS. There are small but significant cold biases at the 13th model level where AMPS has warm and dry biases, and no bias in the lowest 6 model levels where AMPS has significant cold biases. Cold surface biases in AMPS and

GFS are consistent with the inversion occurrence biases showing temperature inversions are forecast too frequently in both models. In comparison, ERA forecasts approximately the correct number of inversions, which is consistent with the mean temperature biases since it has smaller biases than AMPS and GFS.

MSESS for 6 hour forecasts of dew point are not significantly different between the three models although ERA has on average larger skill than AMPS and GFS (Figure 7.5g). There are no mean biases in ERA below the 13th model level while there are mean biases in AMPS and GFS at these levels. Above the 8th model level AMPS and GFS biases are positive while below the 5th model level they are negative. Mean biases are zero for all models between these levels simply because the AMPS and GFS biases are varying from negative in the upper levels to positive in the lower levels. There are no significant differences between inversion strength and depth biases (not shown). However there are differences in inversion occurrence. At 6 hours all models over-forecast the number of inversions, consistent with the small negative dew point biases at the surface. GFS has the worst results at night but the best during day while ERA and AMPS have similar results. At the 0 hour lead time AMPS over-forecasts by 0.4 although this is possibly a bi-product of the small sample size at early lead times. Below the 5th model level, of the 3 models, ERA has negative standard deviation biases while AMPS and GFS have no significant standard deviation biases although there are no significant differences between the models (Figure 7.5e). Errors are smallest in ERA and this difference is significant below the 5th model level. Thus increased skill in ERA for dew points over ice is a product of lower mean biases and lower errors although these results are not significant.

There are no significant differences between any of the models with respect to wind speed MSESS, however ERA has notably greater MSESS values than either AMPS or GFS above the 4th model level (Figure 7.5l). This higher skill coincides

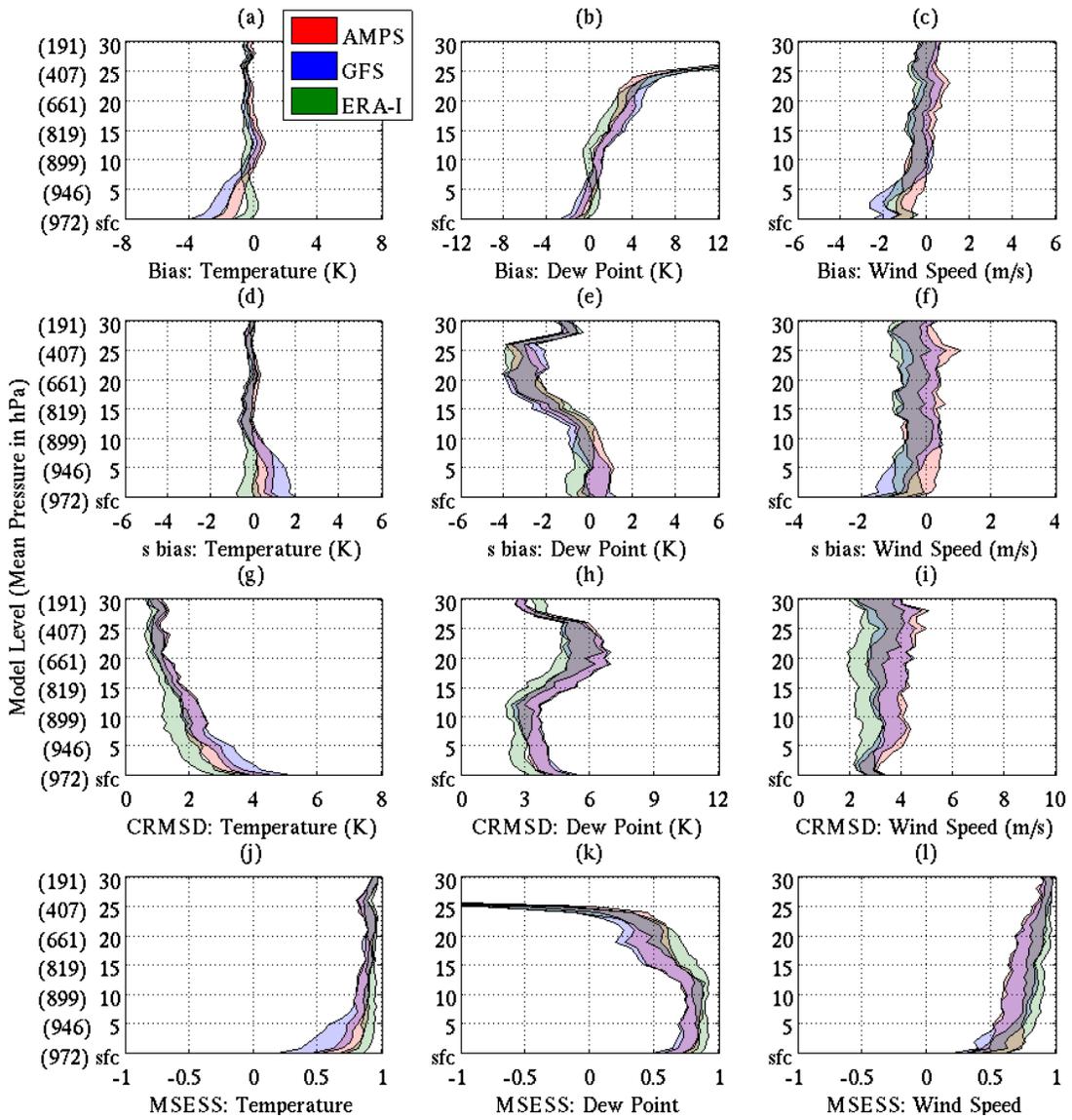


Figure 7.5: As in Figure 7.1 but for the IC subset.

with notably lower errors in ERA than either AMPS or GFS although none of these differences are significant (Figure 7.5i). Throughout most of the column there are no significant mean biases in wind speed. However, between the 2nd and 5th model levels there are significantly lower biases in AMPS when compared to GFS and there are notably smaller biases in AMPS when compared to ERA. Again, this corresponds to the level of the LLWM suggesting AMPS is more reliable at forecasting the higher wind speeds. There are no significant biases in any of the models for LLWM maximum wind speed or height. However all models have relative occurrence that is 0.1 - 0.2 too low for LLWMs at 6 hours. There are no significant differences in regard to wind speed standard deviation biases.

MSESS for 72 hour forecasts of temperature are significantly different between AMPS and GFS only below the 3rd model level (Figure 7.6j). At the surface AMPS skill is average (MSESS of 0.4 - 0.65) while GFS skill is poor (MSESS of -0.3 - 0.4). Significant differences in skill at the low-levels coincide with significant differences in mean biases (Figure 7.6a; AMPS mean biases are -1 - 0.3 while GFS mean biases are -5 - -3) and significant differences in standard deviation biases (Figure 7.6d; AMPS mean standard deviation biases are -0.8 - 0.3 while GFS mean standard deviation biases are 1 - 2.4). These statistics indicates that low-level forecasts of AMPS temperature improve upon GFS forecasts through much smaller bias and better variability. Further, AMPS forecasts of temperature inversion occurrence are more accurate than GFS forecasts (Figures 6.6a and 6.6b). The frequency of inversion occurrence in GFS is double that of the observed at night and 0.2 more than the observed during day in contrast to AMPS with accurate forecasts during the day and only a 0.2 bias during night. This is consistent with the mean biases where GFS has a strong cold bias at the surface implying the likelihood of more inversions. There are no significant difference in CRMSDs between the 2 models (Figure 7.6g).

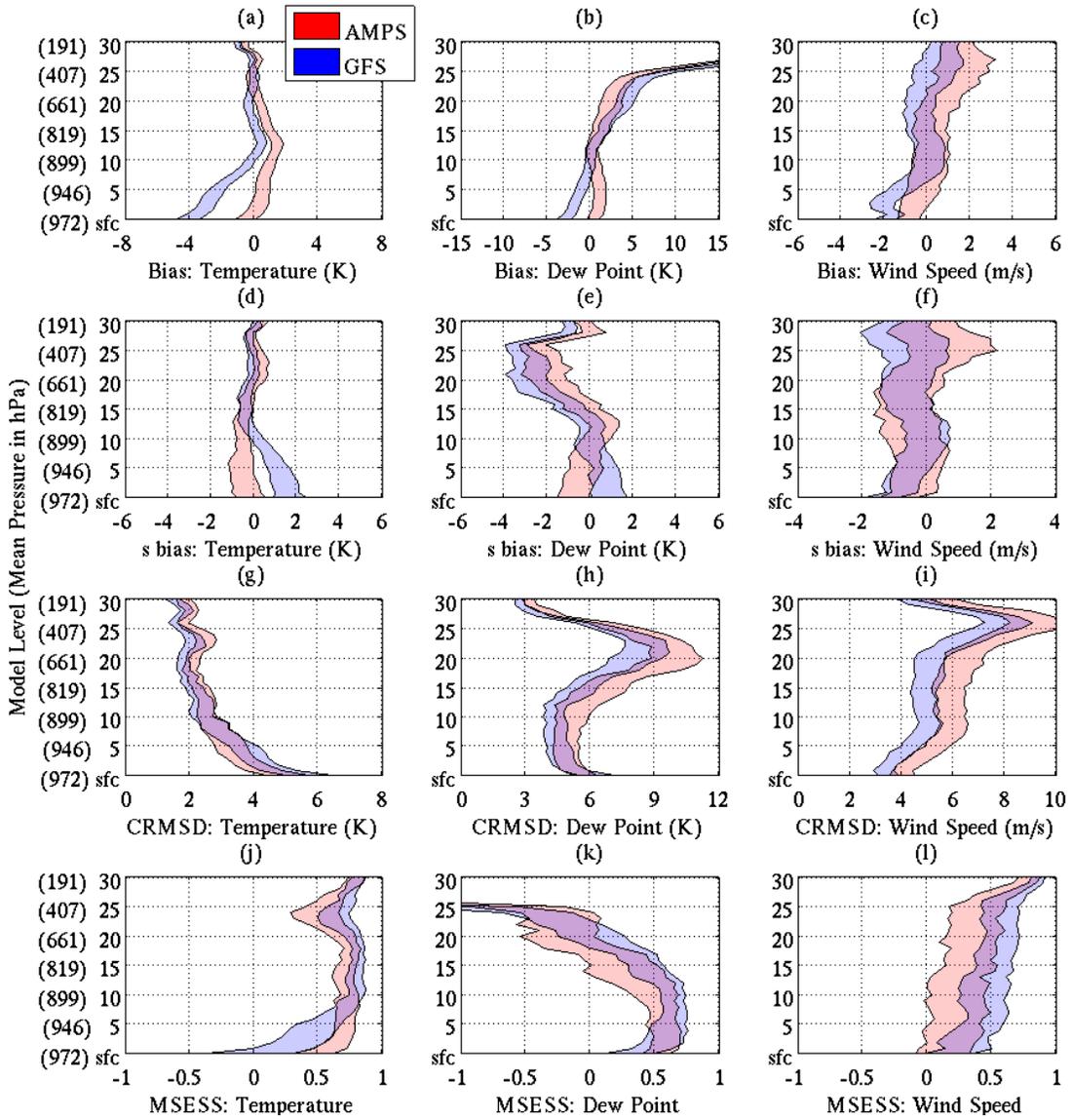


Figure 7.6: As in Figure 7.1 but for the IC subset and forecasts at 72 hours lead time.

MSESS for 72 hour forecasts of AMPS and GFS dew point are not significantly different although GFS has on average higher skill above the 3rd model level and AMPS has on average higher skill below the 3rd model level (Figure 7.6k). There are significant differences between the mean biases for AMPS and GFS (Figure 7.6). Below the 7th model level mean biases in GFS are dry while in AMPS they are moist but smaller in magnitude. Above the 15th model level, mean biases in GFS become moist and larger than AMPS moist biases (although not significant). The cold low-level mean biases in GFS are consistent with inversion occurrence statistics at later lead times (24 hours or later) where GFS has 0.1 - 0.3 relative occurrence more than the observed and AMPS relative occurrence errors are minimal (Figure 6.6c and 6.6d). Standard deviation biases at the surface significantly different; they are negative in AMPS and positive in GFS. Above this, there are no significant differences (Figure 7.6e). Although the differences are not significant CRMSDs are smaller in GFS possibly accounting for the increased skill in GFS above the 5th model level (Figure 7.6h).

There are no significant differences in MSESS between AMPS and GFS for 72 hour forecasts of wind speed (Figure 7.6l). However, it is notable that GFS has relatively higher skill than AMPS. This higher skill in GFS coincides with notably lower errors in GFS relative to AMPS, although again this difference is not significant (Figure 7.6i). The only significant difference between the two models occurs in the mean biases between the 2nd and 4th model levels (Figure 7.6c). Here the low-level slow biases in both models are weaker in AMPS than GFS. Similar to the HP and LE subsets these biases are co-located with the LLWM. Similar to the early lead times there are no significant biases in forecasts of LLWM maximum wind speed or height but there are errors in the relative occurrence of LLWMs. Throughout the later lead times both AMPS and GFS have low relative occurrence, although AMPS relative occurrence is not as low as GFS. At higher

model levels there are no significantly different biases. In addition there are no significantly different biases in standard deviation (Figure 7.6f).

7.4 Discussion

In this chapter we have discussed the differences between 6 and 72 hour forecasts of AMPS, GFS and ERA-Interim. There are a number of instances where AMPS either outperforms or is outperformed by one or both of GFS or ERA. The features that were captured with the best skill in AMPS included some of the key Antarctic low-level features: inversions, especially over the continent, and the LLWM. ERA however, outperformed AMPS and GFS in some instances, specifically over the sea ice.

In the HP subset, at early lead times surface biases of temperature and dew point are positive in all models. Overall AMPS has the smallest surface temperature and dew point biases, with the most accurate variability. In addition there are weak and deep inversion biases in all models at early lead times. AMPS however shows less biased temperature inversion strength than GFS or ERA. In the LE subset results are similar but more varied. Therefore in all models we see that there are similar problems with early lead time forecasts of the low-level temperature and dew point structure. Atlaskin and Vihma (2012) compared the wintertime (snow-covered surface) nocturnal boundary layer structure in Finland for a number of models including GFS and the ECMWF IFS (on which ERA-Interim is designed). They showed that when a strong temperature inversion was observed, the models underestimated it, consistent with the results here and those shown by AMPS. Other results from this study showed that the warm bias in 2 m temperature forecast during periods of observed temperature inversion partly resulted from a warm bias in the initial conditions. The warm bias was due to problems in data assimilation in IFS and GFS. In addition Zhang et al. (2011) showed that ERA

underestimates inversions in Spring, also similar to the results here. In particular, the IFS data assimilation increased the 2 m temperature bias. Fréville et al. (2014) show that the warm biases in ERA over Antarctica are due primarily to an overestimation of surface turbulent fluxes in strongly stratified conditions. Numerical experiments with Crocus, a snowpack model, forced by ERA-Interim showed that a small change in the parameterization of the effects of stability on the surface exchange coefficients can significantly impact the snow surface temperature. Thus they conclude that the ERA-Interim warm bias appears to be likely due to an overestimation of the surface exchange coefficients under very stable conditions. Gallée and Gorodetskaya (2010) ran simulations at Dome C, Antarctica using the MAR (Modèle Atmosphérique Régional). This model also showed weak surface based inversions and warm surface temperatures. Thus the warm biases in AMPS, GFS, and ERA are likely an issue with both surface layer parameterization and data assimilation, an issue that is persistent through the models in this study and others.

At later lead times AMPS and GFS have very different low level biases, with shallow and weak AMPS inversions, and deep and weak GFS inversions. These result in inaccurate temperature and dew point forecasts for both models although overall AMPS has smaller biases atop the inversions as a result of the deep GFS inversions. Atlaskin and Vihma (2012) showed over snow in Finland that 2 m temperature biases in the integrated forecast system (IFS) and GFS increase through lead time, strengthening temperature inversions. This is inconsistent with the results shown for GFS here where inversion biases are constant throughout lead time.

Despite slow maximum wind speed biases, high LLWM maxima in some subsets, low frequency of occurrence, and low variability resulting in low skill for the LLWM in AMPS, a majority of these statistics are worse in GFS and ERA, especially over

the high plateau. This results in better forecast performance in AMPS relative to both ERA and GFS at levels associated with the LLWM. As discussed in Sections 4.1 and 6.6 the LLWMs over Antarctica are often a response to the katabatic forcing produced by the slope of the surface, cooling of the surface layer air, and surface drag. Given that surface biases are larger in GFS in all subsets the surface drag may be higher in GFS resulting in a weaker LLWM. Further, over the high plateau inversions are weaker and deeper in GFS implying weaker stratification. This will lead to weaker LLWMs over all subsets. Thus in GFS the weaker and less skillful forecasts of LLWMs are likely a result of these factors. However, in ERA there are no consistent differences relative to AMPS across all subsets in regard to stratification and surface drag and we can therefore not attribute this to either of these. However, ERA is run at a lower resolution (T255; approximately 80km) than AMPS which may result in more gradual slopes. Thus the most likely cause of weaker LLWMs in ERA is the lower grid spacing.

ERA does not exhibit the warm biases associated with the hypothesized low cloud fraction error in the IC subset. Further dew point biases at these levels are lower indicating improved representation of water vapor over AMPS or GFS. Since such biases have been shown by Bodas-Salcedo et al. (2012) (in the Met Office Unified Model) to be a product of low vertical mixing of stratified shear dominated boundary layers, parameterization of such boundary layers in ERA is likely better than AMPS and GFS. There are also smaller temperature and dew point biases and errors, and lower variability at the surface suggesting improved parameterization of the boundary layer over sea ice in ERA. Further, ERA also uses a 4DVAR data assimilation system with flow-dependent background error covariance that likely spreads information more efficiently than the 3DVAR system utilized by AMPS. Whether the improvement in ERA is due to improved data assimilation

or improved parameterization of the boundary layer should be evaluated through further research.

Overall we show here that the performance of AMPS, GFS, and ERA are comparable except for in the LLWM where AMPS outperforms both models; in inversions over the continent where AMPS captures more accurate inversion tops than GFS; and in the boundary layer over the sea ice where ERA outperforms both AMPS and GFS.

Chapter 8

Conclusions

This study utilizes observations obtained from dropsondes during the Concordiasi field project in order to produce an evaluation of the biases within AMPS forecasts that has vastly improved spacial coverage of observations relative to previous studies. Profiles were grouped into a number of categories dependent on geographic characteristics that include the continental high plateau, the continental low elevations, and ice covered water areas. This methodology allowed the diagnosis of a number of areas where AMPS exhibits biases and other issues with its forecasts including:

- The over-representation of inversions at early lead times including the depth over the continent and the relative occurrence over ice cover. This problem has been linked to the GFS initial conditions and the weakness of the data assimilation system to spread the information provided by sparse observations.
- Multiple issues with parameterizations of the boundary layer including shallow inversions over the continent, warm, moist and slow surface biases and weak katabatic winds. Qualitatively these have been attributed to under-mixing, weak surface fluxes, and high surface drag.

- A warm and dry bias between 750- and 900-hPa likely associated with the previously established problem of an under-representation of cloud in NWP models over Antarctica.

In addition to diagnosing problems with AMPS forecasts, we produced a comparison between Concordiasi, AMPS, GFS, and ERA-Interim in order to further the knowledge of AMPS performance relative to other models. We show that despite issues with the representation of LLWMs associated with the katabatic flow, AMPS provides the best representation. Further, there are many differences between the models in the boundary layer, specifically with inversions and AMPS forecasts of inversions are biased less than GFS. ERA outperforms both AMPS and GFS over the sea ice.

While this study focuses on the benefits of Concordiasi in comparing to- and evaluating AMPS analyses and forecasts, we would also like to emphasize the benefits that Concordiasi can provide in understanding the Antarctic atmosphere. In Section 4.1 we highlighted the key features observed from Concordiasi in each of the surface defined subsets we analyzed. In particular, we noted the strong katabatic flow and the strong surface based inversions atop the high plateau and over the sea ice, none of which have been studied before with such a broad spacial coverage. Thus Concordiasi provides an excellent medium through which to analyze features of the Antarctic atmosphere, that have in the past, only been examined at specific locations with observations that have mainly been confined to the lowest levels of the atmosphere.

Whilst Concordiasi dropsondes provided a vast spacial coverage of observations, a lack of temporal resolution in the observations poses a large limitation for this thesis. We present results in this study only for the September to December period, and only during one year. Bromwich et al. (2013b) evaluate Polar WRF with observations throughout a single year of simulations. They show that statistics

vary between summer and winter. Thus, the results presented in this thesis are only representative of the Spring period. To provide a broad spacial and temporal coverage more observations of this type are required throughout the year. Due to time and monetary limitations, and also the design of the experiment (the stratospheric jet had to be consistent enough for driftsondes not to leave the Antarctic region and this only occurs during Winter and Spring), it was not possible to get good temporal resolution of observations from Concordiasi. Therefore further research should focus on improving the design of such experiments and obtaining more observations similar to Concordiasi throughout the year.

Finally, further research focusing on the performance of AMPS should work to correct the discussed biases through improved treatment of the surface and boundary layer within parameterization schemes. One way to test whether surface biases over the continent are a result of the incorrect parameterization of mixing in the boundary layer would be to perform WRF single column experiments with Concordiasi profiles used to initialize the simulations. Also, development of new data assimilation schemes that spread the information provided by the sparse observational network more efficiently should be a key component to any new developments in AMPS. We note that despite the focus on key biases here, there are few major biases above the surface, and over the ice or water. Also, in most subsets AMPS biases were similar to- or smaller than- those in GFS for our analysis of inversions.

Bibliography

- Agosta, C., V. Favier, G. Krinner, H. Gallee, X. Fettweis, and C. Genthon, 2013: High-resolution modelling of the Antarctic surface mass balance, application for the twentieth, twenty first and twenty second centuries. *Clim. Dynam.*, **41** (11-12), 3247–3260.
- Alexander, S., A. Klekociuk, A. McDonald, and M. Pitts, 2013: Quantifying the role of orographic gravity waves on polar stratospheric cloud occurrence in the Antarctic and the Arctic. *J. Geophys. Res.*, **118** (20), 493–507.
- Andreas, E. L., K. J. Claffy, and A. P. Makshtas, 2000: Low-level atmospheric jets and inversions over the western weddell sea. *Boundary-layer Meteor.*, **97** (3), 459–486.
- Arblaster, J. M. and G. A. Meehl, 2006: Contributions of external forcings to southern annular mode trends. *J. Climate*, **19** (12), 2896–2905.
- Atlaskin, E. and T. Vihma, 2012: Evaluation of nwp results for wintertime nocturnal boundary-layer temperatures over europe and finland. *Quart. J. Roy. Meteor. Soc.*, **138** (667), 1440–1451.
- Barker, D., J. Bray, Y. Guo, X. Huang, Z. Liu, S. Rizvi, and Q. Xiao, 2006: Status report on wrf-arw’s variational data assimilation system (wrf-var). *WRF users’ workshop, Boulder, Colorado*, 19–22.
- Barker, D. M., 2005: Southern high-latitude ensemble data assimilation in the Antarctic mesoscale prediction system. *Mon. Wea. Rev.*, **133** (12), 3431–3449.
- Bengtsson, L., K. I. Hodges, and E. Roeckner, 2006: Storm tracks and climate change. *J. Climate*, **19** (15), 3518–3543.
- Bodas-Salcedo, A., K. Williams, P. Field, and A. Lock, 2012: The Surface Downwelling Solar Radiation Surplus over the Southern Ocean in the Met Office Model: The Role of Midlatitude Cyclone Clouds. *J. Climate*, **25** (21), 7467–7486.
- Bouchard, A., F. Rabier, V. Guidard, and F. Karbou, 2010: Enhancements of Satellite Data Assimilation over Antarctica. *Mon. Wea. Rev.*, **138** (6), 2149–2173.

- Bourassa, M. A., et al., 2013: High-Latitude Ocean and Sea Ice Surface Fluxes: Challenges for Climate Research. *Bull. Amer. Meteor. Soc.*, **94** (3), 403–423.
- Bouttier, F., 1994: A dynamical estimation of forecast error covariances in an assimilation system. *Mon. Wea. Rev.*, **122** (10), 2376–2390.
- Bromwich, D. H., R. L. Fogt, K. I. Hodges, and J. E. Walsh, 2007: A tropospheric assessment of the ERA-40, NCEP, and JRA-25 global reanalyses in the polar regions. *J. Geophys. Res.*, **112** (D10).
- Bromwich, D. H., K. M. Hines, and L.-S. Bai, 2009: Development and testing of polar weather research and forecasting model: 2. Arctic Ocean. *J. Geophys. Res.*, **114** (D8).
- Bromwich, D. H., A. J. Monaghan, K. W. Manning, and J. G. Powers, 2005: Real-time forecasting for the Antarctic: An evaluation of the Antarctic Mesoscale Prediction System (AMPS). *Mon. Wea. Rev.*, **133** (3), 579–603.
- Bromwich, D. H., J. P. Nicolas, A. J. Monaghan, M. A. Lazzara, L. M. Keller, G. A. Weidner, and A. B. Wilson, 2013a: Central West Antarctica among the most rapidly warming regions on Earth. *Nat. Geosci.*, **6** (2), 139–145.
- Bromwich, D. H., F. O. Otieno, K. M. Hines, K. W. Manning, and E. Shilo, 2013b: Comprehensive evaluation of polar weather research and forecasting model performance in the Antarctic. *J. Geophys. Res.*, **118** (2), 274–292.
- Bromwich, D. H., et al., 2012: Tropospheric clouds in Antarctica. *Rev. Geophys.*, **50** (1).
- Cohn, S. A., et al., 2013: Driftsondes: Providing In Situ Long-Duration Dropsonde Observations over Remote Regions. *Bull. Amer. Meteor. Soc.*, **94** (11), 1661–1674.
- Dee, D., et al., 2011: The era-interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137** (656), 553–597.
- Deque, M., C. Dreveton, A. Braun, and D. Cariolle, 1994: The ARPEGE/IFS atmosphere model: a contribution to the French community climate modelling. *Clim. Dynam.*, **10** (4-5), 249–266.
- Fernando, H. and J. Weil, 2010: Whither the stable boundary layer? *Bull. Amer. Meteor. Soc.*, **91** (11), 1475–1484.
- Fogt, R. L. and D. H. Bromwich, 2008: Atmospheric Moisture and Cloud Cover Characteristics Forecast by AMPS. *Wea. Forecasting*, **23** (5), 914–930.

- Fréville, H., E. Brun, G. Picard, N. Tatarinova, L. Arnaud, C. Lanconelli, C. Reijmer, and M. van den Broeke, 2014: Using modis land surface temperatures and the crocus snow model to understand the warm bias of era-interim reanalyses at the surface in antarctica. *Cryo. Disc.*, **8**, 55–84.
- Gallée, H. and I. V. Gorodetskaya, 2010: Validation of a limited area model over dome c, antarctic plateau, during winter. *Clim. Dynam.*, **34** (1), 61–72.
- Guichard, F., D. Parsons, and E. Miller, 2000: Thermodynamic and radiative impact of the correction of sounding humidity bias in the tropics. *J. Climate*, **13** (20), 3611–3624.
- Ha, K.-J., Y.-K. Hyun, H.-M. Oh, K.-E. Kim, and L. Mahrt, 2007: Evaluation of Boundary Layer Similarity Theory for Stable Conditions in CASES-99. *Mon. Wea. Rev.*, **135** (10), 3474–3483.
- Hines, K. M. and D. H. Bromwich, 2008: Development and testing of polar weather research and forecasting (WRF) model. Part I: Greenland ice sheet meteorology. *Mon. Wea. Rev.*, **136** (6), 1971–1989.
- Hines, K. M., D. H. Bromwich, L.-S. Bai, M. Barlage, and A. G. Slater, 2011: Development and Testing of Polar WRF. Part III: Arctic Land. *J. Climate*, **24** (1), 26–48.
- Holtlag, A., et al., 2013: Stable Atmospheric Boundary Layers and Diurnal Cycles: Challenges for Weather and Climate Models. *Bull. Amer. Meteor. Soc.*, **94** (11), 1691–1706.
- Huang, Y., S. T. Siems, M. J. Manton, and G. Thompson, 2014: An evaluation of wrf simulations of clouds over the southern ocean with a-train observations. *Mon. Wea. Rev.*, **142** (2), 647–667.
- Hudson, S. R. and R. E. Brandt, 2005: A look at the surface-based temperature inversion on the antarctic plateau. *J. Climate*, **18** (11), 1673–1696.
- International Association for Antarctic Tour Operations, 2014: Statistics. [Available online at <http://iaato.org/tourism-statistics>].
- IPCC, 2013: *Climate Change 2013: The Physical Science Basis*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.
- Jiménez, P. A. and J. Dudhia, 2012: Improving the representation of resolved and unresolved topographic effects on surface wind in the wrf model. *J. Appl. Meteor. Climatol.*, **51** (2), 300–316.
- Jolliffe, I. T. and D. B. Stephenson, 2012: *Forecast verification: a practitioners guide in atmospheric science*. John Wiley & Sons.

- Kalnay, E., 2003: *Atmospheric modeling, data assimilation, and predictability*. Cambridge university press.
- King, J. C. and J. Turner, 2007: *Antarctic meteorology and climatology*. Cambridge University Press.
- Kuipers Munneke, P., M. Van den Broeke, J. Lenaerts, M. Flanner, A. Gardner, and W. Van de Berg, 2011: A new albedo parameterization for use in climate models over the Antarctic ice sheet. *J. Geophys. Res.*, **116** (D5).
- Kumar, A., S. R. Bhowmik, and A. K. Das, 2012: Implementation of Polar WRF for short range prediction of weather over Maitri region in Antarctica. *J. Earth Syst. Sci.*, **121** (5), 1125–1143.
- Lazzara, M. A., G. A. Weidner, L. M. Keller, J. E. Thom, and J. J. Cassano, 2012: Antarctic Automatic Weather Station Program: 30 Years of Polar Observation. *Bull. Amer. Meteor. Soc.*, **93** (10), 1519–1537.
- Liu, C. and Q. Xiao, 2013: An ensemble-based four-dimensional variational data assimilation scheme. part iii: Antarctic applications with advanced research wrf using real data. *Mon. Wea. Rev.*, **141** (8), 2721–2739.
- Mayewski, P. A., et al., 2009: State of the Antarctic and Southern Ocean climate system. *Rev. Geophys.*, **47** (1).
- McCuen, R. H., Z. Knight, and A. G. Cutter, 2006: Evaluation of the nash–sutcliffe efficiency index. *J. Hydrologic. Eng.*, **11** (6), 597–602.
- McLandress, C., T. G. Shepherd, S. Polavarapu, and S. R. Beagley, 2012: Is Missing Orographic Gravity Wave Drag near 60°CS the Cause of the Stratospheric Zonal Wind Biases in Chemistry-Climate Models? *J. Atmos. Sci.*, **69** (3), 802–818.
- Mesoscale & Microscale Meteorology Division, University Center for Atmospheric Research, 2014: Antarctic Mesoscale Prediction System Information. [Available online at <http://www.mmm.ucar.edu/rt/amps/information/information.html>].
- Michel, Y. and T. Auligné, 2010: Inhomogeneous background error modeling and estimation over antarctica. *Mon. Wea. Rev.*, **138** (6), 2229–2252.
- Miloshevich, L. M., H. Vömel, D. N. Whiteman, and T. Leblanc, 2009: Accuracy assessment and correction of Vaisala RS92 radiosonde water vapor measurements. *J. Geophys. Res.*, **114** (D11).
- Monaghan, A. J., D. H. Bromwich, J. G. Powers, and K. W. Manning, 2005: The Climate of the McMurdo, Antarctica, Region as Represented by One Year of Forecasts from the Antarctic Mesoscale Prediction System. *J. Climate*, **18** (8), 1174–1189.

- Monaghan, A. J., D. H. Bromwich, and D. P. Schneider, 2008: Twentieth century Antarctic air temperature and snowfall simulations by IPCC climate models. *Geophys. Res. Lett.*, **35** (7).
- National Center for Atmospheric Research, Earth Observing Laboratory, 2011: Concordiasi 2010 Quality Controlled Driftsonde Data Set. [Available online at <http://data.eol.ucar.edu/datafile/nph-get/221.002/readme.Concordiasi.driftsonde.pdf>].
- National Snow and Ice Data Center, 2014: Passive Microwave Sea Ice Products. [Available online at <http://nsidc.org/data/seaice/pm.html>].
- Nigro, M. A., J. J. Cassano, M. A. Lazzara, and L. M. Keller, 2012: Case study of a barrier wind corner jet off the coast of the Prince Olav Mountains, Antarctica. *Mon. Wea. Rev.*, **140** (7), 2044–2063.
- NOAA, Environmental Modelling Center, 2014: Global Forecasting System. [Available online at <http://www.emc.ncep.noaa.gov/index.php?branch=GFS>].
- Nordeng, T. E., G. Brunet, and J. Caughey, 2007: Improvement of weather forecasts in polar regions. *Bull. World Meteor. Org.*, **56** (4), 250–256.
- Oppenheimer, M., 1998: Global warming and the stability of the west antarctic ice sheet. *Nature*, **393** (6683), 325–332.
- Pavelsky, T. M., J. Boé, A. Hall, and E. J. Fetzer, 2011: Atmospheric inversion strength over polar oceans in winter regulated by sea ice. *Clim. Dynam.*, **36** (5–6), 945–955.
- Pendlebury, S. F., N. D. Adams, T. L. Hart, and J. Turner, 2003: Numerical weather prediction model performance over high southern latitudes. *Mon. Wea. Rev.*, **131** (2), 335–353.
- Pietroni, I., S. Argentini, and I. Petenko, 2014: One year of surface-based temperature inversions at dome c, antarctica. *Boundary-layer Meteor.*, **150** (1), 131–151.
- Plougonven, R., A. Hertzog, and L. Guez, 2013: Gravity waves over Antarctica and the Southern Ocean: consistent momentum fluxes in mesoscale simulations and stratospheric balloon observations. *Quart. J. Roy. Meteor. Soc.*, **139** (670), 101–118.
- Powers, J. G., 2007: Numerical prediction of an Antarctic severe wind event with the Weather Research and Forecasting (WRF) model. *Mon. Wea. Rev.*, **135** (9), 3134–3157.

- Powers, J. G., K. W. Manning, D. H. Bromwich, J. J. Cassano, and A. M. Cayette, 2012: A decade of Antarctic science support through AMPS. *Bull. Amer. Meteor. Soc.*, **93** (11), 1699–1712.
- Pritchard, H., S. Ligtenberg, H. Fricker, D. Vaughan, M. Van den Broeke, and L. Padman, 2012: Antarctic ice-sheet loss driven by basal melting of ice shelves. *Nature*, **484** (7395), 502–505.
- Rabier, F., et al., 2013: The Concordiasi Field Experiment over Antarctica: First Results from Innovative Atmospheric Measurements. *Bull. Amer. Meteor. Soc.*, **94** (3), ES17–ES20.
- Russell, A. and G. R. McGregor, 2010: Southern hemisphere atmospheric circulation: impacts on antarctic climate and reconstructions from antarctic ice core data. *Climatic Change*, **99** (1-2), 155–192.
- Schlosser, E., M. Duda, J. Powers, and K. W. Manning, 2008: Precipitation regime of Dronning Maud Land, Antarctica, derived from Antarctic Mesoscale Prediction System (AMPS) archive data. *J. Geophys. Res.*, **113** (D24).
- Seefeldt, M. W. and J. J. Cassano, 2012: A description of the Ross Ice Shelf air stream (RAS) through the use of self-organizing maps (SOMs). *J. Geophys. Res.*, **117** (D9).
- Son, S.-W., N. F. Tandon, L. M. Polvani, and D. W. Waugh, 2009: Ozone hole and Southern Hemisphere climate change. *Geophys. Res. Lett.*, **36** (15).
- Steig, E. J., et al., 2013: Recent climate and ice-sheet changes in west antarctica compared with the past 2,000 years. *Nat. Geosci.*, **6** (5), 372–375.
- Steinhoff, D. F., D. H. Bromwich, and A. Monaghan, 2012: Dynamics of the foehn mechanism in the McMurdo Dry Valleys of Antarctica from Polar WRF. *Quart. J. Roy. Meteor. Soc.*, **139** (675), 1615–1631.
- Tastula, E.-M. and T. Vihma, 2011: WRF model experiments on the Antarctic atmosphere in winter. *Mon. Wea. Rev.*, **139** (4), 1279–1291.
- Tastula, E.-M., T. Vihma, and E. L. Andreas, 2012: Evaluation of Polar WRF from Modeling the Atmospheric Boundary Layer over Antarctic Sea Ice in Autumn and Winter. *Mon. Wea. Rev.*, **140** (12), 3919–3935.
- Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, **106** (D7), 7183–7192.
- Tilley, J. S. and D. H. Bromwich, 2005: Short-to medium-range numerical weather prediction in the arctic and antarctic. *Bull. Amer. Meteor. Soc.*, **86** (7), 983–988.

- Turner, J., T. Lachlan-Cope, S. Colwell, G. Marshall, and W. Connolley, 2006: Significant warming of the antarctic winter troposphere. *Science*, **311** (5769), 1914–1917.
- Turner, J. and G. J. Marshall, 2011: *Climate change in the polar regions*. Cambridge University Press.
- Uotila, P., A. Pezza, J. Cassano, K. Keay, and A. Lynch, 2009: A comparison of low pressure system statistics derived from a high-resolution NWP output and three reanalysis products over the Southern Ocean. *J. Geophys. Res.*, **114** (D17).
- Vaisalla, 2013: RS92 Technical Data. [Available online at <http://www.vaisala.com/Vaisala%20Documents/Brochures%20and%20Datashets/RS92SGP-Datasheet-B210358EN-F-LOW.pdf>].
- Vaughan, D. G., et al., 2003: Recent rapid regional climate warming on the Antarctic Peninsula. *Climatic Change*, **60** (3), 243–274.
- Vihma, T., T. Kilpeläinen, M. Manninen, A. Sjöblom, E. Jakobson, T. Palo, J. Jaagus, and M. Maturilli, 2012: Characteristics of temperature and humidity inversions and low-level jets over svalbard fjords in spring. *Adv. Meteor.*, **2011**.
- Vömel, H., et al., 2007: Radiation dry bias of the Vaisala RS92 humidity sensor. *J. Atmos. Oceanic Technol.*, **24** (6), 953–963.
- Wang, J., T. Hock, S. A. Cohn, C. Martin, N. Potts, T. Reale, B. Sun, and F. Tilley, 2013a: Unprecedented upper-air dropsonde observations over antarctica from the 2010 concordiasi experiment: Validation of satellite-retrieved temperature profiles. *Geophys. Res. Lett.*, **40** (6), 1231–1236.
- Wang, J., L. Zhang, A. Dai, F. Immler, M. Sommer, and H. Vömel, 2013b: Radiation Dry Bias Correction of Vaisala RS92 Humidity Data and Its Impacts on Historical Radiosonde Data. *J. Atmos. Oceanic Technol.*, **30** (2), 197–214.
- Wilks, D. S., 2011: *Statistical methods in the atmospheric sciences*, Vol. 100. Academic press.
- World Meteorological Organization, 2013: World Weather Research Program, Polar Prediction Project Science Plan. [Available online at http://www.wmo.int/pages/prog/arep/wwrp/new/documents/Final_WWRP_PPPScience_Plan.pdf].
- Zhang, Y., D. J. Seidel, J.-C. Golaz, C. Deser, and R. A. Tomas, 2011: Climatological characteristics of arctic and antarctic surface-based inversions*. *J. Climate*, **24** (19), 5167–5186.