Fall 2009

Evaluating the performance of model-derived wind forecasts at Summit, Greenland during the summer of 2008

Jeffrey A. Luxford
University of New Hampshire, Durham

EVALUATING THE PERFORMANCE OF MODEL-DERIVED WIND FORECASTS AT SUMMIT, GREENLAND DURING THE SUMMER OF 2008

BY

JEFFREY A. LUXFORD
Baccalaureate of Science, St. Cloud State University, 2001

THESIS

Submitted to the University of New Hampshire
In Partial Fulfillment of
the Requirements for the Degree of

Master of Sciience
In
Earth Sciences

September, 2009

This thesis has been examined and approved.

Jack Dibb

Thesis Director, Dr. Jack E. Dibb
Research Associate Professor of Earth Sciences and EOS

Carolyn Jordan

Dr. Carolyn Jordan
Research Scientist
Climate Change Research Center

Robert J Griffin

Dr. Robert Griffin
Associate Professor of Civil and Environmental Engineering
Rice University

20 AUGUST 2009
Date

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to thank Dr. Jack Dibb for the tremendous opportunity to work at Summit, Greenland during the summer of 2008. In addition to being able to visit a part of the world that very few people have experienced, I appreciated the opportunity to conduct research in a very challenging environment. Thank you for your patience and guidance during my time at UNH, Jack! The advice and friendship of committee members Dr. Carolyn Jordan and Dr. Robert Griffin went a long way in making this document possible. I don't think Carolyn knew what she was getting into when agreeing to be on my committee, but she did a remarkable job keeping me focused and motivated to reach this day. I am grateful that the Fuelberg group at Florida State University provided me with WRF data in order to enhance my research. I would especially like to recognize the efforts of Walter Sessions, who explained various details and offered significant insight over the course of many, many emails. Gong-Do Hwang was invaluable when it came to helping me extract the model data that I needed. His help let me circumvent many stressful hours of programming! Finally, a sincere thank you to all of my family and friends who have supported me while I completed this journey...can't wait to see where the next step will lead. This research was made possible by the National Science Foundation Division of Atmospheric Sciences' support of the GSHOX program.

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

EVALUATING THE PERFORMANCE OF MODEL-DERIVED WIND FORECASTS AT SUMMIT, GREENLAND DURING THE SUMMER OF 2008

by

Jeffrey A. Luxford
University of New Hampshire, September, 2009

The study of atmospheric chemistry relies on data from weather models to determine how air parcels move through the atmosphere. The performance of weather models can be negatively impacted by a number of factors, many of which are present at Summit, Greenland. Meteorological observations at Summit during the summer of 2008 included profiles of wind conditions over the lowest part of the troposphere. Data from the profiles were compared to output from two weather models. It was found that both models had difficulty predicting wind speed at the 650 hPa level, roughly 300-400 m above ground level. Analysis at other sites indicated that the poor performance at Summit was primarily due to 650 hPa being in or near the planetary boundary layer (PBL), but a lack of observational data going into model initialization and sparse regional data were also factors.

CHAPTER I

BACKGROUND

Recent projections from the Intergovernmental Panel on Climate Change have underestimated the rate of increase in global mean temperature (Liu et al., 2008). While they have been correct in predicting that the most significant changes would occur in the Arctic (Graversen et al., 2008), finer details have been poorly resolved due to uncertainties that result from trying to represent an environmental system by using a computer model (e.g. Cubasch et al., 1992; Hasselmann et al., 1993; Wigley and Raper, 2001). Understanding those uncertainties as they relate to the operation of weather models in the Arctic was the motivation behind this research. More specifically, the focus was on quantifying the performance of model-derived wind forecasts.

Understanding modeling uncertainties in the Arctic has become more pressing in recent years due to interest in research subjects such as the stratospheric ozone hole (e.g. Brune et al., 1991), Arctic haze (e.g. Schnell, 1984), and pollution transport (e.g. Wofsy et al., 1992). International research campaigns like Arctic Research of the Composition of the Troposphere from Aircraft and Satellites (ARCTAS) (Jacob et al., 2009) and POLar study using Aircraft, Remote sensing, surface measurements and modeling of Climate, chemistry, Aerosols, and Transport (POLARCAT) (http://www.polarcat.no/motivation/polarcat_white_paper.pdf/view) have been motivated by the need to reduce these uncertainties.

A basic component of transport modeling is the background meteorological data, especially wind conditions. The wind conditions come from a weather model that, in turn, is ultimately based on meteorological observations. Therefore, fewer observational data will increase the uncertainty in all subsequent steps in the modeling process (Dreher, 2006, Yussouf and Stensrud, 2007).

The Arctic exhibits a low density of meteorological observations, especially when compared to mid-latitudes. This is apparent when looking at the sites where measurements of meteorological variables are made by rawinsondes, instrument packages that sample the vertical extent of the atmosphere (Figure 1). Despite the usefulness of this data it is impractical, if not impossible, to gather it on sufficiently dense temporal and spatial scales to continuously and accurately describe the state of the entire atmosphere. Instead, the available observations are interpolated across the weather model's domain in order to establish the initial conditions for the 00-hour forecast. In addition to prediction of meteorological conditions, weather models have been used to better understand the state of the Arctic atmosphere and its impact on the conditions at the surface both directly (Martin and Moore, 2006) and indirectly (Hutterli et al., 2007).

One location that has been the site of a lot of atmospheric research is the Greenland Environmental Observatory at Summit (hereafter referred to as Summit), located in central Greenland. Established in 1989 as a joint venture by the National Science Foundation of the United States and the Danish Commission for Scientific Research, Summit was founded to aid in the recovery of the GISP2 ice core (e.g. Grootes et al., 1993; Crowley, 2000; Stuiver et al., 1995; O'Brien et al., 1995; Taylor et al., 1993). Since that time, the focus has shifted to studying the composition of snow in the firn layer (e.g. Slater et al., 2001; Galbavy et al., 2007; Swanson et al., 2002) and understanding the atmospheric dynamics and chemistry that occur in the lowest part of the troposphere. (e.g. Hutterli et al., 2007; Sjostedt et al., 2007; Jacobi et al., 2002). This includes the surface-based layer that is capped by a thermal inversion and is known as the planetary boundary layer (PBL) Summit is not a permanent rawinsonde site, so vertical profile data is only collected as needed. In 2008 measurements were made as part of the Greenland Summit Halogen experiment (part of the summertime POLARCAT campaign), but these occurred on a vertical scale much smaller than that of a rawinsonde profile. The primary purpose for making these measurements was to better understand how the PBL thickness changes diurnally and what impact that has on atmospheric chemistry at Summit.

When considering questions related to atmospheric chemistry, transport pathways to and from the observation site can provide useful information that is not attainable through on-site analysis. Computer models can be used to determine air parcel pathways to (back-trajectories) and from (forward trajectories) a specific point. Of these two types, back-trajectories are most often calculated because they can help to identify or confirm air mass source regions. The two most common computer models used for generating back-trajectories are FLEXPART (http://zardoz.nilu.no/~andreas/flexpart/flexpart8.pdf) and the Hybrid Single Particle Lagrangian Integrated Trajectory Model (HySPLIT) (http://www.arl.noaa.gov/HYSPLIT_info.php).

One example of the potential differences between FLEXPART and HySPLIT output for particle pathways to Summit is shown in Figure 2. The left-hand side shows output from FLEXPART, where the numbers indicate the location of the center of the pathway and how many days it took to reach Summit from that point. The right-hand side shows the movement of particles that arrived at 10, 500, and 1500 meters above ground level at Summit as determined by HySPLIT. All three paths track very close to the surface during the first 60 hours of the period, showing the importance of model performance near ground level.

When comparing the two models, they exhibit distinct differences in particle residence time over the ice sheet and, more strikingly, they produce paths that travel towards Summit in opposite directions (counter-clockwise in FLEXPART, clockwise in HySPLIT). Note that a comprehensive examination of back-trajectories was not attempted, so it is unknown if this represents the typical degree of difference between these two models. Rather, this comparison is offered to illustrate the possibility that one or both of these transport models have substantial difficulty dealing with the physical conditions at Summit.

While these two particular models take their background meteorological parameters from different sources, all transport models rely on meteorological fields derived from weather forecast models and observational data. Again, this means that the accuracy of weather models is tied to all research that deals with atmospheric transport.

The accuracy of a synoptic model in a remote region has been given some consideration by Adhikary et al. (2009) in connection with the Intercontinental Chemical Transport Experiment-Phase B. The area of interest was the atmosphere over the Pacific Ocean a short distance from the west coast of the United States, so analytical measurements were made aboard aircraft. The main objective was to examine the transport of various molecular species in order to better understand their source regions and the chemical and physical processes that occurred during transport. The Weather

Research and Forecasting (WRF) model was used to help meet these objectives, but the lack of weather observations over the open ocean raised concern that their model may not be reliable. In order to verify model predictions, measurements of meteorological variables such as temperature and wind speed were made aboard research aircraft. Results indicated that the model was producing data that correlated well with observations of temperature and wind speed at different altitudes. However, in this case individual observations were not compared to corresponding model output. Instead, verification was indicated by a comparison of model and observational means (Figure 3). While this was a beneficial analysis, a direct comparison of values at specific forecast sites and times would provide a better test of weather model accuracy across regions with sparse observational data.

Global climate models (GCMs) have larger domains and coarser resolution when compared to synoptic models, but they operate using the same basic dynamics. Walsh et al. (2008) have shown that GCM output of sea level pressure, surface temperature, and precipitation across nested domains covering Greenland and Alaska was no more or less accurate than across larger domains such as the Arctic or Northern Hemisphere. While it is encouraging that GCMs are not systematically worse over Arctic sub-regions, verification of synoptic models in the Arctic has not been given the same consideration. This may be due to a lack of corresponding observations and sparse human settlement, but as more research takes place in this region, an improved understanding of the capabilities and limits of weather modeling in the Arctic will be needed.

Unlike most of the Arctic, which has an elevation near mean sea level (MSL) and is relatively flat, Greenland presents an anomalous barrier to mean geostrophic flow

(Martin and Moore, 2006; Renfrew et al., 2008). Summit lies at the apex of the Greenland ice sheet, located at 72.5°N, 38.5°W and approximately 3200 m above MSL. This unique location raises the question of how well the weather models are able to determine the state of the atmosphere above Summit. Possible factors influencing model performance are: 1) the inability to adequately resolve conditions in and near the PBL, 2) a lack of observational data at Summit being incorporated into model initialization 3) sparse data coverage around and upstream of Summit, and 4) Summit's location at a high elevation and near a sharp orographic gradient.

In order to address the question of model accuracy, direct comparisons will be made between observational data collected at Summit and two different weather models, the WRF and the Global Forecast System (GFS). The primary concern is that output from synoptic models may be less accurate at Summit due to one or more of the reasons listed in the previous paragraph. Given that many current research projects utilize models to study atmospheric transport, wind parameters were used to evaluate model output. Meteorological data at Summit was collected with a tethersonde, which routinely made measurements up to 500 m above ground level. To assess model accuracy at Summit, first the observations were compared to GFS and WRF output. Archived data from rawinsonde observations were available at other sites and were also compared directly to GFS and WRF output. Then the model accuracy at Summit was compared to the accuracy at the rawinsonde sites. However, based on the potential causes for model error at Summit, no individual rawinsonde site had all of the same characteristics. Therefore, it was necessary to evaluate a number of sites that possessed some of these attributes (Table 1).

CHAPTER II

DATA

Tethersonde

A tethersonde was used to make high-frequency measurements of the lowest part of the atmosphere, with a focus on the time periods when the PBL thickness was most likely to be changing. The tethersonde flights were made during the summer of 2008 by scientists from the University of Houston and other members of the Greenland Summit Halogen team. There were a total of 135 flights between June 6 and July 9, 2008, with flights usually taking between 15 and 25 minutes.

The tethersonde setup consisted of a helium-filled balloon attached to approximately 600 m of cable that was spooled on an electric winch. A Vaisala Tethersonde TTS111 instrument package was suspended beneath the balloon and recorded data as the cable was unwound and the balloon rose. Measurements of air temperature, dew point, atmospheric pressure, wind speed, and wind direction were recorded every second during each flight. Wind speed and direction were observed directly by the anemometer and compass within the instrument package. Measurements of these two variables were contaminated by the mechanical retrieval of the tethersonde, with observations showing enhanced speed and direction biased towards the location of the winch. Because of these issues, data recording was stopped at the maximum altitude of the flight. Data was transmitted to a receiving station on the ground and processed with software provided by the instrument manufacturer. The height of the observations was calculated by the software using the hydrostatic equation and a ground level of 3200 m above MSL.

Rawinsonde

The primary difference between a tethersonde and a rawinsonde is that the latter is untethered. Rawinsondes often reach pressure levels above 100 hPa, or more than 16000 m above MSL, so reported observations span most of the atmosphere. Launches are made at many weather observation posts across the globe (See Figure 1 for a partial map.) and they occur at 0000 and 1200 GMT each day. The primary purpose of these measurements is to provide data for the initialization of meteorological models, which tie model output to the measured physical state of the atmosphere. Data from the rawinsondes include the same variables observed by the tethersonde, although wind speed and direction are derived from global positioning system tracking and not measured directly. Measurements are reported at certain mandatory pressure levels during each flight and also where the slope of the temperature profile changes. The data collected from each sounding is available through the University of Wyoming's weather page (http://weather.uwyo.edu/upperair/sounding.html).

Forecast Models

The version of the GFS model that was used for this study was the GFS4, hereafter referred to as GFS. It was run four times each day by the National Oceanic and Atmospheric Administration and stored on the National Operational Model Archive and Distribution System (NOMADS) website (http://nomads.ncdc.noaa.gov/data/gfs4/). The domain was global, with grid spacing of approximately 0.5° latitude x 0.5° longitude (Table 2). The model time-step was 3 hours, with a total run time of 180 hours.

The WRF model output that was used was generated by the Fuelberg group at Florida State University in support of flight activities for the ARCTAS campaign. The domain was a polar-projected 45km x 45km grid, centered on 90°N and extended to 5.4°N (Table 2). The WRF was initialized every day at 1200 GMT by taking data from the current 00-hour GFS analysis field and interpolating it across the WRF domain. The WRF was then run in forecast mode for a 12-hour 'spin-up' period. Model output for these 12 forecast hours was discarded and the subsequent data valid from 0000 to 2300 GMT the next day was preserved as the operational forecast for ARCTAS activities.

Data Limitations

GFS and WRF model data were available at 25 hPa intervals. Surface pressure levels at Summit ranged between 690 and 675 hPa and the tethersonde never rose above 610 hPa. Because very few tethersonde flights surpassed 625 hPa and 675 hPa was often less than 100 m above ground level, direct comparisons to Summit observations were limited to 650 hPa or about 300-400 m above ground level.

There were multiple factors limiting the total number of tethersonde flights, including: the research objective, environmental conditions, and technical reasons. The primary objective of the tethersonde measurements was to document the movement of the top of the PBL. Because its movement was most often linked to the diurnal fluctuation of solar energy, there was not an even distribution of flights across all hours of the day (Figure 4). Also, tethersonde flights were suspended when surface speeds exceeded 10 m s^{-1} in order to ensure the safety of the balloon and persons operating the tethersonde. This led to observations at Summit being biased towards low wind speeds.

Lastly, equipment problems also occurred that inhibited sampling, most notably during the second and fourth weeks of the campaign.

In addition to wind speed, observations of wind direction at Summit were going to be compared to model-derived forecasts. However, during the second week of measurements at Summit it was noted that the tethersonde's compass was reporting inaccurate wind direction throughout the duration of each flight. By the end of the third week of observations it was discovered that metal near the launch site led to the inaccurate compass readings during tethersonde flights and the situation was immediately rectified. Unfortunately it could not be determined when the erroneous data began so only there are only 8 flights corresponding to rawinsonde launch times that have reliable wind direction data. As a result, no meaningful analysis of wind direction accuracy could be conducted.

Finally, it has been noted that Summit's unique geographic setting was not comparable to any other data site. Because there was no analog among rawinsonde sites that would allow for a direct comparison to model forecast skill at Summit, data from a suite of rawinsonde locations was examined (Table 1). Among the study sites were four Arctic locations near Summit's latitude: Barrow, Cambridge Bay, Resolute, and Aasiaat. Points west of Summit were chosen because atmospheric flow generally moves from west to east, so model performance would be known prior to the mean flow being disrupted by the orography of Greenland. The Pas, in central Manitoba, Canada, was selected due to its remote location (8000 km from the nearest adjacent rawinsonde site) and position somewhat close to a large change in orography (1100 km east of the Rocky Mountains). There were no rawinsonde sites at Summit's elevation, so Flagstaff, AZ was selected as the most comparable (2000 m above MSL). Finally, an "idealized" spot was considered, where the location did not possess any of the potentially negative aspects that were present at Summit. In this case, the rawinsonde site at Gray, ME was downstream from several nearby observation sites, had data incorporated into weather model initializations, was located by no large topographic features, and was near MSL.

CHAPTER III

METHODOLOGY

Data Extraction

Data during each tethersonde flight was recorded at one-second intervals. There were often brief periods during flights when the barometer recorded an increase in pressure, meaning variations in the tethersonde's buoyancy would cause it to drop in elevation. As they were a potential source of contamination in the wind speed data, these particular observations were avoided. For each tethersonde flight that rose above 650 hPa, the observation closest to 650 hPa that occurred immediately prior to the first observation above 650 hPa was selected. This ensured that the instrument was ascending at the time of the observation.

A total of 122 tethersonde flights yielded data near 650 hPa. All of the selected observations occurred within 0.37 hPa of 650 hPa and for more than half of the flights the observations were within 0.10 hPa. Each tethersonde observation was linked to the closest GFS timestep, meaning all tethersonde data was not more than 90 minutes from its corresponding GFS data. This also meant that there were often multiple observations during a particular 3-hour window. In these cases, the flights closest to the timing of the GFS forecast were selected for comparison and other observations were not considered. This yielded a total of 80 observations in the GFS comparison dataset. When making comparisons between Summit observations and WRF data, the one-hour WRF timesteps meant that the 122 flights which reached 650 hPa were included.

As previously discussed, tethersonde flights were not always coincident with the timing of global rawinsonde launches. In order to align tethersonde data and rawinsonde data, tethersonde measurements made within 90 minutes of 0000 and 1200 GMT were tied to the corresponding rawinsonde data. If more than one tethersonde flight occurred within the 3-hour window around a particular rawinsonde launch, the flight closest to the target time was used for comparison.

Many rawinsonde profiles did not contain data for comparison to model output at 650 hPa because that pressure is not a required reporting level. In order to maximize the number of soundings available for comparison, soundings with observations at 650 ± 5 hPa were identified and the observation closest to 650 hPa was selected for comparison to model data. Other pressures selected in this study were mandatory reporting levels, so data extraction was not necessary at 925 and 500 hPa.

GFS data on the NOMADS website was stored in grib2 format as separate files for every 3-hour time step. Each file contained all model data for the entire vertical and horizontal domain, with information sorted by vertical level and stored in two-dimensional arrays. Grid points closest to the observation sites were identified and data extraction was performed using a UNIX program that included the wgrib2 decoder available on the Climate Prediction Center (CPC) website (http://www.cpc.ncep.noaa.gov/products/wesley/). This routine was performed for each desired grid point and at the desired pressure levels. Because wind vectors were in their u and v components, the Pythagorean theorem was used to calculate the wind speed.

Because the WRF files from Florida State University's Fuelberg group came in grib format, a separate UNIX program was written to sort information into a onedimensional array. Once more the model gridpoints closest to the observation sites were identified and the desired information was extracted. The UNIX program included the wgrib program, which is also available on the CPC website. Like the GFS, WRF wind vector data is stored in u and v components, so wind speed had to be calculated.

Analysis

The GFS and WRF wind speeds were plotted against corresponding observational data from all sites in Table 1 in order to determine the correlation between measurements to model output. At Summit, a total of 42 flights that reached 650 hPa were not counted in the comparison to GFS output because they took place during the same 3-hour window as another tethersonde observation (see previous section on extracting tethersonde data). Examination of the plot led to the identification of a significant outlier in the Summit data (Figure 5). The outlier was the maximum measured wind speed at 650 hPa, 17.1 m s^{-1}, and it fell more than three standard deviations beyond the mean of all Summit observations (2.58 and 5.84 m s^{-1}, respectively). Weather conditions prevented subsequent flights on this day, but the observation was believed to be accurate based on available evidence. The strong wind speed at 650 hPa was recorded while surface wind speeds approached 10 m s^{-1} and these measurements occurred as a vigorous synoptic system moved over Summit. Additionally, rawinsonde data upstream from Summit indicated peak wind speeds were greater than 13 m s^{-1} at 650 hPa near The Pas. Despite indications that the observation was accurate, the impact of that measurement on model performance statistics at Summit was significant enough to make some calculations including and excluding the outlier, such that any conclusions do not appear to be unduly biased by this one observation. The impact of the outlier was more evident in the GFS statistics due to a worse forecast and smaller dataset compared to the WRF. Therefore, certain statistics for comparisons of GFS output to Summit observations were reported both with and without the outlier.

In order to compare model accuracy at each of the sites, the RMSE of model output was calculated and divided by the mean observed wind speed at a particular location. The resulting ratio accounted for the range in mean wind speeds between sites, making a small RMSE less important at a site where wind speeds were strong compared to a site where speeds were light. The ratio of error to mean observed speed was calculated for each site at each forecast hour, which allowed for direct comparisons between sites, forecast times, and models.

CHAPTER IV

RESULTS

Summit Compared to Model Data

Tethersonde wind speeds at 650 hPa were compared to the earliest corresponding forecast from the GFS in order to gauge how well the model initial conditions represented the state of the atmosphere (Figure 5). A linear regression through the dataset including(excluding) the outlier had a slope of 0.63(0.68) and an R^2 of 0.33(0.48). Forcing the regression through the intercept resulted in a slope of 0.83(0.81) and an R^2 of 0.29(0.46).

Results similar to the GFS analysis were found when the tethersonde observations were compared to corresponding WRF data (Figure 6). A linear regression fit to the entire dataset including the outlier had an R^2 of 0.09 and a slope of 0.28, which increased to 0.70 when the line was fit through the origin.

The possibility that model accuracy depends on time of day was considered, and the data was separated into two subsets based on visual evidence of when the top of the PBL lowered at night (i.e. the formation of a ground-based fog layer) and temperature profiles indicating when the top of the PBL began to rise during the morning. The period with less solar input, what was effectively nighttime during the Arctic summer, showed improved model accuracy (Figure 7). The slope of the regression line for the nighttime period was nearly twice that of the daytime (0.38 compared to 0.21) and the offset of the line was smaller by 0.4 m s^{-1}.

The GFS and WRF did not predict the 650 hPa wind speed at Summit very well, with RMSE roughly one-half of the mean wind speed and linear regressions that exhibited a poor correlation to the observed data. While these results supported the theory that models have difficulty handling the conditions at Summit, comparisons to other observational sites were needed in order to understand why this was happening.

Inter-Site Comparisons

After isolating the tethersonde observations that occurred around the time of rawinsonde launches, plots of observed wind speed compared to GFS wind speed were made for each forecast verification time at each observation site. For this comparison the outlier at Summit was excluded, but statistics are reported in Table 3. The regression, correlation, and RMSE at the rawinsonde sites trended towards those observed at Summit over time, in most cases becoming comparable by the 72-hour forecast (Table 3). This trend was illustrated by plots from Cambridge Bay, Barrow, and Summit, based on plots at 00, 24, and 72-hours (Figure 8). The slope of the regression lines decreased noticeably at Cambridge Bay (1.01, 1.05, 0.63) and Barrow (0.89, 0.84, 0.61), but not at Summit (0.59, 0.54, 0.52). The same variation was observed with respect to R^2 values at Cambridge Bay (0.96, 0.74, 0.32), Barrow (0.78, 0.71, 0.44), and Summit (0.30, 0.29, 0.34).

The proximity of 650 hPa to Summit's surface (generally 300-400 m above ground level) and the suspension of tethersonde flights during high winds led to the mean observed wind speed at Summit being smaller than the rawinsonde data from other sites reported here (Figure 9). As described previously, RMSE was normalized by the mean observed wind speed at each site so a smaller ratio indicates a better forecast.

The GFS performance was determined at all observational sites at 650 hPa over a period of 72 hours (Figure 10). Including the outlier at Summit, the error was about 50% of the wind speed during the first 36 hours, but during that period other sites saw their error ratio increase from roughly 20% to 30-35%. The impact of Summit's outlier is evident with its removal, but even then Summit was worse than the other locations for the first 24 hours.

The PBL influence at Summit at 650 hPa could be an important reason for poor model performance, where friction near ground level caused lighter mean wind speeds compared to 650 hPa wind speeds at the rawinsonde sites. Model performance was examined at a higher pressure at the rawinsonde sites in order to determine how the GFS dealt with the PBL at those locations. While the top of the PBL varied in height from site to site, during the summer months it moved through the 925 hPa level at all rawinsonde locations (except Flagstaff, where that level was below ground). Mean observed wind speeds near the PBL ranged from 4.6 m s^{-1} at Aasiaat to 7.5 m s^{-1} at Cambridge Bay (Figure 11) This differed from 650 hPa, where wind speeds at rawinsonde sites were 9.1 to 14.0 m s^{-1} compared to 6.4 m s^{-1} at Summit (including the outlier) (Figure 9).

Model performance at 925 hPa degraded significantly from the 00-hour forecast to the 12-hour forecast at the rawinsonde sites, after which the accuracy remained more or less constant out to 72 hours (Figure 12). By the 12-hour forecast the RMSE to mean wind speed ratio at Barrow, The Pas, and Aasiaat had risen above 0.40 and from 36-72 hours Barrow was higher than 0.50, meaning the GFS was less accurate there than at Summit when the outlier was included. This supported the hypothesis that the GFS had difficulty predicting wind speeds close to and within the PBL.

The influence of the PBL was removed in the 500 hPa comparison, where observational data was more than 5000 m above MSL (Figure 13). Recall that no Summit observations were available at this pressure level. Hence, analysis here focused only on the rawinsonde sites. The same RMSE to mean wind speed ratio was calculated at 12-hour intervals and most sites showed a relatively consistent decline in forecast accuracy. GFS accuracy at most sites was equal to or better than that seen at 650 hPa initially, between 0.10 and 0.20. The ratios increased at a relatively consistent rate and ended up between 0.30 and 0.45, somewhat worse than at 650 hPa. The exception was The Pas, where performance is consistently poor and similar to that observed at Summit. A possible explanation will be presented in the discussion.

Inter-model Comparisons

Differences between the WRF and GFS go beyond those highlighted in Table 2 and range from cloud microphysics to terrain parameterization. That means a difference in output could be attributed to a number of different factors. It is beyond the scope of this paper to examine the specific components and their individual influence on model performance; rather, any differences will be attributed to the whole model.

A comparison between model and observed wind speed correlations for the GFS and WRF at 650 hPa was made, but results from Flagstaff were not included because each timestep had only 5 or 6 observations (Table 4). The difference between GFS and WRF R^2 values at 00 hours ranged from 0.01 to 0.03. The short length of the WRF runs meant that comparisons could only be made out to 24-hours, but even after this short period there was a substantial difference at some sites. While the WRF R^2 at Summit was only 0.05 (excluding the outlier) better than that of the GFS, much larger differences were found at other sites. The 24-hour WRF R^2 at Cambridge Bay was 0.95, 0.88 at The Pas, and 0.85 at Aasiaat. The difference between GFS and WRF correlations ranged from 0.11 to 0.24. Although this is not an explicit measure of model forecast skill, it does hint at the notion that the WRF outperforms the GFS. A better gauge would be looking at the RMSE/mean wind speed ratio to see if there are obvious differences (Table 5).

There was no trend in the differences at 650 hPa between GFS and WRF accuracy for the 00-hour forecast, but the 24-hour WRF performance is better than the GFS at most observation sites (Table 5). Relatively speaking, WRF performance was 30-50% better than the GFS and RMSE ratio improvements ranged from 0.083 at The Pas to 0.131 at Resolute. When this comparison was made for the 00-hour forecasts at the 500hPa level, WRF was slightly worse than GFS at most locations (Table 6). From 00 to 24-hours GFS accuracy declined while WRF performance was relatively steady, which resulted in WRF being more accurate at all sites.

CHAPTER V

DISCUSSION

Based on linear regressions of model output versus observations, it has been shown that predictions of wind speed at Summit, from both GFS and WRF, tended to exceed observed wind speeds at 650 hPa. This did not translate to an evident model bias, however, as data correlations were consistently lower than 0.45. As previously noted, possible causes for this poor model performance are: 1) the top of the PBL being located near 650 hPa, 2) a lack of observational data at Summit being incorporated into model initialization 3) sparse data coverage around and upstream of Summit, and 4) Summit's location at a high elevation and near a sharp orographic gradient.

Because Summit had no analogous observation site among rawinsonde locations, it was necessary to identify a number of sites that possessed some of these attributes (Table 1) in order to test the importance of each potential source of error. In this way model performance at Summit could be compared to model performance at other locations.

Influence of the Planetary Boundary Layer. When evaluating a weather model, a decrease in forecast skill is expected as one proceeds further away from the time of initialization. This is because the size of errors tends to increase with the subsequent calculations at every successive model time step. The lack of this occurring at Summit indicates that there is something unusual occurring at Summit at 650 hPa (Figure 10). Similar behavior is found when examining GFS output compared to rawinsonde observations at 925 hPa (Figure 12), indicating that this behavior is not dependent on local orography or regional data coverage. Instead, the unifying factor among all sites in Figure 12 is that the observations were occurring in or near the PBL. Mean wind speeds are lighter and more uniform at 925 hPa (Figure 11) than at 650 hPa (Figure 9). This is because 650 hPa is in the free troposphere at all rawinsonde locations. This indicates that atmospheric flow was being affected by frictional drag from the earth's surface. Because this results in perturbations that are too small to be resolved by weather models, forecasts of low-level wind conditions tend to be worse than those for other parameters. Jones et al. (2007) found this to be true when they examined how well model ensembles predicted conditions at the surface. They found that the observed nocturnal 10-meter wind speed was not within the ensemble range 57% of the time.

Influence of Observational Data at Summit. A substantial decrease in GFS accuracy between the 00 and 12-hour forecasts is observed at the rawinsonde sites at 925 hPa (Figure 12). Subsequent hours exhibit little or no trend in performance, implying that the relatively good 00-hour forecasts were due to observational data being part of the model initialization. After that time, the benefit was lost as the model had difficulty representing conditions in or near the PBL. This is supported by the fact that a similar decrease of similar magnitude is not present at higher pressure levels, roughly a 0.15 increase at 925 hPa compared to less than 0.10 at 650 and 500 hPa (Figure 10; Figure 13). Based on analysis at 925, 650, and 500 hPa, forecast accuracy at the rawinsonde sites tends to improve with height. Therefore, if observational data at Summit were ingested into a weather model, no improvement in the accuracy of modeled wind speed is expected below 650 hPa after the 00-hour forecast. As one moves higher into the atmosphere model performance would be more likely to improve beyond this initial forecast.

Influence of sparse data coverage. More direct comparisons can be made regarding the lack of observational data being ingested near Summit, focusing on data coverage upstream. In this respect, the rawinsonde sites at Barrow, Aasiaat, and The Pas are most similar to Summit. The possibility that limited regional data was a significant problem at Summit was supported by data at Barrow at 650 hPa, where the 00-hour wind speeds were modeled worse than most other sites and approach the accuracy at Summit at 12 hours (Figure 10). Barrow exhibited an improvement at 24 hours, which may have resulted from the influence of upstream observations across Siberia (Figure 1). The stronger wind speeds at 500 hPa smoothed out this anomaly (Figure 13), whereas the PBL-influenced 925 hPa results showed no positive impact from the upstream sites (Figure 12).

Aasiaat at 925 hPa had GFS wind speeds that were initially almost as inaccurate as Summit's were at 650 hPa (without the outlier), becoming worse than it over the next 12 hours. This was in contrast to Aasiaat at 650 hPa, where there was improvement in the accuracy of the GFS winds at 12 hours. Again, this may have been due to the proximity of upstream rawinsonde sites. Because flow is generally from west to east, the air advected to Aasiaat is usually from a region that features more observation sites and is relatively flat.

At The Pas, GFS accuracy was most obviously poor at 500 hPa (Figure 12). It was suspected that this was a result of being downstream from the Rocky Mountains and their substantial impact on zonal air flow. The one-month period chosen for observation may have been anomalously bad at 500 hPa, but poor model performance was also evident at lower pressure levels.

Influence of elevation. Elevation was also considered, but rawinsonde data was limited. Also, Flagstaff was not a close match because it is more than 1000 m lower than Summit and its observations at 650 hPa were not near the PBL. The limited data in Figure 10 seemed to indicate that altitude was not a substantial factor, as it outperformed sites near MSL. This was contrary to Figure 13, where Flagstaff was one of the worst performers. This indicated that GFS accuracy was not noticeably affected by Flagstaff's elevation. As previously discussed, the variation in wind speed with respect to height would render nearby observations more important at 650 hPa versus 500 hPa. At the higher altitude, the influence of nearby observations would more quickly be advected past Flagstaff and be replaced by the unsampled region over the Pacific Ocean.

Overall. The best results were observed at Gray, ME, both in terms of a low RMSE/mean wind speed ratio and the consistency of that ratio during the 72-hour period (Figure 10; Figure 13). This was expected and it supported the idea that certain desirable attributes of Gray's location resulted in superior model performance. However, even these beneficial factors are not enough to yield consistently better results in or near the PBL (Figure 12). By the 72 hours, the model accuracy at Gray (0.45) was comparable to that at Aasiaat (0.44) and Summit including the outlier (0.46). Accuracy at Gray was significantly worse than Resolute (0.38) and Summit excluding the outlier (0.37). This strongly suggests that the most significant problem in modeling the atmosphere above Summit is trying to represent the near-surface conditions and that poor weather model performance is not a result of limited observational data or orographic conditions at a particular site.

Future Considerations. This study focused on conditions during the summer of 2008, but atmospheric circulations in the Arctic have seasonal tendencies (Ambaum and Hoskins, 2002, Zhou et al., 2002). It is possible that weather model accuracy exhibits a seasonality, improving with the onset of winter as the polar vortex establishes itself and the synoptic storm track shifts southward. Analysis during the winter may show that comparisons between observations and model data at 650 hPa at Summit are more comparable to other sites, as the colder temperatures cause the top of the PBL to remain below that pressure level.

While the lack of accurate data precluded a thorough evaluation of wind direction at Summit, it is likely that model performance would again be negatively impacted at 650 hPa. Accurate measurements of wind direction could be used to evaluate this and either support or refute the theory that particle paths such as those in Figure 2 are an artifact of model errors. As with wind speed, it is expected that weather model performance at higher altitudes would benefit from the incorporation of observations at Summit.

While the scope of this research was limited, it does provide some tentative conclusions and may indicate ways to improve weather model performance at Summit. It was shown that a regional disparity in model accuracy exists, based on the performance at Gray compared to the isolated sites like Barrow and Summit. As such, increasing the density of rawinsonde sites in the Arctic should improve the performance of weather models across the region.

CHAPTER VI

SUMMARY

It was shown that regardless of method used to analyze the accuracy of the model output, wind speeds at Summit were poorly resolved in the models. This was expected, given that there were a number of potential factors that could negatively impact model performance.

The consistency of the error magnitude and lack of forecast improvement at Summit when using a model with higher spatial resolution indicates that the dominant component of the error at 650 hPa is the inability of the weather models to accurately handle conditions in and near the PBL. As such, there would be little improvement in model forecasts near and below the 650 hPa pressure level during the Arctic summer if rawinsonde data at Summit was incorporated into model initialization. The impact would likely be greater at higher altitudes, providing a benefit to general circulation patterns in the free troposphere over Greenland. However, if one is concerned with air parcel trajectories that cross the interior of Greenland near the surface, there will be no improvement due to PBL influence after the initial 00-hour forecast.

REFERENCES

Adhikary, B, et al, Trans-Pacific transport and evolution of aerosols and trace gases from Asia during the INTEX-B field campaign, Atmospheric Chemistry and Physics Discussions, 9, 16381-16439, 2009

Ambaum, M. H. P., Hoskins, B., The NAO troposphere-stratosphere connection, Journal of Climate, 15, 1969-1978, 2002

Brune, W. H., Anderson, J. G., Toohey, D. W., Fahey, D. W., Kawa, S. R., Jones, R. L., McKenna, D. S., Poole, L. R., The potential for ozone depletion in the Arctic polar stratosphere, Science, 252, 1260-1266, 1991

Cubasch, U., Hasselmann, K., Maierreimer, E., Mikolajewicz, U., Santer, B. D., Sausen, R., Time-dependent greenhouse warming computations with a Coupled ocean-atmosphere model, Climate Dynamics, 8, 55-69, 1992

Crowley, T. J., Causes of climate change over the past 1000 years, Science, 289, 270-277, 2000

Dreher, J. G., Manobianco, Manobianco, J., Regional forecast impacts from GEMS Observations, 10th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface, Atlanta, GA, Amer. Meteor. Soc., 2006, Extended abstract available at: http://ams.confex.com/ams/pdfpapers/101509.pdf

Galbavy, E. S., Anastasio, C., Lefer, B., Hall, S. R., Light penetration in the snowpack at Summit, Greenland: Part 1 nitrite and hydrogen peroxide photolysis, Atmospheric Environment, 41, 5091-5100, 2007

Graverson, G. G., Mauritsen, T., Tjernström, M., Källén, E., Svensson, G., Vertical structure of recent Arctic warming, Nature, 541, 53-56, 2008

Grootes, P. M., Stuvier, M., White, J. W. C., Johnsen, S., Jouzel, J., Comparison of oxygen isotope records from the GISP2 and GRIP Greenland ice cores, Nature, 366, 552-554, 1993

Hasselmann, K., Sausen, R., Maierreimer, E., Voss, R., On the cold start problem in Transient simulations with coupled atmosphere-ocean models, Climate Dynamics, 9, 53-61, 1993

Hutterli, M. A., Crueger, T., Fischer, H., Andersen, K. K., Raible, C. C., Stocker, T. F., Siggaard-Andersen, M. L., McConnell, J. R., Bales, R. C., Burkhart, J. F., The influence of regional circulation patterns on wet and dry mineral dust and sea salt deposition over Greenland, Climate Dynamics, 28, 635-647, 2007

Jacob, D. J., et al., The ARCTAS aircraft mission: design and execution, Atmospheric Chemistry and Physics Discussions, 9, 17073-17123, 2009

Jacobi, H. W., Frey, M. M., Hutterli, M. A., Bales, R. C., Schrems, O., Cullen, N. J., Steffen, K., Koehler, C., Measurements of hydrogen peroxide and formaldehyde exchange between the atmosphere and surface snow at Summit, Greenland, Atmospheric Environment, 36, 2619-2628, 2002

Jones, M. S., Colle, A. E., Tongue, J. S., Evaluation of a mesoscale short-range ensemble forecast system over the Northeast United States, Weather and Forecasting, 22, 36-55, 2007

Lui, J., Zhang, Z., Hu, Y., Chen, L., Dai, Y., Ren, X., Assessment of surface air temperature over the Arctic Ocean in reanalysis and IPCC AR4 model simulations with IABP/POLES observations, Journal of Geophysical Research, 113, D10105, doi:10.1029/2007JD009380, 2008

Martin, R., Moore, G. W. K., Transition of a synoptic system to a polar low via interaction with the orography of Greenland, Tellus, 58A, 236-253, 2006

O'Brien, S.M., P.A. Mayewski, L.D. Meeker, D.A. Meese, M.S. Twickler, S.I. Whitlow, Complexity of Holocene climate as reconstructed from a Greenland ice core, Science, 270, 1962-1964, 1995

Renfrew, I. A. et al., The Greenland flow distortion experiment, Bulletin of the American Meteorological Society, 89, 1307-1324, 2008

Schnell, R. C., Arctic haze and the Arctic gas and aerosol sampling program (AGASP), Geophysical Research Letters, 11, 361-364, 1984

Slater, J. F., Dibb, J. E., Keim, B. D., Kahl, J. D. W., Relationships between synoptic-scale transport and interannual variability of inorganic cations in surface snow at Summit, Greenland: 1992-1996, Journal of Geophysical Research, 106, 20897-20912, 2001

Sjostedt, S. J., Huey, L. G., Tanner, D. J., Peischl, J., Chen, G., Dibb, J. E., Lefer, B., Hutterli, M. A., Beyersdorf, A. J., Blake, N. J., Blake, D. R., Sueper, D., Ryerson, T., Burkhart, J., Stohl, A., Observations of hydroxyl and the sum of peroxy radical at Summit, Greenland during summer 2003, Atmospheric Environment, 41, 5122-5137, 2007

Stuiver, M., Grootes, P. M., Braziunas, T. F., The GISP2 $\delta^{18}O$ climate record of the past 16,500 years and the role of the sun, ocean, and volcanoes, Quaternary Research, 44, 341-354, 1995

Swanson, A. L., Blake, N. J., Dibb, J. E., Albert, M. R., Blake, D. R., Rowland, F. S., Photochemically induced production of CH_3Br, CH_3I, C_2H_5I, ethane, and propene within surface snow at Summit, Greenland, Atmospheric Environment, 36, 2671-2682, 2002

Taylor, K. C., Lamorey, G. W., Doyle, G. A., Alley, R. B., Grootes, P. M., Mayewski, P. A., White, J. W. C., Barlow, L. K., The "flickering switch" of late Pleistocene climate change, Nature, 361, 432-436, 1993

Walsh, J. E., Chapman, W. L., Romanovsky, V., Christensen, J. H., Stendel, M., Global climate model performance over Alaska and Greenland, Journal of Climate, 21, 6156-6174, 2008

Wigley, T. M. L, Raper S. C. B., Interpretation of high projections for global-mean Warming, Science, 293, 451-454, 2001

Wofsy, S. C., et al., Atmospheric chemistry in the Arctic and sub-Arctic – influence of natural fires, industrial emissions, and stratospheric inputs, Journal of Geophysical Research – Atmospheres, 97, 16731-16746, 1992

Yussouf, N, Stensrud, D., Bias-corrected short-range ensemble forecasts of near-surface variables during the 2005/06 cool season, Weather and Forecasting, 22, 1274-1286, 2007

Zhou, S., Miller, A. J., Wang, J., Angell, J. K., Downward-propagating temperature anomalies in the preconditions polar stratosphere, Journal of Climate, 15, 781-792, 2002

Table 1
List of sites used in this study, their locations, elevations above MSL, and distances from the nearest forecast model grid point. Locations north of the Arctic Circle are shaded.

			Distance (km)	
Site name	**Location**	**Elevation (m)**	**GFS**	**WRF**
Barrow, Alaska	71.30° N, 156.77° W	4	23.7	12.3
Cambridge Bay, Nunavut, Canada	69.12° N, 105.06° W	26	8.2	3.6
Resolute, Nunavut, Canada	74.69° N, 94.97° W	40	6.5	9.1
Aasiaat, Greenland	68.69° N, 52.84° W	41	19.3	26.2
Summit, Greenland	72.57° N, 38.48° W	3200	3.2	30.1
The Pas, Manitoba, Canada	53.97° N, 101.09° W	274	10.2	7.4
Gray, Maine	43.88° N, 70.25° W	118	29.4	9.1
Flagstaff, Arizona	35.22° N, 111.81° W	2181	29.0	2.9

Table 2
Selected characteristics of the two models examined in this study.

	GFS	**WRF**
Grid point spacing	~ .5° Lat X .5° Lon	45 km X 45 km
Grid type	Gaussian (T254)	Polar Stereographic
Vertical representation	Sigma	Eta
# Vertical layers	64 (non-linear)	50 (non-linear)
Initialization frequency	every 6 hours	every 24 hours
Forecast timestep	3 hours	1 hour

Table 3

Statistical data for GFS wind speed verification at each rawinsonde site and for Summit including and excluding the outlier. Mean wind speed is in m s^{-1}. RMSE, the linear regression equation, and corresponding R^2 are all listed for 00, 24, and 72-hours.

Site	**Mean**	**00h RMSE**	**24h**	**72h**	**00h linear**	**00h R^2**	**24h**	**24h**	**72h**	**72h**
Barrow	8.84	2.10	2.40	3.35	0.89*GFS + 1.22	0.78	0.84*GFS + 1.31	0.71	0.61*GFS + 3.59	0.44
Cambridge Bay	11.89	0.98	2.41	3.91	1.01*GFS - 0.04	0.96	1.05*GFS + 0.07	0.74	0.63*GFS + 5.60	0.32
Resolute	10.25	1.49	2.57	3.48	0.92*GFS + 1.14	0.92	0.86*GFS + 2.95	0.77	0.64*GFS + 3.78	0.54
Aasiaat	9.69	1.37	2.10	3.03	0.95*GFS + 0.62	0.88	0.74*GFS + 2.57	0.71	0.57*GFS + 4.51	0.40
Summit (outlier)	6.29	3.31	3.32	2.95	0.44*GFS + 3.35	0.10	0.40*GFS + 3.63	0.10	0.63*GFS + 2.18	0.29
Summit (no outlier)	5.84	2.21	2.24	2.18	0.59*GFS + 1.84	0.30	0.54*GFS + 2.23	0.29	0.52*GFS + 2.47	0.34
The Pas	9.79	1.37	2.72	3.58	0.96*GFS + 0.07	0.91	0.88*GFS + 0.98	0.64	0.66*GFS + 3.33	0.38
Gray	13.65	1.92	2.64	3.67	0.87*GFS + 2.52	0.84	0.81*GFS + 2.72	0.70	0.64*GFS + 4.53	0.42
Flagstaff	9.23	1.89	1.99	2.83	0.83*GFS + 1.85	0.89	0.92*GFS + 1.52	0.88	1.02*GFS - 0.23	0.77

Table 4
The R^2 values from linear regressions of observed wind speed at 650 hPa versus GFS and WRF output at 00 and 24-hour forecast times.

	00h GFS	00h WRF	24h GFS	24h WRF
Barrow, AK	0.73	0.70	0.60	0.72
Cambridge Bay, NT	0.97	0.96	0.75	0.95
Resolute, NT	0.95	0.94	0.77	0.93
Aasiaat, Greenland	0.90	0.91	0.71	0.85
Summit, Greenland	0.07	0.09	0.02	0.06
Summit (no outlier)	0.35	0.36	0.23	0.28
The Pas, MB	0.90	0.89	0.64	0.88
Gray, ME	0.88	0.86	0.76	0.87

Table 5
RMSE/Mean observed wind speed ratio for observation sites at 650 hPa. (Flagstaff was omitted due to having only 5 or 6 observations for each time step while other sites had more than 10). This ratio shows the magnitude of the forecast error at 00h and 24h compared to the mean observed wind speed at each site. Gray columns show the difference between the two models, with a positive number indicating that the WRF forecast was more accurate.

650 hPa	00h GFS	00h WRF	00h GFS-WRF	24h GFS	24h WRF	24h GFS-WRF
Barrow, AK	0.240	0.320	-0.079	0.276	0.314	-0.038
Cambridge Bay, NT	0.086	0.097	-0.011	0.196	0.106	0.091
Resolute, NT	0.141	0.105	0.036	0.243	0.112	0.131
Aasiaat, Greenland	0.234	0.119	0.115	0.265	0.154	0.111
Summit, Greenland	0.497	0.554	-0.058	0.498	0.524	-0.026
Summit excluding outlier	0.357	0.377	-0.020	0.359	0.358	0.002
The Pas, MB	0.149	0.171	-0.022	0.278	0.172	0.106
Gray, ME	0.150	0.125	0.025	0.204	0.122	0.083

Table 6
RMSE/Mean observed wind speed for all sites except Summit at 500 hPa. This ratio shows the magnitude of the forecast error at 00h and 24h compared to the mean observed wind speed at each site. Gray columns show the difference between the two models, with a positive number indicating that the WRF forecast was more accurate.

	00h		00h	24h		24h
500 hPa	GFS	WRF	GFS-WRF	GFS	WRF	GFS-WRF
Barrow, AK	0.125	0.150	-0.025	0.234	0.147	0.087
Cambridge Bay, NT	0.130	0.155	-0.025	0.249	0.135	0.114
Resolute, NT	0.168	0.187	-0.019	0.265	0.187	0.078
Aasiaat, Greenland	0.127	0.184	-0.057	0.265	0.192	0.073
The Pas, MB	0.382	0.172	0.210	0.404	0.182	0.222
Gray, ME	0.116	0.117	-0.001	0.186	0.126	0.060
Flagstaff, AZ	0.193	0.222	-0.029	0.284	0.247	0.037

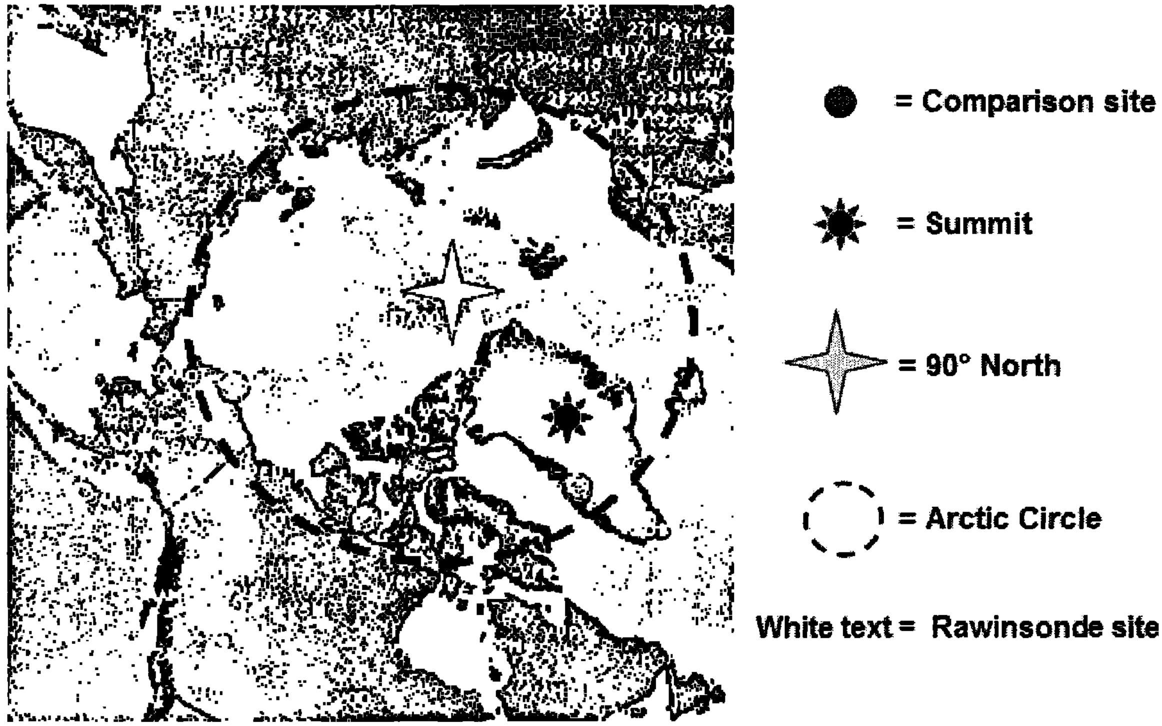

Figure 1
View of the atmospheric sounding stations in and around the Arctic. Alphanumeric character strings denote the location of rawinsonde sites, with circles marking sites located upstream from Summit and near the same latitude.

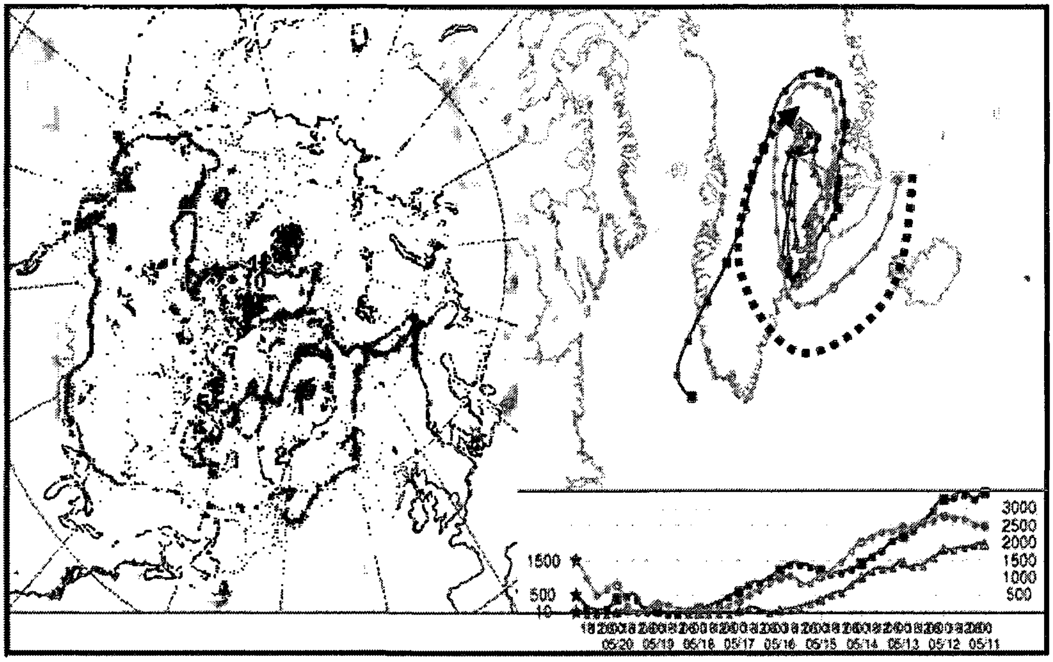

Figure 2
Two sample transport plots, with air reaching Summit on May 21, 2007 at 0000 GMT and spanning a period of 10 days. On the left side is a FLEXPART retroplume, with the broken white line indicating the approximate shape of the plume's path from its source region towards ground level at Summit. On the right is a HySplit back-trajectory

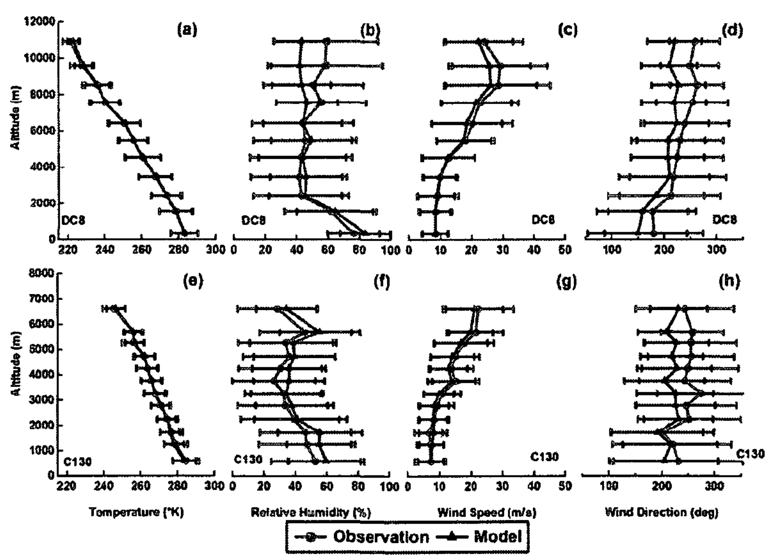

Figure 3
Comparison between meteorological parameters as measured from two different airplanes over the Pacific Ocean near the west coast of the United States (in red) and concurrent WRF model output (in blue). The standard deviation of each measurement is represented by the error bars. (Figure from Adhikary et al., 2009)

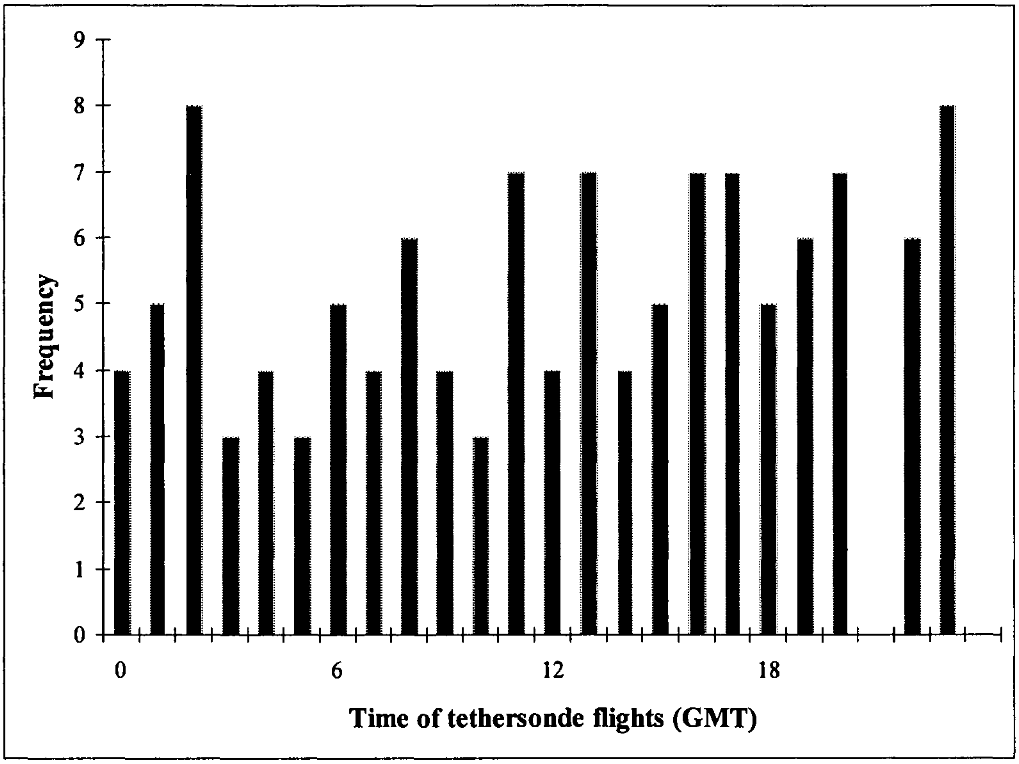

Figure 4
Distribution of all tethersonde flights at Summit that reached a minimum pressure below 650 hPa. The 122 flights occurred between 6 June and 6 July 2008, but were not evenly distributed during the field campaign.

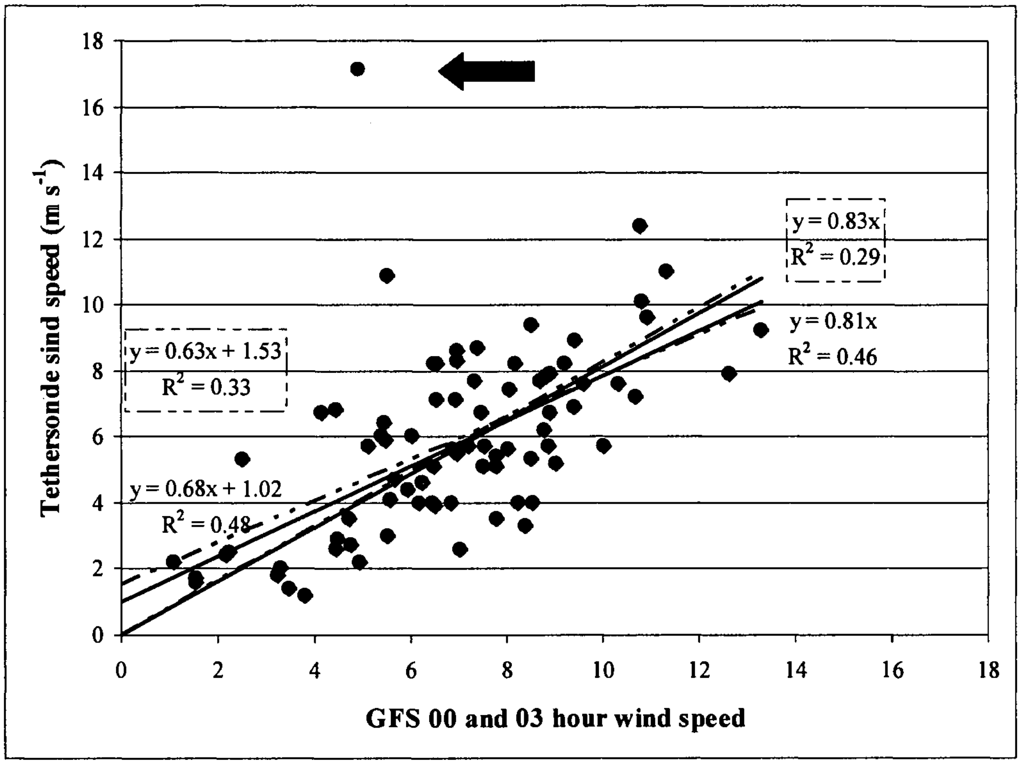

Figure 5
GFS output verifications for 00h and 03h plotted against all valid 650 hPa wind speed observations at Summit, with the gray arrow indicating the outlier. The regression statistics on the right(left) side are for lines forced(not forced) through the intercept. Statistics in the dashed boxes are for the dashed lines, which are regressions through the dataset including the outlier.

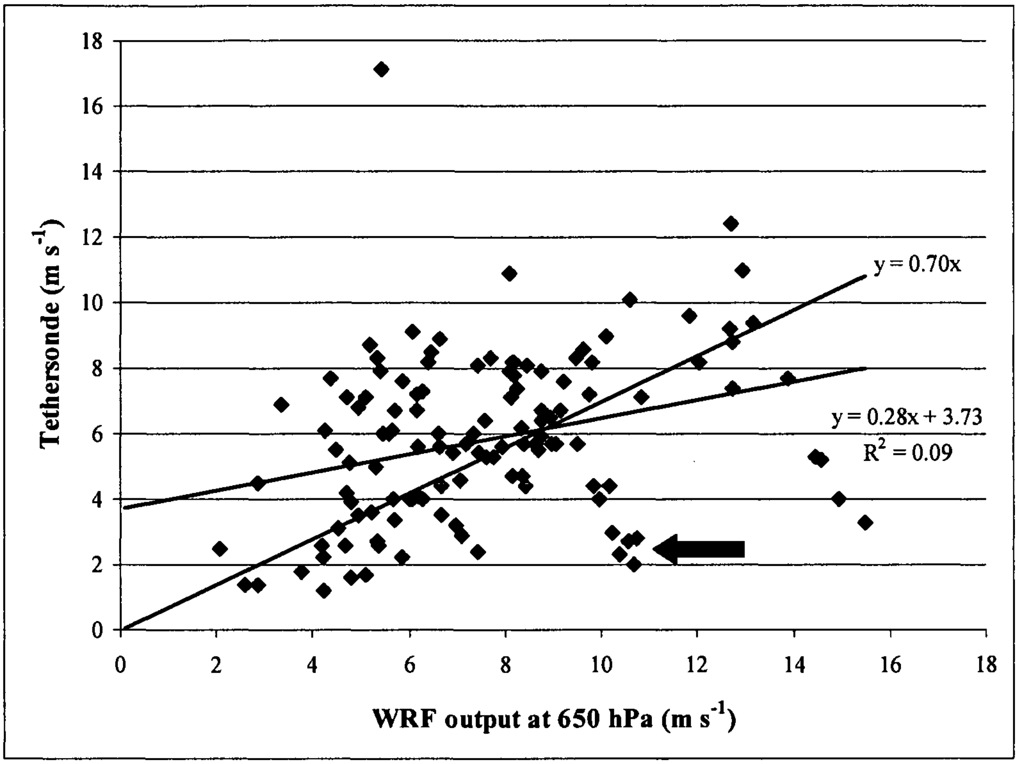

Figure 6
WRF output plotted against observed wind speeds at 650 hPa, including the high wind speed outlier, and covering all WRF forecast times from 00 - 23 hours. The gray arrow indicates a cluster of poor forecasts that occurred during one 6-hour time period.

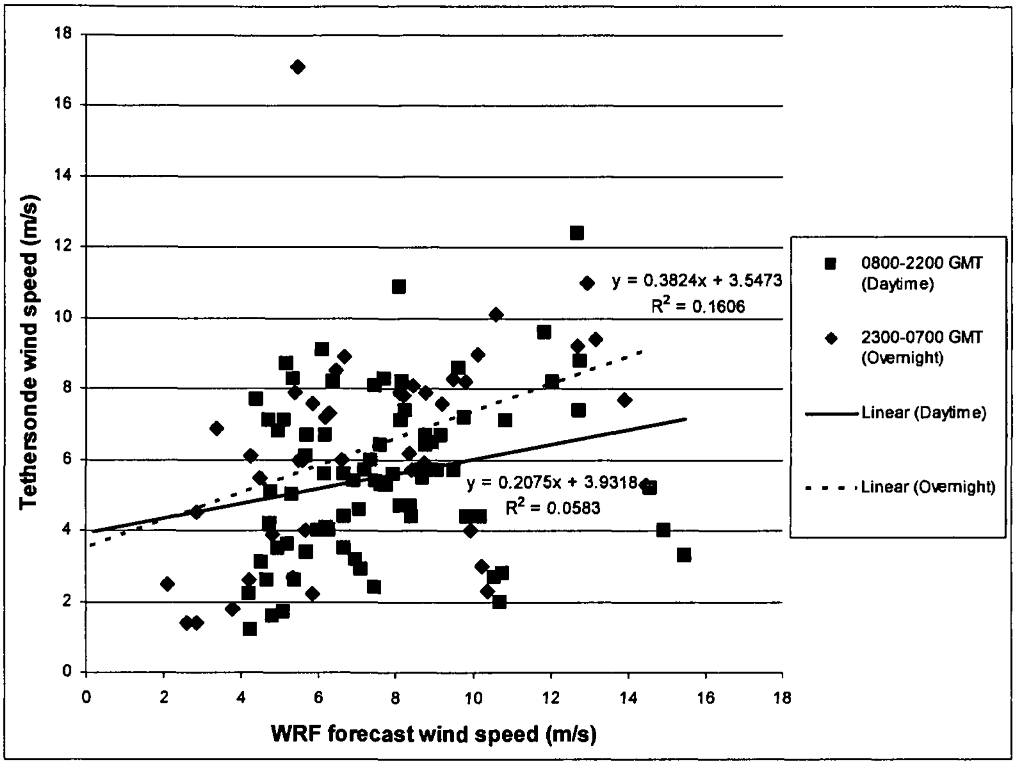

Figure 7
A comparison of "Daytime" and "Overnight" tethersonde observations to WRF model data, including the outlier point. The day period runs from 6am – 8pm local time, starting after the nocturnal inversion began to lift in the morning and the surface temperature began to increase.

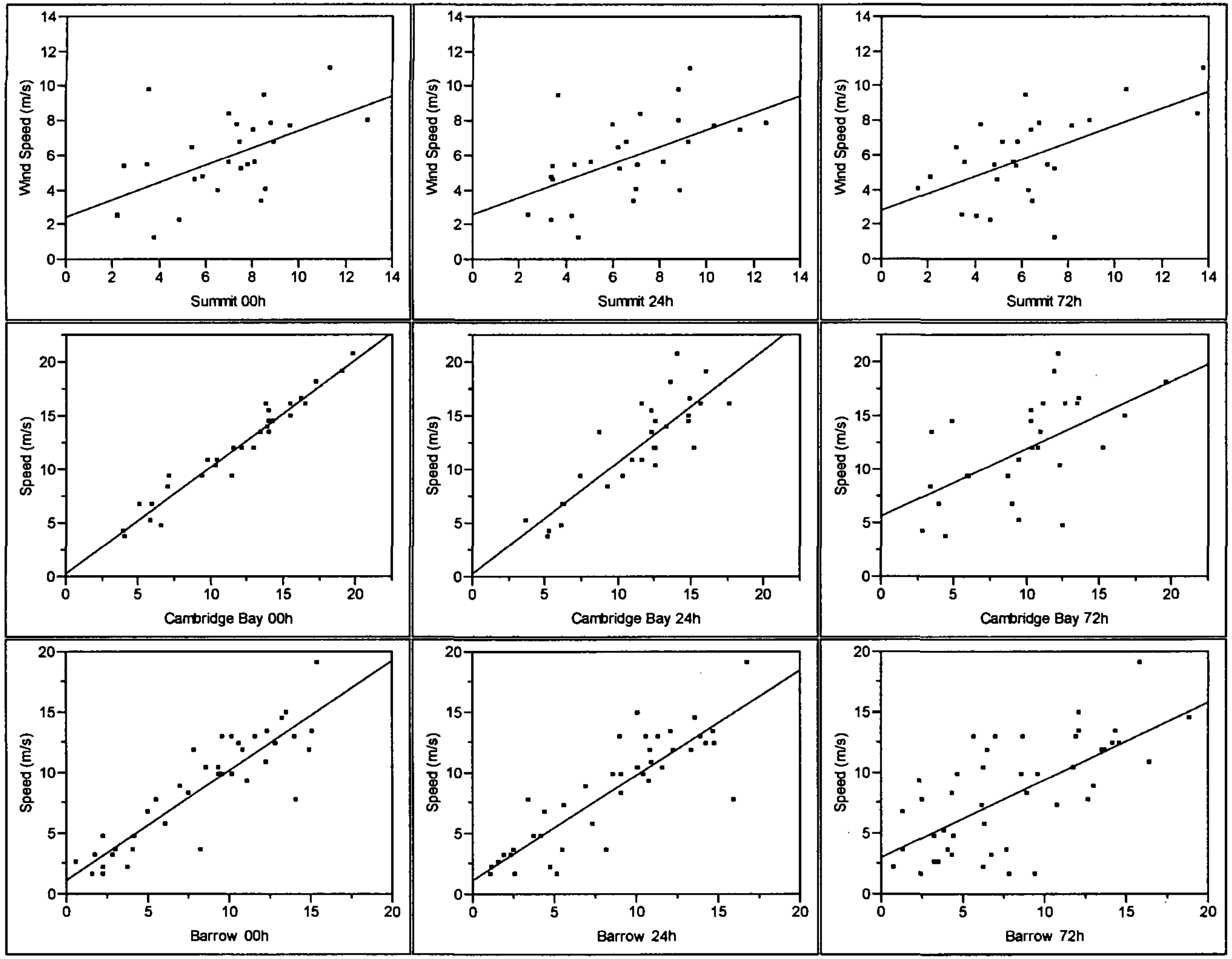

Figure 8

Linear regressions of 650mb wind observations versus 00Z and 12Z GFS output for the 00h, 24h, and 72h forecast time periods from 6 June – 9 July 2008. R^2 values for each regression in parentheses: Summit, no outlier (0.30, 0.29, 0.34), Cambridge Bay (0.96, 0.74, 0.32), and Barrow (0.78, 0.71, 0.44)

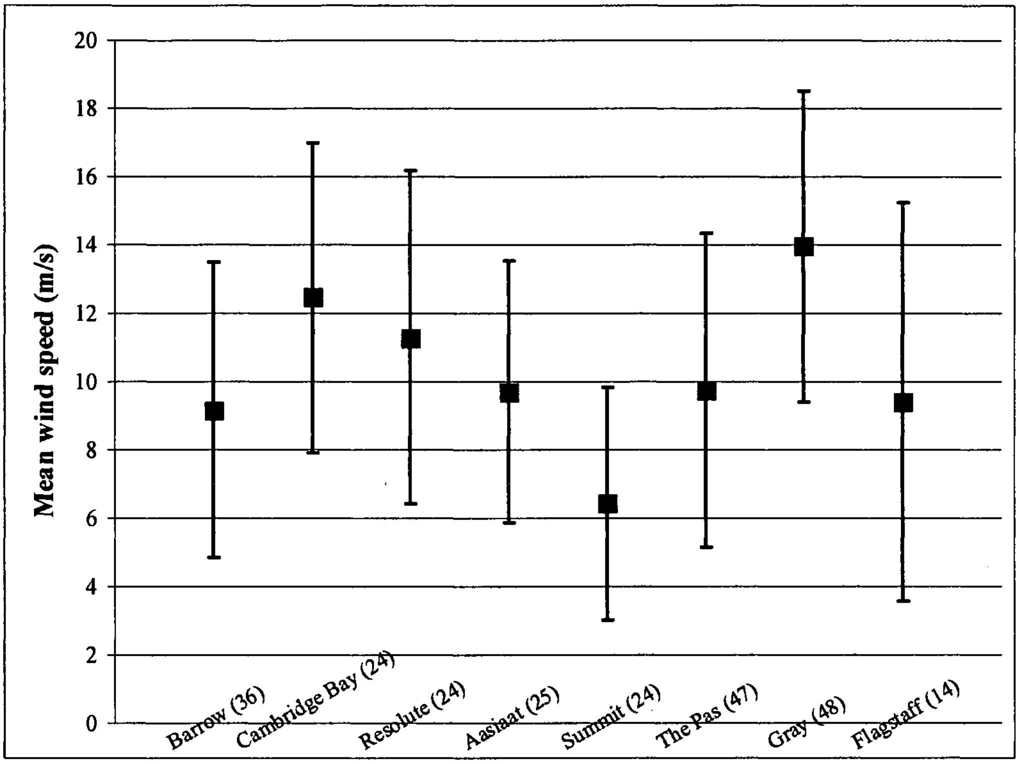

Figure 9
Mean 650 hPa wind speed observed daily at 0000 and 1200 GMT from 0000 6 June through 0000 6 July 2008 for all sites except Summit, where data are from 650 hPa. Error bars indicate one standard deviation from the mean. The number in parentheses is the number of observations that went into determining the mean.

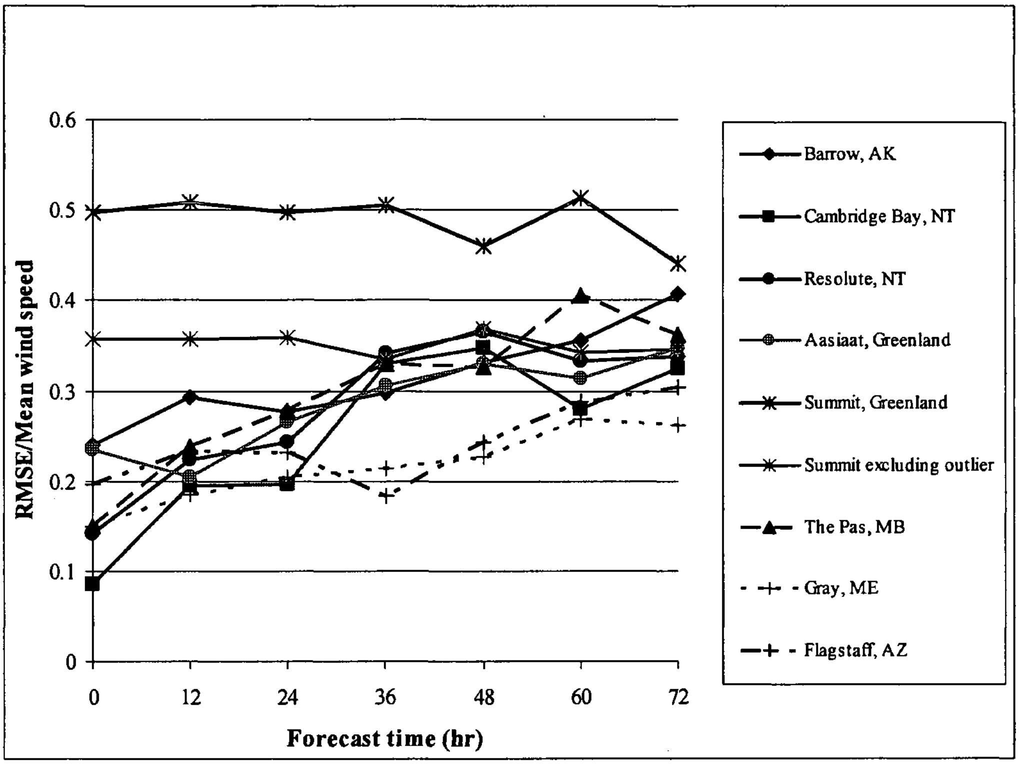

Figure 10
Ratio of RMSE to mean observed wind speed at 650 hPa for all sites. Results are calculated for 12-hour intervals using output from the 00Z and 12Z runs of the GFS.

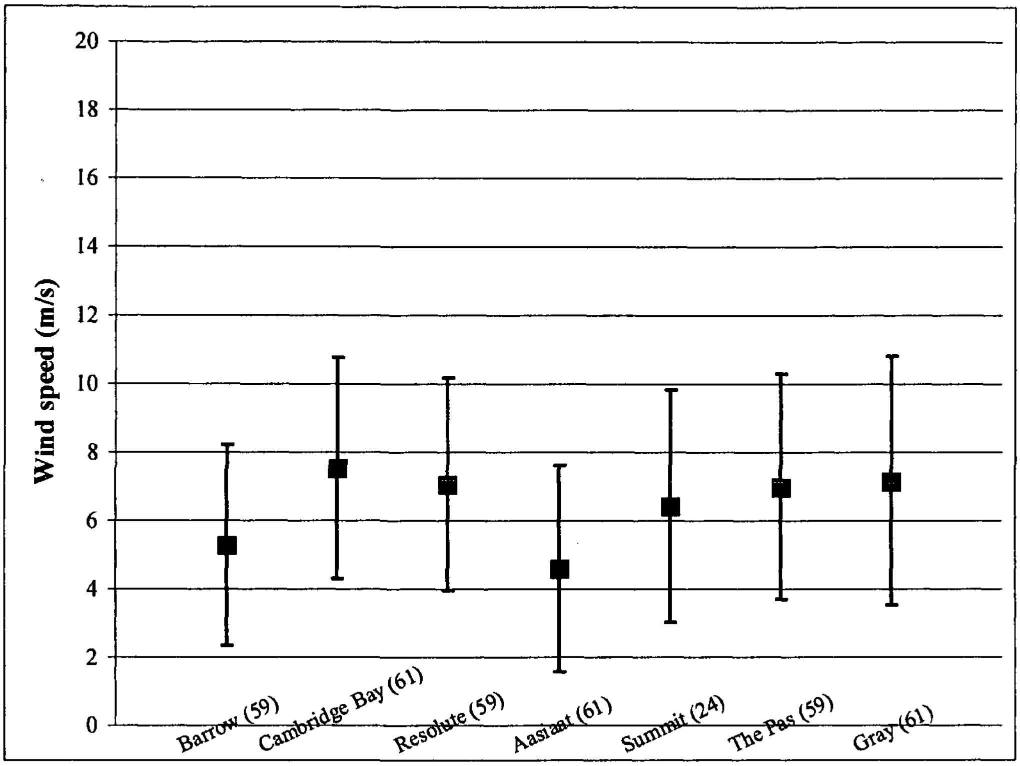

Figure 11
Mean 925 hPa wind speed observed daily at 0000 and 1200 GMT from 0000 6 June through 0000 6 July 2008 for all sites except Summit, where data are from 650 hPa. Error bars indicate one standard deviation from the mean. The number in parentheses is the number of observations that went into determining the mean.

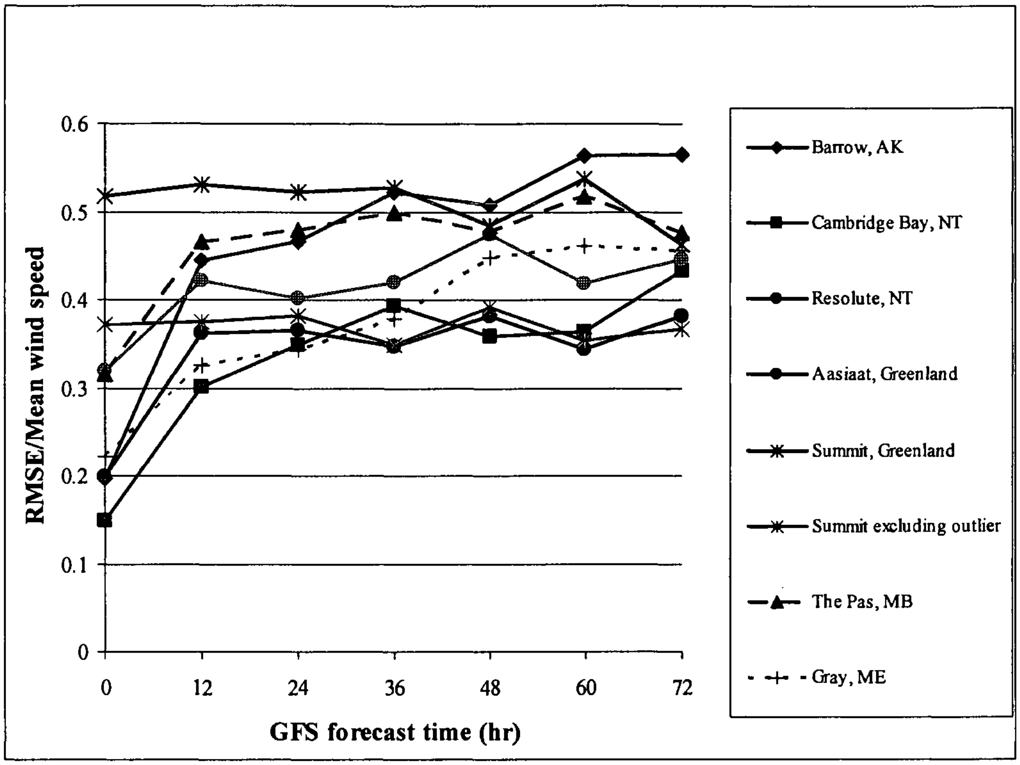

Figure 12
Comparison of model performance near the height of the planetary boundary layer from 00Z 6 June through 00Z 9 July 2008. Ratio of RMSE to mean observed wind speed at 925 hPa for all sites rawinsonde sites below this pressure level. Data from Summit at 650 hPa is also shown. Points are spaced every 12 hours, coinciding with the timing of rawinsonde launches.

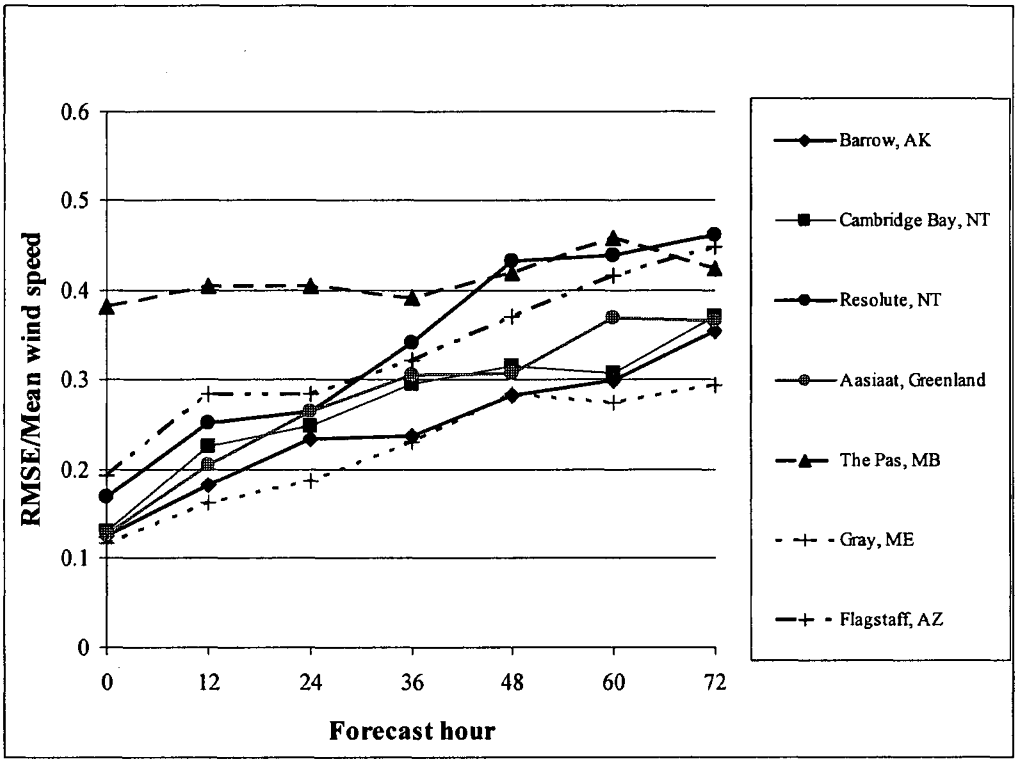

Figure 13
Same as Figure 12, except at 500 hPa.