Statistical Performance of Forecasting Models Across the North Slope of Alaska During the Mixed-Phase Arctic Cloud Experiment

Victor Yannuzzi, Eugene Clothiaux, Johannes Verlinde and Jerry Harrington
The Pennsylvania State University

Introduction

The Mixed-Phase Arctic Cloud Experiment (MPACE) was conducted across the North Slope of Alaska (NSA) from September 27 through October 22, 2004. The experiment was funded by the Department of Energy (DOE)-Atmospheric Radiation Measurement (ARM) program to study the microphysics, radiative properties, thermodynamics, and life cycle of mixed-phase clouds in the Arctic. A focused set of observation stations and soundings were in place across the NSA (figure 1) where conditions could be expected to produce low-level mixed-phase clouds on a consistent basis. Barrow and Oliktok Point were chosen as coastal locations while Atqasuk and Toolik Lake were selected as inland stations.

Figure 1. The MPACE experimental sites.

Observational Data

Vaisala RS-92 radiosondes were launched every six hours at the four sites along the NSA during two Intensive Observation Periods covering October 4-9 and 14-22, 2004. Outside this period soundings were launched sporadically between September 26-October 3 and October 9-14, 2004. High spectral resolution lidar provided by the University of Wisconsin and depolarization lidar provided by the University of Alaska Fairbanks were used in conjunction with Millimeter Cloud Radar (MMCR) and satellite data from the Moderate Resolution Imaging Spectroradiometer (MODIS) to gather information about cloud height, thickness, and radiative properties at Barrow, Alaska. Surface downwelling shortwave, longwave, and total radiation data were available at Barrow and Atqasuk, where outputs were generated every 60 seconds.

Model Output

Model outputs from the National Centers for Environmental Prediction (NCEP) Eta model, the European Center for Medium-Range Weather Forecasts (ECMWF) model, the NASA-Global Modeling and Assimilation Office (GMAO) model, and the Regional Atmospheric Modeling System (RAMS) model each produced data that were tested across the NSA during MPACE. Table 1 shows the specifics for each model.

	Eta	Fine-Resolution ECMWF	Coarse-Resolution ECMWF	GMAO	RAMS
Forecast Time Range	0-84 hours	0-18 hours	0-84 hours	0, 24-120 hours	0-48 hours
Forecast Time Increment	3 hours	6 hours	3 hours	12 hours	3 hours
Horizontal Resolution	40 km x 40 km	0.5° lat x 0.5° lon	0.5° lat x 0.5° lon	1.0° lat x 1.25° lon	4 km x 4 km
Vertical Resolution	25 hPa from 1000-50 hPa	15 levels from 1000 to 10 hPa	7 Levels from 1000 to 100 hPa	26 levels from 1000-100 hPa	11 Levels from 1000 to 100 hPa

Table 1. Model Information pertaining to the models used in the intercomparisons during MPACE.

Methodology

Model outputs from the Eta, ECMWF, GMAO, and RAMS models were compared with sounding data and surface observations during MPACE. Temperature, relative humidity, and u- and v-wind components were tested for all models. Surface downwelling shortwave and longwave radiation, and cloud fraction were tested for the Eta and RAMS models. The test the statistical significance of the model mean bias errors, we used the bootstrap technique discussed in Efron and Tibshirani (1991), Wilks (1997), and Marchand et al. (2006). This test has the unique advantage over other statistical tests such as the t-test in that the datasets to be compared do not have to have a normal distribution or have correlations in time and space that are present in meteorological data.

MPACE Synoptic Events

In examining the MPACE sounding data in conjunction with surface observations, radar, lidar, and satellite imagery we have identified three distinct synoptic regimes during MPACE. Figure 2 shows Eta analyses of mean sea level pressure, temperature, and wind fields from each regime. Synoptic regime I (Sept 24-Oct 1) was marked by unsettled conditions due to the presence of a deep trough aloft that steered several shortwave systems into the NSA (left panel). Synoptic Regime II (Oct 4-14) saw the NSA being under the domain of a large upper level ridge and strong surface high that built in over the rapidly cooling pack-ice (center panel). This kept disturbances out of the area and made for persistent low-level ENE onshore flow that brought in moisture and produced boundary layer stratus clouds throughout the period. Eventually the high moved southeast into Canada, allowing distubances to influence the NSA and ushering in synoptic regime III (Oct 18-23). During this regime (right panel) the NSA was under the influence of a strong and fast developing system that recorded a 42 hPa central pressure drop in 24 hours and peaked at 940 hPa in intensity.

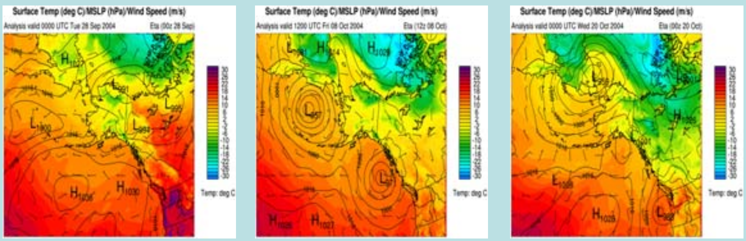

Figure 2. Eta analyses at 0000 UTC on September 28, 2004 (left panel), 1200 UTC on October 8, 2004 (center panel), and 0000 UTC on October 20, 2004 (right panel). Sea-level pressure (contours), temperature (shading), and wind speed and direction (wind barbs) are shown.

Results-Temperature, Moisture, and Wind

Figures 3-6 show the statistically significant mean bias errors at the 80% confidence intervals (CI's) for temperature, relative humidity, u-wind component, and v-wind component for all of the models at each of the four MPACE sites. There are four vertical columns representing each site on the NSA and five horizontal rows representing each specific model. Each individual graph has three columns representing statistically significant errors that occurred in synoptic regime II (left), synoptic regime III (center), and overall during MPACE (right). Synoptic regime I was left out of specific regime comparisons since only 3 soundings were launched during this period. Red bars denote statistically significant positive mean bias errors while blue bars denote statistically significant negative mean bias errors. Green bars indicate areas where models had varied biases depending on forecast hour. The width of the bars is proportional to the magnitude of the mean bias errors. Sample bars are given for each variable tested. An error bar twice as wide as the sample bar represents an error of twice the magnitude. Overall temperature forecasts for most models were good with most mean bias errors under 2.0 K in magnitude and few significant errors (figure 3). Forecasting moisture proved to be most difficult for the models . Significant moist biases can be seen in the mid to upper-troposphere in Eta, fine-resolution ECMWF, GMAO, and RAMS models (figure 4). For the GMAO and RAMS models the moist bias is as high as 40% or more. For wind fields (figures 5 and 6) the GMAO underperforms overall, with the Eta also having several significant errors in the low- to mid-troposphere.

Temperature

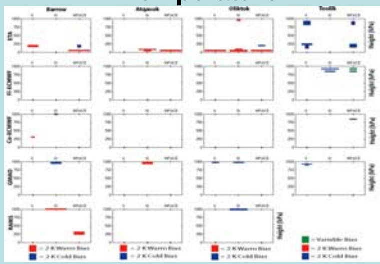

Figure 3. Statistically significant (80% CI's) temperature mean bias errors for the Eta, ECMWF, GMAO, and RAMS models at each of the four MPACE sites. Sample bars represent biases of 2 K.

Relative Humidity

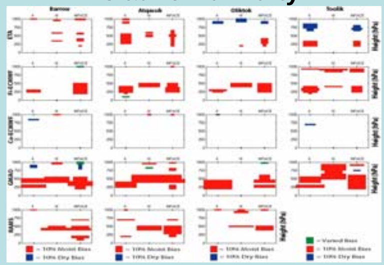

Figure 4. Statistically significant (80% CI's) relative humidity mean bias errors for the Eta, ECMWF, GMAO, and RAMS models at each of the four MPACE sites. Sample bars represent biases of 10%.

U-Wind

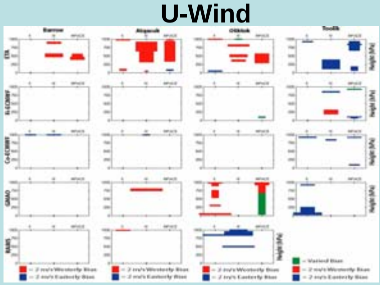

Figure 5. Statistically significant (80% CI's) u-wind component mean bias errors for the Eta, ECMWF, GMAO, and RAMS models at each of the four MPACE sites. Sample bars represent biases of 2 m/s.

V-Wind

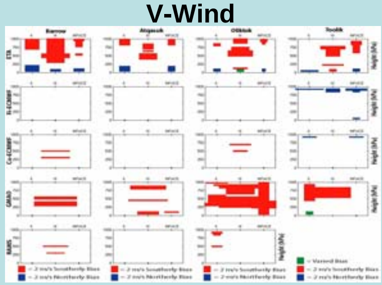

Figure 6. Statistically significant (80% CI's) v-wind component mean bias errors for the Eta, ECMWF, GMAO, and RAMS models at each of the four MPACE sites. Sample bars represent biases of 2 m/s.

Results-Radiation

Figure 7 shows the mean bias errors for shortwave, longwave, and total downwelling surface radiation at Barrow (□) and Atqasuk (°) for the Eta (left panel) and RAMS (right panel) models. Overall, the Eta significantly overestimated shortwave radiation and underestimated longwave radiation at Barrow and Atqasuk. These biases canceled each other out completely at Barrow and left a small but significant negative total radiation bias at Atqasuk. The RAMS had a variable shortwave bias, but underestimated longwave radiation and total radiation at both sites. Errors for both models tended to be larger at Atqasuk for longwave radiation.

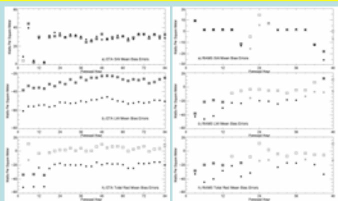

Figure 7. Mean bias errors at Barrow (□) and Atqasuk (°) for the Eta (left) and RAMS (right) models during MPACE. Dark symbols represent significant differences at the 80% CI.

Results-Cloud Fields

Figure 8 shows a time series of cloud fraction throughout the MPACE experiment for the Eta and RAMS models along with the observed cloud fraction. Eta and RAMS both significantly underestimated cloud fraction by 11.8% and 9.5%, respectively. However, RAMS statistically outperformed Eta during the middle of MPACE (synoptic regime II) where it underestimated cloud cover by only 4.5%, versus 12.2% for the Eta.

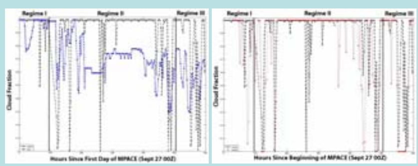

Figure 8. Time series of cloud fraction during MPACE for the Eta (left panel) and RAMS (right panel) models.

References

-Efron, B., and R. Tibshirani, 1991. Statistical data analysis in the computer age. *Science*., **253**, 390-395.
-Marchand, R. et al., 2006. A Bootstrap Technique for Testing the Relationship between Local-Scale Radar Observations of Cloud Occurrence and Large-Scale Atmospheric Fields. *J. Atmos. Sci*., **63**, 2813-2830.
-Wilks, D. S., 1997. Resampling Hypothesis Tests for AutoCorrelated Fields. *J. Climate*., **10**, 65-82.