P2.13 EVALUATING WINTER SEASON PRECIPITATION TYPE IN THE BALITIMORE/WASHINGTON NWSFO REGION

Michelle Farver* and Gregory Jenkins
Howard University Program in Atmospheric Sciences

1. INTRODUCTION

There is considerable value, both economically and for human welfare, on accurately forecasting precipitation type (P-TYPE) for a large metropolitan area during the winter season. Under-forecasting these events can lead to unexpected conditions causing loss of life and/or destruction of property over-forecasting can lead to unnecessary school, government, and business closings causing superfluous economic losses; and inaccurate forecasts of P-TYPE leaves the public with reason to lose confidence in meteorological forecasts. The intent of this study is to quantify uncertainties associated with P-TYPE during winter weather events. The use of the acronym P-TYPE in this paper refer to the type of precipitation observed at surface level and more specifically refers to freezing rain, sleet, and snow. The reference to a winter season in this study has a starting date of December 1st and an ending date of March 31st. The three station used in this study are Reagan National Airport (DCA), Thurgood Marshall Baltimore Washington International Airport (BWI), and Washington Dulles International Airport (IAD.) Observed data from these stations have been collected from the 2003/2004 and 2004/2005 winter seasons and is compared with model data from the same periods.

2. METHODOLOGY

Identifying winter weather events for this study started with referring to the Preliminary Local Climatological Data (PLCD) form (F-6.) Observed data is taken directly from the site and is collected at each of the previously mentioned stations, including some P-TYPE and amounts. Any days exhibiting traces of frozen precipitation, were noted and considered further before being classified as an event. If frozen precipitation was observed for a single day, meaning the previous and the following days did not report frozen precipitation, that single day alone was documented as an event. If two or more consecutive days were documented as having frozen precipitation, those set of days were considered a single event. Also, to eliminate insignificant occurrences, frozen precipitation must have been observed at two or more stations. There were a total of 19 events identified for both winter seasons by this processes.

After identifying specific dates for events, it was necessary to pinpoint initial starting times for each event. This process was done by locating National Climatic Data Center (NCDC) hourly observations, searching through the start dates of each event, and appropriately documenting times, making special effort to rounding the times to the nearest hour if necessary. This set of initial start times are referred to as IST. From these observations, we conducted a separate documentation for the event times in order to comply with the forecasting model runs. These times are referred to as RST in this study. Each events initial P-TYPE start time was rounded to the nearest 6th hour interval on a 24 hours time span. For example, if the initial starting time was 05Z the time was rounded up to 06Z; similarly, if the initial starting time was 20Z the time was rounded down to 18Z. At this point, we have two sets of event starting times; IST, a documentation of initial start times; and RST, the initial start times rounded to comply with the model runs.

Next, archived model predictions were identified for both ETA and GFS models, this process involves the RST set of start times. The ETA model is used to compare events for periods 24 hours prior to RST of an event. The GFS model is used to compare all events for periods 24 hours and 48 hours prior RST of an event. Expediting this task encouraged an analysis software tool, BUFKIT, be used. We used BUFKIT to look at P-TYPE predictions for each model. The RST P-TYPE analysis for each model was compared to the IST observations.

Finally, contingency tables were used to analyze the different P-TYPE events. Hit rate (H), threat score (TS), probability of detection (POD), false alarm rate (FAR), and BIAS value were computed for each event, at all stations, for all three model

---

**Corresponding author address*: Michelle Farver Howard University, Program in Atmospheric Science, Washington, DC 20059; e-mail: mfarver@howard.edu.

scenarios. A forecast was considered to be successful if it occurred within a window of 6 hours of the IST (+/- 3 hours.) The Hit rate and TS are emphasized because these account for the greatest number of frozen events, while the TS and POD are emphasized for sleet and freezing rain because the low frequency of these events.

3. RESULTS

3.1 Snow Event Analysis

The GFS model tended to have the best forecasts for snow events for the two winter seasons between both model scenarios (24 and 48). These two model scenarios produce good hit rate (Table 1a), threat score (Table 1b), and BIAS values (not shown). The GFS 48 averages 0.89 and the GFS 24 averages 0.88 from a best possible value of 1 for TS. The GFS 48 averages 0.94 from a desired score of 1 for BIAS. The GFS 48 also produced some good BIAS values, but proved some small potential of over-forecasting at times.

When compared to the GFS forecasts, the ETA 24 model scenario did not perform as well for hit rates or threat scores (Tables 1a and 1b, respectively) when forecasting snow events for the two winter seasons. Hit rates for this model were low averaging 60% while the threat scores (0.57) were substantially lower than the GFS forecasts. Moreover, the ETA 24 had the tendency to under-forecast snow events with a

	DCA	BWI	IAD	
GFS 24	79	83	89	*HIT*
ETA 24	58	58	63	*RATES*
GFS 48	74	78	95	

Table 1a. Winter season 2003/2004 and 2004/2005 snow hit rate values

	DCA	BWI	IAD	
GFS 24	0.78	0.83	0.88	*THREAT*
ETA 24	0.53	0.58	0.59	*SCORE*
GFS 48	0.71	0.78	0.94	

Table 1b. Winter season 2003/2004 and 2004/2005 snow TS values

BIAS of 0.57 (not shown). Overall forecasts performed better for IAD while the forecast were poor for DCA and a low amount of confirmed "yes" and "no" forecasts. The POD also produced poor scores for snow events averaging 0.57. This model proved to show significant amounts of under-forecasting with its BIAS averaging 0.57 for snow events.

3.2 Sleet Event Analysis

All of the model (GFS 24, GFS 48 and ETA 24) forecasts produced low threat scores for the three locations (Table 2a). The lowest threat scores were found for IAD followed by DCA and BWI. The GFS

	DCA	BWI	IAD	
GFS 24	0.5	0.17	0	*PROBABILITY*
ETA 24	0.71	0.8	0.5	*OF*
GFS 48	0.67	0.5	0.43	*DETECTION*

Table 2a. Winter season 2003/2004 and 2004/2005 sleet POD values

	DCA	BWI	IAD	
GFS 24	0.33	0.13	0	*THREAT*
ETA 24	0.22	0.5	0.38	*SCORE*
GFS 48	0.44	0.44	0.38	

Table 2b. Winter season 2003/2004 and 2004/2005 sleet TS values

4 forecasts produced the lowest threat scores, while the GFS 48 produced the highest threat score. The ETA model however, had the highest scores when considering the POD (Table 2b). The ETA model was not biased and produced a FAR value of 0.50 (not shown). The GFS 24 POD scores were poor for all locations for sleet forecasts.

3.3 Freezing Rain Analysis

Similar to sleet forecasts, all of the models (GFS 24, GFS 48 and ETA 24) forecasts produced low threat scores for the three locations (Table 3a). However, the ETA 24 produced the highest TS scores relative to GFS 24 and GFS 48 models. DCA had the lowest TS values while IAD produced

	DCA	BWI	IAD	
GFS 24	0.33	0.33	0.67	*PROBABILITY*
ETA 24	1	1	1	*OF*
GFS 48	0.33	0.5	0.8	*DETECTION*

Table 3a. Winter season 2003/2004 and 2004/2005 freezing rain POD values

	DCA	BWI	IAD	
GFS 24	0.25	0.25	0.67	*THREAT*
ETA 24	0.57	0.67	0.86	*SCORE*
GFS 48	0.2	0.4	0.57	

Table 3b. Winter season 2003/2004 and 2004/2005 freezing rain TS values

the highest TS values. ETA 24 produced a perfect POD value for all three locations (Table 3b) but this was due to a consistent over forecasting of freezing rain events at all locations (high bias – not shown). Both the GFS 24 and GFS 48 forecasts of freezing rain POD values for DCA and BWI were low but significantly better at IAD (Table 3b). There was a tendency for the GFS 24 forecast to under-predict (low bias) freezing rain at each of the locations (not shown).

5. CONCLUSION

In this study we have examined snow, sleet and freezing rain events for the winter seasons of 2003/2004 and 2004/2005. As expected the models show the best forecast skill for snow events and the worst forecast skills for sleet and freezing rain. The skill for freezing rain is low even for 24-hour forecasts at BWI and DCA airports. Since freezing rain is a significant weather maker because of it ability to disrupt automobile and aircraft traffic we believe that it is necessary to examine a larger number of these events to verify the results of this study. The next step involved in this process is to examine observed and forecasted soundings at IAD and the specific synoptic event in order to determine why the skill of freezing rain and sleet events were poor relative to snow events at the three locations. This should help us to identify the specific biases in each of the models associated with the low skill scores.

6. REFERENCES

Wilks, Daniel S., 1995: Statistical Methods in the atmospheric sciences: an introduction" Academic Press San Diego, California

Yucel, Ismail et el, 1998 : Evaluating NCEP Eta Derived Data against Observations. Monthly Weather Review, Vol 126, No.7, 1977-1991

Black, Thomas L., 1994: The New NMC Mesoscale Eta Model: Description and Forecasting examples. Weather and Forecasting, Vol.9, No.2, 265-278