Verification of the National Oceanic and Atmospheric Administration's $PM_{2.5}$ Forecast for the Shenandoah Valley

Ann G. Albrecht
Center for Energy and Environmental Sustainability, James Madison University
14 Virginia Ave, Staunton, VA 24401

1 ABSTRACT

Since air quality is so important to human health, there is a great deal of interest in being able to forecast it so as to alert the public to periods of poor air quality. The National Oceanic and Atmospheric Administration's /National Center for Environmental Prediction's (NOAA/NCEP) air quality forecast was developed with this use in mind. While such models can be highly useful, they are only so when the forecasts being generated are reasonably accurate and reliable. To this end, models need to be verified by comparing their outputs with real-world data. The objective of this project is to evaluate the accuracy of the NOAA/NCEP air quality forecast model for the Shenandoah Valley region of Virginia. While the NCEP model forecasts concentrations of various air pollutants, our analysis is focused only on fine particulate matter, or $PM_{2.5}$. The questions we are seeking to answer in the course of this study include: 1) How well are the NCEP forecasts predicting the monitored data at various sites in the Shenandoah Valley? 2) Is the forecast predicting better at certain times and/or in certain locations than at others? 3) Is there a bias to the model; i.e., does it have a tendency to over- or under-predict the monitored data? 4) If there is a bias, does it vary depending on season and/or location? To accomplish this objective, we first established where, and of what type, the $PM_{2.5}$ monitoring stations were in the Valley. We found that there were six sites using Federal Reference Method (FRM) monitors to measure $PM_{2.5}$ in the area within which we were working. These sites included: Martinsburg Ball Field (Martinsburg, WV), Big Meadows (Shenandoah National Park), Luray Airport (Luray, VA), Rockingham County VDOT (Harrisonburg, VA), and Raleigh Court Library and Round Hill Montessori School (Roanoke, VA). Two Tapered-Element Oscillating Microbalance (TEOM) monitors were also running, one at Big Meadows and one at Round Hill. After finding the monitors, the time period of the study was determined to be from September 2006 to December 2007. It was confirmed that NCEP was generating forecasts for the same sites and time period. We retrieved the forecast data from the NCEP forecast archives and obtained the corresponding monitored data from the Virginia and West Virginia state environmental agencies for the six sites listed in the paragraph above. After consultation with a statistician, we selected a simple box-plot analysis method applied to a variety of time scales. The monitored data were arranged in a column, in a spread sheet. In the next column we entered the forecast data that corresponded to the same location, day and/or time. In the third column, we subtracted the monitored values from the forecast values to obtain the differences between the two data sets. These differences were used to determine if the model was over- or under-predicting the monitored values. We then made box plots of the differences based on different parameters, i.e. season, month, day of the week, hour of the day. Our conclusions are currently as follows: 1) There was significant variability between the forecast and monitored concentrations of $PM_{2.5}$ for the time period considered. 2) The accuracy of the forecast varied by site, time of the year, day of the week, and hour of the day. 3) The model sometimes under-predicted or over-predicted the

monitored data by 30μg/m^3 or more. 4) $PM_{2.5}$ concentrations at all sites were under-predicted by the model in summer 2007, and over-predicted in fall 2007. 5) Differences between the forecast and monitored data appeared to show patterns when compared by day of the week and hour of the day. 6) It is possible that the Luray and Raleigh Court FRM sites were better predicted than the other four FRM sites. However, a more quantitative analysis would have to be done to support these conclusions.

2 INTRODUCTION

According to Lee, et al., "During 2003, NOAA and the U.S. EPA signed a Memorandum of Agreement to work together to develop a National Air Quality Forecasting (AQF) capability. To meet this goal, NOAA's National Weather Service (NWS), the Office of Atmospheric Research (OAR) and the U.S. EPA developed and evaluated a prototype O_3 [ozone] forecast capability for Northeastern U.S. (Davidson et al, 2004). The NWS/ National Centers for Environmental Prediction (NCEP) North American Meso-scale (NAM) model (Rogers et al, 1996; Janjic 2003) at 12 km was used to drive the EPA Community Multi-scale Air Quality (CMAQ) model (Bynn et al, 1999) to produce up to 48 hour O_3 predictions. From the outset, plans have called for the AQF capability to include particulate matter forecast guidance also. The importance of such a capability is obvious… [Fine] particulate matter suspended in the atmosphere [poses a] hazard to health and impairment to visibility."[1]

The objective of this project is to evaluate the accuracy of the experimental NOAA/NCEP $PM_{2.5}$ forecast model for the Shenandoah Valley region of Virginia. While such models can be highly useful, they are only so when the forecasts being generated are reasonably accurate and reliable. To this end, models need to be verified by comparing their outputs with real-world data.

The questions we were seeking to answer in the course of this study include: 1) How well are the NCEP forecasts predicting the monitored data at various sites in the Shenandoah Valley? 2) Is the forecast predicting better at certain times and/or in certain locations than at others? 3) Is there a bias to the model; i.e., does it have a tendency to over- or under-predict the monitored data? 4) If there is a bias, does it vary depending on season and/or location?

The time period under consideration in the study is from September 2006 to December 2007. As the forecast is still under development, the model and its outputs are constantly changing. Therefore, the results presented here are only representative of the specific time frame referenced above, and cannot be assumed to be indicative of the quality of the forecasts that came before or after that time.

Due to space considerations, this is a shortened version of the original report.

3 MONITORING STATIONS

We were interested in the monitoring stations in the Shenandoah Valley. In the end, only six sites matched the needed criteria, five in Virginia and one in West Virginia. These sites are listed in the following table.

Table 1. Monitoring Sites and Associated Information Used in this Analysis.

AQS Site ID	Latitude	Longitude	County/City	Site Name	Monitor Start Date
540030003	39.4480	-77.9641	Martinsburg, WV	Martinsburg Ball Field	FRM: Feb. 1999
511390004	38.6633	-78.5047	Page County, VA	Luray Airport	FRM: Oct. 1999
511130003	38.5219	-78.4361	Madison County, VA	Big Meadows	TEOM: May 2004 FRM: Sept. 2006*
511650003	38.4773	-78.8190	Rockingham County, VA	VDOT	FRM: Dec. 2006
517700014	37.2561	-79.9850	Roanoke City, VA	Raleigh Court Library	FRM: Jan. 1999
517700015	37.2972	-79.9556	Roanoke City, VA	Round Hill Montessori School	TEOM: Dec. 2004 FRM: Oct. 2006

* The last run for the FRM at Big Meadows was on 8/1/07.

All of the sites listed in Table 1 had FRM monitors at them for the majority of the time between 9/06 and 12/07, and two of them (Round Hill and Big Meadows) also had continuous TEOM monitors.

4 METHODOLOGY

4.1 Overview of Comparison Methodology

We consulted with a statistician in order to determine the best way to compare the monitored and forecast data. A simple box-plot analysis method applied to a variety of time scales was selected.

Both the forecast data and the TEOM monitored data are hourly averages, so those two data sets could be directly compared. The FRM monitored data is a 24-hour average. In order to be able to compare it with the forecast, it was first necessary to average the forecast data over the same 24-hour time period as the FRM. One issue we encountered in this process was that the forecast for the current day does not start until 6am local time. In order to make a 24-hr average we needed to use the hours from midnight to midnight. Luckily, the daily forecasts each run for 48-hours. Therefore, we used the forecast from the previous day to get our midnight to midnight average.

The monitored data was then arranged in a column, in a spreadsheet. In the next column we entered the forecast data that corresponded to the same location, day and/or time. In the third column, we subtracted the monitored values from the forecast values to obtain the differences between the two data sets. These differences were used to determine if the model was over- or

under-predicting the monitored values. We then made box plots of the differences based on different parameters, i.e. season, month, day of the week, hour of the day.

4.2 Data Collection

4.2.1 Forecast Data

The forecast data were collected in two ways. The data from September 2006 to May of 2007 had to be retrieved from the archives of the NCEP node of the NOAA R&D computer complex. The forecasts from June to December of 2007 were posted on the website of NOAA's Environmental Modeling Center (EMC) and were collected directly from that site.

In general, the forecasts were generated by NCEP everyday. However, there were some days when the forecast was not generated, or was not properly archived, and it was therefore not available. In these circumstances, whenever possible the forecast data was filled in using the data from the previous day 48-hour forecast.

4.2.2 Monitored Data

Monitored data (FRM and TEOM) were obtained from the West Virginia Department of Environmental Protection (for Martinsburg, WV) and the Virginia Department of Environmental Quality.

5 RESULTS OF COMPARING FORECAST VS. MONITORED DATA

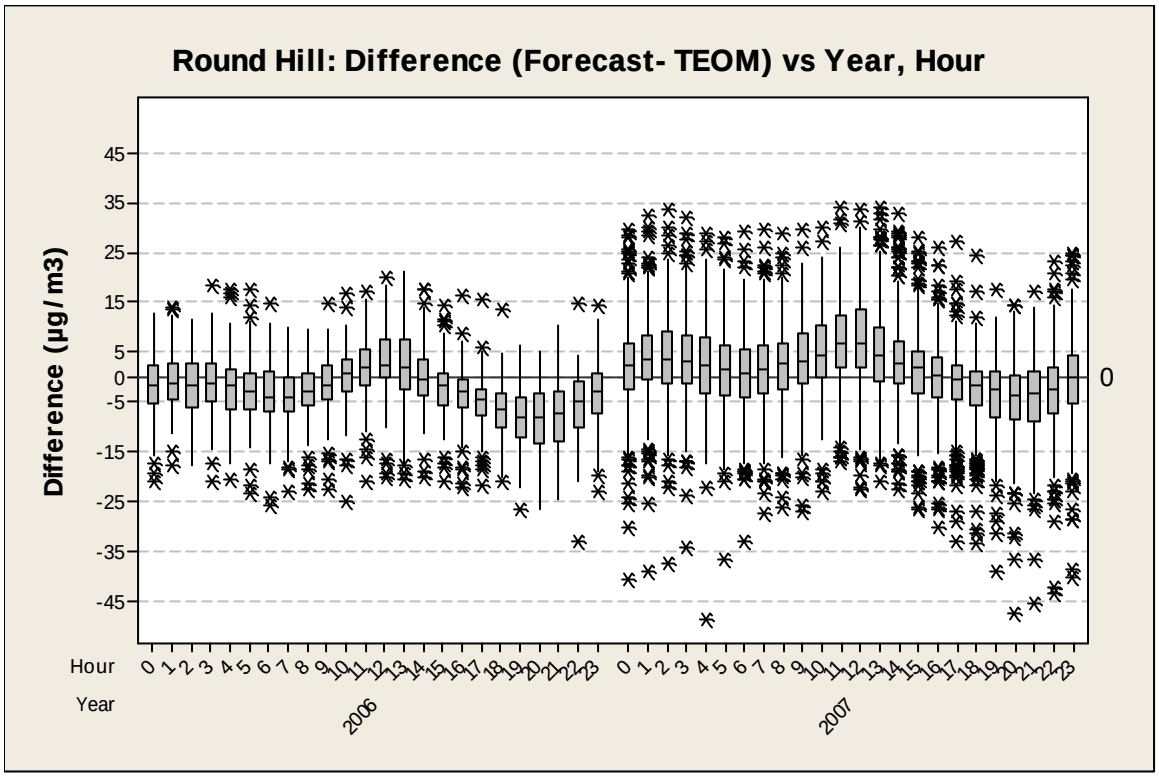
5.1 Hourly Forecasts vs. TEOMs by Hour

The following analysis compares the hourly forecast and TEOM data for the time period of September 2006 through December 2007.

5.1.1 Round Hill

[The way the charts are set up, values above zero indicate that the forecast was predicting a higher value than was measured (over forecast), and those below zero indicate that the forecast was predicting a lower value than was measured (under forecast).]

Figure 1. Box Plot of the Differences between the Forecast and TEOM Data at Round Hill by Year and Hour.



For both years the forecast seems to increase compared to the TEOM data from about 7am through about noon. Then the difference between the forecast and the TEOM data decreases until about 8pm, when the differences start increasing again until 3am. From 3am to 7am the differences decrease.

A similar pattern was observed at the Big Meadows TEOM site.

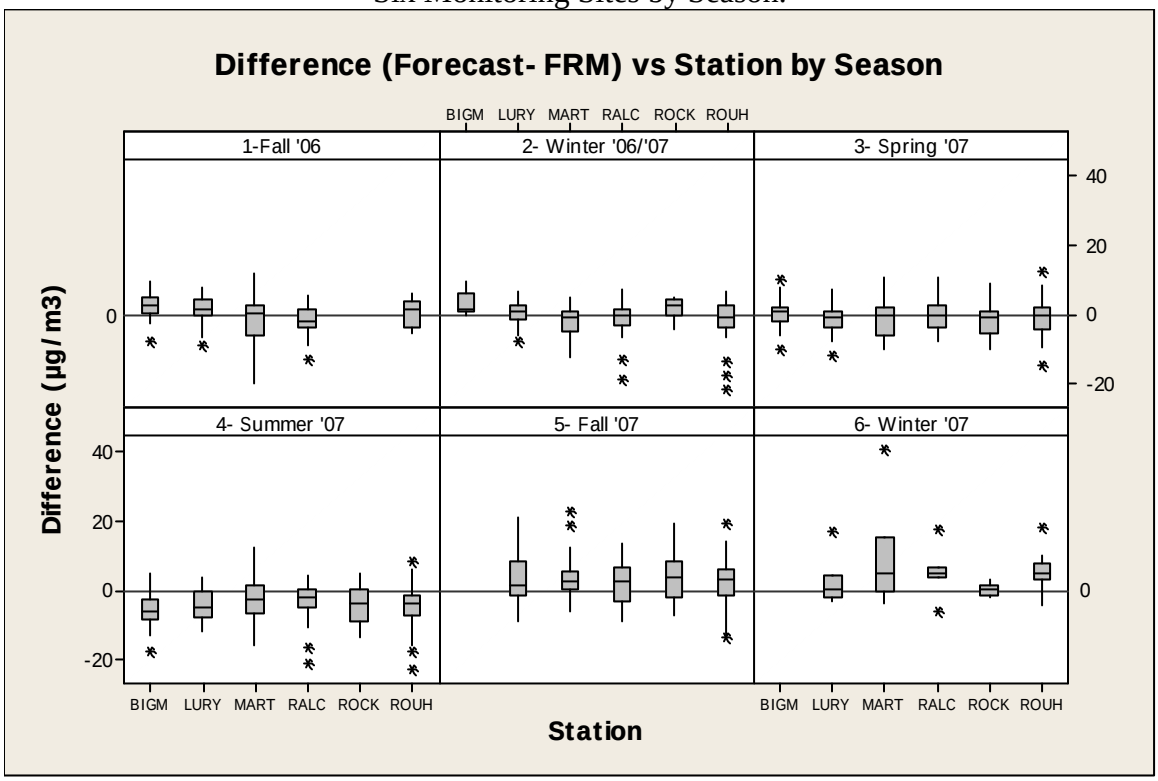
5.2 Comparisons between FRM Monitoring Stations

The following section is a comparison of the results between monitoring stations.

5.2.1 FRM Sites by Season

The following figure shows the forecast minus FRM differences by season at all the FRM sites.

Figure 2. Box Plot Comparison of the Differences between the Forecast and FRM Data at All Six Monitoring Sites by Season.



Fall 2006 was over forecast at three out of five sites— Big Meadows, Luray and Round Hill. The differences were pretty close to zero at Martinsburg (median difference less than 1µg/m^3 from zero). Raleigh Court was under forecast in fall 2006.

Winter 2006/2007 was over forecast at Big Meadows, Luray, and Rockingham, and the median differences were pretty close to zero at Martinsburg, Raleigh Court and Round Hill (median differences less than 1µg/m^3 from zero).

In spring 2007, all sites had median differences that were pretty close to zero (median differences less than 1µg/m^3 from zero).

During the summer of 2007, all sites were under forecast.

In the fall of 2007, all the sites were over forecast (Big Meadows had no data after summer 2007).

Martinsburg, Raleigh Court and Round Hill were all over forecast in the winter of 2007, while Luray and Rockingham had median differences that were close to zero (median differences less than 1µg/m^3 from zero).

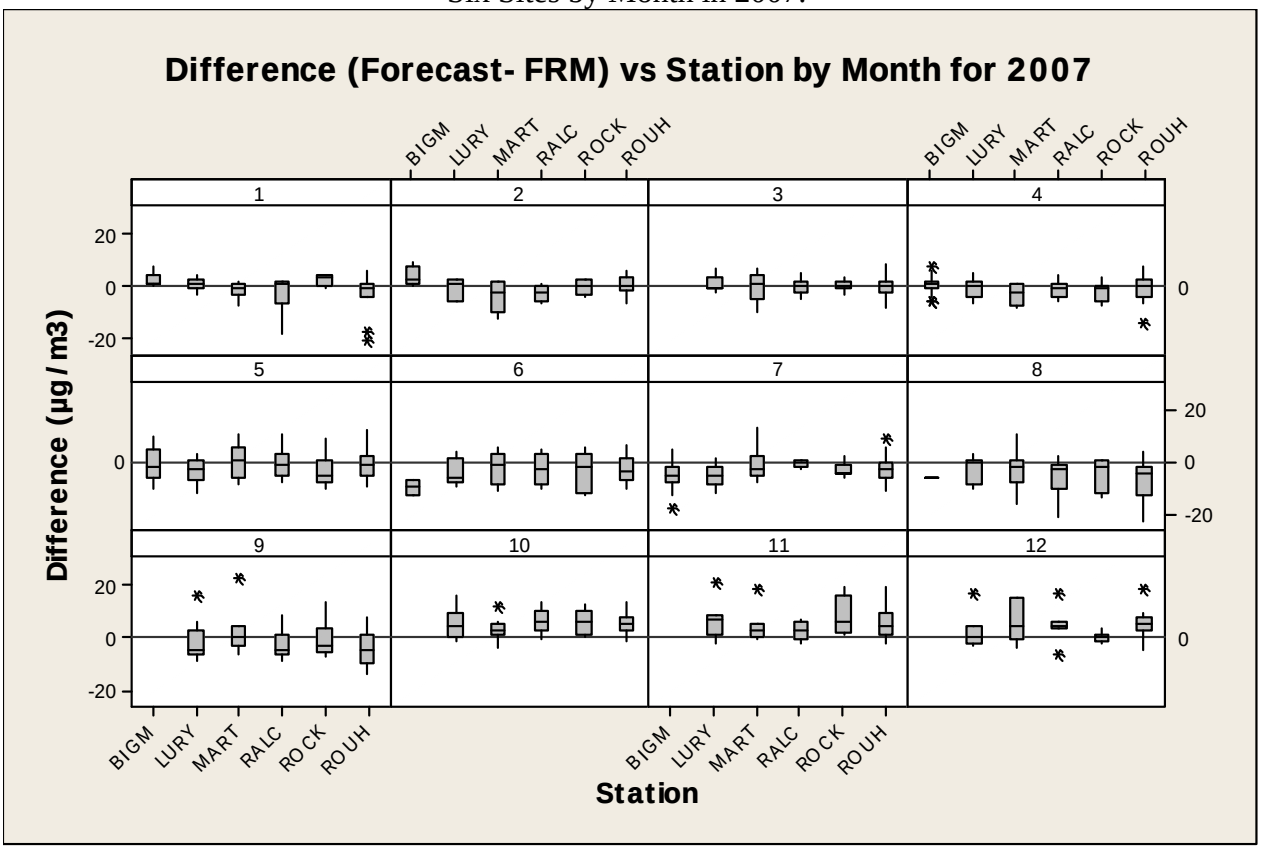
Big Meadows was over forecast 75% of the time in fall 2006 and 100% of the time in winter 2006. Big Meadows and Round Hill were under forecast greater than 75% of the time in

summer 2007, and Luray was under forecast 75% of the time. Martinsburg was over forecast 75% of the time in fall 2007. Raleigh Court and Round Hill were over forecast over 75% of the time in winter 2007.

5.2.2 FRM Sites by Month

The following section discusses the differences between the FRM sites by month. It should be noted that while the analysis encompassed the time between September 2006 and December 2007, only the analysis for 2007 was included here due to space considerations.

Figure 3. Box Plot Comparison of the Differences between the Forecast and FRM Data at All Six Sites by Month in 2007.



January 2007 was over predicted at Big Meadows and Rockingham. Median differences were close to zero at the other four sites (median difference less than or equal to 1µg/m^3 from zero).

February 2007 was over predicted at Big Meadows. Median differences were close to zero at Luray, Rockingham and Round Hill. And, Martinsburg and Raleigh Court were under predicted in February.

There was no data for Big Meadows in March of 2007. All other sites had median differences close to zero that month.

In April 2007 all sites had median differences close to zero except for Martinsburg which was under forecast.

May 2007 was under predicted at all sites except for Martinsburg and Raleigh Court. Those two sites had median differences close to zero.

All sites were under predicted in June 2007. This was also the case in July except for Raleigh Court which had a median value close to zero that month.

There was no data available for Big Meadows in August through December of 2007.

Luray had median differences close to zero in August 2007. All other sites were under predicted.

Martinsburg had median differences close to zero in September 2007. All other sites were under predicted that month.

October and November were over predicted at all sites.

Luray and Raleigh Court had median differences close to zero in December 2007. The other three sites were over forecast for that month.

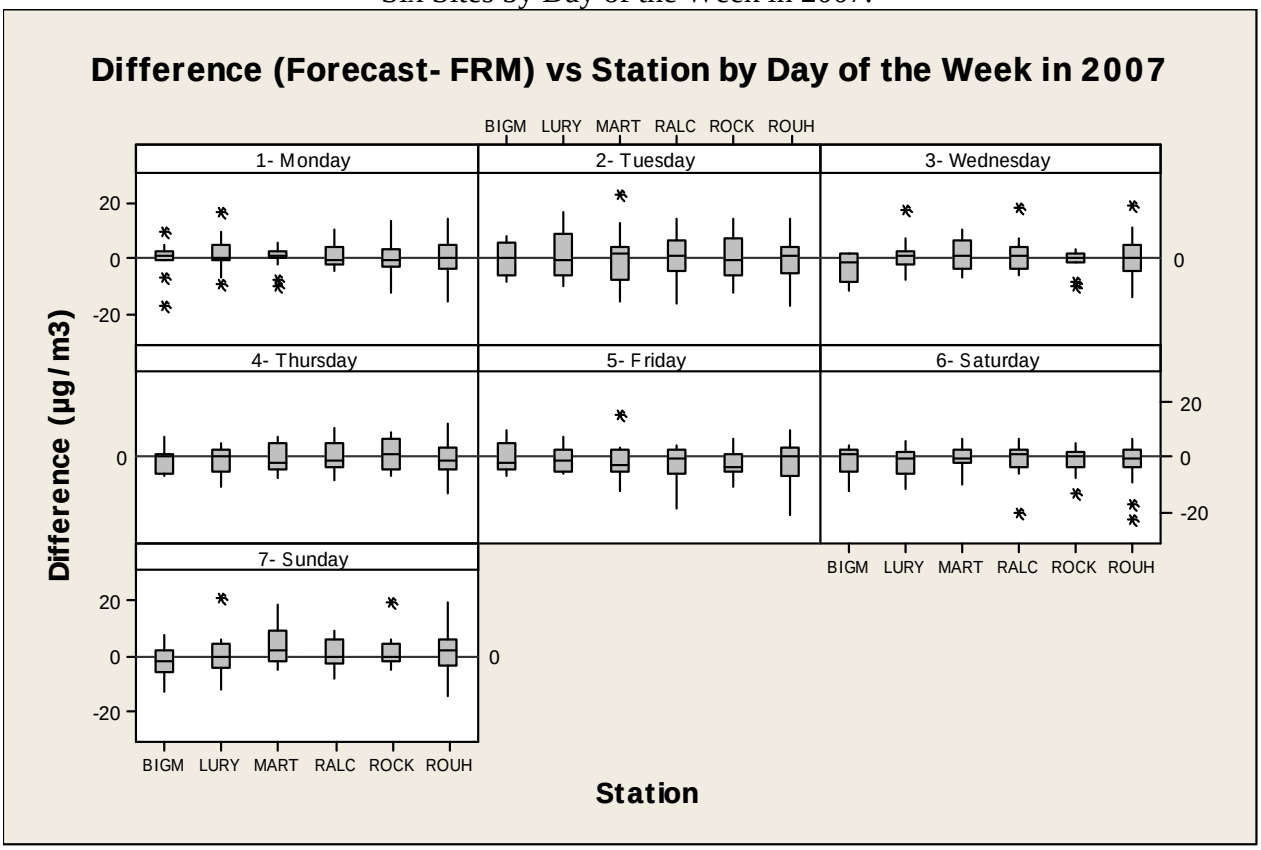
Big Meadows and Rockingham were over predicted 100% of the time in January 2007. Big Meadows was over predicted 100% of the time in February 2007. Rockingham was under forecast 75% of the time in April 2007. Big Meadows was under forecast 100% of the time in June 2007. Big Meadows, Luray, Rockingham, and Round Hill were under forecast at least 75% of the time in July. Raleigh Court and Round Hill were under predicted over 75% of the time in August. All sites were over predicted at least 75% of the time in October 2007. Luray, Martinsburg, Rockingham and Round Hill were over predicted at least 75% of the time in November. Martinsburg, Raleigh Court and Round Hill were over predicted at least 75% of the time in December 2007.

Luray and Raleigh Court had median differences close to zero for seven of the fifteen months under consideration. All the other sites had median differences close to zero for fewer than seven months. This might lead to the conclusion that these two sites were better forecast overall than the other sites. More analysis would have to be done to support this conclusion. It must also be remembered that Big Meadows and Rockingham did not have data for all of the fifteen months.

5.2.3 FRM Sites by Day of the Week

The following two figures show the day of the week differences at each station broken down by year. It should be noted that while the analysis encompassed the time between September 2006 and December 2007, only the analysis for 2007 was included here due to space considerations.

Figure 4. Box Plot Comparison of the Differences between the Forecast and FRM Data at All Six Sites by Day of the Week in 2007.



On Mondays and Saturdays in 2007, all six sites had median differences less than 1µg/m^3 from zero. This also occurred on Wednesdays for five out of six sites.

Fridays, four out of six sites were under predicted by 1µg/m^3 to 4µg/m^3.

All the other days of the week had mixed results.

6 CONCLUSIONS

Our conclusions are currently as follows:

1) There was significant variability between the forecast and monitored concentrations of $PM_{2.5}$ for the time period considered.
2) The accuracy of the forecast varied by site, time of the year, day of the week, and hour of the day.
3) The model sometimes under-predicted or over-predicted the monitored data by 30µg/m^3 or more.
4) $PM_{2.5}$ concentrations at all sites were under-predicted by the model in summer 2007, and over-predicted in fall 2007.

5) Differences between the forecast and monitored data appeared to show patterns when compared by day of the week and hour of the day.
6) It is possible that the Luray and Raleigh Court FRM sites were better predicted than the other four FRM sites. However, more analysis would have to be done to support these conclusions.

7 ACKNOWLEDGEMENTS

This project could not have been completed without the help of a number of people. Thanks to Dr. Gene Tucker, the primary advisor for the project; Jeff McQueen and the rest of the NOAA/NCEP team, without whom I never would have been able to retrieve the archived forecasts and who also helped me in understanding the forecast model; Dr. Mike Deaton for providing statistical advice to the project; Carolyn Stevens (VA DEQ) for providing the monitored $PM_{2.5}$ data for Virginia and for helping to clarify the differences between the types of monitors and the data collection protocols; Pat Adkins (WV DEP) for providing the $PM_{2.5}$ monitored data for the Martinsburg, WV site; and Dr. CJ Brodrick, and Dan Salkovitz (VA DEQ meteorologist) for their review of, and advice on, the project.

8 REFERENCES

1. Lee, P., McKeen, S., McQueen, J., Kang, D., Tsidulko, M., Lu, S., Lin, H., DiMego, G., Seaman, N., & Davidson, P. (2007). *Air Quality forecast using WRF/NMM-CMAQ during the TexAQS*. Retrieved October 15, 2007, from http://ams.confex.com/ams/pdfpapers/120826.pdf.