Particulate Matter and Ammonia Pollution in the Animal Agricultural-Producing Regions of North Carolina: Integrated Ground-Based Measurements and Satellite Analysis

Rebecca Wiegand, William H. Battye, Casey Bray Myers and Viney P. Aneja

Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University, Raleigh, NC 27606, USA; rwiegand94@gmail.com (R.W.); whbattye@ncsu.edu (W.H.B.); cdbray@ncsu.edu (C.B.M.)

* Correspondence: vpaneja@ncsu.edu

Abstract: Intensive animal agriculture is an important part of the US and North Carolina’s (NC’s) economy. Large emissions of ammonia (NH₃) gas emanate from the handling of animal wastes at these operations contributing to the formation of fine particulate matter (PM₂.₅) around the state causing a variety of human health and environmental effects. The objective of this research is to provide the relationship between ammonia, aerosol optical depth and meteorology and its effect on PM₂.₅ concentrations using satellite observations (column ammonia and aerosol optical depth (AOD)) and ground-based meteorological observations. An observational-based multiple linear regression model was derived to predict ground-level PM₂.₅ during the summer months (JJA) from 2008–2017 in New Hanover County, Catawba County and Sampson County. A combination of the Cumberland and Johnston County models for the summer was chosen and validated for Duplin County, NC, then used to predict Sampson County, NC, PM₂.₅ concentrations. The model predicted a total of six 24 h exceedances over the nine-year period. This indicates that there are rural areas of the state that may have air quality issues that are not captured for a lack of measurements. Moreover, PM₂.₅ chemical composition analysis suggests that ammonium is a major component of the PM₂.₅ aerosol.

Keywords: fine particulate matter; ammonia; remote sensing; agricultural emissions; deposition

1. Introduction

Animal feeding operations (AFOs) are agricultural facilities where animals are kept and raised in a small area. While animal agriculture is a large driver of the economy, these farms also create a strain on the environment. The eastern region of North Carolina (NC) has become highly populated with thousands of AFO farms, primarily housing poultry and hogs [1]. NC is ranked 4th in the United States for hog production and 1st in the United States for all poultry production due to the strong turkey production [2]. During 2016, Duplin and Sampson Counties, which are in Eastern NC, contained over 4 million hogs and 100 million chickens [3].

Ammonia (NH₃) emissions from AFOs account for more than half of the reactive nitrogen released into the environment [4,5]. The largest emission source of NH₃ into the atmosphere is from agricultural sources [6,7]. NH₃ plays a significant role in the biogeochemical nitrogen cycle and causes health and environmental effects [8–17]. The manure that is produced by the animals is frequently used as fertilizer and spread on crops throughout the farm, further releasing the pollutant into the atmosphere. NH₃ also has several human health effects including eye, nose and throat irritation, dizziness and headaches [18,19]. Once NH₃ is released into the atmosphere it can then deposit by either rainfall or dry deposition to regional waterbodies or it can react with other compounds in the area to create other pollutants and cause further harm [14–17,20–22]. Gaseous NH₃ in the atmosphere contributes to the formation of airborne fine particulate matter (PM₂.₅)
through reactions with water vapor and other air pollutants, including oxidation products of sulfur dioxide (SO$_2$) or nitrogen oxides (NO and NO$_2$, or NO$_X$) [23]. Ammonium compounds, including ammonium sulfates (NH$_4$HSO$_4$ and (NH$_4$)$_2$SO$_4$) and ammonium nitrate (NH$_4$NO$_3$), make up a large fraction of PM$_{2.5}$ [24] (defined as particulate matter with an aerodynamic diameter of 2.5 microns or smaller). Elevated concentrations of PM$_{2.5}$ are associated with respiratory issues, heart problems and has been linked to premature death [25–27]. Ammonia can be important in the nucleation of new particles [28,29]. These ammoniated particles scatter light, attenuating visibility, and can result in some atmospheric cooling [30]. Particulate matter emissions are released through primary sources (e.g., power plants, transportation) [31] as well as secondary sources (e.g., gas to particle conversion). However, in the United States, regulatory strategies to reduce PM$_{2.5}$ have focused on primary emissions of PM$_{2.5}$ and on reductions of SO$_2$ and NO$_X$ emissions; however, the control of NH$_3$ emissions can also be effective for reducing concentrations of PM$_{2.5}$.

The North Carolina Clean Smokestacks Act has reduced sulfur dioxide emissions by 89% compared to 1998 emissions [32] and the improved regulations on vehicle manufacturers has helped to reduce NO$_x$ emissions. Because these reductions have already been made, any additional effort to reduce both SO$_2$ and NO$_x$ emissions would become less cost effective. Thus, some studies have started to indicate that a reduction in ammonia emissions would be a more cost-effective solution in terms of long-term reductions in particulate matter concentrations [33–40]. However, ammonia is not a criteria pollutant. Moreover, any reduction in particulate matter must be measured at monitoring stations, leading to economic, political and logistical problems.

Recent studies in Europe using detailed chemical transport models and time resolved NH$_3$ emissions illustrate the strong nonlinearity between PM$_{2.5}$ and NH$_3$ emissions, and the reduction in NH$_3$ emissions significantly reduces PM$_{2.5}$ levels in both summer and winter periods [33]. The impact of NH$_3$ emissions on PM$_{2.5}$ depends on meteorological parameters (e.g., temperature, relative humidity), the magnitude of the perturbation to NH$_3$ emissions and the abundance of particulate nitrate (NO$_3^-$), gaseous nitric acid (HNO$_3$) and particulate sulfate (SO$_4^{2-}$ and HSO$_4^-$), which are the products of the oxidation of SO$_2$ and NO$_x$, two byproducts of combustion [31,33,35]. Utilizing this foundational knowledge of key processes and an integrated approach of using satellite measurements (ammonia and aerosol optical depth) and ground-based measurements (meteorology and PM$_{2.5}$), an observational-based statistical model is developed to predict PM$_{2.5}$ concentrations in these rural regions in an effort to advance Earth system predictability. The abundance of NO$_3$T and SO$_4$T was not measured.

Measurements of PM$_{2.5}$ in rural regions (especially in intensively agricultural locations) of the US are in general not conducted and are therefore not available. Recent studies [21,41] have shown that increased ammonia from agriculture leads to increased PM$_{2.5}$. Therefore, the purpose of this work is to estimate PM$_{2.5}$ concentrations in ammonia-rich environments of Eastern NC using a combination of satellite and in situ data. Using satellite data will allow us to develop a method for predicting ground-level PM$_{2.5}$ concentrations in areas of high agricultural influence which normally do not have a ground-based measurement site. By utilizing satellite-derived ammonia concentrations and aerosol optical depth (AOD) along with meteorological parameters, we can predict, with reasonable certainty, PM$_{2.5}$ concentrations across NC. Moreover, a thorough review of the scientific literature did not provide any studies offering predictions of PM$_{2.5}$ in rural regions based on our process-based approach (i.e., coupling of satellite and ground-based measurements). However, a complex numerical air quality model such as the Weather Research and Forecasting-Community Multiscale Air Quality (WRF-CMAQ) or WRF-Chem Modeling System may also be utilized to predict PM$_{2.5}$ and NH$_3$ emission from agricultural land using these air quality models across various regions of the world [42–45].
2. Data and Methodology

2.1. Location Description

AFOs in Eastern NC, while good for the economy, create a strain on the environment. For example, many residences of these high agriculturally active counties have started filling nuisance lawsuits against large hog farms’ odor issues and a general decrease in quality of life for those in the area. For these reasons, our study focuses on various areas in NC, all of which, due to proximity and wind direction, will have a direct influence from these farms. This ultimately will allow for further understanding of how these agricultural areas affect ground-level PM$_{2.5}$ concentrations.

The most recent measurements within the region of our study are for Kinston, NC, in 2007, as part of the U.S. EPA Chemical Speciation Network for PM$_{2.5}$. Ammonium compounds formed a substantial portion of the mass of PM$_{2.5}$ even during the summer months [46], as shown in Table 1 and Figure 1.

**Table 1.** PM$_{2.5}$ speciation in Kinston, NC, during the summer months of 2006 and 2007 *.

<table>
<thead>
<tr>
<th>PM$_{2.5}$ Composition (µg m$^{-3}$)</th>
<th>Summer 2006</th>
<th>Summer 2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ammonium sulfates and ammonium nitrate</td>
<td>7.08</td>
<td>7.34</td>
</tr>
<tr>
<td>Organic carbon</td>
<td>4.44</td>
<td>5.10</td>
</tr>
<tr>
<td>Elemental carbon</td>
<td>0.22</td>
<td>0.24</td>
</tr>
<tr>
<td>Other and unidentified</td>
<td>2.28</td>
<td>2.45</td>
</tr>
<tr>
<td>Total</td>
<td>14.03</td>
<td>15.13</td>
</tr>
<tr>
<td>Fraction ammonium sulfates and nitrate</td>
<td>50%</td>
<td>49%</td>
</tr>
<tr>
<td>Fraction EC/OC</td>
<td>33%</td>
<td>35%</td>
</tr>
</tbody>
</table>


**Figure 1.** Speciation of PM$_{2.5}$ (µg m$^{-3}$) during the summer months of 2007 in Kinston, NC, USA.

Ammonia is directly measured in North Carolina by passive monitors as part of the Ammonia Monitoring Network (AMoN); however, the locations of these monitors is sparse, with only 4 active sites across the state of North Carolina. Moreover, particulate matter is measured at various locations throughout the state with monitors placed and maintained by the Environmental Protection Agency (EPA) and the North Carolina Department of
Environmental Quality (NCDEQ). The location of these monitors (Figure 2) is widespread throughout the state and lacks the ability to monitor at a high spatial resolution. To achieve the spatial resolution needed to accurately monitor PM$_{2.5}$ reductions throughout the state, satellite data may be utilized. Satellites allow us to obtain both the spatial and temporal data resolution needed to accurately determine if there is a relationship between ammonia emissions from agricultural farms in Eastern North Carolina and PM$_{2.5}$ concentrations. Satellite data can also be used to predict PM$_{2.5}$ concentrations in areas with a high agricultural trace gas emission.

![PM$_{2.5}$ Monitoring Sites](image)

**Figure 2.** PM$_{2.5}$ site locations. These sites are set up and maintained by the Environmental Protection Agency (EPA) and the North Carolina Department of Environmental Quality (NCDEQ).

Satellite data are a relatively new data source in air quality monitoring [47–50]. These satellites can be extremely useful in capturing a relatively high-resolution (such as 0.25° by 0.25° or 1 km by 1 km) air quality image over the entire world for a single day or multiple days. For this reason, satellite data can be especially useful in rural areas around the state, where monitoring networks are scarce. Using satellite data will allow us to develop a method for predicting ground-level PM$_{2.5}$ concentrations in areas of high agricultural influence which normally do not have a ground-based measurement site. By utilizing satellite-derived ammonia concentrations and aerosol optical depth (AOD) along with meteorological parameters, we can predict, with reasonable certainty, PM$_{2.5}$ concentrations across North Carolina.

2.2. Satellite-Derived Ammonia Data

Satellite-derived ammonia retrievals from the Infrared Atmospheric Sounding Interferometer (IASI, version 2.1, EUMETSAT) were used with <25% cloud cover for 2008 through 2017. IASI was used because it provides a longer-term repository of measurement data than other satellite platforms, such as the Cross-track Infrared Sounder (CrIS). IASI is a collaboration between the European Organization for the Exploitation of Meteorological Satellites and the National Centre for Space Studies (CNES) or the French government space agency. IASI was launched in 2006 on the MetOp-A satellite, in 2012 on the MetOp-B satellite and finally on the MetOp-C satellite launched in 2018. They are polar orbiting sun synchronous satellites with an orbit altitude at around 817 km above the Earth’s surface. IASI has an orbital period of about 90 min, crossing the equator at around 9:30 a.m. and p.m. local time, with each subsequent pass displaced by about 22.5 degrees of longitude. The satellite measures detailed infrared spectra over a broad angular swath with a spatial resolution of about 12 km [51]. IASI researchers retrieve estimated total atmospheric column loadings of NH$_3$ and other pollutants based on patterns of infrared absorption. Specific information on the development and comparison testing of the algorithm used to derive ammonia measurements is seen in both [51,52].
At the latitude of NC, morning and evening measurements are collected at about 10 a.m. and 10 p.m. local time. For this study, we used the morning pass to perform the analysis because NH$_3$ concentrations are typically higher during this time. This is because during the morning hour, the mixing layer height is lower in the atmosphere, which allows for NH$_3$ concentrations to remain closer to the surface. The data were filtered by the error estimate reported with the data. Only the points that had an error less than 100% were used in this study.

2.3. Satellite-Derived Aerosol Optical Depth

Satellite-based 550 nm AOD retrievals from the Moderate Resolution Imaging Spectroradiometer (MODIS) from 2008–2017 were used in this study. MODIS is onboard National Aeronautics and Space Administration’s (NASA) Aqua and Terra satellites which, similar to IASI, are polar orbiting satellites. Aqua and Terra orbit at around 705 km above the Earth’s surface and pass NC once per day. For this study, MODIS level 2 collection 6.1 high confidence AOD at 10 km $\times$ 10 km measurements from Aqua were used. Retrievals onboard the Aqua satellite were used specifically due to deterioration issues with the MODIS sensor on the Terra satellite. A visual example of these data can be seen in Figure 3. The blank location in the image can be attributed to cloud cover over the area [51]. Another reason for the blank locations in the image is the filtering process of the data. The data were filtered to only include the best quality data, which NASA indicates is a quality assurance flag of 3.

![MODIS Aerosol Optical Depth (0.55 $\mu$m) for july 30, 2011](image)

**Figure 3.** Aerosol optical depth (AOD) retrievals from the MODIS Aqua satellite for 30 July 2011.

2.4. Ground-Based PM$_{2.5}$

The ground-based PM$_{2.5}$ data were obtained from the Environmental Protection Agency’s (EPA) Outdoor Air Data website, developed based on EPA’s Air Quality System (AQS) network (https://www.epa.gov/outdoor-air-quality-data/download-daily-data; accessed on 30 October 2019). The air quality stations around the state are a collaboration between the EPA office in Research Triangle Park and NCDEQ. Two NC sites were chosen (i.e., training data), one in Cumberland County (35.04 N, 78.95 W) and another in Johnston County (35.59 N, 78.46 W) as input data to develop the statistical model. These locations were chosen based on their proximity to high agricultural areas in NC. The PM$_{2.5}$ data were available at each of those sites for 3-day intervals for 2008–2017. PM$_{2.5}$ data were then collected (i.e., test data) from a site in Duplin County (34.95 N, 77.96 W), New Hanover County (34.36 N, 77.84 W) and Catawba County (35.73 N, 81.36 W) to validate the model performance. The monitor used in NC is an active monitor that takes in a predetermined amount of air and pulls it through a filter within the sensor. After 24 h, the filter is then taken out of the sensor and gravimetrically measured and the mass concentration in $\mu$g/m$^3$. 


is reported [52]. PM$_{2.5}$ was the only particulate matter size incorporated into this study because PM$_{10}$ only makes up around less than 20% of the measured PM concentrations in these locations.

2.5. Meteorology Data

Ground-based meteorology data were obtained from the NC State Climate Office for the entire study period from sites in Cumberland, Johnston, Duplin, Sampson, Catawba and New Hanover Counties. Daily averaged temperature in degrees Celsius (°C), pressure in millibar (mb), wind speed in meters per second (m/s) and relative humidity in percent (%) where available from each station and used in the development of the model. Days with reported rainfall were excluded from the final analysis because they would lead to washout events for both the ammonia and AOD data. Figure 4 shows the location of these sites, which for Cumberland, Johnston, Duplin, Catawba and New Hanover are collocated with the ground-level PM$_{2.5}$ sites. The figure also shows the counties with the highest agricultural production as of 2017 and the predominant wind directions to illustrate the agricultural influence on each of these areas are shown in Figure 5.

![Figure 4. The initial locations that we built the model from are seen in blue while the locations we modeled can be seen in black. Total ammonia emissions from agricultural farms are indicated by the shading.](image-url)
Figure 5. The initial locations that we built the model from are seen in blue while the locations we modeled can be seen in black. Animal density is indicated by the shading and the wind roses are used to indicate the predominate wind direction and speed from three sites, New Hanover, Sampson and Catawba Counties.

2.6. Methodology

IASI ammonia retrievals are filtered so that only those retrievals with a relative error less than 100% are used. AOD retrievals are filtered to use only those observations with the best quality rating (quality assurance flag = 3). For the Cumberland County training case, a total of 135 observation days were identified for which NH$_3$ retrievals and AOD retrievals meeting the above filtration criteria could be compared with available ground-level PM$_{2.5}$ measurements. For the Johnston County training case, a total of 142 observation days were identified. The quality assurance flag is given by the algorithm team as their assessment of the data quality [53]. These retrievals are then collocated with the ground-level meteorology sites and the ground-level PM$_{2.5}$ sites. This is carried out through the “average” approach [54,55]. In the averaging method, a 100 km radius of the PM$_{2.5}$ ground site and the ground-level meteorology station are averaged together to receive a single value for the time period, which in this study is one day. This process is repeated every day for the months of June, July and August for the 9-year period. June, July and August were chosen because they would allow for the highest ammonia retrieval as, statistically, the summer months have the highest ammonia concentrations in NC. Days with reported rainfall are excluded from the analysis because this would result in a washout event for both the PM$_{2.5}$ data and the ammonia data as they are typically scavenged by rainfall [41].

At the sites in Cumberland and Johnston Counties, the data were then compiled for each day using averaged meteorology parameters and subsequently averaged by month. A multivariate regression model was run in the Statistical Analysis System (SAS) using the IASI ammonia data, the MODIS AOD data and the meteorology data per month at both locations. We then tested the accuracy (i.e., validated) of our model using data from Duplin County. For the Duplin County test, a total of 175 observation days were identified, for
which NH\textsubscript{3} and AOD retrievals met the filtration criteria, and PM\textsubscript{2.5} measurements were available for comparison. Once the model was tested, PM\textsubscript{2.5} values were then predicted in Catawba County, New Hanover County and Sampson County to see the influence of agriculture on PM\textsubscript{2.5} concentrations across the state and in different eco-regions. For Catawba County, 222 observation days were identified for which NH\textsubscript{3} and AOD retrievals met the filtration criteria, 168 valid observation days were identified for New Hanover County and 490 observation days for Sampson County.

3. Results

Multiple linear regression is a classical and well-known statistical technique to quantify the association of a variable (called the dependent variable) with several other variables (called independent variables). ANOVA or analysis of variance is a statistical technique to gauge the statistical significance and practical usefulness of the linear regression towards explaining the dependent variable. In this analysis, we implemented multiple linear regression and ANOVA to quantify the association of PM\textsubscript{2.5} (the dependent variable) with six independent variables: ammonia, AOD, T, P, WS and RH. Furthermore, we constructed the ANOVA table to study the statistical significance of this multiple linear regression model. Multiple linear regression models were created to assess the ability of remotely sensed data to predict ground-based PM\textsubscript{2.5} concentrations in rural areas that are usually characterized as having high agricultural activity and no ground-based PM\textsubscript{2.5} monitor. The models were created for Cumberland and Johnston Counties in NC (i.e., training data) and then validated from Duplin County NC (i.e., test data) against the ground-based measurements located in the county. Cumberland and Johnston Counties were chosen specifically because of their proximity to Sampson County, which has the highest agricultural activity in that state next to Duplin County, which was chosen as the validation site for this reason. Sampson County could not be used for validation because there is no ground-based PM\textsubscript{2.5} monitor in the county. After reviewing and testing different model inputs (e.g., a variety of meteorological parameters with data from each county of interest), a combination of Cumberland County and Johnston County data for the entire summer period were chosen:

\[
PM_{2.5} = \exp(15.14 + (1.05 \times 10^{-17} \times \text{Ammonia}) + (1.51 \times \text{AOD}) + (0.26 \times T) + (-0.013 \times P) + (-0.040 \times WS) + (-0.013 \times RH))
\]

where PM\textsubscript{2.5} is given in units of \(\mu g m^{-3}\), Ammonia is the total atmospheric column loading of ammonia as retrieved from IASI observations in molecules cm\textsuperscript{-2}, AOD is the MODIS aerosol optical depth, T is temperature in \(^\circ C\), P is pressure in millibars, WS is wind speed in meters per second and RH is relative humidity in percent.

This model gave an \(r^2\) value of 0.43. The specific statistics can be seen in Table 2 (an analysis of variance or ANOVA table). Here, DF stands for “degrees of freedom”, i.e., the number of values that are free to vary as we estimate the parameters of the model. The total sum of squares (TSS) is the sum of the squared difference between each PM\textsubscript{2.5} observation and the sample average, which is decomposed into two parts: the model sum of squares which is the part of the TSS that can be explained by the regression model, and the error sum of squares which is the remaining part of the TSS which cannot be explained by the model. The mean squares are the sum of squares divided by the respective degrees of freedom. The F value is the ratio of the mean model square to the error mean square, which is a classical statistical metric to quantify the explanatory power of the regression model and the final column is the \(p\)-value associated with the F value which indicates the statistical significance of the model (lower means more significant). The F value and the \(p\)-value seen in the table indicate that our model was able to explain a statistically significant portion of the data variation [54,55]. Table 3 is a parameter estimate table. The table indicates the significance of a parameter based on the t statistic used in the model. For our model, the table shows that AOD is the most significant variable in our model. This is to be expected, since AOD results from the presence of PM\textsubscript{2.5} in the atmosphere.
However, AOD is not directly proportional to the mass concentration of PM$_{2.5}$, but is also dependent on the particle size distribution and other properties of the aerosol and the atmosphere. Relative humidity, temperature and ammonia are also significant variables in our model. The intercept is also relatively significant while pressure and wind speed are the least significant. We know this because Pr $> F$ values have a significant range to them. Anything less than 0.01 has strong significance while anything between 0.01 to 0.05 has appreciable significances.

Table 2. Analysis of variance (ANOVA) table.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>$F$ Value</th>
<th>Pr $&gt; F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>6</td>
<td>18.84</td>
<td>3.14027</td>
<td>38.99</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Error</td>
<td>309</td>
<td>24.88</td>
<td>0.08053</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>315</td>
<td>43.73</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Parameter estimates for the model.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter Estimate</th>
<th>Standard Error</th>
<th>Type II SS</th>
<th>$F$ Value</th>
<th>Pr $&gt; F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>15.14</td>
<td>5.07</td>
<td>0.72</td>
<td>8.92</td>
<td>0.0030</td>
</tr>
<tr>
<td>Ammonia</td>
<td>$1.05 \times 10^{-17}$</td>
<td>$2.99 \times 10^{-18}$</td>
<td>0.98</td>
<td>12.21</td>
<td>0.0005</td>
</tr>
<tr>
<td>AOD</td>
<td>1.51</td>
<td>0.16</td>
<td>7.05</td>
<td>87.53</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Temp</td>
<td>0.26</td>
<td>0.007</td>
<td>1.06</td>
<td>13.14</td>
<td>0.0003</td>
</tr>
<tr>
<td>Pressure</td>
<td>$-0.013$</td>
<td>0.005</td>
<td>0.54</td>
<td>6.65</td>
<td>0.0104</td>
</tr>
<tr>
<td>WS</td>
<td>$-0.04$</td>
<td>0.02</td>
<td>0.38</td>
<td>4.75</td>
<td>0.0301</td>
</tr>
<tr>
<td>RH</td>
<td>$-0.013$</td>
<td>0.003</td>
<td>2.12</td>
<td>26.35</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

The model is dominated by data in the range of ~5 to ~25 micrograms per cubic meter. The model Pr $< 0.0001$ and $F$ Value of ~39 (Table 2), coupled with $r^2$ of ~0.5 (Figure 6, Duplin County), suggest that the model is offering a good prediction overall. Moreover, having Pr for each of the variable (Table 3) less than 0.05 suggests the robust contribution of each variable.

The results of the Duplin County prediction can be seen in Figure 6. The figure shows four different scatter plots each representing a summer month (June, July and August) and a total graph that included all three months in one plot. These plots act to illustrate a visual representation of the model performance at this location. Table 4 illustrates the normalized mean bias (NMB) and normalized mean error (NME) values, which are commonly used to assess the performance of models [54,55].
Figure 6. Scatter plots for Duplin County indicating model performance. The black points indicate the modeled versus predicted point and the grey bars indicate the ±1 standard deviation (s.d.) of the modeled dataset. The red line indicated the one-to-one line and the blue line is the best fit line for the dataset. The slope of the best fit line is indicated in blue in the corner of each figure.

Table 4. Normalized mean bias (NMB) and normalized mean error values (NME) for Duplin County by month.

<table>
<thead>
<tr>
<th>Month</th>
<th>Normalized Mean Bias</th>
<th>Normalized Mean Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>June</td>
<td>0.47%</td>
<td>25.31%</td>
</tr>
<tr>
<td>July</td>
<td>13.19%</td>
<td>24.66%</td>
</tr>
<tr>
<td>August</td>
<td>14.59%</td>
<td>25.23%</td>
</tr>
</tbody>
</table>

Once this model was verified for Duplin, it was then used to predict PM$_{2.5}$ concentrations in New Hanover and Catawba Counties. These locations are also affected by high agricultural production areas; however, they have very different meteorological conditions and processes affecting them daily. The results of this analysis can be seen in Figures 7 and 8. Figure 7 shows the results for New Hanover County and illustrates that despite the different meteorological mechanisms, the model can predict PM$_{2.5}$ concentrations for this area at a relatively high degree of accuracy. Table 5 shows the model prediction parameters used for this study and clearly shows the model’s prediction struggle in July. This particularly high value is the result of the model’s dependence on AOD values. A high AOD value will sometimes result in an overpredicted PM$_{2.5}$ value.
Table 5. Normalized mean bias (NMB) and normalized mean error values (NME) for New Hanover County by month.

<table>
<thead>
<tr>
<th>Month</th>
<th>Normalized Mean Bias</th>
<th>Normalized Mean Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>June</td>
<td>17.62%</td>
<td>28.55%</td>
</tr>
<tr>
<td>July</td>
<td>17.70%</td>
<td>43.61%</td>
</tr>
<tr>
<td>August</td>
<td>19.86%</td>
<td>28.04%</td>
</tr>
</tbody>
</table>

Figure 7. Scatter plots for New Hanover County indicating model performance. The black points indicate the modeled versus predicted point, and the grey bars indicate the ±1 standard deviation of the modeled dataset. The red line indicated the one-to-one line and the blue line is the best fit line for the dataset. The slope of the best fit line is indicated in blue in the corner of each figure.
Figure 8. Scatter plots for Catawba County indicating model performance. The black points indicate the modeled versus predicted point, and the grey bars indicate the ±1 standard deviation of the modeled dataset. The red line indicated the one-to-one line and the blue line is the best fit line for the dataset. The slope of the best fit line is indicated in blue in the corner of each figure.

Similarly, Figure 8 shows the results for Catawba County; however, the model consistently under predicts concentrations at this location compared to both New Hanover and Duplin Counties, which can be seen in Table 6. This indicates that the different meteorological mechanisms that dominate in Catawba County are not well understood by the model. Catawba County is also at a different elevation than the other two locations, which could explain some of the inconsistent predictions. Overall, these results suggest that the multiple regression model can predict (at a relatively high certainty) for the eastern portion of NC and loses some capabilities in the western portion of the state likely due to topography.

**Table 6.** Normalized mean bias (NMB) and normalized mean error values (NME) for Catawba County by month.

<table>
<thead>
<tr>
<th>Month</th>
<th>Normalized Mean Bias</th>
<th>Normalized Mean Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>June</td>
<td>–9.68%</td>
<td>23.92%</td>
</tr>
<tr>
<td>July</td>
<td>–13.19%</td>
<td>26.01%</td>
</tr>
<tr>
<td>August</td>
<td>–14.85%</td>
<td>27.10%</td>
</tr>
</tbody>
</table>

We then used the combination model to predict PM$_{2.5}$ values in Sampson County. As previously mentioned, Sampson County has one of the highest concentrations of agricultural activity (both animal and crop) in the state. Due to its rural landscape, there is no PM$_{2.5}$ monitor available and thus no PM$_{2.5}$ data for this location. Our model results showed,
also seen in Figure 9, that the PM$_{2.5}$ concentrations are low; however, the model did indicate that six days out of the nine-year period were over the EPA NAAQS of 35 $\mu$g m$^{-3}$. In order to further investigate how many exceedances were predicted by the model, the normalized mean error values calculated for New Hanover County were used to see how the model’s error would affect the number of exceedances calculated. New Hanover County errors were used because they were the highest of the three errors calculated for the different counties. Given the errors in the model, the number of exceedances could range from ten to three during the ten-year period. This could indicate that areas near agricultural activity could see higher PM$_{2.5}$ values, most of which are not being captured due to monitoring locations. This could also suggest that a reduction in ammonia emissions could have a positive impact on PM$_{2.5}$ concentrations in these high agricultural areas.

![Figure 9](image-url)

**Figure 9.** Histogram plot illustrating the model predictions in Sampson County for June (a), July (b), August (c) and total (d). The blue lines indicate the number of times (y-axis) the PM$_{2.5}$ concentrations reached a certain amount (x-axis). The red line indicates the NAAQS exceedance value.

### 4. Conclusions

We have created a multiple regression model that can predict PM$_{2.5}$ mass concentrations in counties with the highest agricultural activity in NC. High agricultural activity results in high concentration of ammonia being released from agricultural farms. Ammonia is also a precursor to PM$_{2.5}$ formation and may cause an increase in PM$_{2.5}$ concentrations in these areas which could lead to poor air quality and quality of life for those living in the area.

This model was developed using remotely sensed data (column ammonia and AOD) and ground-based meteorology, making it resistant to issues in the scarcity of ground-
based PM$_{2.5}$ monitors around the state. Many of the areas around these agricultural farms are rural and have no monitor in place to track these concentrations daily. Satellite data, however, can introduce other limitations. For example, the instrument cannot directly measure ammonia concentrations at the surface, the concentrations are calculated through an algorithm developed by researchers which introduces limitation on the accuracy of the concentrations from the satellite. The same issues can be said for the AOD data. Despite the limitations, satellite algorithms are improving over time and as they improve, utilizing these data will become more important. Satellite data can also be utilized for the meteorology data if ground-based data are not available. Modern-era retrospective analysis for research and applications (MERRA) data is globally modeled meteorological data based on GEOS-5 atmospheric data assimilation system. The data models ground-based temperature, pressure and wind speed well. However, the data do not contain relative humidity, thus it must be obtained experimentally or mathematically. Moreover, PM$_{2.5}$ chemical composition analysis (Figure 1) suggests that ammonium is a major component of the PM$_{2.5}$ aerosol.

The model was developed (i.e., training data) for the Cumberland and Johnston Counties. The model was validated (i.e., tested) for Duplin County and applied to predict PM$_{2.5}$ in New Hanover, Catawba and Sampson Counties, all of which have a predominate influence from these agricultural farms. The model predicted PM$_{2.5}$ values in Sampson County that were above the National Ambient Air Quality Standard (NAAQS) limit of 35 µg m$^{-3}$. The need to investigate an ammonia reduction strategy due to its effects on PM$_{2.5}$ concentrations in high agricultural areas is becoming more prevalent. If the PM$_{2.5}$ reduction strategies seen in the past, such as the Clean Smokestacks Act and the Clean Air Act, have been as successful as they can be, then reducing ammonia emissions will not only provide air quality improvements, but also a reduction in PM$_{2.5}$-related issues.

Author Contributions: Conceptualization, V.P.A.; Formal analysis, R.W., W.H.B. and C.B.M.; Funding acquisition, V.P.A.; Investigation, R.W., W.H.B., C.B.M. and V.P.A.; Methodology, R.W., W.H.B. and V.P.A.; Project administration, W.H.B. and V.P.A.; Supervision, V.P.A.; Writing —original draft, R.W.; Writing —review & editing, R.W., W.H.B., C.B.M. and V.P.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Informed Consent Statement: Not applicable.

Acknowledgments: We would like to first acknowledge the scientist in the Atmospheric Spectroscopy group at the Université Libre de Bruxelles, Belgium (ULB-LATMOS) for access to the IASI retrievals and advice on using the data. Next, we would like to acknowledge various people at the NC Department of Environmental Quality, Elliot Tardif, Bradley McLamb, Nicolas Witcraft and Joette Steger for their advice and guidance on the PM$_{2.5}$ monitoring data. We thank David Dickey, William Neal Reynolds Distinguished Professor and Srijan Sengupta, Department of Statistics, NC State University, Raleigh, NC 27695, for reviewing and discussing the statistical development in the paper. We thank Swarnali Sanyal, University of Illinois, for her assistance in air quality modeling discussions. We would also like to acknowledge the air quality research group at NC State University for their assistance.

Conflicts of Interest: There are no conflict of interest.

References


36. Stokstad, E. Ammonia pollution from farming may exact hefty health costs. *Science* 2014, 343, 238. [CrossRef]


42. Mazzeo, A.; Zhong, J.; Hood, C.; Smith, S.; Stocker, J.; Cai, X.; Bloss, W.J. Modelling the impact of national vs. local emission reductions in PM$_{2.5}$ in the West Midlands, UK using WRF-CMAQ. *Atmosphere* 2022, 13, 377. [CrossRef]


45. Choi, H.; Sunwoo, Y. Environmental Benefits of Ammonia Reduction in an Agriculture-Dominated Area in South Korea. *Atmosphere* 2022, 13, 384. [CrossRef]


