RESEARCH & DEVELOPMENT

Increasing the Utilization of Weather Data for Safety Applications and Traveler Information

Venkata R. Duddu, Ph.D.
Srinivas S. Pulugurtha, Ph.D., P.E.
Ajinkya S. Mane, M.Tech.
Christopher M. Godfrey, Ph.D.
Matthew D. Eastin, Ph.D.
Jacob Radford, Ph.D. Student

NCDOT Project 2016-13

FHWA/NC/20YY-NN

August 2017

Increasing the Utilization of Weather Data for Safety Applications and Traveler Information

Draft Final Report

Submitted to

North Carolina Department of Transportation
Research and Development

by

Venkata R. Duddu, Ph.D., E.I.
The University of North Carolina at Charlotte

Srinivas S. Pulugurtha, Ph.D., P.E.
The University of North Carolina at Charlotte

Ajinkya S. Mane, M.Tech.
The University of North Carolina at Charlotte

Christopher M. Godfrey, Ph.D.
The University of North Carolina at Asheville

Matthew D. Eastin, Ph.D.
The University of North Carolina at Charlotte

Jacob Radford, Ph.D. Student,
North Carolina State University

Technical Report Documentation Page

<table>
<tr><td colspan="2">1. Report No.
FHWA/NC/2016-13</td><td>2. Government Accession No.</td><td>3. Recipient's Catalog No.</td></tr>
<tr><td colspan="3">4. Title and Subtitle
Increasing the Utilization of Weather Data for Safety Applications and Travellers Information</td><td>5. Report Date
August 16, 2017</td></tr>
<tr><td colspan="3" rowspan="2"></td><td>6. Performing Organization Code</td></tr>
<tr><td colspan="3">7. Author(s)
Venkata R. Duddu, Srinivas S. Pulugurtha, Ajinkya Mane, Christopher Godfrey, Matthew D. Eastin & Jacob Radford</td><td>8. Performing Organization Report No.</td></tr>
<tr><td colspan="3">9. Performing Organization Name and Address
Infrastructure, Design, Environment, and Sustainability (IDEAS) Center
Department of Civil & Environmental Engineering
The University of North Carolina at Charlotte
9201 University City Boulevard
Charlotte, NC 28223 - 0001
Telephone: 704-687-1233
Fax: 704-687-0957
Email: sspulugurtha@uncc.edu</td><td>10. Work Unit No. (TRAIS)</td></tr>
<tr><td colspan="3"></td><td>11. Contract or Grant No.
NCDOT Project # 2016-13</td></tr>
<tr><td colspan="3">12. Sponsoring Agency Name and Address
North Carolina Department of Transportation
Research and Analysis Group
Raney Building, 104 Fayetteville Street
Raleigh, North Carolina 27601</td><td>13. Type of Report and Period Covered
Final Report
Aug. 16, 2015 – July 31, 2017</td></tr>
<tr><td colspan="3"></td><td>14. Sponsoring Agency Code
NCDOT Project # 2016-13</td></tr>
<tr><td colspan="4">Supplementary Notes:</td></tr>
<tr><td colspan="4">16. Abstract
The objectives of this study are 1) to evaluate the quality of available weather data, 2) to develop weather (fog / visibility) prediction models from historical weather data, 3) to predict weather at the route / link level for safety applications (crash analysis) and traveler information, and, 4) to research and recommend technologies for capturing fog / visibility information. The data collected using a visibility sensor is used to validate both statistical and back propogation neural netwok models, and assess the accuracy of numerical weather prediction models such as High-Resolution Rapid Refresh (HRRR) model and satellite-based model.</td></tr>
<tr><td colspan="2">17. Key Words
Fog, Visibility, Prediction, Model</td><td colspan="2">18. Distribution Statement</td></tr>
<tr><td>19. Security Classif. (of this report)
Unclassified</td><td>20. Security Classif. (of this page)
Unclassified</td><td>21. No. of Pages</td><td>22. Price</td></tr>
</table>

DISCLAIMER

The contents of this report reflect the views of the authors and not necessarily the views of the University of North Carolina at Charlotte (UNC Charlotte), the University of North Carolina at Asheville (UNC Asheville) or the North Carolina Department of Transportation (NCDOT). The authors are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of either UNC Charlotte, UNC Asheville, NCDOT or the Federal Highway Administration (FHWA) at the time of publication. This report does not constitute a standard, specification, or regulation.

Executive Summary

Low visibility and foggy conditions create safety issues for travelers. One way to address this and improve safety on roads is to provide real-time information to travelers using dynamic signs, radio broadcasts, or mobile applications. There is a variety of weather information available to help inform the use of traveler information systems, including the High-Resolution Rapid Refresh (HRRR) numerical model, widely-spaced surface weather observing stations across the state of North Carolina, and satellite observations.

Fog is a highly localized phenomenon and it would be difficult and expensive to install visibility sensors every few miles along roads. In an effort to identify appropriate meteorological data sources or new statistical models that may provide accurate and timely visibility information for North Carolina travelers, this study tackles the following objectives:

- Evaluate the quality and utility of available meteorological data,
- Develop statistical fog/visibility prediction models based on historical weather data,
- Evaluate the accuracy of visibility forecasts at the route / link level for use in safety applications (crash analysis) and traveler information, and,
- Research and recommend technologies for assessing fog / visibility conditions and provide real-time fog related weather data for safety applications and traveler information.

Historical surface observations at airports across North Carolina were used to develop, both, regression models and a backward propagation neural network model in an effort to predict future visibility conditions. The authors also explored the use of short-term, high-resolution dynamical forecast model and visibility maps drawn using both Thiessen polygons and inverse distance-weighted interpolation methods. Additionally, the authors have also applied a novel fog-detection algorithm to generate real-time satellite-based fog products across North Carolina.

A visibility sensor was installed between December 2016 and May 2017 at five different locations to capture visibility conditions in different terrains and climate regions across North Carolina, including UNC Charlotte, Shelby, Lenoir, UNC Asheville, and Wilmington. The data collected using visibility sensor was used to validate both statistical and back propogation neural netwok models, and assess the accuracy of numerical weather prediction models. The report concludes with a discussion on available and needed technologies to help evaluate visibility and provide real-time fog related weather data for safety applications and traveler information.

Table of Contents

1. INTRODUCTION 1
2. LITERATURE REVIEW 4
2.1 Fog Prediction Models 4
2.2 Effect of Fog on Traffic Operations and Safety 10
3. WEATHER DATA SOURCES 15
3.1 Sources and Quality Control of Meteorological Data 15
3.2 Integrated Surface Database (ISD) 16
3.3 North American Regional Reanalysis (NARR) 17
3.4 High-Resolution Rapid Refresh (HRRR) 17
3.4 Satellite Data 19
4. METHODOLOGY 20
4.1 Selection of Data Source for Modeling Visibility 20
4.2 Data Processing 20
4.3 Develop Models 23
4.4 Compare and Validate Data 24
5. VISIBILITY PREDICTION MODELS 25
5.1 Linear Regression Models 25
5.2 Back Propagation Neural Network Models 30
5.3 Visibility at Link level 37
5.3.1 Fog Detection and Fog Depth Algorithm 37
5.3.2 Real-Time North Carolina Satellite Products 40
5.3.3 Visibility Maps using Thiessen Polygons 42
5.3.4 Interpolated Visibility Maps 44
6. VALIDATION 48
6.1 Comparison of Visibility from Visibility Sensor with Nearby Weather Stations 48
6.2 Comparison of Visibility Values with All Data Sources 51
6.3 Validation of Back Propagation Neural Network Models 53
7. TECHNOLOGIES AND LIMITATIONS 54
8. CONCLUSIONS 56
8.1 Implementation Plan 58
References 59
Annexure A 65

List of Tables

Table 1 Physical Reasoning Used to Classify Fog Events (Tardif and Rasmussen, 2007)..... 9
Table 2 Classification of Absolute Percentage Error within the Visibility Range 22
Table 3 Data Collection Schedule 24
Table 4 Descriptive Statistics for Visibility Data <15000m..... 27
Table 5 Classification of Visibility Range as per the International Standards 34
Table 6 Validation of Neural Network Model..... 36
Table 7 Weather Monitoring Stations and HRRR Grid Point from Visibility Sensor 48

List of Figures

Figure 1 Dense Fog Observations by Hour of Day during July 1996-2002 (Friedlein, 2004) 7
Figure 2 Algorithm for the Assessment of Fog Probability (Guidard and Tzanos, 2007) 9
Figure 3 Selected ISD Stations 16
Figure 4 NARR Grid Point Locations 17
Figure 5 Two-hour Forecasted Visibility Obtained from HRRR Forecasts 18
Figure 6 Two-hour Air Temperature Obtained from HRRR Forecasts 18
Figure 7 Average Annual Frequency of Visibility Less Than 1000 m 26
Figure 8 Average Annual Frequency of Visibility Less Than 2000 m 26
Figure 9 Nonlinear Model of an Artificial Neuron 32
Figure 10 Neural Network Model Developed in MATLAB 34
Figure 11 Performance of the Neural Network Model 35
Figure 12 Error Histogram for Training, Testing and Validation Data 35
Figure 13 Regression for Training, Validation, Testing and All Combined 36
Figure 14 Visibility Scenario from the Fog Depth Algorithm (February 23, 2017 at 1130 UTC) 38
Figure 15 Output from Fog Depth Algorithm (April 20, 2017 at 1025 UTC) 39
Figure 16 Output from Fog Depth Algorithm (February 8, 2017 at 1230 UTC) 40
Figure 17 Output from Visible Satellite (May 22, 2017 at 1900 UTC) 41
Figure 18 Output from Fog Depth Estimate (May 22, 2017 at 1900 UTC) 41
Figure 19 Output from Shortwave IR Satellite (May 22, 2017 at 1900 UTC) 42
Figure 20 Output from Longwave IR Satellite (May 22, 2017 at 1900 UTC) 42
Figure 21 Visibility using Thiessen Polygon (December 26, 2014 at 1700 UTC) 44
Figure 22 Visibility – 0.75 Degree Lat/Long Grid (December 26, 2014 at 1700 UTC) 45
Figure 23 Visibility – 0.5 Degree Lat/Long Grid (December 26, 2014 at 1700 UTC) 46
Figure 24 Visibility – 0.25 Degree Lat/Long Grid (December 26, 2014 at 1700 UTC) 46
Figure 25 Visibility – 0.1 Degree Lat/Long Grid (December 26, 2014 at 1700 UTC) 47
Figure 26 Visibility – 0.05 Degree Lat/Long Grid (December 26, 2014 at 1700 UTC) 47
Figure 27 Comparison of Visibility from Visibility Sensor with Nearby Weather Stations 50
Figure 28 Comparison of Visibility Values from Visibility Sensors with Nearest HRRR, Regression Model, and Nearest Weather Station 52
Figure 29 Error Histogram of Visibility from BPNN Model Compared with Visibility Sensor .. 53
Figure 30 Visibility Sensor (Installed at UNC Charlotte) 54

1. INTRODUCTION

Weather is an influential factor which affects road safety. Extreme weather conditions such as dense fog and heavy rain can directly affect the level of safety on a road. All crashes that occur in adverse weather conditions (i.e., rain, sleet, snow, fog, severe crosswinds, or blowing snow / sand / debris) or on slick pavement (i.e., wet pavement, snowy / slushy pavement, or icy pavement) are considered weather-related incidents.

According to the Federal Highway Administration (FHWA), there are over 5,870,000 vehicle crashes each year. About 23% (around 1,312,000) of these crashes are weather-related. On average, 6,250 people are killed and over 480,000 people are injured in weather-related crashes each year. Research on crashes due to adverse weather conditions between 1995 and 2001 showed that nearly 80% of injury crashes occur in rain and 17% in snow or sleet, while fatal crashes in rain, sleet or snow, and fog account for 69%, 18%, and 13%, respectively (Goodwin, 2002).

Visibility is critical to the task of driving and the reduction of visibility due to fog is a major traffic operation and safety concern. Fog presents a challenge to motorists and can result in significant safety concerns due to reduced visibility. According to the FHWA, around 31,500 crashes occur each year due to fog, resulting in over 500 fatalities and 11,500 injury crashes. From 2002 to 2012, there were 19,188 reported fog-related crashes in North Carolina (Oliver, 2013).

According to the National Oceanic and Atmospheric Administration (NOAA; Glickman, 2000), fog consists of a collection of suspended water droplets or ice crystals near Earth's surface that leads to a reduction of horizontal visibility below 1 km (0.62 mile). Petterssen (1956) stated that fog is generally whitish in color. However, in the vicinity of local sources of pollution, it may be of dirty yellowish or grayish color. Furthermore, fog was categorized into three subsections with respect to air temperature (T): (1) Liquid fog ($T > -10°C$), (2) mixed phase fog ($-10°C > T > -30^0C$), and (3) ice fog ($T < -30°C$). Liquid fog consists of small water droplets suspended in the air. The liquid fog may be a mixture of water droplets, smoke, and fine dust, which reduces the horizontal range of visibility to less than 1 km. However, according to World Meteorological Organization (WMO; 1996), if the visibility is greater than 1 km, then the suspended particles are referred to as mist.

Fog is categorized as radiation, inversion, and advection fog (Gultepe *et al.* 2007). Radiation fog typically occurs at night under clear skies with light winds and high relative humidity. Inversion fog forms because of a downward extension of a layer of stratus cloud, situated

under the base of a low-level temperature inversion. Advection fogs occur frequently when moist air moves over a cold surface, often near bodies of water.

Several researchers have studied the effect of fog or smoke on traffic safety and indicated that driving in foggy conditions is a potentially dangerous activity. Some of the most recent research works include efforts by Trick *et al.* (2010), Abdel-Aty *et al.* (2011), Hassan & Abdel-Aty (2013), Mueller & Trick (2013), Ahmed *et al.* (2014), Theofilatos & Yannis (2014), and Yan *et al.* (2014). A detailed description of these studies and other research efforts is presented in the literature review section.

Fog has resulted in some terrible crashes over the years, and drivers are often caught unaware by sudden reductions in visibility (Hamilton *et al.* 2014). Applying fog prediction models or technological advances to evaluate or predict visibility conditions during fog events could help to provide advanced warning to motorists. Additionally, accurate meteorological data are essential for collision analyses to establish contributing factors for collisions. However, crash investigations in rural locations may have to rely on observational data from distant weather stations that may have recorded very different meteorological conditions. It might be possible to improve the reliability of such observations through predictive models or via technologies for establishing weather-related contributing factors for collision analyses.

For motorists, the lack of timely data about potentially foggy conditions poses a significant threat. In particular, these conditions can result in severe injuries as some drivers choose much lower speeds than other drivers. Multiple vehicles (i.e., more than two) may collide as a result of these conditions. Since numerical weather prediction models can help to forecast weather conditions that may impact roads, it seems feasible that the extraction of data from these models could help to reduce weather-related crashes through the dissemination of such information to travelers in a timely manner.

Another issue results from the analysis of collision data when weather data are linked to collision reports. As the distance between a point of interest and the nearest weather station increases, the likelihood of accurately assessing the weather condition at the point of interest decreases. This is particularly true during fog events due to the highly localized nature of the phenomenon. Therefore, a methodology for improved analyses and predictions of high-impact weather phenomena (in particular, fog) at the route level and the efficient and economical use of weather observations is critical. The objectives of this study, therefore, are to:

- Evaluate the quality and utility of available meteorological data,

- Develop weather (fog/visibility) prediction models from historical weather data,
- Predict and evaluate the accuracy of forecasted weather at the route / link level for safety applications (crash analysis) and traveler information, and,
- Research and recommend technologies for capturing fog / visibility information and providing real-time fog related weather data for safety applications and traveler information.

A review of past studies on the effect of fog on transportation safety, relevant predictive models, and various sources of historical and real-time meteorological data is presented in the next chapter.

1.1 Organization of Report

This research report is organized as follows. A detailed literature review on fog prediction models and effect of fog on traffic operations and safety is discussed in Chapter 2. Various weather data sources that are used to evaluate visibility are presented in Chapter 3. Data collection, data processing, and methodology adopted to develop the visibility prediction models are discussed in Chapter 4. Chapter 5 discusses the models developed, while validation of the models developed are presented in Chapter 6. Currently available technologies and limitations are discussed in Chapter 7. Conclusions from this research study are presented in Chapter 8.

2. LITERATURE REVIEW

The literature review is presented in two parts. The first part focuses on previous studies related to fog prediction models using various scientific techniques such as numerical modeling, remote sensing, and Geographic Information Systems (GIS). The second part reviews past studies on the impact of fog on transportation operations and safety.

2.1 Fog Prediction Models

Among the early fog studies, Peace (1969) analyzed the geographical distribution of fog within the United States using 256 first-order weather stations and isopleths in regional analyses. Peace (1969) noted that heavy fog occurred, on average, more than 20 days a year based on the observations from 229 of those weather stations. Here, heavy fog was defined as a visibility of less than one quarter of a mile. Similarly, Hardwick (1973) analyzed the annual and monthly geographic distribution of heavy fog to incorporate monthly distributions of fog parameters in forecasting visibility.

The presence of fog generally corresponds with small dewpoint depressions (i.e., the difference between the air temperature and the dewpoint temperature). When the dewpoint depression is less than roughly 2.5-4.0°C, fog typically forms. However, Baker *et al.* (2002) have modeled the occurrence of fog based on the crossover temperature (T_{xover}), which is equal to the minimum dewpoint temperature observed during the warmest daytime hours in the preceding day. In their study, fog is forecasted when the surface air temperature is expected to cool a few degrees below the crossover temperature, rather than a few degrees below the dewpoint temperature. Baker *et al.* (2002) concluded that if the surface temperature is equal to T_{xover}, their model generally forecasts one- to three-mile visibilities in mist. If surface temperatures are less than or equal to –3°F less than T_{xover}, then the model generally forecasts half-mile visibility or less, under the assumption that turbulent mixing does not prevent fog formation.

Past studies have considered various parameters in fog assessments. Meyer *et al.* (1980) showed that visibility in foggy conditions is a function of droplet number concentration (N_d). In early research studies, due to the unavailability of advanced measuring instruments, continuous measurement of fog characteristics was a difficult task. However, Jiusto (1981) suggested that visibility is a function of both droplet size and liquid water content (LWC), concluding that LWC is directly related to droplet size.

Kunkel (1984) developed extinction coefficients and mean droplet terminal velocities in terms of LWC. In their study, 90 hours of droplet data were recorded at 5 m and 30 m above ground level for 11 life cycles of advection fog. The results obtained indicate that the visibility during foggy conditions is directly related only to LWC and the relationship between visibility and LWC is linear.

Researchers have also used advanced numerical modeling techniques for fog prediction. Bergot *et al.* (2005) studied the integration of observations with a one-dimensional numerical model in an attempt to forecast fog and low clouds at Paris's Charles de Gaulle International Airport, France. Their study results indicate that including boundary layer observations within the model will result in a more accurate short-term weather forecast. Gultepe *et al.* (2007) state that one-dimensional models are the simplest approach for numerical simulation of fog. One-dimensional models are based on the assumption that the thermodynamic variables remain horizontally homogeneous. The occurrence of fog using this method generally considers the atmospheric radiation field, turbulent mixing, and the interaction of moisture and heat at the Earth's surface. However, the actual dynamic processes in nature are not considered in this model and it thus failed to produce reasonable predictions for fog events.

Bott *et al.* (1990) developed a one-dimensional radiation fog model for studying the interactions between atmospheric radiative transfer processes and microphysical processes within fog. The model was used to measure and predict fog events. Bott (1991) extended this work by focusing on land-atmosphere interactions. Different physicochemical properties of aerosols and their impact on fog formation at urban, rural, and maritime locations was investigated.

Gultepe *et al.* (2006) formulated a relationship that expresses visibility as a function of LWC and $N_{d.}$ They developed and tested a mesoscale model using 10-hour forecasts on a 0.62 mile 1-km (0.62 mile) nested domain covering the region within 50 km (31 miles) of the Zurich Unique Airport in Switzerland. They found differences in visibility forecasts between the old and new parameterizations of over 50% and concluded that accurate visibility estimates require appropriate forecasts of both LWC and N_d. Similarly, Gultepe and Milbrandt (2007) measured LWC, N_d, and temperature simultaneously via instrumentation in southern Ontario, Canada. Using the Canadian Mesoscale Compressible Community (MC2) model, they showed that visibility is nonlinearly related to both LWC and N_d.

A few researchers have approached the fog forecasting problem from a probabilistic perspective. Müller *et al.* (2007), for example, used an ensemble of one-dimensional fog

parameterization schemes coupled with several three-dimensional numerical weather prediction models to assess the skill of the one-dimensional fog forecasts. They found that fog forecasts are reasonably accurate when horizontal advection is considered.

Statistical or empirical techniques have also been used to forecast fog in the past. Vislocky and Fritsch (1997) developed linear regression fog prediction models, while Hilliker and Fritsch (1999) developed logistic regression fog prediction models. Marzban et al. (2006) compared an artificial neural network approach with both linear and logistic regression methods to forecast ceiling height and visibility, concluding that the artificial neural network approach produces better fog forecasts. Newer approaches include fuzzy logic, which considers the uncertainty in qualitative variables such as dewpoint (dry, moderate, moist and very moist), spread (unsaturated, saturated and very saturated), rate (drying and saturating), wind (too light, excellent and too strong), etc. (e.g., Sujitjorn *et al*. 1994; Murtha 1995; Hansen 2000), and decision-making processes, which follow a graphical approach that follows a set of rules (e.g., Colquhoun 1987). The application of decision-making processes, however, remains challenging, so this approach in fog and visibility forecasting has been limited.

Meyer and Lala (1989) analyzed the seasonal and diurnal timing as well as the five primary synoptic setups leading to radiation fog in Albany, New York. They found that initial relative humidity and nocturnal cooling rate are key parameters for predicting the onset of fog. Further, they noted that radiation fog occurs under clear skies with wind speeds under 1 m s^{-1} as temperatures drop from 2°C to 12°C. Similarly, Friedlein (2004) analyzed dense radiation and advection fog events at Chicago O'Hare International Airport, Illinois. Though fog is prevalent at all hours during winter, analyses indicate that fog occurred most often during the nighttime, predawn, and dawn hours of all seasons. As an example, Figure 1 shows the breakdown of dense fog observations by hour of the day during July 1996–2002. Wind speed was observed to be a major influential factor for fog occurrences, with 84% of fog events occurring at wind speeds of 7 knots (8 mph) or less.

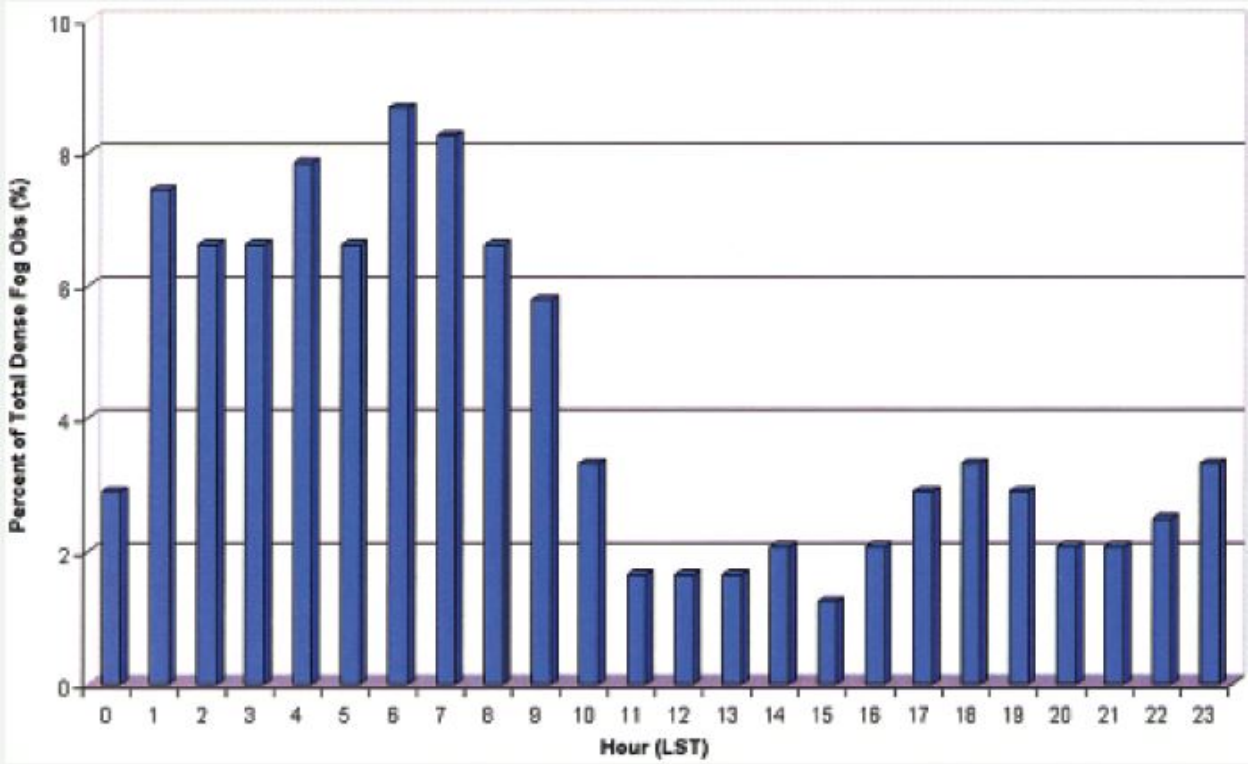


Figure 1 Dense Fog Observations by Hour of Day during July 1996-2002 (Friedlein, 2004)

Ward and Croft (2008) studied fog occurrences in the New York metropolitan area using GIS. They considered 15 fog events during a single winter season and assessed the meteorological conditions as measured by both Automated Surface Observing System (ASOS) sites and satellite imagery. They replicated conditions favorable for fog within an ArcGIS environment by using inverse distance-weighting interpolation techniques.

Several authors have considered satellite data as a fog-detection technique. Ellrod (1995) developed a technique for detection of fog and low clouds at night using multispectral infrared (IR) imagery from Geostationary Operational Environmental Satellites (GOES). Ellrod (1995) used two GOES IR window channels at 10.7-11.2 μm and 3.9 μm to estimate the thickness of ground fog, noting that a patch is fog that is generally thicker in its interior compared with its boundary. A regression relationship between the thickness of fog based on aircraft pilot reports and the brightness difference between two IR windows (11.2 and 3.9 μm) was then developed. It was observed that the two variables are highly correlated with a correlation coefficient of 0.94, though this method may not be able to differentiate between ground fog and altocumulus layers. Similarly, Ellrod and Gultepe (2007) studied low clouds at night using infrared (IR) imagery from GOES, concluding that low ceilings and visibilities tend to occur when the difference between the

surface temperature and the GOES 10.7 μm IR cloud top brightness temperature is less than 3°C. Additionally, Whiffen *et al.* (2004) studied fog-related motor vehicle crashes in Canada and assessed the feasibility of using multispectral satellite imagery for fog detection. They also provided a comprehensive list of potential mitigation strategies to reduce fog-related crashes.

Tardif and Rasmussen (2007) studied the characteristics of fog events in the New York City area, using 20 years of climatological data to classify fog types into five categories with an objection categorization algorithm (Table 1). They defined a fog occurrence as one in which the observed visibility is less than 1.6 km (1 statute mile) during fog, ground fog, or ice fog. They found that the marine environment may affect fog events and, in spring, advection fog tends to occur near coastal areas while radiation fog tends to occur in rural and suburban areas. Additionally, the urban heat island effect may reduce the probability of fog formation. This study presents a clear case where climatological data can help to improve the understanding of fog formation and forecasting methods. Additionally, Guidard and Tzanos (2007) assessed fog probability based on a combination of surface observations and satellite data. It was concluded that a lack of cloud cover, rainfall rates less than 0.2 mm hr^{-1}, 10-meter wind speeds greater than 7 m s^{-1}, and 2-meter relative humidity values below 90% are not compatible with the presence of fog. An algorithm that evaluates the risk of the presence of fog as proposed by Guidard and Tzanos (2007) is shown as Figure 2.

Table 1 Physical Reasoning Used to Classify Fog Events (Tardif and Rasmussen, 2007)

Fog type	Primary physical mechanism	Definition/morphology of fog formation	References
Precipitation (PCP)	Thermodynamical influence of evaporating precipitation	Precipitation observed at the onset of fog or the hour prior	Petterssen (1969)
Radiation (RAD)	Radiative cooling over land	Onset during the night with an observed wind speed below 2.5 m s^{-1} and cooling during the hour prior in absence of a cloud ceiling, or with a cloud base rising concurrently or slight warming in the hour leading to onset if preceded by cooling period (e.g., fog forming between two hourly observations) or cloud ceiling below 100 m if followed shortly by fog at surface (e.g., elevated fog formation)	Taylor (1917), Pilié et al. (1975), Roach et al. (1976), Roach (1995a), Meyer et al. (1986), Meyer and Lala (1990), Baker et al. (2002)
Advection (ADV)	Shear-induced mixing of air parcels of contrasting temperatures as moist, warm air flows over a colder (water, land, or snow) surface	Onset as a "wall" of fog reaches a station, with an observed wind speed greater than 2.5 m s^{-1} and an associated sudden decrease in visibility or sudden appearance of a cloud ceiling below 200 m, followed by fog onset within next 2 h	Roach (1995b), Baars et al. (2003)
Cloud-base lowering (CBL)	Moistening and/or cooling of the layer below boundary layer stratiform clouds and/or prolonged subsidence	Gradual lowering of a cloud ceiling within a 5-h period prior to fog onset, with initial ceiling height below 1 km	Petterssen (1940), Oliver et al. (1978), Pilié et al. (1979), Duynkerke and Hignett (1993), Koračin et al. (2001), Baker et al. (2002)
Morning evaporation (EVP)	Evaporation of surface water and mixing in the surface layer	Increasing temperature and greater increase in dewpoint leading to saturation, within 1 h of sunrise	Arya (2001), Brutsaert (1982)*

* These references do not specifically discuss the formation of fog under the influence of evaporation at sunrise, but rather present comprehensive discussions on the evaporation of water at the earth's surface.

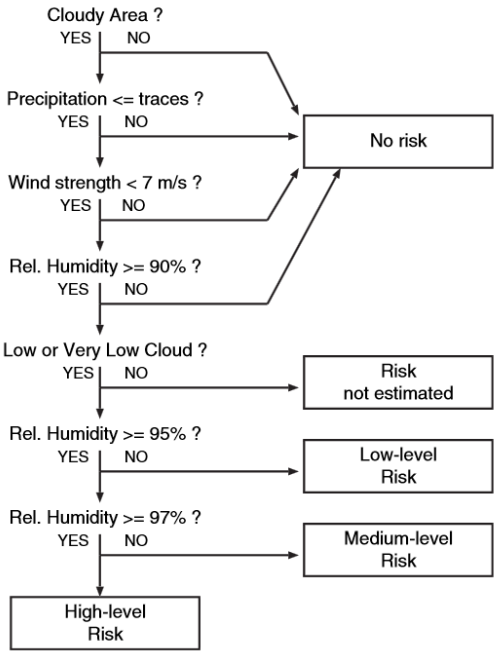


Figure 2 Algorithm for the Assessment of Fog Probability (Guidard and Tzanos, 2007)

Tardif and Rasmussen (2008) also analyzed meteorological factors and scenarios leading to the occurrence of precipitation fog in the New York City area. Their study indicates that 18%

of the analyzed precipitation events corresponded with fog events and that the majority of fog events occurred during light liquid precipitation. Most of the fog events occurred at high elevation stations due to upslope flow and attendant lowering of the cloud base. Since relative humidity is a function of temperature, they divided fog events into those that occurred due to moistening, cooling, moistening and cooling, or static conditions. An analysis of all fog events based on these tendencies indicate that moistening, cooling, moistening and cooling, and static conditions were observed for 42%, 25%, 10%, and 23% of the fog events, respectively.

2.2 Effect of Fog on Traffic Operations and Safety

According to a study in 1999 (Chin *et al.*, 2002), the Oak Ridge National Laboratory has estimated that fog, snow and ice reduces the capacity of United States freeways and principle arterials by more than 11%. They also projected that nearly 544 million vehicle-hours of delay, or 23% of total delay, was caused by these weather events. In 1999, 13% of large truck crashes happened in rain, snow, sleet, hail, or under foggy conditions. Weather events even impact the productivity of the transportation system. Further, winter road maintenance accounts for 24% of road operating costs.

Goodwin (2003a) studied the 2001 crash database of the National Highway Traffic Safety Administration (NHTSA) as well as the Fatality Analysis Reporting System (FARS). The study found that 75%, 18%, and 5% of the crashes occurred in rain, snowfall, and foggy conditions, respectively. Around 2% of crashes occurred during some combination of adverse weather conditions like sleet, rain, and fog. Out of the total number of crashes, 16% resulted in fatalities.

Pisano *et al.* (2008) studied weather-related crashes in the United States between 1995 and 2005. The investigation focused on driver behavior, crash risk, and regional variations in crash types and severity, along with the economic impacts of adverse road weather conditions. It was found that out of all weather-related crashes, 75% of those crashes occurred on wet pavement and that fog is involved in 2% of all crashes. The estimated annual cost of all weather-related crashes is between $22 billion and $51 billion. The study also suggested mitigation measures such as the installation of motorist warning systems, weather-related signal timing, access control, and speed management during adverse weather conditions.

Goodwin (2003b) described mitigation measures adopted in 21 states by using 30 case studies in the United States. Advisory, control, and treatment strategies are generally adopted during adverse weather conditions to improve safety as well as mobility. The report also provided

a detailed description of real-time weather data collection techniques using environmental sensor technologies, along with roadway condition monitoring, to improve safety.

Kang *et al.* (2008) examined the effect of reduced visibility due to fog on car-following performance. They observed that distance headway decreased in the densest fog conditions, root mean square (RMS) velocity error increased with an increase in fog density, and drivers had greater difficulty responding to changes in lead vehicle speed under foggy conditions.

The Utah Department of Transportation (UDOT) provided the public with a "Weather Responsive Traveler Information System" and documented the implementation, findings and lessons learned from the implementation (Gopalakrishna *et al.*, 2013). Their study was mainly aimed at encouraging agencies to be more proactive in managing traffic during inclement weather conditions. The UDOT developed an application that was driven by a citizen reporting system where individuals would report road weather condition with the primary goal to provide travelers with more accurate road weather forecasts and information on travel impacts. The public response to this project was quite positive.

Pisano and Goodwin (2004) studied the effect of weather on various transportation systems. They documented various practices that are currently adopted in the United States to manage traffic under adverse weather conditions, including motorist warning systems and control strategies. Analyzing travel time data and surface weather, Pisano and Goodwin (2004) estimated a 10-12% reduction in average freeway speeds due to low visibility. On the other hand, some motorists respond to poor visibility conditions by slowing down, while others do not. While the general recommendation was to maintain larger headways during poor visibility, many drivers seem to simply follow the taillights of the vehicle ahead.

Travel information plays a significant role in reducing crash risk on roads. Al-Ghamdi (2007) evaluated the performance of a fog detection and warning system installed on a two-lane rural road in southern Saudi Arabia. This warning system includes visibility sensors that automatically activate a variable message sign that posts an advisory speed when hazardous conditions are present. The study involved the collection of vehicle speed, traffic volume, vehicle classification, headway, time of day, and visibility. Activation of the warning system reduced the mean speed in the experimental road section by about 6.5 k.p.h., but the mean speed still remained higher than the posted advisory speed. Drawing from relationships between mean driving speed and the number of crashes, they noted that a speed reduction of only 5 k.p.h. would yield a 15% decrease in the number of crashes.

Abdel-Aty *et al.* (2011) performed a comprehensive analysis of crashes related to visibility obstructions by comparing fog- and smoke-related crashes to clear visibility crashes. They found that lighting conditions adversely affected fog- and smoke-related crashes and that foggy and smoky conditions pose a more deadly threat in terms of crash severity. The odds ratio is as high as 3.24 when compared to clear visibility conditions. By considering variations in crash severity through an assessment of multiple contributing factors, it was found that head-on and rear-end crashes are the most significant crashes under foggy and smoky conditions.

Shi and Tan (2013) analyzed the use of intermittent release measures in heavy fog with an improved cellular automation model in China. They studied the risk of rear-end collisions in fog and the intermittent release measures that are taken to reduce such risk. They developed a cellular automation model that considers driving behaviors in heavy fog. Two main driving behaviors were observed: 1) few drivers may become laggers in the queue as the space headway is greater than the visibility in heavy fog, and, 2) some drivers feel risky when driving in fog and speed to reduce space headways. The study revealed that traffic crash probabilities increase under heavy fog conditions compared with normal weather conditions.

Balagh *et al.* (2014) developed a stochastic approach to model road crashes during winter in an effort to predict crashes based on driving conditions. Since most road safety prediction models are based on deterministic weather forecasts that are not able to capture changes in the likelihood of collision occurrence, this study used a probabilistic forecast model to improve the decision-making process. They developed a logistic regression model based on crash data to examine its ability to predict collision occurrence and compared it with various models that are available in the literature. The model outperformed the existing models and accurately predicted collision occurrences.

Ahmed *et al.* (2014) conducted a feasibility study to examine whether or not meteorological observations collected at airports can be used for real-time road crash risk assessment at locations with recurrent fog problems. Typically, automated sensors at airport locations collect observations at least hourly and those observations are available in near real time. The meteorological observations from eight airports in Florida were paired with crash data in fog-prone counties to show that reductions in visibility as reported at an airport are statistically related to nearby crash occurrences. It was concluded that real-time meteorological data from airports can be used for nearby roads (i.e., within 5 miles) to mitigate the increased risk of limited visibility.

Departments of transportation (DOTs) in several states have implemented warning systems in the past to reduce visibility-related crashes. Following a 193 vehicle crash on the I-10 Bay Bridge, Alabama DOT traffic managers implemented an intelligent transportation system (ITS) that detects fog on the Bay Bridge and other segments of I-10, automatically alters speed limits with variable speed limit signs, and modifies the lane configuration via lane control signs. In the state of California, the California Highway Patrol groups traffic into platoons when visibility is less than 500 feet and leads traffic through affected areas at a safe speed. Similarly, following an analysis of a 95-vehicle pileup on I-77 at the Virginia-North Carolina state line, Oliver (2013) suggested that information sharing across all the DOTs in the state and to travelers via media outlets could help to reduce the number of fog- and weather-related crashes.

Ashley *et al.* (2015) studied the weather-related crash database within the FARS from 1994 to 2011. The study focused on visibility-related (VR) crashes where smoke, fog, or blowing dust were observed at the time of the incident and weather-related vision-obscured (VO) crashes where a driver's vision was obscured by weather and a weather-related vision hazard was reported. Spatiotemporal analyses of these crash types revealed that most fatal crashes occurred during the morning commute, during the cool season, and on state and US routes, while 72% of those fatal crashes occurred without a National Weather Service visibility-related advisory in effect.

Mueller and Trick (2012) studied driving behavior under foggy conditions by considering variables such as a driver's experience, average speed, speed variability, steering variability, and other factors. Under controlled conditions in driving simulators, participants were divided into groups of experienced and novice drivers based on license type and years of driving experience. They found that 25% of young novice drivers had collisions in foggy conditions, with higher speeds and steering variability pegged as the primary variables contributing to collisions.

Yan *et al.* (2014) evaluated the speed control behavior of motorists under foggy conditions by categorizing road condition into three different scenarios of low risk, medium risk, and high risk. The low risk scenario corresponds with simple driving skills related to road alignment such as speed control, braking, steering, staying in the lane, and other vehicle control behavior. The medium risk scenario corresponds with skills related to car following, overtaking, route choice in an unfamiliar road network, and other factors. The high-risk scenario corresponds with a driver's emergency speed response. With the addition of foggy conditions, drivers tend to reduce their speed in all three scenarios. More rear-end collisions occurred in the medium-risk and high-risk scenarios, while crash severity is reduced in the high-risk scenario due to slower speeds.

Theofilatos and Yannis (2014) examined the effect of traffic and weather characteristics on road safety. The results obtained indicate that the presence of precipitation led to higher crash rates, though it does not seem to increase crash severity. It was suggested that the use of real-time meteorological data may address gaps in the current literature, such as in studies covering urban and rural areas rather than major arterial roadways.

Hassan and Abdel-Aty (2013) explored predicting real-time reduced visibility-related crashes on freeways using random forests technique (a data mining technique) and a logistic regression model. They considered factors such as vision obstruction (e.g., fog or heavy rain), crash type, lighting condition, and crash location. The results obtained indicate that higher occupancy rates downstream 10-15 minutes prior to a crash, coupled with an increase in the average speed downstream and upstream 5-10 minutes before the crash, increase the likelihood of a visibility-related crash occurrence. The model was able to predict 69% of crashes correctly, despite having no knowledge of driver factors and errors.

Trick *et al.* (2010) researched the behavior of older drivers using a driving simulator under varying conditions of visibility (e.g., a clear day or fog), traffic density, and navigational challenges. Based on speed adjustment indices, they found that reduced visibility, increased traffic density, and navigational challenges reduce both driving speed and driving performance, noting that the interaction of these challenges produced a stronger response than any individual challenge. However, the combined effects of high-density traffic and navigational challenges reduced driving performance even further in foggy conditions compared with clear conditions.

3. WEATHER DATA SOURCES

This chapter presents various weather data sources, such as historical data, real-time meteorological observations, and short-term forecasts, that are used to evaluate visibility condition. The data sources include the Integrated Surface Database (ISD), the North American Regional Reanalysis (NARR), and High-Resolution Rapid Refresh (HRRR) forecasts.

3.1 Sources and Quality Control of Meteorological Data

Historical meteorological data are critical for the development of models that could assist with short-term fog prediction. The ISD database from the NOAA/National Centers for Environmental Information (NCEI) contains hourly surface observations for over 20,000 locations across the world (Del Greco *et al.* 2006; Smith *et al.* 2011). This dataset contains a merged repository of both manual and automated surface observations from a variety of original data sources, including the ASOS, Automated Weather Observing System (AWOS), surface synoptic observations, and aviation routine weather report (METAR) observations (Smith *et al.* 2011). The database includes hourly observations of 2-m air temperature, dewpoint temperature, precipitation, wind speed, atmospheric pressure, visibility, cloud cover, and present weather conditions. Some stations also report snowfall and snow depth. Before distribution, NCEI applies rigorous quality control procedures to the ISD data (Lott 2004; Smith *et al.* 2011), yet data quality problems still remain in the database (Godfrey 2015). Such data problems can be handled effectively through additional quality-assurance algorithms.

Airport observations provide a good source for real-time meteorological conditions, but the spatial coverage across North Carolina is poor, and each set of observations is only representative of a very small region immediately surrounding the airport (Godfrey 2015). High-resolution numerical weather prediction models, therefore, can provide short-term forecasts of deterministic variables for use as input to historically-derived logistic regression models in near-real time on small spatial grids.

The HRRR model (Pinto *et al.* 2015) is a real-time, 3-km horizontal resolution, hourly-updated, cloud-resolving, and convection allowing numerical weather prediction model that assimilates radar data every 15 minutes. The HRRR model incorporates a variety of observational data through three-dimensional variational data assimilation (3DVAR) techniques. The HRRR output fields include both the horizontal and vertical dimensions. Therefore, it provides the best picture of the current state of the atmosphere across the entire continental United States.

3.2 Integrated Surface Database (ISD)

The ISD is a comprehensive database consisting of information from more than 35,000 weather stations worldwide. Some of them date back to 1901. A total of 14,000 stations are active and collect information such as visibility, wind speed, wind direction, temperature, dewpoint, sea level pressure, station pressure, precipitation, altimeter setting, and cloud data. For this study, 238 ISD locations in and near North Carolina from January 1, 1979 to January 1, 2015 was collected and processed for analysis and modeling. Figure 3 shows the selected ISD stations in and around North Carolina.

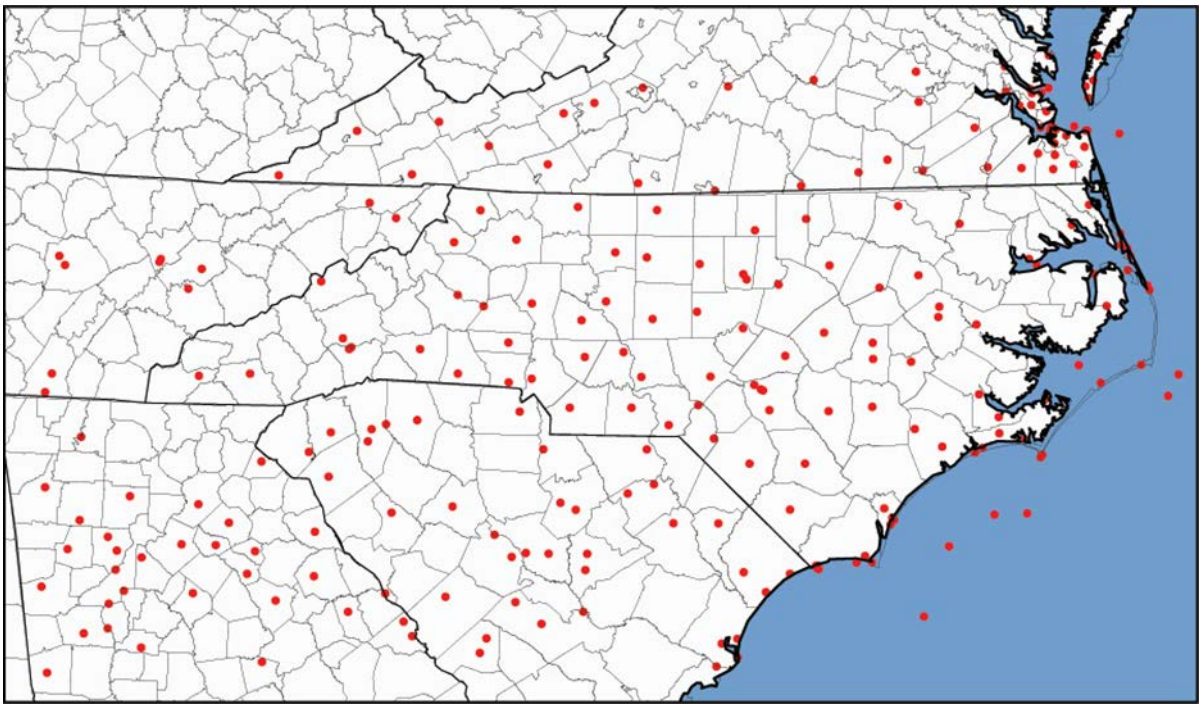


Figure 3 Selected ISD Stations

3.3 North American Regional Reanalysis (NARR)

The NARR assimilates surface and upper-air observations into a 32-km forecast model to produce the most reasonable state of the atmosphere every three hours. Output variables for each grid (Figure 4) include air temperature and dewpoint temperature at 2-m above ground level (AGL), 2-m relative humidity, 10-m zonal (east-west) and meridional (north-south) wind speed, accumulated precipitation, total cloud cover, surface visibility, ceiling height, air temperature at 10-m AGL, air temperature at 30-m AGL, soil temperature in the 0-10 cm and 10-40 cm layers, and moisture availability in the 0-100 cm layer. Even though NARR data is based on a forecast model, the weather data for a given time stamp can only be obtained after the time. For this study, data were available from January 1, 1979 to January 1, 2015. Since the data from NARR were not available for every hour, it was not further explored in the modelling process.

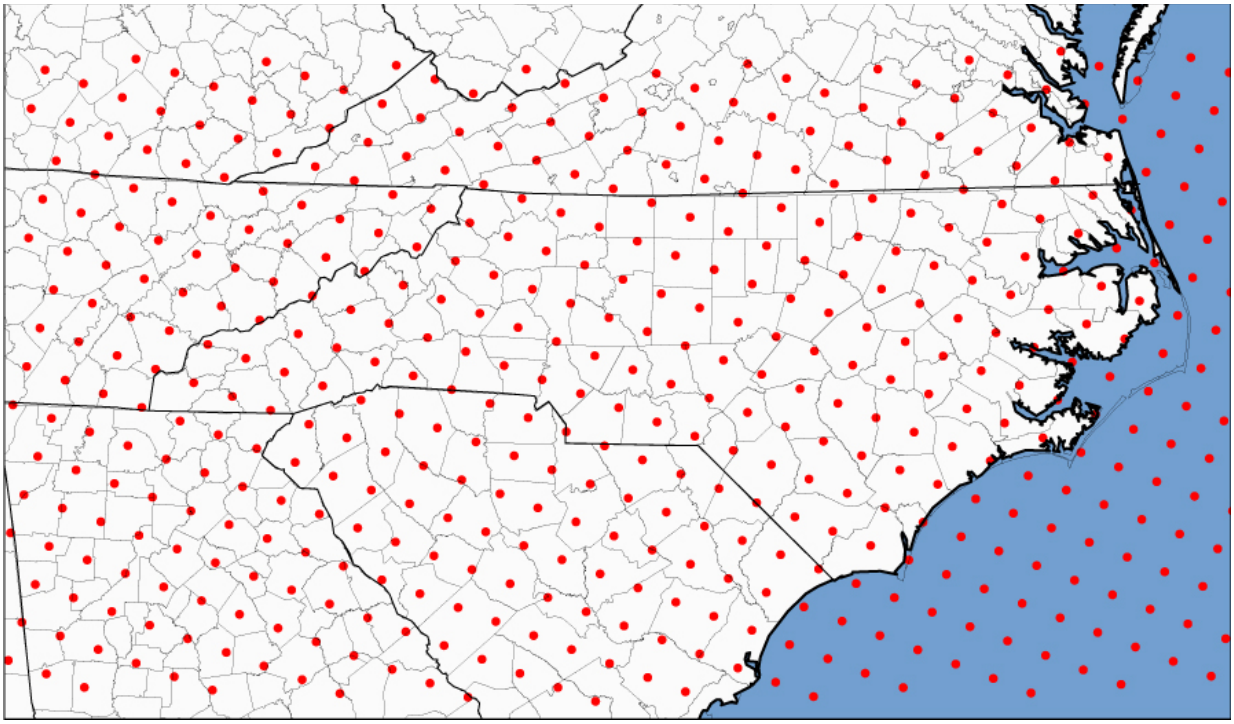


Figure 4 NARR Grid Point Locations

3.4 High-Resolution Rapid Refresh (HRRR)

The HRRR model is a NOAA real-time 3-km resolution model with 3-km radar data assimilation that produces short-term forecasts and are updated hourly. Forecast fields used in the present study include 2-m air temperature, 2-m dewpoint temperature, 2-m relative humidity, 10-m zonal and meridional wind speeds, accumulated precipitation, total cloud cover, surface visibility, ceiling

height, skin temperature, moisture availability, and downward shortwave radiation. Figures 5 and 6 show examples of a two-hour forecast of surface visibility and air temperature in and around the state of North Carolina, respectively.

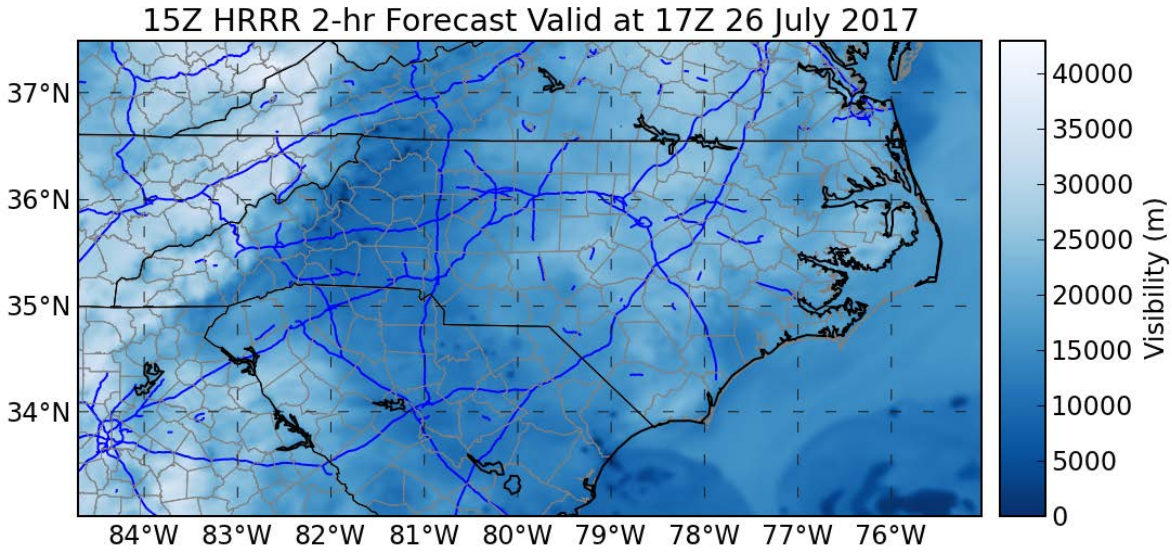


Figure 5 Two-hour Forecasted Visibility Obtained from HRRR Forecasts

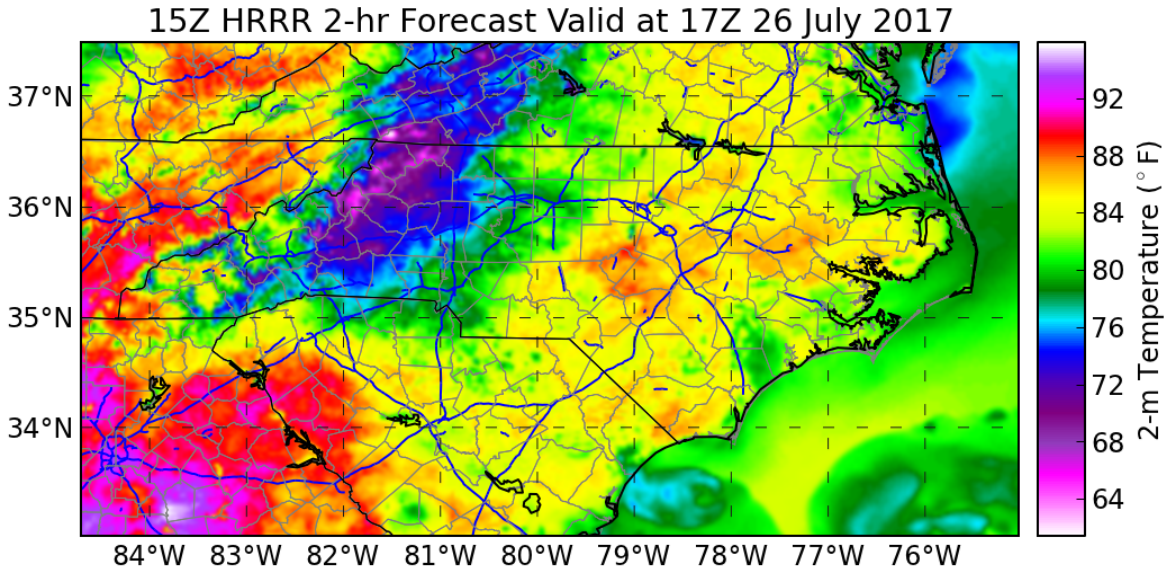


Figure 6 Two-hour Air Temperature Obtained from HRRR Forecasts

3.4 Satellite Data

Satellite data are the novel dataset used to identify fog at any location. Fog depth algorithm and real-time fog depth imagery are the two products developed by NOAA to continuously capture the metrological data. In the fog depth product, information from GOES is primarily used to detect fog or low clouds. From this information, the temperature difference between Infrared Band 2 (3.9 μm) and Band 4 (10.7 μm) helps detect low clouds / fog. In addition, the real-time fog product is the most recent product to detect visibility for every one hour. Visibility is updated at 1-km resolution during daylight and at 4-km resolution at nighttime. A detailed description of satellite data sources is explained in section 4.3.

4. METHODOLOGY

The methodology adopted for analysis and modeling is described in this chapter.

4.1 Selection of Data Source for Modeling Visibility

ISD data source is based on the data collected at the weather stations. It is an actual representation of weather condition at a specific location. However, weather stations are generally located at important locations only. HRRR is a forecasting model where different weather parameters are forecasted at grid points that are spaced distinctly. Further, NARR data source is a 32-km (medium resolution) model that ingests actual observations for every three hours. However, ISD and HRRR data sources provide the information for every one-hour interval. NARR data source was not considered since it only provides information for every three hours.

To select the data source among ISD and HRRR for the development of visibility prediction models, it is necessary to compare the information obtained from these two different data sources. Visibility values obtained from both the data sources (ISD and HRRR) were compared using Absolute Percentage Error (APE).

$$\text{Absolute Percentage Error} = \left|\frac{Actual - Predicted}{Actual}\right| \times 100 \quad (4.1)$$

In Equation (4.1), weather station data (ISD) was considered as the actual value and HRRR model output was considered as the predicted value.

ISD and HRRR data from July 6, 2016 to August 15, 2016 was compared to check the accuracy of HRRR dataset. Table 2 indicates the classification of APE while comparing ISD and HRRR data sources. The results obtained indicate that majority of the errors were less than <50% (moderately high) and between the range of 15,000 m and 2,000 m. In addition, ISD represents the actual weather condition at a specific location. Therefore, ISD dataset was used for visibility prediction model development.

4.2 Data Processing

Data processing is an important step before the development of visibility prediction models. The data was processed to remove any missing values and outliers from the data. Further, raw data was

acquired in comma separated file format and size of each database was more than 500MB. Hence, Microsoft SQL Server 2016 was used for data processing.

Firstly, ISD dataset was imported and the database was generated in SQL server. Secondly, all the missing values and outliers were deleted from the database by using SQL Server queries. Moreover, literature review suggested that rainfall in previous hours and time-of-the-day could influence the occurrence of low visibility conditions. Therefore, two sets of dichotomous variables: rainfall and time-of-the-day were added to the database. Dichotomous variables, such as rainfall in past three hours (Rain 3hr), three to six hours (Rain 6hr), six to twelve hours (Rain 12hr), and twelve to twenty-four hours (Rain 24hr) were added to the database. Further, to address time-of-the-day, six dichotomous variables, such as, time from 12 AM to 4AM (Time 4am), 4 AM to 8 AM (Time 8am), 8 AM to 12 PM (Time 12pm), 12 PM to 4 PM (Time 4 pm), 4 PM to 8 PM (Time 8 pm), and 8pm to 12 AM (Time 12 am) were added to the database.

Similarly, the presence of water body within the vicinity of weather station may influence the occurrence of fog. Therefore, all the water bodies in the state of North Carolina was acquired and imported in ArcGIS shapefile format using latitude and longitude coordinates. To check the presence of water body within the vicinity of the weather station, 1-mile buffer was generated around each weather station and the generated buffers were intersected with the water body shapefile. Finally, in the database, a dichotomous variable named as “Water” was created to represent the presence of water body within a mile of weather station. Final dataset consists of hourly information of all the weather stations in the state of North Carolina and also the additional variables such as rainfall, time-of-the-day and the presence of water body.

In addition to the above database, one more database was generated by selecting only the weather stations which are located within a mile of road network (state routes, US routes and interstates). As one of the objectives of this study is to inform the road users about the low visibility condition, selecting the weather station within a mile of roadway network could help improve the performance visibility prediction models.

Table 2 Classification of Absolute Percentage Error within the Visibility Range

Range	#Sample	Absolute Percent Error													Sum (%)
		<5	5 - 10	10 -20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100	100-200	200-300	>300	
43000-40000	14	-	-	-	-	-	-	-	-	-	-	-	0.07%	-	-	0.07
40000-35000	65	-	-	-	-	-	-	-	-	-	-	-	0.32%	-	-	0.32
35000-30000	322	-	-	-	-	-	-	-	-	-	0.26%	0.67%	0.63%	0.01%	0.01%	1.58
30000-25000	804	-	-	-	-	-	-	0.74%	1.43%	1.12%	0.60%	-	0.04%	-	0.01%	3.94
25000-20000	1217	-	-	-	1.22%	1.87%	1.88%	0.82%	-	0.01%	0.01%	0.01%	0.06%	0.02%	0.06%	5.97
20000-15000	1989	3.70%	2.24%	2.45%	0.84%	0.06%	0.08%	0.04%	0.02%	0.03%	0.01%	0.02%	0.10%	0.05%	0.11%	9.76
15000-10000	6438	0.30%	1.83%	6.34%	9.92%	12.23%	0.06%	0.07%	0.15%	0.06%	0.05%	0.01%	0.22%	0.10%	0.23%	31.59
10000-5000	8028	1.03%	0.30%	1.98%	3.21%	10.47%	17.61%	1.58%	1.16%	0.12%	0.08%	0.07%	0.59%	0.21%	0.99%	39.39
5000-3000	469	0.03%	0.02%	0.03%	0.05%	0.02%	0.11%	0.13%	0.36%	1.29%	0.15%	0.00%	0.03%	0.02%	0.04%	2.30
3000-2000	195	0.00%	0.00%	0.01%	0.01%	0.02%	0.03%	0.01%	0.04%	0.09%	0.67%	0.01%	0.01%	0.01%	0.02%	0.96
2000-1000	206	0.00%	0.01%	0.02%	0.01%	0.01%	0.02%	0.01%	0.05%	0.04%	0.34%	0.46%	0.01%	0.01%	-	1.01
1000-0	558	0.01%	0.00%	0.01%	0.01%	0.01%	0.03%	0.02%	0.03%	0.07%	0.13%	2.35%	0.00%	-	0.04%	2.74
0	77	-	-	-	-	-	-	-	-	-	-	0.37%	-	-	0.00%	0.38
Total # Sample	20382															100.00

4.3 Develop Models

Fog is a highly localized phenomenon. Therefore, predictions and nowcasts for the occurrence of fog are extremely difficult without direct observations. Variables such as air temperature, humidity, pressure, dew point temperature, cloud cover, cloud height, and wind speed were used to develop a statistical prediction model (linear regression models) and forecast the likelihood of fog and/or low visibility at any location across North Carolina. The models incorporated hourly weather observations from the nearest geocoded stations, high-resolution topographic maps of the study area, and prescribed distributions of near-surface atmospheric lapse rates (which will account for temperature changes resulting from differences in elevation between a given location and its nearest weather station) to predict visibility. Also, similar to the regression models, the neural network models were also developed for predicting the range of visibility.

To evaluate visibility at any given location (point or route), the geocoded weather station data was used to divide the entire study area (North Carolina) into Thiessen or proximal zones. This process was done by generating Thiessen polygons, by triangulating the point input features (location of weather stations) into a triangulated irregular network (TIN) that meets the Delaunay criterion. Each Thiessen polygon consists of only one weather station. The developed Thiessen or proximal polygons indicate that any location (route or point) within a Thiessen polygon is closer to its associated weather station than to any other weather station in the proximity. The weather (fog/visibility) observed at a given weather station can be assigned for all the routes in the polygon associated with the weather station. This method could help evaluate the presence of fog or reduced visibility at any given point in the entire state of North Carolina, given that data from all weather stations in the state are available. However, creating proximity zones and assessing fog or visibility accurately at the route / link level could be cumbersome and would lead to inaccurate results if the number of weather stations is observed to be too low or if data are missing.

Given that many weather station sites do not report hourly observations continuously without some gaps, it is more likely that spatial interpolation methods would be necessary to evaluate weather condition at locations distant from a weather station site. Kriging is the most reasonable geospatial interpolation method, provided adjustments are made to the temperatures using observed or modeled low-level lapse rates. Spatially-interpolated meteorological variables was combined with prediction models to improve the accuracy of weather data for safety analysis

and traveler information. Analysis using the proximal zones was compared with analysis from an interpolation method to evaluate the effectiveness of both methods.

4.4 Compare and Validate Data

To evaluate the accuracy of the visibility prediction models (by comparing the predictions with actual observations at various locations), a visibility sensor was purchased. Five different locations that are geographically distributed across the state of North Carolina were selected for data collection; 1) UNC Charlotte Campus, Charlotte, NC, 2) Lincoln County Maintenance Yard (NCDOT), 3) Salt Bin, Caldwell County Maintenance Yard (NCDOT), 4) UNC Ashville Campus, and 5) NCDOT Division Office at Wilmington. The locations were selected such that they cover piedmont, mountainous and coastal areas. The visibility sensor was installed at each location and the data was retrieved after a minimum of 20 days. Table 3 shows the schedule of data collection at all the five locations.

The data collected from the visibility sensors was processed to compare and evaluate the accuracy of developed models (regression models and neural network models), ISD dataset, and HRRR dataset.

Table 3 Data Collection Schedule

Location	Date Installed	Date Uninstalled	Total Number of Days
UNC Charlotte Campus	12/15/2017	2/2/2017	47
Lincoln County Maintenance Yard (NCDOT)	2/2/2017	2/23/2017	21
Salt Bin, Caldwell County Maintenance Yard (NCDOT)	2/23/2017	3/23/2017	27
UNC Ashville Campus	3/23/2017	4/13/2017	21
NCDOT Division Office at Wilmington	4/20/2017	5/10/2017	21

5. VISIBILITY PREDICTION MODELS

This chapter presents various models that are developed to predict visibility. Four different methods were adopted to predict the visibility. They are:

- Linear regression models
- Neural network models
- Thiessen polygon method
- Geographical interpolation method

Each of these methods and their results are discussed in detail in the following sections.

5.1 Linear Regression Models

Regression-based prediction models were developed using weather data obtained from weather stations as it represents the actual scenario of weather at different locations. Four years of weather data from January 2011 to December 2014 was used for model development. Two sets of regression models were developed. The first set of regression model was developed based on the data obtained from all weather stations in the state of North Carolina, while the second set of regression models were developed based on the data obtained from weather stations that are located within a mile of the road network in North Carolina. The visibility values were observed to have a huge range (0 m to 15,000 m) and were also observed to vary significantly with elevation. Figure 7 and figure 8 show the average annual frequency of visibility less than 1000 m and 2000 m observed at weather stations in North Carolina, respectively. From the figures, it can be observed that the frequency of low visibility condition (< 41 times in year) is very low in most parts of North Carolina. However, in mountainous regions the frequency of low visibility conditions is observed to be higher compared other areas. Therefore, multiple regression models were developed based on visibility (<15,000 m, <10,000 m, <5,000 m, and <2,000 m) and elevation (<50 m, 50 - 250 m, 250 - 750 m, and >750 m). Overall, twenty different models were developed for each range of visibility considering all stations data and data from weather stations near the road network separately. In this process, data was first classified based on the different ranges of visibility and was further classified based on the respective elevations.

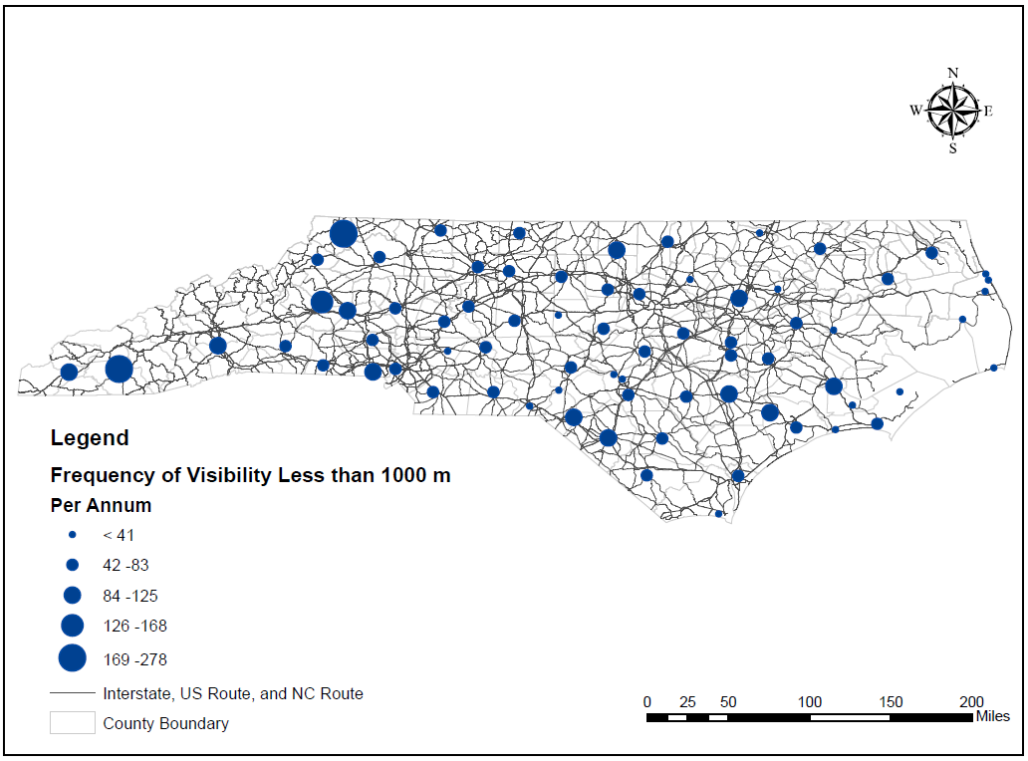


Figure 7 Average Annual Frequency of Visibility Less Than 1000 m

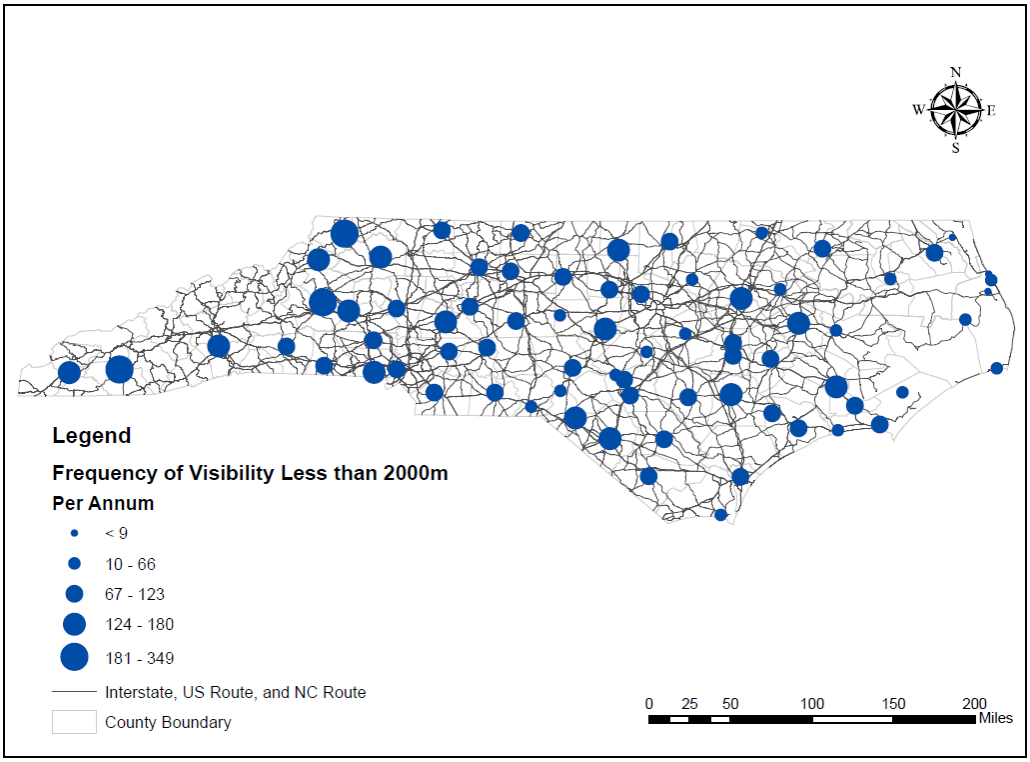


Figure 8 Average Annual Frequency of Visibility Less Than 2000 m

The methodology includes the following steps:

1. Selection of predictor variables contributing to visibility
2. Development of regression models

5.1.1 Selection of Predictor Variables Contributing to Visibility

The descriptive statistics for the weather data with visibility less than 15,000 m is presented in Table 4. Variables such as elevation, cloud cover, speed cover, precipitation, and a difference of air temp and dew point temperature were considered as the independent / predictor variables for model development. Along with these variables, two dichotomous or binary variables such as rainfall and time-of-the-day were considered as independent variables in the model development. In addition, the presence of water within a mile of weather station was captured using ArcGIS and considered as a binary variable.

Table 4 Descriptive Statistics for Visibility Data <15000m

Variables	Mean	Std. Dev.	Min	Max	Variable Type
Visibility	8,718.46	4003	0	14,484	Continuous
Elevation	183.86	200.3	0	969.3	
Cloud Cover	53.15	46.55	0	100	
Wind Speed	1.55	2.02	0	26.3	
Precipitation	0.3135	1.61	0	78.5	
Air Temperature (–) Dew Point Temperature	1.8569	2.87	0	54	
Time 4 am	-	-	0	1	Dichotomous
Time 8 am	-	-	0	1	
Time 12 pm	-	-	0	1	
Time 4 pm	-	-	0	1	
Time 8 pm	-	-	0	1	
Time 12 am	-	-	0	1	
Rain 3 hr	-	-	0	1	
Rain 6 hr	-	-	0	1	
Rain 12 hr	-	-	0	1	
Rain 24 hr	-	-	0	1	
Presence of Water Bodies	-	-	0	1	

5.1.2 Development of Linear Regression Models

Ordinary Least Squares (OLS) and Weighted Least Squares (WLS) regression models were developed using STATA considering data for all weather stations in North Carolina and weather stations near the road network; twenty regression models based on visibility and elevation ranges. All the developed linear regression models are presented in Annexure A. Independent variables with a level of significance (p-value) less than 0.05 (at a 95% confidence level) were considered to have a statistically significant effect on the dependent variable (visibility). Statistical measures such as R-square, Adjusted R-square, AIC, and Root Mean Square Error (RMSE) were computed to evaluate the performance of the models developed.

5.1.2.1 Visibility Less Than 15,000 m

OLS and WLS regression models were developed for predicting visibility less than 15,000 meters. A total of five OLS regression models were developed. The first model was developed by considering all the data. Other four models were developed for different elevations (<50 m, 50 m to 250 m, 250 m to 750 m, and >750 m) above Mean Sea Level (MSL). The results obtained indicate that predictor variables such as elevation, cloud cover, and precipitation are negatively associated with the visibility (Table A1 in Annexure A). From Table A1, the negative coefficient for elevation indicates that the visibility decreases by ~ 76m for every 100 m increase in elevation. Similarly, the visibility decreases by 183 m with 1 mm increase in precipitation, while the visibility decreases by 32 m with 1% increase in cloud cover. All other predictor variables that are considered in the model were observed to be positively associated with visibility. The positive coefficient for wind speed indicates that the visibility increases by ~230 m for every 1m/s increase in wind speed. This could be attributed to boundary-layer mixing during higher wind speeds resulting in reduced humidity leading to good visibility conditions. Likewise, rainfall in past three hours, three to six hours, and twelve to twenty four hours are positively associated with visibility indicating that past rainfall during these time ranges will not have an effect on visibility conditions. However, rainfall during past six to twelve hours is observed to be negatively associated with visibility.

WLS models were also developed for predicting visibility less than 15,000 meters (Table A2). Similar to OLS models, five WLS regression models were developed. The first model was developed by considering all the data. Other four models were developed for different elevations

(Elevation <50 m, 50 m to 250 m, 250 m to 750 m, and >750 m). Table A2 shows the WLS models to predict visibility considering all data. Similar to OLS models, the WLS models also indicate that the predictor variables such as elevation, cloud cover, and precipitation are negatively associated with visibility and remaining all other variables are observed to be positively associated with the visibility. Moreover, the coefficient of predictor variables in WLS models is approximately consistent with respective OLS models. Even though the statistical measures of performance for both types of models are acceptable, the WLS model with an R-squared value of 0.99 is observed to be better compared to the OLS model with R-squared value of 0.86. Also, adjusted R-squared value is observed to be higher for WLS model (0.99) when compared to OLS model (0.86). In addition, AIC and Root Mean Square Error (RMSE) are found to be lower for the WLS model compared to the OLS model. Similar observations were noted in case of the other four models for different elevations. Therefore, WLS models were preferred over OLS models in further analysis.

Similarly, OLS and WLS regression models were developed by considering the weather station data within 1 mile of the road network (Table A3 and Table A4, respectively). Approximately in all the models, the sign of the coefficients was found to be consistent when compared with the all-weather station models (Table A1 and Table A2 in Appendix A). The AIC value decreased substantially for WLS regression models that are developed by considering the weather stations near to roads as compared to WLS regression models developed by considering data from all the weather stations in the State of North Carolina. This indicates that WLS regression models may yield better estimates when weather station data near to roads were considered for model development. However, RMSE increased up to 50% when predicting visibility.

Similar to the earlier models, models were developed by considering all the data with visibility less than 10,000 m, 5,000 m, and 2,000 m. Likewise, models were developed for different elevations (elevation less than 50 m, between 50 m to 250 m, between 250 m to 750 m, and greater than 750 m) for each range of visibility. In addition, OLS and WLS regression models were also developed by considering the weather stations within 1 mile of the road network. Overall, sixteen additional models were developed as shown in Appendix A.

In all the developed OLS and WLS regression models, cloud cover was observed to be negatively associated with visibility (Appendix A). Further, in most of the OLS and WLS regression models where elevation, precipitation, and water were found to be significant, they were

observed to be negatively associated with the visibility. Conversely, in developed regression models where wind speed was found to be significant; it was positively associated with the visibility. Also, “tair_dew” (difference between air temperature and dew temperature) variable was found to be positively associated with the visibility in the majority of models. However, for some of the lower range visibility models (visibility < 5,000 m and < 2,000 m), “tair_dew” variable was observed to be negatively associated with the visibility. In all the visibility prediction regression models, even though the coefficient of predictor variables was different, the sign of coefficients was observed to be consistent. The total number of observations considered for developing regression models for lower visibility conditions (<2,000 m and <5,000 m) was observed to be very low compared to sample size for other models (<15,000 m visibility). Therefore, to eliminate any bias and to have an improved model performance, WLS regression model developed considering sample from all weather station with visibility <15,000 m (Table A2) was used for validation purpose.

5.2 Back Propagation Neural Network Models

Artificial intelligence (AI) is an ability of a system that can independently perform tasks normally requiring human or animal intelligence (Nilsson, 1971). It is a system with an ability to learn, adapt and improve. The first known AI system is “Turing Machine” invented by Allen Turing in 1950. In the subsequent years, the research in the field of AI has grown rapidly and is sub-divided into many different areas based on their applicability to various fields in science and technology. Some of the applications that are commonly used in the field of civil engineering include Expert Systems, Genetic Algorithms, Intelligent Agents, Neural Networks, Logic Programming, and Fuzzy Logic. Each of the above-mentioned applications are used based on the type of problem to be addressed.

Neural networks were chosen to develop the methods and models in this report. A brief description of neural networks and how it effectively helps solve the problem are discussed in the following sections.

5.2.1 Neural Networks

Artificial neural networks, also called as neural networks, is a computational model that mimic at least partially the structure and functions of brains and nervous systems of living beings (Cichocki & Unbehauen, 1993). In general, a neural network is a computational model composed of simple

processing elements called neurons or nodes, which are interconnected by links with weights that help perform parallel distributed processing in order to solve a desired problem. Neural networks have the ability to learn from the environment and to adapt to it in an interactive manner similar to their biological counterparts.

The interest in the neural networks has grown dramatically in the fields of science and engineering in the last few years. A basic neural network model consists of set of nodes connected by the links that has numeric weights associated with them. Each node has a set of input links from other nodes and a set of output links to other some nodes. The nodes from the input links are connected to an activation function to compute the activation level at the next time step.

Figure 9 shows a model of an artificial neuron. The inputs to the artificial neuron are expressed in the form individual vector components, given as x_i, for i = 1, 2, 3…, n. The entire input is given as a vector signal $x \in \Re^{n\times 1}$, where, $x = [x_1, x_2, x_3, \ldots, x_n]^T$. Each input neuron '$x_i$' is connected to the neuron 'q' through a link called synapse, which is associated with a synaptic weight 'W_{qi}'. The neuron 'q' receives an input from 'x_i' as the product of the individual input vector component 'x_i' and the weight 'W_{qi}' associated with it. Since, there are multiple inputs to the neuron 'q', all these inputs are multiplied with their respective synaptic weights and then summed as $u_q = \sum_{i=1}^{n} W_{qi}x_i$. The threshold or bias (-ve of threshold) θ_q is externally applied, usually to lower the cumulative input to the activation function. The activation function shown in the Figure 9 helps define the output 'y_q' for a given input 'u_q'. From Figure 9, the output of the neuron 'q' can be written as, $y_q = f(v_q) = f(u_q - \theta_q) = f(\sum_{i=1}^{n} W_{qi}x_i - \theta_q)$. For no threshold scenario, $y_q = f(v_q) = f(u_q) = f(\sum_{i=1}^{n} W_{qi}x_i)$. The above explanation on neural networks is based on the work by Ham & Kostanic (2001).

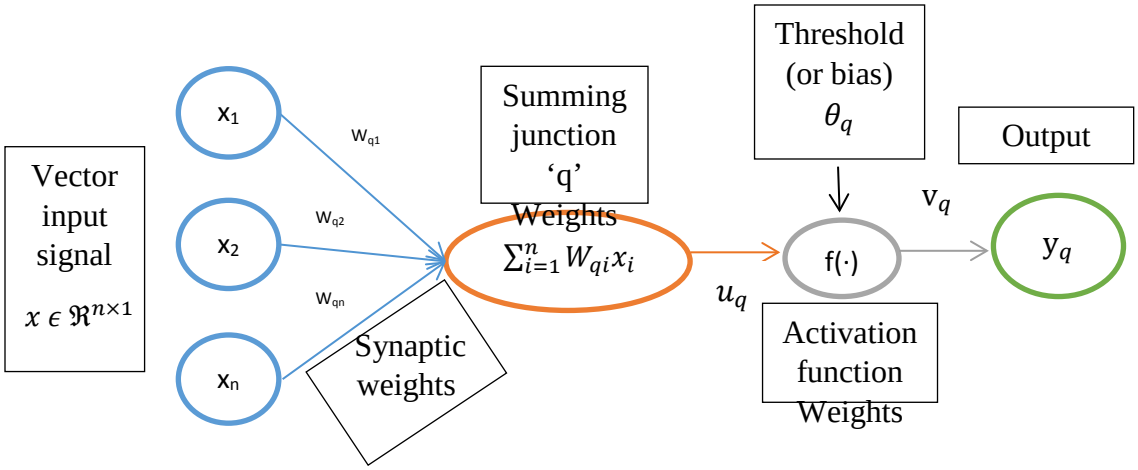


Figure 9 Nonlinear Model of an Artificial Neuron

One of the most popular neural networks is a layered neural network with a back-propagation (BP) learning algorithm where the weights are adjusted based on least mean square error of the output. Neural networks could also be used for prediction purposes by using BP architecture. The use of BP architecture in the neural networks let the network learn an approximation of mapping (pattern) between inputs and outputs by updating its synaptic weights along with error minimization in order identify the implicit rules and relationship between the inputs and outputs. Along with the prediction, neural networks can also be used for various other applications such as optimization, forecasting, associative memory, function approximation, clustering, data compression, speech recognition, non-linear system modeling and control, pattern classification, feature extraction, solutions to matrix algebra problems and differential equations (Ham & Kostanic, 2001).

A multi-layer feed-forward back-propagation network using the BP learning algorithm, which can also be referred to as a back-propagation neural network (BPNN) model, was developed. Typically, a BPNN model consists of three layers: input layer, hidden layer and output layer.

The independent variables used in the statistical models have to be given as the input vector to the network. Therefore, the input layer has the number of neurons each of which corresponds to an independent variable in the model. The output layer has the number of neurons equal to the number of outputs. For example, the output layer of the model has a neuron each indicating the visibility as the dependent variable. Tangent sigmoid function and purelin function are used as transfer functions for hidden layer and output layer with Back-propagation function as training

function. BP is used to calculate the Jacobian 'J'. Each variable is adjusted according to Levenberg-Marquardt as follows (MATLAB, 2016):

$$\text{Error } (E) = \frac{1}{n}\sum_{i=1}^{n}(\hat{c}_i - c_i)^2 \quad (5.1)$$

$$\text{Error gradient } (g) = J^t E \quad (5.2)$$

$$\text{Hessian Matrix } (H) = J^t J \quad (5.3)$$

$$\text{Cost Function } (C) = \beta * E_d + \alpha * E_w \quad (5.4)$$

where, E_d and E_w are sum-squared errors and sum-squared weights, respectively.

$$(H + \lambda \mathrm{I})\delta = g \quad (5.5)$$

where, $\hat{c}_i$ and c_i are predicted and observed visibility values, 'I' is an identity matrix, β *and* $\alpha$are scale parameters, and λ is the damping factor and is adjusted based on sum-squared errors.

Equation (5.5) is solved to calculate 'δ' and the weights are updated based on the value of 'δ'. MATLAB was used to build the BPNN and train the network. The BPNN was trained such that the error term shown in Equation (5.1) and the cost function shown in Equation (5.2) are minimized. The readers can refer to Liang (2003), Liang (2005), MATLAB (2016) and Xie et al. (2007) for more detailed description of Bayesian regularization, Levenberg-Marquardt algorithm, and BPNN.

The BPNN was trained using the data from all weather stations in North Carolina from 2011 to 2014. All the variables obtained from station data such as elevation, relative humidity, wind speed, precipitation, month of the year, presence of water bodies, cloud cover, ceiling height, time-of-the-day (divided into 4 hour intervals), rainfall in past 3 hours, rainfall in past 3-6 hours, rainfall in past 6-12 hours and rainfall in past 12-24 hours are considered as input variables for the regression. The dependent variable (visibility) was divided into eight different classes (Table 5) based on the visibility range, as per the international standards:

The number of neurons in the hidden layers is not limited to a fixed number. Therefore, appropriate number of neurons was evaluated by changing the number of neurons in the hidden layer until the network performs well after training. Figure 10 shows the neural network used to train the data. The network has two hidden layers with 10 and 5 neurons in each layer, respectively.

Table 5 Classification of Visibility Range as per the International Standards

Visibility	Description
< 40 m	Dense Fog (1)
40 to 200 m	Thick Fog (2)
200 to 1000 m	Fog (3)
1 to 2 km	Mist/Haze (4)
2 to 4 km	Poor Visibility (5)
4 to 10 km	Moderate Visibility (6)
10 to 40 km	Good visibility (7)
> 40 km	Excellent visibility (8)

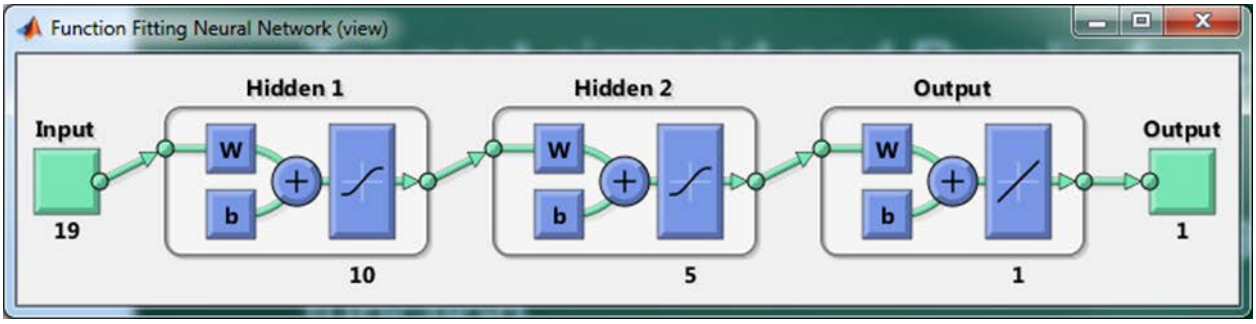


Figure 10 Neural Network Model Developed in MATLAB

The performance of the network was evaluated by computing the mean squared errors. Figure 11 shows the performance of the neural network model. Figure 12 shows the error histogram for training, testing and validation data. Most of the errors are observed to be within ±1 indicating, very low errors in predicting the visibility from the model developed. Figure 13 shows the regression for training, validation, testing and all combined. The R-squared value is observed to be 0.73 for all the four cases. The higher R-squared value indicates better predictive capability of the develop neural network model.

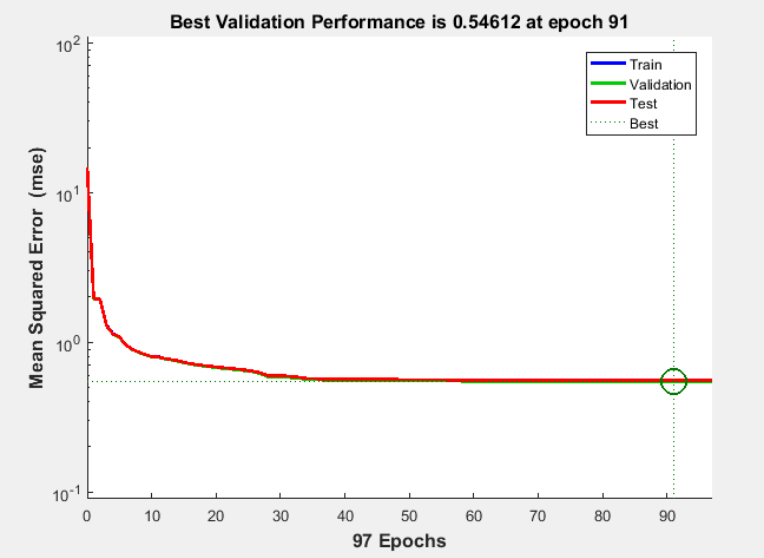


Figure 11 Performance of the Neural Network Model

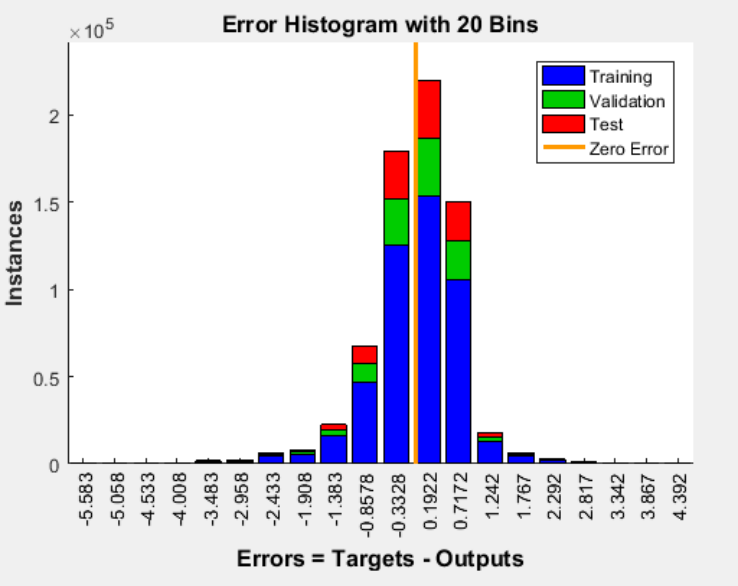


Figure 12 Error Histogram for Training, Testing and Validation Data

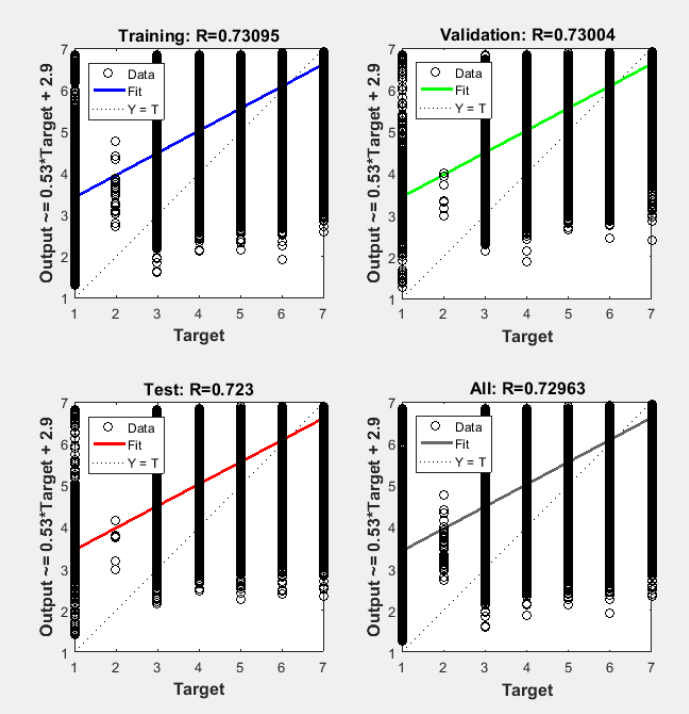


Figure 13 Regression for Training, Validation, Testing and All Combined

The validation of the developed neural network model was performed using the data for years 2010 and 2015. Table 6 shows validation performance of the neural network model developed. From Table 6, the neural network model is observed to have better predictive capability for higher visibility conditions. However, for lower visibility conditions, it is observed to predict the most nearest range indicating a very low error in prediction. For example, in case of visibility range 40 to 200 m, 60% of the times the model has predicted a range on 200 to 1,000 m instead. Similarly, for less than 40 m visibility conditions, the model has predicted 40 to 200 m and 200 to 1,000 m, 7% and 93% of times, respectively.

Table 6 Validation of Neural Network Model

Absolute Error				
Visibility	**0**	**1**	**≥ 2**	**Total**
10 to 40 km (7)	90%	10%	0%	100%
4 to 10 km (6)	57%	41%	2%	100%
2 to 4 km (5)	18%	72%	10%	100%
1 to 2 km (4)	22%	38%	40%	100%
200 to 1000 m (3)	48%	37%	15%	100%
40 to 200 m (2)	0%	60%	40%	100%
< 40 m (1)	0%	7%	93%	100%

5.3 Visibility at Link level

This section provides four different algorithms/methods that were explored to identify the visibility at the route / link level: 1) satellite-based fog depth algorithm, 2) real-time fog depth imagery, 3) visibility maps based on Thiessen polygons, and, 4) visibility maps based on interpolation.

5.3.1 Fog Detection and Fog Depth Algorithm

The NOAA / NESDIS Forecast Products Development Team has developed an experimental satellite-based fog detection algorithm that produces fog depth color enhancement based on the temperature difference between GOES IR Band 2 (3.9 μm) and Band 4 (10.7 μm). This algorithm identifies the approximate depth of low-level clouds and fog. Weinrab and Han (2011) developed equations to convert GOES Variable Format (GVAR) infrared data to both scene radiance and brightness temperature. In this study, an algorithm to produce real-time fog detection imagery and associated data files based on GOES-13 satellite imagery over North Carolina was adopted.

The following examples illustrate the applicability of the fog detection product. The color steps show increasing fog depth, from green (0 - 200 m) to black (> 500 m). Light blue areas are cirrus clouds and gray areas are cloud-free. The fog depths obtained from this product are valid only for single cloud layers. After 1500 UTC, (i.e., during daylight hours), the fog detection algorithm can no longer identify fog due to the strong influence of reflected solar radiation on the radiance measured in GOES IR Band 2 (shortwave infrared).

Example 1

On the morning of February 23, 2017, dense fog was noted in the French Broad River valley near Arden, North Carolina in Buncombe County. Figure 14 indicates that fog with a depth of less than 200 m is present in the valley, with deeper pockets of fog from I-40 in Buncombe County northward through Madison County. The 1144 UTC observation at the nearby Asheville Regional Airport (KAVL) corroborates this satellite-derived estimate. The ASOS station reports calm winds, fog, visibility of ~ 402 m (0.25 miles), and an indefinite ceiling through the fog with a vertical visibility of ~ 61m (200 feet).

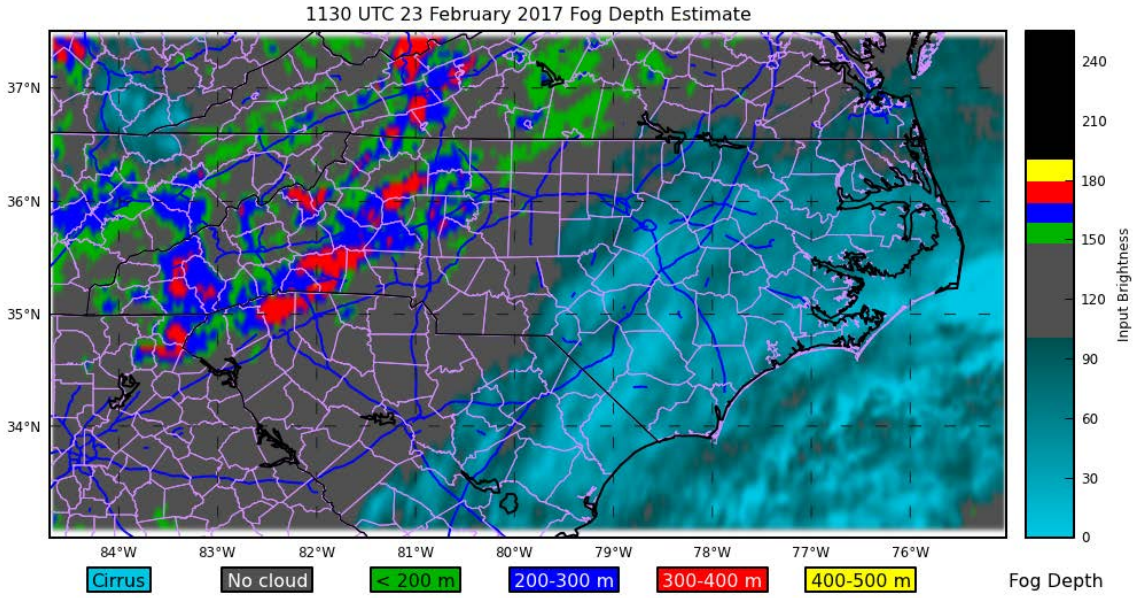


Figure 14 Visibility Scenario from the Fog Depth Algorithm (February 23, 2017 at 1130 UTC)

Example 2

On the morning of April 20, 2017, the National Weather Service issued a dense fog advisory for much of the Piedmont region of North Carolina and portions of the western mountains. The fog product (Figure 15) indicates widespread fog throughout the region. The Charlotte Douglas International Airport (KCLT) is within the dark blue region in the plot. The 1014 UTC ASOS observations here indicate fog, visibility of 1/4 statute miles, calm winds, and an indefinite ceiling through the fog with a vertical visibility of 100 feet. The Piedmont Triad International Airport (GSO) near Greensboro is in the green region of the plot. The 1027 UTC observation here indicates a 5-knot (5.758 mph) wind from 250°, visibility of 1.5 statute miles, mist (i.e., light fog), and overcast skies at 200 feet.

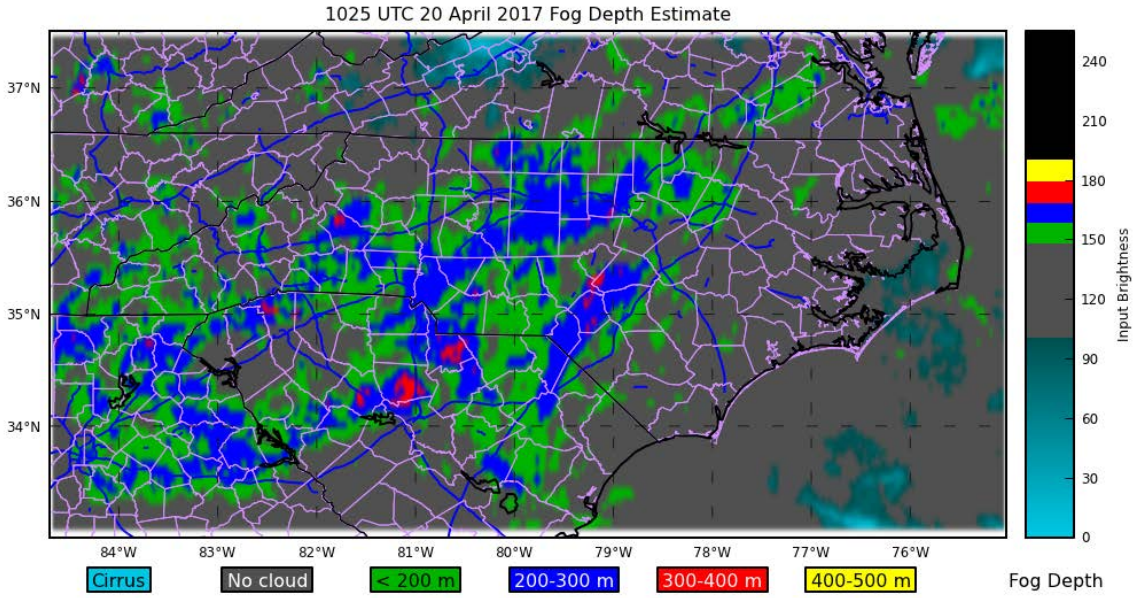


Figure 15 Output from Fog Depth Algorithm (April 20, 2017 at 1025 UTC)

Example 3

Dense fog was noted in the French Broad River valley near Arden, North Carolina on the morning of February 8, 2017. The 1246 UTC observations at the Asheville Regional Airport (KAVL) indicate calm winds, visibility of 1/4 statute miles, and fog with vertical visibility of 100 feet. Figure 16 shows the fog product for this event and indicates that cirrus clouds over Buncombe County are obscuring the low-level fog. This example shows that the fog product is only valid for single cloud layers. The 1235 UTC observations at the Smith-Reynolds Airport (KINT) near Winston-Salem, however, also indicate a reduced visibility of 1.75 statute miles in mist (i.e., light fog) and a ceiling of 200 feet. The fog depth product, Figure 16, indicates a low-level cloud depth of 300-400 meters at this location.

As a whole, fog detection product is unique and a useful product for locating early morning, evening, and night time fog with high spatial resolution.

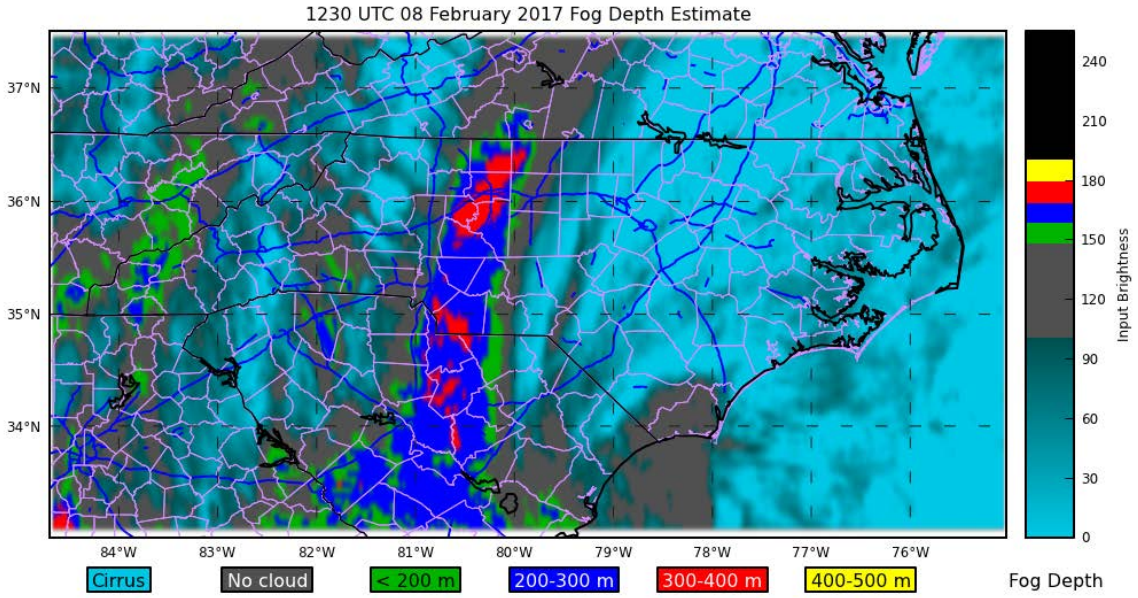


Figure 16 Output from Fog Depth Algorithm (February 8, 2017 at 1230 UTC)

5.3.2 Real-Time North Carolina Satellite Products

Imagery can be produced in near real time from GOES-13 GVAR (GOES Variable Format) data. In the present experimental form, images (examples, Figure 17, Figure 18, Figure 19 and Figure 20) are available approximately every hour. Visible imagery at 1-km resolution is updated only during daylight hours. Infrared imagery is available at 4-km resolution. Note that the fog depth product is not valid during daylight hours after 1500 UTC. The most recent satellite data is just feed into to generate the fog product. The visible imagery combined with the infrared imagery is useful for noting the presence of low-level clouds. However, information from these products alone cannot provide anything definitive about surface fog during daylight hours.

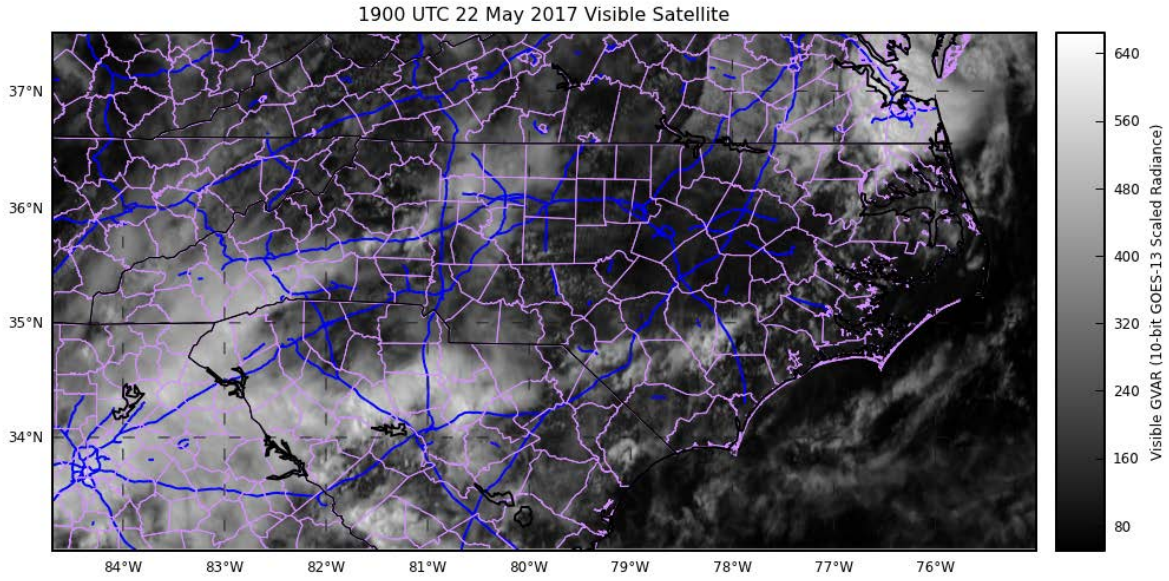


Figure 17 Output from Visible Satellite (May 22, 2017 at 1900 UTC)

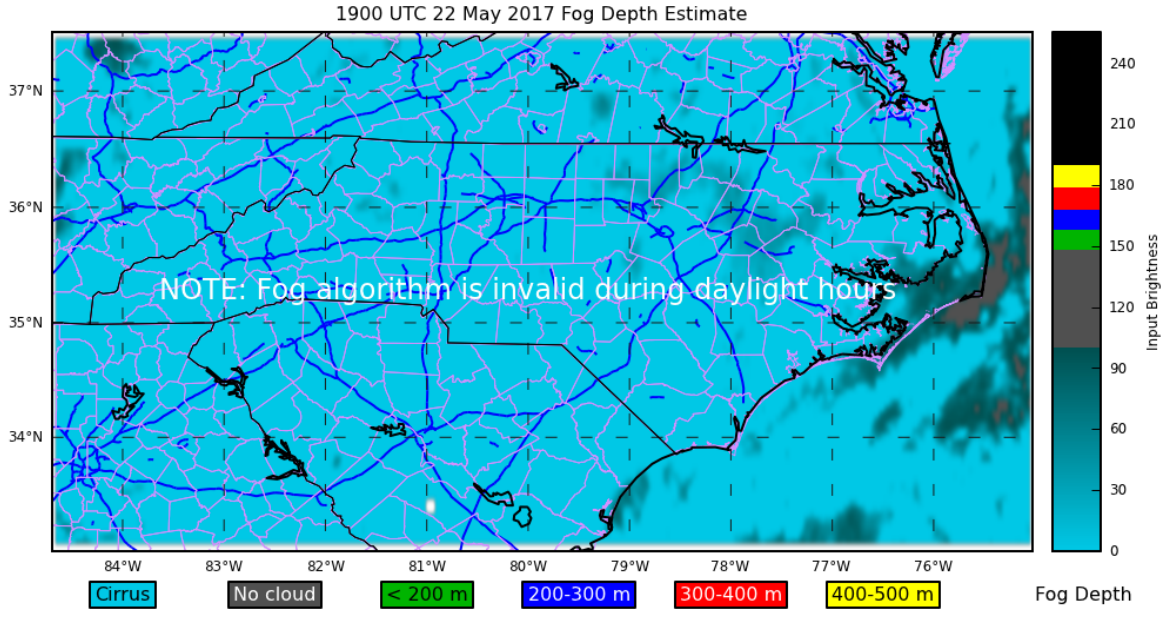


Figure 18 Output from Fog Depth Estimate (May 22, 2017 at 1900 UTC)

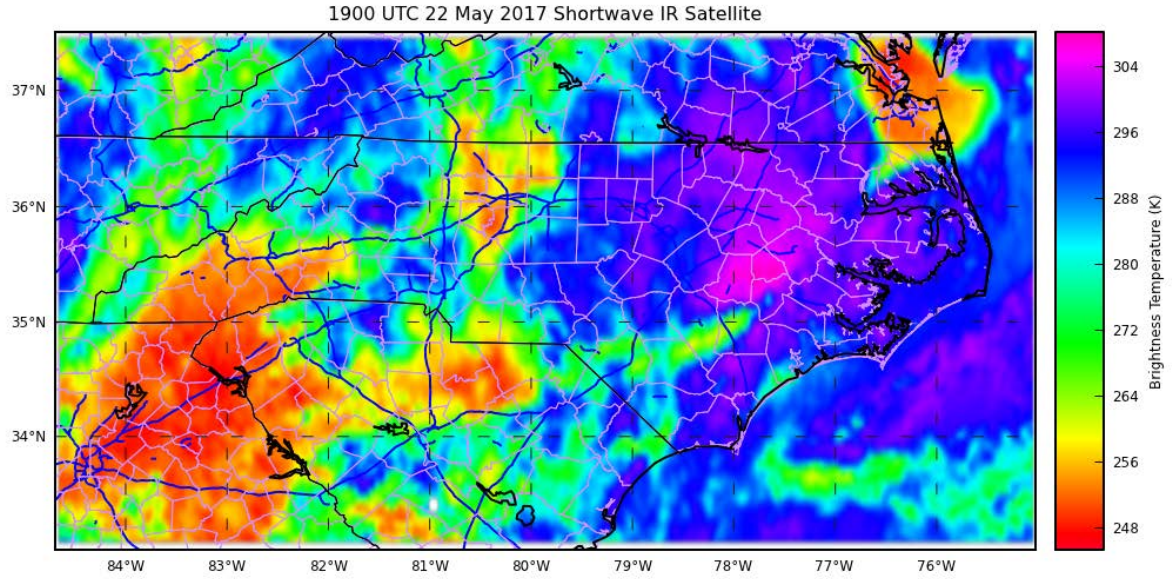


Figure 19 Output from Shortwave IR Satellite (May 22, 2017 at 1900 UTC)

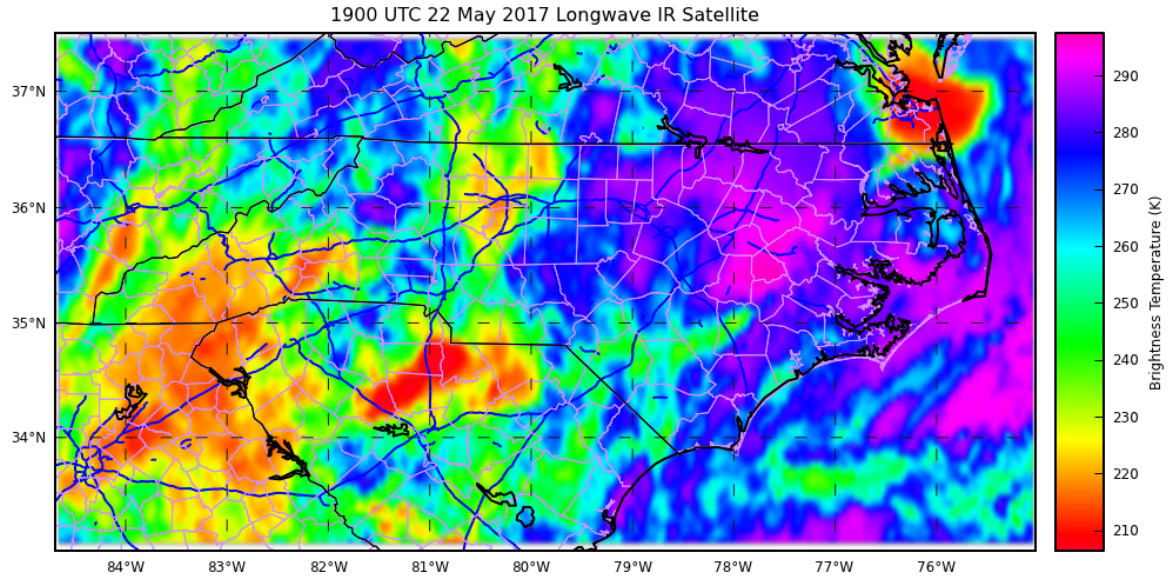


Figure 20 Output from Longwave IR Satellite (May 22, 2017 at 1900 UTC)

5.3.3 Visibility Maps using Thiessen Polygons

A GIS-based method is adopted to predict or nowcast weather conditions at the route / link level for dissemination of information to drivers in a timely manner. In this method, the geocoded weather station data is used to divide the entire study area into Thiessen or proximal zones. This

process is done by generating Thiessen polygons, by triangulating the point input features (location of weather stations) into a triangulated irregular network (TIN) that meets the Delaunay criterion. Each Thiessen polygon consists of only one weather station. The developed Thiessen or proximal polygons indicate that, any location (route or point) within a Thiessen polygon is closer to its associated weather station than to any other weather station in the proximity. The weather (fog / visibility) observed at a given weather station can be assigned to all the routes / links in the polygon associated with the weather station. Figure 21 represents the visibility condition for North Carolina and areas surrounding North Carolina on December 26, 2014 at 1700 UTC. At this particular date and time, visibility maps represent the visibility less than 5,000 m near Goldsboro, Tarboro, Greenville and New Bern in North Carolina. It is represented in fluorescent blue color in Figure 21. Further, the visibility is between 5,000 m and 10,000 m in Roaring Gap, North Carolina (Dark green color). Visibility between 10,000 m to 15,000 is represented in light green color. In all other places in North Carolina, visibility was more than 15000m at this particular date and time (Orange color).

The Thiessen polygon method helps evaluate the presence of fog or reduced visibility at any given point in the entire state of North Carolina, given that data from all weather stations in the state are available. Also, fog is a localized phenomenon and size of the theissen polygons is dependent on the distance between the weather stations. Generally, weather stations are located at airports and are located far from each other. Therefore, generated theissen polygons are very large compared to the localized phenomenon of visibility. Hence, creating proximity zones and assessing fog or visibility accurately at the route / link level could be cumbersome and would lead to inaccurate results.

Since fog is a local phenomenon and weather stations are located far from each other, it is more likely that spatial interpolation methods would be help better evaluate weather condition at locations distant from an observation site. This method is discussed next.

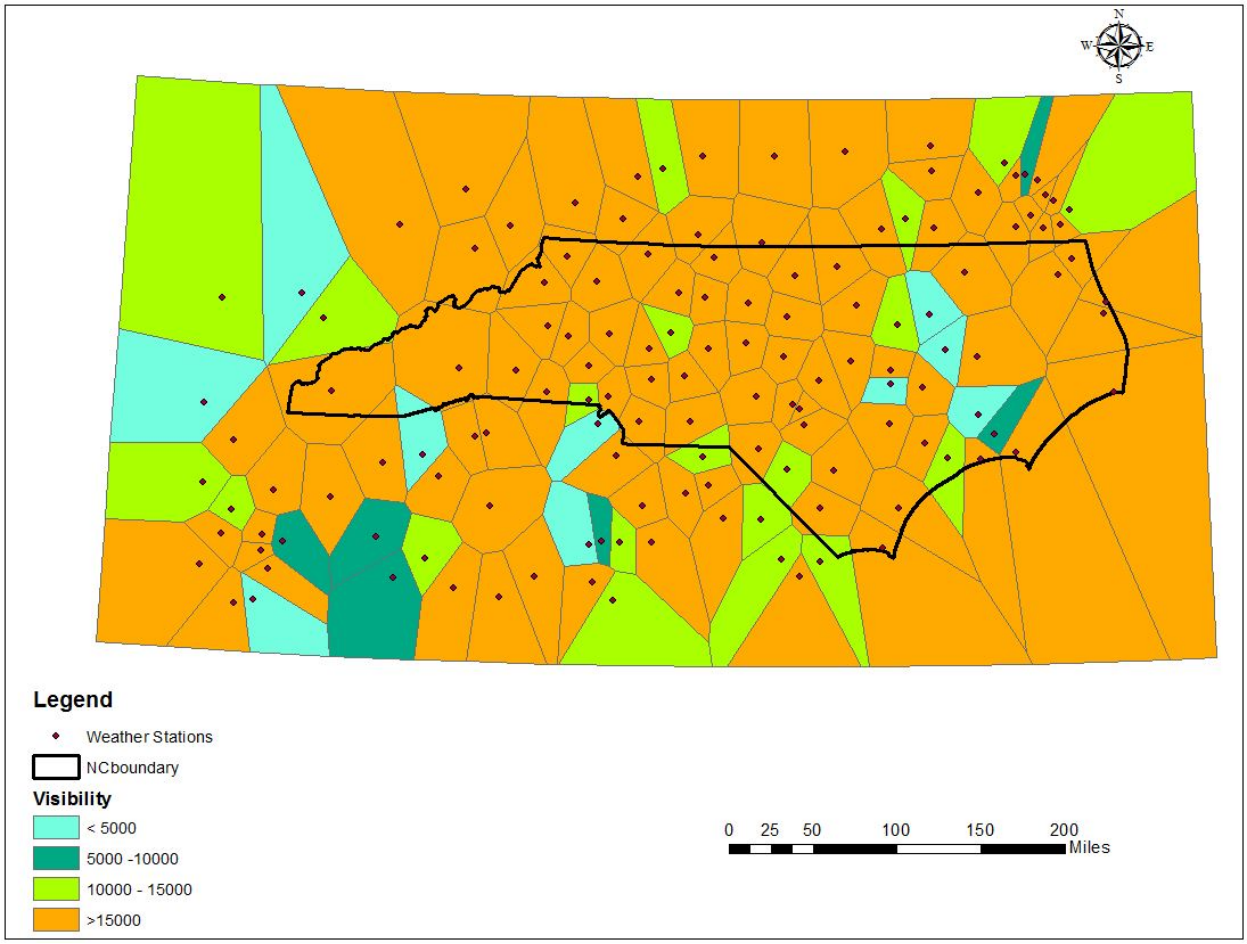


Figure 21 Visibility using Thiessen Polygon (December 26, 2014 at 1700 UTC)

5.3.4 Interpolated Visibility Maps

The visibility at an unknown location in a given area could be computed through spatial interpolation. An inverse distance weighted interpolation method was used to evaluate the visibility at an unknown location based on the visibility information obtained from nearby weather stations. One of the assumptions of Inverse Distance Weighted (IDW) interpolation is that locations that are close to one another are more alike than those that are farther apart. Therefore, to predict visibility at any given location, IDW uses the visibility values surrounding the prediction location. The visibility values closest to the prediction location will have more influence on the predicted value than those farther away from the prediction location. This indicates that each visibility value has a local influence by giving greater weights to locations closest to the prediction location, and the weights diminish with an increase in the distance from the prediction location.

R-statistical software was used to develop visibility maps for the state of North Carolina based on the visibility information obtained from the weather stations. Firstly, the coordinates of

each weather station were imported to create a spatial object of weather stations for North Carolina. A grid or extent (boundary) is defined such that it covers the entire state of North Carolina and its surroundings. The visibility was interpolated at each grid point using inverse distance weighting method based on the visibility from the weather stations. As the size of the grid decreases, the accuracy in the interpolation increases. Therefore, various grid sizes such as 0.75 Degree Lat/Long Grid (83 Km) – Figure 22, 0.50 Degree Lat/Long Grid (55 Km) – Figure 23, 0.25 Degree Lat/Long Grid (28 Km) – Figure 24, 0.10 Degree Lat/Long Grid (11 Km) – Figure 25, and 0.05 Degree Lat/Long Grid (5 Km) – Figure 26 were developed and tested. The figures indicate that as the grid size decreases, the accuracy in spatial representation of visibility increases. However, visibility maps are observed to be constant without any significant changes with grid sizes less than 0.25 Degree Lat/Long Grid (28 Km). Therefore, for North Carolina, a grid size of 0.25 Degree Lat/Long Grid (28 Km) is recommended to evaluate visibility at route / link level. Similar to section 5.3.3, florescent blue grids indicate lower visibility, whereas orange color grids indicate higher visibility (>15,000 m).

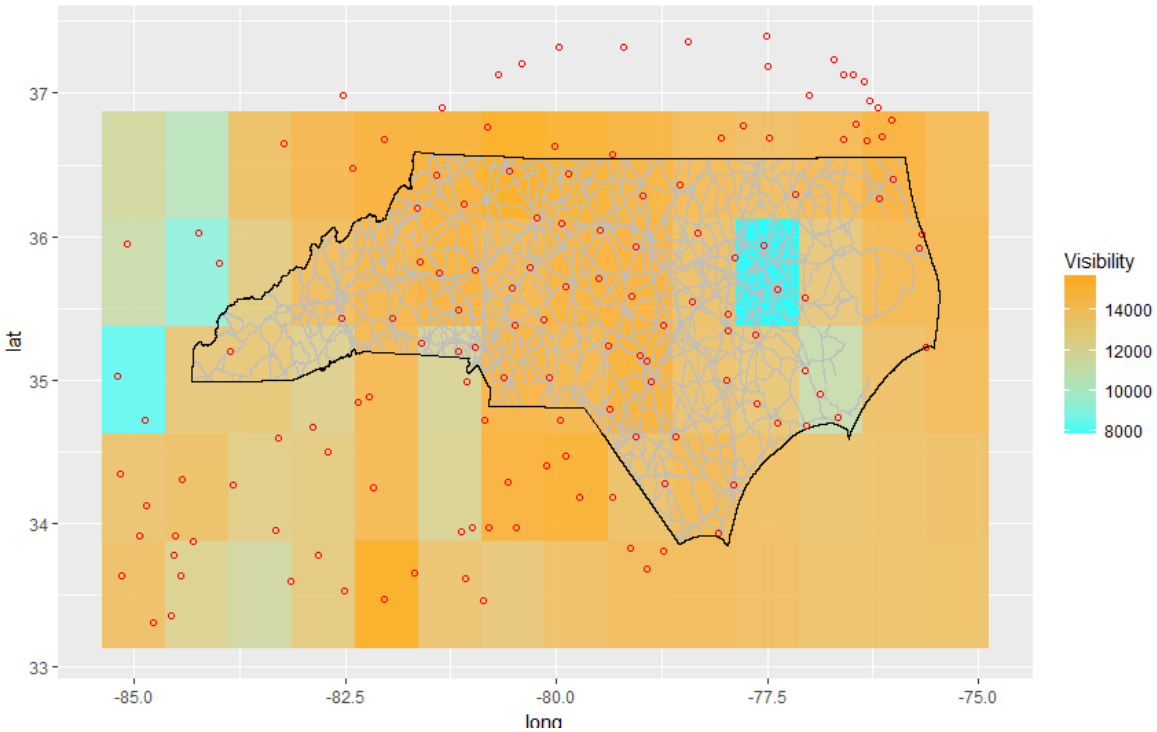


Figure 22 Visibility – 0.75 Degree Lat/Long Grid (December 26, 2014 at 1700 UTC)

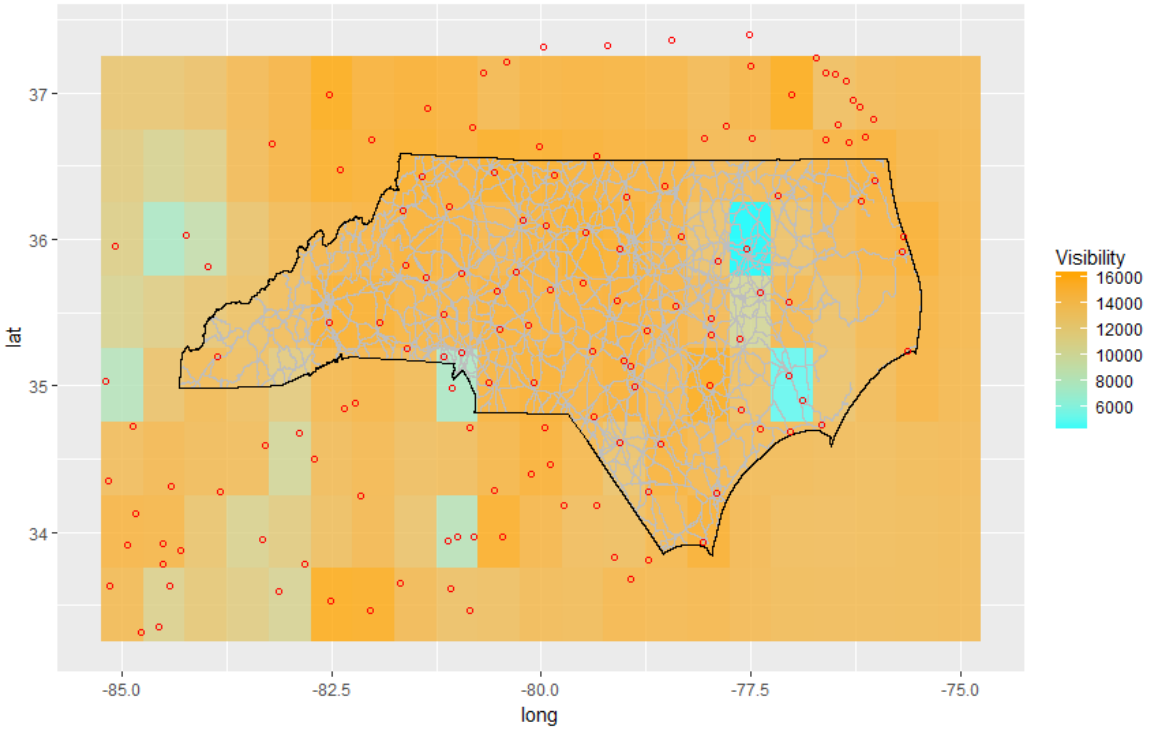


Figure 23 Visibility – 0.5 Degree Lat/Long Grid (December 26, 2014 at 1700 UTC)

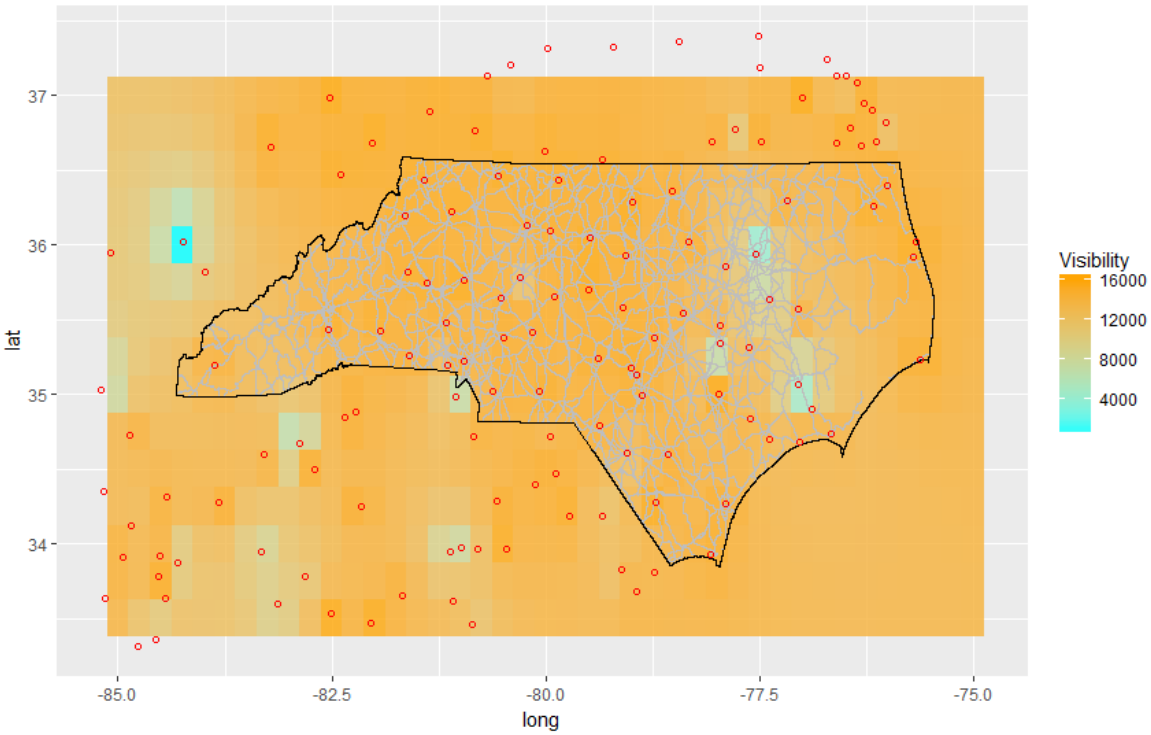


Figure 24 Visibility – 0.25 Degree Lat/Long Grid (December 26, 2014 at 1700 UTC)

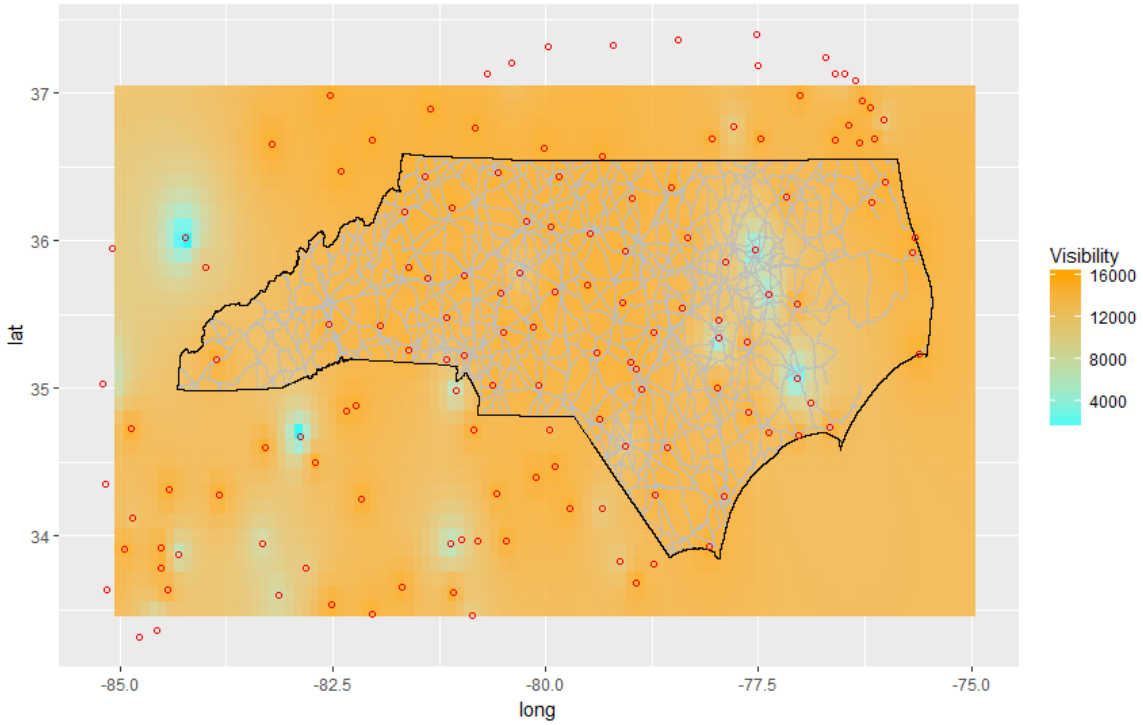


Figure 25 Visibility – 0.1 Degree Lat/Long Grid (December 26, 2014 at 1700 UTC)

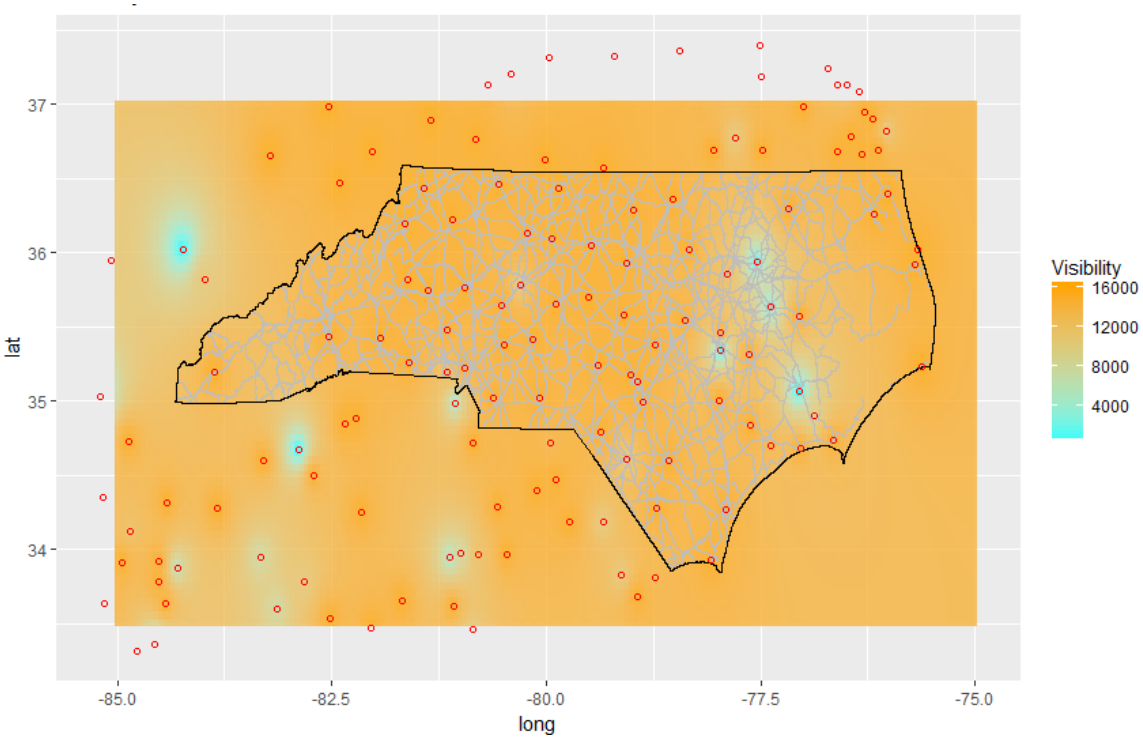


Figure 26 Visibility – 0.05 Degree Lat/Long Grid (December 26, 2014 at 1700 UTC)

6. VALIDATION

This chapter presents the validation of developed models and also a comparison of visibility information obtained from different data sources. This could help to quantify the applicability of developed models and different data sources based on their accuracy in predictions.

Validation was conducted in three different ways. All the three ways of validation are discussed next. As explained in "Methodology" chapter, for validation purpose, visibility data was collected using visibility sensor at five different locations (Table 7). Also, visibility data collected from visibility sensor was considered as the actual value of visibility since it provides the actual nearby visibility scenario at the installed location. Table 7 illustrates the nearby weather stations and their distance from the visibility sensor. It also shows the distance between the visibility sensor and nearest HRRR grid point.

Table 7 Weather Monitoring Stations and HRRR Grid Point from Visibility Sensor

Visibility Sensor	Weather Monitoring Station	Distance between Sensor and Weather Station (Miles)	Distance between Sensor and Nearest HRRR Grid Point (Miles)
UNC Charlotte	Charlotte/Douglas International Airport (CLT)	13.23	1.052
	Monroe Airport	21.29	
	Gastonia Municipal Airport	24.63	
Shelby	Lincolnton–Lincoln County Regional Airport	4.26	0.396
	Hickory Regional Airport	19.55	
	Gastonia Municipal Airport	20.39	
	Shelby Municipal Airport	25.80	
Lenoir	Boone Airport	6.86	0.906
	Hickory Regional Airport	28.65	
UNC Asheville	Asheville Regional Airport	13.21	0.405
Wilmington	Wilmington International Airport	5.40	0.805

6.1 Comparison of Visibility from Visibility Sensor with Nearby Weather Stations

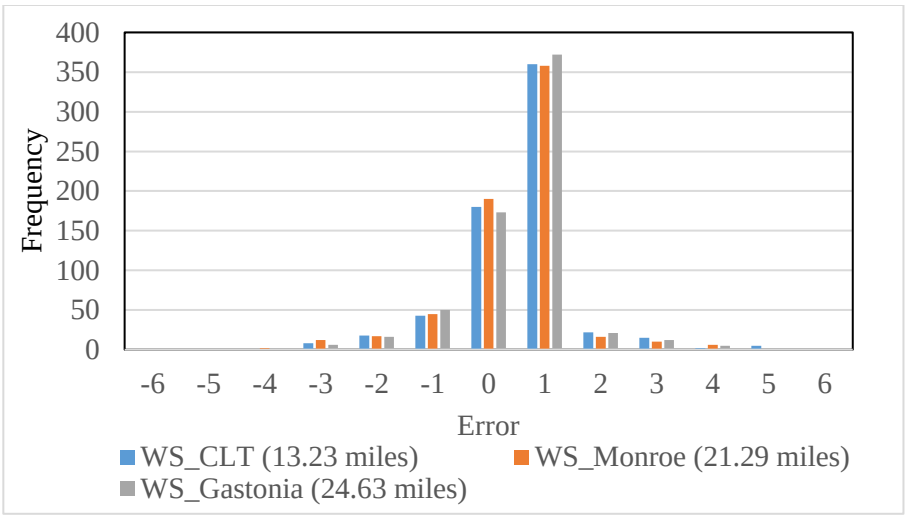
Both, visibility sensor (VS) and weather monitoring stations (WS) gather actual visibility values at the installed locations. However, it is difficult to install visibility sensor or obtain visibility

information from weather monitoring stations at each and every given location (say, a given point or a road link). Therefore, comparisons between visibility values from the visibility sensor and nearby weather stations may shed some light on the spatial correspondence between regional visibility observations.

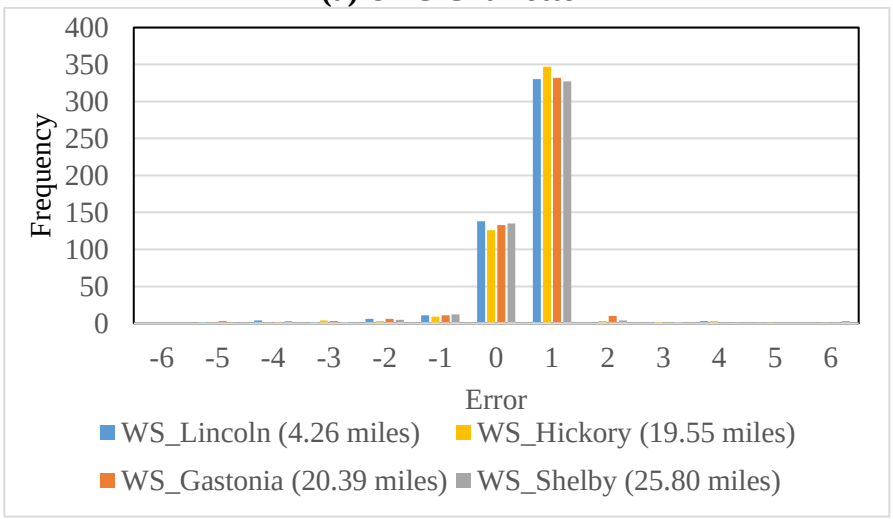
Weather stations within the range of 30 miles from visibility sensors were selected to compare the visibility values. Visibility sensor installed at UNC Charlotte, Shelby and Lenoir were considered for this analysis. Table 7 presents nearby weather stations for these locations. UNC Asheville and Wilmington data were not considered since only one weather station is located within the 30 miles from the installed location.

Firstly, hourly visibility values from nearby weather stations were collected for the distinct dates, as presented in Table 7. All the missing data points were removed from the dataset. As per the international standards (Table 5), visibility values from visibility sensors and weather stations were classified into eight categories. Errors were computed by taking a difference between the visibility categories and are presented in Figure 27.

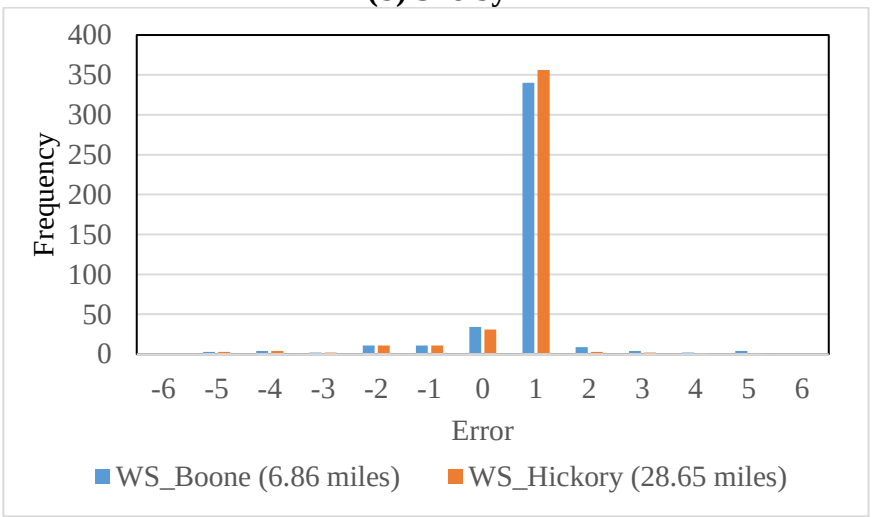
Figure 27 (a) illustrates the error histogram of visibility from weather station compared with visibility sensor at UNC Charlotte. Here, positive error indicates that visibility information obtained from visibility sensor is higher than the visibility information obtained from weather station. Also, negative error indicates that visibility information obtained from visibility sensor is lower than the visibility information obtained from weather station. Further, zero error indicates that the visibility information obtained from both visibility sensor and weather station falls in the same visibility category. Most of the samples (~ 350 samples) were observed to have +1 error. The grey line indicates errors obtained by comparing visibility values at UNC Charlotte with visibility values from Gastonia weather station. The blue line and the orange line indicates errors from comparison of visibility values at UNC Charlotte with weather stations at Charlotte Douglas International Airport (KCLT) and Monroe airport, respectively. CLT weather station (13.23 miles) is closest to the visibility sensor installed at UNC Charlotte. Monroe and Gastonia weather stations are located ~ 21 miles and ~ 24 miles from the visibility sensor installed at UNC Charlotte, respectively. Similar results were observed when visibility sensor located at Shelby was compared with nearby weather stations [Figure 27 (b) and (c)].



(a) UNC Charlotte



(b) Shelby



(c) Lenoir

Figure 27 Comparison of Visibility from Visibility Sensor with Nearby Weather Stations

6.2 Comparison of Visibility Values with All Data Sources

Visibility information obtained from visibility sensor was compared with visibility computed from the WLS regression model (RM), nearest HRRR grid point, and nearest weather monitoring station. This comparison provides the applicability between the developed model, HRRR dataset, and ISD dataset to help identify the best source to evaluate visibility values at route / link level.

Visibility sensor / weather station installed near the road link indicate the actual visibility scenario in the nearby area. Weather stations are generally installed at the airport and widely distributed across the state. However, it is a difficult and costly to install visibility sensor or have a weather station at every 5 miles / 10 miles apart on the roads. HRRR dataset provides forecasted visibility values for every hour and for grid sizes as small as 2-miles.

Regression model requires various weather parameters to predict the visibility at a given location. Therefore, it is necessary to compare visibility from all the three sources with visibility from visibility sensor to identify the best source for evaluating visibility at route / link level.

Four visibility sensor locations (UNC Charlotte, Lenoir, Shelby and UNC Asheville) were considered for the comparison. For each location, nearest HRRR grid point and the nearest weather station were considered. The weather data from these sources were collected for distinct dates which are summarized as shown in Table 7. For the regression model, weather parameters such as air temperature, dew point temperature, cloud cover, wind speed, precipitation, and elevation were collected from the nearest weather station. Hourly visibility was predicted for the distinct dates and for respective visibility sensor location using these parameters as an input to the regression model (Table 2). Similar to the earlier comparison, visibility values were classified into eight categories as per the international classification (Table 5). Further, visibility values obtained from HRRR, weather station, and regression model were compared with respect to visibility values from the sensor by computing the errors. Errors were computed for all the four locations by taking a difference between the visibility categories and are presented in Figure 28.

Figure 28 (a) illustrates the error histogram of visibility from HRRR, regression model, and weather station compared with visibility sensor at UNC Charlotte. Nearest HRRR grid point is located at 1.052 miles from the installed visibility sensor location at UNC Charlotte. However, the nearest weather station is located 13.23 miles away from the sensor at UNC Charlotte. In this analysis, positive error indicates that the visibility at the visibility sensor location is higher than the visibility from HRRR, regression model, and weather station. Also, negative error indicates

that the visibility at the visibility sensor location is lower than the visibility from HRRR, regression model, and weather station. Further, zero error indicates that the visibility at the visibility sensor location and visibility from HRRR, regression model, and weather station fall in the same visibility category.

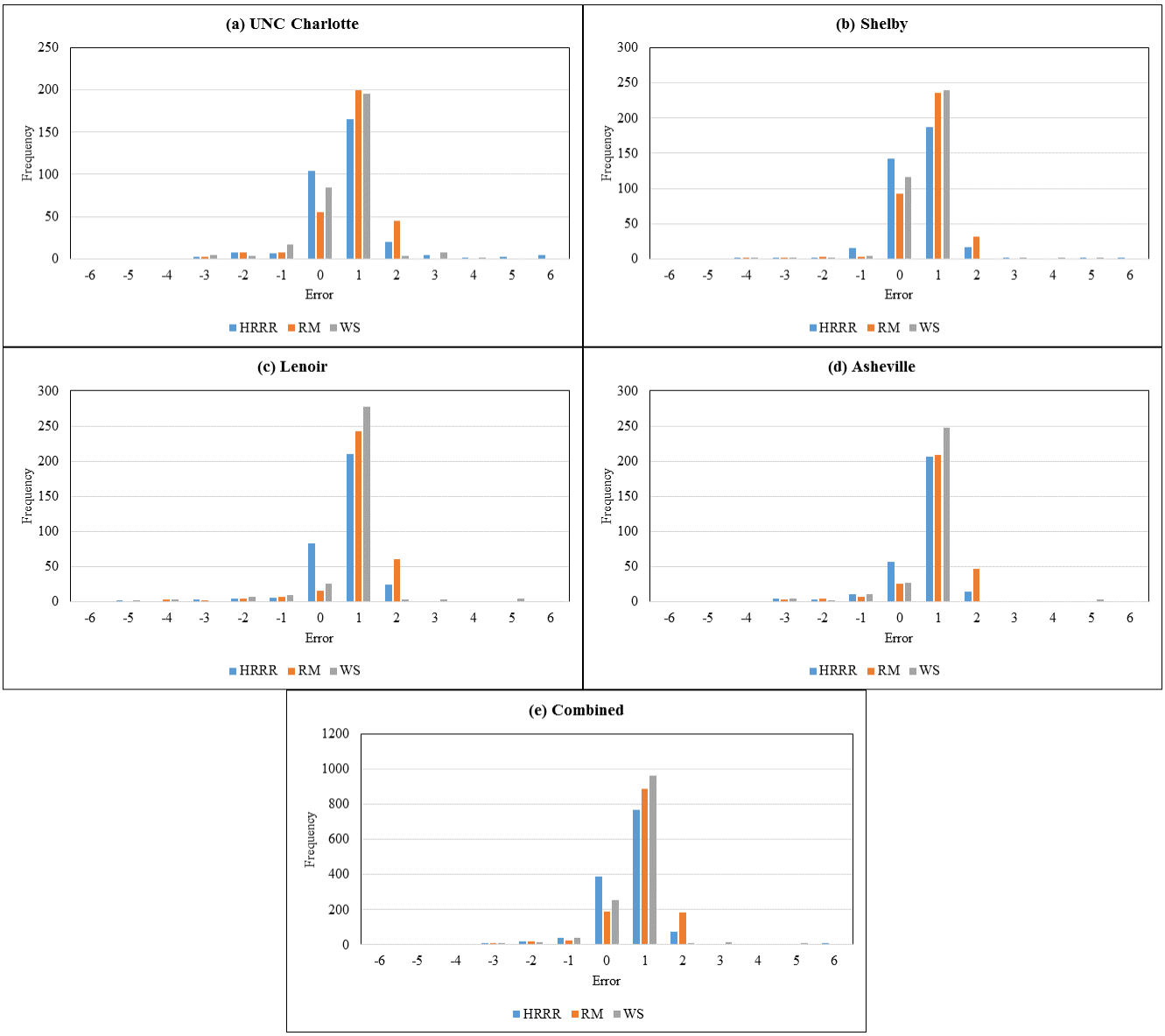


Figure 28 Comparison of Visibility Values from Visibility Sensors with Nearest HRRR, Regression Model, and Nearest Weather Station

Most of the samples (~ 200 samples) were observed to have +1 error. Also, the number of samples with zero error were higher for HRRR dataset compared to weather station and regression model. Similar results were observed for all the four locations (Figure 28 (b), Figure 28 (c) &

Figure 28 (d)]. Therefore, HRRR model is observed to be a better predictive model compared to all other models considered in this study.

6.3 Validation of Back Propagation Neural Network Models

The weather data obtained from the nearest weather station of each location where visibility sensor was installed was used to validate neural network models. Overall, data with 938 samples were obtained from weather stations closest to the visibility after removing any missing records. The data was used as input to the neural network model to compute the classification of visibility as shown earlier in Table 5. Similarly, the visibility information collected using visibility sensor was also classified into eight categories according to international standards as shown in Table 5. The visibility classification from the neural network model output was compared with the actual visibility obtained from visibility sensors. Figure 29 shows the error histogram of visibility from neural network model compared with visibility sensor. Most of the samples (700 samples or 75% of samples) are observed to have zero error and more than 95% of the samples are with in ±2 categories from the actual visibility obtained from the visibility sensor location.

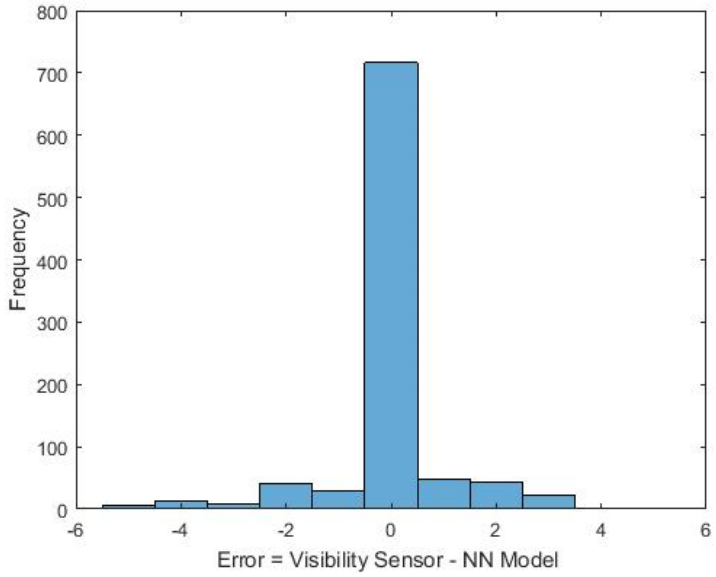


Figure 29 Error Histogram of Visibility from BPNN Model Compared with Visibility Sensor

7. TECHNOLOGIES AND LIMITATIONS

There are several companies such as Campbell Scientific, Vaisala, and Belfort instrument which provide visibility sensors for installation at a given specific location. In this study, CS120A was purchased from Campbell Scientific which costs around $8,000 per unit (Figure 30).

Figure 30 Visibility Sensor (Installed at UNC Charlotte)

The procured instrument currently captures only visibility up to 75,000 m for every 1 minute. The limitations of instrument procured are: 1) it can capture only visibility information, 2) it

requires AC power to operate (battery / solar power option is also available), 3) data has to be manually downloaded, and, 4) it cannot be easily transported from one place to another.

The regression models are only applicable to weather information collected from the weather monitoring stations in the state of North Carolina. To evaluate the accuracy of regression models, the visibility sensor equipped with the add-on features that collect temperature, dew point temperature, precipitation and cloud cover is required. With each add-on, the price of the instrument increases. These instrument store the data, but the captured data has to be downloaded manually for monitoring purpose. This visibility sensor has a limitation to transfer the data to a distinct place wirelessly. However, wireless transmission of real-time data can also be done by purchasing and installing a communications add-on to the sensor.

Other notable manufactures include Vaisala, which is currently being used by NCDOT Asheville division to monitor visibility i.e. PWD 10. This instrument has a measurement range of 10 m to 2000 m. PWD 10 is most preferred for monitoring visibility on roads. For higher visibility ranges, FS 11 by Vaisala can be used with a range of 5 m to 75,000 m. These instruments can also be equipped with solar panel for power supply. However, the cost of the equipment could go up to $25,000 per unit.

8. CONCLUSIONS

Visibility is critical to the task of driving. Fog presents a challenge to motorists and can result in significant safety concerns due to reduced visibility and has resulted in terrible crashes because drivers are often caught unaware by sudden reductions in visibility. Various historical weather data sources, such as ISD, HRRR forecasts, and visibility sensors were explored to predict visibility conditions. Data obtained from these sources was processed considering variables such as wind speed, elevation, relative humidity, temperature, dewpoint, sea level pressure, station pressure, precipitation, cloud cover, ceiling height data, rainfall in the past three hours (Rain 3hr), three to six hours (Rain 6hr), six to twelve hours (Rain 12hr) and twelve to twenty-four hours (Rain 24hr), and presence of water bodies to predict visibility. Statistical and back-propagation neural network techniques were used in developing models and were validated for performance evaluation. Along with these models, the Thiessen polygon method and inverse distance weighted interpolation maps were developed to help predict visibility at route / link level.

The results obtained indicate that weighted linear regression models were statistically significant in predicting visibility at a 95% confidence level. The weighted linear regression model developed for visibility less than 15,000 m (which considered data from all weather information) has better performance compared to all other regression models. In addition, visibility sensors were installed and weather information from nearest weather station was used to validate data sources and the developed models. Visibility values obtained from the models were compared with visibility information obtained from the visibility sensor to validate the developed models.

The results obtained from model validation indicate that HRRR model has better performance in predicting the visibility. For validation of regression model, the primary challenge is that it requires weather information such as air temperature, dew point, precipitation, elevation, and cloud cover to predict the visibility at a given location. Generally, for all the locations, nearest weather station is located at more than 5 miles apart. Also, the procured visibility sensor has a limited capability which can collect only visibility information. Different add-ons for the sensor are required to capture information, such as temperature, dew point temperature, rainfall, cloud cover information. Ideally, for validation of regression model, information such as air temperature, dew point, precipitation, elevation, and cloud cover at the installed location of visibility sensor is necessary.

Backward propagation neural network was observed to outperform all other models developed in predicting visibility. The neural networks are fast and do not need any formulas or conditions. Their adaptive nature helps them adapt to the data variations and learn input characteristics yielding better results. However, unlike statistical methods, the neural network model is a black box model that has a predictive value solely based on observations but does not provide any explanation. Therefore, statistical models are more appropriate when one wants to understand the role of predictor variables.

HRRR model is an important source that can be considered as it helps predict visibility dynamically for the next two hours. Also, it can provide visibility information for future timestamps at a route / link level. Out of the all models developed and available data sources, HRRR dataset is observed to be more useful for government agencies to provide dynamic forecasts of visibility information to travelers through radio / mobile application / dynamic signs. Similar to HRRR, satellite-based fog detection product is a novel product to forecast visibility in real-time.

To provide with a real-time visibility information, it is recommended to perform a feasibility study that explores suitability of off-the-shelf low-cost visibility sensors or build a low-cost visibility sensor which can help provide dynamic visibility information to travelers. A further study on identifying hot-spots of low visibility areas which could have an impact on traffic safety is recommended. Identifying such hot-spots will help agencies prioritize and use the limited resources available efficiently. Further, with advancement in modelling methods and technologies, several products such as HRRR and satellite-based fog data sources could be put to better use for predicting and informing road users with visibility information at link-level. The cost of hosting such a server with tools for downloading / uploading and disseminating visibility information is estimated around $50,000, plus annual maintenance and upgrades.

Currently available off-the-shelf instruments are expensive and can be installed at a limited number of locations. There is need to explore or develop a new low cost visibility sensor that are solar powered and can transmit information wirelessly to enable dynamic massages for road users. The development of such low cost visibility sensors and their applicability at route / link level merit research and investigation.

8.1 Implementation Plan

Based on the findings from this research, it is recommended that NCDOT use HRRR and satellite-based fog data sources to provide with visibility information for travelers. As HRRR can provide with 2-hour forecasts, it can be used to provide advanced warning to travelers through dynamic message signs. The statistical and neural network models developed in this research can be used by NCDOT to forecast visibility based on long-term weather forecast (> 2 hours) data obtained from weather stations. It helps NCDOT to evaluate long-term forecasts of visibility. It is suggested that current available technologies for evaluating visibility be tested at low visibility hot-spots, which could have an impact on traffic safety.

To provide with dynamic messages on low visibility conditions to road users, large amounts of weather data that can be obtained from HRRR and GOES (satellite-based) data sources has to be processed. To process such data, it is recommended to install a stand-alone sever that is capable of downloading / uploading and processing the weather data to identify both low visibility hot-spots and duration of such adverse weather conditions in the state of North Carolina. The processed information can be further transmitted to travelers using dynamic massage signs on roadways to provide them with advanced warnings related low-visibility conditions.

References

1. Abdel-Aty, M., Ekram, A. A., Huang, H., & Choi, K. (2011). A study on crashes related to visibility obstruction due to fog and smoke. Accident Analysis & Prevention Journal, 43(5), 1730-1737.
2. Ahmed, M. M., Abdel-Aty, M., Lee, J., & Yu, R. (2014). Real-time assessment of fog-related crashes using airport weather data: A feasibility analysis. Accident Analysis & Prevention Journal, 72, 309-317.
3. Al-Ghamdi, A. S. (2007). Experimental evaluation of fog warning system. Accident Analysis & Prevention Journal, 39(6), 1065-1072.
4. Ashley, W. S., Strader, S., Dziubla, D. C. & A. Haberlie. (2015). Driving blind: weather-related vision hazards and fatal motor Vehicle Crashes. Bulletin of the American Meteorological Society, 96, 755–778. doi: http://dx.doi.org/10.1175/BAMS-D-14-00026.1.
5. Baker, R., Cramer, J., & Peters, J. (2002). Radiation fog: UPS airlines conceptual models and forecast methods. In 10th conference on aviation, range, and aerospace, Portland, OR, American Meteorological Society.
6. Balagh, A. K. G., Naderkhani, F., & Makis, V. (2014). Highway accident modeling and forecasting in winter. Transportation Research Part A: Policy and Practice, 59, 384-396.
7. Bergot, T., Carrer, D., Noilhan, J., & Bougeault, P. (2005). Improved site-specific numerical prediction of fog and low clouds: a feasibility study. Weather and Forecasting, 20(4), 627-646.
8. Bott, A. (1991). On the influence of the physico-chemical properties of aerosols on the life cycle of radiation fogs. Boundary-Layer Meteorology, 56, 1-31.
9. Bott, A., Sievers, U., & Zdunkowski, W. (1990). A radiation fog model with a detailed treatment of the interaction between radiative transfer and fog microphysics. Journal of the Atmospheric Sciences, 47, 2153-2166.
10. Chin, S. M., Franzese, O., Greene, D. L., Hwang, H. L., & Gibson, R. C. (2002). Temporary losses of highway capacity and impacts on performance (No. ORNL/TM-2002/3). Oak Ridge, TN: Oak Ridge National Laboratory.
11. Cochocki, A., & Unbehauen, R. (1993). Neural networks for optimization and signal processing. John Wiley & Sons, Inc.

12. Colquhoun, J. R. (1987). A decision tree method for forecasting thunderstorms and tornadoes. Weather and Forecasting, 2, 337-345.
13. Del Greco, S. A., Lott, J. N., Hawkins, S. K., Baldwin, R., Anders, D. D., Ray, R., Dellinger, D., Jones, P., & Smith, F.. (2006). Surface data integration at NOAA's national climatic data center: data format, processing, QC, and product generation. International Conference on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology (IIPS), Atlanta, GA, American Meteorological Society.
14. Ellrod, G. P. (1995). Advances in the detection and analysis of fog at night using GOES multispectral infrared imagery. Weather Forecasting, 10, 606-619.
15. Ellrod, G. P., & Gultepe, I. (2007). Inferring low cloud base heights at night for aviation using satellite infrared and surface temperature data. Journal of Pure and Applied Geophysics, 164, 1193-1205.
16. Friedlein, M. T. (2004). Dense fog climatology: Chicago O'hare international airport July 1996-April 2002. Bulletin of the American Meteorological Society, 85(4), 515-517.
17. Glickman, T. S. (1995). 2000: Glossary of meteorology, 2nd Edition, American Meteorological Society.
18. Godfrey, C. M. (2015) Improved climatic data for mechanistic-empirical pavement design. Report FHWA/NC/2014-01, North Carolina Department of Transportation, University of North Carolina at Asheville, Asheville, NC.
19. Goodwin, L. (2003a). Weather-related crashes on U.S. highways in 2001. Federal Highway Administration/Mitretek Systems, Inc., Available at www.ops.fhwa.dot.gov/weather/resources/publications/fhwa/01crashanalyspaperv2.doc.
20. Goodwin, L. (2003b) Best practices for road weather management, version 2.0. Tech. Rep. FHWA-OP-03-081. Available at: http://ops.fhwa.dot.gov/weather/best_practices/CaseStudiesFINALv2-RPT.pdf.
21. Goodwin, L. C. (2002). Analysis of weather-related crashes on US highways. Prepared for the FHWA Road Weather Management Program. Available at: http://ops.fhwa.dot.gov/weather/resources/publications/fhwa/crashanalysis2001.pdf.
22. Gopalakrishna, D., Cluett, C., Kitchener, F., & Sturges, L. (2013). Utah DOT weather responsive traveler information system. Last Accessed on 12/08/2015 at http://ntl.bts.gov/lib/51000/51000/51067/F9C62201.pdf.

23. Guidard, V., & Tzanos, D. (2007). Analysis of fog probability from a combination of satellite and ground observation data. Pure and Applied Geophysics, 164(6-7), 1207-1220.
24. Gultepe, I. & Milbrandt, J. A. (2007). Microphysical observations and mesoscale model simulation of a warm fog case during FRAM project. Journal of Pure and Applied Geophysics, 164, 1161-1178.
25. Gultepe, I., Muller, M. D., & Boybeyi, Z. (2006). A new visibility parameterization for warm-fog applications in numerical weather prediction models. Journal of Applied Meteorology and Climatology, 45, 1469-1480.
26. Gultepe, I., Tardif, R., Michaelides, S. C., Cermak, J., Bott, A., Bendix, J., Müller, M. D., Pagowski, M., Hansen, B., Ellrod, G., Jacobs, W., Toth, G., & Cober, S. G. (2007). Fog research: a review of past achievements and future perspectives. In Fog and Boundary Layer Clouds: Fog Visibility and Forecasting, Springer, 1121-1159.
27. Hansen, B. K. (2000). Analog forecasting of ceiling and visibility using fuzzy sets. In 2nd Conference on Artificial Intelligence, American Meteorological Society, 1-7.
28. Hardwick, W. C. (1973) Monthly fog frequency in the continental United States. Monthly Weather Review, 101, 763–766.
29. Hassan, H. M., & Abdel-Aty, M. A. (2013) Predicting reduced visibility related crashes on freeways using real-time traffic flow data. Journal of Safety Research, 45, 29-36.
30. Ham, F. M., & Kostanic, I. (2000). Principles of neurocomputing for science and engineering. McGraw-Hill Higher Education.
31. Hamilton, B., Tefft, B., Arnold, L., & Grabowski, J. (2014). Hidden highways: Fog and traffic crashes on America's roads. AAA Foundation for Traffic Safety, Washington, DC.
32. Hilliker, J. L. & Fritsch, J. M. (1999). An observations-based statistical system for warm-season hourly probabilistic forecasts of low ceiling at the San Francisco International Airport. Journal of Applied Meteorology, 38, 1692-1705.
33. Jiusto, J. E. (1981). Fog Structure. In Clouds: Their formation, optical properties, end effects, P. V. Hobbs and A. Deepak, Eds., Academic Press, 187-239.
34. Kang, J., Ni, R., & Andersen, G. (2008). Effects of reduced visibility from fog on car-following performance. Transportation Research Record: Journal of the Transportation Research Board, 2069, 9-15.

35. Kunkel, B. A. (1984). Parameterization of droplet terminal velocity and extinction coefficient in fog models. Journal of Climate and Applied Meteorology, 23, 34-41.
36. Liang, F. (2003). An effective bayesian neural network classifier with a comparison study to support vector machine. Neural Computation, 15(8), 1959-1989.
37. Liang, F. (2005). Bayesian neural networks for nonlinear time series forecasting. Statistics and Computing, 15(1), 13-29.
38. Lott, J. N. (2004). The quality control of the integrated surface hourly database. Presented in 14th Conference on Applied Climatology, American Meteorological Society, Seattle, WA.
39. Marzban, C., Leyton, S., & Colman B. (2007). Ceiling and visibility forecasts via neural networks. Weather Forecasting, 22,466-479.
40. MATLAB Neural Network Toolbox, 2016 (a). The MathWorks Inc., Natick, MA.
41. Meyer, M. B., & Lala, G. G. (1990). Climatological aspects of radiation fog occurrence at Albany, New York. Journal of Climate, 3(5), 577-586.
42. Meyer, M. B., Jiusto, J. E., & Lala, G. G. (1980). Measurements of visual range and radiation-fog (haze) microphysics. Journal of Atmospheric Science, 37, 622-629.
43. Mueller, A. S., & Trick, L. M. (2012). Driving in fog: the effects of driving experience and visibility on speed compensation and hazard avoidance. Accident Analysis & Prevention Journal, 48, 472-479.
44. Müller, M. D., Schmutz, C. & Parlow, E. (2007). A one dimensional ensemble forecast and assimilation system for fog prediction. Journal of Pure and Applied Geophysics, 164, 1241–1264.
45. Murtha, J. (1995). Applications of fuzzy logic in operational meteorology. Scientific Services and Professional Development Newsletter, Canadian Forces Weather Service, 42-54.
46. Nilsson, N. J. (1971). Problem-solving methods in artificial intelligence. McGraw-Hill: New York.
47. Oliver, C. J. (2013). Fog related crashes statewide study of fog related crashes in North Carolina. Prepared by Traffic Safety Unit Transportation Mobility and Safety Division, Highways North Carolina Department of Transportation.
48. Peace, R. L. (1969). Heavy-fog regions in the conterminous United States. Monthly Weather Review, 97, 116-123.

49. Petterssen, S. (1956). Weather analysis and forecasting. Second Edition, Vol. 2, (McGraw-Hill Publ. Inc., New York.
50. Pinto, J. O., Grim, J. A., & Steiner, M. (2015). Assessment of the High-Resolution Rapid Refresh Model's ability to predict mesoscale convective systems using object-based evaluation. Weather and Forecasting, 30(4), 892-913.
51. Pisano, P. A., Goodwin, L. C., & Rossetti, M. A. (2008). U.S. highway crashes in adverse road weather conditions. Proc. 24th Conf. on Int. Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology, New Orleans, LA.
52. Shi, J., & Tan, J. (2013). Effect analysis of intermittent release measures in heavy fog weather with an improved CA model. Discrete Dynamics in Nature and Society, Vol. 2013, Article ID 812562, 7 pages.
53. Smith, A., Lott, N., & Vose, R. (2011). The integrated surface database: Recent developments and partnerships. Bulletin of the American Meteorological Society, 92(6), 704-708.
54. Sujitjorn, S., Sookjaras, P., & Wainikorn, W. (1994). An expert system to forecast visibility in Don- Muang Air Force Base. IEEE International Conference on Systems, Man and Cybernetics (Humans, Information and Technology), New York.
55. Tardif, R., & Rasmussen, R. M. (2007). Event-based climatology and typology of fog in the New York City region. Journal of Applied Meteorology and Climatology, 46(8), 1141-1168.
56. Tardif, R., & Rasmussen, R. M. (2008). Process-oriented analysis of environmental conditions associated with precipitation fog events in the New York City region. Journal of Applied Meteorology and Climatology, 47(6), 1681-1703.
57. Theofilatos, A., & Yannis, G. (2014) A review of the effect of traffic and weather characteristics on road safety. Accident Analysis & Prevention Journal, 72, 244-256.
58. Trick, L. M., Toxopeus, R., & Wilson, D. (2010). The effects of visibility conditions, traffic density, and navigational challenge on speed compensation and driving performance in older adults. Accident Analysis & Prevention Journal, 42, 1661-1671.
59. Vislocky, R. L., & Fritsch, J. M. (1997). An automated, observations-based system for short-term prediction of ceiling and visibility. Weather and Forecasting, 12(1), 31-43.
60. Ward, B., & Croft, P. J. (2008). Use of GIS to examine winter fog occurrences. National Weather Association, Electronic Journal of Operational Meteorology, 33.

61. Weinreb, M., Johnson, J. X., & Han, D. (2011). Conversion of GVAR infrared data to scene radiance or temperature. NOAA NESDIS Office of Satellite Operations.
62. Whiffen, B., Delannoy, P. & Siok, S. (2004). Fog: impact on road transportation and mitigation options. National Highway Visibility Conference, Madison, WI, 18-19 May.
63. World Meteorological Organization (WMO) (1966). International Meteorological Vocabulary. Geneva, Switzerland.
64. Xie, Y., Lord, D., & Zhang, Y. (2007). Predicting motor vehicle collisions using Bayesian neural network models: An empirical analysis. Accident Analysis & Prevention Journal, 39(5), 922-933.
65. Yan, X., Li, X., Liu, Y., & Zhao, J. (2014). Effects of foggy conditions on drivers' speed control behaviors at different risk levels. Journal of Safety Science, 68, 275-287.

Annexure A

Table A1 OLS Regression Model for Visibility Data <15000m

Variable	All		Elevation < 50 m		Elevation 50 m to 250 m		Elevation 250 m to 750 m		Elevation > 750 m	
	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value
Elevation	-0.76	<0.01	-6.79	<0.01	1.58	<0.01	-0.62	<0.01	-2.74	<0.01
Cloud cover	-32.07	<0.01	-30.46	<0.01	-31.21	<0.01	-35.15	<0.01	-40.11	<0.01
m10wspd	231.90	<0.01	265.33	<0.01	273.84	<0.01	250.82	<0.01	120.49	<0.01
Precipitation	-183.39	<0.01	-219.76	<0.01	-159.00	<0.01	-185.38	<0.01	-74.06	<0.01
tair_dew	268.18	<0.01	199.31	<0.01	298.13	<0.01	363.39	<0.01	277.89	<0.01
Time 4 am	10,148.41	<0.01	10,084.07	<0.01	9,842.95	<0.01	10,214.81	<0.01	11,820.29	<0.01
Time 8 am	9,745.73	<0.01	9,540.62	<0.01	9,518.98	<0.01	9,914.41	<0.01	11,302.41	<0.01
Time 12 pm	9,292.91	<0.01	9,106.77	<0.01	9,008.21	<0.01	9,474.77	<0.01	10,865.77	<0.01
Time 4 pm	9,639.20	<0.01	9,356.07	<0.01	9,461.49	<0.01	9,735.17	<0.01	11,275.40	<0.01
Time 8 pm	9,190.01	<0.01	9,175.67	<0.01	8,819.86	<0.01	9,300.62	<0.01	10,964.43	<0.01
Time 12 am	9,520.59	<0.01	9,582.64	<0.01	9,094.74	<0.01	9,528.11	<0.01	11,274.07	<0.01
Rain 3 hr	711.86	<0.01	679.21	<0.01	557.05	<0.01	753.21	<0.01	1,212.77	<0.01
Rain 6 hr	93.73	<0.01	102.47	<0.01	89.48	<0.01	69.91	0.02	-	-
Rain 12 hr	-44.87	<0.01	159.78	<0.01	-118.56	<0.01	-273.55	<0.01	-185.96	<0.01
Rain 24 hr	50.15	<0.01	166.28	<0.01	-	-	-	-	174.51	<0.01
Water	-	-	114.83	<0.01	-46.65	0.04	133.37	<0.01	-	-
No. of Observations	685,375		244,448		259,202		158,126		23,599	
R-Square	0.86		0.87		0.86		0.84		0.80	
Adj. R-Square	0.86		0.87		0.86		0.84		0.80	
AIC	13,200,000		4,674,395		4,983,872		3,038,998		451,866	
RMSE	3,567.00		3,436.30		3,622.50		3,606.30		3,479.40	

Table A2 WLS Regression Model for Visibility Data <15000m

Variable	All		Elevation < 50 m		Elevation 50 m to 250 m		Elevation 250 m to 750 m		Elevation > 750 m	
	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value
Elevation	-0.78	<0.01	-7.59	<0.01	1.50	<0.01	-0.67	<0.01	-3.05	<0.01
Cloud cover	-32.38	<0.01	-30.43	<0.01	-31.66	<0.01	-35.77	<0.01	-41.65	<0.01
m10wspd	234.85	<0.01	262.50	<0.01	274.46	<0.01	246.38	<0.01	122.74	<0.01
Precipitation	-184.56	<0.01	-228.52	<0.01	-166.44	<0.01	-192.18	<0.01	-85.40	<0.01
tair_dew	265.58	<0.01	191.97	<0.01	296.01	<0.01	362.05	<0.01	285.38	<0.01
Time 4 am	10,247.42	<0.01	10,211.02	<0.01	9,986.03	<0.01	10,350.68	<0.01	12,215.94	<0.01
Time 8 am	9,833.03	<0.01	9,641.01	<0.01	9,643.31	<0.01	10,042.75	<0.01	11,631.98	<0.01
Time 12 pm	9,367.32	<0.01	9,184.14	<0.01	9,091.50	<0.01	9,588.56	<0.01	11,157.89	<0.01
Time 4 pm	9,700.69	<0.01	9,426.37	<0.01	9,544.33	<0.01	9,856.87	<0.01	11,624.32	<0.01
Time 8 pm	9,199.83	<0.01	9,239.34	<0.01	8,901.71	<0.01	9,404.95	<0.01	11,274.35	<0.01
Time 12 am	9,578.33	<0.01	9,672.88	<0.01	9,186.11	<0.01	9,629.27	<0.01	11,585.38	<0.01
Rain 3 hr	683.66	<0.01	652.33	<0.01	545.01	<0.01	762.64	<0.01	1,288.19	<0.01
Rain 6 hr	74.74	<0.01	99.86	<0.01	84.92	-	50.92	<0.01	-	-
Rain 12 hr	-48.45	<0.01	159.26	<0.01	-157.62	<0.01	-324.04	<0.01	-201.17	<0.01
Rain 24 hr	45.96	<0.01	166.97	<0.01	-	-	-	-	203.30	<0.01
Water	-	-	97.51	<0.01	-67.99	<0.01	115.88	<0.01	-	-
No. of Observations	685,375		244,448		259,202		158,126		23,599	
R-Square	0.99		0.99		0.98		0.98		0.95	
Adj. R-Square	0.99		0.99		0.98		0.98		0.95	
AIC	11,200,000		366,836		4,468,495		2,684,371		415,281	
RMSE	825.40		808.27		1,340.60		1,175.10		1,602.80	

Table A3 OLS Regression Model for Visibility Data <15000m (Near to Roads)

Variable	All		Elevation < 50 m		Elevation 50 m to 250 m		Elevation 250 m to 750 m		Elevation > 750 m	
	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value
Elevation	-1.09	<0.01	-	-	-0.92	<0.01	-1.48	<0.01	-17.37	<0.01
Cloud cover	-32.52	<0.01	-34.09	<0.01	-28.33	<0.01	-35.44	<0.01	-36.97	<0.01
m10wspd	209.86	<0.01	282.80	<0.01	231.95	<0.01	240.39	<0.01	-	-
Precipitation	-161.89	<0.01	-216.12	<0.01	-141.82	<0.01	-128.12	<0.01	-105.95	<0.01
tair_dew	432.77	<0.01	294.88	<0.01	608.58	<0.01	513.25	<0.01	427.63	<0.01
Time 4 am	9,851.07	<0.01	9,679.63	<0.01	9,675.72	<0.01	10,238.53	<0.01	25,359.16	<0.01
Time 8 am	9,488.88	<0.01	9,147.41	<0.01	9,527.57	<0.01	9,910.24	<0.01	24,998.97	<0.01
Time 12 pm	8,980.28	<0.01	8,670.09	<0.01	8,983.95	<0.01	9,327.96	<0.01	24,636.44	<0.01
Time 4 pm	9,271.29	<0.01	9,041.36	<0.01	9,134.62	<0.01	9,632.86	<0.01	24,716.27	<0.01
Time 8 pm	8,762.38	<0.01	8,672.40	<0.01	8,174.97	<0.01	9,454.76	<0.01	24,695.88	<0.01
Time 12 am	9,131.83	<0.01	9,110.62	<0.01	8,605.57	<0.01	9,618.31	<0.01	24,799.25	<0.01
Rain 3 hr	779.91	<0.01	933.00	<0.01	399.94	<0.01	914.04	<0.01	776.27	<0.01
Rain 6 hr	69.80	<0.01	151.12	<0.01	-	-	-	-	-198.67	<0.01
Rain 12 hr	-	-	271.61	<0.01	-87.21	0.02	-349.60	<0.01	-	-
Rain 24 hr	94.18	<0.01	237.95	<0.01	-	-	-137.93	<0.01	375.40	<0.01
Water	259.19	<0.01	367.98	<0.01	206.83	<0.01	-	-	-	-
No. of Observations	183,377		71,615		59,531		42,334		9,897	
R-Square	0.85		0.87		0.86		0.84		0.82	
Adj. R-Square	0.85		0.87		0.86		0.84		0.82	
AIC	3,523,477		1,376,872		1,146,357		808,167		188,501	
RMSE	3,598.30		3,619		3,670.60		3,381		3,360.00	

Table A4 WLS Regression Model for Visibility Data <15000m (Near to Roads)

Variable	All		Elevation < 50 m		Elevation 50 m to 250 m		Elevation 250 m to 750 m		Elevation > 750 m	
	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value
Elevation	-1.10	<0.01	-	-	-1.18	<0.01	-1.50	<0.01	-18.23	<0.01
Cloud cover	-33.01	<0.01	-34.47	<0.01	-28.23	<0.01	-36.02	<0.01	-38.64	<0.01
m10wspd	210.64	<0.01	279.38	<0.01	230.35	<0.01	240.47	<0.01	-	-
Precipitation	-168.84	<0.01	-225.91	<0.01	-148.92	<0.01	-133.94	<0.01	-109.88	<0.01
tair_dew	436.03	<0.01	292.43	<0.01	614.36	<0.01	524.95	<0.01	442.70	<0.01
Time 4 am	9,945.09	<0.01	9,810.18	<0.01	9,820.39	<0.01	10,342.91	<0.01	26,260.36	<0.01
Time 8 am	9,572.02	<0.01	9,239.20	<0.01	9,647.69	<0.01	9,986.01	<0.01	25,883.67	<0.01
Time 12 pm	9,031.51	<0.01	8,742.02	<0.01	9,074.99	<0.01	9,378.47	<0.01	25,454.09	<0.01
Time 4 pm	9,319.43	<0.01	9,119.31	<0.01	9,197.60	<0.01	9,705.09	<0.01	25,545.07	<0.01
Time 8 pm	8,805.10	<0.01	8,761.18	<0.01	8,229.73	<0.01	9,519.71	<0.01	25,502.93	<0.01
Time 12 am	9,187.82	<0.01	9,186.21	<0.01	8,692.96	<0.01	9,673.67	<0.01	25,625.09	<0.01
Rain 3 hr	779.30	<0.01	933.70	<0.01	370.77	<0.01	902.89	<0.01	841.47	<0.01
Rain 6 hr	55.68	<0.01	150.13	<0.01	-	-	-	-	-234.02	<0.01
Rain 12 hr	-	-	259.16	<0.01	-127.64	<0.01	-373.40	<0.01	-	-
Rain 24 hr	95.76	<0.01	237.83	<0.01	-	-	-139.14	<0.01	413.21	<0.01
Water	269.34	<0.01	369.33	<0.01	205.99	<0.01	-	-	-	-
No. of Observations	18,377		71,615		59,531		42,334		9,897	
R-Square	0.98		0.98		0.98		0.97		0.96	
Adj. R-Square	0.98		0.98		0.98		0.97		0.96	
AIC	3,141,427		1,237,751		1,032,004		735,429		172,251	
RMSE	1,269.70		1,370.10		1,406.30		1,430.10		1,454.70	

Table A5 OLS Regression Model for Visibility Data <10000m

Variable	All		Elevation < 50 m		Elevation 50 m to 250 m		Elevation 250 m to 750 m		Elevation > 750 m	
	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value
Elevation	-0.39	<0.01	-		1.89	<0.01	-0.05	<0.01	-0.83	<0.01
Cloud cover	-27.90	<0.01	-28.20	<0.01	-26.92	<0.01	-28.91	<0.01	-33.01	<0.01
m10wspd	171.60	<0.01	181.64	<0.01	205.36	<0.01	192.98	<0.01	132.37	<0.01
Precipitation	-72.42	<0.01	-99.64	<0.01	-63.34	<0.01	-57.13	<0.01	-	-
tair_dew	126.20	<0.01	107.67	<0.01	147.31	<0.01	137.40	<0.01	117.39	<0.01
Time 4 am	7,060.14	<0.01	6,840.46	<0.01	6,707.08	<0.01	7,170.20	<0.01	7,597.80	<0.01
Time 8 am	6,839.58	<0.01	6,560.42	<0.01	6,551.50	<0.01	6,942.90	<0.01	7,376.72	<0.01
Time 12 pm	6,693.67	<0.01	6,488.45	<0.01	6,363.08	<0.01	6,759.93	<0.01	7,093.93	<0.01
Time 4 pm	7,027.37	<0.01	6,890.60	<0.01	6,717.75	<0.01	6,976.27	<0.01	7,335.88	<0.01
Time 8 pm	6,878.63	<0.01	6,790.81	<0.01	6,444.54	<0.01	6,945.61	<0.01	7,193.21	<0.01
Time 12 am	6,979.21	<0.01	6,930.48	<0.01	6,504.79	<0.01	7,039.19	<0.01	7,388.89	<0.01
Rain 3 hr	916.54	<0.01	1,002.32	<0.01	801.65	<0.01	845.83	<0.01	1,130.51	<0.01
Rain 6 hr	158.21	<0.01	184.84	<0.01	155.20	<0.01	126.25	<0.01	-	-
Rain 12 hr	50.85	<0.01	199.27	<0.01	46.63	0.02	-148.63	<0.01	-102.69	0.03
Rain 24 hr	-	-	169.87	<0.01	-41.18	0.02	-119.56	<0.01	-	-
Water	-222.53	<0.01	-219.11	<0.01	-224.95	<0.01	-	-	-	-
No. of Observations	361,182		117,233		137,447		90,637		15,865	
R-Square	0.83		0.84		0.83		0.81		0.79	
Adj. R-Square	0.83		0.84		0.83		0.81		0.79	
AIC	6,701,768		2,173,896		2,550,679		1,681,680		292,863	
RMSE	2,587.90		2,572.60		2,590.90		2,586.30		2,466.40	

Table A6 WLS Regression Model for Visibility Data <10000m

Variable	All		Elevation < 50 m		Elevation 50 m to 250 m		Elevation 250 m to 750 m		Elevation > 750 m	
	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value
Elevation	-0.44	<0.01	-	-	2.01	<0.01	-0.06	0.03	-0.79	<0.01
Cloud cover	-28.39	<0.01	-28.87	<0.01	-27.63	<0.01	-29.73	<0.01	-34.03	<0.01
m10wspd	178.65	<0.01	189.37	<0.01	211.70	<0.01	199.64	<0.01	138.62	<0.01
Precipitation	-74.99	<0.01	-102.16	<0.01	-64.49	<0.01	-58.03	<0.01	-	-
tair_dew	127.14	<0.01	102.90	<0.01	145.39	<0.01	134.62	<0.01	116.94	<0.01
Time 4 am	7,109.28	<0.01	6,911.71	<0.01	6,759.31	<0.01	7,240.62	<0.01	7,627.63	<0.01
Time 8 am	6,897.88	<0.01	6,596.70	<0.01	6,613.05	<0.01	7,019.88	<0.01	7,960.07	<0.01
Time 12 pm	6,722.66	<0.01	6,519.74	<0.01	6,390.47	<0.01	6,816.74	<0.01	7,080.83	<0.01
Time 4 pm	7,071.53	<0.01	6,940.13	<0.01	6,768.97	<0.01	7,043.13	<0.01	7,340.95	<0.01
Time 8 pm	6,901.08	<0.01	6,827.12	<0.01	6,468.13	<0.01	7,003.21	<0.01	7,193.06	<0.01
Time 12 am	7,011.97	<0.01	6,964.96	<0.01	6,549.90	<0.01	7,098.79	<0.01	7,405.52	<0.01
Rain 3 hr	926.88	<0.01	1,021.40	<0.01	812.61	<0.01	849.06	<0.01	1,176.74	<0.01
Rain 6 hr	143.17	<0.01	190.18	<0.01	141.72	<0.01	123.49	<0.01	-	-
Rain 12 hr	55.87	<0.01	208.90	<0.01	41.51	<0.01	-178.46	<0.01	-95.62	<0.01
Rain 24 hr	-	-	177.17	<0.01	-50.67	<0.01	-110.93	<0.01	-	-
Water	-234.49	<0.01	-219.00	<0.01	-241.94	<0.01	-	-	-	-
No. of Observations	361,182		117,233		137,447		90,637		15,865	
R-Square	0.98		0.98		0.97		0.97		0.96	
Adj. R-Square	0.98		0.98		0.97		0.97		0.96	
AIC	5,254,038		1,846,543		2,292,924		1,485,337		264,234	
RMSE	348.78		636.84		1,014.50		875.56		1,000.50	

Table A7 OLS Regression Model for Visibility Data <10000m (Near to Roads)

Variable	All		Elevation < 50 m		Elevation 50 m to 250 m		Elevation 250 m to 750 m		Elevation > 750 m	
	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value
Elevation	-0.49	<0.01	-	-	1.03	<0.01	-0.64	<0.01	-13.87	<0.01
Cloud cover	-28.15	<0.01	-28.82	<0.01	-25.06	<0.01	-31.11	<0.01	-30.48	<0.01
m10wspd	175.74	<0.01	219.64	<0.01	195.00	<0.01	167.83	<0.01	54.24	<0.01
Precipitation	-60.44	<0.01	-101.71	<0.01	-41.13	<0.01	-42.16	<0.01	-40.13	0.04
tair_dew	208.83	<0.01	134.75	<0.01	382.33	<0.01	263.27	<0.01	275.58	<0.01
Time 4 am	6,865.89	<0.01	6,686.13	<0.01	6,295.22	<0.01	7,390.74	<0.01	19,683.49	<0.01
Time 8 am	6,722.82	<0.01	6,412.11	<0.01	6,335.12	<0.01	7,241.17	<0.01	19,524.11	<0.01
Time 12 pm	6,520.76	<0.01	6,294.97	<0.01	6,190.45	<0.01	6,829.43	<0.01	19,326.89	<0.01
Time 4 pm	6,798.39	<0.01	6,634.50	<0.01	6,400.06	<0.01	7,118.41	<0.01	19,224.01	<0.01
Time 8 pm	6,581.29	<0.01	6,411.67	<0.01	5,806.57	<0.01	7,273.34	<0.01	19,272.53	<0.01
Time 12 am	6,723.76	<0.01	6,705.21	<0.01	5,838.03	<0.01	7,355.68	<0.01	19,386.16	<0.01
Rain 3 hr	986.64	<0.01	1,130.13	<0.01	657.96	<0.01	1,097.47	<0.01	835.14	<0.01
Rain 6 hr	156.28	<0.01	227.71	<0.01	155.47	<0.01	127.88	<0.01	-	-
Rain 12 hr	61.84	<0.01	262.39	<0.01	-	-	-202.26	<0.01	-	-
Rain 24 hr	57.88	<0.01	223.63	<0.01	66.11	0.05	-181.92	<0.01	216.16	<0.01
Water	-181.77	<0.01	-140.86	<0.01	-139.24	<0.01	-	-	-	-
No. of Observations	101,847		36,708		32,531		25,839		6,769	
R-Square	0.82		0.83		0.82		0.82		0.81	
Adj. R-Square	0.82		0.83		0.82		0.82		0.81	
AIC	1,890,508		682,266		605,237		476,734		124,376.80	
RMSE	2,597.00		2,628		2,652.60		2,455		2,362.20	

Table A8 WLS Regression Model for Visibility Data <10000m (Near to Roads)

Variable	All		Elevation < 50 m		Elevation 50 m to 250 m		Elevation 250 m to 750 m		Elevation > 750 m	
	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value
Elevation	-0.50	<0.01	-	-	1.02	<0.01	-0.68	<0.01	-14.04	<0.01
Cloud cover	-28.86	<0.01	-29.78	<0.01	-25.61	<0.01	-31.91	<0.01	-30.72	<0.01
m10wspd	183.96	<0.01	230.16	<0.01	204.84	<0.01	165.42	<0.01	54.35	<0.01
Precipitation	-62.40	<0.01	-101.15	<0.01	-45.97	<0.01	-42.21	<0.01	-49.54	<0.01
tair_dew	221.09	<0.01	133.92	<0.01	403.70	<0.01	298.94	<0.01	275.05	<0.01
Time 4 am	6,898.32	<0.01	6,694.60	<0.01	6,300.22	<0.01	7,469.19	<0.01	19,850.55	<0.01
Time 8 am	6,763.11	<0.01	6,437.51	<0.01	6,365.48	<0.01	7,313.28	<0.01	19,694.79	<0.01
Time 12 pm	6,542.06	<0.01	6,319.06	<0.01	6,201.69	<0.01	6,841.30	<0.01	19,504.57	<0.01
Time 4 pm	6,824.36	<0.01	6,672.49	<0.01	6,422.91	<0.01	7,167.27	<0.01	19,373.14	<0.01
Time 8 pm	6,590.24	<0.01	6,441.66	<0.01	5,804.14	<0.01	7,325.95	<0.01	19,422.28	<0.01
Time 12 am	6,741.61	<0.01	6,734.72	<0.01	5,836.11	<0.01	7,423.99	<0.01	19,552.91	<0.01
Rain 3 hr	1,000.87	<0.01	1,149.07	<0.01	688.43	<0.01	1,108.02	<0.01	859.68	<0.01
Rain 6 hr	158.00	<0.01	209.72	<0.01	129.67	<0.01	138.20	<0.01	-	-
Rain 12 hr	63.19	<0.01	271.57	<0.01	-	-	-211.21	<0.01	-	-
Rain 24 hr	66.97	<0.01	218.13	<0.01	58.35	<0.01	-200.60	<0.01	219.05	<0.01
Water	-195.69	<0.01	-118.89	<0.01	-144.38	<0.01	-	-	-	-
No. of Observations	101,847		36,708		32,531		25,839		6,769	
R-Square	0.97		0.99		0.98		0.97		0.99	
Adj. R-Square	0.97		0.99		0.98		0.97		0.99	
AIC	1,688,427		595,369		531,956		426,776		105,086	
RMSE	962.97		804.63		860.03		933.68		568.16	

Table A9 OLS Regression Model for Visibility Data <5000m

Variable	All		Elevation < 50 m		Elevation 50 m to 250 m		Elevation 250 m to 750 m		Elevation > 750 m	
	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value
Elevation	0.06	<0.01	5.90	<0.01	1.54	<0.01	-	-	-	-
Cloud cover	-17.87	<0.01	-17.12	<0.01	-17.33	<0.01	-19.63	<0.01	-24.52	<0.01
m10wspd	109.61	<0.01	101.52	<0.01	137.48	<0.01	130.29	<0.01	114.44	<0.01
Precipitation	-6.62	<0.01	-19.51	<0.01	-	-	-	-	-	-
tair_dew	-14.59	<0.01	-12.94	<0.01	-8.24	0.04	-20.46	<0.01	54.87	<0.01
Time 4 am	3,956.05	<0.01	3,541.96	<0.01	3,744.69	<0.01	4,332.98	<0.01	4,355.17	<0.01
Time 8 am	3,795.35	<0.01	3,406.09	<0.01	3,585.98	<0.01	4,161.01	<0.01	4,224.34	<0.01
Time 12 pm	3,730.89	<0.01	3,438.09	<0.01	3,499.48	<0.01	3,990.90	<0.01	4,140.02	<0.01
Time 4 pm	4,038.83	<0.01	3,812.21	<0.01	3,797.77	<0.01	4,257.24	<0.01	4,293.55	<0.01
Time 8 pm	4,224.96	<0.01	3,916.37	<0.01	3,986.86	<0.01	4,515.61	<0.01	4,382.63	<0.01
Time 12 am	4,179.28	<0.01	3,928.52	<0.01	3,894.24	<0.01	4,462.16	<0.01	4,457.43	<0.01
Rain 3 hr	696.22	<0.01	754.85	<0.01	612.67	<0.01	658.82	<0.01	879.57	<0.01
Rain 6 hr	198.93	<0.01	193.97	<0.01	211.37	<0.01	202.77	<0.01	-	-
Rain 12 hr	62.29	<0.01	196.46	<0.01	53.15	<0.01	-69.18	<0.01	-	-
Rain 24 hr	-45.67	<0.01	94.78	<0.01	-61.42	<0.01	-179.96	<0.01	-	-
Water	-231.30	<0.01	-308.40	<0.01	-127.30	<0.01	-113.81	<0.01	-	-
No. of Observations	163,213		48,951		60,442		44,703		9,117	
R-Square	0.79		0.79		0.81		0.80		0.77	
Adj. R-Square	0.79		0.79		0.81		0.80		0.77	
AIC	2,845,300		852,055		1,053,069		778,812		159,183	
RMSE	1,476.70		1,456.90		1,469.00		1,468.30		1,496.00	

Table A10 WLS Regression Model for Visibility Data <5000m

Variable	All		Elevation < 50 m		Elevation 50 m to 250 m		Elevation 250 m to 750 m		Elevation > 750 m	
	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value
Elevation	0.09	<0.01	6.32	<0.01	1.64	<0.01	-	-	-	-
Cloud cover	-18.76	<0.01	-18.45	<0.01	-18.51	<0.01	-20.39	<0.01	-25.13	<0.01
m10wspd	112.19	<0.01	105.74	<0.01	143.10	<0.01	137.30	<0.01	118.44	<0.01
Precipitation	-7.87	<0.01	-22.74	<0.01	-	-	-	-	-	-
tair_dew	-14.87	<0.01	-15.98	<0.01	-11.51	<0.01	-19.27	<0.01	52.86	<0.01
Time 4 am	4,006.41	<0.01	3,626.47	<0.01	3,809.63	<0.01	4,381.15	<0.01	4,400.86	<0.01
Time 8 am	3,866.88	<0.01	3,465.54	<0.01	3,644.32	<0.01	4,202.68	<0.01	4,259.34	<0.01
Time 12 pm	3,781.91	<0.01	3,504.71	<0.01	3,554.47	<0.01	4,041.32	<0.01	4,127.13	<0.01
Time 4 pm	4,085.93	<0.01	3,894.71	<0.01	3,883.76	<0.01	4,308.11	<0.01	4,332.60	<0.01
Time 8 pm	4,295.41	<0.01	4,011.86	<0.01	4,067.09	<0.01	4,578.40	<0.01	4,392.95	<0.01
Time 12 am	4,247.53	<0.01	4,022.37	<0.01	3,977.91	<0.01	4,523.72	<0.01	4,496.25	<0.01
Rain 3 hr	735.63	<0.01	814.40	<0.01	647.13	<0.01	697.50	<0.01	928.77	<0.01
Rain 6 hr	206.85	<0.01	200.41	<0.01	217.62	<0.01	204.66	<0.01	-	-
Rain 12 hr	55.34	<0.01	202.04	<0.01	56.21	<0.01	-71.68	<0.01	-	-
Rain 24 hr	-47.05	<0.01	103.02	<0.01	-68.91	<0.01	-197.12	<0.01	-	-
Water	-249.03	<0.01	-344.62	<0.01	-135.75	<0.01	-134.60	<0.01	-	-
No. of Observations	163,213		48,951		60,442		44,703		9,117	
R-Square	0.97		0.95		0.96		0.96		0.96	
Adj. R-Square	0.97		0.95		0.96		0.96		0.96	
AIC	2,515,898		772,895		936,818		696,850		140,030	
RMSE	538.30		649.02		562		587.06		523.29	

Table A11 OLS Regression Model for Visibility Data <5000m (Near to Roads)

Variable	All		Elevation < 50 m		Elevation 50 m to 250 m		Elevation 250 m to 750 m		Elevation > 750 m	
	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value
Elevation	-	-	-1.54	0.04	3.50	<0.01	-0.35	<0.01	-11.92	<0.01
Cloud cover	-18.14	<0.01	-16.55	<0.01	-16.33	<0.01	-24.37	<0.01	-20.57	<0.01
m10wspd	118.75	<0.01	107.30	<0.01	163.22	<0.01	114.42	<0.01	68.10	<0.01
Precipitation	-	-	-10.72	0.02	-	-	-	-	-	-
tair_dew	15.25	<0.01	17.06	0.03	37.19	<0.01	-	-	134.33	<0.01
Time 4 am	3,884.68	<0.01	3,683.15	<0.01	3,013.27	<0.01	4,836.86	<0.01	15,265.24	<0.01
Time 8 am	3,757.55	<0.01	3,480.71	<0.01	2,989.43	<0.01	4,638.25	<0.01	15,298.57	<0.01
Time 12 pm	3,690.34	<0.01	3,535.34	<0.01	2,986.91	<0.01	4,352.19	<0.01	15,226.36	<0.01
Time 4 pm	3,980.76	<0.01	3,917.06	<0.01	3,255.69	<0.01	4,602.52	<0.01	15,280.81	<0.01
Time 8 pm	4,141.13	<0.01	4,031.86	<0.01	3,237.23	<0.01	5,031.24	<0.01	15,370.37	<0.01
Time 12 am	4,075.38	<0.01	4,090.94	<0.01	3,061.52	<0.01	5,001.94	<0.01	15,258.44	<0.01
Rain 3 hr	718.88	<0.01	703.84	<0.01	493.85	<0.01	933.22	<0.01	724.73	<0.01
Rain 6 hr	204.57	<0.01	249.44	<0.01	198.09	<0.01	177.10	<0.01	-	-
Rain 12 hr	90.56	<0.01	217.31	<0.01	124.58	<0.01	-92.42	<0.01	-	-
Rain 24 hr	-	-	66.62	0.02	-	-	-106.91	<0.01	-	-
Water	-382.00	<0.01	-369.09	<0.01	-177.14	<0.01	-	-	-	-
No. of Observations	47,849		16,028		14,963		12,964		3,894	
R-Square	0.79		0.79		0.80		0.80		0.79	
Adj. R-Square	0.79		0.79		0.80		0.80		0.79	
AIC	834,011		279,060		260,837		225,244		67,610	
RMSE	1,474.30		1,460		1,475.40		1,434		1,423.60	

Table A12 WLS Regression Model for Visibility Data <5000m (Near to Roads)

Variable	All		Elevation < 50 m		Elevation 50 m to 250 m		Elevation 250 m to 750 m		Elevation > 750 m	
	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value
Elevation	-	-	-1.64	<0.01	3.73	<0.01	-0.44	<0.01	-12.61	<0.01
Cloud cover	-19.36	<0.01	-17.82	<0.01	-17.37	<0.01	-25.36	<0.01	-21.62	<0.01
m10wspd	123.84	<0.01	116.01	<0.01	169.63	<0.01	116.03	<0.01	70.90	<0.01
Precipitation	-	-	-11.47	<0.01	-	-	-	-	-	-
tair_dew	14.96	<0.01	14.01	<0.01	42.79	<0.01	-	-	127.71	<0.01
Time 4 am	3,954.84	<0.01	3,762.29	<0.01	3,042.41	<0.01	4,925.00	<0.01	15,975.83	<0.01
Time 8 am	3,827.46	<0.01	3,526.27	<0.01	3,031.71	<0.01	4,732.54	<0.01	16,006.92	<0.01
Time 12 pm	3,755.81	<0.01	3,595.82	<0.01	3,006.90	<0.01	4,410.19	<0.01	15,944.89	<0.01
Time 4 pm	4,069.04	<0.01	4,009.87	<0.01	3,299.47	<0.01	4,687.41	<0.01	15,982.99	<0.01
Time 8 pm	4,227.37	<0.01	4,124.37	<0.01	3,278.01	<0.01	5,137.49	<0.01	16,081.01	<0.01
Time 12 am	4,167.70	<0.01	4,160.07	<0.01	3,103.45	<0.01	5,111.04	<0.01	15,972.71	<0.01
Rain 3 hr	764.13	<0.01	756.75	<0.01	525.06	<0.01	1,003.59	<0.01	742.26	<0.01
Rain 6 hr	198.27	<0.01	258.78	<0.01	198.58	<0.01	171.13	<0.01	-	-
Rain 12 hr	109.46	<0.01	221.70	<0.01	113.89	<0.01	-89.43	<0.01	-	-
Rain 24 hr	-	-	64.56	<0.01	-	-	-115.57	<0.01	-	-
Water	-413.85	<0.01	-392.74	<0.01	-201.22	<0.01	-	-	-	-
No. of Observations	47,849		16,028		14,963		12,964		3,894	
R-Square	0.96		0.98		0.97		0.98		0.95	
Adj. R-Square	0.96		0.98		0.97		0.98		0.95	
AIC	748,164		243,483		228,051		195,046		61,166	
RMSE	601.18		481.12		493.29		447.26		622	

Table A13 OLS Regression Model for Visibility Data <2000m

Variable	All		Elevation < 50 m		Elevation 50 m to 250 m		Elevation 250 m to 750 m		Elevation > 750 m	
	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value
Elevation	-0.13	<0.01	-0.58	<0.01	0.49	<0.01	-0.19	<0.01	0.85	<0.01
Cloud cover	-2.78	<0.01	-2.66	<0.01	-2.88	<0.01	-3.39	<0.01	-34.66	<0.01
m10wspd	33.65	<0.01	30.14	<0.01	40.18	<0.01	30.80	<0.01	56.38	<0.01
Precipitation	6.44	<0.01	3.38	<0.01	5.78	<0.01	13.54	<0.01	36.39	<0.01
tair_dew	-	-	-4.90	<0.01	-	-	-	-	66.70	<0.01
Time 4 am	1,133.31	<0.01	1,106.04	<0.01	1,057.32	<0.01	1,268.88	<0.01	3,270.99	<0.01
Time 8 am	1,090.59	<0.01	1,065.39	<0.01	1,023.24	<0.01	1,208.81	<0.01	3,317.70	<0.01
Time 12 pm	1,001.42	<0.01	992.13	<0.01	924.23	<0.01	1,112.77	<0.01	3,220.28	<0.01
Time 4 pm	988.33	<0.01	961.19	<0.01	898.89	<0.01	1,117.53	<0.01	3,196.27	<0.01
Time 8 pm	1,153.26	<0.01	1,132.55	<0.01	1,069.03	<0.01	1,308.94	<0.01	3,252.65	<0.01
Time 12 am	1,146.76	<0.01	1,153.92	<0.01	1,025.41	<0.01	1,312.28	<0.01	3,234.74	<0.01
Rain 3 hr	137.14	<0.01	102.13	<0.01	110.72	<0.01	156.78	<0.01	219.15	<0.01
Rain 6 hr	58.60	<0.01	75.76	<0.01	45.23	<0.01	75.17	<0.01	-	-
Rain 12 hr	18.94	<0.01	30.46	<0.01	21.71	0.01	-	-	-	-
Rain 24 hr	-	-	42.75	<0.01	-	-	-52.69	<0.01	-57.10	<0.01
Water	-41.79	<0.01	-47.31	<0.01	-	-	-66.86	<0.01	-	-
No. of Observations	54,232		17,125		19,144		14,590		3,373	
R-Square	0.76		0.78		0.77		0.76		0.66	
Adj. R-Square	0.76		0.78		0.77		0.76		0.66	
AIC	821,396		258,336		289,282		221,156		51,702	
RMSE	470.56		456.35		462.27		473.24		514.49	

Table 14 WLS Regression Model for Visibility Data <5000m

Variable	All		Elevation < 50 m		Elevation 50 m to 250 m		Elevation 250 m to 750 m		Elevation > 750 m	
	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value
Elevation	-0.14	<0.01	-0.63	<0.01	0.44	<0.01	-0.19	<0.01	0.93	<0.01
Cloud cover	-2.92	<0.01	-2.72	<0.01	-3.09	<0.01	-3.70	<0.01	-36.16	<0.01
m10wspd	35.10	<0.01	29.98	<0.01	41.39	<0.01	32.24	<0.01	61.23	<0.01
Precipitation	6.89	<0.01	4.29	<0.01	6.98	<0.01	14.34	<0.01	40.65	<0.01
tair_dew	-	-	-3.75	<0.01	-	-	-	-	65.18	<0.01
Time 4 am	1,145.82	<0.01	1,103.98	<0.01	1,078.67	<0.01	1,296.25	<0.01	3,307.29	<0.01
Time 8 am	1,103.25	<0.01	1,069.16	<0.01	1,041.65	<0.01	1,239.83	<0.01	3,362.56	<0.01
Time 12 pm	1,010.26	<0.01	995.74	<0.01	947.86	<0.01	1,134.96	<0.01	3,274.85	<0.01
Time 4 pm	992.61	<0.01	961.29	<0.01	912.90	<0.01	1,143.75	<0.01	3,239.56	<0.01
Time 8 pm	1,174.59	<0.01	1,137.23	<0.01	1,099.22	<0.01	1,359.14	<0.01	3,304.63	<0.01
Time 12 am	1,157.48	<0.01	1,160.04	<0.01	1,050.00	<0.01	1,346.66	<0.01	3,275.99	<0.01
Rain 3 hr	135.92	<0.01	102.04	<0.01	111.51	<0.01	159.57	<0.01	250.79	<0.01
Rain 6 hr	57.59	<0.01	77.53	<0.01	44.61	<0.01	74.71	<0.01	-	-
Rain 12 hr	17.93	<0.01	29.65	<0.01	19.22	<0.01	-	-	-	-
Rain 24 hr	-	-	42.13	<0.01	-	-	-55.72	<0.01	-49.05	<0.01
Water	-45.20	<0.01	-47.48	<0.01	-	-	-76.24	<0.01	-	-
No. of Observations	54,232		17,125		19,144		14,590		3,373	
R-Square	0.98		0.97		0.96		0.96		0.93	
Adj. R-Square	0.98		0.97		0.96		0.96		0.93	
AIC	665,843		222,987		249,388		192,932		45,662	
RMSE	112.14		162.58		163		179.89		210.16	

Table A15 OLS Regression Model for Visibility Data <2000m (Near to Roads)

Variable	All		Elevation < 50 m		Elevation 50 m to 250 m		Elevation 250 m to 750 m		Elevation > 750 m	
	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value
Elevation	-0.03	0.02	-0.96	<0.01	1.54	<0.01	-	-	-3.75	<0.01
Cloud cover	-2.21	<0.01	-2.28	<0.01	-2.07	<0.01	-6.09	<0.01	-33.41	<0.01
m10wspd	32.93	<0.01	26.21	<0.01	33.98	<0.01	32.18	<0.01	44.38	<0.01
Precipitation	10.19	<0.01	8.00	<0.01	-	-	20.28	<0.01	28.00	<0.01
tair_dew	-	-	21.94	<0.01	-27.37	<0.01	-	-	85.90	<0.01
Time 4 am	1,032.90	<0.01	1,040.41	<0.01	791.83	<0.01	1,425.15	<0.01	7,462.57	<0.01
Time 8 am	1,003.15	<0.01	1,013.48	<0.01	741.41	<0.01	1,372.85	<0.01	7,572.18	<0.01
Time 12 pm	902.55	<0.01	935.58	<0.01	657.76	<0.01	1,235.84	<0.01	7,495.51	<0.01
Time 4 pm	863.87	<0.01	861.75	<0.01	653.28	<0.01	1,198.26	<0.01	7,461.19	<0.01
Time 8 pm	1,013.51	<0.01	1,025.61	<0.01	757.16	<0.01	1,442.44	<0.01	7,525.68	<0.01
Time 12 am	1,016.46	<0.01	1,132.93	<0.01	731.93	<0.01	1,474.37	<0.01	7,377.35	<0.01
Rain 3 hr	145.06	<0.01	102.49	<0.01	131.49	<0.01	169.85	<0.01	272.25	<0.01
Rain 6 hr	51.69	<0.01	86.38	<0.01	-	-	76.97	<0.01	-	-
Rain 12 hr	22.05	<0.01	-	-	-	-	-	-	-	-
Rain 24 hr	-	-	39.43	<0.01	-	-	-	-	-	-
Water	-	-	-	-	47.03	<0.01	-	-	-	-
No. of Observations	16,508		5,818		4,760		4,619		1,311	
R-Square	0.74		0.77		0.75		0.72		0.71	
Adj. R-Square	0.74		0.77		0.75		0.72		0.71	
AIC	250,903		87,763		72,320		70,124		20,185	
RMSE	483.04		456		481.30		479		531.13	

Table A16 WLS Regression Model for Visibility Data <2000m (Near to Roads)

Variable	All		Elevation < 50 m		Elevation 50 m to 250 m		Elevation 250 m to 750 m		Elevation > 750 m	
	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value
Elevation	-0.03	<0.01	-1.05	<0.01	1.71	<0.01	-	-	-4.55	<0.01
Cloud cover	-2.14	<0.01	-2.10	<0.01	-2.36	<0.01	-7.24	<0.01	-35.37	<0.01
m10wspd	34.57	<0.01	25.65	<0.01	37.12	<0.01	36.62	<0.01	43.43	<0.01
Precipitation	11.09	<0.01	9.62	<0.01	-	-	20.80	<0.01	24.94	<0.01
tair_dew	-	-	22.87	<0.01	-22.06	<0.01	-	-	96.73	<0.01
Time 4 am	1,022.58	<0.01	1,021.27	<0.01	780.40	<0.01	1,523.79	<0.01	8,377.35	<0.01
Time 8 am	994.35	<0.01	993.84	<0.01	715.10	<0.01	1,480.63	<0.01	8,494.15	<0.01
Time 12 pm	887.23	<0.01	916.62	<0.01	635.89	<0.01	1,325.42	<0.01	8,422.79	<0.01
Time 4 pm	838.68	<0.01	832.01	<0.01	631.07	<0.01	1,274.96	<0.01	8,388.28	<0.01
Time 8 pm	1,012.60	<0.01	1,010.51	<0.01	728.70	<0.01	1,531.63	<0.01	8,440.80	<0.01
Time 12 am	1,004.58	<0.01	1,112.75	<0.01	703.43	<0.01	1,581.43	<0.01	8,319.78	<0.01
Rain 3 hr	144.49	<0.01	100.23	<0.01	143.49	<0.01	192.33	<0.01	330.23	<0.01
Rain 6 hr	47.20	<0.01	87.54	<0.01	-	-	74.69	<0.01	-	-
Rain 12 hr	22.84	<0.01	-	-	-	-	-	-	-	-
Rain 24 hr	-	-	39.54	<0.01	-	-	-	-	-	-
Water	-	-	-	-	45.50	<0.01	-	-	-	-
No. of Observations	16,508		5,818		4,760		4,619		1,311	
R-Square	0.97		0.97		0.95		0.92		0.94	
Adj. R-Square	0.97		0.97		0.95		0.92		0.93	
AIC	213,342		74,906		63,663		63,806		18,187	
RMSE	154.85		151.00		193.86		241.49		248	