A STUDY OF SPATIAL DISTRIBUTION OF OZONE AND ITS CORRELATION WITH METEOROLOGICAL PARAMETERS AND PRECURSORS

By

DHARMIK VADEL

Bachelor of Science in Civil Engineering

Calicut University

Kerala, India

2002

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
December, 2005

A STUDY OF SPATIAL DISTRIBUTION OF OZONE

AND ITS CORRELATION WITH METEOROLOGICAL

PARAMETERS AND PRECURSORS

Thesis Approved:

Dr. DEE ANN SANDERS

Graduate Advisor

Dr. JOHN N. VEENSTRA

DR. GREGORY G. WILBER

Dr. A. GORDON EMSLIE

Dean of the Graduate College

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my advisor, Dr Dee Ann Sanders for her guidance and support throughout this research program, and my graduate study. Her enthusiasm and expertise was a constant source of motivation for this work.

I extend my sincere thanks to my thesis committee member Dr John N Veenstra and Dr Gregory G Wilber for their support during my graduate study.

This research would not have been possible without extensive air quality data provided by Technology Transfer Network of EPA and Oklahoma Mesonet System.

I wish to thank Aruna for all her help and support during my thesis. I also thank all my friends for their help.

I would like to thank my parents Rasik Vadel and Jyoti Vadel for their consistent love and support. I thank my brother Hardik Vadel for supporting me whenever I need.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

The air we breathe is continuously being polluted by human activities such as driving cars and trucks, burning coal, oil, and other fossil fuels, and manufacturing chemicals. Even smaller everyday activities such as dry cleaning, degreasing and painting operations add gases and particles to the air we breathe. As per the U.S. Environmental Protection Agency (EPA), an adult breathes approximately 3000 gallons of air everyday, and children are even more susceptible to air pollution since they breathe more air per pound of body weight. Air pollution thus is a problem for all of us. Health effects from exposure to high levels of air pollution include irritation to eyes and throat and breathing difficulties. Long term exposure to air pollution may cause cancer, respiratory problems and in extreme cases, it can even cause death.

The Clean Air Act requires EPA to set the National Ambient Air Quality Standards (NAAQS) for pollutants considered harmful to public health and the environment. The National Ambient Air Quality Standards are set for six principal pollutants, which are called “criteria” pollutants. Ozone, O_3, is classified as one of these criteria pollutants. Ozone pollution in the lower troposphere is a major issue in several urban and suburban areas. The original NAAQS for ozone was a 1-hr average of 0.12 part per million (ppm), not to be exceeded more than once a year. In July 1997, EPA revised the standard to an 8-hr average of 0.08 ppm that is based on a 3-year average of the fourth-highest daily maximum 8-hr average ozone concentrations measured at each monitor within an area (EPA, 1998). Since, the ozone concentrations are measured to the nearest part per billion (ppb), the smallest concentration required to exceed the 1-hr and 8-hr ozone standard are 125 ppb and 85 ppb respectively (Lin et al., 2001). Ozone is not emitted directly into the air, but is formed by the chemical reactions between oxides of nitrogen (NO_x) and volatile organic compounds (VOC) in the presence of sunlight (EPA, 2003). Typically, the necessary conditions for ozone formation are high precursor emissions, heat and sunlight. Hence, high concentrations are generally observed during summer months. Ozone poses a variety of health problems including chest pain, coughing, and can worsen bronchitis, emphysema, and asthma. Long term ozone exposure may also permanently scare the lung tissue (EPA, 2003). Ground level ozone also damages vegetation and ecosystems. EPA estimates the cost of damage related to reduction in crop production due to ozone at $500 million per year. It is thus very important to understand how both the chemical processes and meteorological conditions affect the ozone formation.

Presently Oklahoma is in compliance with the NAAQS. But there are a few sites in Oklahoma that are approaching ozone levels in excess of the NAAQS. This study focuses on identifying the areas in Oklahoma with high levels of ozone and studying the effects of local meteorology on these ozone levels for the time period of 1998 to 2000. The work was conducted in three parts:

- First part is the analyses of temporal distribution of ozone that includes studying 1-hr and 8-hr exceedance patterns and identifying episodes of high ozone levels to study the relation between various meteorological parameters and ozone.

- Second part is performing geostatistical analyses to study the diurnal variation of ozone and conducting spatial analysis to identify areas having high probability of ozone standard exceedances.
- Finally regression analysis was performed to develop forecasting models for three types of responses. These responses are the next day 1-hr ozone, 8-hr ozone and daily maximum 1-hr ozone concentration.

2. LITERATURE REVIEW

2.1 Processes affecting ozone concentrations

A thorough knowledge of basic ozone formation processes will help understand the ozone trends and forecast the effects of weather on ozone and emissions. Ozone is a secondary pollutant formed in the atmosphere through photochemical reactions involving nitrogen oxides (NOx) and volatile organic compounds (VOC) in the presence of sunlight (Zhang et al., 1999). Ozone concentrations usually build in stagnant atmospheric conditions such as light surface winds and high air temperature. However, local conditions are not solely responsible for ozone formation. Transport of ozone and/or precursors can also cause high ozone levels. Thus, role of meteorological parameters is as important as chemical reactions for understanding ground level ozone trends.

Precursors have both anthropogenic (man-made) and biogenic (natural) sources. Some of the major sources of NO_x and VOC include emissions from industrial facilities and electric utilities, motor vehicle exhaust, gasoline vapors and chemical solvents (EPA, 2003). Trees, plants and other vegetation also emit VOCs. In the presence of ultra violet radiation (hv), oxygen (O_2) and oxides of nitrogen react to form ozone. Atkinson (2000) explains the ozone formation process with and without the presence of VOCs. The NO titration actually destroys ground level ozone:

$$NO + O_3 \rightarrow NO_2 + O_2 \qquad (1)$$

NO_2 thus formed reacts with oxygen (ambient O_2 and O_2 from NO titration) in the presence of sunlight to form ozone (O_3):

$$NO_2 + O_2 + hv \rightarrow O_3 + NO \qquad (2)$$

In the presence of VOCs, the degradation reactions of VOCs lead to formation of intermediate RO_2 and HO_2 radicals. These radicals convert NO to NO_2 that finally photolyze to form ozone. The two step process involved in ozone formation with VOCs is:

$$VOCs + NO \rightarrow NO_2 \qquad (3)$$

$$NO_2 + O_2 + hv \rightarrow O_3 + NO \qquad (4)$$

Figure 1 shows the ozone formation process with and without VOC s as described by Atkinson (2000).

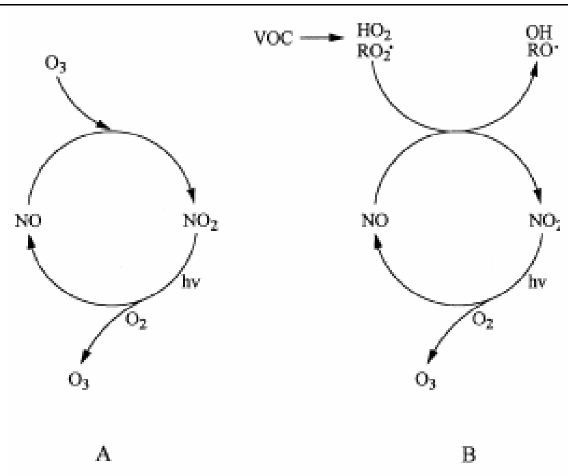


FIGURE 1: Ozone formation process in (a) the absence of VOCs, (b) the presence of VOCs (Atkinson, 2000)

Ozone concentrations are generally higher during afternoons (10:00 am - 5:00 pm) due to ozone-favoring conditions such as high temperature and ozone formation from NO_2. Ozone is monitored at 15 stations located statewide in different counties as per 2000 data. Ozone concentrations are recorded hourly all throughout the year. Figure 2 shows diurnal variation of ozone for five sites in Oklahoma. The ozone concentration for each hour is averaged for the summer season (May to September) for each site. All the five sites show a similar diurnal pattern confirming the above concept of higher ozone concentration in the afternoon hours.

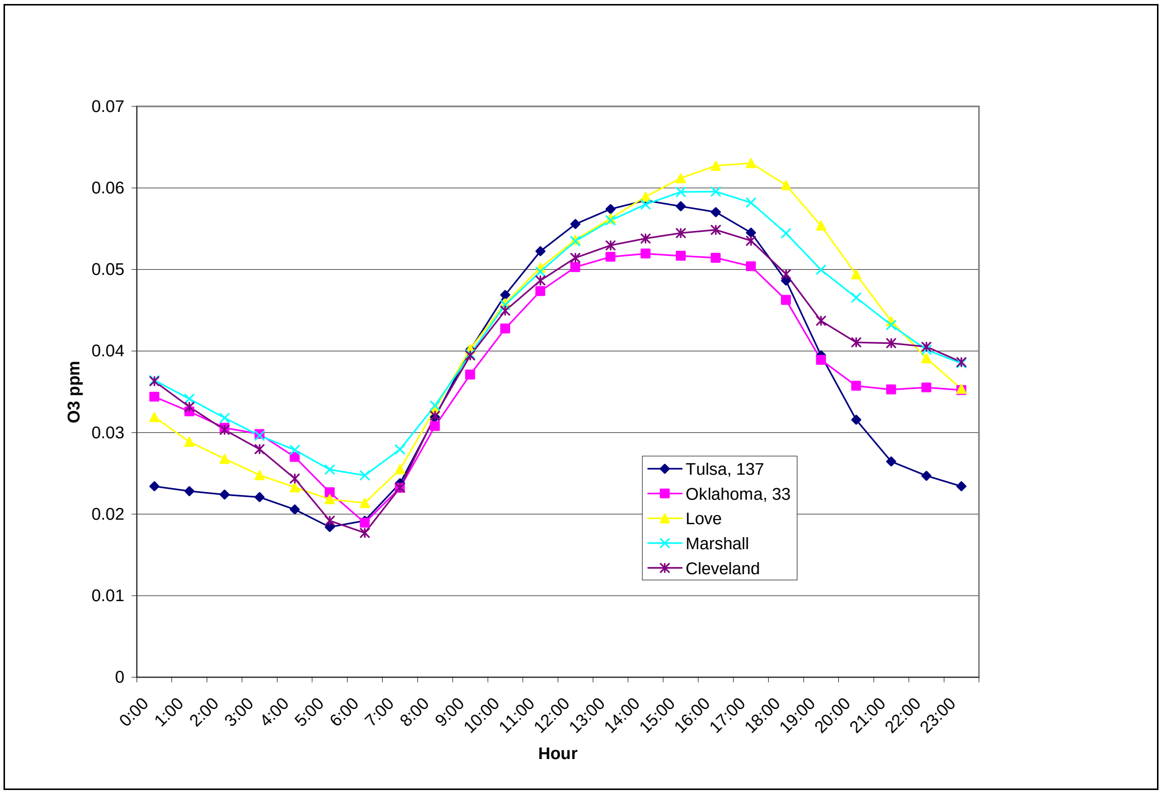


FIGURE 2: Diurnal variation of ozone averaged over May to September, 2000 at 5 sites in Oklahoma (Raw data retrieved from U.S. EPA, Technology Transfer Network (TTN), AQS)

Figure 3 shows a comparison of daily variations of O_3 and NO_2. A lag was observed between occurrence of peak NO_2 concentration and peak O_3 concentration value. This lag time in peak O_3 concentration occurs because NO_2 reacts with sunlight to form O_3. The data in Figure 3 also shows that overall NO_2 concentrations are much lower than O_3 concentrations.

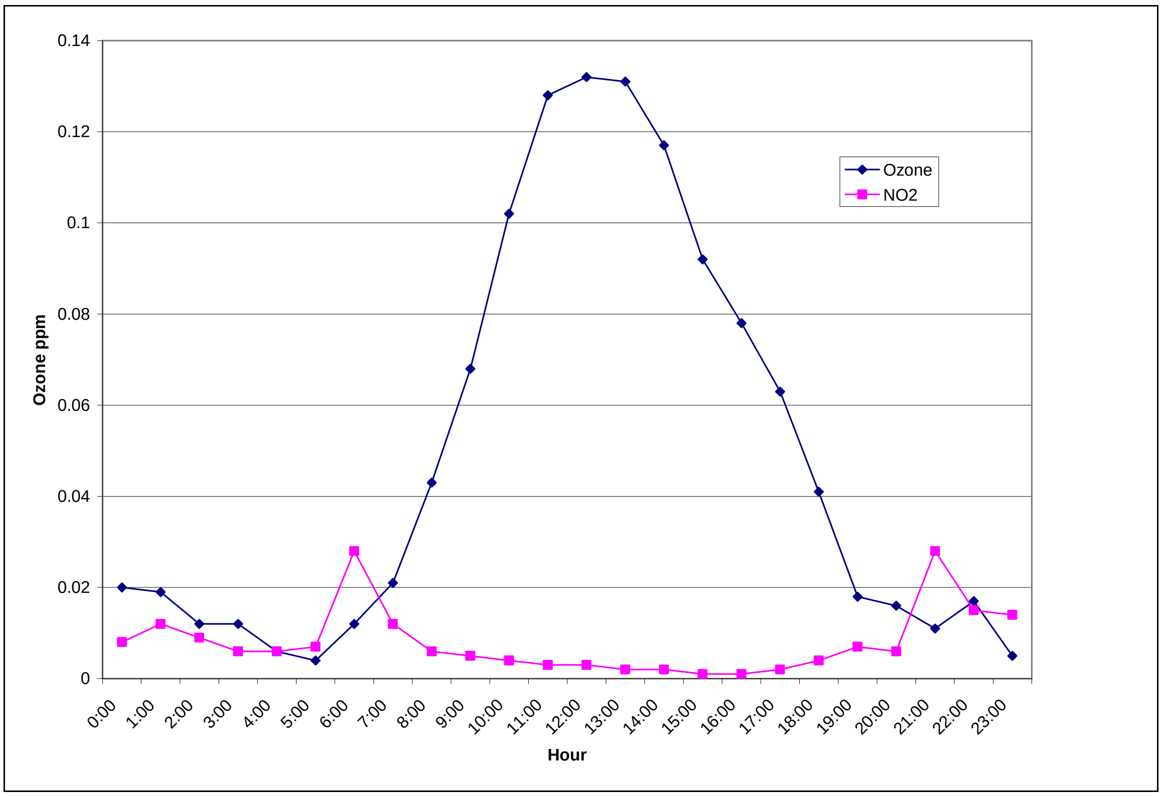


FIGURE 3: Daily variation of O_3 and NO_2 concentrations at Skiatook, Oklahoma September 2, 2000 (Raw data retrieved from U.S. EPA, TTN AQS)

An in depth understanding of variations in meteorological parameters and their impact on ozone concentrations is important in order to choose the right parameters to forecast ozone. A basic knowledge of the physical significance of these parameters in ozone formation is equally important to assess the trends generated from available data. The following discussion explains briefly the influence of various weather parameters on ozone (Ludwig et al., 1995):

- Sunlight - Ultra violet radiation is needed for ozone photochemistry.
- Temperature - Higher temperatures increase the photochemical reaction rates. In addition, evaporative emissions of VOCs increase with high temperature. Biogenic emissions also increase because of high temperature.
- Vertical temperature variation - Vertical mixing is governed by the phenomenon of adiabatic lapse rate (temperature change by height). Strong stable conditions tend to confine the pollutant emissions and ozone concentrations. In addition, aloft temperature inversions inhibit vertical mixing by aiding the confining of emissions.
- Wind - Surface winds serve the basic process of ventilation. Lower wind speeds cause the accumulation of emissions, thus making higher concentration of precursors available for ozone formation. Also, the ozone concentrations formed are retained in stagnant air conditions. Aloft winds affect the ozone formation in a different manner. These winds are primarily responsible for regional transport of air emissions. Ozone and precursors are transported into a region overnight or in early mornings.

These mechanisms control the local ozone concentrations and thus determining the amount of influence of these parameters is necessary in forecasting ozone.

2.2 Concepts in Geostatistics

Geostatistics is the statistical procedure for developing a correlation of spatially distributed random variables and for performing interpolation and areal estimation of these variables (Copper and Istok, 1988). Geostatistics was applied here to study the spatial distribution of ozone in Oklahoma. The hourly data recorded at ozone monitoring stations set up by EPA were used to study the diurnal variations in ozone. It is hence important to understand the basics of geostatistics.

A variogram is a basic tool in geostatistics for describing spatial correlation. A variogram or semivariogram is a graphical representation of a mathematical function called semivariance versus the distance between data values. The semivariance can be calculated using the following equation given by Isaaks and Srivastava (1989)

$$\gamma(h) = \left[\frac{1}{2N(h)}\right]\sum_{i=1}^{N(h)}[z(x_{i+h}) - z(x_i)]^2 \qquad (5)$$

where, N (h) = the number of sample pairs separated by the vector h, z (x_i) is the sample data value at measurement point x_i. There are several variogram models that are available such as linear, spherical, and gaussian. The variogram model is selected from a set of mathematical functions that describe spatial relationships. The model is selected based on comparison of the shape of the experimental variogram and that of the mathematical function. Figure 4 shows a sample spherical semivariogram. A semivariogram is a plot of semivariance versus lag distance.

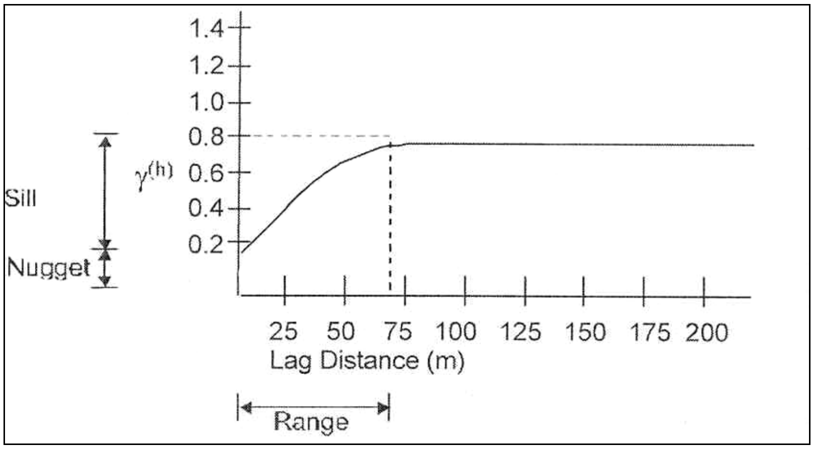


FIGURE 4: A typical spherical semivariogram (PINCOCK Perspectives, 2001)

The important features of a semivariogram, as shown in Figure 4, are range, sill, and nugget, which are explained below:

Range - With the increase in separation distance between pairs, the corresponding variogram value will also generally increase. Eventually, the curve asymptotes, and there is no further increase in the variogram value with increase in separation distance between pairs. The distance at which the semivariogram reaches this constant value is called the range.

Sill - The constant value that the variogram achieves at the range is called the sill.

Nugget - The intercept on the semivariance or Y-axis at zero or very small separation distance is known as nugget. Ideally the semivariance at zero separation distance should be zero, but because of several factors, such as sampling error and short scale variability, there may be a vertical jump from the value of zero at the origin for extremely small separation distances.

Lag - Lag is the average distance for grouping points for variogram calculation (i.e. minimum distance between sample points).

The development of geostatistics has resulted in the generation of an estimation technique commonly known as **"kriging."** Kriging is a geostatistical interpolation method used to estimate the quantity of interest at an unsampled location based on nearby measurements. Kriging is generally associated with the acronym B.L.U.E. for "best linear unbiased estimator." (Isaaks and Srivastava, 1989) Ordinary kriging is "linear" because it estimates the unknown samples values based on weighted linear combinations of the available data; it is "unbiased" since it tries to achieve mean residual or error (m_R) equal to 0; it is "best" because it aims in getting minimum variance of the errors (σ^2_R). (Isaaks and Srivastava, 1989).

Indicator kriging – This procedure determines the probability of data values in a given location to exceed a threshold value (also known as indicator kriging cut-off value), by using the samples in the neighborhood (Isaaks & Srivastava, 1989). Indicator kriging is performed by coding the data values as 1's for those exceeding the threshold values, and 0's for those data values below the threshold value. These indicators are then analyzed to determine their spatial directional variability with a series of experimental variograms. By assessing these variograms, orientations of greatest and least spatial continuity are defined. Corresponding to these directions, variogram models are fitted to the experimental variogram. Then using ordinary kriging for indicator values, the probabilities of exceeding the threshold values at desired grid locations are determined. Geostatistical software provides visual outputs in the form of contour maps of areas with different probabilities of exceeding the threshold values. The software GS+ (Geo

Environmental Sciences version 7.0) was used in this research for generating probability maps of areas exceeding the ozone standards.

2.3 Concepts in Regression Analysis

Regression is a statistical technique used to determine relationships between variables. For ozone forecasting, regression equations are used to determine the relationship of response (i.e. ozone concentrations) with predictors (i.e. temperature, wind speed, relative humidity, etc.). Statistical parameters such as R^2 value and p-value are used to choose the best regression model. Since more than one variable affects ozone production, a multivariate linear regression (MLR) is used to develop a regression equation to forecast ozone. A MLR model can be given in a generic form as:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + \varepsilon \qquad (6)$$

where y is the response variable

β_0 is the intercept

β_1 is the slope coefficient for the first explanatory variable

β_2 is the slope coefficient for the second explanatory variable

β_k is the slope coefficient for the kth explanatory variable, and

ε is the remaining unexplained noise in the data (error) (Helsel & Hirsch, 1995).

There are three types of procedures for selecting the best MLR model, viz. forward selection, backward elimination and stepwise regression. **Forward selection** starts with an intercept and adds variables to the equation one at a time. Once a variable is selected it stays in the model. Evaluation for parameter to be selected in the model is done using partial F or t statistics. The variable with the highest significant partial F or t-statistic is included, and the process repeats until either all available variables are included or no new variables are significant. **Backward elimination** starts by forming a regression equation using all the variables. It starts eliminating variables with the least partial - F statistic (lowest |t|). It stops its procedure when all the remaining variables are significant. There are certain variables that individually may not have a significant impact on the model, but might have a combined significance with another variable. Neither of the approach capitalizes on this situation. **Stepwise regression** combines the concept of both forward selection and backward elimination. It alternates between adding and removing variables, checking the significance of individual variables within and outside the model. Variables that show initial significance while entering the model will be removed later if they lose significance (Helsel & Hirsch, 1995).

2.4 Past tropospheric ozone studies

G.C.Tiao (1983) plotted monthly averages of ozone from January 1958 to 1978 that showed that there is a strong seasonal pattern over time. He found there was a phase difference in the seasonal pattern between the northern and southern hemisphere. This was evident from the findings that, for northern sites, maximum ozone values were found in March and April and minimum concentrations were found in September and October, and the opposite was true for the southern sites. Several authors including Angell and Korshover (1973), London and Kelly (1974), and Komhyr et al. (1971), have used linear regression to study the ozone trends in their analysis.

Ludwig et al. (1995) performed cluster and empirical orthogonal function (EOF) analysis to study the spatial patterns of daily ozone maxima associated with violations in Pinnacles National Monument in Monterey Bay Unified Air Pollution Control District. This study focused on finding answers to questions such as what meteorological conditions were associated with the NAAQS exceedances in the Pinnacle area, whether the exceedances were mainly because of local meteorological conditions or because of regional transport, and finally studying the ozone patterns associated with the high-ozone concentrations. Their analysis used daily maximum, hour-average ozone concentrations for various sites in the chosen area. Meteorological parameters of focus were temperature, wind, pressure, and temperature difference between coastal area and central valley, depth of mixed layer, and inversion. Their findings suggest that ozone violations were attributed to aloft transport winds and stable inversion layer.

An association of meteorological parameters temperature, pressure, dewpoint, and wind speed with diurnal maximum ozone concentration (DMOC) was studied by Vukovich (1995) for the eastern United States. The only persistent relationship between ozone and meteorological parameters was found between ozone and surface wind speed. Surface wind speeds were found to be negatively correlated with ozone. It has also been shown by Vukovich (1994) that O_3 concentrations increase with temperature for rural and urban sites.

Previous efforts to relate ozone concentration data to surface meteorological variables found that temperature, wind speed, relative humidity, and cloud cover were important variables (National Research Council, 1991). Bloomfield et al. (1996) modeled ozone concentrations using meteorological parameters such as wind direction, dew point temperature, and sea level pressure. Profiles of regression surfaces generated indicated that ozone concentrations had a characteristic functional relationship with temperature but the scale of response was dependent on the relative humidity. A nonlinear parametric model was developed to account the effect of seasonal dependence, and it was found that 80% of the variance in ozone concentration data was explained by incorporating the seasonal trend factor into the model. The proposed Neural Networks by Chaloulakou et al. (2003) provided better forecast of summertime ozone concentrations over multiple Linear Regression models based on same set of input variables. Their neural networks could also be employed to calculate the maximum ozone concentrations for days that had unavailable data.

Temperature was the most significant meteorological parameter affecting ozone concentrations for the studies conducted by Olszyna et al. (1997). Their study found that a linear combination of temperature and sum of all nitrogen oxide compounds (NO_y) provided a better association with O_3 as compared to their impact individually. The number of O_3 molecules produced per NO_y molecule present varied from 4 to 14 for a temperature range of (22 °C – 33 °C), while temperature had negligible effect on the number of O_3 molecules produced per NO_x molecule consumed. They concluded that at higher temperature O_3 would form near the emission sources while at lower temperatures the same amount of O_3 forms eventually at further distance downwind of the source.

The use of generalized additive models and cluster analysis showed that daily maximum surface temperature, the daily mean v component of the surface wind and total daily global radiation were the most important covariates associated to high ozone concentrations (Davis et al., 1999). The wind components were identified as three components – the west-east component (u, positive from the west), south-north component (v, positive from the south) and opaque cloud cover (opcov).

Regression-based methods are aimed primarily at forecasting or performing trend analysis and to some extent studying the underlying mechanisms. Thompson et al. (2001) suggest that use of statistical methodology should be process driven. A thorough understanding of underlying mechanisms should be the basis of selecting a statistical procedure for analysis. The choice of methodology also depends on the type of analysis (e.g. trend analysis, forecasting or assessing changes in mean level etc.) desired. Thus it is very important to set forth the objectives of the study and the type of analysis required before selection of a statistical procedure.

Considerable past efforts have been made to determine the dependence of ozone on volatile organic compounds (VOC) and oxides of nitrogen (NO_x) in different urban and non-urban regions. A study of weekend versus weekday ozone concentrations provides a means for empirically investigating the impact of VOC and NO_x reductions on ozone formation (Blanchard et al., 2001). The study was based on the fact that the relative ozone precursor concentrations were significantly different on weekends than on weekdays. Their studies for central California showed that most VOC-limited sites exhibited higher weekend peak ozone concentrations, whereas NO_x-limited sites generally had higher weekday peak ozone levels.

Lin et al. (2001) examined the long term trends in exceedances of the new and the old ozone standards in the United States over the past two decades. They focused their trend analysis on summer months (June-August) based on their findings that these summer months accounted for 81% of all the exceedances in Northeast, 60% in the Southeast, 70% in the Northwest, and 58% in the Southwest.

Vukovich (2003) showed that temperature and high water vapor were necessary, but not a sufficient condition for high ozone to be found in Baltimore-Washington corridor. They suggest that sufficiency condition is satisfied when significant amounts of solar radiation reach the site and when stagnation conditions prevail at the same time. They used three data sets - complete 15-year summer data set, 15-year summer data set and 15-year data set that included only those days with daily maximum ozone concentrations (DMOC) greater than or equal to 100 ppb (DMOC≥100 ppb). For the complete data set, temperature and dew point were the most significant meteorological parameters, for the summer data set of 15 years, temperature and sky cover were the significant meteorological parameters and for the data set with the conditions of DMOC≥100 ppb the meteorological parameters of highest influence were surface wind speed and sky cover. A regression model of daily mean ozone concentrations on the meteorological parameters like temperature, relative humidity, and wind speed can explain approximately 70% of the ozone variability (Tarasova and Karpetchko, 2003).

Qin et al. (2004) analyzed the annual trends, temporal and spatial distributions and weekend effect of exceedances for both the 1 hr and 8 hr O_3 NAAQS in California south coast Air Quality Management District (AQMD). The found that the 8 hr O_3 air quality standard greatly increased the number of non-attainment episodes in AQMD in the period 1997-2001.

Casado et al. (1994) applied geostatistical and visualization procedures to analyze hourly ozone measurements collected from 29 stations in the southeastern United States. They generated semivariogram for every hour of the day to study the diurnal pattern of ozone. Local factors were dominant over the ozone production process during daytime hours and regional factors and small scale uneven variations became more dominant in the ozone depletion process during the evening hours.

3. SITE INFORMATION

EPA has monitoring stations for measuring ground level ozone and oxides of nitrogen and nitrogen dioxide at various sites in Oklahoma. These stations record hourly observations for ozone, NO_x and NO_2. These data are available in the Technology Transfer Network, Air Quality System (TTN AQS) of EPA. Data from three years – 1998, 1999, and 2000 – were used for analysis. Figure 5 shows the location of the ozone monitoring stations in Oklahoma. A description of site specific information such as the latitude and longitudes, elevation above mean sea level (MSL), type of site and site identification is given in Table 1.

Meteorological data were obtained from Oklahoma Mesonet system. The Oklahoma Mesonet consists of over 110 automated stations covering Oklahoma. At every Mesonet site environmental parameters are measured using set of instruments on or near a 10m tall tower. Recordings are made every five minutes and then transmitted to a central facility 24 hours a day throughout the year. Figure 6 shows a typical Oklahoma Mesonet site with approximate positions of various recording instruments used. Table 2 lists the variables measured at Mesonet site, the type of sensors used, the standard units and a brief description of measurement technique.

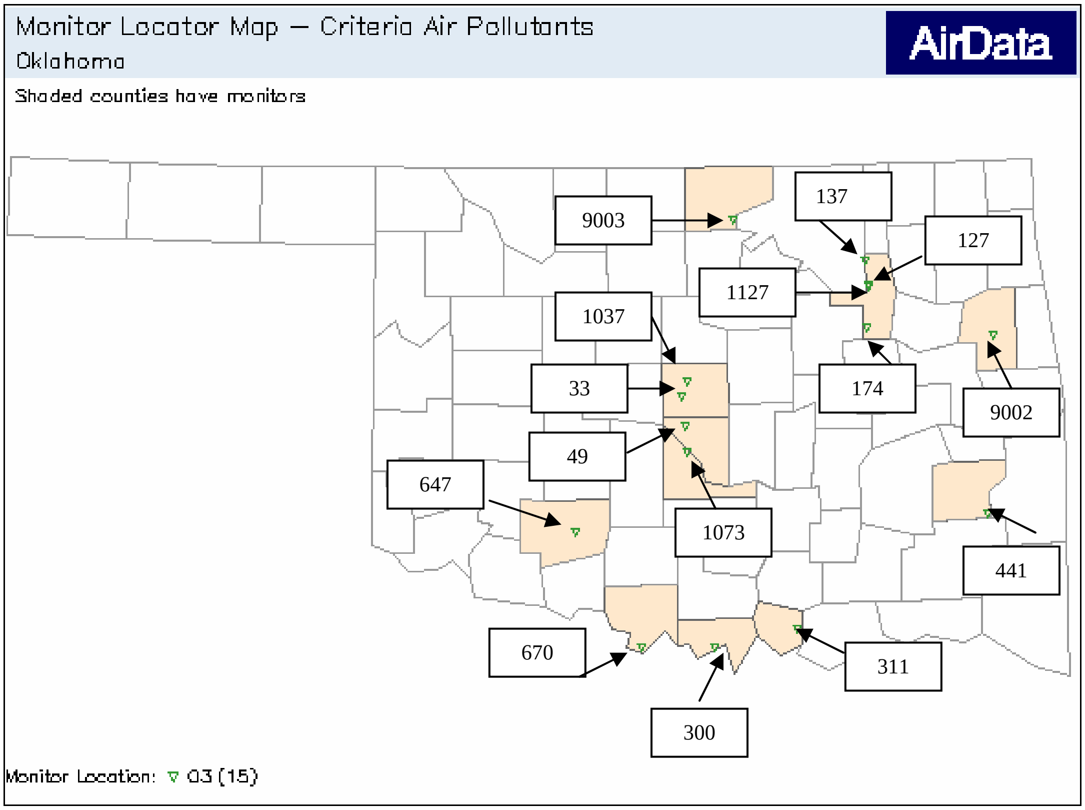


FIGURE 5: Location of ozone monitors in Oklahoma, 2000 (U.S. EPA Office of Air and Radiation, AQS)
Note: Numbers in the boxes indicate Site Identification numbers.

TABLE 1: Description of ozone monitoring sites in Oklahoma, 2000 (U.S. EPA Office of Air and Radiation, AQS)

Site Identification (ID)	County	Type of Setting	Site Elevation Above MSL	Latitude	Longitude
9002	Cherokee	RURAL	238	35.855	-94.986111
49	Cleveland	SUBURBAN	1170	35.32	-97.488333
647	Comanche	RURAL	347	34.648889	-98.366111
670	Jefferson	RURAL	0	33.894167	605673
9003	Kay	RURAL	335	36.662778	-97.074444
441	Latimer	RURAL	0	34.709722	-95.0725
300	Love	RURAL	1	33.8825	-97.275833
1073	McClain	RURAL	0	35.161389	-97.473889
311	Marshall	RURAL	205	33.995	-96.630278
33	Oklahoma	URBAN AND CENTER CITY	379	35.516944	-97.505833
1037	Oklahoma	SUBURBAN	354	35.612778	-97.472222
127	Tulsa	SUBURBAN	198	36.198056	-95.973889
137	Tulsa	SUBURBAN	200	36.373889	-96
174	Tulsa	RURAL	221	35.938889	-96.0025
1127	Tulsa	URBAN AND CENTER CITY	167	36.206602	-95.976584

TABLE 2: Summary of relevant variables measured at Mesonet site (Oklahoma Mesonet)

Variable	Type	Sensor used	Standard Units	Description
TAIR - Air temperature	Core variable	Thermistor	degrees Celsius	The average temperature measured at 1.5 m (4.9 ft) above the ground.
WSPD - Average wind speed	Core variable	Propeller vane anemometer	meter per second	The average wind speed without regard to direction measured at 10 m (32.8 ft) above the ground.
WDIR - Average vector wind direction	Core variable	Propeller vane anemometer	described in degrees in a circle (from north), where north is 0 degrees, east is 90 degrees, south is 180 degrees and west is 270 degrees	The vector average wind direction at a height of 10 m (32.8 ft) above the ground. Note that wind direction always describes the direction from which the wind is blowing and not toward which the wind is blowing.
RELH - Relative humidity	Core variable	Sorption sensor	described as a percentage	The average relative humidity at a height of 1.5 m (4.9 ft) above the ground.
PRES - Station pressure	Core variable	Barometer	millibars (mb)	The average station pressure.
SRAD - Solar radiation	Core variable	Pyranometer	watts per square meter	The average amount of solar radiation received at the sensor. A tripod supports the sensor at a height of 1.8 m (5.9 ft) at the southern end of the site, a position selected to avoid reflections from the tower and to ensure that no shadow from the tower obscures the sensor.

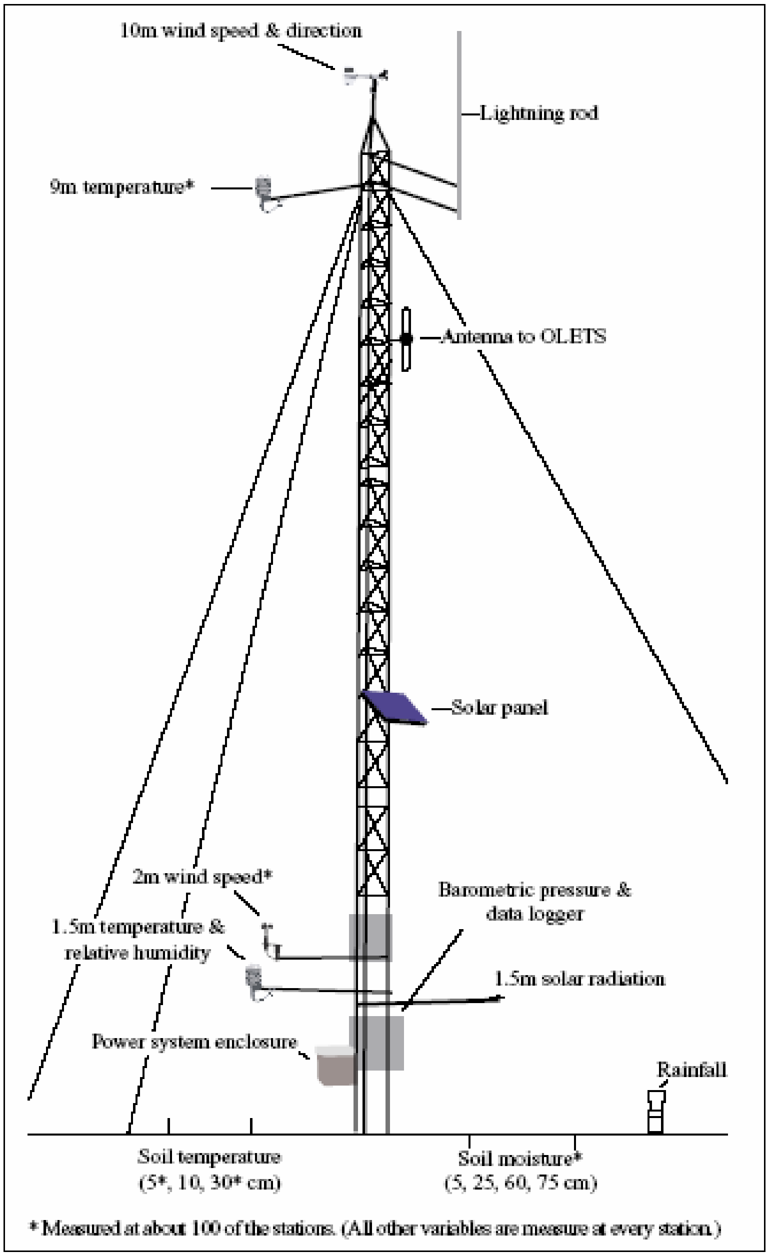


FIGURE 6: Schematic of an Oklahoma Mesonet site (Oklahoma Mesonet)

4. TEMPORAL DISTRIBUTION OF OZONE

4.1 Method

Several studies have shown that the 8-hr standard is more difficult to achieve than the 1-hr standard. Studies conducted by Velasco et al. (2000) showed that five new non-attainment areas for the 8-hr standard were identified in California during 1997 to 1999. The existing status of Oklahoma is complete attainment for 1-hr and 8-hr ground level ozone. However, a few sites in Oklahoma have shown high ozone concentrations. Exceedance in this section is defined as those days with 1-hr ozone concentrations higher than 0.12 ppm and/or 8-hr ozone concentrations higher than 0.085 ppm for the year considered in the study. These exceedances are not identified by averaging the ozone concentration values over three year period. However, a cut off value of 0.12 ppm was chosen for identifying 1-hr ozone exceedance days since the number days exceeding 1-hr ozone concentration value of 0.125 ppm in Oklahoma were very few. Hence using the 0.12 ppm value did identify a few episodes to compare it with the 8-hr exceedance episodes. 1-hr ozone data for 1998-2000 was analyzed separately for each year to identify high ozone episodes. The 8-hr O_3 concentrations were calculated as moving averages of 1-hr O_3 values. Exceedances were identified by applying an ‘If’ criteria to all the 1-hr O_3 values in Microsoft Excel.

The 1-hr O_3 exceedances were identified by the 'If' criterion that checks each 1-hr O_3 value with the cut-off value of 0.12 ppm. If the 1-hr value was higher than 0.12 ppm, it was identified as exceedance. A similar criterion was applied for 8-hr O_3 values, but the cut-off value in this case was 0.085 ppm. Multiple exceedances in a day were not accounted separately. Thus in case of multiple exceedances only one exceedance was counted per day. Ozone patterns were also analyzed to study the following distributions:

- Distribution of ozone exceedances by day of the week.
- Distribution of ozone exceedances by hour of the day.
- Distribution of ozone exceedances by month of the year.

The aim of conducting this analysis is to provide background information required to build a detailed forecasting model to predict ozone concentration in Oklahoma. An attempt is made here to develop a basic multilinear regression model for Skiatook, a suburban site in Tulsa. Description of the model development and the procedure used to develop these models is provided in chapter 5.

4.2 Results and Discussion

Analysis was performed to compare the number of exceedance episodes that occurred following the implementation of 1-hr and 8-hr ozone standards. Figure 7 shows the total number of 1-hr O_3 and 8-hr O_3 exceedance days that occurred annually from 1998 to 2000.

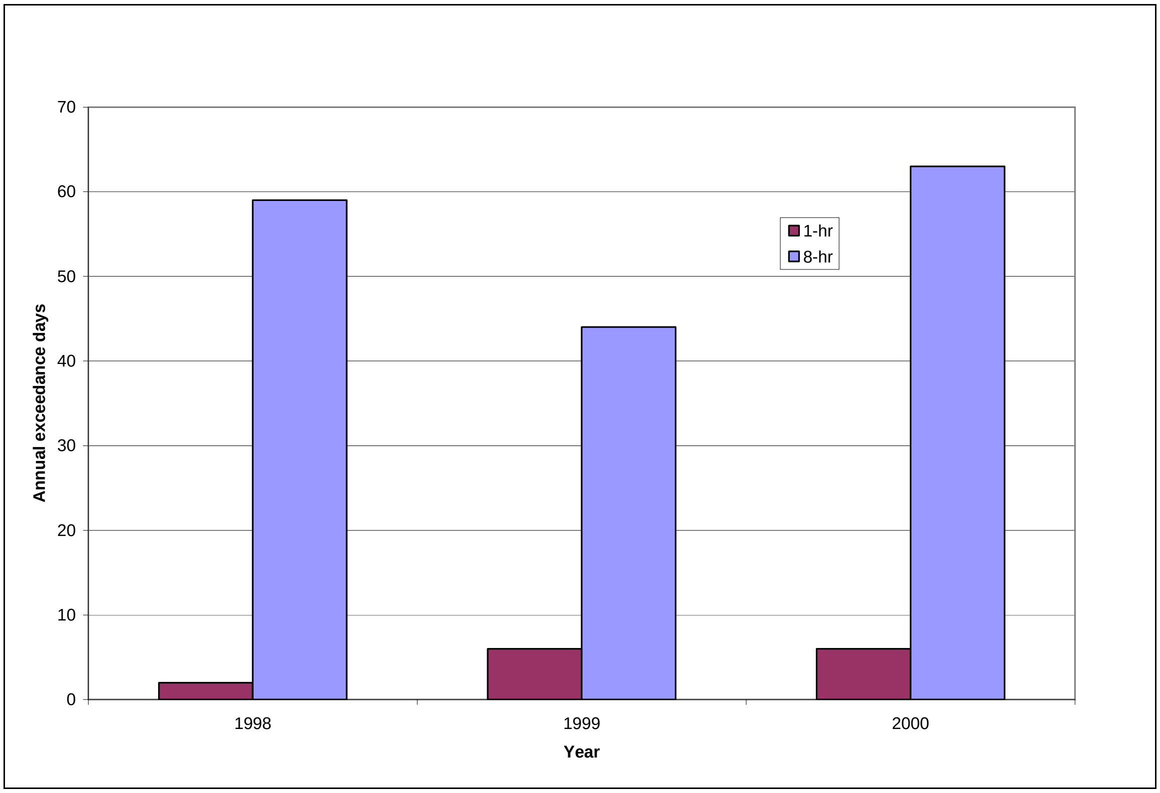


FIGURE 7: Distribution of total annual 1-hr and 8-hr exceedance days from 1998-2000 in Oklahoma (Raw data retrieved from U.S. EPA TTN AQS)

It is observed that the number of 8-hr exceedance days is noticeably higher as compared to the number of 1-hr exceedance days. A county wise distribution of the number of exceedances of 1-hr and 8-hr O_3 standard is shown in Figure 8. These exceedances are actually obtained by analyzing data recorded at monitoring sites in each county. Most of the counties have only one site, except for Oklahoma and Tulsa counties. Hence the graph is labeled as county-wise distribution of exceedances with the actual exceedances being observed at the sites in the counties.

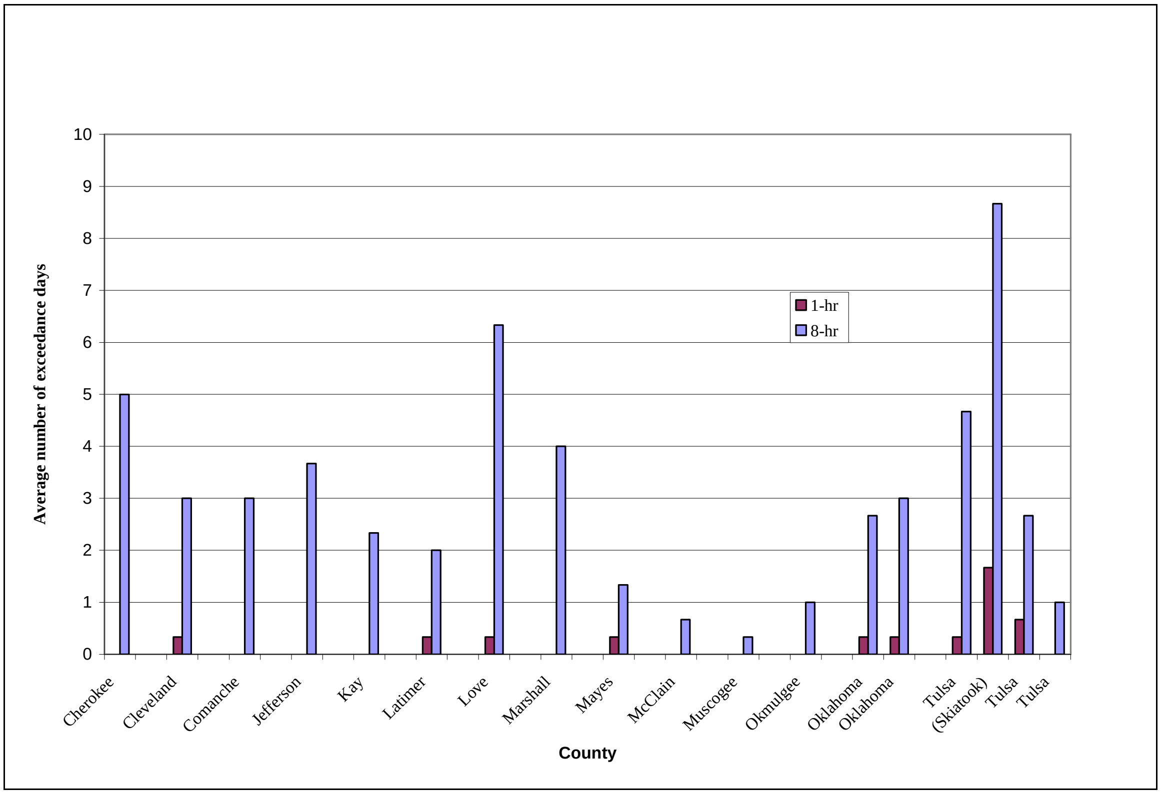


FIGURE 8: County-wise distribution of average 1-hr and 8-hr ozone exceedance days from 1998-2000 in Oklahoma (Raw data retrieved from U.S. EPA, TTN AQS)

Skiatook (Site Id-137), a site in Tulsa, was found to have the highest number of 1-hr and 8-hr exceedances. The 8-hr exceedances were approximately five times the 1-hr exceedances for Skiatook. Other sites that had high number of exceedances were Cherokee and Love. This study supports the work of several other authors that the 8-hr O_3 standard is more difficult to achieve than the 1-hr O_3 standard. Sather et al. (2000) showed that exceedances of the 8-hr O_3 standard were about four times the exceedances of the 1-hr O_3 in Dallas in the summer of 1998. Y.Qin et al. (2004) conducted similar work in the California South Coast Air Quality Management District. They found that the number of exceedances of 8-hr O_3 standard were about 13.5% higher than exceedances of 1-hr standard. The statistics presented above indicate the importance of Skiatook as a potential site to exceed the ozone standard in the future. Spatial analysis to further indicate high exceedances in Skiatook is provided later in section 5.1. The data of Skiatook was hence used for conducting detailed analyses to study the effect of meteorological parameters on peak ozone days.

Once a good understanding of the ozone formation process is developed, it is important to identify the trends in ozone in the specific area where an ozone forecasting model is to be developed. The result of the above analyses indicates the importance of Skiatook, since it has the maximum days of 1-hr and 8-hr ozone standard exceedances. The temporal distribution of ozone was thus studied for the site in Skiatook. The analysis was focused to answer questions relevant to ozone exceedances such as:

- During what months are the 1-hr and 8-hr ozone exceedance likely to occur?
- At what time of the day the highest ozone concentrations occur?
- Do ozone concentration peaks vary by day of the week?

Figure 9 shows the distribution of average annual days of 1-hr and 8-hr ozone exceedances by month for Skiatook. The entire occurrence of exceedance days were found to be in July, August and September. This is very much expected since ozone conducive conditions such as high surface temperature and high solar radiation are generally available during the summer months of May to September. Figure 10 shows the distribution of average monthly maximum temperature and average monthly highest temperatures recorded in Skiatook. A general high temperature trend during the period of May to October was observed. This plot thus supports the explanation made above of having ozone favoring conditions such as high temperature during the months of July to September.

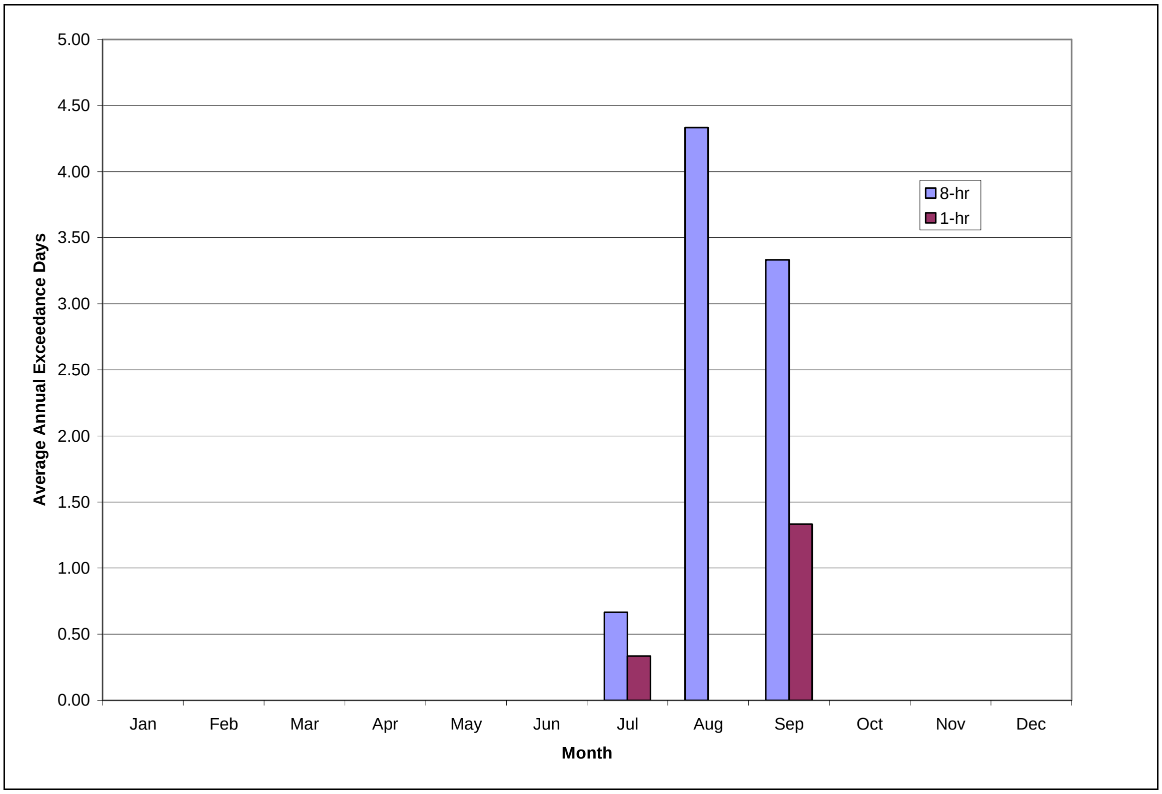


FIGURE 9: Distribution of average number of days with 8-hr and 1-hr exceedances by month from 1998-2000 for Skiatook

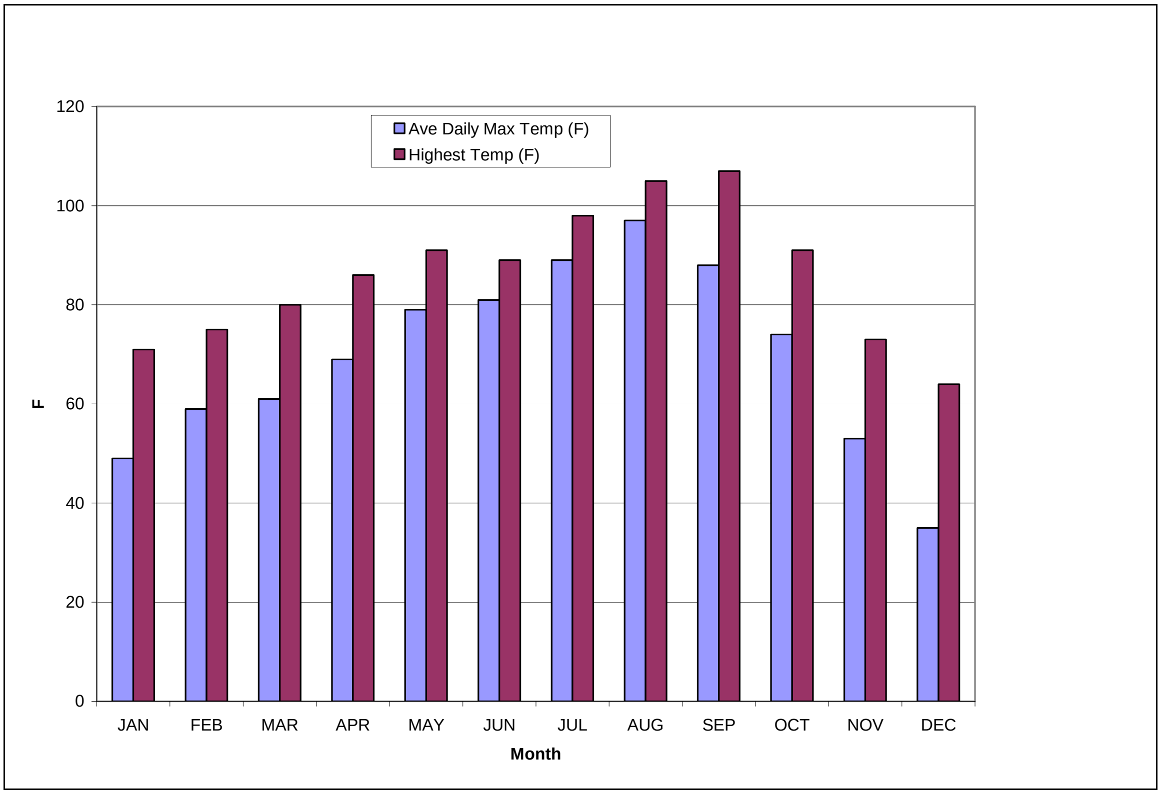


FIGURE 10: Month-wise distribution of average daily maximum temperature and highest recorded temperature in Skiatook, 2000

Figure 11 shows a distribution of 8-hr ozone exceedances by hour of day. 8-hr exceedances were generally observed during late mornings and afternoons around 10:00 a.m. to 5:00 p.m. The maximum number of exceedances took place at 2:00 pm. Ozone forms in presence of sunlight and hence as the day progresses around afternoon a considerable amount of NO_x has been utilized to produce ozone in presence of sunlight. Thus, higher ozone concentrations are observed during the afternoon.

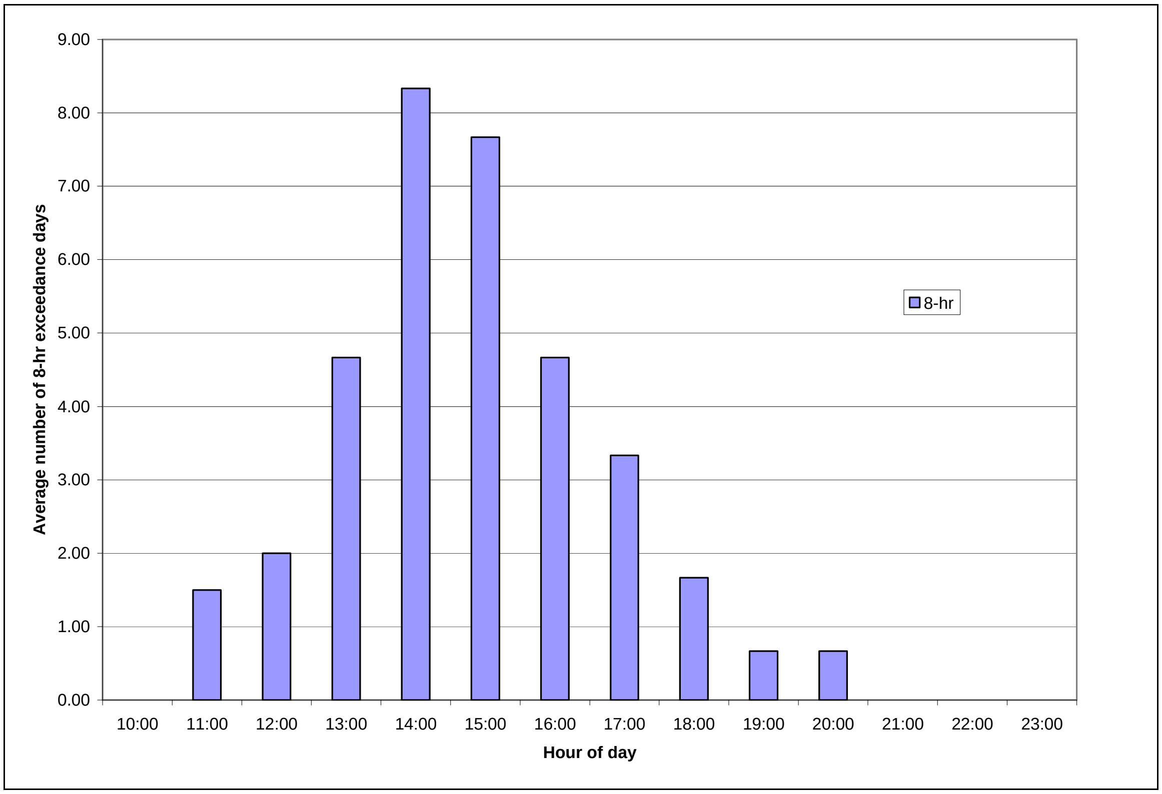


FIGURE 11: Distribution by hour of occurrence of 8-hr average ozone concentration on days that exceeded 0.085 ppm from 1998 - 2000 in Skiatook

Another important aspect to be studied in order to develop a detailed ozone forecasting model is the occurrence pattern of high ozone concentrations on the day of the week. Figure 12 displays the distribution of the occurrence of 1-hr and 8-hr ozone exceedances by day of week for Skiatook. Tuesdays, Fridays and Saturdays showed higher exceedances as compared to other days of the week. The occurrence pattern of ozone exceedances for Skiatook is also compared to that of the entire state. Figure 13 shows exceedance day trends by day of week for Oklahoma. A similar pattern of Tuesdays and Saturdays showing higher exceedances was observed for the state. However, no specific weekend-weekday effect was observed. A few episodes of high ozone concentrations were identified by assessing the raw ozone data available for Skiatook. An investigation was also carried out to find the impact of meteorological parameters on days with high ozone concentrations.

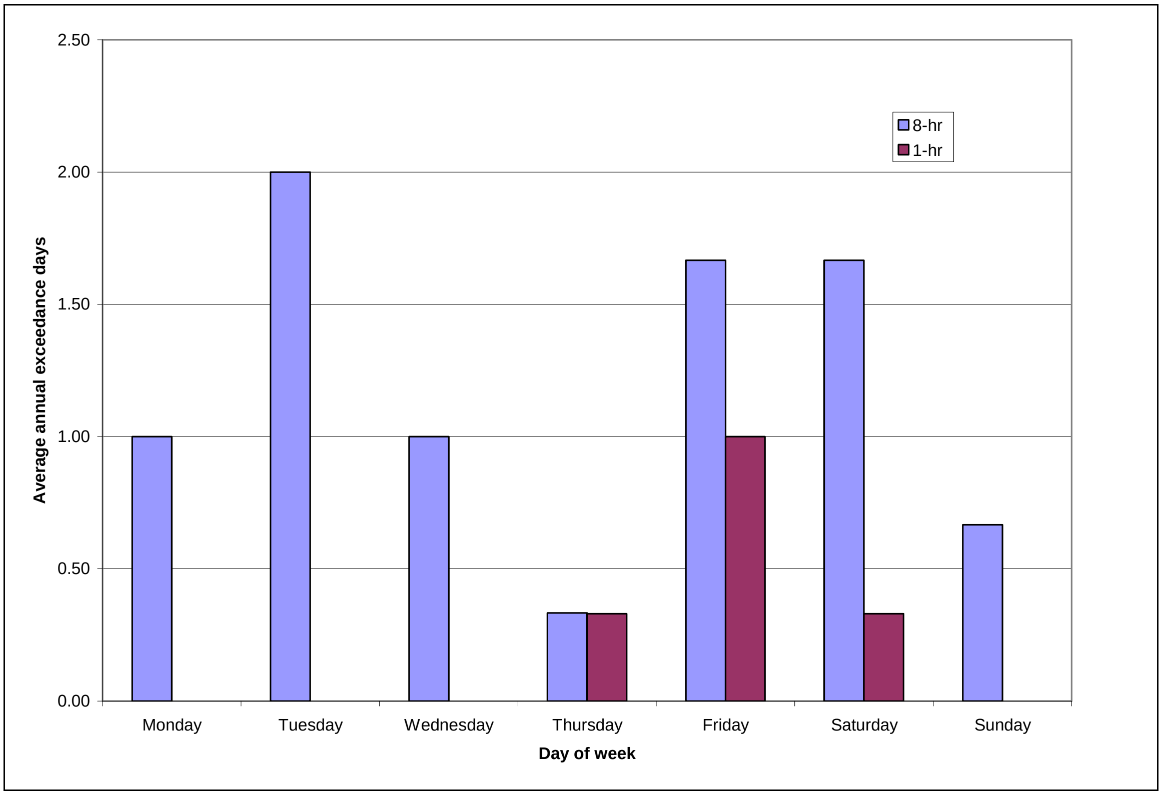


FIGURE 12: Distribution of the average number of 8-hr and 1-hr exceedances by day of week from 1998-2000 for Skiatook

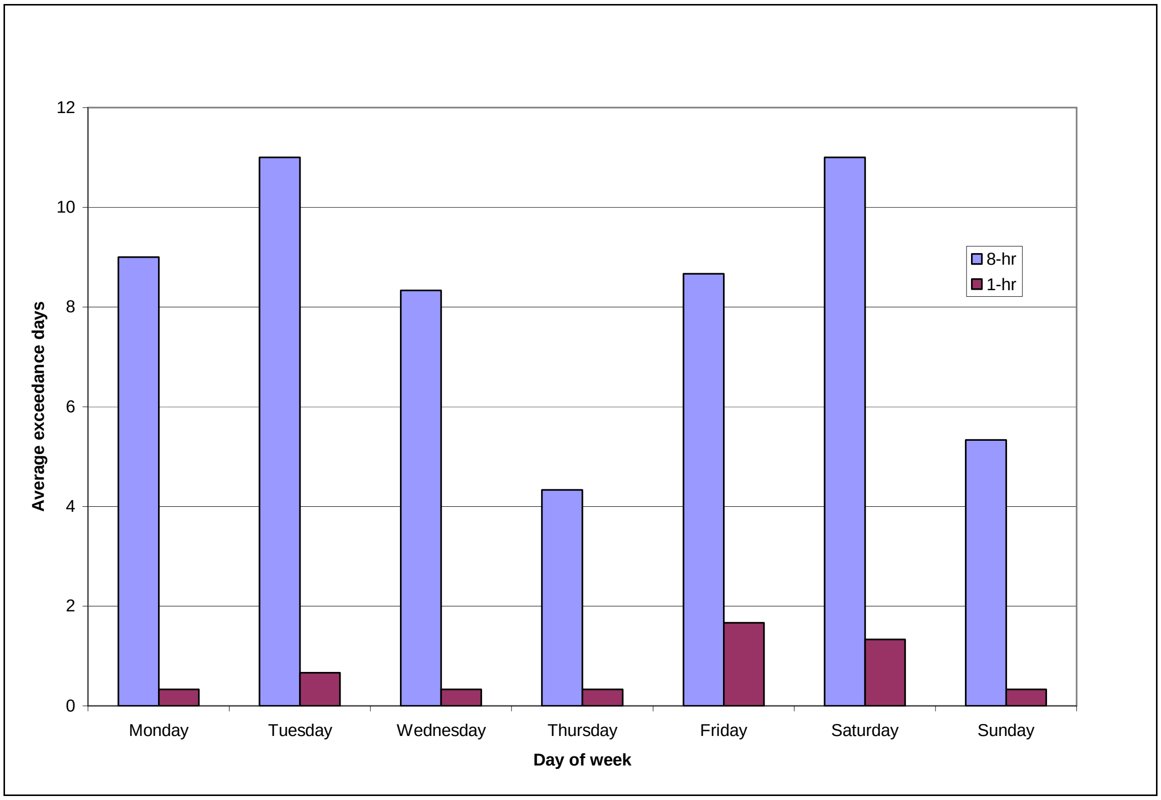


FIGURE 13: Distribution of the average number of 8-hr and 1-hr exceedances by day of week from 1998-2000 for Oklahoma

Refer to Appendix A for the raw data used to generate the graphs in the above section.

4.3 Ozone episodes in Skiatook

Three episodes with high ozone concentration days were chosen for further analysis. The selection was made by identifying days having higher exceedances. Table 3 lists the three episodes chosen for this analysis.

TABLE 3: Selected O_3 Episode Dates

YEAR	STARTING DATE	ENDING DATE
1998	September 1	September 7
1999	August 26	September 1
2000	September 1	September 7

4.3.1 Episode I: September 1-September 7, 1998

Figure 14 displays daily maximum 1-hr and 8-hr ozone concentrations for the week. Five 8-hr ozone exceedances and one 1-hr ozone exceedance were observed during this week. The variation of daily maximum temperature, wind speed, solar radiation, pressure and relative humidity were studied to investigate any noticeable relationship with high ozone concentrations. Figure 15 shows the variation of daily maximum temperature and daily maximum relative humidity for the week while Figure 16 shows the daily maximum wind speed, solar radiation and pressure. A week high 1-hr ozone concentration of 0.129 ppm was observed on the 4th September. Monthly high daily maximum temperature of 105 °F and lowest daily maximum wind speed of 14.7 mph for the week were observed on 4th September, that explain the occurrence of high ozone concentration. Pressure values during the week remained almost constant around 28.8 in. and solar radiation was fairly constant throughout the week with an average of 21.9 MJ/m^2, except for a low value of 18.4 MJ/m^2 on September 1. Daily maximum relative humidity was fairly low with second lowest value of 71%. The daily maximum temperature values during this week were well above the monthly average daily maximum value of 89 °F that help understand high ozone concentrations during the week. Temperature thus is one meteorological parameter that should be considered in explaining high ozone concentrations. It has been shown that O_3 production in rural and urban environments increases with temperature (Vukovich, 1994). Higher temperatures may not be the direct cause of high ozone concentrations, but are surrogate for other meteorological parameters such as atmospheric stability, wind speed, and solar intensity (Olszyna et al., 1997). Refer Appendix A for raw data used to generate the bar graphs in this section.

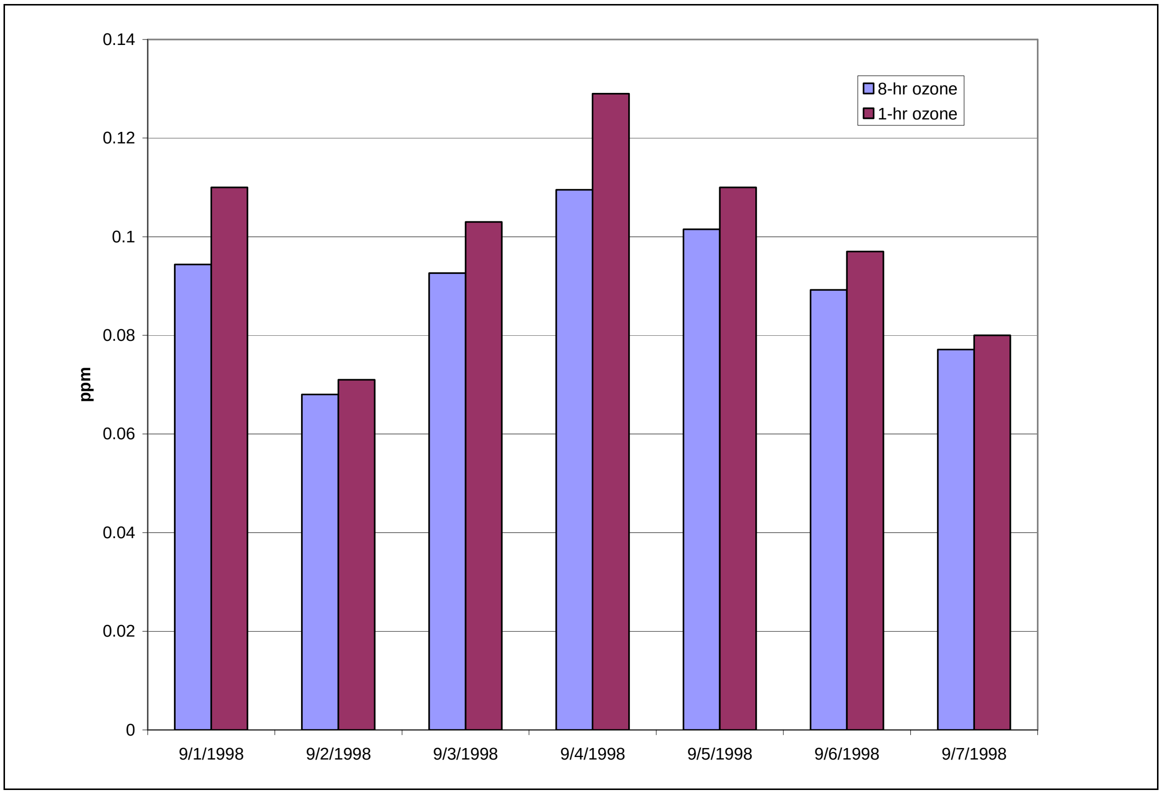


FIGURE 14: Distribution of daily maximum 1-hr and 8-hr ozone from September 1 to 7, 1998 for Skiatook

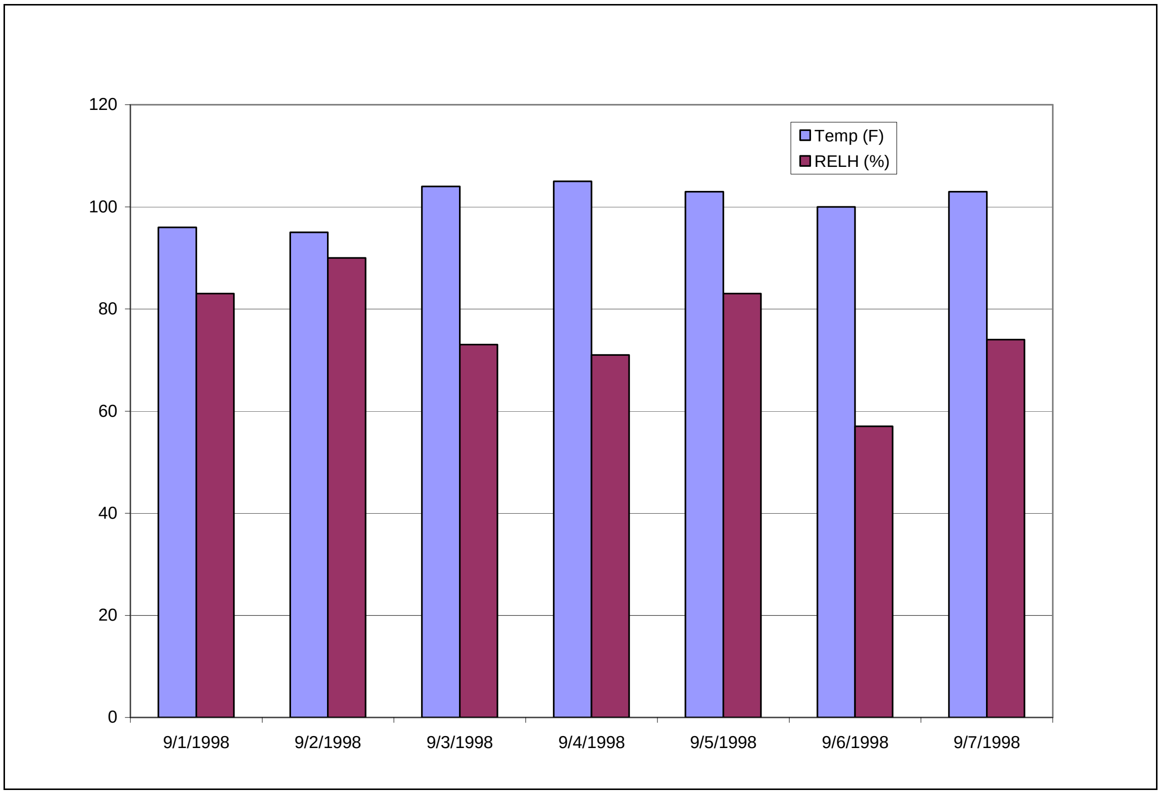


FIGURE 15: Daily maximum temperature and solar radiation from September 1 to 7, 1998 for Skiatook

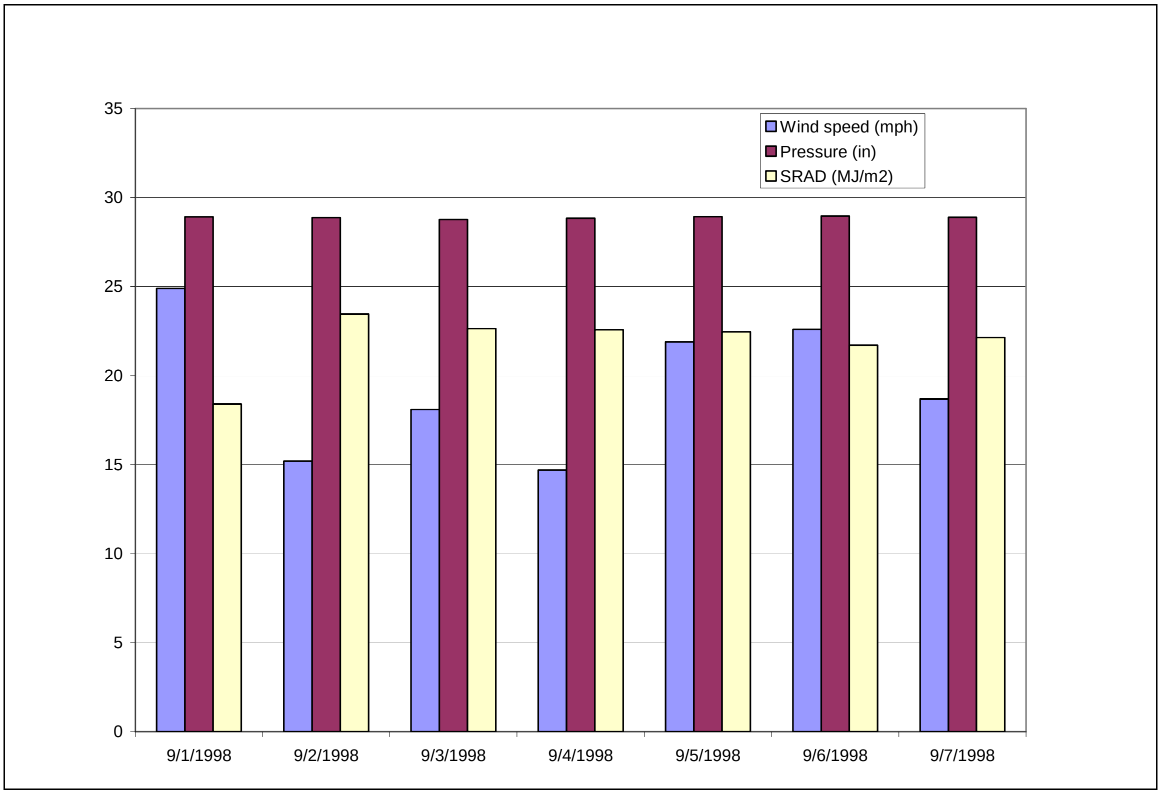


FIGURE 16: Daily maximum wind speed, station pressure and solar radiation from September 1 to 7, 1998 for Skiatook

4.3.2 Episode II: August 26 to September 1, 1999

Four exceedances of 8-hr ozone exceedances were observed during this week. Figure 17 shows the plot of 1-hr and 8-hr ozone values for the week. One hour ozone values for the entire week were below 0.12 ppm. 29th and 30th August recorded highest 8-hr ozone concentration values of 0.094 ppm each. High pressures values of 29.02 in. and 29.03 in. were recorded during 29th and 30th August respectively. A high pressure in the area is considered to help create stagnant atmosphere that helps build higher ozone. The temperatures during these two days were 96 °F and 95 °F, respectively, which were below the weekly average daily maximum of 97.57 °F. However, unlike Episode I, temperature was not found to be related to high ozone concentrations. Figure 18 and 19 show the variation of different meteorological parameters during the week. Temperature, relative humidity, solar radiation and wind speed did not show any noticeable relationship that could explain the high ozone concentration in these areas. The study of this episode thus stresses the need to investigate other possible causes for high ozone episodes. Processes such as regional ozone transport and/or transport of precursors should thus be studied along with the variation of meteorological parameters, to find the true phenomena causing high ozone concentration. Refer to Appendix A to view the raw data used to generate the bar graphs in this section.

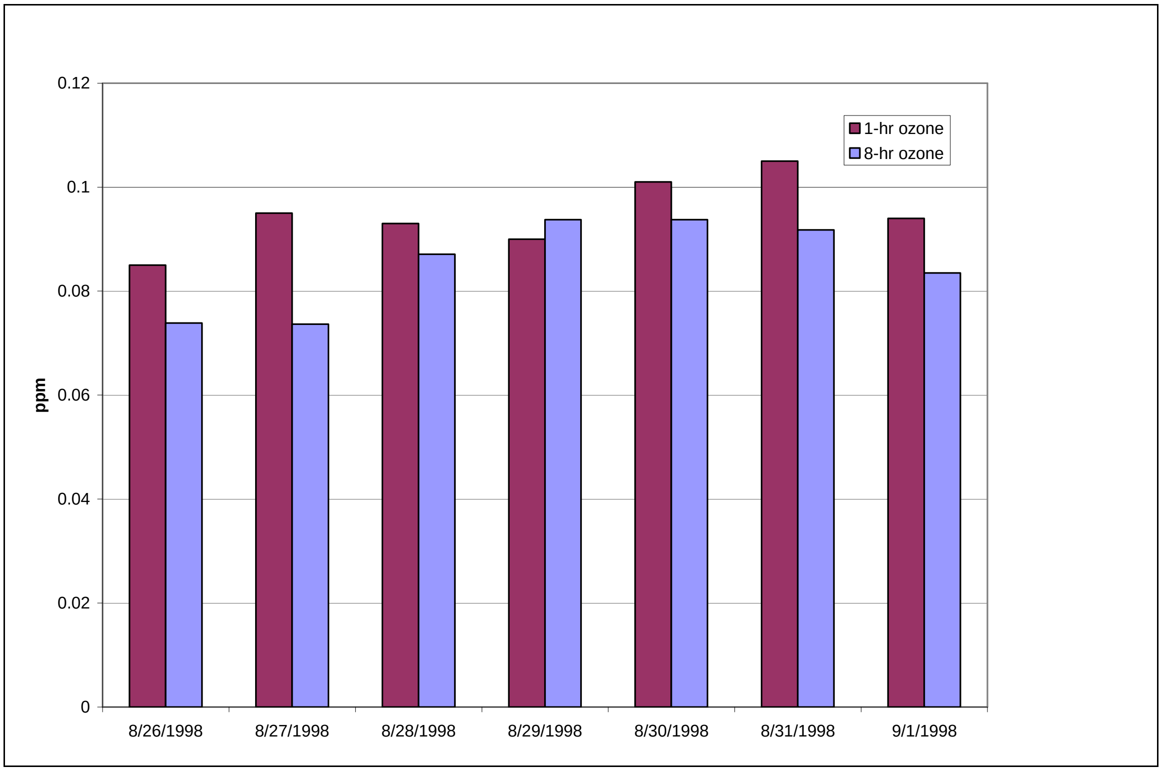


FIGURE 17: Distribution of daily maximum 1-hr and 8-hr ozone from August 26 to September 1, 1999 for Skiatook

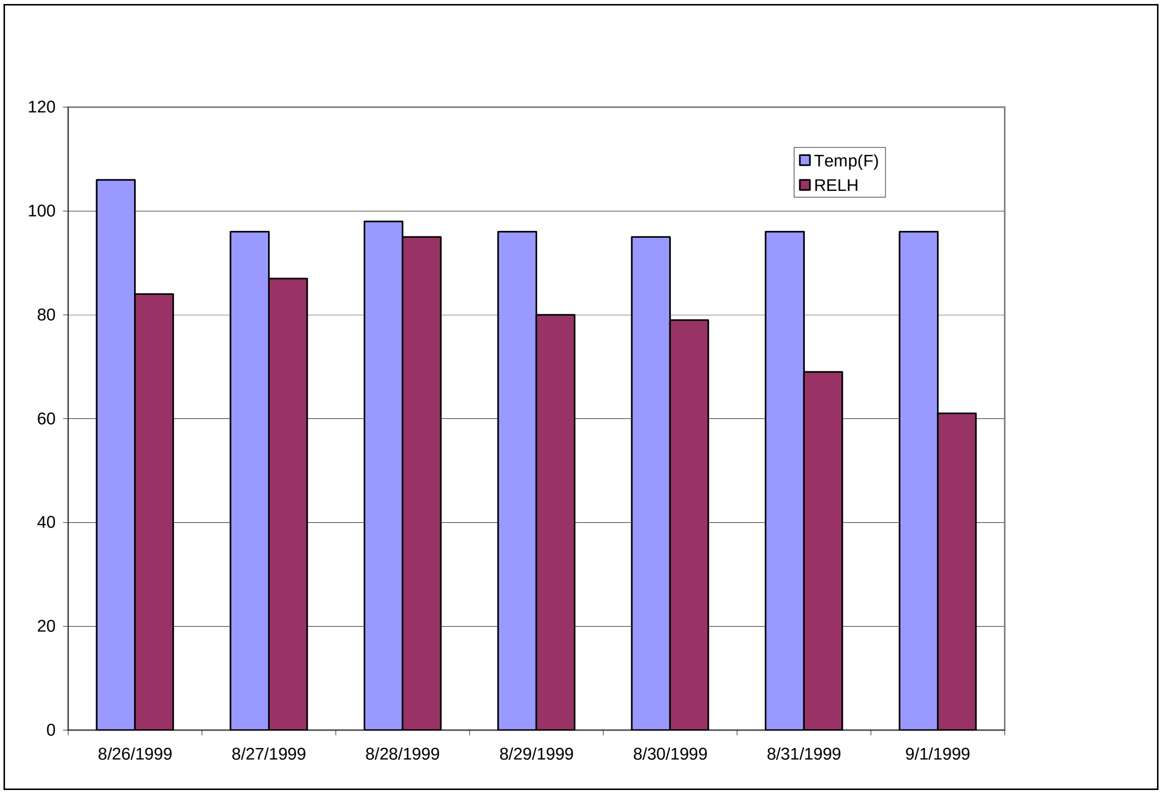


FIGURE 18: Daily maximum temperature and maximum relative humidity from August 26 to September 1, 1999 for Skiatook

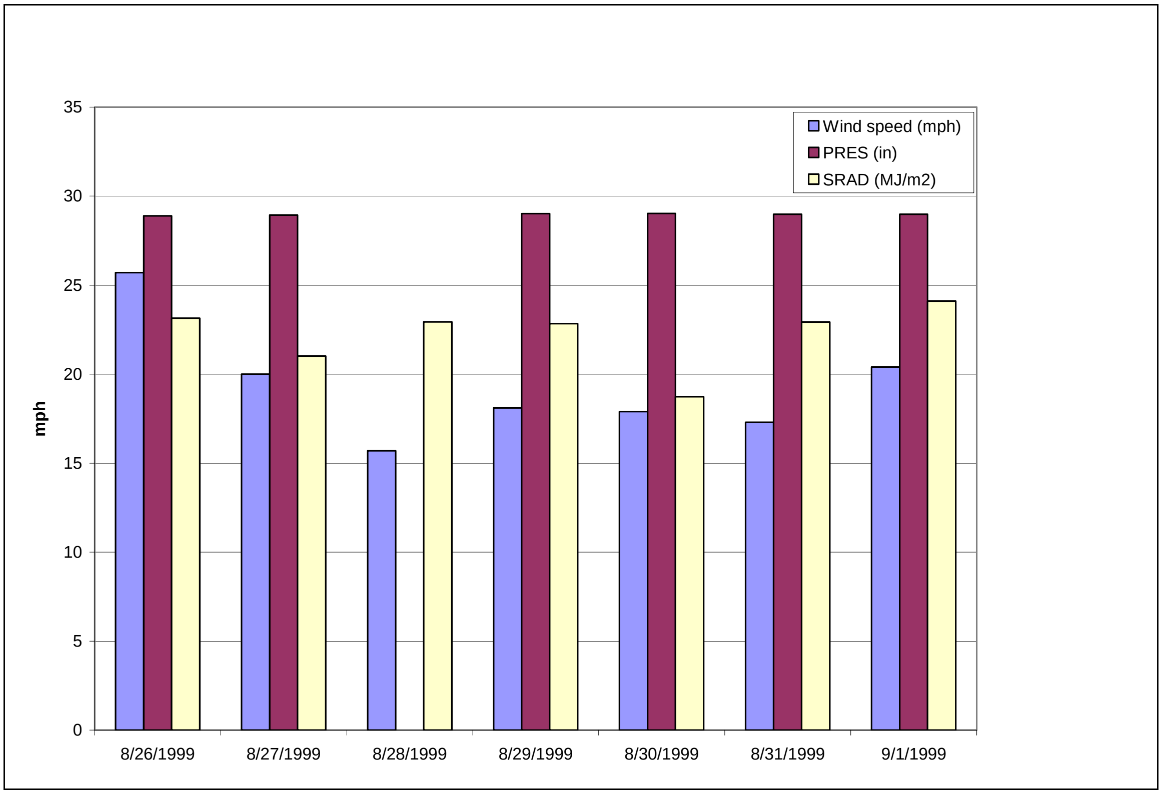


FIGURE 19: Daily maximum wind speed, station pressure and solar radiation from August 26 to September 1, 1999 for Skiatook

4.3.3 Episode III: September 1 to 7, 2000

Figure 20 shows the daily maximum 1-hr and 8-hr ozone concentrations recorded during this week. Two days exceeded the 1-hr ozone concentration value of 0.12 ppm. Monthly high temperature value of 107 °F was recorded on 1st and 2nd September that had 1-hr ozone concentrations of 0.127 ppm and 0.137 ppm respectively. This explanation is in agreement with that provided in Episode I. Daily maximum relative humidity recorded during these days were 52% and 58% respectively that again help in explaining high ozone occurrence. Figure 21 shows the bar graph showing daily maximum temperature and daily maximum relative humidity recorded during this week. Figure 22 shows the daily maximum wind speed, station pressure and solar radiation during each of these days. Solar radiation does not appear to have a noticeable effect on high ozone concentration. Wind speed values on September 1 and 2 were 19.8 and 16.1 mph. These are less than the weekly average wind speed value of 19.9 mph. Generally lower wind speeds results in stagnant atmosphere that are conducive for high ozone concentrations. Refer to Appendix A to review data used to generate bar charts in this section.

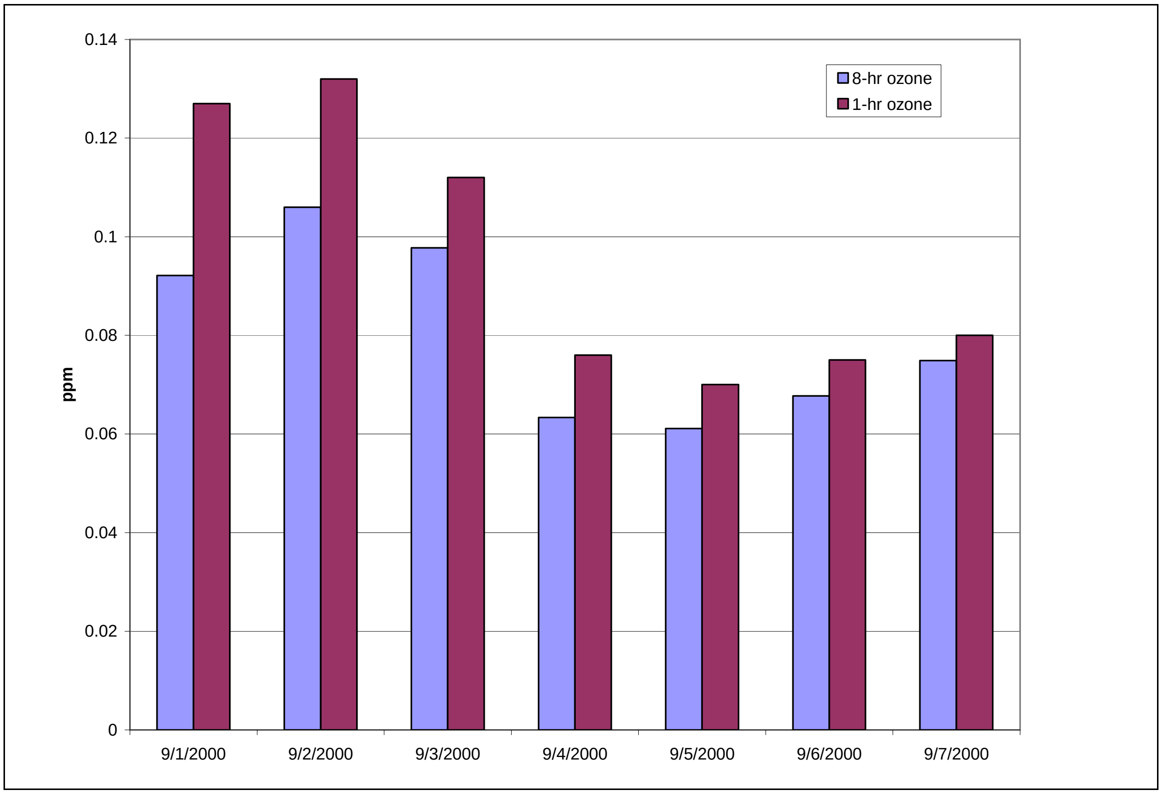


FIGURE 20: Distribution of daily maximum 1-hr and 8-hr ozone from September 1 to 7, 2000 for Skiatook

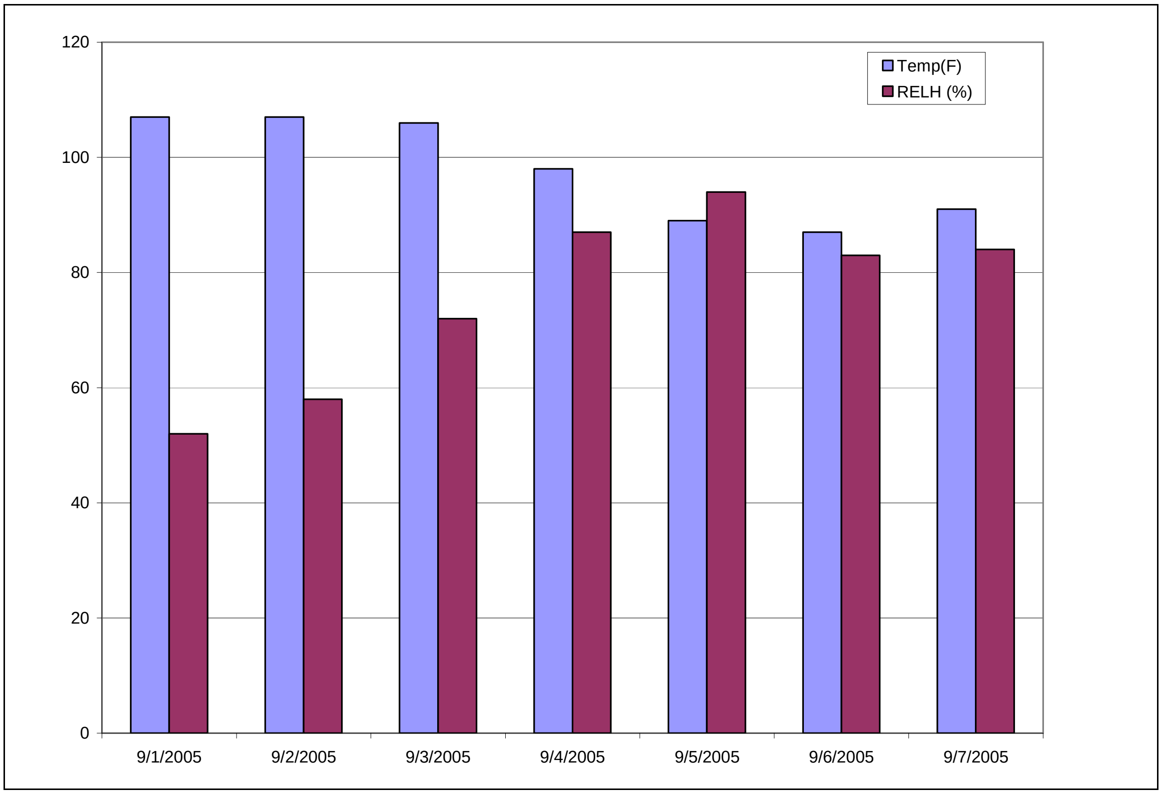


FIGURE 21: Daily maximum temperature and maximum relative humidity from September 1 to 7, 2000 for Skiatook

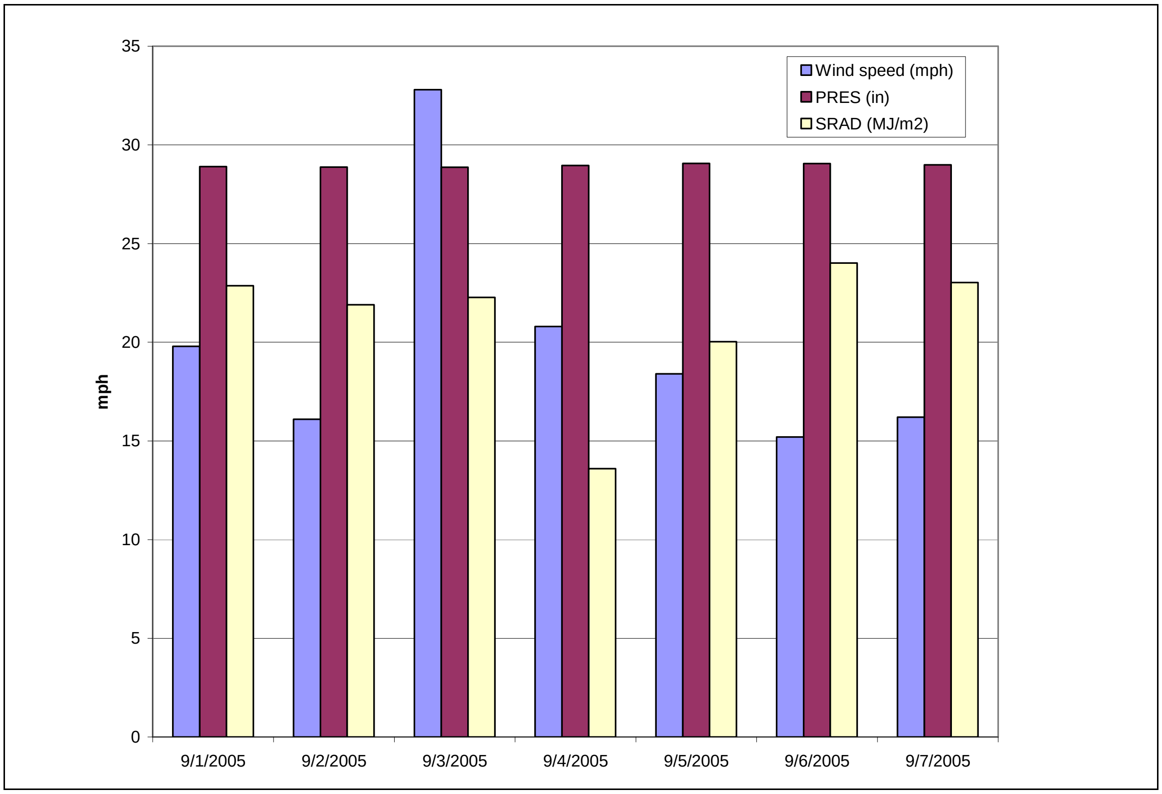


FIGURE 22: Daily maximum wind speed, station pressure and solar radiation from September 1 to 7, 2000 for Skiatook

5. GEOSTATISTICAL ANALYSIS

5.1 Spatial analysis for exceedances of ozone standard

5.1.1 Method

The objective of this section is to produce findings to support those made in the earlier section regarding the high exceedances in Skiatook, and to focus the further study using data from Skiatook. As per the NAAQS for 8-hr ozone, an area is under non-attainment if the 4^{th} maximum 8-hr ozone value for a monitor within an area averaged over a three year period exceeds 0.08 ppm. The objective of this study is to generate contour maps showing probabilities of exceeding the 4^{th} maximum 8-hr ozone values for areas in Oklahoma. These maps will help visually identify areas that have higher probabilities of exceeding the standard. The contour maps were generated using the software GS+ 7.0, that uses the concept of indicator kriging. The threshold value selected here was 0.085 ppm. This is the minimum value required to exceed the standard of 0.08 ppm. The data of the 4^{th} maximum 8-hr ozone values for the years 1998-2000, for the various ozone monitors was taken from Office of Air and Radiation (OAR), EPA. The X and the Y coordinates were taken as the latitudes and longitudes for each monitoring station. The software applies the concept of indicator kriging to assign values of 1’s and 0’s to areas, based on whether the concentrations exceed the threshold value or not.

In the next step, the software assigns probabilities of exceeding the threshold value using ordinary kriging and interpolates these values for the unknown locations in the area under consideration based on values of the known locations. This area is automatically defined by the software based on the number of monitoring stations available. Figure 24 shows the spatial location of various monitoring stations in Oklahoma using the latitudes and longitudes as coordinates.

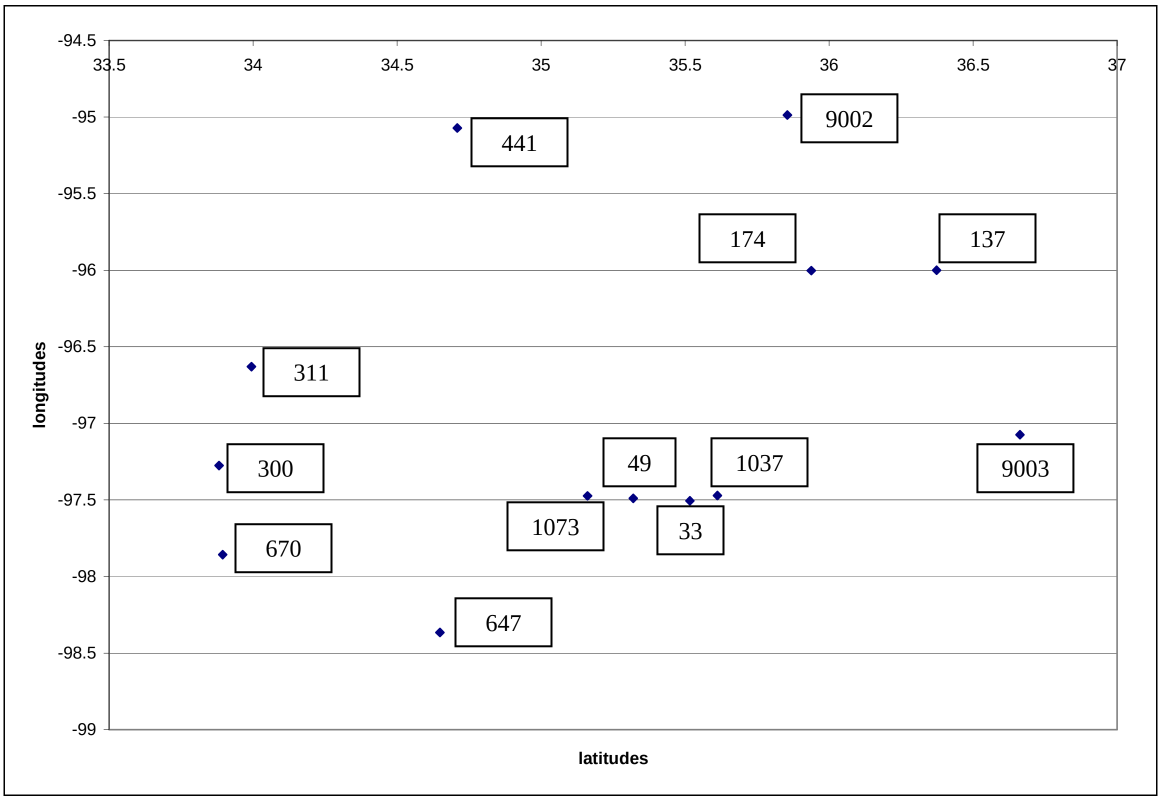


FIGURE 24: Location of monitors based on their latitudes and longitudes

5.1.2 Results and Discussion

The output of the indicator kriging performed for the 4th maximum 8-hr ozone concentration values is in the form of a contour map. Figures 25, 26 and 27 are the contour maps for the indicator kriging performed using the 4th maximum 8-hr average ozone concentrations data for monitoring sites in the years 1998 to 2000 respectively.

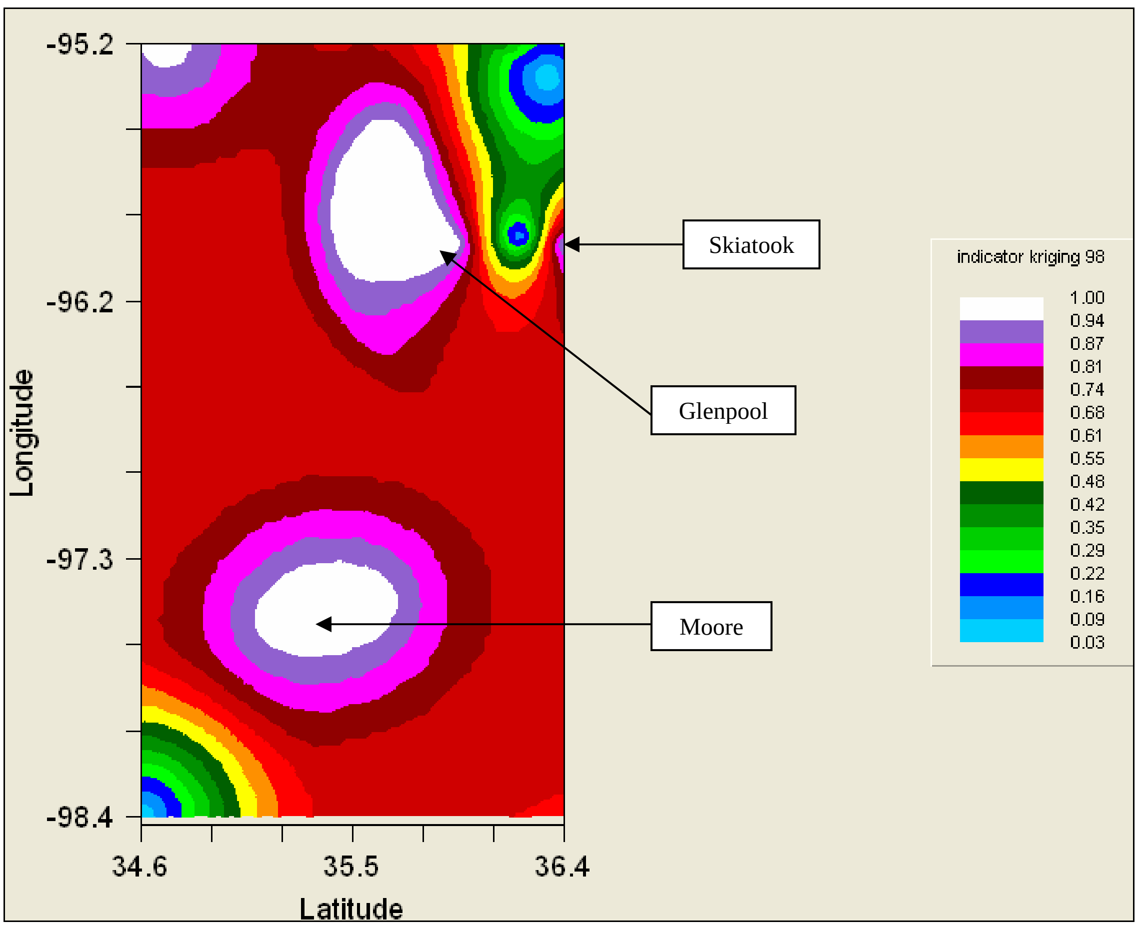


FIGURE 25: Contour map of 4th maximum 8-hr ozone value exceeding 0.08 ppm in Oklahoma, 1998

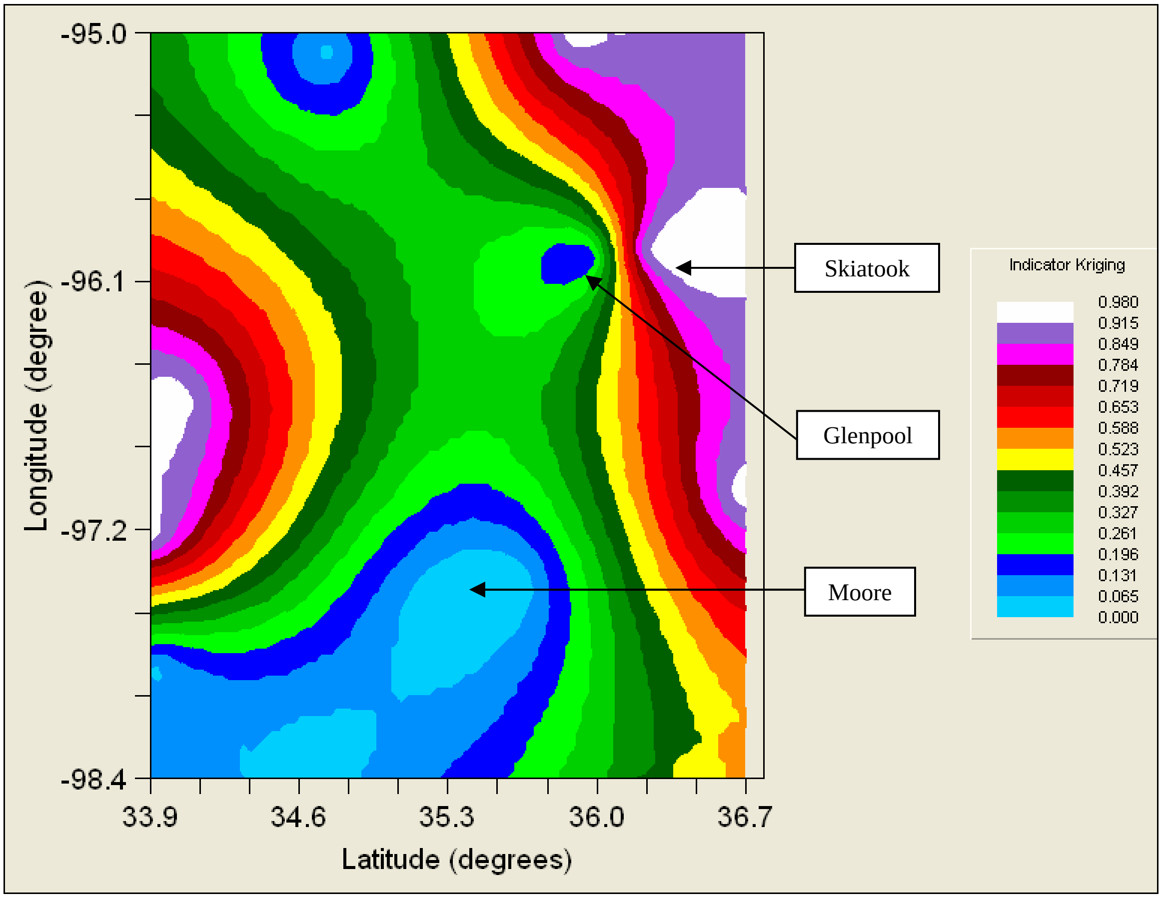


FIGURE 26: Contour map of 4^{th} maximum 8-hr ozone value exceeding 0.08 ppm in Oklahoma, 1999

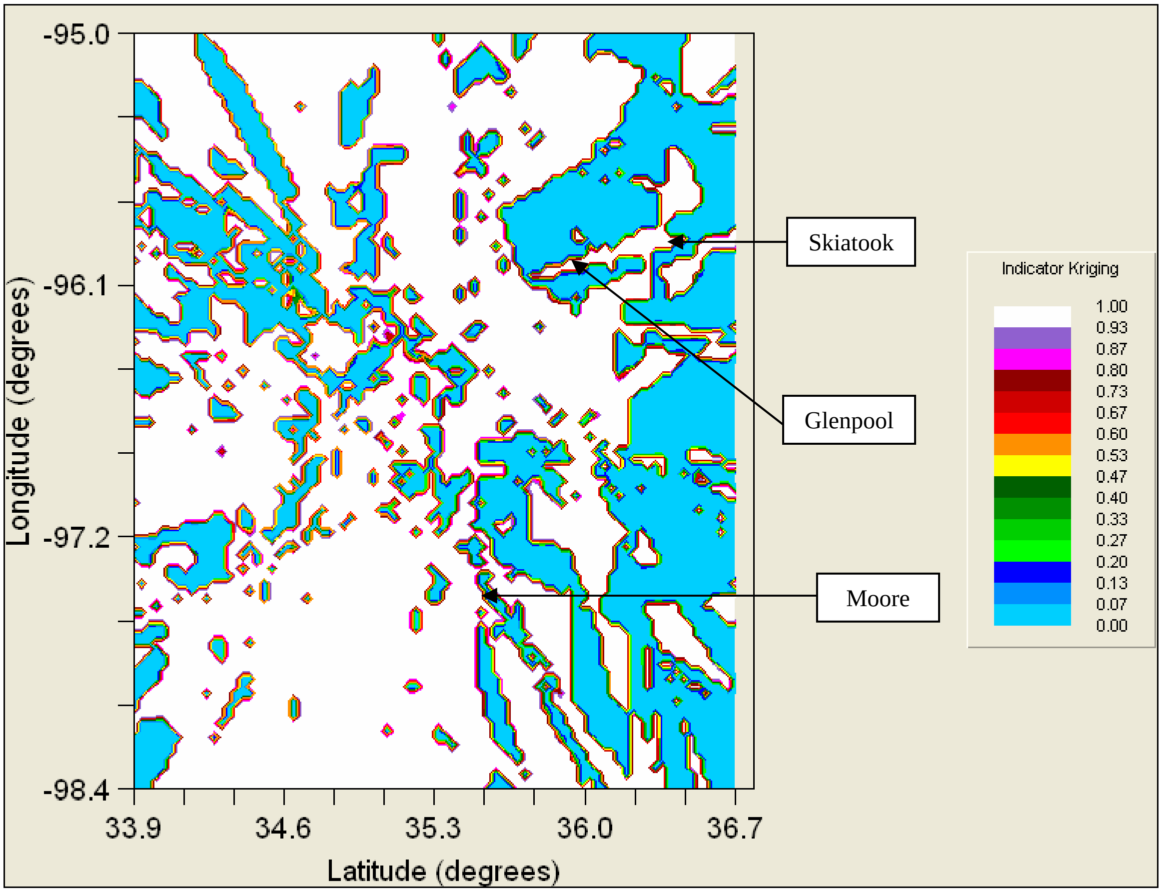


FIGURE 27: Contour map of 4th maximum 8-hr ozone value exceeding 0.08 ppm in Oklahoma, 2000

The approximate locations of ozone monitoring sites in Skiatook, Glenpool and Moore are shown in each of the above shown three figures. As seen in Figure 25, in 1998 Skiatook showed a probability range of 0.87 to 0.94 for its 4th highest 8-hr ozone value to exceed the threshold value of 0.085 ppm. Sites in Glenpool (Tulsa county) and Moore (Cleveland county) were the other areas that showed high probability of exceeding the threshold value. Skiatook showed a probability close to 1 of exceeding the 0.085 ppm threshold value for the remaining years 1999 and 2000. This can be seen in Figures 26 and 26 respectively. Skiatook, thus was a site that had high probability of exceeding the 8-hr ozone standard for all the three years. A similar conclusion was made while observing Figure 8. (pp. 27) that showed that Skiatook had the highest number of 1-hr and 8-hr ozone exceedances averaged over the three years of study. Hence, these conclusions helped to focus the regression analysis described in section 6 for data from Skiatook. The forecasting models developed in section 6, used the data for Skiatook.

Figure 27 that show the contour map for 2000, does not display a smooth transition from low to high probabilities. In 2000 the 4th maximum 8-hr average ozone concentration values had large variations. The values on the lower side of 0.085 ppm were very low and the higher values were well above the 0.085 value. Thus, the software is not able to give an output with smooth transition of probabilities from high to low. Also, as observed earlier in Figure 7 (pp. 26), maximum 8-hr exceedances were observed in the year 2000. The output from this software can be improved if higher data points are available.

5.2 Spatial analysis for diurnal variation of ozone

5.2.1 Method

The second objective of this analysis was to investigate the spatial structure of the 1-hr ozone concentrations during different hours of the day. Ozone data collected at 15 monitoring stations in Oklahoma for the year 2000 were considered for the analysis. One of the monitoring stations in Mohawk, Tulsa (Site ID – 127) was omitted from the analysis since it had no data available for August and September. High ozone concentration values were observed during these two months in 2000, and hence the data for these months were chosen for studying the diurnal patterns in ozone. 24 groups of data, each representing measured values at a given hour of the day for all the 14 monitoring stations were formed. Figure 28 shows a plot of diurnal pattern of station averaged ozone concentration for the month of August, September and of combined August and September. A similar pattern of diurnal variation for the three data sets suggests that similar hour-specific chemical and meteorological processes existed during these individual months and in the combined data. High ozone concentrations were observed during day time hours and low ozone concentrations were seen during evening and night hours. This increase in day time ozone concentrations can be attributed to the photochemical production of ozone. Also in the morning, the nocturnal inversion layer breaks due to aloft air mixing downward and thus is followed by increase in ozone concentration (Casado et al., 1994). At night the ozone gets used in reactions with oxides of nitrogen and also gets deposited on surface of shallow inversion layer (Casado et al. 1994). Once the hourly data sets were formed a semivariogram for each set was generated. The software GS+ was used for generating semivariograms. A variogram is a basic tool in geostatistics for describing spatial correlation. A variogram or semivariogram is a graphical representation of a mathematical function called semivariance versus the distance between data values. A semivariogram is a plot of semivariance versus lag distance. The important features of a semivariogram are the range, sill, nugget, and lag.

Range – The distance at which the semivariogram curve asymptotes and no further increase in variogram value is observed with increase in separation distance between pairs, is called range. Sill is the constant value that the variogram achieves at the range. Nugget is the intercept on the semivariance or Y-axis at zero or very small separation distance. Lag is the average distance for grouping points for variogram calculation (i.e. minimum distance between sample points).

The software automatically selects the model that best fits the data set. Several models such as Linear, Spherical and Gaussian are used for semivariograms. For the sake of consistency in all hours, spherical models were used for all the data sets. This helps in having a nugget, sill and range as three different parameters for each data set.

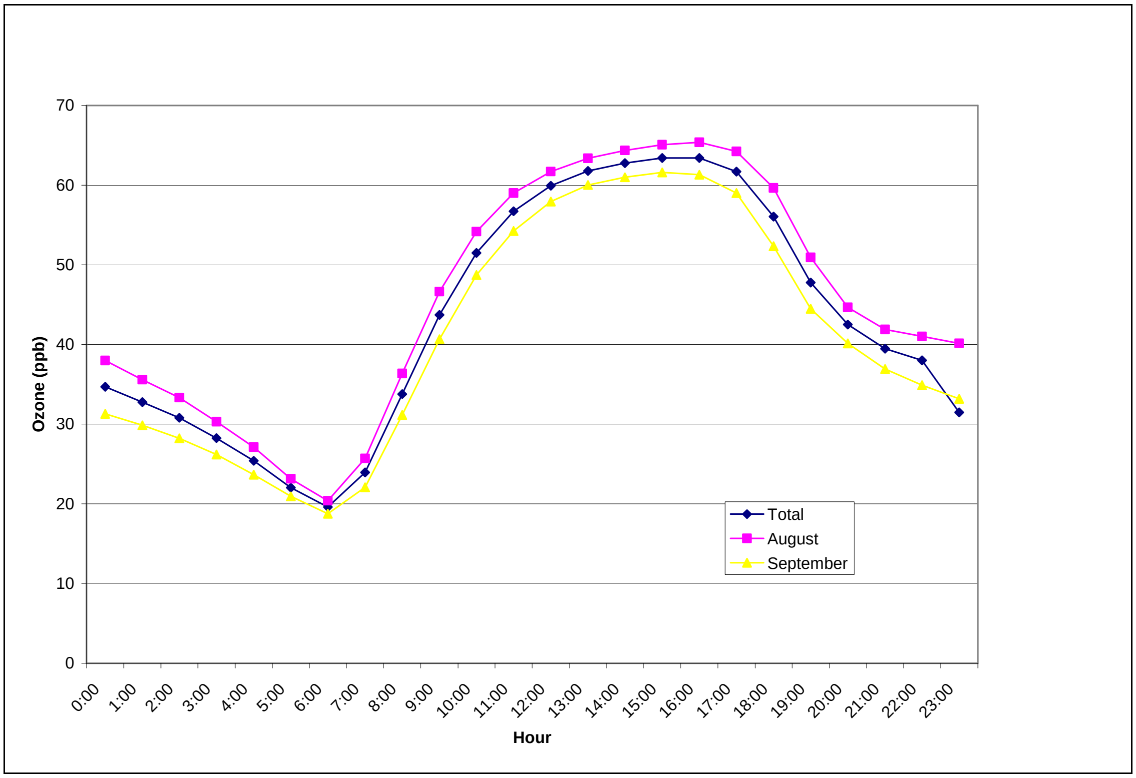


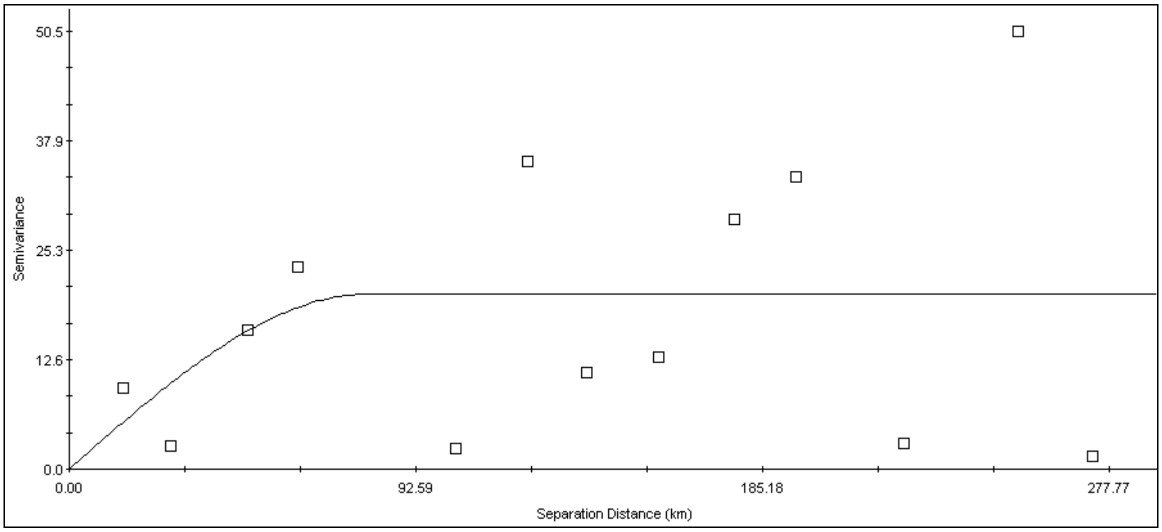
FIGURE 28: The August, September and August-September 2000, station-averaged ozone diurnal patterns

5.2.2 Results and Discussion

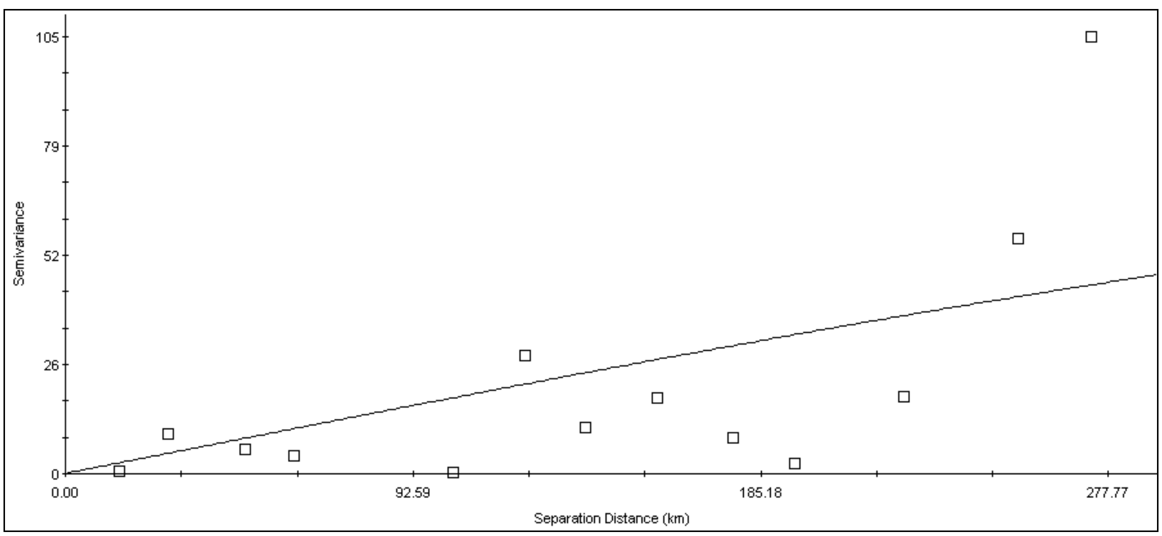
Table 4 lists the fitted parameters of all hour-specific variogram models while Figure 29 illustrates two fitted semivariograms for different hours of the day. Figure 30 displays a plot of variation in range with respect to hour of the day.

TABLE 4: Parameters of the fitted hourly semivariogram models (Spherical-type semivariogram models are used for each data set)

Hour	**Nugget**	**Sill**	**Range (km)**
0	0.1	41	610.9
1	0.01	29.94	610.9
2	0.01	11.33	175.9
3	0.01	11.94	174.1
4	0.72	11.47	146.6
5	0.01	10.53	14.3
6	0.01	10.36	14.3
7	0.58	26.49	610.9
8	0.10	40.33	610.9
9	0.01	31.69	610.9
10	0.01	24.11	610.9
11	0.21	17.48	445.6
12	0.15	17.61	395.5
13	0.53	17.43	335
14	1.61	18.66	273.6
15	4.65	20.19	268
16	4.65	20.19	268
17	0.01	20.19	79.7
18	0.01	26.86	74.4
19	0.1	37.95	68.2
20	0.1	34.71	68.7
21	0.1	89.3	610.9
22	0.01	72.2	610.9
23	1	442.7	610.9



a) Semivariogram for 5:00 p.m.



b) Semivariogram for 10:00 p.m.

FIGURE 29: Fitted semivariograms for different hours of the day for Skiatook, for data of August-September, 2000: Semivariogram for a) 5:00 p.m.; b) 10:00 p.m.

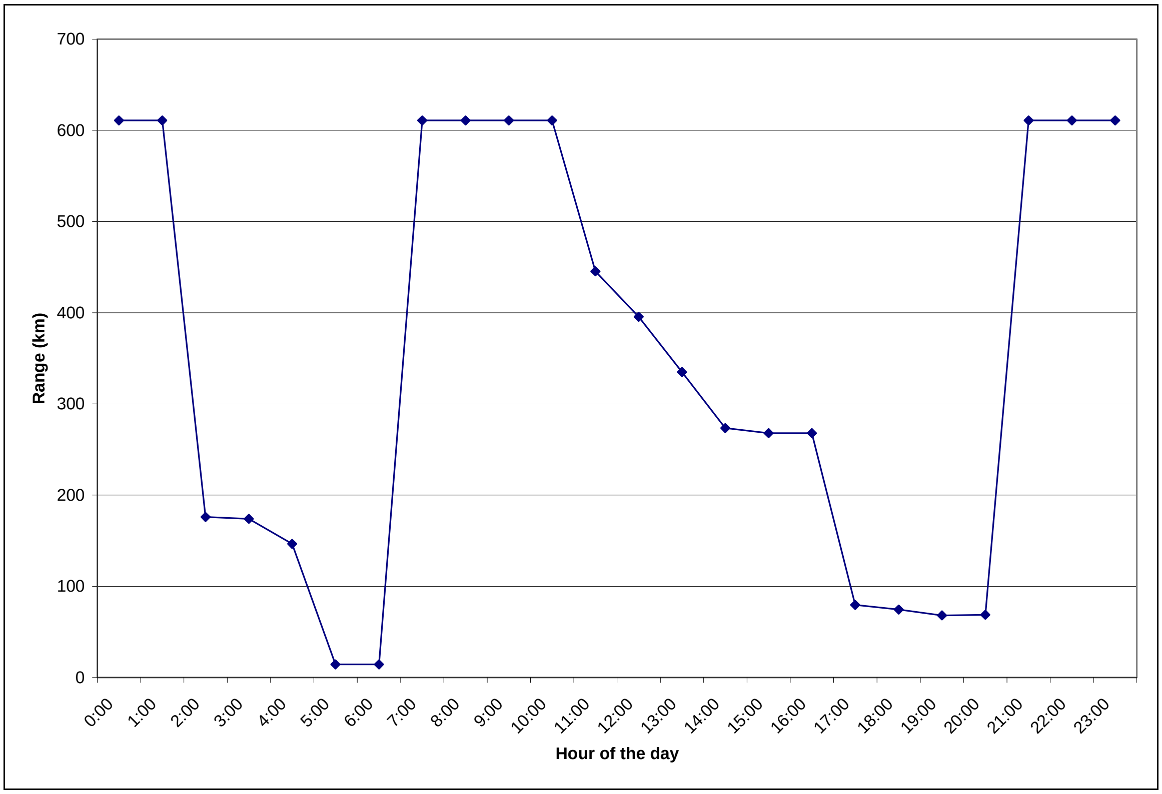


FIGURE 30: The fitted semivariogram range in km vs. the hour of the day

The results observed in Table 4 and Figure 30 exhibit hour-specific spatial characteristics of measured ozone data. Considerable differences were encountered between the spatial structures of daylight and evening hours. The ranges varied from a low of 68.2 km at 7:00 p.m. to a high of 610.9 at midnight and early morning hours. The Day time hours from 11:00 a.m. to 7:00 p.m. show short range effect, suggesting that local factors in the surrounding of the site contribute towards ozone generation. Semivariograms for late evening hours from 9:00 p.m. and onwards show long range effect. It indicates that regional factors dominate in causing low ozone concentrations during night hours. The ranges for early morning hours were very low. No possible reason was found to explain this behavior. Casado et al. (1994) had shown similar results for their study of

Southeastern United States. They found a regular transition of lower to higher ranges from daytime hours to evenings and night hours. The ranges for their work varied from a low of 140 km at 4:00 p.m. to a high of 360 km at 1:00 a.m. The possible reason for unexplainable behavior in Oklahoma is the lower number of monitoring stations. A higher number of monitoring stations would have helped generate better semivariograms. This would have resulted in a better plot of range vs. hour to explain the diurnal variation of ozone.

6. REGRESSION ANALYSIS

6.1 Method

The study of spatial distribution of high ozone for the 1998 – 2000 data shows the importance of Skiatook, a suburban site in Tulsa, with regards to the number of 1-hr and 8-hr ozone exceedances. The regression model was thus specifically developed for Skiatook. Three types of response were desired from the regression models to be developed. These response variables are the next day

- 1-hr ozone concentration
- 8-hr ozone concentration
- Daily maximum 1-hr ozone concentration (DMOC)

The data from 1998 to 2000 were used for developing the model. Data from 10:00 a.m. to 5:00 p.m. for the months of May to September were used considering the conditions favorable to high ozone production. Based on review of literature and with the knowledge of parameters involved in ozone formation, nitrogen dioxide, temperature, wind speed, solar radiation, pressure, and relative humidity were selected as predictor variables for the analysis. Also high ozone episodes analyzed earlier have shown at some point of time, there was certain amount of variation relative to variation in the above mentioned parameters. The degree of variation was however, not determined. Also previous day ozone concentrations were used as predictors.

Skiatook has an ozone monitor and a Mesonet station, but does not have a NO_x or NO_2 monitor. Glenpool is the only site close to Skiatook in Tulsa county with a NO_2 monitor, and it was assumed that similar concentrations of NO_2 may be found in Skiatook. Hence, the NO_2 data from Glenpool were used for the entire analysis for Skiatook. The NO_x data set has a large number of missing values as compared to NO_2 data. So, it was preferred to use NO_2 as one of the predictors instead of NO_x. There were several missing values in the NO_2 data as well, and since actual NO_2 concentration data at Skiatook were not available, the predicting ability of the model was expected to be affected. However, it is important to include NO_2 data in the analysis since it is essential for ozone formation.

A stepwise regression procedure was applied at a 95% confidence level using the software MINITAB 14. This procedure runs several trials with different combinations of predictor variables. The output shows the regression coefficients and R-squared (R^2) value for each set of predictor variables used. The user thus has the option to select the set of predictor variables desired as per the site specific conditions. The step wise regression procedure performs various runs with the aim to maximize the R^2 value. The procedure applied by the statistical software used also maintains the assumption of multiple linear regression of having independent predictor variables. This is done by checking correlation between predictor variables. In the event of having highly correlated variables, the variable with lower effect on the response variation is eliminated. Table 5 lists the response variables and the number and type of predictor variables used for model runs. Table 6 shows the mean and standard deviation of the parameters used as predictors for regression.

TABLE 5: Desired response and predictor variables used for regression models

Model Number	Response Variable	Number of Predictor Variables	Predictor Variables
1	1-hr ozone	6	Temperature (TAIR), Relative humidity(RELH), Wind speed (WSPD), Solar radiation (SRAD), Pressure (PRES), and Nitrogen dioxide (NO_2)
2	1-hr ozone	7	Previous day 1-hr ozone (PREV1 O_3), TAIR, RELH, WSPD, SRAD, and NO_2
3	8-hr ozone	6	TAIR, RELH, WSPD, SRAD, PRES, NO_2
4	8-hr ozone	7	Previous day 8-hr ozone (PREV8 O_3), TAIR, RELH, WSPD, PRES, NO_2
5	Daily maximum 1-hr ozone	7	Daily Maximum 1-hr ozone concentration (DMOC), Maximum temperature (TMAX), Average temperature (TAVE), Average relative humidity (RELHave), PRES, SRAD, Average wind speed (WSPDAve)

TABLE 6: Statistical properties for predictor variables used in the regression analyses

Predictor variables	Mean	Standard Deviation	Number of observations N
NO_2	0.004844	0.004307	3362
Relative Humidity	76.87	15.73	3537
Temperature	22.76	5.112	3537
Wind speed	3.123	1.481	3537
Pressure	981.5	4.493	3537
Solar radiation	283.1	302.0	3537
Previous day 1-hr ozone	0.05669	0.01750	3536
Previous day 8-hr ozone	0.05229	0.01517	3551

6.2 Results and Discussion

As described earlier, the software automatically checks the correlation between predictor variables to satisfy the assumption of having independent predictors for multiple linear regressions. Table 7 shows the Pearson correlation coefficients between predictor variables.

TABLE 7: Correlation coefficients of predictor variables for 10:00 am to 5:00 pm data, May to September of 1998-2000, Skiatook

	RELH	**TAIR**	**WSPD**	**PRES**	**SRAD**
TAIR	-0.484				
WSPD	-0.230	0.223			
PRES	-0.106	-0.100	-0.295		
SRAD	-0.592	0.556	0.237	0.120	
NO_2	0.079	-0.057	-0.135	0.059	-0.058

Comparing the correlations among the various predictor variables, temperature and solar radiation were found to be well correlated with a Pearson coefficient of 0.556. However, the correlation coefficient values between any two variables used in the analysis were not high enough to be considered as highly correlated variables, and hence were not discarded from use in the regression analysis. Relative humidity is negatively correlated to solar radiation and temperature. The negative sign indicates that with an increase in relative humidity a decrease in the both temperature and solar radiation will be observed. Regression was thus performed for the three responses desired using the five models with different sets of predictor variables given in Table 5. Table 8 gives the regression equation obtained and the R^2 values for each model. Refer to Appendix B for stepwise output from software Minitab 14 obtained for each of the Model.

TABLE 8: Regression equations for 1-hr and 8-hr ozone forecasting models developed

Model number	**Model Equation**	**R^2 value %**
1	-0.12392 + 0.00097 TAIR - 0.00382 WSPD - 0.00041 RELH - 0.00001 SRAD - 0.250 NO_2 + 0.00021 PRES	29.86
2	-0.15232 + 0.00043 TAIR + 0.410 PREV1 O_3 - 0.00307 WSPD -0.00025 RELH - 0.323 NO_2 + 0.00021 PRES	42.28
3	-0.03651 + 0.00091 TAIR - 0.00033 RELH - 0.00304 WSPD - 0.224 NO_2 + 0.00011 PRES	33.65
4	-0.09247 + 0.00037 TAIR + 0.481 PREV8 O_3 - 0.00229 WSPD - 0.0002 RELH - 0.281 NO_2 + 0.00014 PRES + 0.000001 SRAD	50.82
5	0.62843 + 0.00078 TMAX –0.00266 WSPD – 0.00036 RELH – 0.0203 PRES	45.68

Model 1 has a final R^2 value of 29.86. The use of temperature, wind speed and relative humidity resulted in a R^2 value of 28.67. Subsequent addition of the remaining parameters caused the remaining increase in the R^2 value. Refer to appendix B for the actual output obtained using the software Minitab. This helps understand that temperature, wind speed and relative humidity are important meteorological parameters in ozone formation for Skiatook. A similar observation was made in the analysis of high ozone episodes that showed variation of ozone was noticeably related mostly to variations in temperature and on a smaller extent to variations in relative humidity and wind speed.

Model 2 uses previous day 1-hr ozone concentration in addition to the meteorological parameters as predictors. A higher R^2 value of 42.28 was achieved in this model compared to the previous one. The previous day ozone concentration is thus, considered to have a considerable effect on next day ozone forecasting. It is very much possible that the portions of previous day ozone gets carried over to the next day and adds up to the photochemical ozone formed in the existing day.

Models 3 and 4 were developed for forecasting next day 8-hr ozone. An increase of 17.17 in the R^2 value was noted with the addition of previous day 8-hr ozone concentration as predictor. Observations were similar to 1-hr models with the initial steps of regression using temperature, wind speed and relative humidity showing an R^2 value of 33.17, followed by a final value of 33.65 using all the predictors for model 3. Feister and Balzer (1991) have also studied the dependence of ozone concentrations on meteorology, and found that previous day's ozone concentration was the most important variable in their short term forecasting.

Model 5 uses maximum temperature in the first step of regression that gives a R^2 value of 34.38. Subsequent addition of remaining predictors used in Model 5 shows a final R^2 value of 46.58. Thus, temperature can be considered to have a greater effect on ozone concentration as compared to other meteorological parameters. This conclusion was also supported by observations made in earlier analysis of high ozone episodes wherein peaks in ozone concentrations were observed on days that recorded highest daily maximum temperatures. The model 5 eliminates solar radiation from final output. The possible reason for this is its high correlation with temperature and lesser effect on ozone concentrations. This was also observed in the analysis of high ozone episodes, wherein no noticeable effect of variation in solar radiation was observed on ozone concentrations.

Davis and Speckman (1999) developed a model to predict mamimum and 8-hr average ozone in Houston. They used opaque cloud cover (averaged over the period of 10 a.m. to 5 p.m.), yesterday's maximum ozone, today's maximum temperature and morning mixing depth as predictor variables for their model. R^2 values for 8-hr average forecasts ranged from 0.66 to 0.73 while R^2 values for maximum ozone ranged from 0.61 to 0.68. Narasimhan et al. (2000) used neural networks to develop ozone forecasting models. They obtained a correlation coefficient of 0.77 using eight surface meteorological input variables and NO_2, a coefficient of 0.82 by incorporating ozone concentrations of previous 3 days and finally a correlation coefficient of 0.88 by incorporation of upper-air data. Their work takes into account the non-linearity in ozone formation as well.

6.3 Validation

The model validation was performed using data for August, 2001. Data for 1-hr ozone, 8-hr ozone, NO_2, and other meteorological parameters were retrieved from their respective sources. The validation was performed by calculating response variables based on the regression equations obtained. It is not recommended to judge the quality of a model simply based on R^2 value. A basic analysis of residuals is important to assess the predicting ability of a model. Residuals are obtained by subtracting the predicted values from the actual values of the response variable (Helsel and Hirsch, 1995). In order to assess the quality of a model, residuals plots are generated for each model. A residuals plot is a scatter plot with residuals on the Y axis and the predicted values on the X axis. Two elements of the plot help in deciding the predicting ability of the model. One is to check for presence of any curvature in the data. Curvature implies the model has biased residuals. In other words, residuals are either mostly positive or mostly negative. This can be seen as values predicted either mostly higher than actual values or lower than actual values. Two, a model should be checked for heteroscedasticity. Heteroscedasticity can be defined as non-constant variance (Helsel and Hirsch, 1995). Lower heteroscedasticity implies relatively constant variance of predicted values. In summary, a model is considered to have good performance if the residuals plot shows low variance and low curvature. Figures 31 – 35 show the residuals plot for the five models developed.

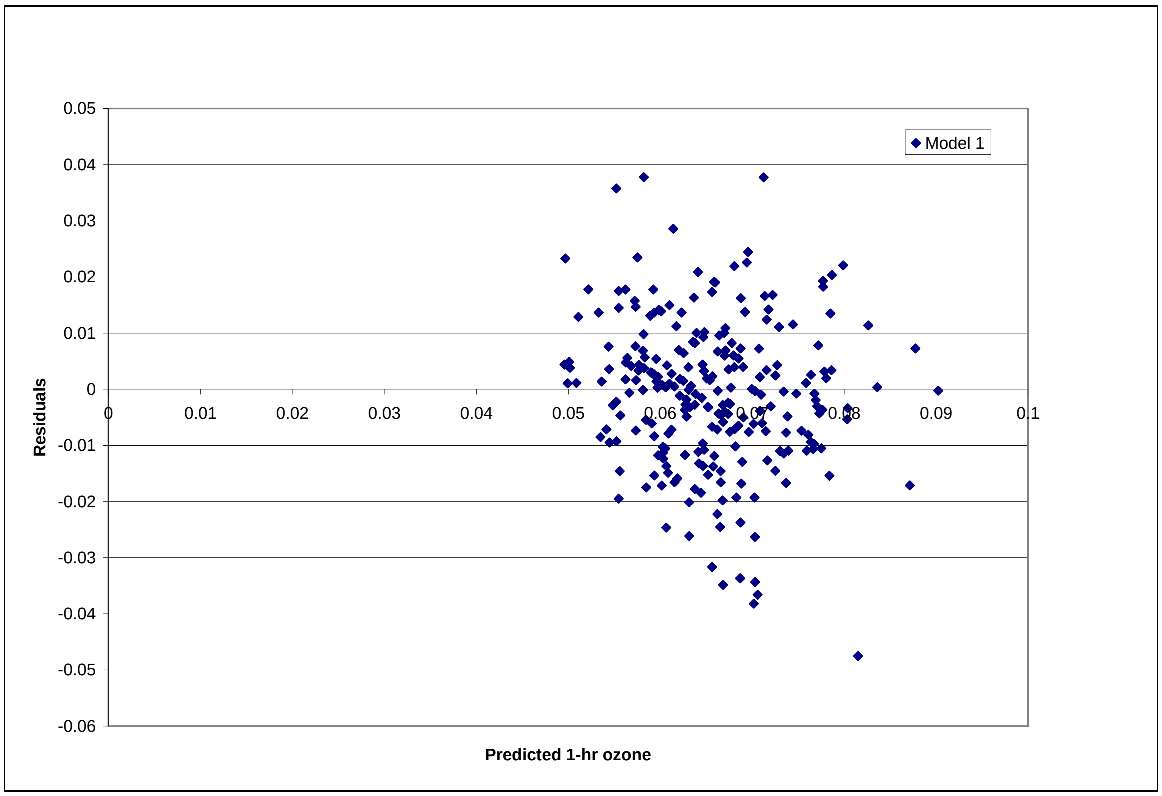


FIGURE 31: Residuals plot for Model 1 (Response is 1-hr ozone)

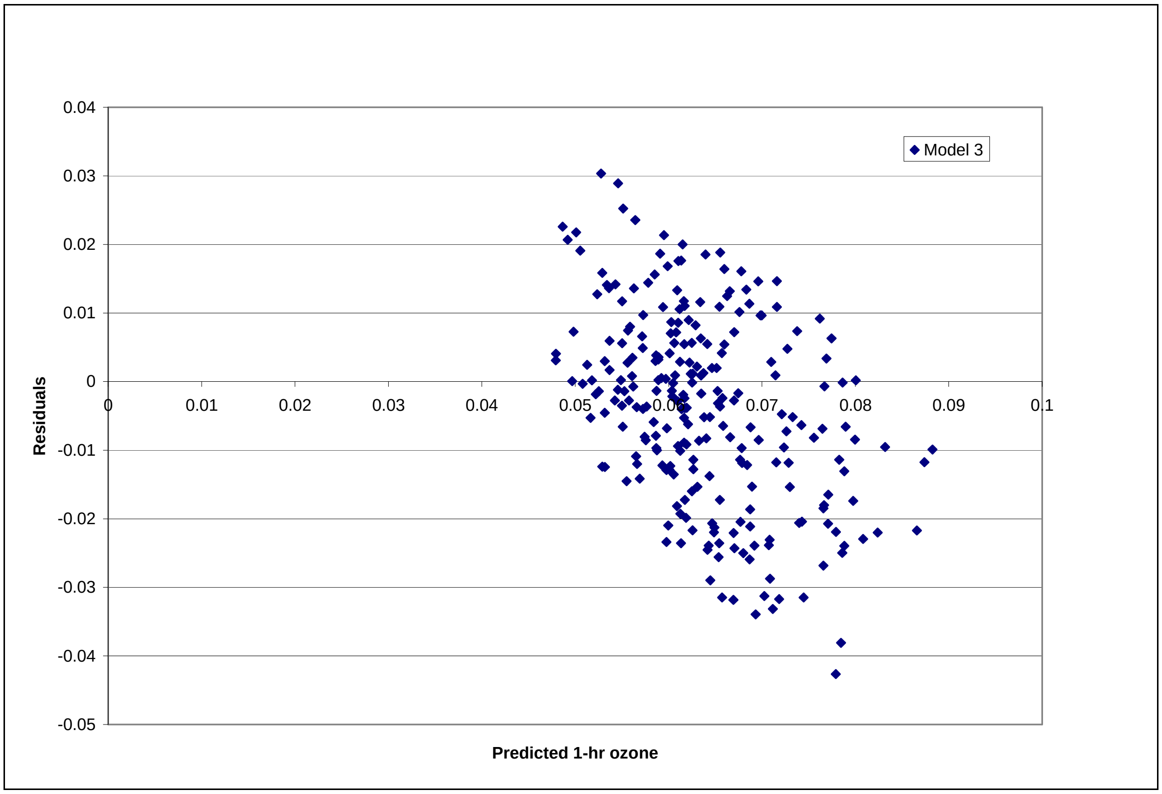


FIGURE 32: Residuals plot for Model 2 (Response is 1-hr ozone)

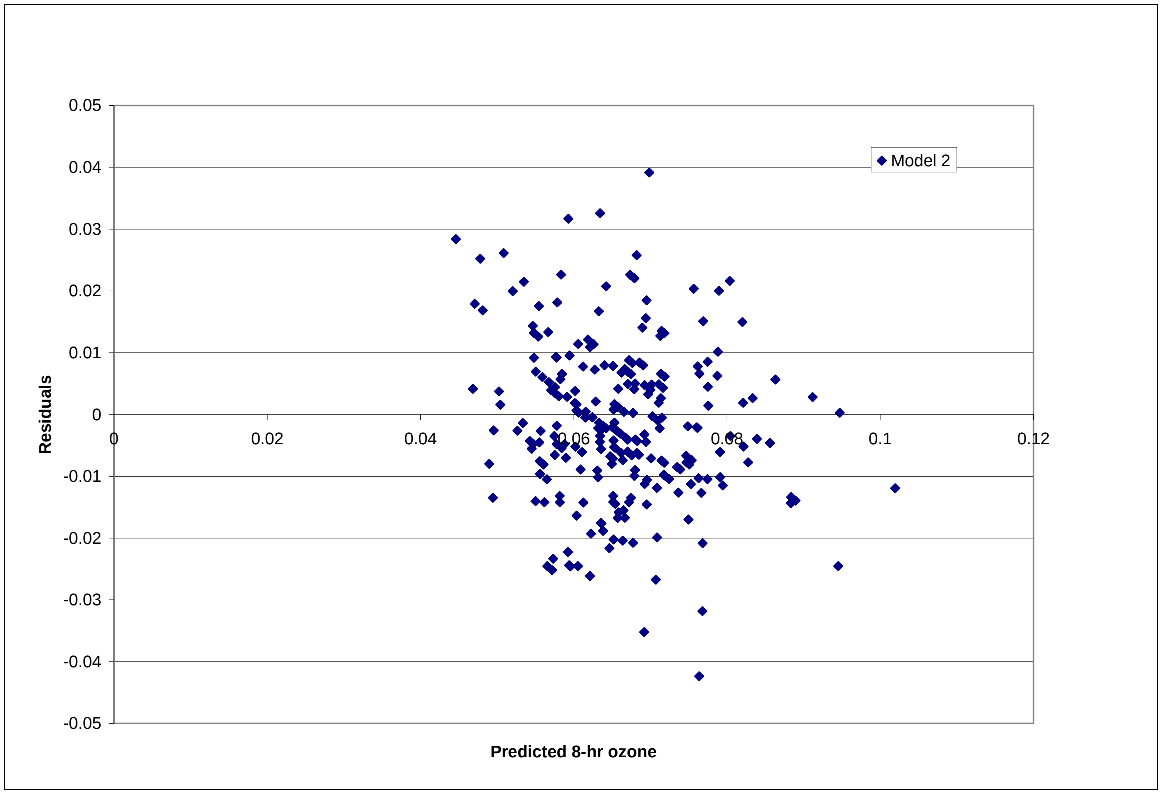


FIGURE 33: Residuals plot for Model 3 (Response is 8-hr ozone)

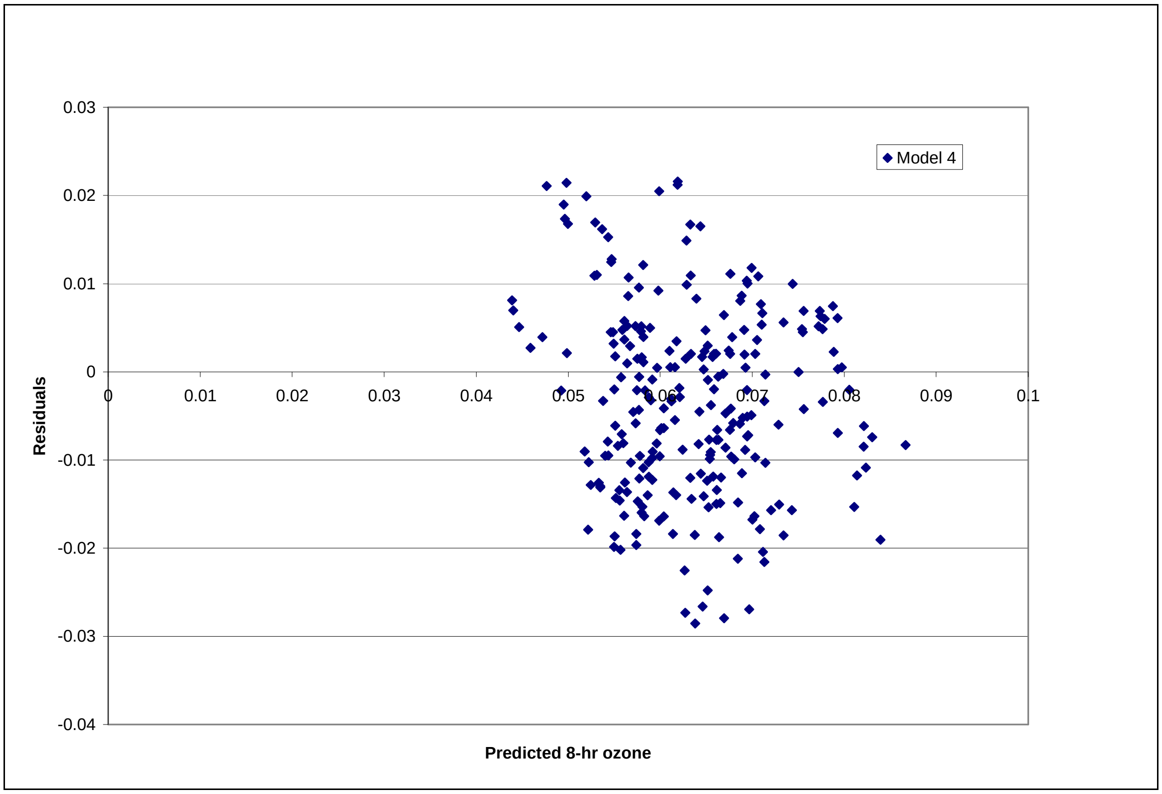


FIGURE 34: Residuals plot for Model 4 (Response is 8-hr ozone)

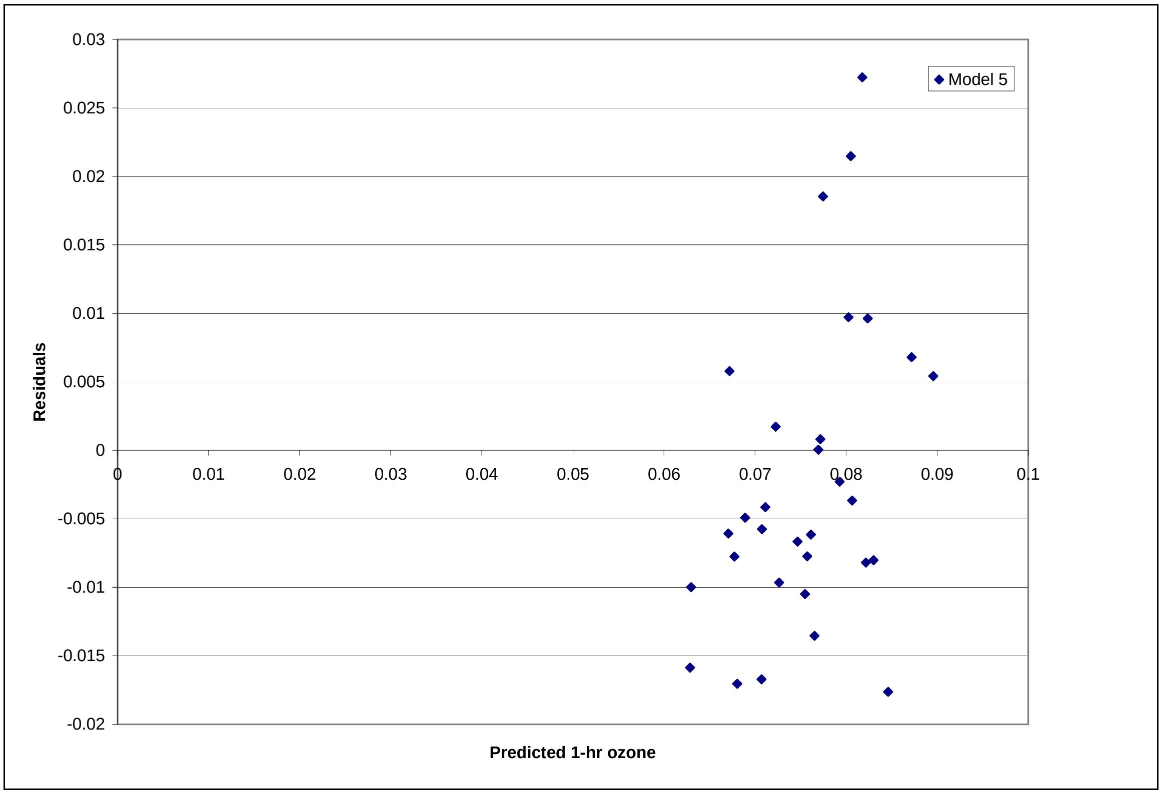


FIGURE 35: Residuals plot for Model 5 (Response is 1-hr ozone)

A common observation from the residuals plot of each model is high variance, suggesting that the predicting ability of the models developed here is not very good. A criterion was established to check how well the model forecasts the ozone values. Predicted values were checked if they were within ± 0.005 ppm of the actual values. A percentage of the total validation data that satisfies the above criterion was calculated for each model. Table 9 gives these percentages for each model calculated by applying the above criterion to the validation data of August 2001. The total number of data points for the validation data were 248. Model 1 shows the highest percentage of data that satisfies the above mentioned criterion. Model 4 has a high R^2 value, but it forecasts ozone concentrations generally higher than actual values.

Residuals plot of Model 1, Figure 31, shows the least curvature and lower variance with respect to the residual plots of other models. Figures 32 and 33 that are the residuals plot of Model 3 and 4 respectively showed higher negative residuals suggesting that these models forecast higher than actual ozone concentration values. Model 5 was validated for data of August, 2001. The number of data points for validation of Model 5 was 31. As observed in Figure 35, Model 5 has high variance that implies that the forecasting ability of the model is poor. An important conclusion here is that since there is no specific value of R^2 that is considered good or bad, other statistical analysis such as the residual analysis performed here should be done to check the predicting ability of the model. The requirement of the user and availability of the data are factors that would help decide what model is best suited. Since, there is no NO_2 data available for Skiatook, using data from Glenpool has possibly affected the quality of the model. Such a conclusion can be made since a negative regression coefficient of NO_2 was observed in all the models. Day time NO_2 is considered to be useful in ozone formation and hence, ideally a positive regression coefficient should have been observed in the regression equation.

TABLE 9: Forecasting ability of models developed

Model number	**Response variable**	**Percentage of forecasted ozone values within ± 5ppb**	**R^2 value %**
1	1-hr ozone	35.88 %	29.86
2	1-hr ozone	30.6%	42.28
3	8-hr ozone	25.8%	33.65
4	8-hr ozone	28.6%	50.82
5	Daily maximum 1-hr ozone	19.35%	45.68

7. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

This research was focused on conducting statistical analysis of ozone data from 1998 to 2000 for Oklahoma, to identify areas that showed high ozone concentrations and to investigate the effect of meteorological parameters on high ozone concentrations.

7.1 Summary and Conclusions

1) Temporal distribution of ozone:

The temporal analysis conducted using the data from 1998 to 2000 showed that the number of exceedances in 1999 were lesser compared to 1998 and 2000. In 2000, the areas having high ozone concentrations were spread throughout the state. The site in Skiatook showed higher exceedances for the period of study. Temporal analysis of Skiatook showed occurrence of 1-hr and 8-hr ozone exceedances to be concentrated in months of July to September. Also, the exceedances entirely occurred during 10:00 a.m. to 5:00 p.m. hours of the day. These daytime hours are ozone conducive and were confirmed by results shown in Skiatook. The exceedances did not follow a specific weekend-weekday pattern in either Oklahoma or in Skiatook.

Temperature was the only meteorological parameter that showed noticeable effect on high ozone concentrations as observed in two of the three episodes of high ozone concentrations analyzed.

Also the stepwise output for the models developed showed that R^2 values increased significantly by using temperature as a predictor variable and the subsequent addition of the other meteorological parameters and precursors had a lower contribution towards the final R^2 values. Relative humidity and wind speed also had effect on high ozone days, though on a smaller scale compared to temperature.

2) Geostatistical analysis:

The geostatistical investigation confirmed the diurnal pattern of ozone accumulation during the day followed by depletion in the evening hours. Semivariograms were generated by grouping data into hour-specific data files. These fitted variograms suggest that daytime ozone concentrations are controlled by local, short-range influences. The evening ozone concentrations, on the other hand, exhibit a more uneven pattern. However, patterns during late nights were difficult to explain. More data needs to be monitored to observe the true diurnal variation in ozone and actual actual processes governing the variation in ozone in Skiatook.

3) Regression analysis:

Models were developed to obtain 1-hr ozone, 8-hr ozone and daily maximum ozone concentrations as response. Two sets of predictor variables were used for each 1-hr and 8-hr ozone responses and one set of predictor variables was used to forecast daily maximum ozone concentration.

Temperature, wind speed, relative humidity, solar radiation, pressure, and NO_2 were used as predictor variables. In another set previous day ozone concentrations were used in addition to the above listed variables as predictor. A significant increase in the R^2 value was observed with inclusion of previous day ozone concentration among predictor variables. Residuals analysis was performed on each model to check the model quality. The models were also checked for their predicting ability by setting a criterion of predicted values being within ± 0.005 ppm of the actual ozone values.

Model 1 that used temperature, wind speed, relative humidity, solar radiation, pressure, and NO_2, though had the lowest R^2 value, showed highest predicting percentage of 35.88% based on the above criterion.

The unavailability of actual NO_2 data in Skiatook is causing the poor performance of the models developed. It is thus not possible to recommend what model is best suited for Skiatook.

7.2 Recommendations

The models developed here were intended to supply introductory knowledge for forecasting in Skiatook. It is recommended that in order to develop a detailed forecasting model for Skiatook, the effect of the following important processes should also be studied: synoptical weather patterns, aloft transport, local carryover processes, aloft temperature structure and weather patterns associated with cloud cover. Knowledge of these processes along with the findings from above research will help determine the cause of high ozone concentration. This will help in the selection of the right set of predictors to forecast ozone. Also a NO_2 monitor in Skiatook will provide the data representative of the actual conditions in Skiatook, and will improve the forecasting ability of the model in Skiatook.

BIBLIOGRAPHY

Atkinson, Roger, "Atmospheric Chemistry of VOCs and NO_x," Atmospheric Environment, Vol. 34, 2000, pp. 2063-2101.

Blanchard, Charles L. and Fairley, **David,** "Spatial Mapping of VOC and NOx-Limitation of Ozone Formation in Central California," Atmospheric Environment, Vol. 35, 2001, pp. 3861-3873.

Bloomfield, Peter, Royle, Andrew J., Steinberg, Laura J., and Yang, Qing, "Accounting for Meteorological Effects in Measuring Urban Ozone Levels and Trends," Atmospheric Environment, Vol.30, 1996, pp. 3067-3077.

Casado, Luis S., Rouhani, Shahrokh, Cardelino, Carlos A. and Ferrier, Adrian J., "Geostatistical Analysis and Visualization of Hourly Ozone Data," Atmospheric Environment, Vol. 28, 1994, pp. 2105-2118.

Chaloulakou, Archontoula, Saisana, Michaela and Spyrellis, Nikolas, "Comparative Assessment of Neural Networks and Regression Models for forecasting Summertime Ozone in Athens," The Science of the Total Environment, Vol.313, 2003, pp. 1-13.

Cooper, Richard M. and Istok, Jonathan D.,"Geostatistics Applied to Groundwater Contamination. I: Methodology," Journal of Environmental Engineering, Vol. 114, 1988, pp. 270-286.

Davis, J. M. and Speckman, P., "A Model for Predicting Maximum and 8h Average Ozone in Houston," Atmospheric Environment, Vol. 33, 1999, pp. 2487-2500.

Lin, Cynthia C.Y., Jacob, Daniel J. and Fiore, Arlene M., "Trends in Exceedances of the Ozone Air Quality Standard in the Continental United States, 1990-1998," Atmospheric Environment, Vol. 35, 2001, pp. 3217-3228.

Ludwig, F. L., Jiang, Jih-Yih and Chen, Jun, "Classification of Ozone and Weather Patterns Associated with High Ozone Concentrations in the San Francisco and Monterey Bay Areas," Atmospheric Environment, Vol. 29, 1995, pp. 2915-2928.

Olszyna, K. J., Luria, M. and Meagher, J. F., "The Correlation of Temperature and Rural Ozone Levels in Southeastern U.S.A.," Atmospheric Environment, Vol. 31, 1997, pp. 3011-3022.

Rozelle, John, "Geostatistics – What Exactly Does All This Mumbo-Jumbo Mean?," PINCOCK Perspectives, Issue No. 20, 2001, pp. 1-4.

Qin, Y., Tonnesen, G.S. and Wang, Z., "One-hour and Eight-hour Average Ozone in the California South Coast Air Quality Management District: Trends in Peak Values and Sensitivity to Precursors," Atmospheric Environment, Vol. 38, 2004, pp. 2197-2207.

Ryan, William F., "Forecasting Severe Ozone Episodes in the Baltimore Metropolitan Area," Atmospheric Environment, Vol. 29, 1995, pp. 2387-2398.

Tarasova, O.A., Karpetchko, A. Yu., "Accounting for Local Meteorological Effects in the Ozone Time-Series of Lovozero (Kola Peninsula)," Atmospheric Chemistry and Physics Discussions, Vol. 3, 2003, pp. 655-676.

Thompson, Mary Lou, Reynolds, Joel, Cox, Lawrence H., Guttorp, Peter and Sampson, Paul D., "A review of Statistical Methods for the Meteorological Adjustment of Tropospheric Ozone," Atmospheric Environment, Vol. 35, 2001, pp. 617-630.

Tiao, G. C., "Use of Statistical Methods in the Analysis of Environmental Data," The American Statistician, Vol. 37, 1983, pp. 459-470.

Vukovich, Fred M., "Regional-Scale Boundary Layer Ozone Variations in the Eastern United States and their Association with Meteorological Variations," Atmospheric Environment, Vol. 29, 1995, pp. 2259-2273.

U.S. EPA, "Good Up High Bad Nearby," AIRNow, http://www.epa.gov/oar/oaqps/gooduphigh/ozone.html

Vukovich, Fred M. and Sherwell John, "An Examination of the Relationship Between Certain Meteorological Parameters and Surface Ozone Variations in the Baltimore-Washington Corridor," Atmospheric Environment, Vol.37, 2003, pp. 971-981.

Zhang, Jian and Rao, Trivikrama S., " The Role of Vertical Mixing in the Temporal Evolution of Ground-Level Ozone Concentrations," Journal of Applied Meteorology, Vol.38, 1999, pp.1674-1691.

APPENDICES

APPENDIX A

Data sets for temporal distribution of ozone

Total annual 1-hr and 8-hr exceedance days from 1998 - 2000 in Oklahoma

Year	Total annual 1-hr exceedance days	Total annual 8-hr exceedance days
1998	2	59
1999	6	44
2000	6	63

County wise distribution of average 1-hr and 8-hr ozone exceedance days from 1998-2000 in Oklahoma

County	Average exceedances days	
	1-hr	**8-hr**
Cherokee	0	5
Cleveland	0.333333333	3
Comanche	0	3
Jefferson	0	3.666666667
Kay	0	2.333333333
Latimer	0.333333333	2
Love	0.333333333	6.333333333
Marshall	0	4
Mayes	0.333333333	1.333333333
McClain	0	0.666666667
Muscogee	0	0.333333333
Okmulgee	0	1
Oklahoma	0.333333333	2.666666667
Oklahoma	0.333333333	3
Tulsa	0.333333333	4.666666667
(Skiatook)	1.666666667	8.666666667
Tulsa	0.666666667	2.666666667
Tulsa	0	1

Average number of days with 8-hr and 1-hr exceedances by month from 1998-2000 for Skiatook

Month	**Avg annual exceedances**	
	8-hr	**1-hr**
Jan	0.00	0.00
Feb	0.00	0.00
Mar	0.00	0.00
Apr	0.00	0.00
May	0.00	0.00
Jun	0.00	0.00
Jul	0.67	0.33
Aug	4.33	0.00
Sep	3.33	1.33
Oct	0.00	0.00
Nov	0.00	0.00
Dec	0.00	0.00

8-hr average concentration on days that exceeded 0.085 ppm from 1998-2000 in Skiatook by hour of day

Hour	**Average number of exceedances days**	
	8-hr	**1-hr**
10:00		
11:00	1.50	0.67
12:00	2.00	0.67
13:00	4.67	0.67
14:00	8.33	0.33
15:00	7.67	
16:00	4.67	0.33
17:00	3.33	0.33
18:00	1.67	
19:00	0.67	
20:00	0.67	
21:00		
22:00		
23:00		

Average number of 8-hr and 1-hr exceedances by day of the week from 1998-2000 for Skiatook

Day	**Average annual exceedance days**	
	1-hr	**8-hr**
Monday	0	1
Tuesday	0	2
Wednesday	0	1
Thursday	0.33	0.333333333
Friday	1	1.666666667
Saturday	0.33	1.666666667
Sunday	0	0.666666667

Average number of 8-hr and 1-hr exceedances by day of the week from 1998-2000 for Oklahoma

Day	**Average exceedances**	
	8-hr	**1-hr**
Monday	9	0.333333333
Tuesday	11	0.666666667
Wednesday	8.333333333	0.333333333
Thursday	4.333333333	0.333333333
Friday	8.666666667	1.666666667
Saturday	11	1.333333333
Sunday	5.333333333	0.333333333

Average daily maximum temperature and monthly highest recorded temperature for Skiatook in 2000

Month	Ave Daily Max Temp (F)	Monthly Highest Temp (F)
JAN	49	71
FEB	59	75
MAR	61	80
APR	69	86
MAY	79	91
JUN	81	89
JUL	89	98
AUG	97	105
SEP	88	107
OCT	74	91
NOV	53	73
DEC	35	64

Raw data for Episode I: September 1 to 7, 1998

Date	1-hr max Ozone	8-hr max Ozone	TAIR	RELH	PRES	WSPD	SRAD
9/1	0.11	0.094	96	55	28.92	24.9	18.4
9/2	0.071	0.068	95	60	28.87	15.2	23.46
9/3	0.103	0.093	104	47	28.77	18.1	22.65
9/4	0.129	0.109	105	46	28.84	14.7	22.58
9/5	0.11	0.101	103	52	28.93	21.9	22.46
9/6	0.097	0.089	100	42	28.97	22.6	21.71
9/7	0.08	0.077	103	42	28.9	18.7	22.14

Raw data for Episode II: August 26 to September 1, 1999

Date	**1-hr max Ozone**	**8-hr max Ozone**	**TMAX**	**RELH**	**PRES**	**WSPD Max**	**SRAD**
8/26	0.085	0.073	106	84	28.89	25.7	23.14
8/27	0.095	0.074	96	87	28.94	20	21.01
8/28	0.093	0.087	98	95	28.99*	15.7	22.94
8/29	0.09	0.094	96	80	29.02	18.1	22.84
8/30	0.101	0.094	95	79	29.03	17.9	18.73
8/31	0.105	0.092	96	69	28.99	17.3	22.92
9/1	0.094	0.083	96	61	28.98	20.4	24.1

Raw data for Episode III: September 1 to 7, 2000

Date	**1-hr max Ozone**	**8-hr max Ozone**	**TAIR**	**RELH**	**PRES**	**WSPD**	**SRAD**
9/1	0.127	0.092	107	36	28.9	19.8	22.86
9/2	0.132	0.106	107	40	28.88	16.1	21.9
9/3	0.112	0.098	106	40	28.87	32.8	22.27
9/4	0.076	0.063	98	53	28.96	20.8	13.6
9/5	0.07	0.061	89	61	29.06	18.4	20.03
9/6	0.075	0.068	87	59	29.05	15.2	24.02
9/7	0.08	0.075	91	59	28.99	16.2	23.03

APPENDIX B

OUTPUT FROM MINITAB 14 OF STEP WISE REGRESSION FOR MODEL 1

Response is 1-hr ozone on 6 predictors, with N = 3309

Step	1	2	3	4	5	6
Constant	0.03172	0.03780	0.08035	0.08266	0.08389	-0.12392
TAIR	0.00112	0.00134	0.00081	0.00093	0.00092	0.00097
T-Value	20.47	25.14	14.32	15.29	15.29	15.69
P-Value	0.000	0.000	0.000	0.000	0.000	0.000
WSPD		-0.00352	-0.00403	-0.00395	-0.00404	-0.00382
T-Value		-18.88	-22.66	-22.23	-22.60	-20.25
P-Value		0.000	0.000	0.000	0.000	0.000
RELH			-0.00038	-0.00042	-0.00042	-0.00041
T-Value			-20.17	-20.56	-20.43	-19.87
P-Value			0.000	0.000	0.000	0.000
SRAD				-0.00001	-0.00001	-0.00001
T-Value				-5.22	-5.21	-5.78
P-Value				0.000	0.000	0.000
NO2 (ppm)					-0.244	-0.250
T-Value					-4.07	-4.17
P-Value					0.000	0.000
PRES						0.00021
T-Value						3.45
P-Value						0.001
R-Sq	11.25	19.88	28.67	29.25	29.61	29.86

OUTPUT FROM MINITAB 14 OF STEP WISE REGRESSION FOR MODEL 2

Response is 1-hr ozone on 7 predictors, with N = 3270

Step	1	2	3	4	5	6
Constant	0.03100	0.01767	0.02353	0.05561	0.05694	-0.15232
TAIR	0.00115	0.00050	0.00072	0.00041	0.00040	0.00043
T-Value	20.69	9.50	13.69	7.45	7.38	7.82
P-Value	0.000	0.000	0.000	0.000	0.000	0.000
Prev day 1-hr		0.492	0.453	0.402	0.406	0.410
T-Value		31.48	29.71	26.68	27.05	27.28
P-Value		0.000	0.000	0.000	0.000	0.000
WSPD			-0.00272	-0.00318	-0.00328	-0.00307
T-Value			-16.20	-19.26	-19.85	-17.62
P-Value			0.000	0.000	0.000	0.000
RELH				-0.00027	-0.00026	-0.00025
T-Value				-15.35	-15.10	-13.86
P-Value				0.000	0.000	0.000
NO2 (ppm)					-0.316	-0.323
T-Value					-5.77	-5.91
P-Value					0.000	0.000
PRES						0.00021
T-Value						3.85
P-Value						0.000
R-Sq	11.58	32.15	37.20	41.43	42.02	42.28

OUTPUT FROM MINITAB 14 OF STEP WISE REGRESSION FOR MODEL 3

Response is 8-hr ozone on 6 predictors, with N = 3322

Step	1	2	3	4	5
Constant	0.02542	0.05748	0.06775	0.06887	-0.03651
TAIR	0.00120	0.00078	0.00090	0.00090	0.00091
T-Value	26.49	15.59	19.07	19.10	19.22
P-Value	0.000	0.000	0.000	0.000	0.000
RELH		-0.00029	-0.00034	-0.00033	-0.00033
T-Value		-17.71	-21.52	-21.38	-20.42
P-Value		0.000	0.000	0.000	0.000
WSPD			-0.00307	-0.00315	-0.00304
T-Value			-20.65	-21.08	-19.40
P-Value			0.000	0.000	0.000
NO2 (ppm)				-0.221	-0.224
T-Value				-4.40	-4.46
P-Value				0.000	0.000
PRES					0.00011
T-Value					2.12
P-Value					0.034
R-Sq	17.45	24.58	33.17	33.56	33.65

OUTPUT FROM MINITAB 14 OF STEP WISE REGRESSION FOR MODEL 4

Response is 8-hr ozone on 7 predictors, with N = 3296

Step	1	2	3	4	5	6
Constant	0.02488	0.01196	0.01604	0.04357	0.04478	-0.11696
TAIR	0.00122	0.00052	0.00068	0.00042	0.00042	0.00044
T-Value	26.70	12.17	15.78	9.59	9.54	9.92
P-Value	0.000	0.000	0.000	0.000	0.000	0.000
Prev day 8-hr		0.550	0.520	0.469	0.472	0.475
T-Value		37.28	36.00	33.00	33.42	33.63
P-Value		0.000	0.000	0.000	0.000	0.000
WSPD			-0.00194	-0.00232	-0.00241	-0.00224
T-Value			-14.41	-17.67	-18.36	-16.27
P-Value			0.000	0.000	0.000	0.000
RELH				-0.00023	-0.00023	-0.00022
T-Value				-16.67	-16.44	-15.25
P-Value				0.000	0.000	0.000
NO2 (ppm)					-0.276	-0.281
T-Value					-6.35	-6.48
P-Value					0.000	0.000
PRES						0.00016
T-Value						3.75
P-Value						0.000
R-Sq	17.79	42.19	45.62	49.86	50.47	50.68

Step	7
Constant	-0.09247
TAIR	0.00037
T-Value	7.74
P-Value	0.000
Prev day 8-hr	0.481
T-Value	33.75
P-Value	0.000
WSPD	-0.00229
T-Value	-16.54
P-Value	0.000
RELH	-0.00020
T-Value	-12.70
P-Value	0.000
NO2 (ppm)	-0.281
T-Value	-6.50
P-Value	0.000
PRES	0.00014
T-Value	3.10
P-Value	0.002

```
SRAD            0.00000
T-Value            3.06
P-Value           0.002

R-Sq              50.82
```

OUTPUT FROM MINITAB 14 OF STEP WISE REGRESSION FOR MODEL 4

Response is Daily Max 1-hr ozone on 5 predictors, with N = 545

Step	1	2	3	4
Constant	-0.02577	-0.01395	0.02050	0.62843
TMAX	0.00104	0.00106	0.00091	0.00078
T-Value	16.87	17.93	14.84	11.18
P-Value	0.000	0.000	0.000	0.000
WSPD		-0.00200	-0.00218	-0.00266
T-Value		-6.99	-7.85	-8.75
P-Value		0.000	0.000	0.000
RELH			-0.00030	-0.00036
T-Value			-6.62	-7.59
P-Value			0.000	0.000
PRES				-0.0203
T-Value				-3.68
P-Value				0.000
R-Sq	34.38	39.81	44.32	45.6

VITA

Dharmik Vadel

Candidate for the Degree of

Master of Science

Thesis: A STUDY OF SPATIAL DISTRIBUTION OF OZONE AND ITS CORRELATION WITH METEOROLOGICAL PARAMETERS AND PRECURSORS

Major Field: Environmental Engineering

Biographical:

Personal Data: Born in Bhavnagar, Gujarat in India on May 4th, 1980, the son of Rasik and Jyoti Vadel.

Education: Graduated from Fatima Convent High School, Bhavnagar, Gujarat in May 1997; received Bachelor of Technology degree in Civil Engineering from National Institute of Technology (formerly Regional Engineering College), Calicut, India in June 2002. Completed the requirements for the Master of Science degree with a major in Environmental Engineering at Oklahoma State University in December, 2005.

Experience: Worked as a Lecturer at Shantilal Shah Engineering College, Bhavnagar, India after completion of undergraduate studies; employed by Saurashtra Cement Ltd., Bhavnagar, India; worked in capacity of a Graduate Research Assistant under the guidance of Dr. Dee Ann Sanders in the department of Environmental Engineering from January – December, 2004; worked as a Graduate Lab Assistant with the College of Engineering Architecture and Technology Laboratories in Spring 2005 and Fall 2005 terms; employed as a summer intern at REHAU Inc., Leesburg in the Summer of 2005.

Professional Memberships: American Society of Civil Engineers.

Name: Dharmik R Vadel Date of Degree: December, 2005

Institution: Oklahoma State University Location: Stillwater, Oklahoma

Title of Study: A STUDY OF SPATIAL DISTRIBUTION OF OZONE AND ITS CORRELATION WITH METEOROLOGICAL PARAMETERS AND PRECURSORS

Pages in Study: 88 Candidate for the Degree of Master of Science

Major Field: Environmental Engineering

Scope and Method of Study:
The existing status of Oklahoma is complete attainment with the 1-hr and 8-hr ozone standards set by the NAAQS. However, Skiatook a suburban site in Tulsa county showed higher exceedances of ozone standards in the years 1998 to 2000. Temporal analysis was performed to study ozone exceedance patterns and geostatistical analysis was performed to study diurnal variation of ozone in Oklahoma and specifically in Skiatook. An investigation was performed on high ozone episodes in Skiatook to study the effects of temperature, relative humidity, wind speed, solar radiation and pressure on ozone concentrations. Multi-linear regression analysis was performed to forecast next day 1-hr ozone, 8-hr ozone, and daily maximum 1-hr ozone in Skiatook. The data from 10:00 a.m. to 5:00 p.m. for the months of May to September from 1998 – 2000 was used for the analysis. The forecasted values obtained from these models were tested to be within ± 0.005 ppm of the actual ozone values. The geostatiscital analysis was performed using software GS+ and the regression analysis was performed using MINITAB 14.

Findings and Conclusions:
Temperature was the only meteorological parameter found to have noticeable effect on high ozone concentrations in Skiatook. The model using the meteorological parameters and nitrogen dioxide as predictors showed the highest percentage (35.88 %) of satisfying the criterion of forecasted values being within ± 0.005 ppm of the actual ozone values. It was found that using the previous day ozone concentration as a predictor helped achieve a higher R^2 value for ozone forecast. The unavailability of actual nitrogen dioxide concentrations in Skiatook resulted in poor forecasts using the models developed.

ADVISER'S APPROVAL: Dr. DEE ANN SANDERS