An analysis of the relationship between the planetary boundary layer and surface level ozone concentrations in Phoenix, Arizona

David G. Lopez

School of Geographical Sciences and Urban Planning
Arizona State University

Honors Thesis - Spring 2023
Barrett Honors College
Arizona State University

Readers
R.S. Cerveny
R.C. Balling, Jr.

Table of Contents

Abstract

This study examines the linkage between surface level ozone and planetary boundary layer meteorological variables in the Phoenix Metropolitan region during the summer North American Monsoon period for years 2010 through 2020. Data used in this study was obtained and derived from both 1200 UTC radiosonde observations launched from the Phoenix National Weather Service office, and 8-hour average ozone concentration measurements from Maricopa County monitoring stations. Specific boundary layer meteorological variables examined in this study included inversion temperature, mixing level pressure, mixing level height, and the surface level variables of temperature, dew point temperature, pressure, wind speed, and meridional and zonal wind directions. The daily maximum, 8-hour average ozone concentrations among all Maricopa County monitoring stations were used in this study. To determine ozone's linkage to meteorological variables, normality tests, determination of Pearson product moment correlation coefficient and/or the Spearman rank correlation coefficient, and the discriminative Student's two-sided t-test statistic between ozone exceedance and non-exceedance days were used. Statistically significant coefficients indicate weak negative correlations between surface level ozone and surface level pressure, and mixing level pressure, and weak positive correlations between surface level ozone and surface level temperature, surface level zonal wind direction, mixing level height, and inversion temperature. These correlations were linear for surface level pressure, surface level temperature, and inversion temperature. The two-sided Student's t-test statistic indicates a significant difference in the mean on ozone exceedance and non-exceedance days for surface level temperature, and the upper-air variables of mixing level height, mixing level pressure, and inversion temperature. Both correlations and differences in the mean of upper-air variables showed statistically significant results. These findings suggest that further research should be completed to determine the forecasting ability of morning sounding analyses on surface level ozone in locations exhibiting similar emissions and geographic features as the Phoenix Valley.

1. Introduction

1.1 Ozone as a forecast problem

In the lower troposphere, ozone formation occurs in a daily cycle, dependent on concentrations of the primary pollutants of nitrogen oxides and volatile organic compounds when in the presence of ultraviolet radiation. Surface level ozone is a secondary pollutant as it is not directly emitted to the atmosphere. It is considered as the most damaging air pollutant (Sicard et al., 2016) because it presents several adverse health effects on humans (American Lung Association, 2022) and the environment (National Research Council, 2004). As such, in the US, the Environmental Protection Agency (EPA) regulates surface level ozone with a National Ambient Air Quality Standard (NAAQS). However, surface level ozone is highly dependent on both local geography and emissions, making it difficult to regulate and forecast.

In Arizona, the Department of Environmental Quality (ADEQ) is responsible for forecasting surface level ozone concentrations and setting statewide plans to mitigate the number of days reaching the ozone exceedance threshold (NAAQS) set by the EPA. Surface level ozone concentrations in the Phoenix Metropolitan region tend to be elevated during the summer North American Monsoon (NAM), with frequent ozone exceedances at stations across the valley. A linkage between meteorological variables (with known forecast techniques) and surface level ozone concentrations in the Phoenix Metropolitan region during NAM season, could increase the set of tools for air quality forecasters at ADEQ to use.

1.2 Research question and hypothesis

Given the importance of ozone to health and therefore the value of air quality forecasting, the following research question can be posed: Can boundary layer meteorological variables be used as a predictor for surface level ozone concentrations and ozone exceedance days in the Phoenix Metropolitan region during the North American Monsoon? To examine that question, I will conduct a set of statistical tests. These include the determination of normality, calculation of correlation coefficients, and implementation of Student's two-sided t-tests to determine if there is a significant difference in the mean of each variable on ozone exceedance and non-exceedance days. I hypothesize that those analyses will show that several near-surface weather variables, as measured by radiosondes, will demonstrate a statistical relationship to surface level ozone variations during the NAM season.

1.2 Structure of honors thesis

To initiate this research, I review the important literature associated with surface level ozone, giving special interest on its linkages to climate, the Southwest, and the NAM season.

I then construct a data set of both meteorological variables within the planetary boundary layer and surface level ozone concentrations within the Phoenix Metropolitan region for NAM seasons from the year 2010 through 2020. Weather observations are derived from radiosonde measurements launched from the Phoenix National Weather Service office. Specific boundary layer meteorological variables examined in this study include inversion temperature, mixing level pressure, mixing level height, and the surface level variables of temperature, dew point temperature, pressure, wind speed, and meridional and zonal wind directions. Ozone data is collected as 8-hour ozone concentration observations from Maricopa County monitoring stations.

To test my hypothesis, I then conduct a set of statistical tests. Depending on the normality of each variable, I determine the Pearson product moment correlation coefficient and/or the Spearman rank correlation coefficient to ozone. The ozone data set includes a column to classify ozone exceedance days, allowing me to use the discriminative Student's two-sided t-test statistic to determine if there was a difference in the mean value of variables on ozone exceedance and non-exceedance days.

Following the statistical analysis, I interpret the results with the goal of establishing the validity or non-validity of my research hypothesis. Finally, I establish the significance of those results with regard to my hypothesis and present suggestions for future research on this subject.

2. Literature Review

2.1. Introduction

In order to understand the importance of the relationship between surface level ozone and meteorological variables, I have conducted a literature review focusing on past research where surface level ozone concentrations have been both correlated and attributed to changes in meteorological variables. I start by presenting surface level ozone basics and the regulation thereof. I then discuss several studies investigating the relationship between climate and ozone, specifically reviewing the impact of inversions and the planetary boundary layer with focus on the US Southwest region. Since my study is focused on the summer in Phoenix, I also review the North American Monsoon.

Ultimately, this research will allow me to test my hypothesis that ozone measurements are statistically linked to variations in Phoenix weather as measured by radiosonde variables. To begin my analysis, I start with reviewing the basic chemistry of surface level ozone.

2.2. Surface Level Ozone Basics

Surface level ozone (SLO), also known as ground level ozone and tropospheric ozone, is considered one of the major scientific challenges associated with urban air pollution (Sillman, 1999). This is due to SLO being largely influenced by anthropogenic sources, and that SLO presents several adverse health effects on humans (American Lung Association, 2022) and the environment (National Research Council, 2004). SLO is a secondary pollutant formed through sets of chemical reaction cycles involving Volatile Organic Compounds (VOCs - hydrocarbons), and Nitrous Oxides ($NOx = NO + NO2$). A diurnal cycle is introduced to both the formation and concentration of SLO, as the NO2 photolytic cycle, the basis for SLO formation, involves photochemistry. The NO2 photolytic cycle is as follows (Hobbs, 2000; Oke, 1987):

$$NO2 + hv \rightarrow NO + O \tag{1.1}$$

$$O + O2 \rightarrow O3 \tag{1.2}$$

$$O3 + NO \rightarrow NO2 + O2 \tag{1.3}$$

Nitrogen dioxide (NO2) photodissociates when in the presence of ultraviolet radiation (hv), in the waveband 0.37-0.42 μm, forming nitric oxide (NO) and atomic oxygen (O) (1.1). The atomic oxygen is highly reactive and combines with ambient molecular oxygen (O2) to form ozone (O3) (1.2). Ozone then reacts with NO to form NO2 and O2 (1.3). The photolytic cycle is a naturally occurring cycle that acts on a near instantaneous time scale, in which no net ozone is created while in the presence of NO (Hobbs, 2000; Oke, 1987). However, when VOCs are introduced into the system, reactions of NO with VOCs become favored, allowing for the buildup of SLO during the day. Additional chains of reactions involving VOCs that ultimately lead to ozone production are highly complex (Sillman, 1999).

Logan (1985) argued that SLO concentrations during the summer are significantly influenced by photochemical production of ozone associated with emissions of NOx, hydrocarbons, and carbon monoxide (CO). Logan's claim has since been supported by research surveyed by Vingarzan (2004), stating that sites exhibiting local ozone production (photochemical production) display a broad summer maximum.

On a shorter timescale, there exists a cycle called the ozone "weekend effect" where SLO concentrations increase on the weekends. This weekly cycle has been observed in several locations including Phoenix (Shutters and Balling, 2006), and notably Los Angeles, California, where the Southern California Air Quality Study took place (Lawson, 2012). This "weekend effect" is believed to be directly related to weekly transportation cycles (Gao, 2007; Shutters and Balling, 2006), and has been observed in several locations globally, primarily in valleys (Sicard et al, 2020). My area of interest, the Phoenix Metropolitan region, is located within a valley.

An example of weekly and diurnal cycles of O3, NOx, and CO concentrations, can be described by observations in Azusa during the summer of 1995 (Figure 1.1). Azusa is a city located in Los Angeles County, California, and the San Gabriel Valley. The ozone "weekend effect" was observed there during that study (Fujita et al., 2012). NO2, and CO show two peaks throughout the day, with a morning maximum and secondary peak in the late afternoon. NO also shows a morning maximum, however instead of a late afternoon peak, concentrations slowly build up from late afternoon throughout the night. As would be expected within the photolytic cycle, O3 concentrations increased throughout the day while NO concentrations decreased to a near daily zero minimum concentration. Weekend O3 concentrations were highest during the weekend while NO, NO2, and CO concentrations were lowest on the weekends during the daytime. NO, NO2, and CO concentrations were highest on the weekends during the nighttime.

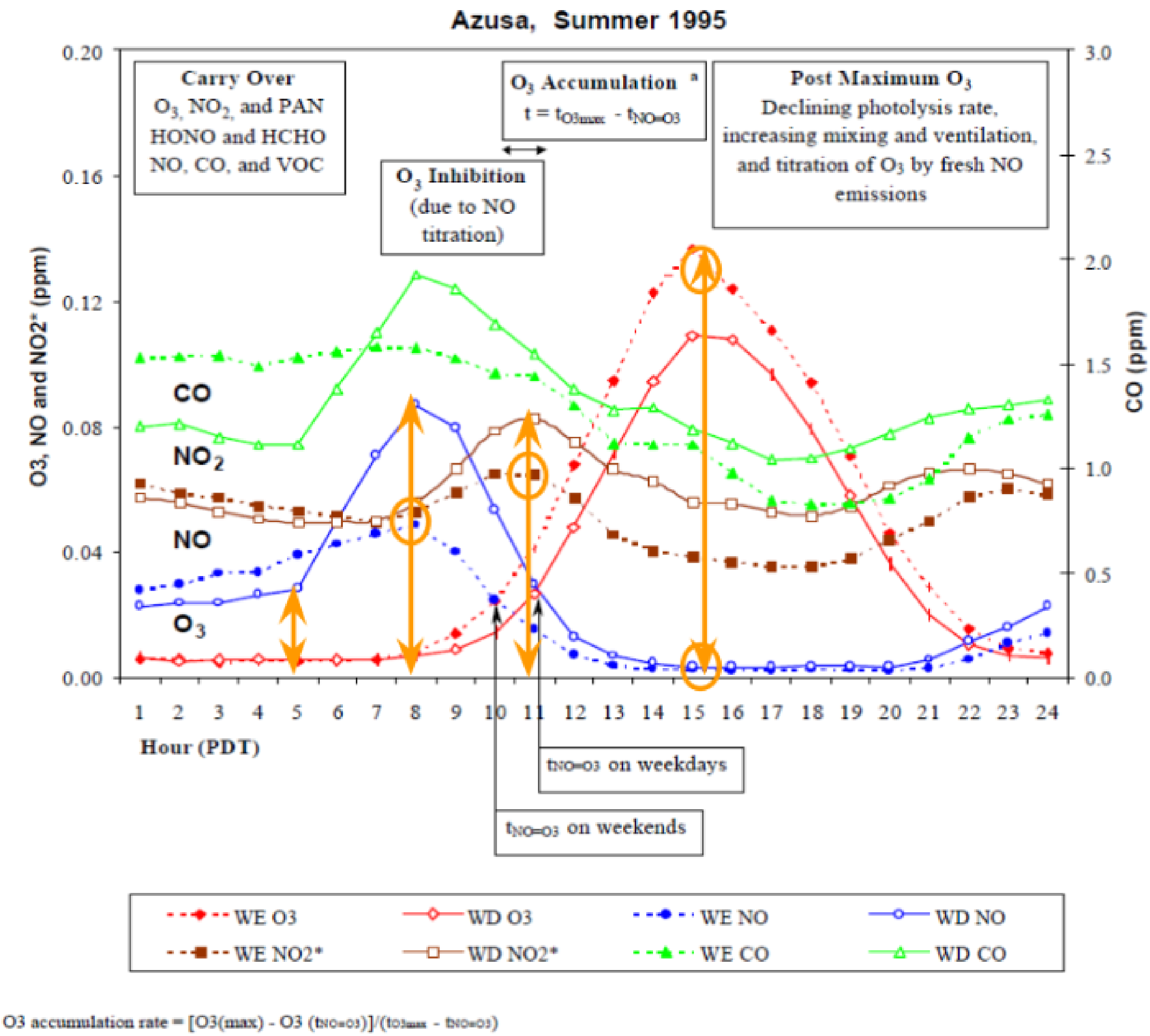

Figure 1.1. Mean hourly summer 1995 diurnal variations of O3 (red lines), NO (blue lines), NO2 (brown lines), and CO (green lines) at Azusa during the weekday (solid lines) and weekend (dashed lines), all measured in parts per million (ppm). Azusa is located in the San Gabriel Valley of Los Angeles County, California. Source: Fujita et al., 2003

2.3. Regulation

In the United States, the Environmental Protection Agency (EPA) regulates SLO concentrations with a National Ambient Air Quality Standard (NAAQS) of 70 ppb or ‘parts per billion’ (United States Environmental Protection Agency, 2022). A location is said to be in violation of the NAAQS for ozone, after it has, in four instances, surpassed an 8-hour average ozone concentration of 70 ppb. The EPA began regulating SLO in 1996. Unfortunately, ozone emission sources cannot be directly regulated as ozone is a secondary pollutant and is not directly emitted to the atmosphere. It is important to monitor ozone due to the adverse health effects on humans (American Lung Association, 2022) and the environment (National Research Council, 2004). The EPA also regulates select ozone precursors, each presenting additional health concerns: carbon monoxide (CO), nitrogen dioxide (NO2), particulate pollution (PM), and sulfur dioxide (SO2) (United States Environmental Protection Agency, 2022).

NO2 is regulated by the EPA with a NAAQS of 53 ppb averaged over a year, and of 100 ppb for the 98th percentile of one hour daily maximum concentrations averaged over three years. NO2, as previously stated, can lead to increased SLO concentrations through photolysis when also in the presence of VOCs (which currently are not regulated by the EPA). NO2 however can also be captured in moisture, alongside sulfur oxides (SOx) to form acid rain (EPA, 1999). SO2 is regulated by the EPA with a NAAQS of 0.5 ppm or 'parts per million' averaged over three hours, not to be exceeded more than once per year, and of 75 ppb for the 98th percentile of one hour daily maximum concentrations averaged over three years. NO2 is the only NOx regulated by the EPA, however understanding and controlling the generation of the entire NOx family of air pollutants is necessary to prevent the unwanted consequences (EPA, 1999). CO is regulated by the EPA with a NAAQS of both 9 ppm averaged over eight hours, and 35 ppm averaged over one hour, not to be exceeded more than once per year. Both CO and SO2 are involved in highly complex chains of reactions involving VOCs that ultimately lead to ozone production (Sillman, 1999). PM is regulated by the EPA in particle sizes of 2.5 μm (PM2.5) and 10 μm (PM10), or 'micrometers.' PM2.5 has a NAAQS of 12.0 $\mu g/m^3$ or 'micrograms per cubic meter' as an annual mean averaged over three years, and of 35.0 $\mu g/m^3$ for the 98th percentile of daily average concentrations averaged over three years. PM10 has a NAAQS of 150 $\mu g/m^3$ for the 98th percentile of daily average concentrations, not to be exceeded more than once per year, on average, over three years. Particulate pollution has been linked to ozone concentrations through its ability to scatter light (Heuss et al., 2012). It has been observed in Los Angeles, California, that ozone concentrations increase when concentrations of particulate pollution are lower (Gao, 2007; California Air Resources Board, 2003).

2.4. Climate and Ozone

Several meteorological variables have been linked to the formation and concentrations of SLO when examined from a climatological perspective (Sillman et al., 1990; Cox and Chu, 1996; Wise and Comrie, 2005; Pearce et al, 2011).

Sillman et al. (1990) used a regional photochemical model to examine the sensitivity of ozone concentrations in rural areas across the United States to concentrations of NOx and hydrocarbons. During this study they conducted simulations with temperatures between 283 K and 303 K, keeping other variables constant, to examine the sensitivity of ozone to temperature. They found that rural ozone concentrations increase by about 8 ppb per increase in temperature of 5 K. Sillman et al. stated that ozone forms more rapidly in conditions with warm temperatures and sunshine.

In a study by Pearce et al. (2011) from 1999 to 2006 in Melbourne, Australia, the strongest variations in maximum 8-hour ozone concentrations observed were related to maximum daily temperature, boundary layer height observed at 4 pm local time, and daily radiation sum, with a positive, negative, and positive respective non-linear relationship. Additional variables under investigation by Pearce et al. included: mean sea-level pressure, water vapor pressure, zonal and meridional wind components, and precipitation.

Wise and Comrie (2005) conducted a study from 1990 to 2003 for five major metropolitan areas across the Southwest United States: Albuquerque, NM; El Paso, TX; Las Vegas, NV; Phoenix, AZ; and Tucson, AZ. They found mixing height, the vertical extent of the lowest part of the planetary boundary layer where friction is dominant, to be the strongest correlated to ozone in the Southwest, an “important contrast to other parts of the United States where temperature has been found to be by far the strongest controlling factor” (Wise and Comrie, 2005, 2977). They found that the relationship between ozone concentrations and mixing height was always positive despite the presumed association between stagnant conditions and ozone buildup. They attribute this relationship to two hypotheses that require further study; increased entrainment of polluted air downward increasing ozone concentrations, and an increase of ozone precursors emitted when mixing level heights are higher resulting in an increase of ozone concentrations. Wise and Comrie noted that mixing heights were not directly measured for Phoenix, but rather a proxy from Tucson soundings were used.

While in the Wise and Comrie (2005) study, mixing height was the strongest correlated variable to ozone concentrations, the ozone concentrations were most strongly influenced by temperature (also strongly correlated to ozone concentrations), with additional influence from wind speed and dew point temperature. However, they did not include seasonal variability within their study, which will be considered in the present study.

2.5. Ozone’s Linkages to Inversions, the Planetary Boundary Layer, the Southwest, and the North American Monsoon

As described by Oke (1978), there are several effects of atmospheric stability on the dispersion of pollutants in the planetary boundary layer. The worst conditions for dispersion occur when there is a temperature inversion and the boundary layer is stable. A subsidence inversion, an inversion due to warming, can form an effective “lid,” resulting in stagnation of the pollutants at the surface below the inversion layer, and a thoroughly murky mixed layer.

2.5.1. Inversions

Oke (1978) also described an inversion that can form due to advection over an urban area at night. Within a city, the heat island effect is capable of maintaining a mixing layer throughout the night (Figure 2.2). Fumigation dispersion acts in tandem with an inversion above a plume, obstructing upward dispersion, where ample buoyant mixing brings the plume contents to the surface (Figure 2.1). Oke also described the potential for a valley fumigation event to occur, with an example of the observed process taking place in the Columbia River valley at Trail, British Columbia (Dean et al., 1944). This process took place after sunrise heating of the valley floor, occurring almost simultaneously along the valley for distances up to 55 km away from the source. While ozone does not build up during the night (due to its dependence on ultraviolet radiation within the photolytic cycle), the containment of ozone precursors during the night with the addition of fumigation in the mixed layer in the morning can lead to increased ozone concentrations throughout the day.

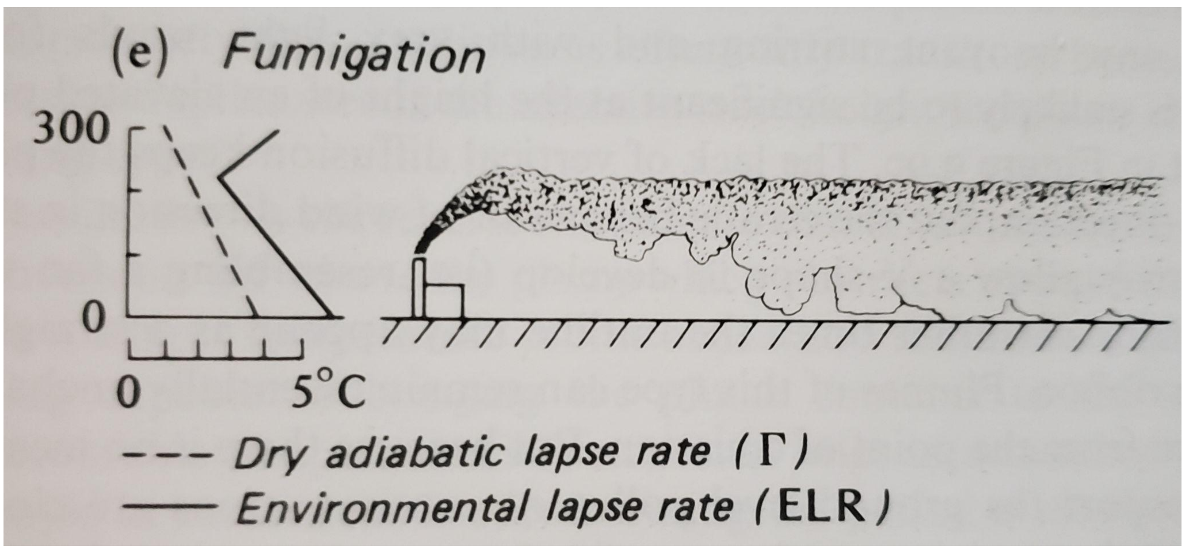

Figure 2.1. Schematic of fumigation dispersion of pollutants from a smoke stack. Source: Oke (1978)

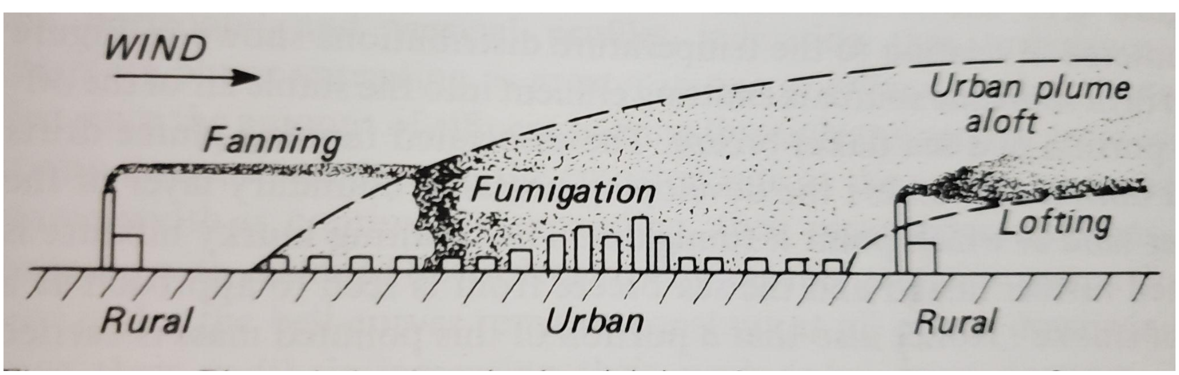

Figure 2.2. Schematic of fumigation within a city at night with clear skies and light winds. Source: Oke (1978)

2.5.2. The Southwest

One way of which ozone concentrations can increase is by regional transport of pollutants. Li et al. (2015) conducted a study investigating the impacts of Southern-California emissions on Phoenix SLO concentrations. Using a high resolution WRF-Chem model they determined for two high exceedance cases (14 May 2012, and 19 July 2005 - 80 ppb or greater at at least 10 Phoenix Metropolitan monitoring stations) that 10 - 30% of Phoenix ozone can be attributed to Southern-California emissions. Li et al. determined that the physical mechanism responsible for this transport to be strong westerly winds advecting ozone within a stable planetary boundary layer at night.

Leffel (2022) completed an honors thesis examining the influence of surface level meteorological factors on surface level ozone concentrations and ozone exceedance and non-exceedance days in the Phoenix nonattainment area for the years 2010 through 2020. Leffel found that the most significant factors associated with the occurrence of ozone exceedance days are wind speed and temperature. Leffel concluded that stable synoptic conditions allow for clearer skies, more sunlight, higher temperatures, and stagnant air, being ideal conditions for an ozone exceedance. It is important to note that Leffel uses 2-meter surface level meteorological observations, while in my study surface level observations are derived from radiosonde launches off of a four-story building, also being confined to the North American Monsoon season.

2.5.3. The Planetary Boundary Layer

Fast et al. (2000), on the evolution of the planetary boundary layer and its effect on air chemistry in the Phoenix area, found that multiple-day buildups of locally produced ozone in May and June of 1998 did not seem to be a major contributing factor to observed ozone concentrations due to a regularly observed ventilation of the boundary layer. However, they noted that particles released during the evening do remain over Phoenix, trapped in the nocturnal boundary layer with relatively light winds, indicating that a buildup of ozone precursors at night may affect ozone production the following day. These findings are fairly synonymous to Logan's (1985) general findings that ozone or its precursors in the boundary layer are most likely to influence the distribution of ozone at higher altitudes in summer, when convective mixing is most effective. As temperatures decrease, the height of the planetary boundary layer (in which mixing occurs) decreases. Inversions also become stronger when present, hindering vertical mixing.

Vingarzan (2004) surveyed literature on ozone precursors from background stations in Canada, the United States, and around the world. A background station is a station where the influences by anthropogenic emission sources are assumed not to be of local origin. During the wintertime a pool of accumulated NOx and hydrocarbons can build up (Holmes et al., 2015; Dibb et al., 2003; Liu et al, 1987). NOx has a longer lifetime in the Winter (Liu et al, 1987). It appears that this build up was the major cause for spring increases in ozone concentrations, alongside an increase in solar radiation during the spring (Vingarzan, 2004), while being synoptically transported by low level winds (Liu et al, 1987).

2.6 The North American Monsoon.

Due to the availability of upper air observations that will be used in this study being confined to the North American Monsoon season for the Phoenix Metropolitan region, it is necessary to introduce past research on the NAM season that may influence patterns in the meteorological observations used in this study.

Historically one of the identifying names used for the NAM was the Arizona monsoon, however the area affected by the NAM is significantly greater than Arizona (Adams and Comrie, 1997). Adams and Comrie (1997) note that the phenomenon is centered over northwestern Mexico, influencing large areas of the Southwestern United States. A general consensus amongst past researchers for the cause of the monsoon is a subtropical ridge providing anticyclonic circulation that advects moisture primarily from the Gulf of California, with some upper-level moisture from the Gulf of Mexico (Adams and Comrie, 1997). In Arizona, the NAM is associated with severe thunderstorms (Maddox et al., 1995), including cloud to ground lightning, and dust storm generation (Brazel and Nickling, 1986).

Granados-Muñoz et al. (2017) found that lightning induced NOx formation during the NAM season has the potential to lead to an increase of 23 ppb in the free troposphere (6 to 10 km above surface) in the Southwestern United States. They further describe the monsoonal anticyclonic winds as a transport mechanism for lightning induced NOx and tropospheric ozone into the Southwestern U.S. An enhancement of ozone concentrations in the Southwestern U.S. occurs when aerosols blow from the maximum lightning induced NOx region of the Southeastern U.S. and Mexico, as well as when aerosols blow from the maximum lightning induced NOx region of the Gulf of California. These two sources align with the NAM moisture sources described by Adams and Comrie (1997).

2.7. Summary

In order to test my hypothesis that boundary layer meteorology influences ozone concentrations in the Phoenix Metropolitan region, I conducted a literature review concerning major aspects of my research. Past research has concluded that several localized meteorological variables may affect surface level ozone concentrations. For example, past research has indicated that surface temperature is significantly and strongly correlated to surface level ozone concentrations in many locations. However, the correlation of surface level ozone to other meteorological variables, such as wind speed and direction, and ozone precursors, have been found to be highly dependent on local geographic features and emission sources. Past research indicates that features observed in the morning within the planetary boundary layer, such as inversion strength and mixing level height, may also influence ozone concentrations throughout the day. A study completed in the Phoenix Metropolitan region using radiosonde data from the Tucson National Weather Service office to calculate mixing level heights, found that mixing height was the most highly correlated variable to ozone concentrations.

The literature review has provided a foundation upon which I can initiate my study of the effects of planetary boundary layer meteorology on ozone concentrations. Therefore, I am now able to proceed to a discussion of the data that I used in this study.

3. Data

3.1. Introduction

There are several considerations that need to be taken into account with regards to the data set that is used in this study. In order to make this study applicable to EPA regulations, an 8-hour average ozone value must be examined. In order to study the effects of boundary layer meteorological variables on ozone concentrations, a data set applicable to the Phoenix Metropolitan region must be available for both ozone and radiosonde observations. In order to make this study applicable to health concerns, it is ideal to study whether or not changes in boundary layer meteorological variables are associated with ozone exceedance days or not.

To continue my study on the relationship between planetary boundary layer meteorology and surface level ozone concentrations, I needed to obtain both meteorological and ozone data for my study area, the Phoenix Metropolitan region. Meteorological data were collected from 1200 UTC (5:00 am LST) radiosonde observations launched by the Salt River Project (SRP) from the Phoenix National Weather Service (NWS) office during the NAM season. 8-hour average surface level ozone data was collected from the EPA's Air Quality System (AQS). Together, a combined data set for the monsoon months (June-September) between years 2010 through 2020 has been compiled and examined for the Phoenix Metropolitan region. The specifics of the ozone and the radiosonde data are given in the following sections.

3.2.1. Ozone

Past research indicates that ozone precursor emissions might be changing (Fujita et al, 2003), and to avoid bias in variance that is not considered in this study, the period of this study was therefore chosen to extend only from years 2010 through 2020. Within this period there have been several changes to the location and objectives of monitoring stations within the Phoenix Metropolitan region. Data from a total of 25 unique ozone monitoring stations in Maricopa County have been used in this study, though a handful of the stations were not active for the whole period of years 2010 through 2020.

In this study I use the maximum, maximum daily 8-hour average ozone concentration across all the Maricopa County monitoring stations (county id: 13). The EPA regulates ozone exceedances with 8-hour average concentrations, and it is therefore more meaningful to study the 8-hour concentrations as opposed to hourly ozone concentrations (which are also reported to the EPA). I use the maximum, maximum daily 8-hour concentration across all stations as opposed to average maximum value across all stations, as past research indicates ozone to often be transported across the Phoenix valley throughout the day. The station that observes the maximum daily ozone concentration is likely observing the buildup of ozone from sources outside of the station's immediate vicinity, representative of secondary emission sources across the Phoenix Metropolitan region.

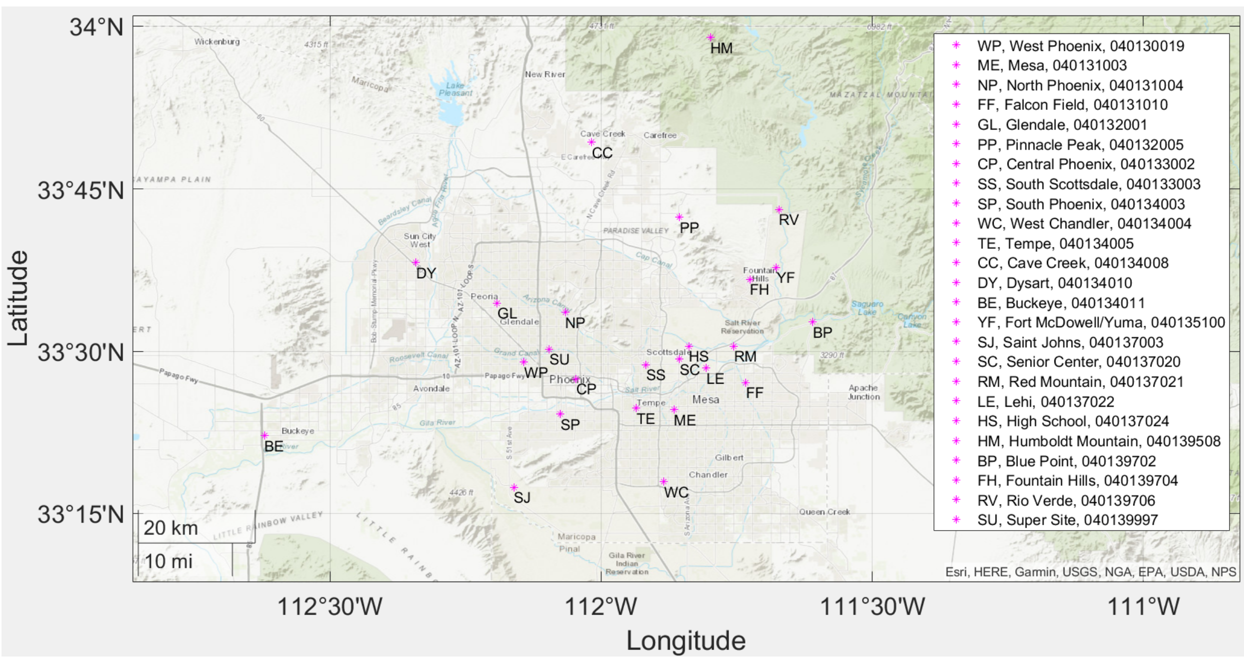

Figure 3.1. Maricopa County ozone monitoring stations.

The EPA defines 8-hour average ozone concentrations as the forward 8-hour average ozone concentration, with at least 75% completeness of the data set. Ozone in Maricopa County, covering the Phoenix Metropolitan region, is monitored by the Maricopa County Air Quality District (MCAQD). The MCAQD is “fully responsible for designing and operating the total air monitoring surveillance system and managing the pollutant data generated” (2020 Air Monitoring Network Review and 2021 Plan, 26). The MCAQD uses a Teledyne API – 400T ozone monitor to measure a continuous ozone concentration with respect to time (2020 Air Monitoring Network Review and 2021 Plan). This method is considered a federal equivalent method (FEM) by the EPA. However, these monitors require frequent calibration during which ozone concentrations are not recorded. Calibrations are staggered across stations and generally take no more than two hours, infrequently conflicting with the calculation of 8-hour ozone concentrations (Davis, personal communication, 09/19/2022). The MCAQD has monitored ozone concentrations for all 25 of the stations presented in figure 3.1. These stations were either active for the whole period or a partial period of the years 2010 through 2020.

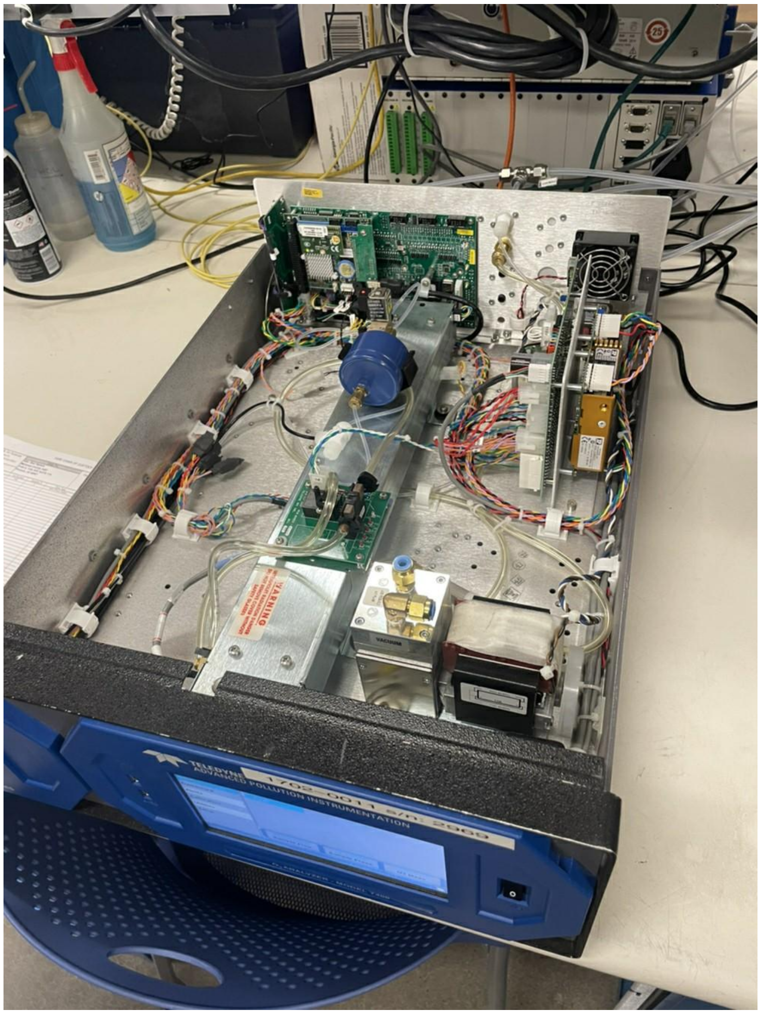

Figure 3.2. Teledyne API – 400T ozone monitor

Eight-hour average ozone data were obtained from the Environmental Protection Agency's Air Quality System (https://aqs.epa.gov/aqsweb/airdata/download_files.html). 8-hour average ozone data has a unique parameter code of 44201. Ozone data were downloaded individually by year, for the years 2010 through 2020. These data sets from the EPA include ozone observations across the United States, Puerto Rico, and the U. S. Virgin Islands, and are reported in these data sets in a unit of parts per million (ppm).

In order to reorganize ozone observations into a data set for Maricopa County, I used R and the corresponding libraries "tidyverse" and "lubridate." First I read the individual yearly files, filtered the data for Arizona's "State.Code" of "4" using the dplyr (included in tidyverse) function "filter," and removed the large original file to prevent cluttering of the workspace. Next, I created a column containing the unique station ID called “site_id.” The EPA’s file contains the state code, county code, and site number, however it is convenient to have one station identifier. I created a column called “dt_local” containing the date-time stamps for each observation. I used the dplyr functions “select” and “rename” to create a data set containing the date-time stamp “dt_local,” the station id “site_id,” and the 8-hour average ozone concentration “O3.” An example of this process for the year 2020 can be seen below. Finally, I used “rbind” to, by row, combine all eleven yearly data sets.

1. us_o3_20 <- read.csv("C://Users/treeo/OneDrive/Documents/R/Thesis/Ozone/8hour_44201_2020.csv", stringsAsFactors = FALSE)
2. AZ_O3_20 <- us_o3_20 %>% filter(State.Code == 4)
3. rm(us_o3_20)
4. AZ_O3_20$site_id <- AZ_O3_20$State.Code*10^7 + AZ_O3_20$County.Code*10^4 + AZ_O3_20$Site.Num
5. AZ_O3_20$dt_local <- format(as.POSIXct(paste(AZ_O3_20$Date.Local, AZ_O3_20$Time.Local), "%Y-%m-%d %H:%M"))
6. AZ_O3_20 <- AZ_O3_20 %>% select(site_id, dt_local, Mean.Excluding.All.Flagged.Data) %>% rename(O3 = Mean.Excluding.All.Flagged.Data)

These steps allowed me to then conveniently make separate columns for “Year,” “Month,” and “Day of Month” for each observation, as well as maximum daily 8-hour ozone concentrations for each station. I then simply took the maximum, maximum 8-hour ozone concentration across all stations for each day. Lastly, the EPA data set reports ozone in parts per million, however to stay consistent with the EPA’s regulation of ozone in parts per billion (ppb), I multiplied the ozone concentration column by 1000 to obtain the maximum 8-hour ozone concentrations in ppb.

3.2.2. Meteorological Variables

Meteorological data were observed by radiosondes launched by the Salt River Project (SRP) during the North American Monsoon season, from the Phoenix National Weather Service (NWS) office. In Arizona, only the Tucson and Flagstaff NWS offices are upper air stations. SRP launches radiosondes during the NAM season to improve forecasting deemed vital for their operations (Svoma, personal communication, 06/08/2022). The Phoenix NWS office is located at the SRP corporate office to the east of Phoenix SkyHarbor International Airport. As past research indicates that features within the planetary boundary layer may influence surface level ozone concentrations, such radiosonde observations are of interest for this study. However, the time period of which this study can be completed is limited by the availability of radiosonde data. SRP only launches radiosondes during the NAM season, and therefore my study is only taking place during the NAM season.

The completeness of radiosonde data for the 2010 through 2020 period during the NAM season is described by the following table (3.1), where the amount of observations per month are presented as the ratio of observations for the period of observations. The completeness of data is

1091/1115 or 97.8 %, calculated as the ratio of total observations to the sum of days within each respective yearly period of observations.

Table 3.1. Period of Phoenix radiosonde observations used in this study by year, and the completeness of observations by month and year for their respective periods.

Year	June	July	August	September	Period	Total Observations
2010	9/10	30/31	31/31	20/22	06/21 - 09/22	90/94 = 95.7 %
2011	18/18	30/31	31/31	23/23	06/13 - 09/23	102/103 = 99.0 %
2012	17/18	29/31	29/31	21/22	06/13 - 09/22	96/101 = 95.0 %
2013	18/18	30/31	31/31	22/22	06/13 - 09/22	101/102 = 99.0 %
2014	22/22	30/31	31/31	21/21	06/09 - 09/21	104/105 = 99.0 %
2015	16/19	31/31	31/31	25/25	06/12 - 09/25	103/106 = 97.2 %
2016	16/18	31/31	30/31	21/22	06/13 - 09/22	98/102 = 96.1 %
2017	16/17	30/31	31/31	21/21	06/14 - 09/21	98/100 = 98.0 %
2018	18/18	31/31	31/31	21/21	06/13 - 09/21	101/101 = 100 %
2019	14/14	30/31	31/31	26/27	06/17 - 09/27	101/103 = 98.1 %
2020	15/16	31/31	31/31	20/20	06/15 - 09/20	97/98 = 99.0 %

All radiosonde data were downloaded from the NOAA/ESRL Radiosonde Database https://ruc.noaa.gov/raobs/) in original FSL format. Original FSL format includes seven columns and several rows with corresponding unique linetypes, or indicators of row types.

---COLUMN NUMBER---

1	2	3	4	5	6	7
LINTYP						
			header lines			
254	HOUR	DAY	MONTH	YEAR	(blank)	(blank)
1	WBAN#	WMO#	LAT **D**	LON **D**	ELEV	RTIME
2	HYDRO	MXWD	TROPL	LINES	TINDEX	SOURCE
3	(blank)	STAID	(blank)	(blank)	SONDE	WSUNITS
			data lines			
9	PRESSURE	HEIGHT	TEMP	DEWPT	WIND DIR	WIND SPD
4						
5						
6						
7						
8						

Figure 3.1. Table of original FSL output format for radiosonde observations. Accessed from https://ruc.noaa.gov/raobs/fsl_format-new.html on 10/13/2022.

```
                    LEGEND

LINTYP: type of identification line
        254 = indicates a new sounding in the output file
          1 = station identification line
          2 = sounding checks line
          3 = station identifier and other indicators line
          4 = mandatory level
          5 = significant level
          6 = wind level (PPBB) (GTS or merged data)
          7 = tropopause level (GTS or merged data)
          8 = maximum wind level (GTS or merged data)
          9 = surface level
```

Figure 3.2. Legend of original FSL output format "linetypes" for radiosonde observations. Accessed from https://ruc.noaa.gov/raobs/fsl_format-new.html on 10/13/2022.

Observational radiosonde data at 1200 UTC for all levels were downloaded in "Original FSL Format" for years 2010 through 2020, using the corresponding station ID of 74626 as a "station series sort", where wind is measured in "tenths of meters per second." The station identifier, 74626, is the World Meteorological Organization station identifier number for the Phoenix radiosonde location.

Data were first converted into a .CSV format through Excel. I used the convert list from WEB, split by digit to non-digit option to convert the list of radiosonde data into an Excel spreadsheet. I then saved the file as a .CSV and read the table into MATLAB for further analysis. Excel added excess columns from the original data set during the conversion process due to periods and words that were contained within the station latitude and longitude information. Excel has also been found to be inconsistent in the number of additional columns generated during its' conversion to a .CSV file. For further analysis of the data in MATLAB, one must both delete the excess columns created by Excel and use the “table2array” function to convert the data from a table data type to an array data type. An example of how this can be done is as follows, where I both saved and converted the column (of index 1) containing the row types into a new matrix that I titled RAW (of column index 1):

```
RAW=[]; % an empty matrix
RAW(:,1)=table2array(EXCEL(:,1)); % Change of row types
```

I reorganized the columns and rows so that I had a continuous data set of columns titled “Day of Month”, “Month”, “Year”, “Time of Launch”, “Temperature”, “Pressure”, “Dew Point Temperature”, “Wind Speed”, “Wind Direction”, and “Height.” This process also included the removal of rows (observations at pressure levels) where temperature was not observed. The units of these variables were of the same initial units listed in the original FSL format. Code for this step is included in the appendix.

Table 3.2. Table of original FSL output format units for radiosonde observations.

Meteorological Variable	Unit
Pressure	Whole millibars
Temperature	Tenths of degrees Celsius
Dew Point Temperature	Tenths of degrees Celsius
Wind Speed	Tenths of meters per second
Wind Direction	Degrees from North
Height	meters

For the column of row identifiers, an identifier of “254” indicates a new sounding in the output file. An identifier of ”9” indicates a row of surface level observations. An identifier of “1”, “2”, and “3” indicates station metadata. All other identifiers contained rows of radiosonde observations. Observations are organized initially from the original FSL format in decreasing order for increasing observation height. I then changed the units from the initial original FSL format units to the SI units listed in the table above. The following code was used to convert matrix RAW2 in original FSL format to matrix RAW3 in SI units:

```
a=1; % index to be counted by row
RAW3=[]; % an empty matrix
while a <= length(RAW2)
   if RAW2(a,1) == 254 % 254 is the row type indicating a new sounding
      RAW3(a,1)=RAW2(a,1);
      a=a+1;
   else
      RAW3(a,1)=RAW2(a,1);
      RAW3(a,2)=RAW2(a,2); % Pressure in mb
      RAW3(a,3)=RAW2(a,3); % Height in m
      RAW3(a,4)=RAW2(a,4)/10; % Temperature in degrees Celsius
      RAW3(a,5)=RAW2(a,5)/10; % Dew Point Temperature in degrees Celsius
      RAW3(a,6)=cosd(RAW2(a,6)-90); % U as a unit vector
      RAW3(a,7)=-sind(RAW2(a,6)-90); % V as a unit vector
      RAW3(a,8)=RAW2(a,7)/10; % Wind Speed in m/s
      RAW3(a,9)=RAW2(a,8); % Time of launch
      RAW3(a,10)=RAW2(a,9); % Day of Month
      RAW3(a,11)=RAW2(a,10); % Month
      RAW3(a,12)=RAW2(a,11); % Year
      a=a+1;
   end
end
```

Using the row identifiers, I extracted the surface level conditions and date time stamps for all observations, inserting them into a final matrix called “MET.” Within this new matrix, mixing level height, mixing level pressure, and inversion temperature, were also inserted by their corresponding date time stamps, and were calculated in the following algorithms:

1. Calculate the change in temperature with respect to height for each pressure level.

```
a=1;
while a <= length(RAW3)
   if RAW3(a,1) == 254
      b=a+1;
      while RAW3(b,1) < 254 && b < length(RAW3)
         c=b+1;
         RAW3(b,13)=(RAW3(c,4)-RAW3(b,4))/(RAW3(c,3)-RAW3(b,3)); % dT/dH
         b=b+1;
      end
      a=a+1;
   else
      a=a+1;
   end
end
```

2. Save the first pressure levels, heights, and temperatures at which an inversion occurred. This saves the surface pressure, height, and temperature when no inversion is present.

```
a=1;
c=1;
while a <= length(RAW3)
   if RAW3(a,1) == 254
      b=a+1;
      d=1; % Condition to activate while loop
      while RAW3(b,1) < 254 && b < length(RAW3) && d == 1
         if RAW3(b,13) > 0
          b=b+1;
         else
            MET(c,12)=RAW3(b,2); % Mixing level pressure in mb
            MET(c,13)=RAW3(b,3); % Mixing level height in m
            MET(c,14)=RAW3(b,4); % Inversion temperature in degrees Celsius
            c=c+1;
            d=2; % Condition to end while loop
         end
      end
      a=a+1;
   else
      a=a+1;
   end
end
```

Meteorological data were then exported from MATLAB using the “writematrix” function. The included MATLAB code in the appendix can be used for any set of radiosonde data, if it is in the same initial format as the data used in my study. As previously stated, Excel has also been found to be inconsistent in the number of additional columns generated during its conversion to a .CSV file. My code also includes the creation of a column for surface level height. Surface level height is a constant value for the Phoenix radiosonde launch site of 384 meters. As surface level ozone concentrations can never change with respect to changes in surface level height, being a constant, it is excluded from table 3.3 and not statistically analyzed.

Table 3.3. Phoenix radiosonde meteorological variables and their corresponding units used in this study.

Meteorological Variable	Unit
Surface Level Pressure	Millibars (mb)
Surface Level Height	Meters (m)
Surface Level Temperature	Degrees Celsius (°C)
Surface Level Dew Point Temperature	Degrees Celsius (°C)
Surface Level Wind Direction “u”	Unit Vector Component
Surface Level Wind Direction “v”	Unit Vector Component
Surface Level Wind Speed	Meters per Second (m/s)
Mixing Level Pressure	Millibars (mb)
Mixing Level Height	Meters (m)
Inversion Level Temperature	Degrees Celsius (°C)

Each row within the final meteorological data set contained a column for each variable listed in table 3.3, in the corresponding units listed in table 3.3, and four corresponding columns containing the date-time stamp. The four date-time stamp columns are “Year,” “Month,” “Day of Month,” and “Launch Time.” Radiosondes are rarely launched at the same time. The 1200 UTC radiosonde observations used in this study actually took place from 1128 UTC to 1235 UTC with an average launch time of 1138 UTC and a standard deviation of 14.0 minutes.

3.4 Exceedances and Combined Data Set

I merged ozone and meteorological data together in the statistical programming language R by “Year”, “Month”, and “Day of Month.” As some statistical analysis requires complete sets of data for each date-time stamp, the rows of data that were not complete were removed for ease of use of the data set. The complete data set followed the same periods as shown by table 3.1. There were maximum 8-hour ozone concentrations in Maricopa County for all days that there were radiosonde observations.

As shown in table 3.1's column 'exceedances,' I categorized the data into two classes, where a category of 'one' (1) signifies that the day exceeded the EPA regulatory 8-hour average ozone concentration value of 70 ppb, while a category of 'zero' (0) signifies that the day was not an exceedance. Table 3.4 shows a sample of the data set, with all the variables used in this study and their corresponding units.

Table 3.4. Sample data set of variables used in this study with corresponding date-time stamps.

Variable	Unit	Sample Data		
Year	—	2010	2010	2010
Month	—	6	6	6
Day of Month	—	21	22	23
Time of Launch	UTC	1145	1131	1134
Ozone Concentration	ppb	63	81	85
Exceedance	1=true, 0=false	0	1	1
Surface Level Pressure	mb	967	968	965
Surface Level Height	m	384	384	384
Surface Level Temperature	°C	22.6	22.8	23
Surface Level Dew Point Temperature	°C	5.6	1.7	-1
Surface Level Wind Direction "u"	Unit component	1	0.985	0.866
Surface Level Wind Direction "v"	Unit component	0	-0.174	-0.5
Surface Level Wind Speed	m/s	3.6	2.5	1.5
Mixing Level Pressure	mb	941	955	953
Mixing Level Height	m	623	502	494
Inversion Level Temperature	°C	27.6	26.6	30

3.3 Summary

My research question involves the analysis of ozone and boundary layer meteorological data (e.g., temperature, dew point, winds). This section has detailed the manipulation and compilation of those data into a single merged data set that addresses specific meteorological variables, e.g., surface temperature, inversion pressure, etc., with the maximum, maximum daily 8-hour average ozone concentration across all the Phoenix Metropolitan monitoring stations for the summer monsoon months from year 2010 through 2020, and a corresponding classification to specify ozone exceedance (> 70 ppb) days. With this data set constructed, it is possible to analyze it to determine what statistical relationship exists between ozone and antecedent monsoonal weather variables.

4. Analysis and Discussion of Results

4.1 Introduction

I have created a data set for days within the Phoenix Metropolitan region's monsoon to determine if statistical relationships exist between a set of boundary layer meteorology variables and surface level ozone. This data set, discussed in the last section, includes a corresponding classification to specify days that exceeded the EPA regulatory 8-hour average ozone concentration value of 70 ppb (United States Environmental Protection Agency, 2022).

To analyze these data, I first must determine the normality in the distribution of each variable to determine which tests are appropriate to use for each variable. This will be done by calculating the mean, standard deviation, skewness, and kurtosis of each variable. I will then calculate the Pearson product moment and Spearman rank correlation coefficients.

The Pearson product moment analysis assumes a linear correlation between two variables that are normally distributed, providing a coefficient between negative one (-1) and positive one (1) for the strength and direction of the linear relationship between both variables (Witte and Witte, 2017). If the coefficient is close to negative one (-1): there is a strong negative linear relationship, positive one (1): there is a strong positive linear relationship, close to zero: there is a weak linear relationship (Bishara, 2015).

Not all variables in this study are expected to be normally distributed nor linearly related to ozone concentrations, and therefore the Spearman rank correlation analysis coefficients will also be calculated. The Spearman rank correlation analysis is a nonparametric form of the Pearson product moment, determining both the strength and direction of the relationship between two variables (Witte and Witte, 2017). Similarly, if the coefficient is close to negative one (-1): there is a strong negative relationship, positive one (1): there is a strong positive relationship, close to zero: there is a weak relationship (Bishara, 2015). For this study, a confidence level of 0.95 will be used for both correlation coefficients in determining statistical significance.

To determine the ability of the meteorological variables to distinguish group variability, I will calculate the Student's two-sided t-test statistic to determine if there is a difference in mean value of variables on ozone exceedance and non-exceedance days. A Student's t-test is a statistical test that gauges the ability for independent variables to discriminate between classes of data, and can be used for non-normally distributed data (Devore, 2011).

The classification of exceedance days included in the data set will be used as the data class, and I will assume a confidence level of 0.95 for these tests. The null hypothesis for each t-test is that there is no difference between the means of corresponding variables on exceedance and non-exceedance days. The alternative hypothesis is that there is a difference between the means of corresponding variables on exceedance and non-exceedance days. This test will establish if there is a significant statistical difference in the means of variables on exceedance and non-exceedance days. If the calculated t-value is outside of the 95% confidence interval, and the p-value is less than or equal to 0.05, I can say with 95% confidence that there is a significant statistical difference in the means of variables on exceedance and non-exceedance days. If this is the case, I can reject the null hypothesis that there is no difference between the means of corresponding variables on exceedance and non-exceedance days. If the t-value is less (greater) than the 95% confidence interval then, I will also be able to say that the mean value for exceedance days is greater (less) than on non-exceedance days. This also would mean that values within one standard error ± of the sample mean between the two sample means do not overlap. This will be visually represented in figures with error bars plotted from the corresponding sample means.

4.2 Influence of Each Variable on Ozone

4.2.1 Ozone

Maximum, maximum daily 8-hour ozone concentrations in the Phoenix Metropolitan region during NAM season for years 2010 through 2020 had a mean value of 65.2 ppb with a standard deviation of 15.3 ppb. It is noteworthy to mention that the threshold of an ozone exceedance day is within one standard deviation of the mean concentration. The skewness value for ozone is 2.56, and the kurtosis value is 14.9. The ozone concentration distribution is right skewed, producing more outliers than in a normal distribution. While this may be problematic for linear correlation analysis, aggregated group analyses (e.g., Student's t-test analyses) should not be affected.

4.2.2 Surface Level Pressure

The 1200 UTC surface level pressure in the Phoenix Metropolitan region during NAM season for years 2010 through 2020 had a mean value of 966 mb with a standard deviation of 2.37 mb. The skewness value for surface level pressure is -0.0786, and the kurtosis value is 3.19.

The surface level pressure distribution is approximately normally distributed, and therefore both a Pearson product moment and a Spearman rank correlation coefficient are calculated to describe surface level pressure's correlation to ozone.

The Pearson product moment coefficient between surface level pressure and ozone is -0.13 with a p-value of $1x10^{-5}$. This indicates a statistically significant, but weak, negative linear correlation between surface level pressure and ozone. The Spearman rank correlation coefficient between surface level pressure and ozone is -0.00771 with a p-value of 0.799. This p-value suggests that the monotonic correlation between surface level pressure and ozone is not statistically important.

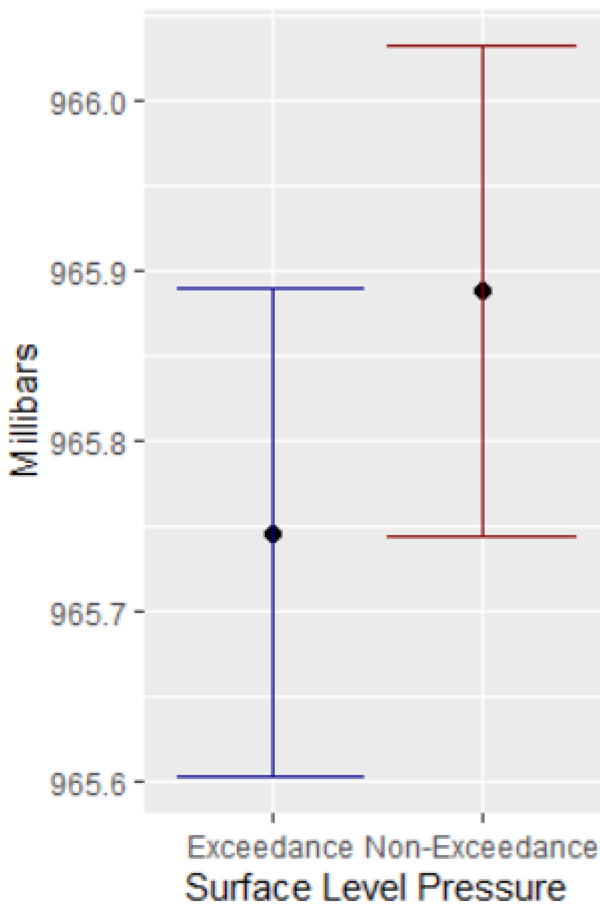

Figure 4.1. Mean surface level pressure in millibars plotted with standard error bars for exceedance days (blue) and non-exceedance days (red).

When I apply aggregate group analysis using a Student's two-pair t-test, I find that, as the t-value is within the 95% confidence interval, I cannot reject the null hypothesis that there is no difference between the mean surface level pressure on exceedance days (965.746 mb) and non-exceedance days (965.888 mb). There is overlap within one standard error of the means between exceedance and non-exceedance days. Monsoonal surface level pressures at 5:00 am MST in Phoenix do not have the ability to discriminate between ozone exceedance and non-exceedances, though lower surface pressure is associated with exceedances.

This finding is interesting, as a higher surface level pressure associated with anticyclonic rotation during the NAM has been previously correlated to higher ozone concentrations in the free troposphere over the Southwestern U.S. (Granados-Muñoz et al., 2017). Leffel (2022) similarly found a negative Pearson product moment coefficient between surface level pressure and ozone in the Phoenix Metropolitan region. However, Leffel found surface level pressure to be a statistically significant discriminator between ozone exceedance and non-exceedances when using average daily surface level pressure, where lower surface level pressure was likewise associated with exceedances.

4.2.3 Surface Level Temperature

The 1200 UTC surface level temperature in the Phoenix Metropolitan region during NAM season for years 2010 through 2020 had a mean value of 28.2 ℃ with a standard deviation of 2.81 ℃. The skewness value for surface level temperature is -0.264, and the kurtosis value is 2.63. The surface level temperature distribution is approximately normally distributed, and therefore both a Pearson product moment and a Spearman rank correlation coefficient are calculated to describe surface level temperature's correlation to ozone.

The Pearson product moment coefficient between surface level temperature and ozone is 0.223 with a p-value of $1x10^{-5}$. This indicates a statistically significant, weak positive and linear correlation between surface level temperature and ozone. The Spearman rank correlation coefficient between surface level temperature and ozone is 0.192 with a p-value of $1x10^{-10}$. This also indicates a statistically significant, but weak, positive correlation between surface level temperature and ozone. Both correlation coefficients are statistically weak, as opposed to the strong positive correlation that Wise and Comrie (2005) found for the Southwest. Although statistically weak, this positive relationship makes sense when considering the photolytic cycle, assuming higher temperatures correlate to higher levels of ultraviolet radiation. The linear relationship found here is interesting, as reaction rates for ozone formation with respect to temperature under various precursor concentrations have previously been found to behave nonlinearly (Khamaganov and Hites, 2001).

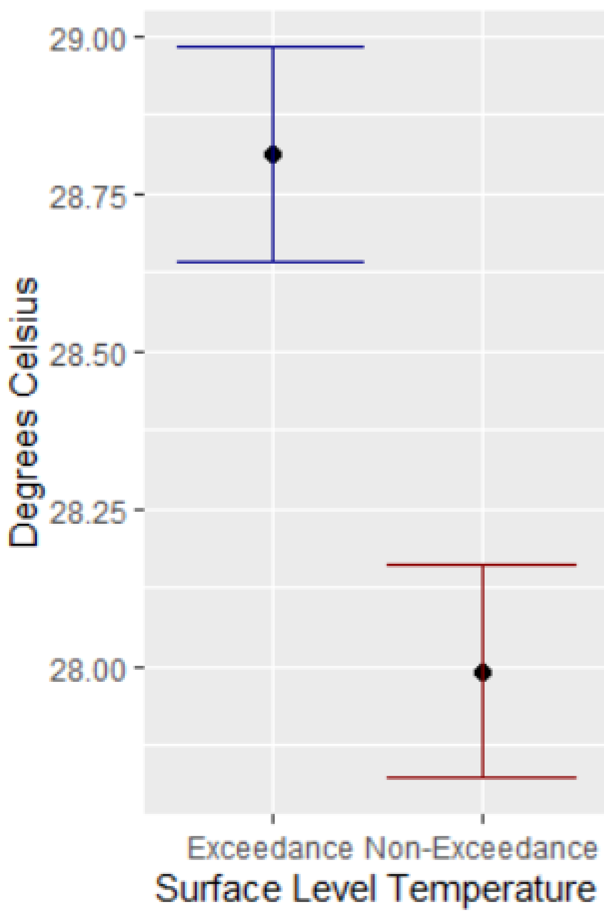

Figure 4.2. Mean surface level temperature in degrees Celsius plotted with standard error bars for exceedance days (blue) and non-exceedance days (red).

As the t-value is less than the 95% confidence interval, I reject the null hypothesis that there is no difference between the mean surface level temperature on exceedance days (28.81 ℃) and non-exceedance days (27.99 ℃). As the p-value of $1x10^{-5}$ is less than my confidence limit of 0.05, there is a clear and significant statistical difference in the mean surface level temperature on ozone exceedance and non-exceedance days. The mean surface level temperature for exceedance days is greater than on non-exceedance days. There is no overlap within one standard error of the means between exceedance and non-exceedance days, and therefore monsoonal surface level temperatures at 5:00 am MST in Phoenix have the ability to discriminate between ozone exceedance and non-exceedances. These findings are similar to Leffel's (2022) findings that daily minimum temperatures are a statistically significant discriminator between ozone exceedance and non-exceedances in the Phoenix Metropolitan region, where higher level surface temperature was likewise associated with exceedances (assuming 5:00 am MST surface level temperatures are a good indicator of daily minimum surface level temperatures). It is also important to remember that the temperature observations used here are near-surface, with radiosondes being launched from the roof of a four-story building.

4.2.4 Surface Level Dew Point Temperature

The 1200 UTC surface level dew point temperature in the Phoenix Metropolitan region during NAM season for years 2010 through 2020 had a mean value of 12.8 ℃ with a standard deviation of 6.07 ℃. The skewness value for surface level temperature is -0.624, and the kurtosis value is 2.47. The surface level temperature distribution is slightly left skewed, producing fewer and less extreme outliers than a normal distribution. Therefore, a Pearson product moment will not be calculated, and it cannot be determined that correlations with surface level dew point temperature are linear.

The Spearman rank correlation coefficient between surface level dew point temperature and ozone is -0.0553 with a p-value of 0.0679. Contrary to Wise and Comrie's (2005) study in the Southwest, this p-value suggests that the monotonic correlation between surface level temperature and ozone is not statistically important at the 95% confidence level. Though, this monotonic correlation is almost significant. However, a relationship may be possible to extract through discriminant analysis.

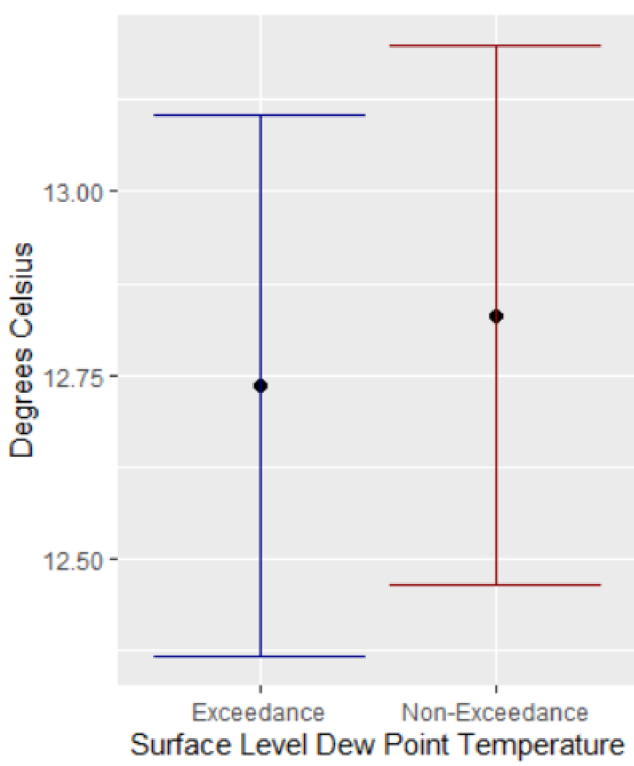

Figure 4.3. Mean surface level dew point temperature in degrees Celsius plotted with standard error bars for exceedance days (blue) and non-exceedance days (red).

When I apply a Student's t-test to the grouped (exceedance/non-exceedance) data, I find that as the t-value is within the 95% confidence interval, I cannot reject the null hypothesis that there is no difference between the mean surface level dew point temperature on exceedance days (12.736 ℃) and non-exceedance days (12.831 ℃). Similar to Leffel's (2022) findings for average daily dew point temperatures, there is overlap within one standard error of the means between exceedance and non-exceedance days while the mean ozone concentration for ozone exceedances is less than that of non-exceedances. Monsoonal surface level dew point temperatures at 5:00 am MST in Phoenix do not have the ability to discriminate between ozone exceedance and non-exceedances. Again, it is important to remember that surface observations used by Leffel are at 2m height, while the radiosonde-derived surface observations used here are only near-surface observations.

4.2.5 Surface Level Wind Speed

The 1200 UTC surface level wind speed in the Phoenix Metropolitan region during NAM season for years 2010 through 2020 had a mean value of 2.22 m/s with a standard deviation of 1.55 m/s. The skewness value for surface level wind speed is 2.58, and the kurtosis value is 23.0. The surface level wind speed distribution is right skewed, producing more outliers than in a normal distribution. Only a Spearman rank correlation coefficient will be calculated as linear correlations cannot be determined for this wind speed data set.

The Spearman rank correlation coefficient between surface level wind speed and ozone is -0.0395 with a p-value of 0.192. This p-value suggests that the monotonic correlation between surface level wind speed and ozone is not statistically important at the 95% confidence level.

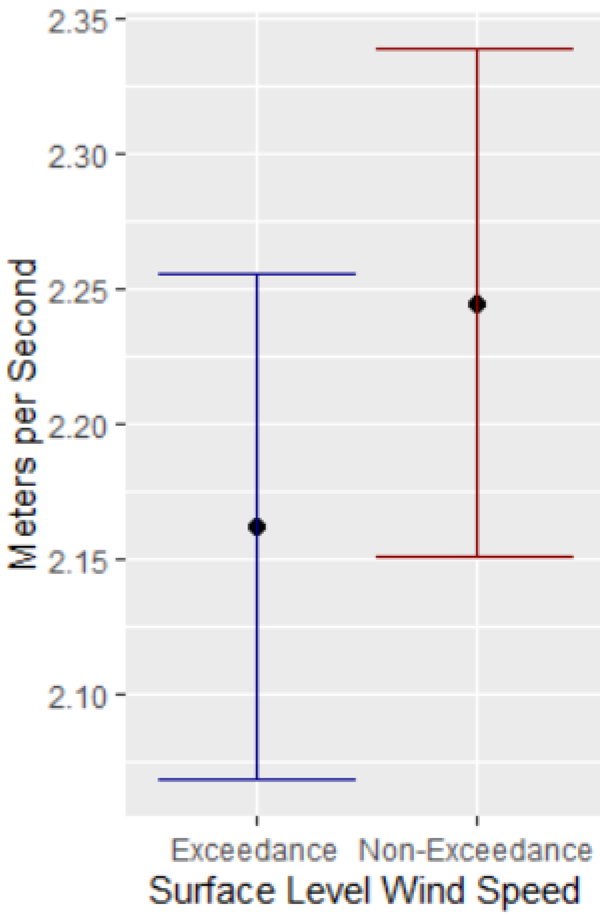

Figure 4.4. Mean surface level wind speed in meters per second plotted with standard error bars for exceedance days (blue) and non-exceedance days (red).

The t-value is within the 95% confidence interval, and I therefore cannot reject the null hypothesis that there is no difference between the mean surface level wind speed on exceedance days (2.162 m/s) and non-exceedance days (2.245 m/s). Monsoonal surface level wind speeds at 5:00 am MST in Phoenix do not have the ability to discriminate between ozone exceedance and non-exceedances as there is overlap within one standard error of the means between exceedance and non-exceedance days.

These findings are interesting as positive correlations between ozone concentrations and wind speeds, indicating both local and regional transport of ozone, have previously been found in the Southwest (Li et al., 2015; Granados-Muñoz et al., 2017; Wise and Comrie, 2015). Leffel (2022) in fact found a highly significant statistical difference in the means of average daily wind speeds between exceedances and non-exceedances.

4.2.6 Surface Level Zonal and Meridional Wind Directions

The 1200 UTC surface level zonal (east-west) wind direction in the Phoenix Metropolitan region during NAM season for years 2010 through 2020 had a mean unit vector value of 0.356, a standard deviation of 0.726, a skewness value of -0.763, and a kurtosis value of 1.98. On the other hand, the meridional (north-south) wind direction had a mean unit vector value of 0.0348, a standard deviation of 0.587, a skewness value of 0.0977, and a kurtosis value of 2.015. Near surface winds at the 5:00 am local time tend to flow zonally from the West, with fewer and less extreme outliers than a normal distribution.

The Spearman rank correlation coefficient between surface level zonal/meridional wind direction and ozone is 0.0623/0.0313 with a p-value of 0.0397/0.302. This indicates a statistically significant, but very weak, positive correlation between surface level zonal wind direction and ozone, with a monotonic correlation between surface level meridional wind direction and ozone that is not statistically important. These correlations suggest that some level of ozone is transported in the direction of the winds, at the near surface height.

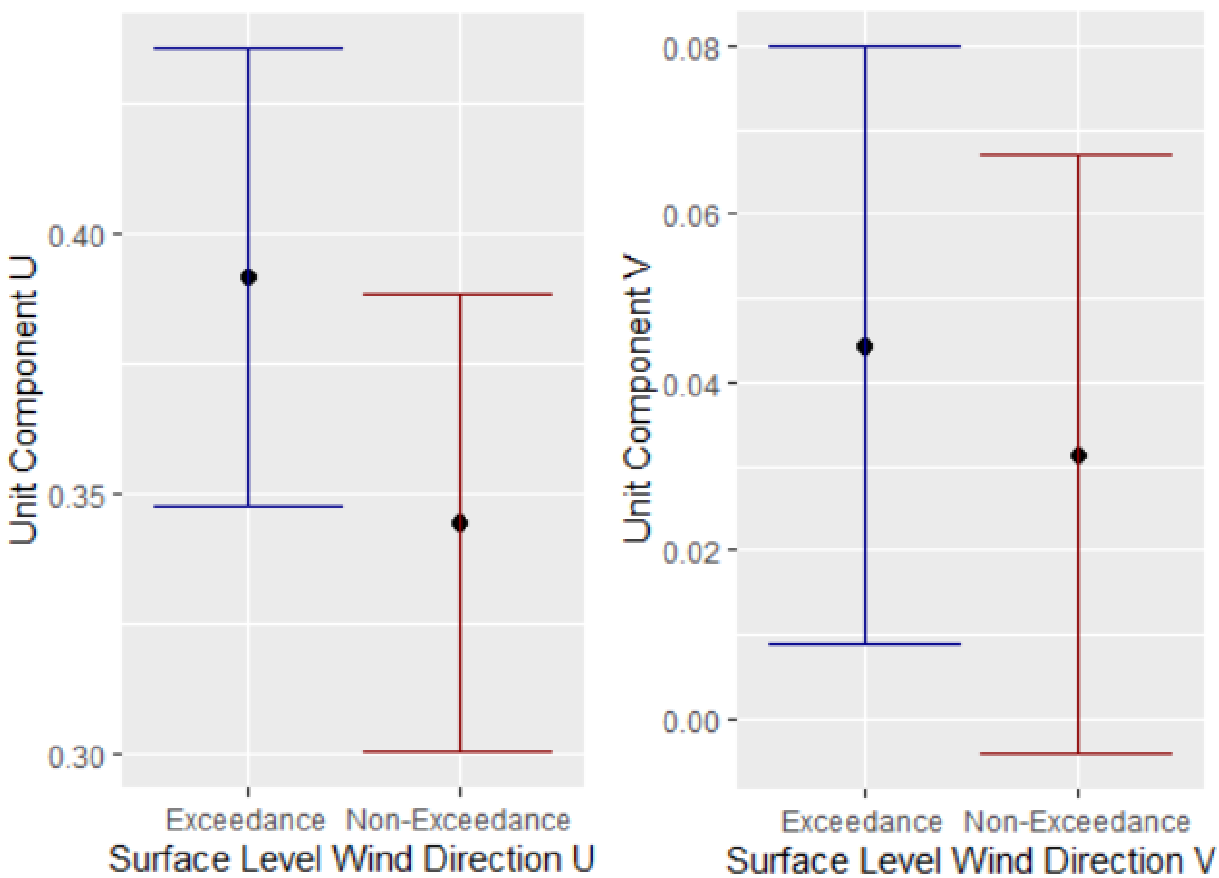

Figure 4.5. Mean surface level zonal (left) and meridional (right) wind direction components plotted with standard error bars for exceedance days (blue) and non-exceedance days (red).

There is overlap within one standard error of the means between exceedance and non-exceedance days for both zonal and meridional winds. I cannot reject the null hypothesis that there is no difference between the mean surface level zonal or meridional wind direction unit components on exceedance days (0.3913/0.0443) and non-exceedance days (0.3441/0.3441), as the t-value is within the 95% confidence interval.

The combined zonal and meridional wind direction results are interesting. Neither monsoonal wind direction measured at 5:00 am MST in Phoenix has the ability to discriminate between ozone exceedance and non-exceedances. Leffel (2022), who examined wind components (wind speed in their vector components) as opposed to unit components (just wind direction), found that only average daily zonal winds can statistically discriminate between ozone exceedances and non-exceedances in the Phoenix Metropolitan region. It is unknown if Leffel's directional findings are controlled by wind speed, which were also found to be significant discriminators.

4.2.7 Mixing Level Pressure

The 1200 UTC mixing level pressure in the Phoenix Metropolitan region during NAM season for years 2010 through 2020 had a mean value of 956 mb with a standard deviation of 11.5 mb. With a skewness value of -0.968 and a kurtosis value is 3.37, the mixing level pressure distribution is left skewed, producing about as many outliers as in a normal distribution. Not being normally distributed, linearity cannot be determined, and only the Spearman rank correlation coefficient is calculated.

The Spearman rank correlation coefficient between mixing level pressure and ozone is -0.146 with a p-value of $1x10^{-6}$. This indicates a highly statistically significant, but weak negative correlation between mixing level pressure and ozone. Surface level ozone concentrations decrease as mixing level pressure increases. This negative correlation suggests that Oke's (1978) description of fumigation dispersion acting in tandem with an inversion, strengthens with a decrease in thickness between the surface and mixing level pressure. This is also supported by the -0.96 Spearman rank correlation coefficient found between mixing level pressure and mixing level height.

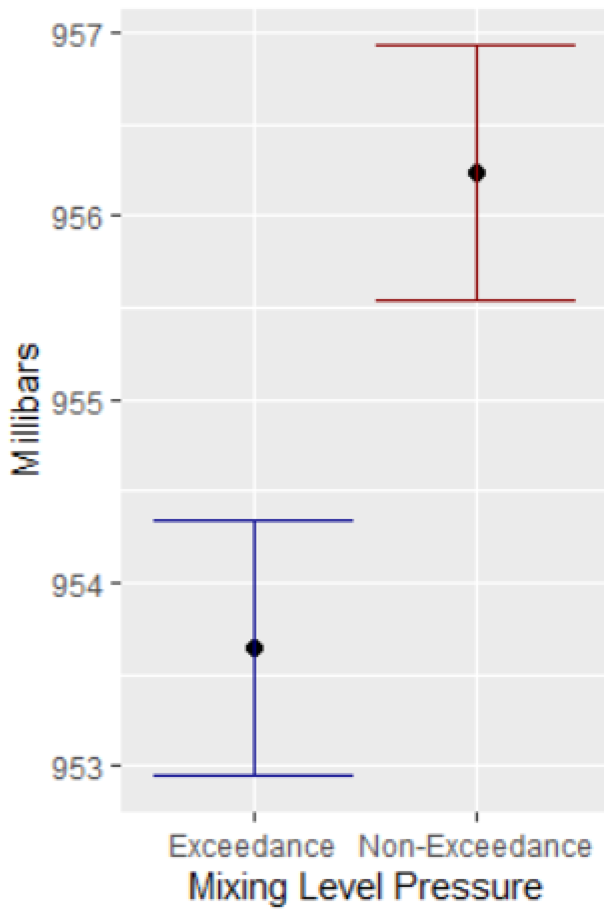

Figure 4.7. Mean mixing level pressure in millibars plotted with standard error bars for exceedance days (blue) and non-exceedance days (red).

As the t-value is less than the 95% confidence interval, I reject the null hypothesis that there is no difference between the mean mixing level pressure on exceedance days (953.64 mb) and non-exceedance days (956.23 mb). As the p-value of 0.002 is less than 0.05, there is a significant statistical difference in the mean mixing level pressure on exceedance and non-exceedance days at the 95% confidence level. The mean mixing level pressure for exceedance days is less than on non-exceedance days. There is no overlap within one standard error of the means between exceedance and non-exceedance days.

4.2.8 Mixing Level Height

The 1200 UTC mixing level height in the Phoenix Metropolitan region during NAM season for years 2010 through 2020 had a mean value of 479 m with a standard deviation of 101 m. The skewness value for mixing level height is 1.14, and the kurtosis value is 3.67. The mixing level height distribution is right skewed, producing slightly more outliers than in a normal distribution.

The Spearman rank correlation coefficient between mixing level height and ozone is 0.168 with a p-value of $1x10^{-8}$. This indicates a highly statistically significant, but weak positive correlation between mixing level height and ozone. Just like mixing level pressure, this negative correlation suggests that Oke's (1978) description of fumigation dispersion acting in tandem with an inversion, strengthens with a decrease in thickness between the surface and mixing level pressure.

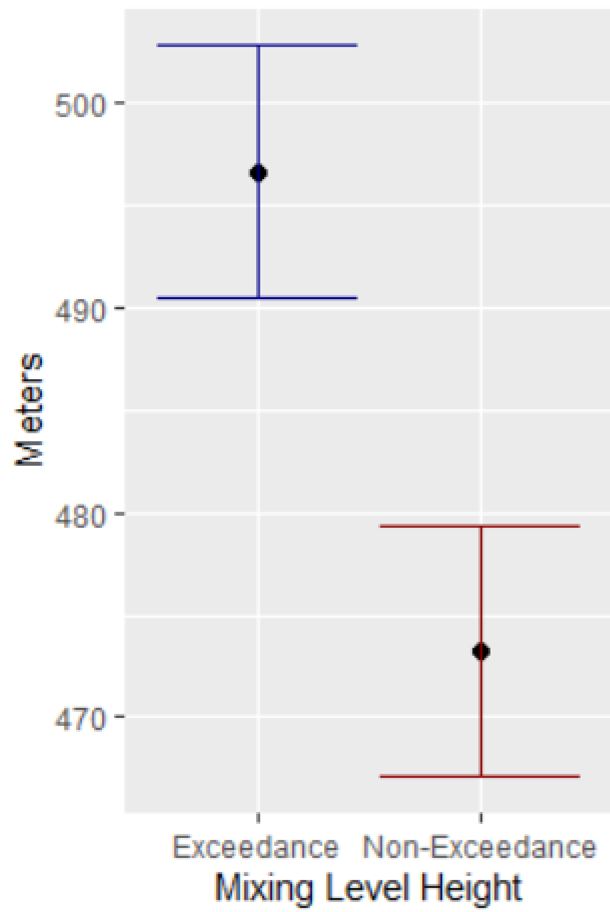

Figure 4.8. Mean mixing level height in meters plotted with standard error bars for exceedance days (blue) and non-exceedance days (red).

As the t-value is less than the 95% confidence interval, I reject the null hypothesis that there is no difference between the mean mixing level height on exceedance days (496.7 m) and non-exceedance days (473.2 m). As the p-value of 0.001 is less than 0.05, there is a significant statistical difference in the mean mixing level height on exceedance and non-exceedance days at the 95% confidence level. The mean mixing level height for exceedance days is greater than on non-exceedance days. There is no overlap within one standard error of the means between exceedance and non-exceedance days.

4.2.9 Inversion Temperature

The 1200 UTC inversion temperature in the Phoenix Metropolitan region during NAM season for years 2010 through 2020 had a mean value of 2.85 ℃ with a standard deviation of 10.3 ℃. The skewness value for inversion temperature is -0.285, and the kurtosis value is 3.04. The inversion temperature distribution is approximately normally distributed, and therefore both a Pearson product moment and a Spearman rank correlation coefficient are calculated to describe surface level pressure's correlation to ozone.

The Pearson product moment coefficient between the inversion temperature and ozone is 0.325 with a p-value of $1x10^{-16}$. This indicates a highly statistically significant, but weak positive linear correlation between the inversion temperature and ozone. The Spearman rank correlation coefficient between the inversion temperature and ozone is 0.282 with a p-value of $1x10^{-16}$. This also indicates a highly statistically significant, but weak positive correlation between the inversion temperature and ozone.

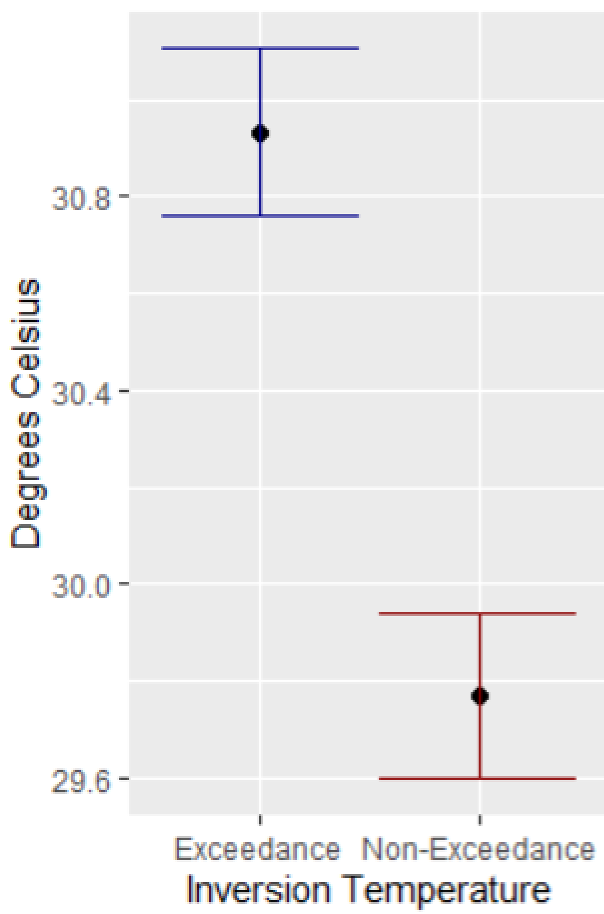

Figure 4.9. Mean inversion temperature in degrees Celsius plotted with standard error bars for exceedance days (blue) and non-exceedance days (red).

As the t-value is less than the 95% confidence interval, I reject the null hypothesis that there is no difference between the mean inversion temperature on exceedance days (30.9 ℃) and non-exceedance days (29.8 ℃). As the p-value of $1x10^{-9}$ is less than 0.05, there is a significant statistical difference in the mean inversion temperature on exceedance and non-exceedance days. The mean inversion temperature for exceedance days is greater than on non-exceedance days. There is no overlap within one standard error of the means between exceedance and non-exceedance days.

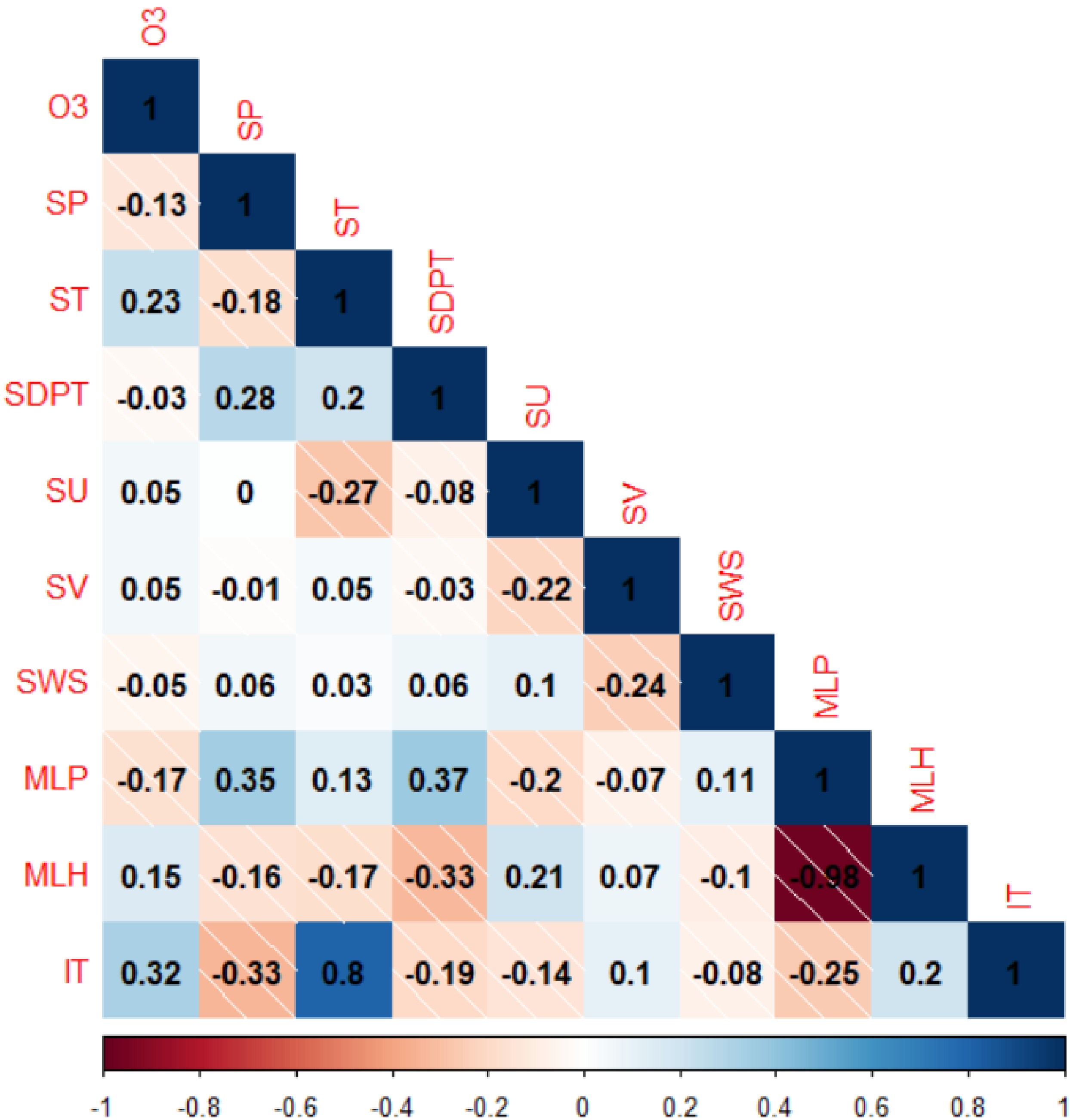

Figure 4.10. Correlation plot of Pearson product moment coefficients.

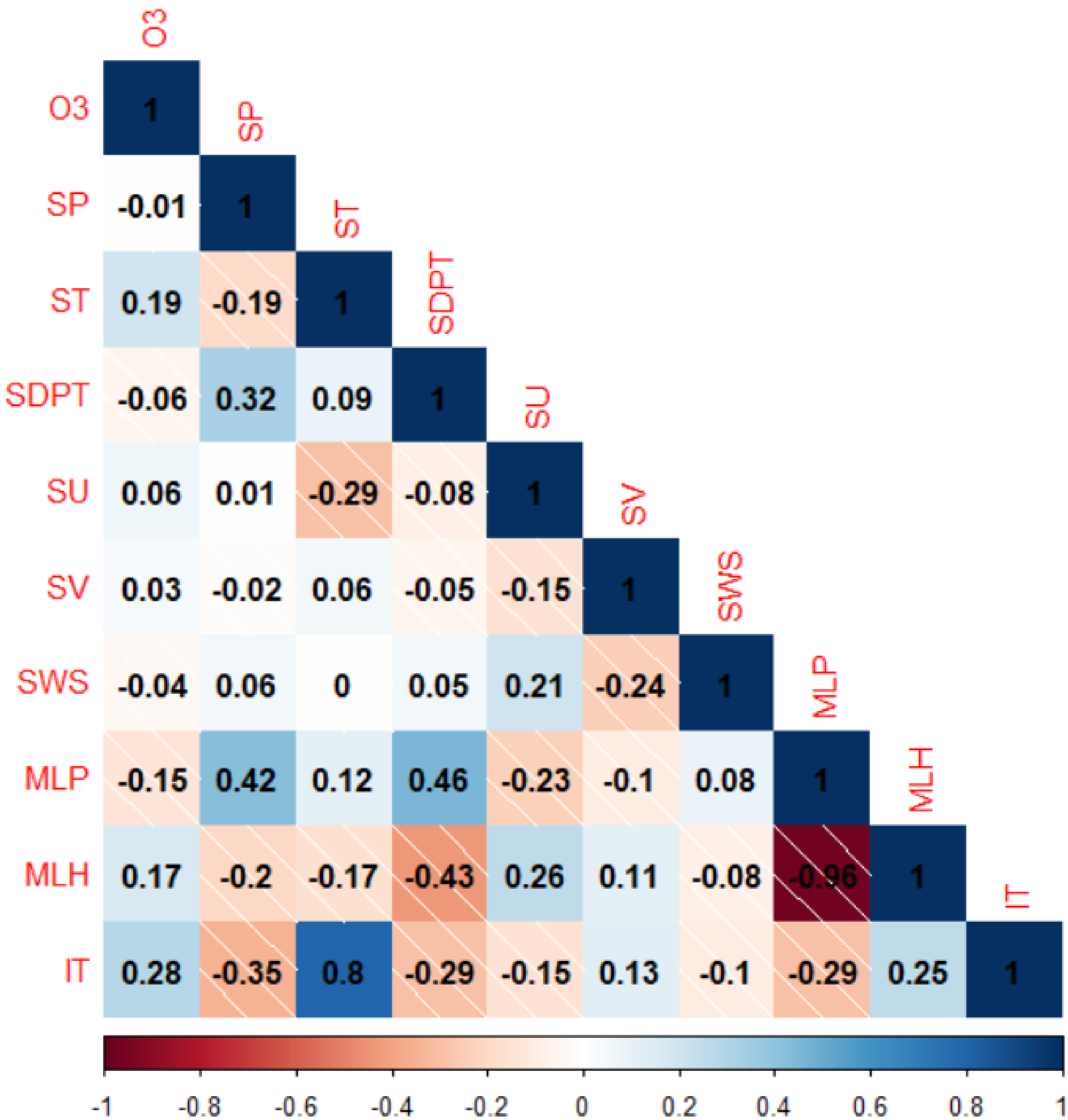

Figure 4.11. Correlation plot of Spearman rank correlation coefficients.

4.3 Summary

Statistical analyses indicate that the three upper-air, boundary layer meteorology variables (mixing level pressure, mixing level height, and inversion temperature), are the most highly significant, and all extremely significantly correlated to daily maximum surface level ozone concentrations. This correlation is very weak and negative for mixing level pressure, very weak and positive for mixing level height, and weak and linearly positive for inversion temperature. There is also a significant difference in the mean value of each of these three variables on ozone exceedance and non-exceedance days. At the 99.7% confidence level, mean mixing level pressure for ozone exceedance days is less than on non-exceedance days. At the 99.8% confidence level, mean mixing level height for ozone exceedance days is greater than on non-exceedance days. Lastly, at a confidence level greater than 99.99%, mean inversion temperature for ozone exceedance days is greater than on non-exceedance days. With these findings, 1200 UTC radiosonde observations have the potential to become a forecasting tool for surface level ozone concentrations and ozone exceedance days for the Phoenix Metropolitan region. Sounding forecast products may also be used as a strong forecasting tool for surface level ozone concentrations and ozone exceedance days for the Phoenix Metropolitan region.

5. Conclusion

Surface level ozone is an EPA regulated ‘criteria air pollutant’ presenting several health concerns to both humans and the environment. Concentrations of surface level ozone, and the number of ozone exceedance days, are the highest during the summer season (aligning with the North American Monsoon season). Consequently, my research question of this study has been to determine the importance and effects of planetary boundary layer meteorology on surface level ozone concentrations in the Phoenix Metropolitan region during the North American Monsoon for years 2010 through 2020.

The Arizona Department of Environmental Quality (ADEQ) forecasts hourly surface level ozone concentrations. Statistically significant relationships between boundary layer meteorology and surface level ozone may aid ADEQ in their forecasts. Surface level ozone concentrations are highly dependent on local geography and emissions. Past studies showed that surface level meteorological variables can be linked to surface level ozone exceedance days in the Phoenix nonattainment area (Leffel, 2022), and that boundary layer meteorology variables can be correlated to surface level ozone concentrations in the Southwestern US region (Wise and Comrie, 2005). This is an initial study examining the importance and effects of planetary boundary layer meteorology on ozone exceedance days within the Phoenix valley and Phoenix Metropolitan region.

The data collected for this study included both 1200 UTC radiosonde observations launched from the Phoenix National Weather Service office during the North American

Monsoon, and 8-hour ozone concentration observations from Maricopa County monitoring stations. Data used in this study was limited to the North American Monsoon season due to the availability of radiosonde observations for the years 2010 through 2020. Specific boundary layer meteorological variables examined in this study included inversion temperature, mixing level pressure, mixing level height, and the surface level variables of temperature, dew point temperature, pressure, wind speed, and meridional and zonal wind directions. To test my hypothesis, I conducted a set of statistical tests, including the determination of the Pearson product moment correlation coefficient, and/or the Spearman correlation coefficient to ozone, as well a Student's two-sided t-test statistic to determine if there was a difference in mean value of variables on ozone exceedance and non-exceedance days.

For the monsoon periods of 2010 to 2020 in the Phoenix Metropolitan region, my statistical analyses established that my three most important findings include:

- Mixing level pressure was weakly and negatively correlated to maximum surface level ozone.
- Mixing level height was weakly and positively correlated to maximum surface level ozone.
- Inversion temperature was weakly and linearly positively correlated to maximum surface level ozone.

The correlations found for these three variables were highly statistically significant. There was also a statistically significant difference in the mean value of each of these three variables on ozone exceedance and non-exceedance days. Respectively, these variables were less, greater, and greater on exceedance days than on non-exceedance days. These findings support my primary hypothesis that boundary layer meteorology influences ozone concentrations in the Phoenix Metropolitan region, and have the potential to become a forecasting tool for surface level ozone concentrations and ozone exceedance days for the Phoenix Metropolitan region.

A continuation of this project would likely involve the determination of how meteorological variables examined in this study interact with each other to influence surface level ozone concentrations. As of now, I have determined the statistical influence of individual variables on ozone. I have not addressed the causative conditions associated with the statistical findings. I cannot, for instance, say whether or not strong inversions paired with stable and stagnant conditions lead to increased ozone concentrations as described by Oke (1978) as a subsidence inversion.

Future research on surface level ozone should also take a closer look at the role of upper-level winds. For example, an examination of the relationship of winds not only to ozone exceedance and non-exceedance days, but also to the location of the monitoring stations recording an ozone exceedance could be useful in creating ozone forecasting algorithms. It is currently believed that the winds are responsible for the buildup of daily surface level ozone concentrations in locations across the Phoenix Metropolitan region. This analysis may provide a more detailed insight into the locations and causes of communities affected by high surface level ozone concentrations. A more theoretical approach to further our understanding of ozone within the planetary boundary layer in the Phoenix Metropolitan region would be to launch frequent ozone sondes, determining the daily progression of ozone concentrations corresponding to the planetary boundary layer. This would provide information on the subsidence and entrainment of ozone within the lower troposphere, a known contributor to surface level ozone concentrations. However, best results may come from stationary sensors, therefore requiring more creative instrumentation, also requiring the formation of a new data set.

This research builds upon past research on surface level ozone in general, ozone in the southwest, and ozone in the planetary boundary layer. Past research tends to focus on surface level meteorological variables linkage to ozone. The findings here indicate that upper-air variables are most highly significantly correlated to surface level ozone concentrations in the Phoenix Metropolitan region. These findings provide insight that more research on surface level ozone in the planetary boundary is warranted. Significant changes in the mean of upper-air meteorological variables in the planetary boundary layer on ozone exceedance and non-exceedance days were found using morning radiosonde observations. These findings may aid forecasters of air quality, a significant problem for the Phoenix Metropolitan region and many densely populated cities globally.

6. References

Adames, D. K., & Comrie, A. C. (1997). The North American Monsoon. *Bulletin of the American Meteorological Society*, *78*(10), 2197–2214. https://doi.org/10.1175/1520-0477(1997)078<2197:TNAM>2.0.CO;2

American Lung Association. (n.d.). State of the Air 2022. Retrieved October 10, 2022, from https://www.lung.org/research/sota

Bishara, A. J., & Hittner, J. B. (2015). Reducing Bias and Error in the Correlation Coefficient Due to Nonnormality. *Educational and Psychological Measurement*, *75*(5), 785. https://doi.org/10.1177/0013164414557639

Brazel, A. J., & Nickling, W. G. (1986). The relationship of weather types to dust storm generation in Arizona (1965–1980). *Journal of Climatology*, *6*(3), 255–275. https://doi.org/10.1002/JOC.3370060303

California Air Resources Board. (2003). The ozone weekend effect in California, staff report. Retrieved October 10, 2022, from http://www.arb.ca.gov/aqd/weekendeffect/weekendeffect.htm

Cox, W. M., & Chu, S. H. (1996). Assessment of interannual ozone variation in urban areas from a climatological perspective. *Atmospheric Environment*, *30*(14), 2615–2625. https://doi.org/10.1016/1352-2310(95)00346-0

Dean, R. S., Swain, R. E., & Mines., U. States. B. of. (1944). Report Submitted to the Trail Smelter Arbitral Tribunal. Retrieved from https://digital.library.unt.edu/ark:/67531/metadc12613/

Devore, J. L. (2011). Probability and Statistics for Engineering and the Sciences. (8th ed.). Boston, MA: Brooks/Cole, Cengage Learning. Retrieved October 10,2022, from https://faculty.ksu.edu.sa/sites/default/files/probability_and_statistics_for_engineering_and_the_sciences.pdf

Dibb, J. E., Talbot, R. W., Scheuer, E., Seid, G., DeBell, L., Lefer, B., & Ridley, B. (2003). Stratospheric influence on the northern North American free troposphere during TOPSE: 7Be as a stratospheric tracer. *Journal of Geophysical Research: Atmospheres*, *108*(D4), 11–1. https://doi.org/10.1029/2001JD001347

Fast, J. D., Doran, J. C., Shaw, W. J., Coulter, R. L., & Martin, T. J. (2000). The evolution of the boundary layer and its effect on air chemistry in the Phoenix area. *Journal of Geophysical Research: Atmospheres*, *105*(D18), 22833–22848. https://doi.org/10.1029/2000JD900289

Fast, Jerome D., Torcolini, J. C., & Redman, R. (2005). Pseudovertical Temperature Profiles and the Urban Heat Island Measured by a Temperature Datalogger Network in Phoenix, Arizona. *Journal of Applied Meteorology and Climatology*, *44*(1), 3–13. https://doi.org/10.1175/JAM-2176.1

Fujita, E. M., Stockwell, W. R., Campbell, D. E., Keislar, R. E., & Lawson, D. R. (2003). Evolution of the Magnitude and Spatial Extent of the Weekend Ozone Effect in California's South Coast Air Basin. *Journal of the Air & Waste Management Association*, *53*(7), 802–815. https://doi.org/10.1080/10473289.2003.10466225

Gao, H. O. (2007). Day of week effects on diurnal ozone/NOx cycles and transportation emissions in Southern California. *Transportation Research Part D: Transport and Environment*, *12*(4), 292–305. https://doi.org/10.1016/J.TRD.2007.03.004

Granados-Muñoz, M. J., Johnson, M. S., & Leblanc, T. (2017). Influence of the North American monsoon on Southern California tropospheric ozone levels during summer in 2013 and 2014. *Geophysical Research Letters*, *44*(12), 6431–6439. https://doi.org/10.1002/2017GL073375

Heuss, J. M., Kahlbaum, D. F., & Wolff, G. T. (2012). Weekday/Weekend Ozone Differences: What Can We Learn from Them? *Http://Dx.Doi.Org/10.1080/10473289.2003.10466227*, *53*(7), 772–788. https://doi.org/10.1080/10473289.2003.10466227

Hobbs, P. V. (2000). *Introduction to Atmospheric Chemistry*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511808913

Holmes, H. A., Sriramasamudram, J. K., Pardyjak, E. R., & Whiteman, C. D. (2015). Turbulent Fluxes and Pollutant Mixing during Wintertime Air Pollution Episodes in Complex Terrain. *Environmental Science and Technology*, *49*(22), 13206–13214. https://doi.org/10.1021/ACS.EST.5B02616/ASSET/IMAGES/LARGE/ES-2015-026164_0006.JPEG

Khamaganov, V. G., & Hites, R. A. (2001). Rate constants for the gas-phase reactions of ozone with isoprene, α- and β-pinene, and limonene as a function of temperature. *Journal of Physical Chemistry A*, *105*(5), 815–822. https://doi.org/10.1021/JP002730Z/ASSET/IMAGES/LARGE/JP002730ZF00007.JPEG

Lawson, D. R. (2012). The Southern California Air Quality Study. *40*(2), 156–165. https://doi.org/10.1080/10473289.1990.10466671

Leffel, J. (2022). An Analysis of the Relationship Between Tropospheric Ozone Pollution and Synoptic Conditions in Phoenix, Arizona, (Honors Thesis). Retrieved from KEEP. (https://keep.lib.asu.edu/items/165418). Tempe, AZ: Arizona State University.

Li, J., Georgescu, M., Hyde, P., Mahalov, A., & Moustaoui, M. (2015). Regional-scale transport of air pollutants: Impacts of Southern California emissions on Phoenix ground-level ozone concentrations. *Atmospheric Chemistry and Physics*, *15*(16), 9345–9360. https://doi.org/10.5194/ACP-15-9345-2015

Liu, S. C., Trainer, M., Fehsenfeld, F. C., Parrish, D. D., Williams, E. J., Fahey, D. W., et al. (1987). Ozone production in the rural troposphere and the implications for regional and global ozone distributions. *Journal of Geophysical Research: Atmospheres*, *92*(D4), 4191–4207. https://doi.org/10.1029/JD092ID04P04191

Logan, J. A. (1985). Tropospheric ozone: seasonal behavior, trends, and anthropogenic influence. *Journal of Geophysical Research*, *90*(D6), 10463–10482. https://doi.org/10.1029/JD090ID06P10463

Maddox, R. A., McCollum, D. M., & Howard, K. W. (1995). Large-Scale Patterns Associated with Severe Summertime Thunderstorms over Central Arizona. *Weather and Forecasting*, *10*(4), 763–778. https://doi.org/https://doi.org/10.1175/1520-0434(1995)010<0763:LSPAWS>2.0.CO;2

NAAQS Table | US EPA. (n.d.). Retrieved October 10, 2022, from https://www.epa.gov/criteria-air-pollutants/naaqs-table

National Research Council (2004). Research priorities for airborne particulate matter: IV. Continuing research progress. Retrieved October 10, 2022, from https://books.google.com/books?hl=en&lr=&id=RrRTAgAAQBAJ&oi=fnd&pg=PT16&ots=VFLAELW5T8&sig=1FQOD67ia5ceXduTLTHp0IKV43k

Nitrogen Oxides (NOx), Why and How They Are Controlled. (1999). Retrieved October 10, 2022, from http://www.epa.gov/ttn/catc

Oke, T. R. (1978). *Boundary Layer Climates*. New York, NY: Methuen & Co. Ltd.

Pearce, J. L., Beringer, J., Nicholls, N., Hyndman, R. J., & Tapper, N. J. (2011). Quantifying the influence of local meteorology on air quality using generalized additive models. *Atmospheric Environment*, *45*(6), 1328–1336. https://doi.org/10.1016/J.ATMOSENV.2010.11.051

Shutters, S. T., & Balling, R. C. (2006). Weekly periodicity of environmental variables in Phoenix, Arizona. *Atmospheric Environment*, *40*(2), 304–310. https://doi.org/10.1016/J.ATMOSENV.2005.09.037

Sicard, P., Serra, R., & Rossello, P. (2016). Spatiotemporal trends in ground-level ozone concentrations and metrics in France over the time period 1999–2012. *Environmental Research*, *149*, 122–144. https://doi.org/10.1016/J.ENVRES.2016.05.014

Sicard, P., Paoletti, E., Agathokleous, E., Araminienė, V., Proietti, C., Coulibaly, F., & de Marco, A. (2020). Ozone weekend effect in cities: Deep insights for urban air pollution control. *Environmental Research*, *191*, 110193. https://doi.org/10.1016/J.ENVRES.2020.110193

Sillman, S., Logan, J. A., & Wofsy, S. C. (1990). The sensitivity of ozone to nitrogen oxides and hydrocarbons in regional ozone episodes. *Journal of Geophysical Research: Atmospheres*, *95*(D2), 1837–1851. https://doi.org/10.1029/JD095ID02P01837

Sillman, Sanford. (1999). The relation between ozone, NOx and hydrocarbons in urban and polluted rural environments. *Atmospheric Environment*, *33*(12), 1821–1845. https://doi.org/10.1016/S1352-2310(98)00345-8

Stewart, C. (2021). *2020 Air Monitoring Network Review and 2021 Plan*. Retrieved October 10, 2022, from https://www.maricopa.gov/DocumentCenter/View/68917/Final-Air-Monitoring-Network-Plan-2020-PDF?bidId=

Vingarzan, R. (2004). A review of surface ozone background levels and trends. *Atmospheric Environment*, *38*(21), 3431–3442. https://doi.org/10.1016/J.ATMOSENV.2004.03.030

Wise, E. K., & Comrie, A. C. (2005). Meteorologically adjusted urban air quality trends in the Southwestern United States. *Atmospheric Environment*, *39*(16), 2969–2980. https://doi.org/10.1016/J.ATMOSENV.2005.01.024

Witte, R. S., Witte, J. S. (2017). Statistics. (Vol. 11). Hoboken, NJ: John Wiley & Sons, Inc. Retrieved October 10,2022, from https://nibmehub.com/opac-service/pdf/read/Statistics%20by%20Robert%20S.%20Witte-%20John%20S.%20Witte.pdf

7. Appendix

A repository containing the Matlab and R code files written for this study can be found on the following GitHub repository: https://github.com/LaForzaDelDestino/Honors-Thesis