Estimating Ambient Ozone Effect of Kansas Rangeland Burning with Receptor Modeling and Regression Analysis

Zifei Liu *, Yang Liu, James P. Murphy and Ronaldo Maghirang

Department of Biological and Agricultural Engineering, Kansas State University, Manhattan, KS 66506, USA; yliu16@k-state.edu (Y.L.); jmurphy@k-state.edu (J.P.M.); maghir@ksu.edu (R.M.)

* Correspondence: zifeiliu@ksu.edu; Tel.: +1-785-532-3587

Academic Editor: Zhongchao Tan
Received: 29 December 2016; Accepted: 29 January 2017; Published: 9 February 2017

Abstract: Prescribed rangeland burning in April is a long-standing practice in the Flint Hills region of eastern Kansas to maintain the tallgrass prairie ecosystem. The smoke plumes originating from these fires increases ambient PM\textsubscript{2.5} concentrations and potentially contributes to ozone (O\textsubscript{3}) exceedances in downwind communities. Source apportionment research using Unmix modeling has been utilized to estimate contributions of Kansas rangeland burning to ambient PM\textsubscript{2.5} concentrations. The objective of this study was to investigate the potential correlations between O\textsubscript{3} and various sources of PM\textsubscript{2.5} that are derived from receptor modeling, and then to specifically estimate contributions of Kansas rangeland burning to ambient O\textsubscript{3} concentrations through regression analysis. Various daily meteorological data were used as predictor variables. Multiple regression models were developed for the eight-hour daily maximum O\textsubscript{3} as well as the daily contributions of the five PM\textsubscript{2.5} source categories that were derived from receptor modeling. Cross correlation was analyzed among residuals of the meteorological regression models for O\textsubscript{3} and the daily contributions of the five PM\textsubscript{2.5} source categories in order to identify the potential hidden correlation between O\textsubscript{3} and PM\textsubscript{2.5}. The model including effects of meteorological variables and episodic contributions from fire and industrial emissions can explain up to 78% of O\textsubscript{3} variability. For non-rainy days in April, the daily average contribution from prescribed rangeland burning to O\textsubscript{3} was 1.8 ppb. On 3% of the days in April, prescribed rangeland burning contributed over 12.7 ppb to O\textsubscript{3}; and on 7% of the days in April, burning contributed more than 7.2 ppb to O\textsubscript{3}. When the intensive burning activities occur in days with high O\textsubscript{3} background due to high solar radiation or O\textsubscript{3} carryover from the previous day, the contributions from these episodic fire emissions could result in O\textsubscript{3} exceedances of the National Ambient Air Quality Standards (NAAQS). The regression models developed in this study demonstrated that the most valuable predictors for O\textsubscript{3} in the Flint Hills region include the O\textsubscript{3} level on the previous day, total solar radiation, difference between daily maximum and minimum air temperature, and levels of episodic fire and industrial emissions. The long term goal is to establish an online O\textsubscript{3} forecasting tool that can assist regulators and land managers in smoke management during the burning season so that the intensive burning activities can be planned to avoid forecasted high O\textsubscript{3} days and thus prevent O\textsubscript{3} exceedance.

Keywords: source apportionment; prescribed burning; smoke; forecasting

1. Introduction

Prescribed rangeland burning is a long-standing practice in Kansas. Burning is used to enhance the nutritional value of native grasses, control woody plants, and is vital to the maintenance of the tallgrass prairie ecosystem. The Flint Hills tallgrass region of eastern Kansas covers 13 Kansas counties,
and approximately 2 million of the 7 million acres of rangeland in the region are burned each year, mostly in April. The smoke from these burnings has contributed to air quality concerns [1,2]. Precursors for ozone (O₃) and PM$_{2.5}$ (particulate matter that less than 2.5 µm in equivalent aerodynamic diameter) are the smoke constituents of concern. In past years, the Kansas Ambient Air Monitoring Network has recorded elevated concentrations of both O₃ and PM$_{2.5}$ in air during the periods of intensive rangeland burning, and the smoke has contributed to O₃ exceedances. The National Ambient Air Quality Standards (NAAQS) are evolving for O₃ and PM$_{2.5}$. In 2012, the primary annual standard of PM$_{2.5}$ was reduced from 15 to 12 µg/m$^3$. In 2015, the eight-hour O₃ standard was reduced from 75 to 70 ppb. The continued lowering of O₃ and PM$_{2.5}$ standards, together with the Regional Haze Regulations, requires changes in air quality management. Air quality regulators are under pressure to quantify all contributing sources at a time of ever-tightening air quality standards. Modeling tools are being developed for better smoke management related to prescribed fires [3,4]. However, management and modeling efforts are limited by a lack of detailed local source profiles of smoke emissions as well as knowledge gaps in formation of O₃ and secondary organic aerosols.

Various receptor modeling approaches have been developed to resolve the contributions of important emission sources to ambient concentrations, such as chemical mass balance (CMB) [5], positive matrix factorization (PMF) [6], and Unmix [7]. Unlike CMB, the Unmix and PMF models do not require externally supplied source profiles, though they require large data quantity for multivariate analysis. When long-term observational data are available, Unmix and PMF have been shown to be able to generate realistic and interpretable results. Our source apportionment research using Unmix modeling has successfully demonstrated that Kansas rangeland burning contributed 39% of the average PM$_{2.5}$ concentration in April [8].

For decades, ground level O₃ has been the most persistent and perhaps the most dangerous air pollutant in the U.S. [9]. Despite huge expenditures for emission control, violation of the air quality standard for O₃ is not rare. Discernible spikes in O₃ concentrations were observed starting around April each year at Konza Prairie, Wichita, and Kansas City, indicating the effects of the annual burning activities occurring in the Flint Hills. At the Konza Prairie monitoring site, there were 23 days which had daily maximum eight-hour O₃ concentration above 70 ppb during the burning season from 2002 to 2013. However, the correlations between O₃ and smoke from Kansas rangeland burning are not well documented. The actual contributions from burning activities and how O₃ concentrations respond to smoke emissions under different meteorological conditions are not well understood. This information is essential for making an effective smoke management plans, air quality policy, and health warnings. Regression analysis has been successfully used in O₃ modeling in many studies [10–17].

O₃ is a secondary pollutant that is formed through reactions between precursor species such as nitrogen oxides (NO$_x$) and volatile organic compounds (VOCs). O₃ and PM$_{2.5}$ controls are traditionally considered separately because their high pollution periods are not concurrent on seasonal timescales, with O₃ usually peaking in summer and PM$_{2.5}$ often peaking in winter [18]. Temperature and relative humidity (RH) exert opposite effects on O₃ and nitrate aerosols. Higher temperature and lower RH promote O₃ formation but cause volatilization of nitrate aerosols. Our hypothesis is that after source contributions to PM$_{2.5}$ are resolved through receptor modeling, the time series pattern of individual source contributions can be characterized and the effects of metrological variables on each individual source contribution can be modeled. In addition, after seasonal and meteorological effects are removed, the hidden correlation between O₃ and PM$_{2.5}$ contributed from smoke will be revealed, since they share precursor compounds in smoke emissions. As a result, the smoke contributions to O₃ can be estimated from the resolved smoke contributions to PM$_{2.5}$ under given meteorological conditions.

The objective of this study is to investigate the potential correlations between O₃ and various sources of PM$_{2.5}$ that are derived from receptor modeling, and then to specifically estimate contributions of Kansas rangeland burning to ambient O₃ concentrations through regression analysis. The long term goal is to establish an online O₃ forecasting tool that can assist regulators and land managers in smoke management during the burning season in an effort to avoid air quality problems.
2. Materials and Methods

2.1. Site and Data

The study utilized data from routine monitoring networks that provide long-term observations of ambient concentrations of PM$_{2.5}$ (including chemical speciation) and O$_3$, including one IMPROVE (Interagency Monitoring of Protected Visual Environments) [19] site, and one CASNET (Clean Air Status and Trends Network) [20] site. The Konza Prairie CASNET site (providing O$_3$ data) is about 60 miles north of the Tallgrass IMPROVE site (providing speciated PM$_{2.5}$ data). The two sites are located near the center of the Flint Hills region (Figure 1), and thus are ideal for evaluating the effect of prescribed rangeland burning on local air quality. The IMPROVE particulate sampler has four sampling modules. Three modules collect PM$_{2.5}$ using Teflon, nylon, and quartz filters; and the other module collects PM$_{10}$ using a Teflon filter. Every three days, a 24-hour filter sample set (Teflon, nylon, and quartz filters sampling in parallel) was collected. A total of 1428 valid filter sample sets from 26 September 2002 to 31 December 2014 at the Tallgrass site were collected and included in the receptor model analysis. All filter samples were analyzed using the IMPROVE protocols [21]. The Teflon filters were analyzed at the University of California-Davis for gravimetric mass and for analysis of elements in PM$_{2.5}$ using Cu-anode X-ray fluorescence (XRF). The denuded nylon filters were analyzed by Research Triangle Institute for nitrate, sulfate, and chloride using ion chromatography. The quartz filters were analyzed by Desert Research Institute for organic and elemental carbon (OC and EC) using the Thermal/Optical Reflectance method. The daily O$_3$ data from 26 September 2002 to 14 July 2011 at the Konza Prairie site was used in the regression analysis ($n = 1071$). Daily meteorological data included in the regression analysis were from the MHK airport (Manhattan Regional Airport), which is about 5 miles north of the Konza Prairie site.

![Map showing National Oceanic and Atmospheric Administration (NOAA) satellite image of Kansas Flint Hills burning activities in 2014 and location of the Konza Prairie and the Tallgrass sampling site (the approximate perimeter of the Flint Hills region is indicated in red lines; the wind rose is from the weather station at MHK airport).](image)

Figure 1. Map showing National Oceanic and Atmospheric Administration (NOAA) satellite image of Kansas Flint Hills burning activities in 2014 and location of the Konza Prairie and the Tallgrass sampling site (the approximate perimeter of the Flint Hills region is indicated in red lines; the wind rose is from the weather station at MHK airport).

2.2. Summary of Previous Results of Receptor Modeling

The U.S. Environmental Protection Agency (US EPA) Unmix6.0 receptor model was applied to the speciated PM$_{2.5}$ data ($n = 1428$) from the Tallgrass site. The modeling identified the following...
five source categories (S1 to S5) that contribute to local ambient PM$_{2.5}$: S1-nitrate/agricultural (22%), S2-primary smoke particles (5%), S3-secondary organic aerosol (29%), S4-sulfate/industrial (30%), and S5-crustal/soil (14%). S2 was characterized by non-soil potassium and OC/EC. S3 was characterized by large OC components indicating a secondary formation. S2 was identified as smoke emissions and mostly included primary smoke particles from vegetative burning. S3 was identified as secondary organic aerosols that can be generated from biogenic, mobile or smoke sources, in which rangeland burning emission was a major contributor. S1 was characterized by high nitrate, which is expected to be fully neutralized ammonium nitrate because of sufficient precursor ammonia from agricultural sources. S4 was characterized by high sulfate and represented secondary sulfate from regional SO$_2$ emissions including coal-fired power plants and industrial processes. S5 represents resuspended soil particles. The contribution of April burning was estimated to be 0.93 µg/m$^3$ as primary smoke particles and 3.79 µg/m$^3$ as secondary aerosols, which in total contributed 39% of the average 24-h PM$_{2.5}$ concentration in April (12.12 µg/m$^3$). Details can be found in [8]. Comparison analysis of source apportionment using PMF and Unmix models at the Tallgrass site demonstrated that the results from the two models were in general agreement [22].

2.3. Methods of O$_3$ and PM$_{2.5}$ Regression Analysis

Metrics of eight-hour daily maximum O$_3$ and 24-hour average PM$_{2.5}$ species were used in the regression analysis because they are closely related to the NAAQS regulated by the US EPA. Various daily meteorological data were used as predictor variables, including daily maximum air temperature, daily minimum air temperature, precipitation, solar radiation, wind speed, wind direction, etc. Multiple regression models were developed for the eight-hour daily maximum O$_3$ as well as the daily contributions of the five PM$_{2.5}$ source categories (S1 to S5) that were derived from receptor modeling. Sinusoidal terms were included in the regression models to describe the long term seasonal effects in addition to the daily effects of various meteorological variables as suggested in Zvyagintsev et al. [23]. A persistence term was included in the O$_3$ modeling as suggested by the US EPA [24]. Residual analysis was conducted for model validation and for potential needs for data transformation. Cross correlation was analyzed among residuals of the regression models using only meteorological data as predictor variables for O$_3$ and the daily contributions of the five PM$_{2.5}$ source categories in order to identify the potential hidden correlation between O$_3$ and PM$_{2.5}$. The O$_3$ simulation models were then enhanced by including identified episodic contributions from PM$_{2.5}$ sources in addition to the effects of meteorological variables The regressions were done stepwise to add and delete terms based on Akaike Information Criterion (AIC) statistics to limit the number of variables by identifying the most important ones and to obtain the best model fit [25]. The REG and MIXED procedures of SAS (SAS for Windows, Version 9.4, SAS Institute, Cary, NC, USA) were used. Significant effects were declared at $p < 0.05$.

3. Results

3.1. Regression Models Using Meteorological Data as Predictor Variables

The environmental and meteorological data used in the regression analysis are presented in Table 1. The O$_3$ data demonstrated a sinusoidal seasonal pattern with a peak in summer and discernable spikes in April. Opposite seasonal patterns were observed for PM$_{2.5}$ from individual source categories S1 and S4. S1 contributions were generally high in winter, possibly due to meteorological drivers such as low wind speeds, low temperatures, high pressures and stagnant conditions. In contrast, S4 contributions were generally high in summer and low in winter, which is consistent with the effect of photochemistry and the presence of oxidants on the formation of secondary sulfate. A long-term decreasing trend was observed for S4, which was likely due to continuous reductions in regional SO$_2$ emissions following the rules such as the Clean Air Nonroad Diesel Rule and the Acid Rain Program for utilities. Both S2 and S3 had consistent spikes in April when intensive prescribed rangeland burning in the Flint Hills
is conducted. From March to October, the prevailing wind direction was from the south, while from November to February, north wind was almost equally frequent.

### Table 1. Environmental and meteorological data used in regression analysis (2002 to 2011).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 8-h daily maximum O$_3$ concentration (ppb)</td>
<td>42.4</td>
<td>13.0</td>
<td>4.0 to 81.0</td>
</tr>
<tr>
<td>2 Daily (24-h) average PM$_{2.5}$ concentration (µg·m$^{-3}$)</td>
<td>7.82</td>
<td>5.41</td>
<td>0.73 to 58.9</td>
</tr>
<tr>
<td>2 Daily PM$_{2.5}$: contribution of S1-nitrate/agricultural (µg·m$^{-3}$)</td>
<td>1.80</td>
<td>2.59</td>
<td>0 to 21.8</td>
</tr>
<tr>
<td>2 Daily PM$_{2.5}$: contribution of S2-primary smoke particles (µg·m$^{-3}$)</td>
<td>0.39</td>
<td>0.78</td>
<td>0 to 13.4</td>
</tr>
<tr>
<td>2 Daily PM$_{2.5}$: contribution of S3-secondary organic aerosol (µg·m$^{-3}$)</td>
<td>2.25</td>
<td>3.62</td>
<td>0 to 72.1</td>
</tr>
<tr>
<td>2 Daily PM$_{2.5}$: contribution of S4-sulfate/industrial (µg·m$^{-3}$)</td>
<td>2.48</td>
<td>2.66</td>
<td>0 to 19.4</td>
</tr>
<tr>
<td>2 Daily PM$_{2.5}$: contribution of S5-crustal/soil (µg·m$^{-3}$)</td>
<td>0.98</td>
<td>1.47</td>
<td>0 to 19.5</td>
</tr>
<tr>
<td>3 Daily maximum air temperature: T$_{\text{max}}$ (°C)</td>
<td>18.9</td>
<td>11.6</td>
<td>−13.2 to 42.3</td>
</tr>
<tr>
<td>3 Daily minimum air temperature: T$_{\text{min}}$ (°C)</td>
<td>6.0</td>
<td>10.9</td>
<td>−25.0 to 26.3</td>
</tr>
<tr>
<td>3 Difference between T$<em>{\text{max}}$ and T$</em>{\text{min}}$: T$_{\text{diff}}$ (°C)</td>
<td>12.9</td>
<td>4.8</td>
<td>0 to 30.4</td>
</tr>
<tr>
<td>3 Daily precipitation (mm)</td>
<td>2.5</td>
<td>9.7</td>
<td>0 to 128.8</td>
</tr>
<tr>
<td>3 Daily total solar radiation: L (Langley)</td>
<td>323.2</td>
<td>172.6</td>
<td>24.2 to 780.4</td>
</tr>
<tr>
<td>3 Daily average wind speed: V (m·s$^{-1}$)</td>
<td>3.8</td>
<td>2.1</td>
<td>0.6 to 16.0</td>
</tr>
<tr>
<td>3 Daily prevailing wind direction in degrees</td>
<td>0.98</td>
<td>1.47</td>
<td>0 to 19.5</td>
</tr>
</tbody>
</table>

1 O$_3$ data are from the Konza Prairie CASNET [20] site; 2 PM$_{2.5}$ data are from the Tallgrass IMPROVE [19] site; 3 Meteorological data are from the MHK airport.

Regression models for O$_3$ and individual PM$_{2.5}$ source contributions using meteorological data as predictor variables are listed in Table 2. For both O$_3$ and PM$_{2.5}$, significant differences were observed between days with and without precipitation. Therefore, regression models were developed for non-rainy days (daily precipitation = 0 mm) and rainy days (daily precipitation > 0 mm). Relatively higher $r^2$ was achieved for models of S1 in comparison with other source categories indicating that S1 is mainly driven by meteorological variables. In addition to the long term seasonal cycle (the sine terms were highest in March and lowest in September), S1 was negatively affected by daily total solar radiation, daily maximum air temperature, and daily average wind speed. Models of S2 and S3 had relatively low $r^2$ since variation of S2 and S3 were mainly caused by episodic emissions of smoke instead of meteorological variables. S2 had very low background and very weak seasonal effects. In contrast, S3 demonstrated a similar long-term seasonal cycle as S1, indicating that secondary organic aerosols and nitrate aerosols may share some common causes. S3 was found to be more sensitive to temperature than S2. S4 exhibited a different long term seasonal cycle (the sine terms were highest in July and lowest in December). For non-rainy days, S4 was negatively affected by daily total solar radiation and daily maximum air temperature as S1, but S4 also increased with increasing daily minimum air temperature. In a simpler model, S4 actually increased with decreasing difference between daily maximum and minimum air temperature. The O$_3$ concentrations were not only affected by weather conditions on the same day, but also significantly affected by the O$_3$ concentration from the previous day. In the O$_3$ models, the persistent term and the sine term determined the background O$_3$ level. The sine terms were highest in June and lowest in November, reflecting the combined effects of seasonal patterns of O$_3$ precursors and climate conditions. The daily variation of O$_3$ was significantly affected by daily total solar radiation and the difference between daily maximum and minimum air temperature. Higher solar radiation and higher temperature difference resulted in higher O$_3$ increase in addition to the background O$_3$. Solar radiation and temperature difference were partially correlated ($r^2 = 0.13$) since air was heated by solar radiation, while their relationship was likely to be affected by conditions of air such as moisture content. On rainy days, the effect of the previous day’s O$_3$ was much lower than on non-rainy days, while the sensitivity of O$_3$ to solar radiation and air temperature was similar or higher.
Table 2. Regression models using meteorological data as predictor variables.

<table>
<thead>
<tr>
<th>Regression Models</th>
<th>( r^2 )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>For non-rainy days ((n = 799))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( S_1(d) = 6.09 + 1.39 \sin\left(\frac{2\pi(d+32)}{365}\right) - 0.050T_{max} - 0.010L + 1.06 \times 10^{-5}L^2 - 0.47V + 0.028V^2 )</td>
<td>0.40</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>( S_2(d) = 0.04 + 0.37 \sin\left(\frac{2\pi(d+22)}{365}\right) + 0.019T_{max} )</td>
<td>0.08</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>( S_3(d) = 0.07 + 1.27 \sin\left(\frac{2\pi(d+31)}{365}\right) + 0.12T_{max} )</td>
<td>0.08</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>( S_4(d) = 3.72 + 0.43 \sin\left(\frac{2\pi(d+364)}{365}\right) - 0.13T_{max} + 0.15T_{min} - 1.39 \times 10^{-5}L^2 + 3.03 \times 10^{-4}LT_{max} )</td>
<td>0.24</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>( S_5(d) = -0.01 + 0.38 \sin\left(\frac{2\pi(d+76)}{365}\right) + 0.020T_{max}^2 )</td>
<td>0.21</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>( O_3(d) = 0.51O_3(d)<em>0 + 12.6 + 4.31 \sin\left(\frac{2\pi(d+280)}{365}\right) + 0.026L - 4.62 \times 10^{-5}L^2 + 0.00152LT</em>{diff} )</td>
<td>0.75</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

| For rainy days \((n = 272)\) |
|-------------------|------|------|
| \( S_1(d) = 3.47 + 0.95 \sin\left(\frac{2\pi(d+32)}{365}\right) - 0.094T_{max} - 0.0078V^2 \) | 0.53 | <0.01 |
| \( S_2(d) = 0.01 + 0.20 \sin\left(\frac{2\pi(d+22)}{365}\right) + 0.017T_{max} \) | 0.06 | <0.01 |
| \( S_3(d) = 0.44 + 0.020 \sin\left(\frac{2\pi(d+31)}{365}\right) + 0.068T_{max} \) | 0.20 | <0.01 |
| \( S_4(d) = 2.05 + 0.25 \sin\left(\frac{2\pi(d+364)}{365}\right) - 2.10 \times 10^{-5}L^2 + 4.03 \times 10^{-4}LT_{max} \) | 0.10 | <0.01 |
| \( S_5(d) = -0.02 + 0.13 \sin\left(\frac{2\pi(d+76)}{365}\right) + 0.0020T_{max}^2 \) | 0.15 | <0.01 |
| \( O_3(d) = 0.39O_3(d)_0 + 13.9 + 3.58 \sin\left(\frac{2\pi(d+280)}{365}\right) + 0.029L - 5.16 \times 10^{-5}L^2 + 0.00186LT_{diff} \) | 0.65 | <0.01 |

Note: \( S_1(d) \) to \( S_5(d) \) are daily PM \(_{2.5}\) contributions of each of the five individual source categories in \( \mu g\cdot m^{-3}\); \( O_3(d) \) is eight-hour daily maximum \( O_3 \) concentration in ppb; \( O_3(d)_0 \) is eight-hour daily maximum \( O_3 \) concentration on the previous day in ppb; \( d \) corresponds to the day of the year; \( T_{max} \) is the daily maximum air temperature in \( ^\circ C \); \( T_{min} \) is the daily minimum air temperature in \( ^\circ C \); \( T_{diff} \) is the difference between the daily maximum and minimum air temperature in \( ^\circ C \); \( L \) is daily total solar radiation in Langley; and \( V \) is daily average wind speed in m s\(^{-1}\).  

3.2. Residual Analysis

Occasional high positive residuals were observed for the above regression models of \( S_2, S_3, S_4 \) and \( S_5 \), indicating that contributions of these source categories were largely affected by strength of episodic emissions of these sources which was not accounted in the above models. The strength of episodic emissions of these sources can be measured by the residuals of these models. Occasional high positive residuals were also observed for the regression models of \( O_3 \). It was found that the residuals of \( O_3 \) significantly increased with increasing residuals of \( S_3 \) and \( S_4 \), indicating that the episodic emissions of \( S_3 \) and \( S_4 \) both significantly contributed to the variation of \( O_3 \). The high positive residuals of \( S_2 \) and \( S_3 \) both occurred consistently in April, the burning season (Figure 2), confirming that \( S_2 \) and \( S_3 \) were both heavily associated with emissions from rangeland burning. \( S_3 \) correlated with \( S_2 \) with a Pearson correlation coefficient of 0.49 \([8]\). The spikes in \( S_3 \) and \( S_2 \) did not always occur on the same day because the secondary aerosols have different pathways with the primary smoke emissions. The residuals of \( O_3 \) correlate well with the residuals of \( S_3 \) but not the residuals of \( S_2 \), reflecting the fact that both \( O_3 \) and secondary aerosols are highly dependent on emissions of volatile organic compounds. High positive residuals were also consistently observed in early July for \( S_2 \) but not for \( S_3 \), which may suggest that the Fourth of July fireworks contributed a lot of primary aerosols but not as much secondary organic aerosols. For \( S_4 \), high positive residuals were mainly observed in summer, from late June to late September. Figure 3 demonstrated that the residuals of \( S_2 \) and \( S_3 \) were relatively low when wind direction was SW, W, and NW, suggesting rangeland burning had limited impact on the Tallgrass site under these wind directions. This is reasonable since the Tallgrass site is located near the west edge of the Flint Hills burning area. For \( S_4 \), high positive residuals were mainly observed when wind direction was E, SE, or S. This was in agreement with the fact that seven out of the eight existing coal-fired power plants in Kansas are located to the east or southeast of the Flint Hills region. Industrial sources in Wichita could also affect the Tallgrass site under south wind conditions.
eight existing coal-fired power plants in Kansas are located to the east or southeast of the Flint Hills region.

Industrial sources in Wichita could also affect the Tallgrass site under south wind conditions.

Figure 2. Effect of date on model residual of (a) S2—primary smoke particles; (b) S3—secondary organic aerosol; and (c) S4—sulfate/industrial.
Figure 3. Effect of wind direction on model residual of (a) S2—primary smoke particles; (b) S3—secondary organic aerosol; and (c) S4—sulfate/industrial.
3.3. Modeling O₃ by Integrating Episodic Contributions of S3 and S4

The following O₃ prediction models were developed by integrating episodic contributions of S3 and S4 into the meteorological-based O₃ models in Table 2. In the following models, O₃(d) is the eight-hour daily maximum O₃ concentration in ppb; d corresponds to the day of the year; \( T_{\text{diff}} \) is the difference between the daily maximum and minimum air temperature in °C; L is daily total solar radiation in Langley; R₃ is the residual of the meteorological S3 models; and R₄ is the residual of the meteorological S4 models. R₃ and R₄ were used to estimate episodic contributions of S3 and S4 on O₃. The interaction term \( R₃ \times R₄ \) was found to be significant \((p < 0.01)\) and thus was included in the models, which represented the contribution of the interaction between fire emissions and industrial emissions in the process of formation of O₃.

For non-rainy days \((n = 799)\),

\[
O₃(d) = 0.44O₃(d)_0 + 16.0 + 5.17 \sin\left(\frac{2\pi(d+283)}{365}\right) + 0.021L - 4.19 \times 10^{-5}L^2 + 0.0161LT_{\text{diff}} + 0.525R₃ + 0.424R₄ + 0.0208R₃ \times R₄ \quad r^2 = 0.78, \ p < 0.01
\]  

(1)

For rainy days \((n = 272)\),

\[
O₃(d) = 0.36O₃(d)_0 + 14.1 + 4.17 \sin\left(\frac{2\pi(d+283)}{365}\right) + 0.037L - 5.73 \times 10^{-5}L^2 + 0.0161LT_{\text{diff}} + 0.772R₄ \quad r^2 = 0.69, \ p < 0.01
\]  

(2)

For non-rainy days, the model obtained an \( r^2 \) as high as 0.78. Measured and predicted O₃ at the Konza Prairie site are compared in Figure 4, which demonstrated that the models performed well in predicting the O₃ spikes during the burning seasons. Contributions of different terms in the O₃ model for non-rainy days were estimated in Table 3. Based on the model, on a typical non-rainy day in April with average previous day O₃ level and average total solar radiation, the background eight-hour daily maximum O₃ without episodic fire contribution is 50.9 ppb. When the daily solar radiation is higher than 660 Langley, the background O₃ will be more than 62.2 ppb. The average daily contribution of episodic fire emissions to O₃ was estimated to be 1.8 ppb. However, on an intensive burning day, the daily contribution of fire emissions to O₃ was much higher. When \( R₃ > 24.2 \ \mu g \cdot m^{-3} \) (two-sigma uncertainties from the average), the daily contribution of fire emissions to O₃ was estimated to be more than 12.7 ppb. When \( R₃ > 13.8 \ \mu g \cdot m^{-3} \) (one-sigma uncertainty from the average), the daily contribution of fire emissions to O₃ was estimated to be more than 7.2 ppb. Contributions from these episodic fire emissions could result in O₃ exceedance when solar radiation on the same day is high. When O₃ levels from the previous day are high, or when there is also contribution from episodic industrial emissions, the total O₃ level will be higher. From 2003 to 2011, 3% of the days in April had \( R₃ \) higher than 24.2 \ \mu g \cdot m^{-3}, \) and 7% of the days had \( R₃ \) higher than 13.8 \ \mu g \cdot m^{-3} \) as results of intensive prescribed burning. In June, the background O₃ was about 3 ppb higher than in April due to seasonal effect, and episodic industrial emissions are a major cause of daily variation of O₃. Around 5.5% of the days in June had \( R₄ \) higher than 6.33 \ \mu g \cdot m^{-3} \) (two-sigma uncertainties from the average), which could contribute over 2.7 ppb to O₃ levels, and increase chances of O₃ exceedance when solar radiation and O₃ on the previous day are high.
Figure 4. Measured and predicted ozone levels at the Konza Prairie site (2003–2011).
Table 3. Contributions of different terms in the model to the eight-hour daily maximum \( O_3 \) concentrations for non-rainy days at various scenarios.

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>( O_3 ) Due to Previous Day and Seasonal Effect (ppb)</th>
<th>( O_3 ) Due to Solar Radiation (ppb)</th>
<th>( O_3 ) Due to ( R_4 ) (Episodic Industrial Emissions) (ppb)</th>
<th>( O_3 ) Due to ( R_3 ) (Episodic Fire Emissions) (ppb)</th>
<th>( O_3 ) Due to Interactions of ( R_3 ) and ( R_4 ) (ppb)</th>
<th>Total ( O_3 ) (ppb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average in April: ( O_3(d)<em>0 = 51.6 ) ppb, ( L = 358 ) Langley, ( T</em>{\text{diff}} = 13.8 ) °C, ( R_3 = 3.52 \mu g \cdot m^{-3} ), ( R_4 = 0 \mu g \cdot m^{-3} )</td>
<td>40.7</td>
<td>10.2</td>
<td>0</td>
<td>1.8</td>
<td>0</td>
<td>52.8</td>
</tr>
<tr>
<td>Average plus two-sigma uncertainties in April: ( O_3(d)<em>0 = 73.4 ) ppb, ( L = 660 ) Langley, ( T</em>{\text{diff}} = 24.2 ) °C, ( R_3 = 24.2 \mu g \cdot m^{-3} ), ( R_4 = 6.01 \mu g \cdot m^{-3} )</td>
<td>50.3</td>
<td>21.5</td>
<td>2.5</td>
<td>12.7</td>
<td>3.0</td>
<td>90.1</td>
</tr>
<tr>
<td>Average in June: ( O_3(d)<em>0 = 51.1 ) ppb, ( L = 494 ) Langley, ( T</em>{\text{diff}} = 12.6 ) °C, ( R_3 = 0 \mu g \cdot m^{-3} ), ( R_4 = 0 \mu g \cdot m^{-3} )</td>
<td>43.6</td>
<td>10.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>53.9</td>
</tr>
<tr>
<td>Average plus two-sigma uncertainties in June: ( O_3(d)<em>0 = 70.8 ) ppb, ( L = 777 ) Langley, ( T</em>{\text{diff}} = 19.8 ) °C, ( R_3 = 2.18 \mu g \cdot m^{-3} ), ( R_4 = 6.33 \mu g \cdot m^{-3} )</td>
<td>52.3</td>
<td>16.0</td>
<td>2.7</td>
<td>1.1</td>
<td>0.3</td>
<td>72.4</td>
</tr>
</tbody>
</table>
4. Discussion

Smoke from rangeland burning is unlike most other pollutant sources that can use a control device to reduce emissions. Burn management is the key to reduce impact of smoke. There is a significant need for air quality modeling in a collaborative smoke management and an air quality plan. Practical smoke modeling and air quality forecasting tools can assist regulators and land managers to determine timing and size of a burn to avoid air quality problems.

Results of regression analysis demonstrated that ambient O\textsubscript{3} was significantly affected by episodic fire and industrial emissions. The regression model including effects of meteorological variables and episodic fire and industrial emissions can explain up to 78% of O\textsubscript{3} variability. The average of the eight-hour daily maximum O\textsubscript{3} in April from 2002 to 2011 at the Konza Prairie site was 51.6 ± 11.0 ppb. For non-rainy days in April, the daily average contribution from prescribed rangeland burning to O\textsubscript{3} was only 1.8 ppb. Variation of background O\textsubscript{3} without burning contributions was mainly determined by solar radiation and O\textsubscript{3} carryover from the previous day. Based on regression analysis of the historical data, on 3% of the days in April, prescribed rangeland burning contributed more than 12.7 ppb to O\textsubscript{3}; and on 7% of the days in April, burning contributed more than 7.2 ppb to O\textsubscript{3}. When intensive burning activities occur in days with high O\textsubscript{3} background due to high solar radiation or O\textsubscript{3} carryover from the previous day, the contributions from these episodic fire emissions could result in O\textsubscript{3} exceedances of the NAAQS. Smoke management will be significantly improved if the intensive burning activities can be planned to avoid forecasted high O\textsubscript{3} days by using practical forecasting tools. The regression models developed in this study demonstrated that the most valuable predictors for O\textsubscript{3} in the Flint Hills region include the O\textsubscript{3} level on the previous day, total solar radiation, difference between daily maximum and minimum air temperature, and levels of episodic fire and industrial emissions. Depending on availability of data, the predictor of difference between daily maximum and minimum air temperature may be replaced by a measure of air moisture content, since temperature is actually determined by solar radiation and air moisture content, and RH has been identified as a successful predictor in other statistical O\textsubscript{3} models [23]. In a practical forecasting model, O\textsubscript{3} level on the next day can be predicted based on current measurements of O\textsubscript{3} concentrations, and forecasts of weather conditions. Forecasts of episodic fire and industrial emissions are also critical in O\textsubscript{3} forecasting during the burning season.

Research is ongoing to expand the regression study to include environmental datasets at more monitoring sites that may be affected by rangeland burning in Kansas, such as Kansas City and Wichita. More relevant variables, such as measures of air stability and air moisture content, may be added to the regression models. In some sites, data may be stratified by seasons or meteorological variables, such as wind direction to improve regression performance. Various modeling techniques, such as machine learning with random forest algorithm, will be utilized to research the best models. Satellite data on fire radiative power have been used to estimate burn area of large-scale open fires [26,27]. Aerosol products retrieved from satellite data, such as aerosol optical depth (AOD) from Geostationary Operational Environmental Satellite (GOES) and Moderate Resolution Imaging Spectroradiometer (MODIS), aerosol composition, and aerosol vertical distribution from Cloud Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO), have been used to evaluate results of air quality forecasting [28–30]. Future directions will be focused on improving the performance of the regression forecasting models by incorporating real-time satellite data as model input so that the long-term goal of establishing an online O\textsubscript{3} forecasting tool can be achieved.

5. Conclusions

The contribution of Kansas rangeland burning to ambient O\textsubscript{3} was estimated by establishing an O\textsubscript{3} simulation model through receptor modeling and regression analysis. The results confirmed the hypothesis and revealed the significant correlations between O\textsubscript{3} and secondary PM\textsubscript{2.5} from smoke after seasonal and meteorological effects were removed. The necessity of O\textsubscript{3} forecasting during the burning season was demonstrated. The regression analysis identified the O\textsubscript{3} level on the previous day, total solar radiation, difference between daily maximum and minimum air temperature, emission
levels of episodic fire, and industrial emissions as the most valuable predictors for ambient $O_3$ in the Flint Hills region, and thus established a foundation for future efforts of $O_3$ forecasting in the region.

**Acknowledgments:** IMPROVE is a collaborative association of state, tribal, and federal agencies, and international partners. U.S. Environmental Protection Agency is the primary funding source, with contracting and research support from the National Park Service. This study is contribution No. 17-245-J from the Kansas Agricultural Experiment Station.

**Author Contributions:** Zifei Liu and Yang Liu processed and analyzed the data. Zifei Liu wrote the paper. James P. Murphy and Ronaldo Maghirang reviewed and edited the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

2. Liu, Z. *Air Quality Concerns of Prescribed Range Burning in Kansas*; Kansas State University Agricultural Experiment Station and Cooperative Extension Service: Manhattan, KS, USA, 2014.


