Decadal Trends and Variability in Intermountain West Surface Ozone near Oil and Gas Extraction Fields

Ying Zhou[1], Huiting Mao[1], and Barkley C. Sive[2]

[1]Department of Chemistry, State University of New York College of Environmental Science and Forestry, Syracuse, NY, 13210, USA
[2]National Park Service, Air Resources Division, Lakewood, CO 80225, USA

Corresponding author: H. Mao (hmao@esf.edu)

Contents of this File

Text: Section S1 to S8

Figures: S1 to S3

Tables: S1 to S8

S1. Locations of selected sites

Long-term ozone (O_3) observations were available at 18 rural sites in the U.S. Intermountain West and 11 sites were located within 100 km of the shale play (Figure 1 and Table S1). Data were obtained from the National Park Service (NPS), Clean Air Status and Trends Network (CASTNET), and Wyoming Department of Environmental Quality (WDEQ).

S2. Identification of reference sites

The backward trajectory cluster analysis was used to identify sites that were under the minimal influence of oil and natural gas (O&NG) extractions (Figure 1). ZION, GRCA, GRBA, and PEFO are located in the downwind area of Las Vegas (Figures S1c-f). Cluster analysis indicated that ~20% – 50% of air masses likely came from the direction of Las Vegas. In addition, ~30% of air masses at GRCA (Cluster 4) and PEFO (Cluster 3) passed Phoenix. Therefore, surface O_3 at ZION, GRCA, GRBA, and PEFO could be significantly influenced by the nearby anthropogenic emissions. At BADL, Cluster 3 indicated that 32% of air masses transported from the periphery of Powder River Basin (Figure S1g). In comparison, nearly all clusters reaching YELL and CRMO were found be from the upwind areas of O&NG extraction. Therefore, YELL and CRMO were used as reference sites to investigate the decadal O_3 change.

Table S1. Names and locations of sites used in this study.

Site Abbreviation	Park Unit/Measurement Network	Site	Latitude	Longitude	Elevation (m)	Start Date	End Date	Year-round Measurement
CANY	Canyonlands National Park	Island in the Sky	38.46	-109.82	1809	7/1/1992	12/30/2015	1993 - 2015
DINO	Dinosaur National Monument	West Entrance Housing	40.29	-108.94	1463	4/1/2005	12/30/2015	2011 - 2015
MEEK	Meeker	Plant Science	40.00	-107.85	1994	1/1/2010	12/30/2015	2010 - 2015
MEVE	Mesa Verde National Park	Resource Management Area	37.20	-108.49	2165	3/1/1993	12/30/2015	1995 - 2015
RANG	Rangely	Golf Course	40.09	-108.76	1655	8/1/2010	12/30/2015	2011 - 2015
ROMO	Rocky Mountain National Park	Long's Peak	40.28	-105.55	2743	4/1/1987	12/30/2015	1998 - 2015
WICA	Wind Cave National Park	Visitor Center	43.56	-103.48	1292	12/31/2003	12/30/2015	2005 - 2008; 2010 - 2014
CAMP	WDEQ	Campbell, WY	44.15	-105.53	1427	1/1/2005	12/31/2015	2005 - 2015
CNTL	CASTNET	Centennial, WY	41.36	-106.24	3175	5/9/1989	12/31/2015	1990 - 2015
GTHC	CASTNET	Gothic, CO	38.96	-106.99	2915	5/13/1989	12/31/2015	1990 - 2015
PNDE	CASTNET	Pinedale, WY	42.93	-109.79	2386	10/21/1988	12/31/2015	1989 - 1994; 1996 - 2015
BADL	Badlands National Park	Visitor Center	43.74	-101.94	739	10/1/1987	12/31/2014	1988 - 1991; 2004 - 2014
CRMO	Craters of the Moon National Monument & Preserve	Visitor Center	43.46	-113.56	1815	9/24/1992	12/30/2015	1993 - 2004; 2007 - 2015
GRBA	Great Basin National Park	Maintenance Yard	39.01	-114.22	2060	8/24/1993	12/30/2015	1994 - 2015
GRCA	Grand Canyon National Park	The Abyss	36.06	-112.18	2073	1/1/1993	12/30/2015	1993 - 2015
PEFO	Petrified Forest National Park	South Entrance	34.82	-109.89	1723	1/1/2002	12/30/2015	2003 - 2015
YELL	Yellowstone National Park	Water Tank	44.56	-110.40	2400	6/1/1996	12/30/2015	1997 - 2015
ZION	Zion National Park	Dalton's Wash	37.20	-113.15	1213	1/1/2004	12/30/2015	2004 - 2015

* MEEK and RANG are not located in national parks, but measurements at these two sites are also maintained by the NPS.

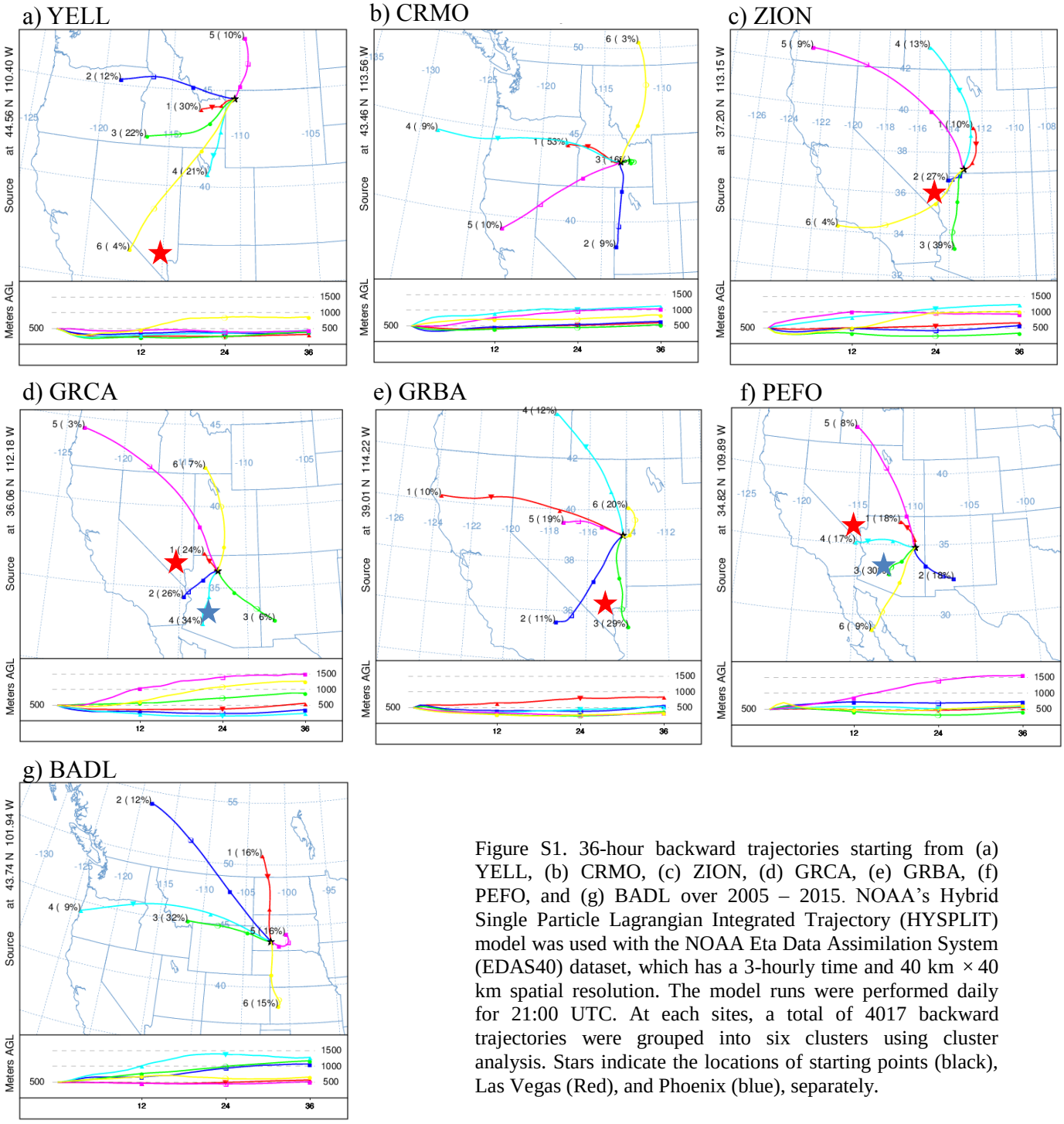

Figure S1. 36-hour backward trajectories starting from (a) YELL, (b) CRMO, (c) ZION, (d) GRCA, (e) GRBA, (f) PEFO, and (g) BADL over 2005 – 2015. NOAA's Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model was used with the NOAA Eta Data Assimilation System (EDAS40) dataset, which has a 3-hourly time and 40 km × 40 km spatial resolution. The model runs were performed daily for 21:00 UTC. At each sites, a total of 4017 backward trajectories were grouped into six clusters using cluster analysis. Stars indicate the locations of starting points (black), Las Vegas (Red), and Phoenix (blue), separately.

S3. Trends in surface ozone at each site

The ozone design value (ODV) is defined as the 3 year running mean of the annual fourth-highest daily maximum 8 hour average (DM8HA) O_3 concentration. We calculated the annual fourth highest DM8HA (A4DM8HA) for the 13 monitoring sites and their trends were examined through ordinary linear least-square regression before and after 2005. The trends were also calculated separately for 5^{th}, 50^{th}, 95^{th} percentiles of DM8HA O_3 in summer and winter.

Table S2. Trends in the A4DM8HA O_3 at each site before and after 2005. Boldfaced numbers indicate p-value≤0.10.

Site	Time Period	Trends	Time Period	Trends (ppbv yr^{-1})	Average (ppbv)
CANY	1993 - 2004	**0.55 (0.06)**	2005 - 2015	**-0.54 (0.02)**	68.8
CAMP			2005 - 2015	-0.44 (0.25)	63.9
DINO			2011 - 2015		**82.7**
MEEK			2010 - 2015		64.1
RANG			2011 - 2015		**72.2**
MEVE	1995 - 2004	0.45 (0.25)	2005 - 2015	**-0.76 (<0.01)**	69.4
ROMO	1998 - 2004	0.41 (0.78)	2005 - 2015	-0.46 (0.23)	**75.1**
WICA			2005 - 2014	**-1.21 (0.05)**	64.5
CTNL	1990 - 2004	**0.65 (<0.01)**	2005 - 2015	-0.06 (0.78)	68.4
GTHC	1990 - 2004	0.17 (0.46)	2005 - 2015	-0.16 (0.64)	68.1
PNDE	1989 - 2004	**0.33 (0.03)**	2005 - 2015	-0.08 (0.75)	66.8
CRMO	1993 - 2004	**0.79 (0.06)**	2007 - 2015	-0.50 (0.23)	63.1
YELL	1997 - 2004	0.30 (0.57)	2005 - 2015	-0.17 (0.58)	64.8

Table S3. Trends in seasonal 5^{th}, 50^{th}, and 95^{th} percentiles of DM8HA O_3 at each site in winter and summer over 2005 – 2015. Boldfaced numbers indicate p-value ≤0.10. Note that wintertime O_3 in a certain year referred to O_3 in January, February, and December in the previous year.

	Winter				Summer			
Site	Time Period	5th	50th	95th	Time Period	5th	50th	95th
CANY	2006 - 2015	0.05	-0.07	-0.17	2005 - 2015	**-0.53**	**-0.51**	-0.40
CAMP	2006 - 2015	0.17	0.12	-0.17	2005 - 2015	-0.12	-0.32	-0.51
DINO	-				2005 - 2015	-0.16	-0.11	0.09
MEVE	2006 - 2015	-0.41	-0.22	**-0.49**	2005 - 2015	-0.51	**-0.63**	**-0.61**
ROMO	2006 - 2015	0.22	-0.04	-0.21	2005 - 2015	0.13	-0.13	-0.34
WICA	2006 - 2014	-0.23	-0.16	-0.30	2005 - 2014	-0.56	-0.98	**-1.38**
CTNL	2006 - 2015	0.15	0.13	-0.08	2005 - 2015	-0.03	-0.07	0.14
GTHC	2006 - 2015	-0.11	**-0.24**	**-0.36**	2005 - 2015	-0.25	-0.31	-0.23
PNDE	2006 - 2015	0.12	0.10	-0.69	2005 - 2015	0.26	-0.07	-0.10
CRMO	2008 - 2015	-0.71	**-1.18**	**-1.10**	2007 - 2015	**-0.14**	**-0.34**	**-0.46**
YELL	2006 - 2015	-0.06	-0.10	-0.55	2005 - 2015	-0.25	-0.39	-0.11

S4. Emissions in selected counties

Emissions of NOx and VOCs were obtained from EPA National Emission Inventory (NEI).

Table S4. NO_x emissions in selected counties

			NOx emission (Ton)_O&NG Extraction				NOx emission (Ton)_total			
County	State	Basin	2005	2008	2011	2014	2005	2008	2011	2014
San Juan	New Mexico	San Juan Basin	263	351	14504	13906	80997	35626	44806	41902
Emery	Utah	Paradox Basin	0	0	88	158	30282	32814	22215	20729
Uintah	Utah	Uintah-Piceance Basin	0	0	10033	7412	8698	2049	11897	9407
Campbell	Wyoming	Powder River Basin	0.33	2407	505	2301	12142	22195	44429	18702
Sublette	Wyoming	Greater Green River Basin	0	5977	2501	4188	2369	9091	4970	5792
Jackson	Colorado	North Park Basin	0	0	2	107	243	442	632	573
Rio Blanco	Colorado	Uintah-Piceance Basin	7	190	1434	4021	2020	3914	5027	6997
Weld	Colorado	Denver-Julesburg Basin	0	76	12478	17892	13112	20088	32696	33275
Custer	South Dakota		0	0	2.25	0.45	1156	1309	1727	1373

Table S5. VOC emissions in selected counties

			VOC emission (Ton)_O&NG Extraction				VOC emission (Ton)_total			
County	State	Basin	2005	2008	2011	2014	2005	2008	2011	2014
San Juan	New Mexico	San Juan Basin	142	512	22089	32819	8858	58095	88840	97096
Emery	Utah	Paradox Basin	0	0	459	549	1715	43441	46945	37107
Uintah	Utah	Uintah-Piceance Basin	0	0	76502	86915	4545	40044	116207	126578
Campbell	Wyoming	Powder River Basin	52	2697	6703	8559	9558	35514	48870	34363
Sublette	Wyoming	Greater Green River Basin	5	15863	9079	58304	1558	70276	45282	86146
Jackson	Colorado	North Park Basin	28	57	516	688	1125	17402	20813	12850
Rio Blanco	Colorado	Uintah-Piceance Basin	0.4	1092	23432	6141	1647	34518	57809	38601
Weld	Colorado	Denver-Julesburg Basin	0	713	104473	91709	12846	52991	150982	116146
Custer	South Dakota		0	0	12	1	1789	27132	29169	19210

S5. Wintertime O_3 at Edmonton, Alberta, Canada

Ozonesonde data at Edmonton were obtained from the World Ozone and Ultraviolet Radiation Data Center (http://woudc.org/data/products/ozonesonde/). Significant negative correlation (r = -0.49, p = 0.01) was found between lower tropospheric O_3 at Edmonton and the AO index over the winter of 1988 – 2014 (Figure S2).

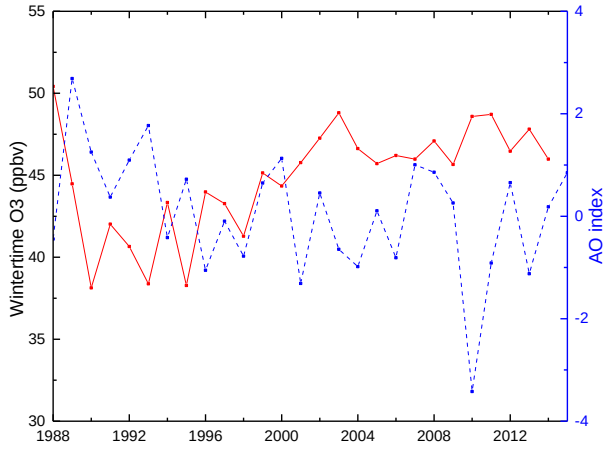

Figure S2.Time series of lower tropospheric (2 – 5 km) O_3 at Edmonton (53.55 ℃N, 114.10 ℃W) and the AO index in winter.

S6. Fire index

The Lagrangian dispersion model, Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) model, was used to estimate O_3 concentrations determined by wildfire emissions. HYSPLIT simulated the long-range and mesoscale transport, diffusion, and dry and wet deposition of gas or particles (Draxler & Hess, 1998). The National Center for Environmental Prediction (NCEP) Eta Data Assimilation System (EDAS40) datasets, with a 3-hourly time and 40 km × 40 km spatial resolution, were used to drive the Lagrangian dispersion model. For each day over summer 2005 – 2015, 250,000 particles were released at a constant hourly rate during the first 24 h at a height of 100 m over each measurement site. A decay half-

life of five days occurred to the particle mass (Lu et al., 2016) and the particles were traced backward for five days. The retroplumes of each sample were calculated daily in $0.25° \times 0.25°$ horizontal resolution from the surface to 5 km. In total, we have computed over 4950 HYSPLIT retroplumes for each Intermountain West site.

The Global Fire Emissions Database version 4.1 (GFED4s) were used to provide information of biomass burning emissions. GFED4s had a spatial resolution of $0.25° \times 0.25°$ and contained monthly burned area, fire carbon and dry matter emissions (DM), as well as daily fraction and contributions of different fire types to the total emissions since 2003. The fire types included savanna, grassland, and shrubland fires, boreal forest fires, temperature forest fires, tropical deforestation and degradation fires, peatland fires, and agricultural waste burning. Biomass burning was complex and the emissions depended greatly on the ecosystem type (Jaffe et al., 2008). Akagi et al. (2011) found that wildfire emissions had a relatively high VOC/NO_x ratio of ~10 – 100 and O_3 production in smoke plumes was very sensitive to NO_x concentrations. Different from previous studies (Jaffe et al., 2008; Lu et al., 2016) using monthly or daily wildfire burned areas, we estimated daily wildfire NO_x emissions using monthly DM emissions, daily fraction, and fractional contribution of fire types in combination of emission factors (Akagi et al. 2011). Based on Lu et al. (2016), we then calculated a fire index (FI) as the product of daily HYSPLIT residence time and daily wildfire NO_x emission, in units of g NO_x m^{-3}. The sum of FI over the 5-day period was defined as total fire index (TFI):

$$\mathrm{FI}(n) = \sum_i \sum_j \sum_k \mathrm{E_{DM}}(i, j) \times \mathrm{t}_r(i, j, n) \times \mathrm{F_{daily}}(i, j, n) \times \mathrm{F_{type}}(k)$$

$$\mathrm{TFI} = \sum_{n=1}^{5} \mathrm{FI}(n)$$

where $E_{DM}(i, j)$ was the monthly wildfire dry matter emission in the model grid cell i (longitude) and j (latitude), $F_{daily}(i, j, n)$ was the daily fraction of wildfire emissions, $F_{type}(k)$ was the NO_x

emission factor from fire type *k*, $t_{r(i, j, n)}$ was HYSPLIT calculated daily residence time, and *n* defined the backward day in the 5-day period. The summertime wildfire NO_x impacts were computed by averaging the daily TFI index.

S7. Impacts of wildfire emissions on summertime O_3

In contrast to wintertime O_3 levels, summertime median DM8HA O_3 displayed significant interannual variations with large differences between sites (Figure 3). The TFI, calculated with HYSPLIT dispersion model and wildfire NO_x emissions, varied at each site over 2005 – 2015 (Figure S3). In 2012, the western U.S. experienced widespread drought with hot weather causing frequent wildfires across the region (Abeleira & Farmer, 2017). The decadal highest TFI value (0.013 – 0.052 g NO_x m^{-3}) in summer 2012 were observed at CRMO and YELL, as well as 9 other sites (DINO, MEEK, RANG, ROMO, WICA, CAMP, CNTL, PNDE, and GTHC) (Figure S3), which was mostly consistent with the observed decadal highest summertime median DM8HA O_3 at the same sites except WICA (54 – 65 ppbv).

Significant positive correlations between summertime O_3 and TFI were found at the reference site YELL, as well as at CANY, CAMP, DINO, MEVE, WICA, CNTL, GTHC, and PNDE during their respective time periods over 2005 – 2015 (Table 2). It should be noted that significant correlation was found for seasonal 75^{th} or 95^{th} percentiles at YELL, CANY, GTHC, PNDE, and MEVE. Strong correlations, albeit not significant, were also found between TFI and summertime 95^{th} percentile DM8HA O_3 at RANG (Table S6). As indicated by data at the reference site YELL, the wildfire emissions had a larger impact on high O_3 levels in summer over the Intermountain West. While wintertime O_3 at CANY, DINO, and RANG was strongly impacted by photochemical production from O&NG emissions within the basins (Section 5), during the summer, the interannual variability of O_3 at 8 out of 11 sites near the O&NG

extraction fields appeared to be predominantly impacted by photochemical production from wildfire emissions.

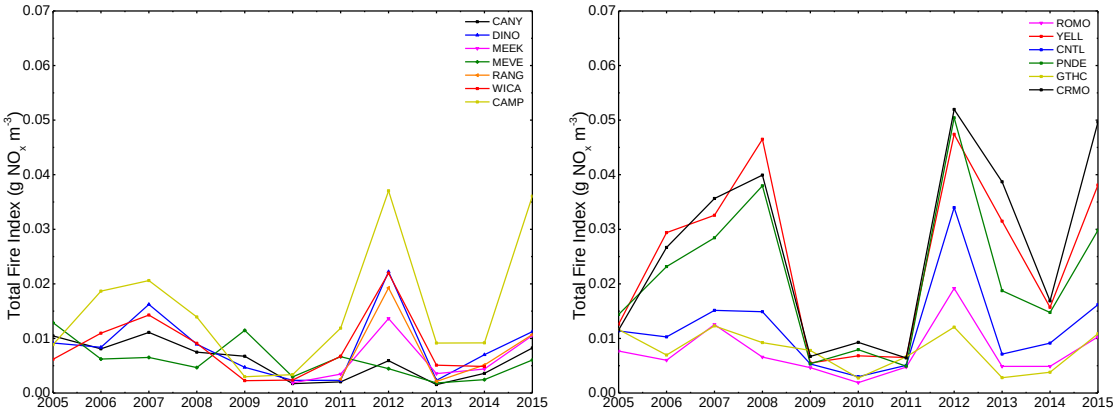

Figure S3. Time series of summertime total fire index at each site.

Table S6. Correlation coefficient (r) and p-value in parenthesis between the pairs of variables over the summer of 2005 – 2015. Boldfaced numbers indicate p-value ≤0.10.

		r(O_3 vs TFI)			Partial r	Regression	r(O_3 vs Relative Humidity)			Partial r
Site	Time Period	5th	50th	95th	95th	Intercept_50th	5th	50th	95th	5th
CANY	2005 - 2015	0.20 (0.55)	0.47 (0.14)	0.37 (0.26)	0.32 (0.40)	**54.01**	**-0.65 (0.08)**	-0.50 (0.11)	-0.01 (0.97)	**-0.61 (0.07)**
CAMP	2005 - 2015	0.33 (0.33)	**0.57 (0.06)**	0.50 (0.12)	0.22 (0.57)	**50.08**	-0.45 (0.19)	**-0.92 (<0.01)**	**-0.66 (0.04)**	**-0.88 (<0.01)**
DINO	2005 - 2015	**0.62 (0.04)**	**0.58 (0.06)**	**0.62 (0.04)**	**0.67 (0.10)**	**52.42**	-0.14 (0.72)	-0.18 (0.64)	-0.21 (0.59)	**0.74 (0.06)**
MEEK	2010 - 2015	0.13 (0.81)	0.01 (0.99)	0.45 (0.37)	0.35 (0.56)	**51.28**	**-0.77 (0.06)**	-0.71 (0.11)	-0.31 (0.55)	**-0.85 (0.07)**
RANG	2011 - 2015	0.34 (0.57)	0.26 (0.67)	0.76 (0.24)	0.07 (0.92)	**52.58**	**-0.90 (0.03)**	-0.62 (0.26)	-0.32 (0.60)	**-0.91 (0.09)**
MEVE	2005 - 2015	0.39 (0.24)	0.50 (0.11)	**0.66 (0.03)**	0.42 (0.26)	**53.21**	-0.49 (0.12)	-0.46 (0.14)	-0.36 (0.28)	-0.43(0.25)
ROMO	2005 - 2015	0.28 (0.39)	0.40 (0.22)	0.40 (0.21)	0.53 (0.12)	**54.89**	-0.40 (0.22)	**-0.67 (0.02)**	**-0.80 (<0.01)**	**-0.58 (0.09)**
WICA	2005 - 2014	**0.65 (0.04)**	**0.71 (0.02)**	**0.57 (0.08)**	-0.42 (0.30)	**46.18**	**-0.71 (0.02)**	**-0.97 (<0.01)**	**-0.90 (<0.01)**	**-0.79 (0.02)**
CNTL	2005 - 2015	**0.59 (0.06)**	0.49 (0.12)	0.36 (0.28)	-0.13 (0.83)	**53.69**	-0.69 (0.13)	**-0.74 (0.09)**	-0.38 (0.45)	-0.44 (0.46)
GTHC	2005 -2015	0.28 (0.40)	**0.55 (0.08)**	**0.49 (0.03)**	**0.67 (0.09)**	**47.76**	0.24 (0.65)	-0.19 (0.72)	-0.21 (0.69)	0.45 (0.45)
PNDE	2005 - 2015	0.46 (0.16)	0.41 (0.21)	**0.53 (0.09)**	-0.06 (0.89)	**52.96**	**-0.58 (0.10)**	-0.67 (0.55)	**-0.78 (0.01)**	**-0.70 (0.05)**
CRMO	2007 -2015	-0.45 (0.22)	0.33 (0.38)	0.29 (0.44)	-0.26 (0.58)	**50.58**	0.36 (0.48)	**-0.89 (0.02)**	**-0.91 (0.01)**	**-0.87 (0.05)**
YELL	2005 - 2014	0.01 (0.49)	0.14 (0.34)	**0.50 (0.05)**	-0.03 (0.95)	**49.57**	**-0.56 (0.04)**	**-0.70 (0.01)**	**-0.81 (<0.01)**	**-0.70 (0.03)**

Regression equations were used to quantify the relationship between seasonal median DM8HA O_3 and TFI, where the intercepts from the equations represented decadal summertime median DM8HA O_3 values in the absence of fires (Jaffe et al., 2008). No significant positive correlation was found between summertime O_3 at ROMO and TFI, while the O_3 concentration without wildfires (55 ppbv), based on the intercept value, was found to be the highest at this site. As stated in Section 4, ROMO was influenced by the high O_3 concentrations from the southeast

over 2005 – 2015 (Figure 5). Summertime O_3 at ROMO was under strong influence of frequent transport of highly polluted air masses emerging from the greater Denver area, which likely dominated over the impact of wildfire emissions during the decade of the study period. Reddy and Pfister (2016) also found significant correlations between July DM8HA O_3 and 500hPa heights, particularly in areas of elevated terrain near urban sources of O_3 precursors.

S8. BOXMOX Model Simulations

S8.1 Field Campaigns

Surface observations were obtained from five campaigns to constrain model simulations (Tables S7–8). The Nitrogen, Aerosol Composition, and Halogens on a Tall Tower (NACHTT) campaign was conducted at the National Oceanic and Atmospheric Administration's Boulder Atmospheric Observatory (BAO). BAO is located in a primarily agricultural region and is very close to major urban areas. It is ~35 km north of Denver and ~30 km east of Boulder. In addition, the site is located within the Denver-Julesburg Basin, which is an active oil and gas exploration and production region. The Uintah Basin Winter Ozone Studies (UBWOS) were a set of field campaigns held at Horse Pool during January and February of 2012, 2013, and 2014. Horse Pool is a remote site, located within the oil and gas basin of northeastern Utah. In summer 2014, the NSF Front Range Air Pollution and Photochemistry Experiment (FRAPPÉ) and the NASA Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality (DISCOVER-AQ) field campaigns conducted aircraft, mobile, and ground-based measurements over 15 locations across the Front Range. In this study, we used field measurements at ROMO-LP. All campaign data were available at https://esrl.noaa.gov/csd /groups/csd7/measurements/.

Table S7. Name and locations of campaign sites used in this study

Campaign	Time Period	State	Site	Elevation/km	Latitude	Longitude
NACHTT	02182011-03132011	Colorado	BAO	1.58	40.05	-105.01
UBWOS2012	01172012-02222012	Utah	Horse Pool	1.53	40.14	-109.47
UBWOS2013	02012013-02212013	Utah	Horse Pool	1.53	40.14	-109.47
UBWOS2014	01182014-02142014	Utah	Horse Pool	1.53	40.14	-109.47
FRAPPÉ	07172014-09032014	Colorado	ROMO-LP	2.74	40.28	-105.55

S8.2 Photolysis Rates

The model used the NCAR TUV radiation model to calculate photolysis frequencies, with inputs for latitude, longitude, altitude, Julian day, temperature, surface pressure, total O_3 column, and albedo. Inputs of surface temperature and pressure were constrained to measurements from field campaigns, while total O_3 column was derived from averaged OMI data. The TUV calculated photolysis frequencies were then scaled to the observed j(NO_2), except for j(O^1D). Measurement of j(NO_2) was not available for UBWOS2013, UBWOS2014, and FRAPPÉ. Instead, total downwelling radiation measurements from these three campaigns were used to calculate photolysis frequencies by comparison with data from UBWOS2012 (Edwards et al. 2014). Polynomial regression was used to find the relationship between downwelling solar radiation versus j(NO_2) using data from UBWOS2012 (E1). Then, j(NO_2) during UBWOS2013, UBWOS2014, and FRAPPÉ was calculated using the derived equation (E1).

$$j(NO_2) = 5.509\times10^{-6} + 1.425\times10^{-5}\times Radiation_{downwelling} - 4.760\times10^{-9}\times Radiation_{downwelling}^{2} \quad \text{E1}$$

S8.3 Turbulent Mixing

Mixing ratios of CO, CH_4, NO, NO_2, and non-methane VOCs were constrained to observations by introducing turbulent mixing. This was represented by adding the background concentration of a species outside the box (Knote et al., 2015).

$$\frac{\partial c_{box}}{\partial t} = \frac{1}{\tau}(c_{bg} - c_{box}(t = t_0))$$

$$c_{box}(t = t_0 + \Delta t) = c_{box}(t = t_0) + \frac{\partial c_{box}}{\partial t} \cdot \Delta t$$

Where t_0 was the beginning of the time step, τ the mixing time-scale, c_{bg} the background concentrations of a species i outside the box, and c_{box} the initial concentration of the species i.

S8.4 Physical loss

First order rate constants were used to simulate all non-chemical loss in the model due to surface deposition. Dry deposition rates of 0.4 cm/s were used for O_3 over the continent and 0.07 cm/s over the snow, based on Hauglustaine et al. (1994).

TableS8. Chemical observations used to inform the box model analysis in this analysis.

	Observational technique				
Compound	NACHTT	UBWOS2012	UBWOS2013	UBWOS2014	FRAPPÉ
O_3	CRDS	CRDS	CRDS	CRDS	
H_2O					
NO	CRDS	CRDS	CRDS	CRDS	No_noxbox
NO_2	CRDS	CRDS	CRDS	CRDS	NO2_Noxbox
N2O5	CRDS	CRDS	CRDS	CRDS	
HNO_3	CIMS	CIMS	CIMS	HR-TOF-CIMS	
HONO	CIMS	CIMS	CIMS	HR-TOF-CIMS	
CO	Thermo 48C	VUF	VUF	VUF	
Alkanes					
Methane	1.8 ppmv	CRDS	CRDS	CRDS	1.8 ppmv
Ethane	GC-MS*	GC-MS	GC-FID	GC-MS	GC-MS*
Propane	GC-MS*	GC-MS	GC-FID	GC-MS	GC-MS*
n-Butane	GC-MS*	GC-MS	GC-FID	GC-MS	GC-MS*
iso-Butane	GC-MS*	GC-MS	GC-FID	GC-MS	GC-MS*
n-Pentane	GC-MS*	GC-MS	GC-FID	GC-MS	GC-MS*
iso-Pentane	GC-MS*	GC-MS	GC-FID	GC-MS	GC-MS*
n-Hexane	GC-MS*	GC-MS	GC-FID	GC-MS	GC-MS*
n-Heptane	GC-MS*	GC-MS	GC-FID	GC-MS	GC-MS*
n-Octane	GC-MS*	GC-MS		GC-MS	GC-MS*
n-Nonane	GC-MS*	GC-MS		GC-MS	GC-MS*
n-Decane	GC-MS*	GC-MS		GC-MS	
Undecane	GC-MS	GC-MS			

2-Methylpentane	GC-MS*	GC-MS	GC-FID	GC-MS	
3-Methylpentane	GC-MS*	GC-MS	GC-FID	GC-MS	
2,2-Dimethylbutane	GC-MS*	GC-MS	GC-FID	GC-MS	
2,3-Dimethylbutane	GC-MS*				
2-Methylhexane	GC-MS*				
3-Methylhexane				GC-MS	
Cyclohexane	GC-MS*	GC-MS	PTR-MS[1]	GC-MS	
Neopentane	GC-MS*	GC-MS	GC-FID		
Aromatics					
Benzene	GC-MS*	GC-MS	GC-FID	GC-MS	GC-MS*
Toluene	GC-MS*	GC-MS	PTR-MS	GC-MS	GC-MS*
Ethyl Benzene	GC-MS*	GC-MS	PTR-MS[1]	GC-MS	GC-MS*
n-Propyl Benzene	GC-MS*	GC-MS	PTR-MS[1]	GC-MS	
iso-Propyl Benzene	GC-MS*	GC-MS	PTR-MS[1]	GC-MS	
m-Xylene	GC-MS*	GC-MS	PTR-MS[1]	GC-MS	GC-MS*
p-Xylene	GC-MS*	GC-MS		GC-MS	GC-MS*
o-Xylene	GC-MS*	GC-MS	PTR-MS[1]	GC-MS	GC-MS*
1,3,5-Trimethylbenzene	GC-MS*	GC-MS	PTR-MS[1]	GC-MS	
1,2,3-Trimethylbenzene	GC-MS*	GC-MS	PTR-MS[1]	GC-MS	
1,2,4-Trimethylbenzene	GC-MS*	GC-MS	PTR-MS[1]	GC-MS	
m-Ethyltoluene	GC-MS*			GC-MS	
p-Ethyltoluene	GC-MS*			GC-MS	
o-Ethyltoluene	GC-MS*	GC-MS	PTR-MS[1]	GC-MS	
Styrene	GC-MS*	GC-MS			
Alkenes and Alkynes					
Ethyne (Acetylene)	GC-MS*	GC-MS	GC-FID	GC-MS	GC-MS*
Ethene	GC-MS*	GC-MS	GC-FID	GC-MS	GC-MS*
Propene	GC-MS*	GC-MS	GC-FID	GC-MS	GC-MS*
1-Butene	GC-MS*				GC-MS*
trans-2-Butene	GC-MS*				GC-MS*
cis-2-Butene	GC-MS*				GC-MS*
iso-Butene					GC-MS*
1-Pentene	GC-MS*				
trans-2-Pentene	GC-MS*				
cis-2-Pentene	GC-MS*				
2-Methyl-1-Butene	GC-MS*				
2-Methyl-2-Butene	GC-MS*				
1,3-Butadiene		GC-MS			
Aldehydes and Ketones					
Acetone	GC-MS	GC-MS	PTR-TOF-MS	GC-MS	PTR-MS
2-Butanone (Methyl Ethyl	GC-MS	GC-MS	PTR-TOF-MS	GC-MS	PTR-MS

Ketone)					
Formaldehyde	PTR-MS	PTR-MS	PTR-MS	PTR-TOF-MS	
Acetaldehyde	GC-MS	GC-MS	PTR-TOF-MS	GC-MS	PTR-MS
Propanal	GC-MS	GC-MS		GC-MS	
Butanal		GC-MS		GC-MS	
Hexanal		GC-MS		GC-MS	
2-Propenal (Acrolein)					
Benzaldehyde		GC-MS			
Alcohols					
Methanol	GC-MS	GC-MS	PTR-TOF-MS	GC-MS	PTR-MS
Ethanol	GC-MS	GC-MS			
iso-Propanol	GC-MS				
1-Butanol	GC-MS				
Biogenics					
Isoprene	GC-MS	GC-MS			GC-MS*
Methyl Vinyl Ketone	GC-MS	GC-MS	PTR-TOF-MS	PTR-TOF-MS	PTR-MS
Methacrolein	GC-MS	GC-MS			PTR-MS
α-Pinene	GC-MS*		PTR-TOF-MS		GC-MS*
β-Pinene	GC-MS*		PTR-TOF-MS		GC-MS*
Limonene	GC-MS*				
Alkyl Nitrates					
Methyl Nitrate	GC-MS*	GC-MS			GC-MS*
Ethyl Nitrate	GC-MS*	GC-MS		GC-MS	GC-MS*
n-Propyl Nitrate	GC-MS*	GC-MS		GC-MS	GC-MS*
iso-Propyl Nitrate	GC-MS*	GC-MS		GC-MS	GC-MS*
n-Butyl Nitrate				GC-MS	
2-Butyl Nitrate	GC-MS*				GC-MS*

*indicated that VOCs were collected with canister first and then measured with GC-MS.

References

Edwards, P.M., Brown, S.S., Roberts, J.M., Ahmadov, R., Banta, R.M., DeGouw, J. A., Dubé, W.P., Field, R. A., Flynn, J.H., Gilman, J.B., Graus, M., Helmig, D., Koss, A., Langford, A.O., Lefer, B.L., Lerner, B.M., Li, R., Li, S.-M., McKeen, S. A., Murphy, S.M., Parrish, D.D., Senff, C.J., Soltis, J., Stutz, J., Sweeney, C., Thompson, C.R., Trainer, M.K., Tsai, C., Veres, P.R., Washenfelder, R. a., Warneke, C., Wild, R.J., Young, C.J., Yuan, B., Zamora, R: High winter ozone pollution from carbonyl photolysis in an oil and gas basin. Nature, 514, 351-354, https://doi.org/10.1038/nature13767, 2014.

Hauglustaine, D.A., Granier, C., Brasseur, G.P., and Mégie, G.: The importance of atmospheric chemistry in the calculation of radiative forcing on the climate system. Journal of Geophysical Research, 99, doi:10.1029/93JD02987.issn:0148-0227, 1994.

Knote, C., Tuccella, P., Curci, G., Emmons, L., Orlando, J.J., Madronich, S., Baró, R., Jiménez-Guerrero, P., Luecken, D., Hogrefe, C., Forkel, R., Werhahn, J., Hirtl, M., Pérez, J.L., San José, R., Giordano, L., Brunner, D., Yahya, K., Zhang, Y.: Influence of the choice of gas-phase mechanism on predictions of key gaseous pollutants during the AQMEII phase-2 intercomparison. Atmospheric Environment, 115, 553-568, https://doi.org/10.1016/j.atmosenv.2014.11.066, 2015.

Levelt, P.F., Hilsenrath, E., Leppelmeier, G.W., Oord, G.H.J. van den, Bhartia, P.K., Tamminen, J., Haan, J.F., and de Veefkind, J.P.: Science objectives of the ozone monitoring instrument. IEEE Trans. Geosci. Remote Sens. 44, 1199-1208, https://doi.org/10.1109/TGRS.2006.872336, 2006.