

# Changes in the frequency and return level of high ozone pollution events over the eastern United States following emission controls

H E Rieder<sup>1,2</sup>, A M Fiore<sup>1,3</sup>, L M Polvani<sup>1,2,3</sup>, J-F Lamarque<sup>4</sup> and Y Fang<sup>5</sup>

<sup>1</sup> Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY, USA

<sup>2</sup> Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY, USA

<sup>3</sup> Department of Earth and Environmental Sciences, Columbia University, New York, NY, USA

<sup>4</sup> National Center for Atmospheric Research, Boulder, CO, USA

<sup>5</sup> Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, NJ, USA

E-mail: [hrieder@ldeo.columbia.edu](mailto:hrieder@ldeo.columbia.edu)

Received 19 September 2012

Accepted for publication 8 January 2013

Published 25 January 2013

Online at [stacks.iop.org/ERL/8/014012](http://stacks.iop.org/ERL/8/014012)

## Abstract

In order to quantify the impact of recent efforts to abate surface ozone ( $O_3$ ) pollution, we analyze changes in the frequency and return level of summertime (JJA) high surface  $O_3$  events over the eastern United States (US) from 1988–1998 to 1999–2009. We apply methods from extreme value theory (EVT) to maximum daily 8-hour average ozone (MDA8  $O_3$ ) observed by the Clean Air Status and Trends Network (CASTNet) and define  $O_3$  extremes as days on which MDA8  $O_3$  exceeds a threshold of 75 ppb (MDA8  $O_3 > 75$ ). Over the eastern US, we find that the number of summer days with MDA8  $O_3 > 75$  declined on average by about a factor of two from 1988–1998 to 1999–2009. The applied generalized Pareto distribution (GPD) fits the high tail of MDA8  $O_3$  much better than a Gaussian distribution and enables the derivation of probabilistic return levels (describing the probability of exceeding a value  $x$  within a time window  $T$ ) for high  $O_3$  pollution events. This new approach confirms the significant decline in both frequency and magnitude of high  $O_3$  pollution events over the eastern US during recent years reported in prior studies. Our analysis of 1-yr and 5-yr return levels at each station demonstrates the strong impact of changes in air quality regulations and subsequent control measures (e.g., the ‘NO<sub>x</sub> SIP Call’), as the 5-yr return levels of the period 1999–2009 correspond roughly to the 1-yr return levels of the earlier time period (1988–1998). Regionally, the return levels dropped between 1988–1998 and 1999–2009 by about 8 ppb in the Mid-Atlantic (MA) and Great Lakes (GL) regions, while the strongest decline, about 13 ppb, is observed in the Northeast (NE) region. Nearly all stations (21 out of 23) have 1-yr return levels well below 100 ppb and 5-yr return levels well below 110 ppb in 1999–2009. Decreases in eastern US  $O_3$  pollution are largest after full implementation of the nitrogen oxide (NO<sub>x</sub>) reductions under the ‘NO<sub>x</sub> SIP Call’. We conclude that the application of EVT methods provides a useful approach for quantifying return levels of high  $O_3$  pollution in probabilistic terms, which may help to guide long-term air quality planning.

**Keywords:** ozone, air pollution, air quality, United States, extreme value theory



Content from this work may be used under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 licence](http://creativecommons.org/licenses/by-nc-sa/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

## 1. Introduction

In surface air, ozone ( $O_3$ ) concentrations affect human health (e.g., Bell *et al* 2005, 2004, Confalonieri *et al* 2007, US EPA

2008), the environment (e.g., Arneeth *et al* 2010, Confalonieri *et al* 2007, Mauzerall and Wang 2001) and climate (e.g., Hegerl *et al* 2007), motivating a need to better understand the drivers of O<sub>3</sub> variability and trends, including changes in the frequency and magnitude of high O<sub>3</sub> pollution events.

Tropospheric O<sub>3</sub> is produced through a series of chemical reactions involving sunlight and O<sub>3</sub> precursors such as volatile organic compounds (VOCs), oxides of nitrogen (NO<sub>x</sub>), carbon monoxide (CO) and methane (CH<sub>4</sub>). Those precursors are emitted from anthropogenic (e.g., traffic and industry) and natural (e.g., soil, trees, lightning) sources (e.g., US EPA 2008, NRC 1991). Determining the impact on observed surface O<sub>3</sub> changes due to changes in anthropogenic and/or natural emissions versus meteorological variability is complicated (e.g., Porter *et al* 2001, Bloomer *et al* 2010).

In particular, atmospheric dynamics strongly influences the inter-annual variability of high O<sub>3</sub> pollution episodes (e.g., Leibensperger *et al* 2008, Liao *et al* 2009). Over the eastern United States (US), as over Europe, observations demonstrate that high O<sub>3</sub> pollution events are typically associated with summertime stagnation events (heat waves) (e.g., Logan 1989, Ordóñez *et al* 2010, Tai *et al* 2010, Vautard *et al* 2005, Vieno *et al* 2010, Vukovich 1995). Therefore a period of hot and dry days can significantly enhance O<sub>3</sub> levels (including through feedbacks from the biosphere by altering biogenic emissions and/or deposition) even if anthropogenic emissions of O<sub>3</sub> precursors do not increase (US EPA 2008).

In the US, the thresholds for the National Ambient Air Quality Standard (NAAQS) for O<sub>3</sub>, intended to protect public health and welfare, have been lowered repeatedly over the past few decades, i.e., becoming more stringent. In 1997 the US Environmental Protection Agency (EPA) set the NAAQS for maximum daily 8-hour average ozone (MDA8 O<sub>3</sub>) to 84 ppb. Following these changes, substantial NO<sub>x</sub> emission controls were implemented (i.e., the NO<sub>x</sub> State Implementation Plan, from here on referred to as 'NO<sub>x</sub> SIP Call'), leading to ~50% decreases in eastern US NO<sub>x</sub> emissions from point sources between 1999 and 2004 (e.g., Frost *et al* 2006). In 2008 the EPA further lowered the MDA8 O<sub>3</sub> NAAQS threshold to the current value of 75 ppb.

Several studies have determined trends in air pollutants by applying a variety of standard statistical models (e.g., Barmpadimos *et al* 2011, Bruno *et al* 2004, Holland *et al* 1999, Hooyberghs *et al* 2005, Zolghadri and Cazaurang 2006) or extreme value (EVT) methods (e.g., Eastoe 2009, Eastoe and Tawn 2009, Smith and Shively 1995). However, in the latter set of studies, air pollution data was mostly used as sample data in statistical applications, while in the former set, there is a notable lack of application of extreme statistics methodologies.

Here we bridge this gap by applying EVT methods to CASTNet data to quantify the extent to which emission controls, implemented to attain compliance with lower NAAQS thresholds have lowered the frequency of ozone extremes (i.e., days with MDA8 O<sub>3</sub> above 75 ppb, hereafter referred to as MDA8 O<sub>3</sub> > 75). Our approach extends earlier studies reporting a declining frequency of mean values and high air pollution events following emission controls over

the eastern US (Bloomer *et al* 2010, 2009, Fang *et al* 2010, Koumoutsaris and Bey 2012, Lin *et al* 2001, Porter *et al* 2001, Wolff *et al* 2001). Further, our work supports earlier air quality model projections of decreases in surface O<sub>3</sub> over the eastern United States resulting from NO<sub>x</sub> emission controls (e.g., Hogrefe and Rao 2001).

Application of EVT specifically allows us to derive probabilistic return levels for ozone extremes above the NAAQS. The T-year return level  $R^T$ , describes the probability of exceeding a value  $x$  within a time window  $T$  (see section 2), and serves as an illustrative metric for quantifying the decreases in eastern US O<sub>3</sub> pollution due to effective emission controls. This simple metric could be considered in air quality planning, and for assessing historical and future changes in extreme O<sub>3</sub> events.

## 2. Data and methods

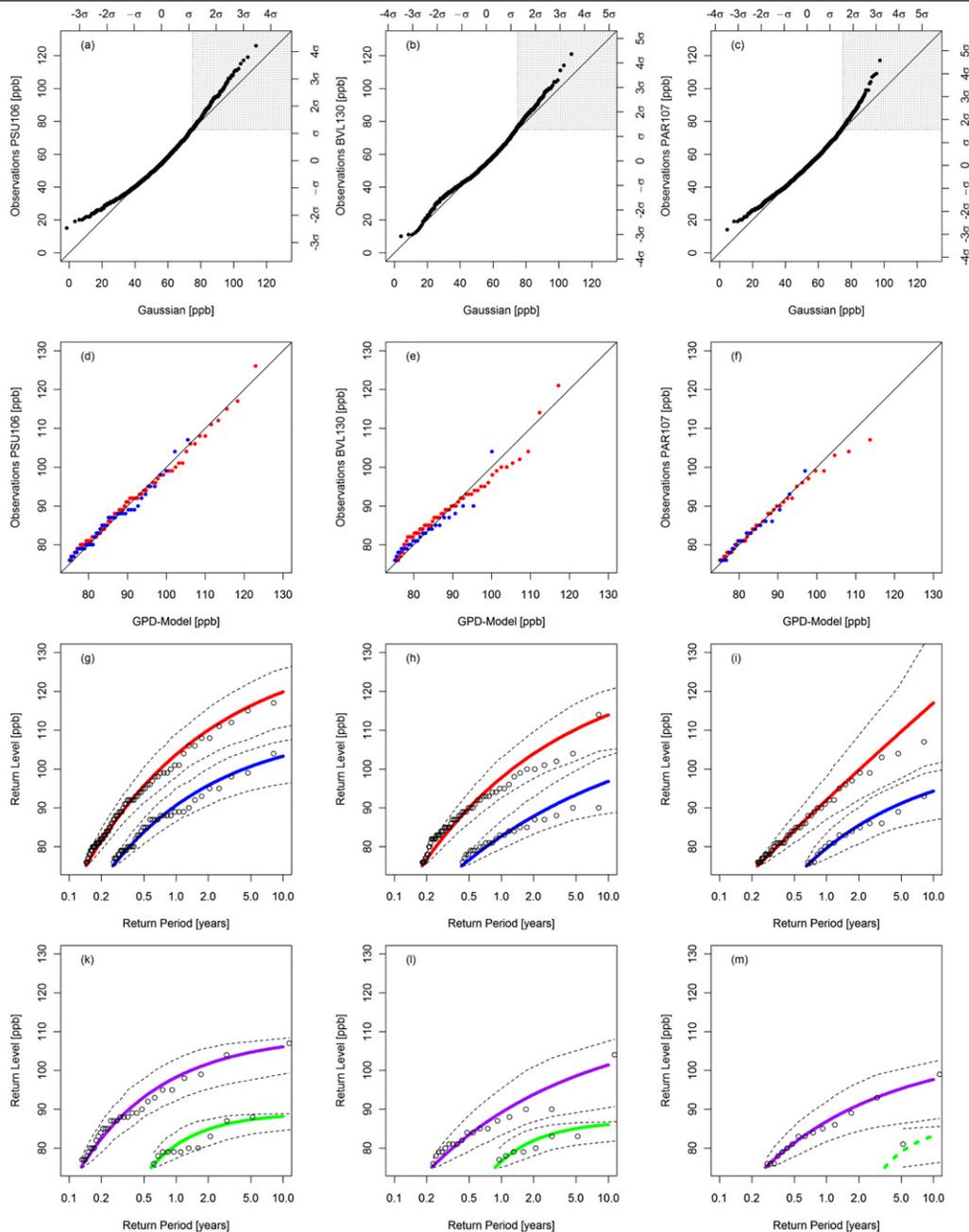
The EPA CASTNet network ([www.epa.gov/castnet](http://www.epa.gov/castnet)) has been operating since 1987 (Clarke *et al* 1997) and comprises currently more than 80 stations within the contiguous United States, Alaska and Canada (US EPA 2008). MDA8 O<sub>3</sub> data is obtained from the EPA's CASTNet data archive (<http://java.epa.gov/castnet/clearsession.do>). The CASTNet sites are representative of rural background concentrations. As the individual sites should not be directly influenced by large changes in local emission sources, CASTNet provides a regionally representative view of US surface O<sub>3</sub>, against which future projections and emission controls can be assessed.

We focus on the eastern US and analyze (i) the number of days with MDA8 O<sub>3</sub> > 75 ppb, (ii) changes in the frequency of extreme events during 1988–1998, 1999–2009 and 2003–2009, and (iii) changes in the return levels of high O<sub>3</sub> events at individual stations and on a broader regional basis. For the regional analysis we follow the selection and spatial grouping of Rasmussen *et al* (2012), splitting the eastern US into three regions of interest: the Great Lakes (GL), the Mid-Atlantic (MA) region and the Northeast (NE). Due to varying record lengths at individual sites, we restrict the analysis to sites with at least twenty years of data within 1988–2009, which limits the dataset to 23 stations within the eastern US (table 1).

We begin by selecting representative stations for the three eastern US regions (GL, MA, NE) based on summertime (JJA) mean MDA8 O<sub>3</sub> and standard deviations (to describe temporal variability) and the number of days with MDA8 O<sub>3</sub> > 75 ppb. Stations are considered to be regionally representative if (i) mean MDA8 O<sub>3</sub> is within ±6 ppb of the regional average, (ii) the standard deviation is within ±3 ppb of the regional average and (iii) the number of days with MDA8 O<sub>3</sub> > 75 ppb is close to the regional average, for both the 1988–1998 and 1999–2009 time periods. The results of the statistical analysis (see table 1) show that Bondville (BVL130) for the GL region, Penn State (PSU106) for the NE region and Parsons (PAR107) for the MA region are regionally representative, roughly consistent with earlier analyses of Reidmiller *et al* (2009) and Rasmussen *et al* (2012). We select these three stations to illustrate the application of extreme value modeling, before applying the methods to all 23 sites.

**Table 1.** Geographic information (site ID, latitude and longitude) and O<sub>3</sub> statistics at Clean Air Status and Trends Network (CASTNet) sites meeting the selection criteria of at least 20 years of data between 1988 and 2009. All statistics are calculated from summertime (JJA) maximum daily 8-hour average (MDA8) ozone (O<sub>3</sub>): mean values—with standard deviations (SD) given in parentheses, number (#) and percentage (%) of days with MDA8 O<sub>3</sub> above the National Ambient Air Quality Standard (NAAQS) for O<sub>3</sub> of 75 ppb (MDA8 O<sub>3</sub> > 75) are shown separately for 1988–1998, 1999–2009 and 2003–2009. \* indicates regional representative stations. Regional values for the Northeast (NE), Great Lakes (GL) and Mid-Atlantic (MA) regions are calculated as the average values over all stations within the region.

Region/station (state)	Site ID	Latitude (°N)	Longitude (°W)	Mean MDA8 O <sub>3</sub> (SD)			# days MDA8 O <sub>3</sub> > 75			% days MDA8 O <sub>3</sub> > 75		
				1988– 1998	1999– 2009	2003– 2009	1988– 1998	1999– 2009	2003– 2009	1988– 1998	1999– 2009	2003– 2009
Northeast				54.2 (16.2)	47.8 (14.1)	45.3 (12.2)	137	62	17	14.8	6.4	2.7
Arendtsville (PA)	ARE128	39.9	77.3	64.5 (18.0)	56.8 (15.0)	54.0 (12.5)	241	100	21	26.1	10.3	3.4
Ashland (ME)	ASH135	46.6	68.4	37.6 (11.1)	33.6 (10.4)	31.9 (9.5)	6	3	0	0.7	0.3	0.0
Connecticut Hill (NY)	CTH110	42.4	76.7	56.0 (15.3)	48.7 (13.2)	46.2 (11.4)	110	35	8	11.2	3.8	1.3
M.K. Goddard (PA)	MKG113	41.4	80.1	59.8 (18.4)	51.9 (15.1)	49.5 (13.4)	189	75	20	19.7	7.8	3.3
Penn State * (PA)	PSU106	40.7	77.9	59.7 (17.6)	52.9 (14.9)	49.8 (12.7)	164	77	17	17.2	7.7	2.7
Washington Crossing (NJ)	WSP144	40.3	74.9	62.9 (21.2)	56.3 (19.4)	52.8 (16.0)	236	144	50	27.1	14.5	7.9
Woodstock (NH)	WST109	43.9	71.7	38.7 (11.8)	34.8 (11.0)	33.2 (10.0)	12	1	0	1.3	0.1	0.0
Great Lakes				57.1 (15.2)	51.9 (13.5)	49.5 (12.4)	129	60	18	14.0	6.2	2.9
Alhambra (IL)	ALH157	38.9	89.6	60.9 (15.3)	55.9 (14.1)	53.7 (13.2)	166	94	38	18.0	9.4	6.0
Ann Arbor (MI)	ANA115	42.4	83.9	54.1 (16.4)	50.1 (14.7)	47.7 (13.4)	98	62	18	11.1	6.6	3.0
Bondville * (IL)	BVL130	40.1	88.4	59.4 (15.7)	52.4 (13.2)	50.0 (12.5)	154	57	18	16.1	5.8	2.8
Lykens (OH)	LYK123	40.9	83.0	59.4 (14.8)	54.1 (14.1)	51.2 (12.3)	115	77	18	14.6	7.7	2.8
Oxford (OH)	OXF122	39.5	84.7	62.6 (17.0)	54.2 (14.0)	51.3 (12.9)	218	72	22	22.8	7.4	3.5
Perkinstown (WI)	PRK134	45.2	90.6	42.6 (12.0)	41.6 (11.5)	40.5 (10.9)	3	9	2	0.3	1.2	0.4
Salamonie Reservoir (IN)	SAL133	40.8	85.7	57.6 (15.3)	52.7 (13.2)	50.0 (11.7)	116	51	10	12.8	5.1	1.6
Vincennes (IN)	VIN140	38.7	87.5	60.2 (15.2)	53.8 (13.1)	51.5 (12.2)	161	57	20	16.4	6.0	3.2
Mid-Atlantic				57.3 (15.5)	51.6 (14.2)	48.4 (12.2)	110	62	17	12.2	6.6	2.8
Beltsville (MD)	BEL116	39.0	76.8	65.1 (20.2)	59.1 (18.3)	56.3 (15.7)	273	166	65	31.1	17.3	11.2
Cedar Creek (WV)	CDR119	38.9	80.8	54.1 (15.5)	44.6 (12.6)	41.1 (10.6)	85	14	1	8.8	1.4	0.2
Edgar Evins (TN)	ESP127	36.0	85.7	53.7 (13.3)	49.4 (12.6)	46.7 (11.2)	45	20	3	5.2	2.1	0.5
Georgia Station (GA)	GAS153	33.2	84.4	56.8 (18.7)	52.6 (17.7)	50.1 (15.6)	157	103	33	17.2	11.4	5.3
Parsons * (WV)	PAR107	39.1	79.7	57.1 (14.6)	49.3 (12.0)	45.9 (10.1)	96	29	4	9.6	2.9	0.6
Prince Edward (VA)	PED108	37.2	78.3	56.8 (14.5)	49.8 (13.0)	46.5 (10.8)	80	32	2	9.8	3.3	0.3
Sand Mountain (AL)	SND152	34.3	86.0	58.2 (15.2)	54.9 (14.7)	51.1 (13.0)	95	86	21	10.8	10.0	3.4
Speedwell (TN)	SPD111	36.5	83.8	56.3 (11.8)	53.2 (12.6)	49.6 (10.4)	46	45	3	5.3	4.5	0.5



**Figure 1.** Quantile–quantile plot comparing summertime (JJA) maximum daily 8-hour average (MDA8) ozone ( $O_3$ ) observations in 1988–2009 with a Gaussian distribution (least-mean-square fitted to the observations) for the regional representative sites (a) Penn State (PSU106), (b) Bondville (BVL130), (c) Parsons (PAR107). Quantile–quantile comparison between observed and GPD model predicted extremes (MDA8  $O_3 > 75$ ) in 1988–1998 (red) and 1999–2009 (blue) for (d) Penn State (PSU106), (e) Bondville (BVL130), (f) Parsons (PAR107). Return level estimates from the fitted GPD for  $O_3$  extremes (MDA8  $O_3 > 75$ ) in 1988–1998 (red) and 1999–2009 (blue) at (g) Penn State (PSU106), (h) Bondville (BVL130), (i) Parsons (PAR107). (k)–(m) as (g)–(i) but for 1999–2002 (purple) and 2003–2009 (green). The gray squares in (a)–(c) mark the area above the 75 ppb National Ambient Air Quality Standard (NAAQS) for  $O_3$ . Straight solid lines in (a)–(f) give the identity line. Red (1988–1998) and blue (1999–2009) curves in (g)–(i) and purple (1999–2002) and green (2003–2009) curves in (k)–(m) give the maximum-likelihood estimated return levels, while black dashed lines mark the corresponding 95% confidence intervals and open circles mark actual events in the individual station records. Note: 2003–2009 return level for PAR107 (panel m) is shown dashed as estimation accuracy is limited by the small fraction of days above the NAAQS during this time period (see table 1). Scale parameter ( $\tau$ ) and shape parameter ( $\xi$ ) estimates are for 1988–1998: PSU106  $\tau = 20.41$ ,  $\xi = -0.35$ ; BVL130  $\tau = 17.25$ ,  $\xi = -0.32$ ; PAR107  $\tau = 11.61$ ,  $\xi = -0.12$ ; for 1999–2009: PSU106  $\tau = 15.00$ ,  $\xi = -0.41$ ; BVL130  $\tau = 9.92$ ,  $\xi = -0.24$ ; PAR107  $\tau = 11.45$ ,  $\xi = -0.38$ ; for 1999–2002: PSU106  $\tau = 19.10$ ,  $\xi = -0.56$ ; BVL130  $\tau = 11.78$ ,  $\xi = -0.30$ ; PAR107  $\tau = 11.30$ ,  $\xi = -0.36$  and for 2003–2009: PSU106  $\tau = 13.78$ ,  $\xi = -0.98$ ; BVL130  $\tau = 12.43$ ,  $\xi = -1.03$ ; PAR107  $\tau = 13.97$ ,  $\xi = -1.27$ .

Figures 1(a)–(c) show that the tails of surface ozone data are clearly non-Gaussian (i.e., data from a Gaussian distribution would lie close to the diagonal line). Proper

analysis of the tails requires extreme value theory (EVT) (e.g., Coles 2001, Coles and Pericchi 2003, Davison and Smith 1990). Peak-over threshold models (POT models), based on

the generalized Pareto distribution (GPD), are frequently applied as they make more efficient use of the information contained in a dataset (e.g., Coles 2001, Davison and Smith 1990) than other types of more selective EVT models (i.e., block-maximum models). In our case, the NAAQS of 75 ppb provides a policy relevant threshold for the POT model, and additionally motivates our choice of this type of extreme value model.

The use of the GPD for modeling exceedances over a high (enough) threshold is justified by arguments on the asymptotic behavior of the data (Pickands 1975) because the GPD is the limiting distribution of a normalized exceedance over a threshold, as the threshold approaches the maximum of the distribution (e.g., Coles 2001). The GPD ( $F(x)$ ) is defined as:

$$F(x) = 1 - \left[ 1 + \xi \frac{x - u}{\tau} \right]^{-1/\xi},$$

$$\tau > 0, x > u, 1 + \xi \frac{x - u}{\tau} > 0,$$

where  $x$  are the observations (here JJA MDA8 O<sub>3</sub> from the individual CASTNet sites)  $u$  is the threshold (here the NAAQS threshold of 75 ppb), and  $\tau$  and  $\xi$  are the scale (describing the spread of the distribution) and shape parameters, respectively. The R ([www.r-project.org](http://www.r-project.org)) POT-package (Ribatet 2007) is used for the EVT analysis in which the GPD parameters are computed by maximum-likelihood estimation. Figures 1(d)–(f) show the improved fit of the GPD to high ozone levels (above the NAAQS), compared to a Gaussian distribution (see figure 1 and values for scale ( $\tau$ ) and shape ( $\xi$ ) parameters in the caption).

The T-year return level  $R^T$  is given as

$$R^T = F_{\xi, u, \tau}^{-1} \left( 1 - \frac{1}{T} \right)$$

and can be directly calculated from the GPD.  $R^T$  is of practical interest as it describes the probability of exceeding a value  $x$  within a time window  $T$ .

### 3. Changes in the frequency of high O<sub>3</sub> events between 1988–1998 and 1999–2009

The high tail of the surface MDA8 O<sub>3</sub> distribution can be adequately fitted by a GPD (figures 1(d)–(f)). In addition, the O<sub>3</sub> NAAQS threshold of 75 ppb is just slightly larger than 1 standard deviation (at almost all 23 stations analyzed, see figures 1(a)–(c) and table 1), a statistical measure frequently used to characterize extreme levels.

During the 1988–1998 time period at 2 out of 7 stations in the NE region (ARE128, WSP144), and 1 out of 8 stations in the GL region (OXF122) and MA region (BEL116) 20% (or more) of all summer days were MDA8 O<sub>3</sub> > 75 (table 1). If we restrict our analysis to stations with an average of 10% (or more) extreme days during June–August in the 1988–1998 period these numbers increase to 5 out of 7 stations in the NE region (ARE128, CTH110, MKG113, PSU116, WSP144), 7 out of 8 stations in the GL region (ALH157, ANA115, BVL130, LYK123, OXF122, SAL133, VIN140) and 3 out of 8 stations in the MA region (BEL116, GAS153, SND152).

The three stations with the largest fraction of summer days with MDA8 O<sub>3</sub> > 75 for the 1988–1998 period are ARE128, BEL116 and WSP144 (>25% or more). The smallest fraction (<1%) occurred at ASH135 and PRK134, which are the most remote and northern sites in the CASTNet data base for the eastern US.

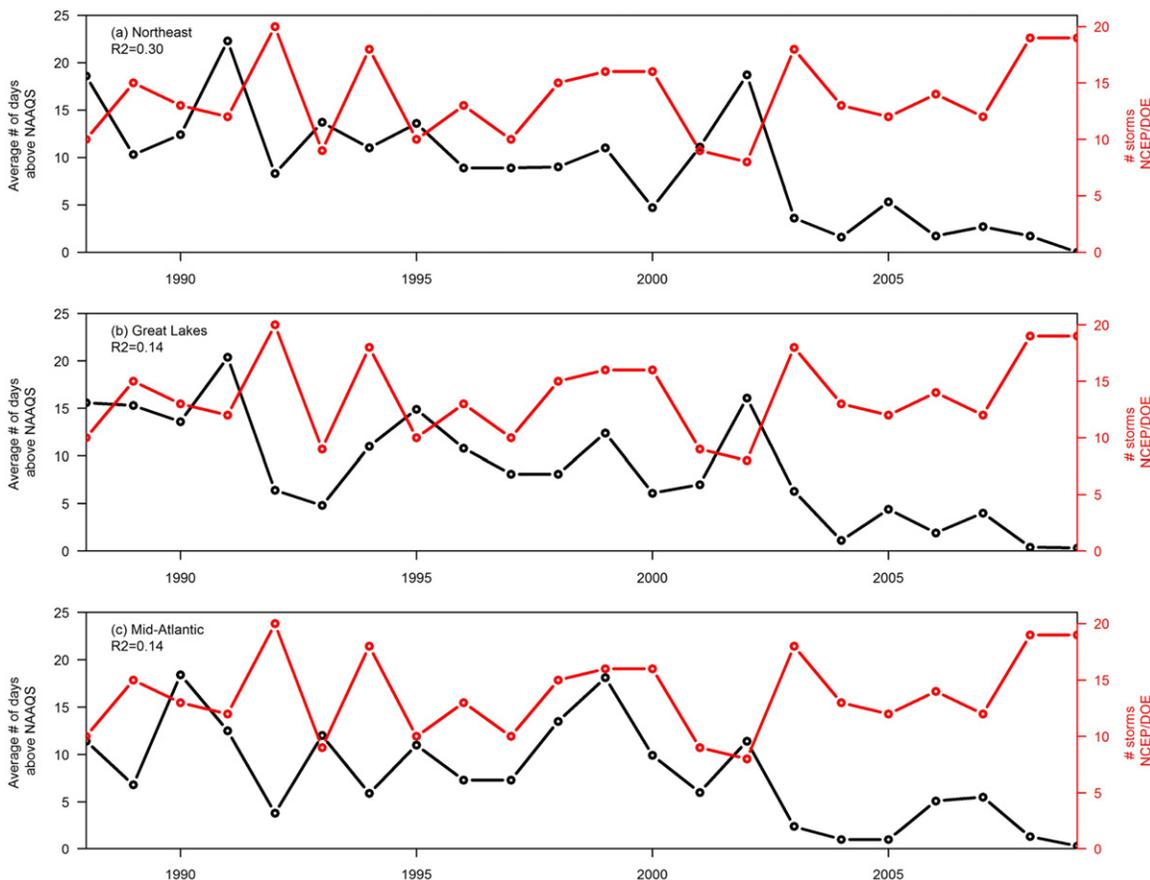
Following the implementation of the ‘NO<sub>x</sub> SIP Call’, the frequency of days with MDA8 O<sub>3</sub> > 75 decreases substantially; notably from 1988–1998 to 1999–2009. In the earlier period the average number of days with MDA8 O<sub>3</sub> > 75 was 137 in the NE region, 129 in the GL region and 110 in the MA region, those numbers decreased to 62, 60, and 62, respectively, in the latter period (table 1). This corresponds to a decrease in days with MDA8 O<sub>3</sub> > 75 by about a factor of 2 in all regions. By 1999–2009, no station remains with more than 20% of all days with MDA8 O<sub>3</sub> > 75, and a total of only 5 out of 23 stations now have an occurrence rate of extremes of 10% or higher (NE: ARE128 and WSP144; MA: BEL116, GAS153 and SND152).

The implementation of the ‘NO<sub>x</sub> SIP Call’ led to a strong decline in O<sub>3</sub> precursor emissions in the eastern US (Napolitano *et al* 2007), with the NO<sub>x</sub> reductions phased in between 1999 and 2002. Therefore we additionally split the time period 1999–2009 in an early (1999–2002) and late (2003–2009) time window to analyze changes in eastern US O<sub>3</sub> pollution following the ‘NO<sub>x</sub> SIP Call’.

The statistical analysis (see table 1) shows that on a regional basis during 1999–2009 more than half of the days (between 53 and 58%) with MDA8 O<sub>3</sub> > 75 occurred during the early phase of the ‘NO<sub>x</sub> SIP Call’ (1999–2002). At over half of the individual stations, two thirds or more of the days with MDA8 O<sub>3</sub> > 75 during 1999–2009 occurred during the early period while NO<sub>x</sub> reductions were still being implemented.

Our results show that recent efforts to abate surface O<sub>3</sub> pollution, e.g., the ‘NO<sub>x</sub> SIP Call’, have been highly effective in the eastern US, evident from the strong decline in the number and frequency of days with MDA8 O<sub>3</sub> > 75 (table 1). This decline occurs, however, in the context of substantial year-to-year variability in the number of days with MDA8 O<sub>3</sub> > 75 (figure 2; see also Leibensperger *et al* (2008)). Over the Northeastern US, recent work indicates that the frequency of storms moving through the region is a key driver of the inter-annual variability of high O<sub>3</sub> pollution episodes (Leibensperger *et al* 2008).

To analyze the contribution of these storms to the inter-annual variability of the number of days with MDA8 O<sub>3</sub> > 75 we perform a correlation analysis. We correlate the frequency of storms moving through the Great Lakes Storm Track Region each summer as calculated in Turner *et al* (2013) from the NCEP/DOE Reanalysis with the regional summer average of days with MDA8 O<sub>3</sub> > 75 in this analysis for the 1988–2009 period (figure 2). Inter-annual variability in storm frequency can explain up to 30% of the inter-annual variability in days with MDA8 O<sub>3</sub> > 75, with the largest correlation occurring over the Northeastern US, as expected based on the work of Leibensperger *et al* (2008). Over the 21st century, changes in the frequency of mid-latitude



**Figure 2.** Annual average of the number (#) of summer (JJA) days with maximum daily 8-hour average (MDA8) ozone ( $O_3$ ) above the NAAQS of 75 ppb (black) in 1988–2009 and number (#) of mid-latitude storms crossing the Great Lakes Storm Track Region, derived from the NCEP/DOE Reanalysis (red) (Turner *et al* 2013) for (a) the Northeast region, (b) the Great Lakes region and (c) the Mid-Atlantic region.

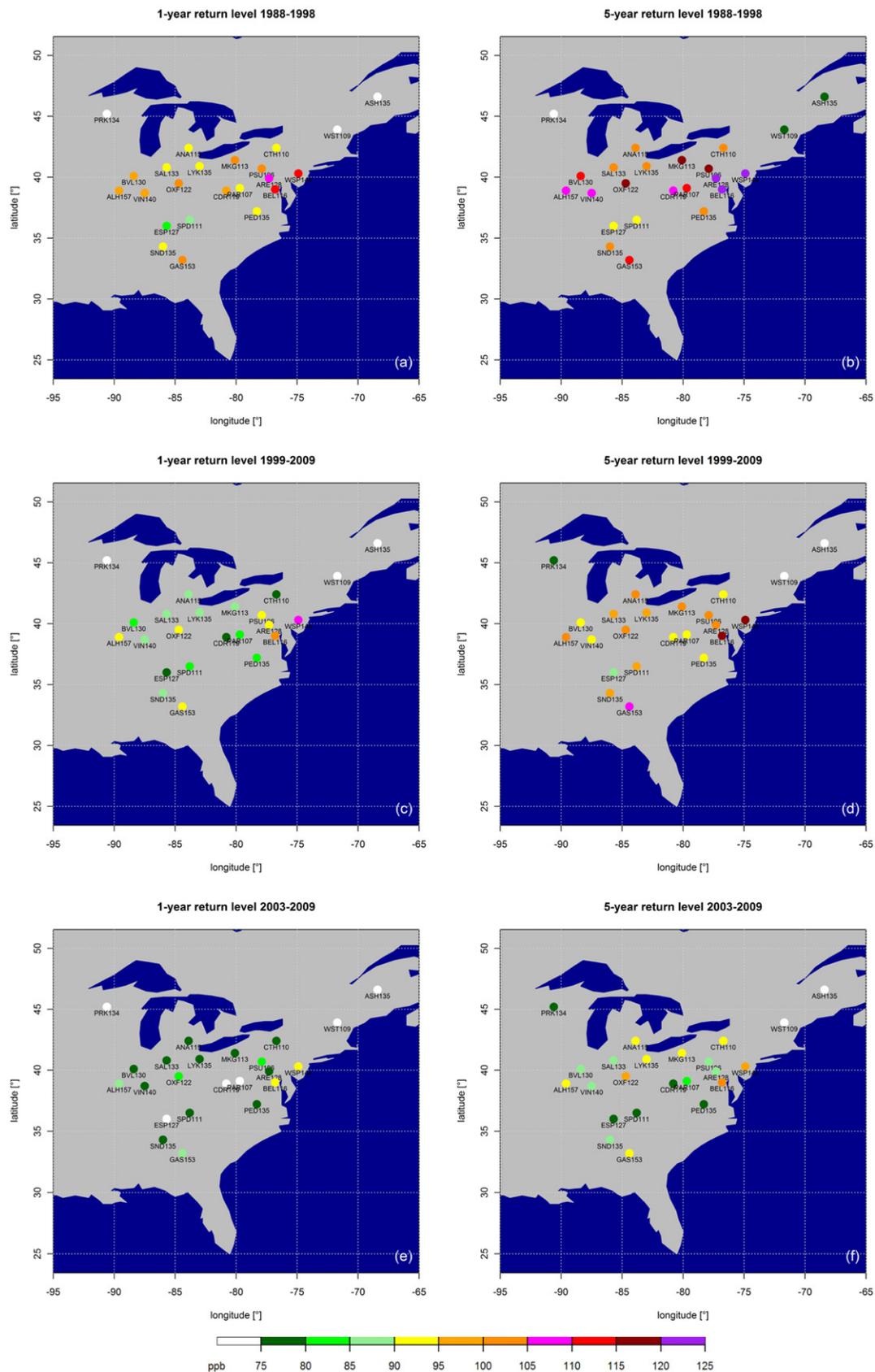
storms possibly driven by a warmer climate (e.g., Lang and Waugh 2011, Turner *et al* 2013) could contribute to changes in the inter-annual variability and/or long-term trend in the frequency of high  $O_3$  episodes.

Ambient air temperature is a major meteorological parameter influencing surface ozone levels (US EPA 2006); some of the relationship with temperature reflects its co-variation with air stagnation and is thus included in our correlation with storm passages. However, temperature can also directly affect the  $O_3$  production chemistry and precursor emissions. Specifically, Bloomer *et al* (2010) analyzed seasonal and diurnal variations and trends in  $O_3$  and temperature at 5 CASTNet sites (BEL116, CTH110, GAS153, PSU106 and WST109) to show that summer temperatures are warming during the time periods where ozone is decreasing. As it is well understood that ozone in polluted regions increases with increasing ambient air temperatures (Fiore *et al* (2012) and references therein; Jacob and Winner (2009)), the decreasing  $O_3$  trends over the Northeastern US are attributed to emission reductions rather than changes in weather and climate (Bloomer *et al* 2010). These findings are consistent with our conclusion that storm passages may contribute to inter-annual variability but that the downward ozone trend is driven by the  $NO_x$  emission reductions that were implemented between 1999 and 2002.

#### 4. Changes in the probabilistic return level of high $O_3$ events

As outlined in section 2, a theoretical return level of MDA8  $O_3$  for time  $T$  ( $R^T$ ), can be calculated directly from the fitted GPD at each station. We calculate,  $R^T$ , of MDA8  $O_3 > 75$  separately for the three regionally representative sites PSU106, BVL130 and PAR107 for 1988–1998 and 1999–2009, respectively (figures 1(g)–(i)). The probability of high ozone days strongly declines from 1988–1998 a value of 104 ppb occurred once a year, lengthening to once in 4–5 years for the later period. Similar changes in  $R^T$  occur at the other two stations. In figures 1(k)–(m) we analyze changes in the  $O_3$  return levels during 1999–2009 at each site by further separating into earlier (1999–2002) and later (2003–2009) periods. The return levels estimated over 1999–2002 are similar to those of the entire 1999–2009 time period, while those estimated for the 2003–2009 subset are substantially lower. We conclude that the emission controls implemented under the ‘ $NO_x$  SIP Call’ between 1999 and 2002 significantly decreased the return level and frequency of high  $O_3$  pollution events, particularly after 2002.

Finally, we extend the analysis illustrated for the three sites in figure 1 to our entire dataset, grouping MDA8  $O_3$



**Figure 3.** Return levels of summertime (JJA) maximum daily 8-hour average (MDA8) ozone ( $O_3$ ) for the 23 sites selected from the eastern United States Clean Air Status and Trends Network (CASTNet): (a) 1-yr return levels in 1988–1998, (b) as (a) but for 5-yr return levels, (c), (d) as (a), (b) but for 1999–2009, (e), (f) as (a), (b) but for 2003–2009.

return levels into 5 ppb bins, spanning from 75 to 125 ppb. Figure 3 shows the corresponding 1-yr and 5-yr return levels at each station for 1988–1998 (panels (a)–(b)), 1999–2009 (panels (c)–(d)) and 2003–2009 (panels (e)–(f)). The 1-yr return levels dropped between 2 and 16 ppb from 1988–1998 to 1999–2009 at the individual sites (see figures 3(a), (c)). Almost all stations (except for BEL116 (MA) and WSP144 (NE)) show 1-yr return levels well below 100 ppb for the 1999–2009 period. At the regional scale, we find an average decline of the 1-yr return level of about 8 ppb for the MA and GL regions while the strongest reductions occur at the NE stations where the return levels dropped on average by about 13 ppb.

The changes in the probabilistic 5-yr return levels (figures 3(b), (d)) essentially mirror those of the 1-yr return levels. At the individual stations the 5-yr MDA8 O<sub>3</sub> return levels dropped by a similar range as the 1-yr return levels. At some sites (ARE128, MKG113) the 5-yr return levels dropped by up to 18 ppb, and almost all stations (except for BEL116 (MA) and WSP144 (NE)) have 5-yr return levels well below 110 ppb for the 1999–2009 period.

The improvements achieved with the ‘NO<sub>x</sub> SIP Call’ for eastern US O<sub>3</sub> pollution are even more obvious when we compare the probabilistic return levels of the entire 1999–2009 time period (figures 3(c), (d)) with the 2003–2009 subset (figures 3(e), (f)). 1-yr return levels estimated over 2003–2009 are on average about 9 ppb in the NE region, 8 ppb in the MA region and 6 ppb in the GL region lower compared to those of the entire 1999–2009 time period. For the 5-yr return levels these changes are 10 ppb (NE), 7 ppb (MA) and 12 ppb (GL), respectively. However, the uncertainty in these estimates is larger than for return levels calculated over 1988–1998 and 1999–2009 due to the smaller number of days with MDA8 O<sub>3</sub> above the NAAQS.

In summary, both 1- and 5-yr return levels decreased over the eastern United States between 1988–1998 and 1999–2009, with the strongest decreases occurring during the latest years (2003–2009) of the record. The largest reductions in high O<sub>3</sub> concentrations occurred in the NE region, although the ambient concentrations and return levels are still slightly higher than those in the GL or MA regions.

## 5. Discussion and conclusions

We have applied extreme value methods to the eastern United States (US) Clean Air Status and Trends Network (CASTNet) maximum daily 8-hour average (MDA8) ozone (O<sub>3</sub>) data. Our goal is to assess whether tightening of the O<sub>3</sub> National Ambient Air Quality Standard (NAAQS) and subsequent regional emission controls (e.g., the ‘NO<sub>x</sub> SIP Call’) have led to significant changes in the frequency and probabilistic return level of high O<sub>3</sub> pollution events over the eastern US. While prior studies have demonstrated decreases in O<sub>3</sub> following the ‘NO<sub>x</sub> SIP Call’ (Bloomer *et al* 2010, 2009, Fang *et al* 2010, Koumoutsaris and Bey 2012, Lin *et al* 2001, Porter *et al* 2001, Wolff *et al* 2001), our methodology enables us to quantify probabilistic return levels, a simple metric that may be useful

for air quality planning and for communicating the success of emission controls.

We show that both the frequency and magnitude of high O<sub>3</sub> pollution events have significantly declined in recent years. On a regional basis, the number of days above the NAAQS of 75 ppb (MDA8 O<sub>3</sub> > 75) declined by about a factor of 2 from 1988–1998 to 1999–2009 throughout the three regions analyzed (Northeast (NE), Great Lakes (GL), Mid-Atlantic (MA)).

Application of methods from extreme value theory (EVT) shows that the generalized Pareto distribution (GPD) better fits the high tail of MDA8 O<sub>3</sub> than a Gaussian distribution. Further, fitting a GPD allows direct calculation of probabilistic return levels for high O<sub>3</sub> pollution events. Our analysis of 1-yr and 5-yr return levels at individual stations illustrates the strong impact of changes in the NAAQS and subsequent control measures (the ‘NO<sub>x</sub> SIP Call’) on air quality in the eastern US, as the 5-yr return levels for 1999–2009 correspond roughly to the 1-yr return levels of 1988–1998. This confirms earlier modeling work that simulated improved O<sub>3</sub> air quality due to NO<sub>x</sub> emission controls (e.g., Hogrefe and Rao 2001). The decreases in extreme O<sub>3</sub> pollution events achieved by the ‘NO<sub>x</sub> SIP Call’ over the eastern US are even more pronounced when we compare the return levels of the 2003–2009 subset with those of the entire 1999–2009 time period. The 1-yr return levels estimated over the last years of the record (2003–2009) are on average about 9 ppb in the NE region, 8 ppb in the MA region and 6 ppb in the GL region lower than those estimated over 1999–2009. The changes in 5-yr return levels are 10 ppb (NE), 7 ppb (MA) and 12 ppb (GL), respectively.

The potential sensitivity of air pollution to climate change is receiving attention, including how the frequency, duration and intensity of air pollution events will evolve during the course of the 21st century (e.g., the recent reviews of Fiore *et al* 2012, Isaksen *et al* 2009, Jacob and Winner 2009, Weaver *et al* 2009). Our results and methodology may serve as a possible benchmark for the evaluation and assessment of future changes in extreme O<sub>3</sub> pollution events as projected by chemistry-climate and chemistry-transport model simulations.

## Acknowledgments

This article was made possible by EPA-STAR grant 83520601. Its contents are solely the responsibility of the grantee and do not necessarily represent the official view of the EPA. Further, the EPA does not endorse the purchase of any commercial products or services mentioned in the publication. HER acknowledges also partial funding through a postdoctoral fellowship grant (no. PEZP2-134426) of the Swiss National Science Foundation (SNF). The National Center for Atmospheric Research (NCAR) is operated by the University Corporation for Atmospheric Research (UCAR) under sponsorship of the National Science Foundation (NSF). The authors are grateful to Alexander J Turner for providing data on the number of mid-latitude storms crossing the Great Lakes Storm Track Region, derived from NCEP/DOE

Reanalysis data. Helpful comments by two anonymous reviewers are greatly appreciated.

**References**

Arneeth A *et al* 2010 Terrestrial biogeochemical feedbacks in the climate system *Nature Geosci.* **3** 525–32

Barnpadimos I, Hueglin C, Keller J, Henne S and Prevot A S H 2011 Influence of meteorology on PM (10) trends and variability in Switzerland from 1991 to 2008 *Atmos. Chem. Phys.* **11** 1813–35

Bell M L, Dominici F and Samet J M 2005 A meta-analysis of time-series studies of ozone and mortality with comparison to the national morbidity, mortality, and air pollution study *Epidemiology* **16** 436–45

Bell M L, McDermott A, Zeger S L, Samet J M and Dominici F 2004 Ozone and short-term mortality in 95 US urban communities, 1987–2000 *Jama-J. Am. Med. Assoc.* **292** 2372–8

Bloomer B J, Stehr J W, Piety C A, Salawitch R J and Dickerson R R 2009 Observed relationships of ozone air pollution with temperature and emissions *Geophys. Res. Lett.* **36** L09803

Bloomer B J, Vinnikov K Y and Dickerson R R 2010 Changes in seasonal and diurnal cycles of ozone and temperature in the eastern US *Atmos. Environ.* **44** 2543–51

Bruno F, Cocchi D and Trivisano C 2004 Forecasting daily high ozone concentrations by classification trees *Environmetrics* **15** 141–53

Clarke J F, Edgerton E S and Martin B E 1997 Dry deposition calculations for the clean air status and trends network *Atmos. Environ.* **31** 3667–78

Coles S 2001 *An Introduction to Statistical Modeling of Extreme Values (Springer Series in Statistics)* (London: Springer)

Coles S and Pericchi L 2003 Anticipating catastrophes through extreme value modelling *J. R. Stat. Soc. Ser. C* **52** 405–16

Confalonieri U, Menne B, Akhtar R, Ebi K L, Hauengue M, Kovats R S, Revich B and Woodward A 2007 Human health *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* ed M L Parry, O F Canziani, J P Palutikof, P J van der Linden and C E Hanson (Cambridge: Cambridge University Press) pp 391–431

Davison A C and Smith R L 1990 Models for exceedances over high thresholds *J. R. Stat. Soc. Ser. B* **52** 393–442 (with discussion)

Eastoe E F 2009 A hierarchical model for non-stationary multivariate extremes: a case study of surface-level ozone and NO<sub>x</sub> data in the UK *Environmetrics* **20** 428–44

Eastoe E F and Tawn J A 2009 Modelling non-stationary extremes with application to surface level ozone *J. R. Stat. Soc. Ser. C* **58** 25–45

Fang Y, Fiore A M, Horowitz L W, Levy H, Hu Y and Russell A G 2010 Sensitivity of the NO<sub>y</sub> budget over the United States to anthropogenic and lightning NO<sub>x</sub> in summer *J. Geophys. Res.—Atmos.* **115** D18312

Fiore A M *et al* 2012 Global air quality and climate *Chem. Soc. Rev.* **41** 6663–83

Frost G J *et al* 2006 Effects of changing power plant NO(x) emissions on ozone in the eastern United States: Proof of concept *J. Geophys. Res. Atmos.* **111** D12306

Hegerl G C, Zwiers F W, Branconnot P, Gillet N P, Luo Y, Marengo Orsini J A, Nicholls N, Penner J E and Stott P A 2007 Understanding and attributing climate change *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* ed S Solomon, D Qin, M Manning, Z Chen, M Marquis, K B Averyt, M Tignor and H L Miller (Cambridge: Cambridge University Press) pp 663–745

Hogrefe C and Rao S T 2001 Demonstrating attainment of the air quality standards: Integration of observations and model predictions into the probabilistic framework *J. Air Waste Manag. Assoc.* **51** 1060–72

Holland D M, Principe P P and Vorburger L 1999 Rural ozone: trends and exceedances at CASTNet sites *Environ. Sci. Technol.* **33** 43–8

Hooyberghs J, Mensink C, Dumont G, Fierens F and Brasseur O 2005 A neural network forecast for daily average PM10 concentrations in Belgium *Atmos. Environ.* **39** 3279–89

Isaksen I S A *et al* 2009 Atmospheric composition change: climate-chemistry interactions *Atmos. Environ.* **43** 5138–92

Jacob D J and Winner D A 2009 Effect of climate change on air quality *Atmos. Environ.* **43** 51–63

Koumoutsaris S and Bey I 2012 Can a global model reproduce observed trends in summertime surface ozone levels? *Atmos. Chem. Phys.* **12** 6983–98

Lang C and Waugh D W 2011 Impact of climate change on the frequency of Northern Hemisphere summer cyclones *J. Geophys. Res.—Atmos.* **116** D04103

Leibensperger E M, Mickley L J and Jacob D J 2008 Sensitivity of US air quality to mid-latitude cyclone frequency and implications of 1980–2006 climate change *Atmos. Chem. Phys.* **8** 7075–86

Liao H, Zhang Y, Cheng W T, Raes F and Seinfeld J 2009 Effect of chemistry-aerosol-climate coupling on predictions of future climate and future levels of tropospheric ozone and aerosols *J. Geophys. Res.—Atmos.* **114** D10306

Lin C Y C, Jacob D J and Fiore A M 2001 Trends in exceedances of the ozone air quality standard in the continental United States, 1980–1998 *Atmos. Environ.* **35** 3217–28

Logan J A 1989 Ozone in rural areas of the United States *J. Geophys. Res.* **94** 8511–32

Mauzerall D L and Wang X P 2001 Protecting agricultural crops from the effects of tropospheric ozone exposure: Reconciling science and standard setting in the United States, Europe, and Asia *Ann. Rev. Energy Environ.* **26** 237–68

Napolitano S, Stevens G, Schreifels J and Culligan K 2007 The NO<sub>x</sub> budget trading program: a collaborative, innovative approach to solving a regional air pollution problem *Electr. J.* **20** 65–76

NRC 1991 *Rethinking the Ozone Problem in Rural and Regional Air Pollution. Report* (Washington, DC: National Academy Press)

Ordóñez C *et al* 2010 Global model simulations of air pollution during the 2003 European heat wave *Atmos. Chem. Phys.* **10** 789–815

Pickands J 1975 Statistical-inference using extreme order statistics *Ann. Stat.* **3** 119–31

Porter P S, Rao S T, Zurbenko I G, Dunker A M and Wolff G T 2001 Ozone air quality over North America: part II—an analysis of trend detection and attribution techniques *J. Air Waste Manag. Assoc.* **51** 283–306

Rasmussen D J, Fiore A M, Naik V, Horowitz L W, McGinnis S J and Schultz M G 2012 Surface ozone-temperature relationships in the eastern US: a monthly climatology for evaluating chemistry-climate models *Atmos. Environ.* **47** 142–53

Reidmiller D R *et al* 2009 The influence of foreign vs. North American emissions on surface ozone in the US *Atmos. Chem. Phys.* **9** 5027–42

Ribatet M 2007 POT: modelling peaks over a threshold *R News* **7** 34–6 (<http://cran.r-project.org/doc/Rnews/Rnews.2007-1.pdf>)

- Smith R L and Shively T S 1995 Point process approach to modeling trends in tropospheric ozone based on exceedances of a high-threshold *Atmos. Environ.* **29** 3489–99
- Tai A P K, Mickley L J and Jacob D J 2010 Correlations between fine particulate matter (PM<sub>2.5</sub>) and meteorological variables in the United States: implications for the sensitivity of PM<sub>2.5</sub> to climate change *Atmos. Environ.* **44** 3976–84
- Turner A J, Fiore A, Horowitz L W, Naik V and Bauer M 2013 Summertime cyclones over the Great Lakes Storm Track from 1860–2100: variability, trends and association with ozone pollution *Atmos. Chem. Phys.* **13** 565–78
- US EPA 2006 *Air Quality Criteria for Ozone and Related Photochemical Oxidants* (EPA/600/R-05/004aF-cf) (Washington, DC: US Environmental Protection Agency)
- US EPA 2008 *Clean Air Status and Trends Network (CASTNET)—2008 Annual Report* (Contract No. EP-W-09-028) (Washington, DC: US Environmental Protection Agency)
- Vautard R, Honoré C, Beekmann M and Rouil L 2005 Simulation of ozone during the August 2003 heat wave and emission control scenarios *Atmos. Environ.* **39** 2957–67
- Vieno M *et al* 2010 Modelling surface ozone during the 2003 heat-wave in the UK *Atmos. Chem. Phys.* **10** 7963–78
- Vukovich F M 1995 Regional-scale boundary layer ozone variations in the eastern United States and their association with meteorological variations *Atmos. Environ.* **29** 2259–73
- Weaver C P *et al* 2009 A preliminary synthesis of modeled climate change impacts on us regional ozone concentrations *Bull. Am. Meteorol. Soc.* **90** 1843–63
- Wolff G T, Dunker A M, Rao S T, Porter P S and Zurbenko I G 2001 Ozone air quality over North America: part I—review of reported trends *J. Air Waste Manag. Assoc.* **51** 273–82
- Zolghadri A and Cazaurang F 2006 Adaptive nonlinear state-space modelling for the prediction of daily mean PM<sub>10</sub> concentrations *Environ. Modell. Software* **21** 885–94