Supporting Information

Text S1. Model-measurement comparison of annual mean ozone concentrations in 2000

Figures S3 and S4 show comparisons between observations and model simulations of annual mean ozone concentrations in 2000 at 13 baseline sites (10 in the NH and 3 in the SH). These concentrations are derived from quadratic fits to the measurements and model results over the complete measurement records and the 1985-2014 model results. The absolute values of the CMIP6 model simulations are improved over the CMIP5 simulations, with similar spatial correlation. The more recent simulations overestimate the observations by an average of 3.5 ppb, with the individual mean model offsets varying between -1 and +9 ppb, while the earlier results overestimated annual mean ozone by 4 to 13 ppb (average overestimate of 9 ppb). This CMIP5 model assessment is consistent with the overestimates previously found at the NH baseline sites (Parrish et al., 2014), and the closer agreement found in the SH (Cooper et al., 2014; Parrish et al., 2016). Turnock et al. (2020) compare five of these same CMIP6 model simulations to observations from the TOAR database, and also find comparable model overestimates of ozone at rural sites in the NH.

Table S1. Monthly and annual average baseline ozone concentrations at the U.S. Pacific Coast MBL.

year	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec	annual
1987											33.50	28.93	
1988	32.06	31.45	38.30	40.08	36.26	24.92	19.70	21.85	29.96	25.09	31.84	25.64	29.76
1989	28.16	32.33	32.77	34.19	34.32	29.09	21.86	25.18	34.37	32.68	30.94	28.38	30.36
1990	29.54	30.96	34.88	30.39	32.11	24.66	22.55	21.90	28.84	31.05	28.39	29.89	28.76
1991	30.29	30.50	36.63	39.53	35.74	28.99	22.07	21.83	27.31	31.41	28.85	27.41	30.05
1992	27.50	33.50	32.85	36.75	35.11	26.75	21.47	24.34	26.84	31.30	26.92	29.19	29.38
1993	28.46	34.59	31.87	34.60	33.13	26.33	22.52	24.85	25.60	28.36	31.21	31.84	29.45
1994	30.29	33.00	37.21	37.42	34.11	27.20	20.56	22.88	23.06	32.45	32.56	29.83	30.05
1995	31.75	35.89	40.87	40.08	34.57				30.81	30.79	32.31	33.09	
1996	32.50	32.33	36.75	36.83	37.34				29.00	29.67	26.55	31.83	
1997	34.20	31.07	39.13	40.86	35.93				14.09		35.80	29.92	
1998	31.38	32.25	36.82	37.73	31.20				29.04	31.88		32.94	
1999	27.50	36.88	38.09	38.33	34.41				27.69	33.13	32.00	33.26	
2000	34.74		38.41	39.61	41.79				30.11	26.07	39.14		
2001	34.60	29.67	41.39	41.30	40.04				27.58	33.63	30.44	37.23	
2002	36.63	36.45	29.52	29.20	24.69	21.73	23.46	26.24	24.71	27.24	23.70	27.84	27.62
2003	27.15	33.23	40.70	46.44	42.51	36.63	25.89	29.02	29.50	30.86	33.81	35.27	34.25
2004	37.71	36.83	40.42	39.70	36.71	34.40	25.21	24.92	33.01	32.28	31.25	33.97	33.87

2005	30.66	35.61	41.93	44.25	38.27	32.60	25.45	25.32	29.51	29.79	35.25	33.41	33.50
2006	41.95	39.13	44.92	42.41	42.21	34.24	26.16	29.47	29.96	34.58	41.54	33.61	36.68
2007	33.24	39.33	40.83	44.04	41.91	31.89	24.61	27.81	33.78	30.92	34.73	37.00	35.01
2008	37.51	35.51	42.11	45.74	40.19	32.64	25.45	23.36	30.81	31.44	30.11	34.92	34.15
2009	30.37	40.67	41.52	43.58	35.14	29.60	24.25	25.12	30.17	31.58	32.59	32.09	33.06
2010	35.92	33.82	40.90	43.13	38.48	29.29	24.46	24.75	24.50	32.96	31.31	40.88	33.37
2011	26.71	34.32	37.71	41.27	39.95	28.93	20.34	23.40	26.38	28.22	31.85	25.89	30.41
2012	27.24	33.69	40.26	37.34	37.05	26.56	22.64	24.13	28.97	30.58	30.61	34.84	31.16
2013	32.57	33.83	34.26	38.08	34.60	24.89	22.10	20.71	25.78	26.38	29.89	30.17	29.44
2014	28.07	30.88	31.61	33.18	31.47	24.28	19.29	18.64	22.87	23.83	28.13	29.76	26.83
2015	22.55	29.94	32.93	36.68	34.90	24.50	20.75	19.13	22.90	24.50	28.67	31.77	27.44
2016	37.99	29.06	34.98	35.03	30.38	24.82	18.12	17.87	18.80	23.46	31.96	29.55	27.67
2017	29.86	36.89	35.80										

Table S2. References for model descriptions and citations for the simulation results from the ESMs that participated in the CMIP6 exercise, and were used in this analysis.

Model	Reference	Data Citation
BCC-ESM1	Wu et al., 2019; 2020	Zhang et al., 2018; 2019
CESM2-WACCM	Emmons et al., 2020; Gettelman et al., 2019; Tilmes et al., 2019	Danabasoglu, 2019a; 2019b; 2019c
GFDL-ESM4	Dunne, 2020; Horowitz, 2020	Horowitz et al., 2018; John et al., 2018; Krasting et al., 2018
GISS-E2-1-H	Bauer et al., 2020	GISS, NASA Goddard Institute for Space Studies (NASA/GISS) (2019). NASA-GISS GISS-E2-1-H model output prepared for CMIP6 CMIP amip. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.1421
MRI-ESM2-0	Yukimoto et al., 2019a; Oshima et al., 2020	Yukimoto et al., 2019b
UKESM1-0-LL	Sellar et al., 2019; Archibald et al., 2020b	Good et al., 2019; Tang et al., 2019

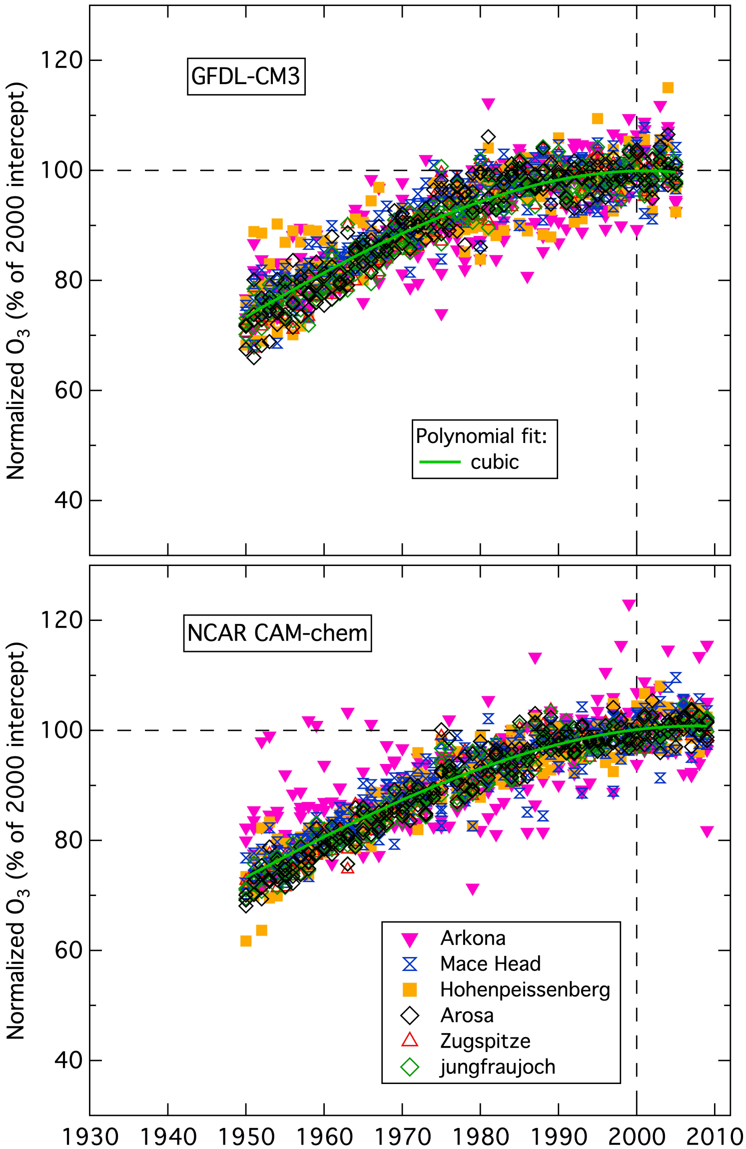

Figure S1. Normalized, seasonal mean ozone simulated at the six baseline representative European sites considered by Parrish et al. (2014). The simulations are from the GFDL-CM3 (upper panel) and NCAR CAM-chem (lower panel) models. Each graph includes cubic polynomial fits (green curves) to all seasonal means.

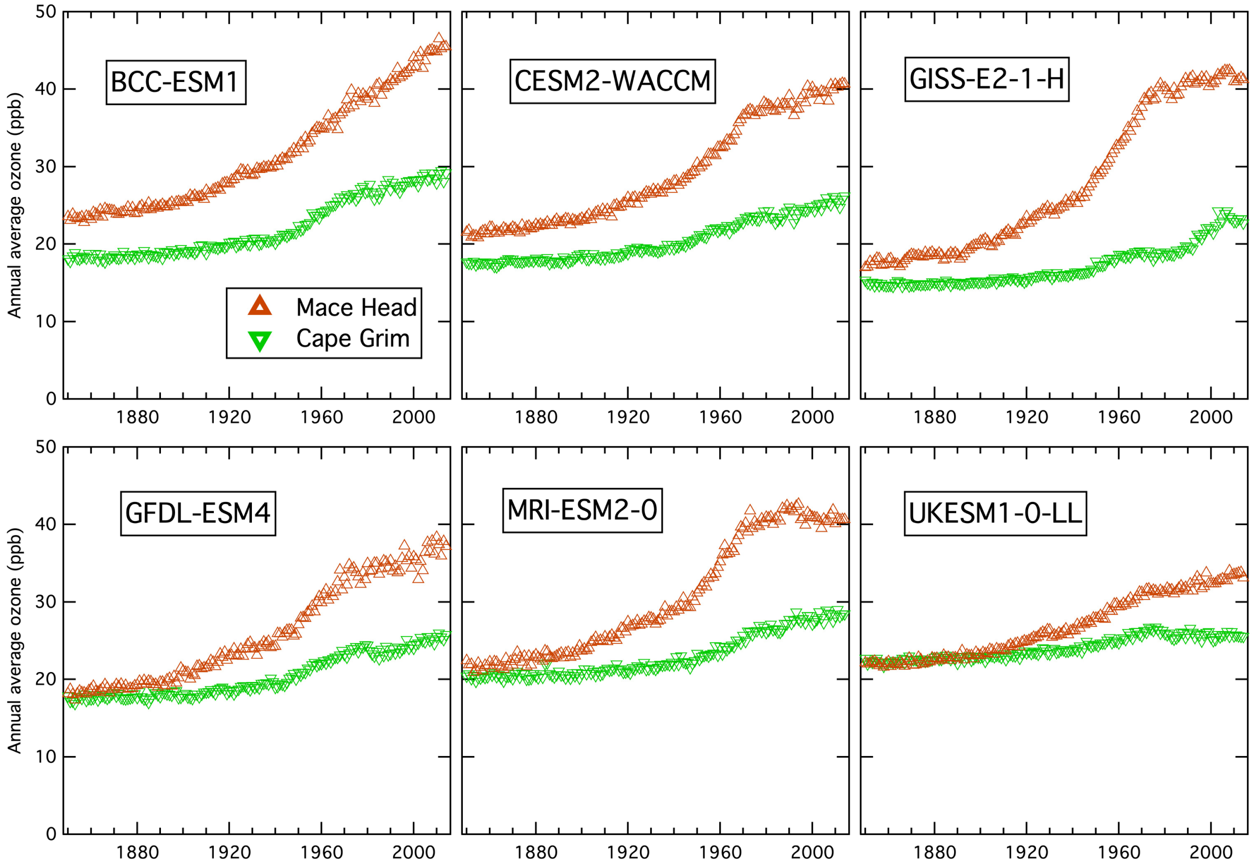

Figure S2. Simulated long-term changes in annual mean ozone mixing ratios at Mace Head, Ireland and Cape Grim, Australia from 1850-2014. The results are from six ESMs (identified in the annotations) that participated in the CMIP6 exercise.

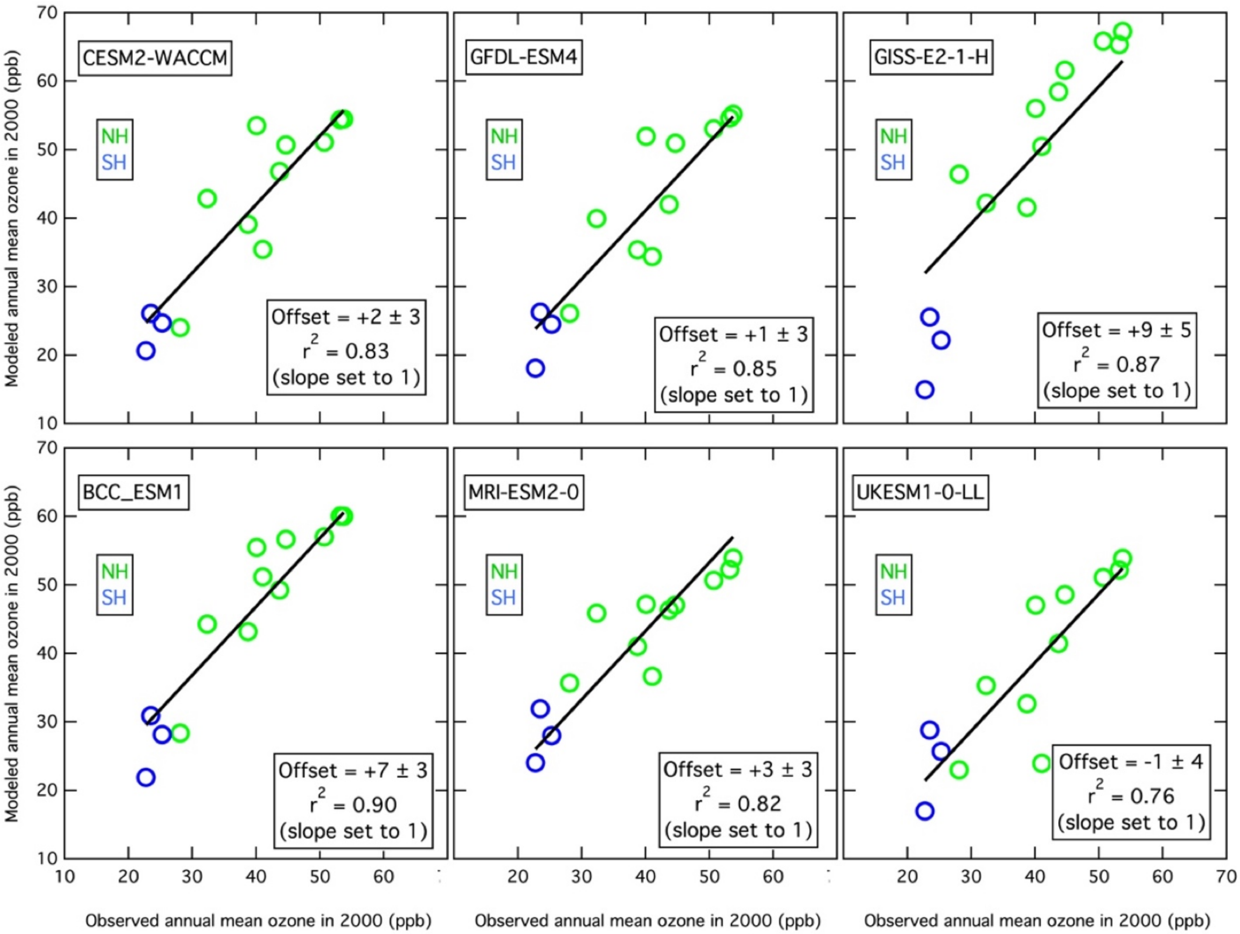

Figure S3. Model-measurement comparison of mean annual ozone in 2000 at 10 NH (green symbols) and 3 SH (blue symbols) baseline sites. The simulations are results from six ESMs (identified in the annotations) that participated in the CMIP6 exercise. Annotations give the square of the linear correlation coefficient and the offset with 95% confidence interval between the simulations and observations calculated; the offset is calculated from a standard linear regression with the slope held at unity. The baseline sites are the 10 NH surface sites discussed by Parrish et al., 2012; 2014, and the 3 SH mid-latitude sites discussed by Cooper et al. 2014 (Cape Point, South Africa; Cape Grim, Australia; and Ushuaia, Argentina).

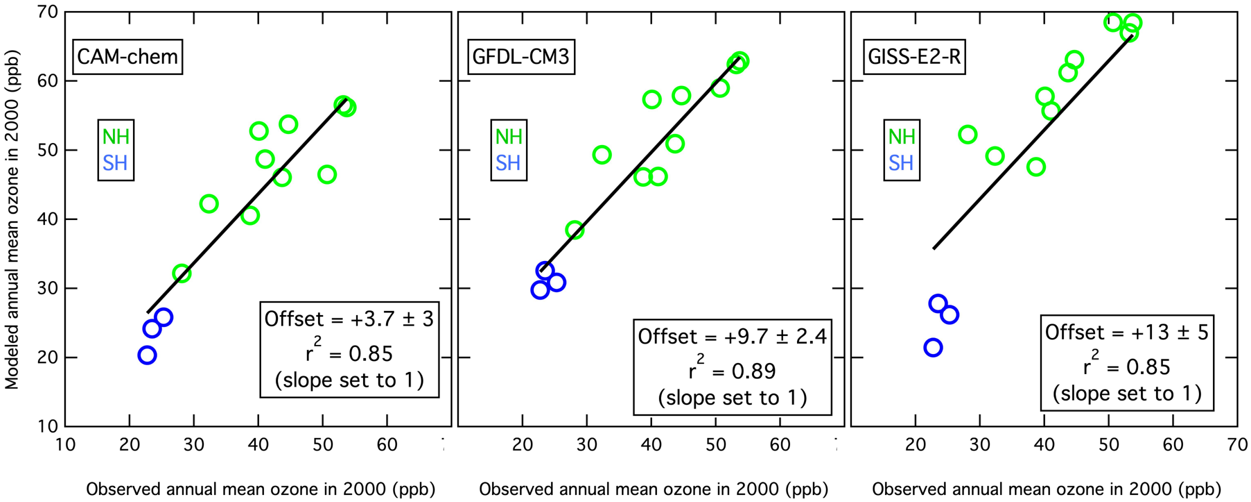

Figure S4. Model-measurement comparison of mean annual ozone in 2000 at 10 NH and 3 SH baseline sites identified in the caption of Figure S3. The simulations are the results from three CCMs (identified in the annotations) that participated in the CMIP5 exercise. Annotations give the square of the linear correlation coefficient and the offset with 95% confidence interval between the simulations and observations; the offset is calculated from a standard linear regression with the slope held at unity.