Supplementary Materials For

Mortality Burdens in California Due to Air Pollution Attributable to Local and Non-local Emissions

Configuration and Performance of WRF-Chem

We simulated the $PM_{2.5}$ and O_3 concentrations using the WRF-Chem model (version 3.9.1). WRF-Chem has been widely used by research groups around the world for a broad range of applications in atmospheric chemistry and air quality. A general description of the model can be found at https://www2.acom.ucar.edu/wrf-chem. The physical options included the National Center for Environmental Prediction, Oregon State University, Air Force, and Hydrologic Research Lab's (NOAH) land-surface module (Chen and Dudhia, 2001), the Yonsei University (YSU) PBL scheme (Hong et al., 2006), the Grell-Freitas cumulus scheme (Grell and Freitas, 2014), the Morrison double-moment scheme for cloud microphysics (Morrison et al., 2009), and the Fu-Liou-Gu (FLG) radiative transfer scheme (Fu and Liou, 1992; Gu et al., 2011, 2006; Zhao et al., 2016). With respect to the chemical scheme, we employed an extended Carbon Bond 2005 (CB05) (Yarwood et al., 2005) with chlorine chemistry (Sarwar et al., 2008) coupled with the Modal for Aerosol Dynamics in Europe/Volatility Basis Set (MADE/VBS) (Ahmadov et al., 2012; Wang et al., 2015). MADE/VBS uses a modal aerosol size representation and an advanced secondary organic aerosol (SOA) module based on the VBS approach. The aqueous-phase chemistry was based on the AQChem module used in the Community Multiscale Air Quality (CMAQ) model (Wang et al., 2015). This model also considered aerosol direct radiative effects and first and second aerosol indirect effects on grid-scale clouds following our previous study (Zhao et al., 2017b). The biogenic emissions were calculated online using the Model of Emissions of Gases and Aerosols from Nature (MEGAN) (Guenther et al., 2006). Dust emissions were calculated online following Zhao et al., (2010), as based on the Goddard Chemical Aerosol Radiation Transport (GOCART) dust emission scheme (Ginoux et al., 2001). Sea-salt emission calculation followed previous studies (Gong, 2003; C. Zhao et al., 2013). The wildfire emissions were calculated using the Brazilian Biomass Burning Emission Model (3BEM) (Longo et al., 2010) with input from the Moderate Resolution Imaging Spectroradiometer (MODIS) fire product (NASA, 2018). The plume rise of wildfire was calculated online following previous studies (Freitas et al., 2010, 2007).

We compare the meteorological predictions with observational data obtained from the National Climatic Data Center (NCDC), where hourly or 3-hour observations of wind speed at 10 m (WS10), temperature at 2 m (T2), and water vapor mixing ratio at 2 m (Q2) are available for 416 sites distributed within the modeling domain. We apply a number of statistical indices to quantitatively evaluate the model performance, as summarized in **Table S1**. These indices include mean observation (Mean OBS), mean simulation (Mean SIM), mean bias (MB), gross error (GE), root mean square error (RMSE) and index of agreement (IOA), which are defined by Emery et al., (2001). In general, the model predictions agree fairly well with surface meteorological observations. The performance statistics for WS10, T2 and Q2 are within the benchmark ranges proposed by Emery et al., (2001), except that the MB for T2 in April and the GE for T2 slightly

exceed the ranges. Note that these benchmark values are proposed based on the performance of a series of model simulations with four dimensional data assimilation (FDDA). Nevertheless, FDDA is not utilized here to allow full aerosol-radiation-cloud interactions, therefore slight exceedance in some cases is deemed reasonable.

We compared the simulated and observed monthly average $PM_{2.5}$ concentrations and monthly average of daily maximum 8-h O_3 concentrations in each individual monitoring site (**Figure S1**). The model reproduces the magnitude and spatial distribution of $PM_{2.5}$ concentrations fairly well, with an overall slight underestimation of 6%. A number of statistical indices were applied to quantitatively assess the model performance based on previous studies (Boylan and Russell, 2006; Yu et al., 2006; B. Zhao et al., 2013; Zhao et al., 2017a, 2017b). Statistics of model performance for daily maximum 1-h, 8-h average, and $PM_{2.5}$ concentrations, as well as PM chemical predictions are summarized in **Table S2**. The indices include Mean OBS, Mean SIM, normalized mean bias (NMB), normalized mean error (NME), mean fractional bias (MFB), and mean fractional error (MFE), as documented in previous studies (Boylan and Russell, 2006; Yu et al., 2006). For O_3, the model is able to capture the spatial variability but slightly underestimates the daily maximum 1-h and 8-h average O_3 concentrations by 15% and 12% on average, respectively. The underestimation occurs in nearly all months but the NMBs are all within ±20%. Regarding to predictions of $PM_{2.5}$ and its chemical components, the statistics in **Table S2** indicated reasonably good model-measurement agreement. The NMBs of $PM_{2.5}$ concentration range from -24% in July to +10% in April. The performance statistics for $PM_{2.5}$ meet the model performance criteria (i.e MFB, within ±60% and MFE, ≤ 75%) proposed by Boylan and Russell (Boylan and Russell, 2006) in all months, and meet the model performance goal (MFB within ±30% and MFE ≤ 50%) in all months except January, indicating an overall good model-measurement agreement. The BC concentrations are significantly overestimated by 105%, probably attributable to the absence of BC aging in the model, leading to a reduced fraction of hydrophilic BC and thus reduced wet deposition. The OC concentrations are underestimated by 29%, which is a common problem in most chemical transport models (Heald et al., 2005; Zhao et al., 2016). Although the inclusion of multi-generational aging of SOA based on the VBS scheme has significantly reduced the underestimation (Ahmadov et al., 2012), some important chemical processes, such as the photo-oxidation of primary organic aerosols and intermediate volatility organic compounds Zhao et al., (2016, 2015), are still missing in model, likely accounting for the underestimation that still exists. The concentrations of SO_4^{2-} and NO_3^- are either overestimated or underestimated, depending on the simulation period. The NMBs are all within ±35% except for the SO_4^{2-} in January, which is overestimated by 62% possibly due to uncertainties in emission inventory or aerosol chemistry. A sensitivity run using the Regional Atmospheric Chemistry Mechanism (RACM) scheme shows improved O_3 simulation results (mean NMBs are -8% and -5% for daily maximum 1-h and 8-h average O_3 concentrations, respectively), although the model performance for particulate matter is not as good as the current simulation results.

Table S1. Statistics of model performance for meteorological predictions.

Variable	Index	Unit	Jan	Apr	Jul	Oct	Benchmark
Wind speed (WS10)	Mean OBS	(m/s)	4.24	4.54	3.9	4.04	
	Mean SIM	(m/s)	4.13	4.06	3.41	3.73	
	MB	(m/s)	-0.12	-0.49	-0.50	-0.30	$\leqslant \pm 0.5$
	GE	(m/s)	1.76	1.69	1.53	1.53	$\leqslant 2$
	IOA		0.75	0.74	0.66	0.75	$\geqslant 0.6$
Temperature (T2)	Mean OBS	(K)	276.6	285.65	296.34	285.94	
	Mean SIM	(K)	276.97	284.98	295.95	286.16	
	MB	(K)	0.37	-0.68	-0.39	0.22	$\leqslant \pm 0.5$
	GE	(K)	2.77	2.31	2.27	2.37	$\leqslant 2$
	IOA		0.94	0.95	0.96	0.96	$\geqslant 0.8$
Water vapor mixing ratio (Q2)	Mean OBS	(g/kg)	3.21	4.83	8.83	5.27	
	Mean SIM	(g/kg)	3.28	4.88	8.89	5.21	
	MB	(g/kg)	0.07	0.05	0.06	-0.06	$\leqslant \pm 1$
	GE	(g/kg)	0.64	0.77	1.35	0.82	$\leqslant 2$
	IOA		0.89	0.91	0.86	0.93	$\geqslant 0.6$

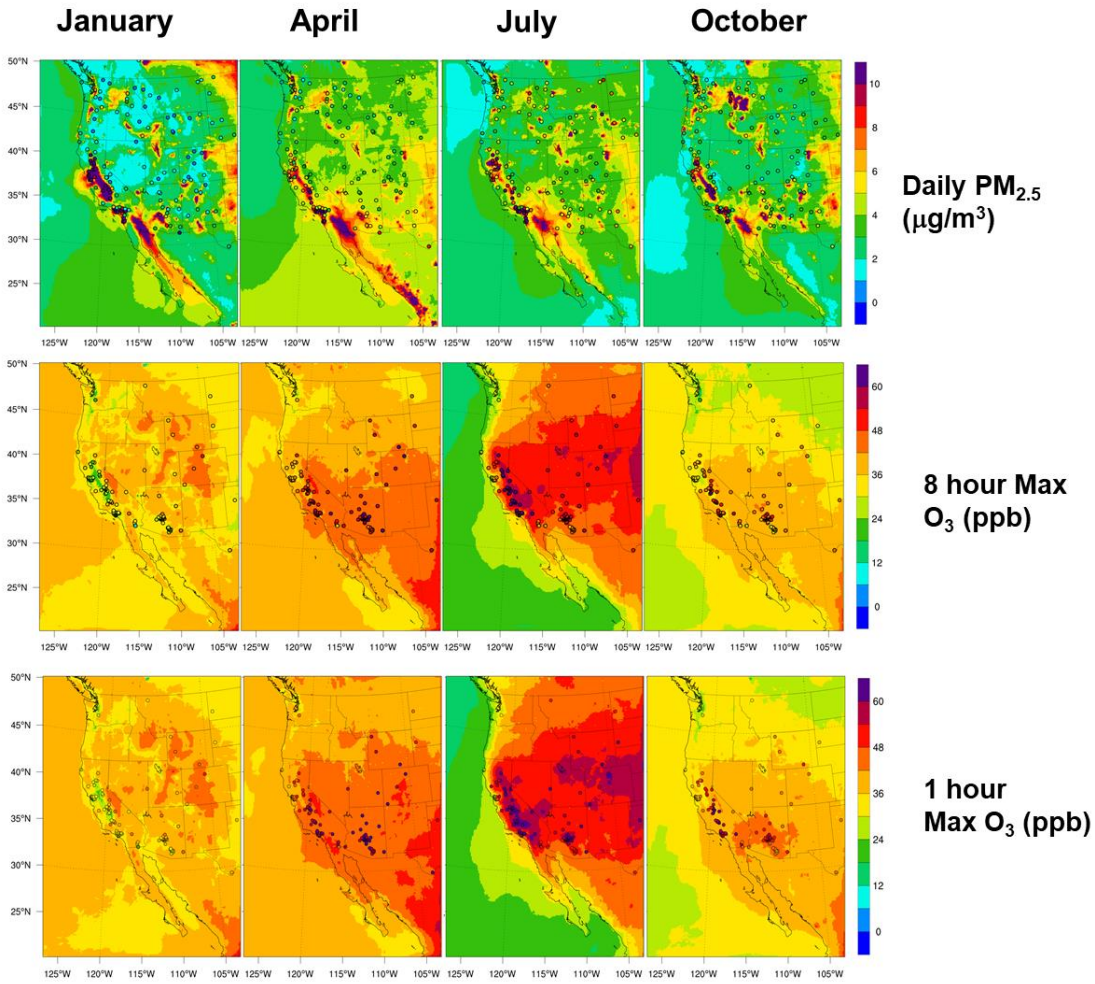

Figure S1. Observed (dots) and simulated (contours) monthly mean $PM_{2.5}$ concentrations (upper row), monthly mean daily maximum 8-h O_3 concentrations (middle row), and monthly mean daily maximum 1-h O_3 concentrations (bottom row), in January , April, July , and October 2012.

Table S2. Statistics of model performance for chemical predictions.

	Month	Site number	Mean SIM	Mean OBS	NMB (%)	NME (%)	MFB (%)	MFE (%)
1-h O_3	Jan	139	70.9	74.6	-5	18	-6	21
	Apr	165	94.9	117.9	-20	21	-21	23
	Jul	169	108.4	124.0	-13	18	-13	19
	Oct	170	82.3	100.5	-18	22	-19	23
	Average		89.1	104.2	-15	20	-15	22
8-h O_3	Jan	139	64.7	64.2	1	21	1	26
	Apr	165	87.8	107.9	-19	21	-20	23
	Jul	169	99.7	111.6	-11	18	-11	20
	Oct	170	75.6	88.6	-15	22	-15	24
	Average		81.9	93.1	-12	20	-11	23
$PM_{2.5}$	Jan	154	5.8	5.8	-1	57	37	65
	Apr	157	5.7	5.2	10	49	27	51
	Jul	154	5.1	6.7	-24	41	-19	43
	Oct	155	5.3	5.5	-4	52	4	50
	Average		5.5	5.8	-6	49	12	52
BC	Jan	118	0.3	0.2	74	118	59	81
	Apr	119	0.3	0.1	177	190	70	78
	Jul	121	0.3	0.2	76	115	29	57
	Oct	116	0.3	0.1	125	147	52	70
	Average		0.3	0.2	105	137	52	72
OC	Jan	117	0.4	0.6	-39	69	-48	74
	Apr	119	0.4	0.5	-20	58	-30	60
	Jul	121	0.7	1.2	-42	64	-52	77
	Oct	116	0.8	0.8	-6	79	-31	73
	Average		0.6	0.8	-29	68	-40	71
NO_3^-	Jan	66	1.0	0.9	12	84	65	110
	Apr	90	0.3	0.5	-32	75	-66	109
	Jul	32	0.3	0.5	-34	73	-90	126
	Oct	22	0.8	0.7	14	74	-24	100
	Average		0.6	0.7	-5	77	-29	111
SO_4^{2-}	Jan	125	0.5	0.3	62	97	68	78
	Apr	125	0.9	0.7	23	64	35	59
	Jul	122	0.9	0.9	-2	50	10	51
	Oct	123	0.7	0.5	23	53	33	52
	Average		0.7	0.6	18	60	37	60

Note: Observations of O_3 are obtained from the Air Quality System (AQS)(U.S. EPA, n.d.), the Interagency Monitoring of Protected Visual Environments (IMPROVE)(U.S. EPA, n.d.), and the Clean Air Status and Trends Network (CASTNET)(U.S. EPA, 2017), while those of $PM_{2.5}$ and its chemical constituents are from AQS and IMPROVE.

Population and baseline mortality rates

In this study, we used the default demographic information in the BenMAP, which is from census block level ACS 5-year estimates from 2012 to 2016 (U.S. EPA, 2018). Custom demographic information was not used because the default database already has detailed population information in 2012 from ACS. To further process the population data, we used BenMAP to calculate a crosswalk, which converts the 2012 demographic data in the default grid (i.e., EPA 12 km CMAQ) to our self-designed 12 km grid covering western United State (U.S. EPA, 2018). The processed grids contain 2012 population information by 5-year age bins from 0-99 years old.

For baseline mortality rates, we used the default numbers in the BenMAP. The default BenMAP baseline mortality rates were based on 2012-2014 county-level mortality data from the Centers for Disease Control WONDER database (http://wonder.cdc.gov) and ACS population data described in the previous paragraph (U.S. EPA, 2018). Baseline age-, cause-, and county-specific mortality rates were generated using the following formula:

$$R_{i,j,k} = \frac{D_{i,j,k\,(2012)} + D_{i,j,k\,(2013)} + D_{i,j,k\,(2014)}}{P_{i,j,k\,(2012)} + P_{i,j,k\,(2013)} + P_{i,j,k\,(2014)}}$$

where $R_{i,j,k}$ is the mortality rate for age group i, cause j, and county k; D is the death count from WONDER; and P is the population from ACS (U.S. EPA, 2018). This method calculates average mortality rates based on three consecutive years instead of any single year, and can provide more conservative and reliable mortality rate estimates. BenMAP provides seven mortality causes: all cause, non-accidental, respiratory, chronic lung, lung cancer, ischemic heart disease, and cardio-pulmonary. Our study mainly used this data based to analyze all-cause and respiratory mortality in 2012 in California.

Raw data for baseline mortality rates in California for CVD, NCD, and LRI were downloaded from the GBD tool (Institute for Health Metrics and Evaluation, 2018). We calculated age-, and cause-specific mortality rates in 2012 in California using the following formula:

$$R_{i,j} = \frac{R_{i,j\,(2011)} + R_{i,j\,(2012)} + R_{i,j\,(2013)}}{3}$$

where $R_{i,j,}$is the mortality rate for age group i, and cause j.

CRF information

This study applied concentration response functions (CRFs) for $PM_{2.5}$ and O_3-associated mortality from several major epidemiological studies. Three studies (Jerrett et al., 2009; Krewski et al., 2009; Turner et al., 2016) were based on the American Cancer Society's Cancer Prevention Study II (CPS-II) cohort. The Hoek et al., (2013) study is a meta-analysis of multiple $PM_{2.5}$ studies. The IER developed by Burnett et al., (2014) is an integrated model incorporating cohort studies from multiple countries estimating mortality from specific causes, and be widely used for burden of disease studies. And the GEMM recently developed by Burnett et al., (2018) is an update to the IER, and also incorporates cohort studies from multiple countries. Information regarding to the CRFs used in this study were provided in **Table S3**.

Table S3 sensitivity analysis for $PM_{2.5}$ CRF

Study	**Air pollutant**	**Cause of death**	**HR (95% CI) per 10μg/m^3 change in $PM_{2.5}$**	**Notes**
Krewski (2009)	$PM_{2.5}$	All-cause	1.06 (1.02–1.10)	CPS-II study for U.S. population
Hoek (2013)	$PM_{2.5}$	All-cause	1.06 (1.04–1.08)	Meta-analysis
Burnett (2018)	$PM_{2.5}$	NCD + LIR	$HR(z)=\exp\{\theta\log(z/\alpha+1)/(1+\exp\{-(z-\mu)/\nu\})\}$	$z=\max(0, PM_{2.5}-2.4)$ $\Theta=0.1231$ $\alpha=1.5$ $\mu = 10.4$ $\nu = 25.9$
Burnett (2014)	$PM_{2.5}$	IHD, Cerebral, COPD and Lung cancer	Multiple	Used BenMAP default formula (U.S. EPA, 2018)
Turner (2016)	O_3	All-cause	1.02 (1.01–1.04)	CPS-II study for U.S. population
Turner (2016)	O_3	CVD	1.03 (1.01–1.05)	
Turner (2016)	O_3	Respiratory	1.12 (1.08–1.16)	
Turner (2016)	O_3	All-cause	1.02 (1.02–1.03)	Summer season only
Jerrett (2009)	O_3	Respiratory	1.04 (1.01–1.07)	CPS-II study, between mortality and 1-h max O_3

Table S4 Selected mortality estimates (95% CI) associated with ambient $PM_{2.5}$ and O_3 contributed by different emission groups in 2012, based on CRFs from Burnett (2018) for $PM_{2.5}$ and Turner (2016) for O_3.

Emission Groups	$PM_{2.5}$	O_3			
	NCD+LRI, year-round	All-cause, year-round	All-cause, summer	CVD, year-round	Respiratory, year-round
California in-State Anthropogenic	14,000 (9,900 – 18, 400)	-240 (-60 - -400)	771 (389 - 1146)	-140 (-45 - -240)	-100 (-40 - -150)
Anthropogenic Emissions from the Western U.S., except California	1,000 (700-1,200)	380 (100 - 670)	136 (68 - 204)	180 (60 - 300)	200 (80- 300)
Natural Emissions from the Western U.S.	3,500 (2,500-4,500)	930 (240 – 1,650)	560 (281 - 835)	430 (150 - 710)	500 (200 - 750)
All Emissions from Outside of the Western U.S.	8,200 (5,800 – 11,000)	12,600 (5,800 – 2,1800)	3,843 (1,960 – 5,657)	5,900 (2,000 – 9,500)	6,700 (2,800 – 10,000)

Table S5 Monetized health loss associated with $PM_{2.5}$ and O_3, disaggregated by emission groups

Emission Group	Monetized Health Loss (billions of 2012 USD)		
	$PM_{2.5}$	O_3	Total
California in-State Anthropogenic	153.8 (113.8 - 194.9)	-2.0 (-0.5 - -3.6)	151.8 (133.3 – 191.3)
Anthropogenic Emissions from the Western U.S., except California	4.0 (3.0 – 5.1)	3.2 (0.8 - 5.6)	7.2 (3.8 - 10.7)
Natural Emissions from the Western U.S.	17.1 (12.7 – 21.7)	8.0 (2 - 13.8)	25.1 (14.7 – 35.5)
All Emissions from Outside of the Western U.S.	55.6 (40.9 - 70.8)	106.0 (48.9 - 182.8)	161.6 (89.8 – 253.6)

Table S6 Population-weighted excess ambient O_3 (daily 8-hour maximum) due to in-state emissions at selected counties and California statewide.

County	Region	ΔO_3 (ppb)			Adult population
		Annual	Winter	Summer	
Los Angeles	Southern California	- 5.4	- 15	2.1	5,732,579
San Diego	Southern California	0.6	- 5.5	5.4	1,821,172
Orange	Southern California	- 0.8	- 9.4	6.5	1,788,568
Riverside	Southern California	2.0	- 5.8	11	1,247,742
Santa Clara	Bay Area	- 2.5	- 8.3	1.2	1,039,730
Alameda	Bay Area	- 4.3	- 9.5	- 1.5	967,497
Contra Costa	Bay Area	- 2.7	- 7.1	- 1.0	721,668
Sacramento	Sacramento Valley	0.7	- 6.4	8.5	719,713
Fresno	San Joaquin Valley	4.4	- 7.5	18	449,856
Statewide		- 0.8	- 8.1	5.8	21,372,052

ΔO_3 is the difference between baseline O3 concentrations and O3 concentrations without in-state anthropogenic emissions

The selection is based on population and region. Adult population includes population aged 30-99 in 2012, according to American Community Survey (ACS) 5-year estimates.