Simulation of Neighborhood-Scale Air Quality with two-way coupled WRF-CMAQ over Southern Lake Michigan-Chicago Region

Anastasia Montgomery[1], Jordan Schnell[2], Zachariah Adelman[3], Mark Janssen[4], and Daniel E Horton[1]

[1]Northwestern University
[2]Cooperative Institute for Research in Environmental Sciences at the University of Colorado Boulder NOAA/Global Systems Laboratory
[3]University of North Carolina at Chapel Hill
[4]Lake Michigan Air Directors Consortium

November 22, 2022

Abstract

The southern Lake Michigan region of the United States, home to Chicago, Milwaukee, and other densely populated Midwestern cities, frequently experiences high pollutant episodes with unevenly distributed exposure and health burdens. Using the two-way coupled Weather Research Forecast and Community Multiscale Air Quality Model (WRF-CMAQ), we investigate criteria pollutants over a southern Lake Michigan domain using 1.3 and 4 km resolution hindcast simulations. We assess WRF-CMAQ's performance using data from the National Climate Data Center and EPA Air Quality System. Our 1.3 km simulation slightly improves on the 4 km simulation's meteorological and chemical performance while also resolving key details in areas of high exposure and impact, i.e., urban environments. At 1.3 km, we find that most air quality-relevant meteorological components of WRF-CMAQ perform at or above community benchmarks. WRF-CMAQ's chemical performance also largely meets community standards, with substantial nuance depending on the performance metric and component assessed. For example, hourly simulated NO2 and O3 are highly correlated with observations ($r > 0.6$) while PM2.5 is less so ($r = 0.4$). Similarly, hourly simulated NO2 and PM2.5 have low biases (<10%), whereas O3 biases are larger (<30%). Simulated spatial pollutant patterns show distinct urban-rural footprints, with urban NO2 and PM2.5 20-60% higher than rural, and urban O3 6% lower. We use our 1.3 km simulations to resolve high-pollution areas within individual urban neighborhoods and characterize changes in O3 regimes across tight spatial gradients. Our findings demonstrate both the benefits and limitations of high-resolution simulations, particularly over urban settings.

Hosted file

`montgomery_etal_2022_supplementformatted_v2.docx` available at https://authorea.com/users/537405/articles/599320-simulation-of-neighborhood-scale-air-quality-with-two-way-coupled-wrf-cmaq-over-southern-lake-michigan-chicago-region

Simulation of Neighborhood-Scale Air Quality with two-way coupled WRF-CMAQ over Southern Lake Michigan-Chicago Region

Anastasia Montgomery[1], Jordan L. Schnell[2,3], Zachariah Adelman[4], Mark Janssen[4], Daniel E. Horton[1,3]

[1]Department of Earth and Planetary Sciences, Northwestern University, Evanston, IL, USA

[2]Institute For Sustainability and Energy at Northwestern (ISEN), Northwestern University, Evanston, IL, USA

[3]Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado-Boulder, Boulder, CO, USA

[4]Lake Michigan Air Directors Consortium (LADCO), Rosemont, IL, USA

Corresponding author: Anastasia Montgomery (stacy@earth.northwestern.edu)

Key Points:

- We perform nested air quality simulations over the Midwestern US, with the inner domain resolution at the neighborhood-scale (1.3 km)
- NO_2 and $PM_{2.5}$ hotspots are simulated to be adjacent to major roadways, with substantial pollutant heterogeneity found within urban settings
- Simulated ozone regime (VOC:NOx ratio) found to vary seasonally and over tight spatial gradients

Abstract

The southern Lake Michigan region of the United States, home to Chicago, Milwaukee, and other densely populated Midwestern cities, frequently experiences high pollutant episodes with unevenly distributed exposure and health burdens. Using the two-way coupled Weather Research Forecast and Community Multiscale Air Quality Model (WRF-CMAQ), we investigate criteria pollutants over a southern Lake Michigan domain using 1.3 and 4 km resolution hindcast simulations. We assess WRF-CMAQ's performance using data from the National Climate Data Center and EPA Air Quality System. Our 1.3 km simulation slightly improves on the 4 km simulation's meteorological and chemical performance while also resolving key details in areas of high exposure and impact, i.e., urban environments. At 1.3 km, we find that most air quality-relevant meteorological components of WRF-CMAQ perform at or above community benchmarks. WRF-CMAQ's chemical performance also largely meets community standards, with substantial nuance depending on the performance metric and component assessed. For example, hourly simulated NO_2 and O_3 are highly correlated with observations ($r > 0.6$) while $PM_{2.5}$ is less so ($r = 0.4$). Similarly, hourly simulated NO_2 and $PM_{2.5}$ have low biases (<10%), whereas O_3 biases are larger (>30%). Simulated spatial pollutant patterns show distinct urban-rural footprints, with urban NO_2 and $PM_{2.5}$ 20-60% higher than rural, and urban O_3 6% lower. We use our 1.3 km simulations to resolve high-pollution areas within individual urban neighborhoods and characterize changes in O_3 regimes across tight spatial gradients. Our findings demonstrate both the benefits and limitations of high-resolution simulations, particularly over urban settings.

1 Introduction

Exposure to poor air quality in the U.S. has been found to exacerbate respiratory diseases (Kurt et al., 2016), drive disparate health burdens in racial minority populations (Jbaily et al., 2022; Tessum et al., 2021), and contribute to ~100,000 premature deaths annually (Goodkind et al., 2019). Given the substantial public health burden associated with exposure to poor air quality, it is essential to resolve pollutant exposure at high spatiotemporal resolutions – particularly for use in the design of amelioration and abatement strategies. Indeed, pollutant exposure in high population settings, i.e., urban environments, can vary widely, which can contribute to disparities in health outcomes on a neighborhood-by-neighborhood basis across individual cities (Alexeeff et al., 2018; Goodkind et al., 2019; O'Leary & Lemke, 2014; Southerland et al., 2021). Determining the relationship between heterogeneous pollutant exposure and disparate health effects is challenging given observational constraints. For example, regulatory-grade air quality monitoring stations are relatively sparse and therefore spatial coverage is limited, particularly in urban settings. Observing platforms that do have better spatial coverage, e.g., remote sensed satellite observations, often have temporal limitations such as making only one observation a day in the case of polar orbiters (Goldberg et al., 2021). Given the need to resolve pollutants across impact-relevant scales (Clark et al., 2022), researchers often turn to physics- and chemistry-based Chemical Transport Models (CTMs) which allow for the spatial heterogeneity of pollutants to be estimated at high temporal resolutions in areas that are otherwise unmonitored (Hu et al., 2019; Southerland et al., 2021).

State-of-the-science CTMs resolve pollutants at geospatially and temporally continuous scales. For example, the Community Multiscale Air Quality modeling system (CMAQ; Byun & Schere, 2006) was developed by the U.S. Environmental Protection Agency (EPA) to study the complex

interactions of pollutants and meteorology and increase our understanding of atmospheric processes. Over time, the spatial resolution of CTMs like CMAQ have increased in resolution as computational costs decrease and spatially-defined inputs are resolved at finer scales (Gan et al., 2016). Higher resolution CTM studies have the potential to simulate more accurate meteorology, emissions, and pollutant concentrations than coarser resolution models (Fountoukis et al., 2013; Gan et al., 2016; Torres-Vazquez et al., 2022). However, some simulated meteorological and chemical variables may show lower model performance at finer resolutions because of incomplete characterizations of complex terrain and limitations in the planetary boundary layer formation (Tran et al., 2018; Zhang et al., 2014a). Notably, higher resolution studies benefit epidemiological studies which identify the health impacts at more health-relevant scales (Jiang and Yoo, 2018; Thompson et al., 2014; Southerland et al., 2022).

Here, given the potential benefits of high resolution CTM studies, i.e., resolved pollutant heterogeneities and hotspots over urban settings and higher-fidelity hindcast simulations, we utilize WRF-CMAQ to characterize pollutant concentrations over a southern Lake Michigan domain, a region in the central midwestern U.S., which includes the major populations centers of Chicago, IL and Milwaukee, WI. Previous modeling studies have focused on this region due to the atmospheric complexities associated with Lake Michigan and high O_3 pollution in the region (Abdi-Oskouei et al., 2020; Doak et al., 2021; Dye et al., 1995; Foley et al., 2011). By and large, air quality in this region has been improving due to emission controls and the outsourcing of industry and manufacturing (Jing et al., 2014, 2017, IEPA, 2019). However, pockets of poor air quality persist, particularly in and downwind of urban centers like Chicago, which has been in EPA 8-hour O_3 National Ambient Air Quality Standards (NAAQS) non-attainment status since 2004 (EPA, 2021).

From a regional perspective, poor air quality in Midwestern summers is often associated with warm stagnant air masses (Jing et al., 2017; Schnell & Prather, 2017; Tai et al., 2010), while near-surface winter pollution is largely restricted to particulate matter accumulations associated with temperature inversions (Hand et al., 2012). However, at local scales, local geography, meteorology, and emissions often play a synergistic role. For example, in Chicago, Illinois, the U.S. Midwest's most populated city, the coastal geography, micro-meteorology, and high-emitting urban footprint combine to create an active atmospheric regime that often facilitates accumulation of primary pollutants and/or the precursors of secondary pollutants. Indeed, Chicago's O_3 NAAQS non-attainment status is a direct result of interacting emissions, geography, and meteorology – particularly the interaction of precursor emissions with Lake Michigan's lake breeze (Abdi-Oskouei et al., 2020; Doak et al., 2021; Dye et al., 1995; Foley et al., 2011). Because the formation of O_3 is generally dependent on the ratio of precursor emissions, i.e., nitrogen oxides (NOx) and volatile organic compounds (VOCs), the EPA has restricted NOx emissions (EPA, 2019). However, previous studies have found that Chicago is in a transitional or VOC-limited regime (Jin et al., 2017; Jing et al., 2014; Lamsal et al., 2015; Lin et al., 2010) – suggesting a limitation to the efficacy of emissions controls that only consider NOx and do not also reduce VOCs. Further, studies also indicate that O_3, a secondary pollutant, often forms over Lake Michigan, and is transported ashore via the lake breeze circulation (Abdi-Oskouei et al., 2020; Doak et al., 2021; Dye et al., 1995; Foley et al., 2011).

Similar to O_3, the concentration of NO_2 also depends on meteorological factors such as winds and temperature (Harkey et al., 2015). In satellite analyses, Chicago appears as a large source of NO_2 pollution to the greater Midwest (Goldberg et al., 2021), a factor that contributes to the

formation and elevated concentration of downwind/rural O_3. In addition, NO_2 can be a precursor to $PM_{2.5}$ formation through the oxidation of NO_2 to nitrate. While Chicago is currently in compliance with $PM_{2.5}$ standards, $PM_{2.5}$ is elevated in comparison with rural areas due to the confluence of transportation, energy generation, and industrial emissions, as well as atmospheric transport, and secondary formation processes (Zhang et al., 2014b). Given all of the above complexities, the fine-scale characterization of Chicagoland air quality is complicated. Criteria pollutants such as NO_2, O_3 and $PM_{2.5}$ are all intricately linked through their emission sources, transport, accumulation, and secondary reactions. All of these factors speak to the need to resolve highly interactive geography, chemistry, and meteorology to accurately characterize the region's air quality.

Given the importance of meteorological conditions to air quality, particularly local-scale conditions in Chicago, we use a high-spatial resolution numerical model that includes atmospheric meteorology, chemistry, and components of their interactions and feedbacks, i.e., the two way-coupled Weather Research and Forecasting-Community Multiscale Air Quality modeling system (WRF-CMAQ). Neighborhood-scale simulations are made possible by the Lake Michigan Air Director's Consortium (LADCO) spatial surrogate dataset (LADCO, 2022), which defines the mapping of regional and county-level emission information. LADCO spatial surrogates are used in the Sparse Matrix Operating Kernel of Emissions (SMOKE) processing system (B.H. Baek & Seppanen, 2018) with the U.S. EPA Beta modeling platform (Eyth et al., 2019) to produce emission data for our 1.3 km grid. We use this emissions dataset in WRF-CMAQ to simulate 4 months representative of the 4 meteorological seasons and characterize pollutant concentrations over a central-Midwestern and Chicago-centric domain.

2 Methods

2.1 CTM Simulations and Domains

We performed CTM simulations using the two-way coupled Community Multi-scale Air Quality (CMAQ, v5.2; (Byun & Schere, 2006)) and Weather Research and Forecasting (WRF, v3.8; (Skamarock et al., 2008)) modeling system (WRF-CMAQ; (Wong et al., 2012)). The two-way configuration of WRF-CMAQ allows feedbacks between simulated aerosols and WRF's shortwave radiation scheme. To perform WRF-CMAQ simulations, we follow the methodology of Wong et al. (2012): (1) we produce dynamically downscaled meteorology with stand-alone WRF simulations, (2) we then use the stand-alone WRF output to create meteorologically-informed emissions data using the Sparse Matrix Operating Kernel of Emissions (SMOKE), and lastly (3) we run the coupled WRF-CMAQ model, incorporating the meteorologically-informed SMOKE emissions data.

To generate boundary and initial conditions and facilitate the production of meteorologically-informed emissions data, we first perform a stand-alone WRF simulation to generate three-dimensional meteorology in nested domains with 12 km (CONUS; d01), 4 km (Midwest; d02), and 1.3 km (southern Lake Michigan; d03) resolutions (Figure 1a). We use a 10-day spin-up period and simulate four months – August 2018, October 2018, January 2019, and April 2019 – using a 60, 20, and 6-second timestep for the 12, 4, and 1.3 km domains respectively. To allow soil moisture and soil temperature variables to reach a state of statistical equilibrium with observational constraints, we turn on the soil moisture initialization option during the 10 day spin-up (Pleim & Xiu, 2003). We run WRF with 35 vertical layers from the surface to 30 hPa

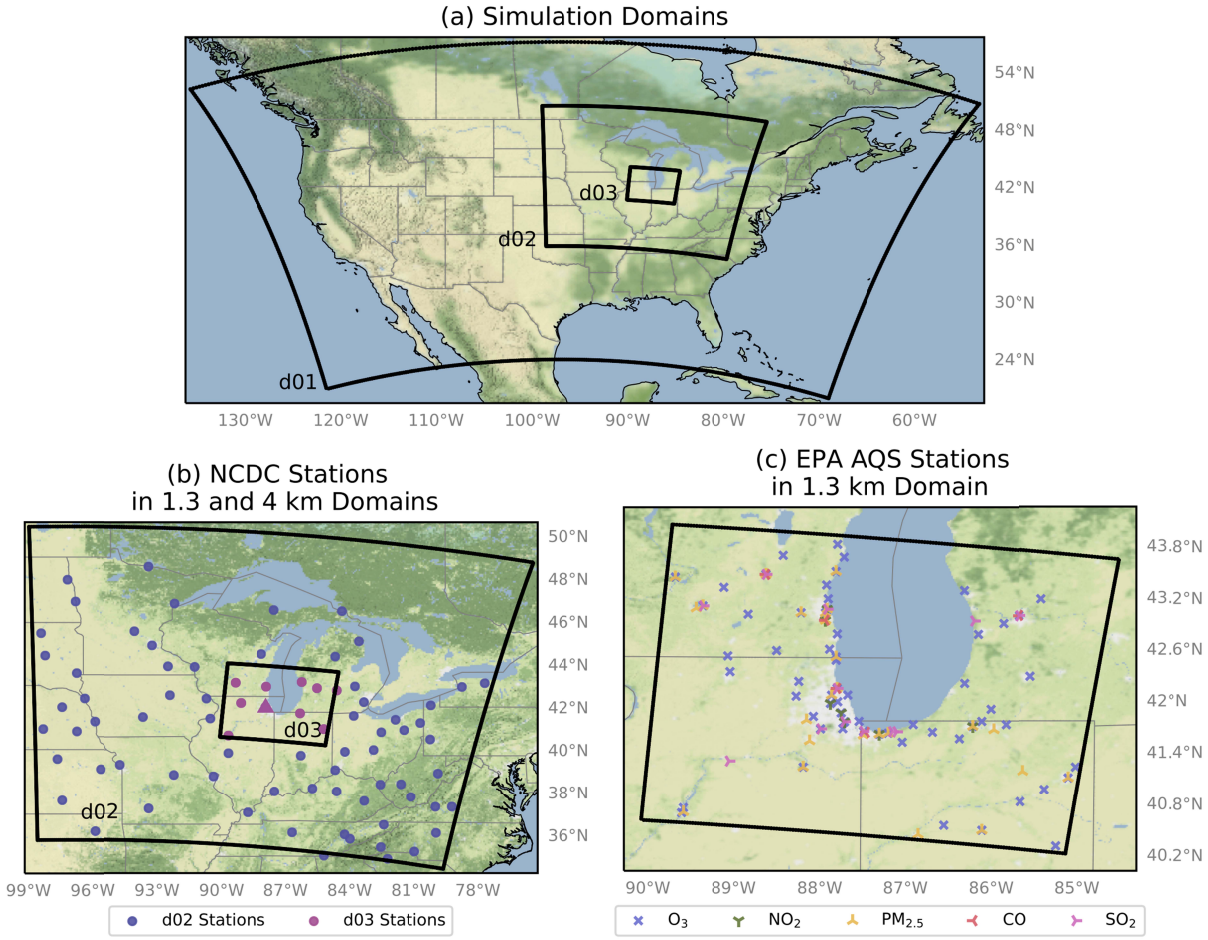


Figure 1. The nested model domains and observation sites. (a) Spatial footprint of nested model domains. We perform stand-alone WRF simulations at 12 km horizontal resolution in the d01 domain. Dynamically downscaled two-way coupled WRF-CMAQ simulations are performed in the d02 and d03 domains at 4 km and 1.3 km horizontal resolutions. (b) We use Local Climatological (LCD) Data from the National Climatic Data Center (NCDC) for meteorological validation of the 4 km (d02, purple) and 1.3 km (d03, pink) domains. The station located at Chicago-O'Hare is indicated by a triangle in the center of the d03 domain in panel (b). (c) EPA Air Quality System (AQS) stations are used to validate simulated pollutants. AQS stations are colored by chemical species and several AQS sensors for different chemical species are co-located, thus overlap on the map.

with a lowest model level thickness of ~20 m. Initial conditions and 3 hourly lateral boundary conditions for the 12 km domain are sourced from the North Atlantic Regional Reanalysis (NARR; Mesinger et al., 2006). Simulated WRF meteorology is nudged toward reanalysis using Four-Dimensional Data Assimilation (FDDA) at the surface and above the boundary layer, with nudging coefficients of 3×10^{-4} s-1 for temperature and winds, and 1×10-4 s-1 for the water vapor mixing ratio. We incorporate the land cover product from the National Land Cover Database (NLCD; Dewitz, 2021) at a 9 arc-second resolution. For the WRF physics options, we select the Morrison 2-moment microphysics scheme (Morrison et al., 2009), version 2 of the Kain-Fritsch (KF2) cumulus cloud parameterization for the 12- and 4-km simulations (Kain, 2004), the

Asymmetric Convective Model version 2 (ACM2) for the planetary boundary layer (Pleim, 2007), and the Pleim-Xiu land surface model (Xiu & Pleim, 2001) with soil moisture and temperature nudging (Pleim & Gilliam, 2009; Pleim & Xiu, 2003). We use the Rapid Radiative Transfer Model for GCMs (RRTMG; Clough et al., 2005) for both our shortwave and longwave radiation schemes.

To create 4 km and 1.3 km emissions inputs for use in WRF-CMAQ, we processed the EPA's 2016 Beta emissions modeling platform with the SMOKE software. We process the 2016v7.2 National Emissions Inventory (Eyth et al., 2019) using the 2016 SMOKE Beta Platform, relying on 4 km spatial surrogates provided by CMAS (CMAS, 2022) and 1.3 km spatial surrogates provided by LADCO (LADCO, 2022). The spatial surrogates map county-level emissions inventories to model grid cells by using the geographic attributes of the modeling area (such as population, industry, and economic activity). As meteorology is a key factor for vehicle emissions (e.g., cold starts and hoteling), we integrate the stand-alone WRF-simulated meteorology into the MOVES version developed for the 2016 beta platform (EPA, 2015). We use SMOKE to create emissions for the on-road, point, and nonpoint sectors. We calculate biogenic emissions (BEIS), lightning NOx emissions, and windblown dust "inline" during the coupled WRF-CMAQ simulation.

To ultimately simulate atmospheric pollutants, we run the two-way coupled version of WRF-CMAQ. We first run coupled WRF-CMAQ over the 4 km domain using the meteorological boundary conditions from the 12 km stand-alone WRF simulation, nudging from NARR (3-hourly resolution) and chemical boundary and initial conditions from CAM-Chem (Emmons et al., 2020; The CESM2 Development Team, 2019). We run the coupled 4 km simulation with an 18-second time step and 10-minute radiation time step, with CMAQ coupled every 8 WRF steps. To approach the neighborhood-scale, we dynamically downscale the 4 km coupled WRF-CMAQ simulation to provide meteorological and chemical boundary conditions for the nested 1.33 km simulation. For the 1.33 km WRF-CMAQ simulation, we run WRF with a 6 second time step and 5-minute radiation time step, with CMAQ coupled every 8 WRF steps. Both the 4 km and 1.3 km simulations integrate the Carbon Bond Mechanism version 6 and aerosol module version 6 with aqueous chemistry (cb6r3_ae6_aq) to create atmospheric constituents.

3 Results

To present our two-way coupled WRF-CMAQ simulations and highlight their ability to resolve neighborhood-scale air quality, we begin with evaluations of the model's domain-wide meteorological and chemical performance across temporal scales. We then highlight the utility of simulations that resolve air quality within individual neighborhoods, by conducting an in-depth analysis of intra-urban air quality by characterizing pollutant heterogeneities across Chicago, IL and their interactions with fine-scale meteorological features, local infrastructure, emissions sources, and the temporal distribution of emissions. Lastly, we assess the benefits, and in some cases disbenefits, of higher spatial resolution for model-observation fidelity performance.

We begin by comparing model-simulated air quality and metrological data from our highest resolution 1.3 km domain (d03; Figure 1) to ground-based observations. We evaluate model performance for each simulated month. Our air quality performance evaluation primarily focuses on O_3, NO_2, and $PM_{2.5}$, although other EPA criteria pollutants (i.e., CO and SO_2) are also

discussed. We evaluate model fidelity to meteorological and air pollutant observations using the following performance metrics: mean observation (μ_d), mean prediction (μ_p), normalized mean bias (*NMB*), normalized mean error (*NME*), correlation coefficients (*r*), mean error (*ME*), mean bias (*MB*), and root mean squared error (*RMSE*) as defined in Table S1. By normalizing model-simulated variables, the statistical performance of our simulations can be compared to similar model simulations performed over locations with different meteorology, emission profiles, and chemical regimes.

3.1 WRF-CMAQ Meteorological Performance

To assess the performance of the two-way coupled WRF-CMAQ meteorological output over the 1.3 km domain, we compare model simulated variables to ground-based measurements of meteorological conditions. We use hourly observational data from METAR stations aggregated by the National Climate Data Center (NCDC) (Figure 1b). We focus on 2-m temperature (T2) and relative humidity (RH) at 2 m, and wind speed (WS) and wind direction (WD) at 10 m, each of which is important to the fate and transport of atmospheric pollutants. We evaluated model fidelity using the performance recommendations outlined in Table S2 (Emery et al. 2001). Model-observation comparisons occur where WRF grid cells contain NCDC stations (Figure 1b). The 1.3 km domain contains 10 NCDC stations, which allows for model-observation comparison and assessment at 0.01% of the simulation grid cells (90,720 total). We also assess the model's meteorological performance within Chicago city limits, which has a single NCDC station.

In Table 1, we summarize the model's 1.3 km domain (d03) hourly meteorological performance against observations. For each month, WRF-CMAQ simulated T2, WD, and WS meet the correlation performance criteria suggested by Emery et al., 2001 (Table S2). Emery et al. (2001) do not make RH performance recommendations. WRF performance is best when simulating T2; model-station agreements have low biases and errors, though January 2019 and August 2018 have slightly higher biases than Emery et al.'s suggested benchmark (Table 1). Model simulations have a consistent warm bias across seasons, with the highest biases in August 2018 and January 2019 (*MB* = 0.8 °C), and highest mean errors in January 2019 and April 2019 (ME = 1.9 °C). Simulated RH is also highly correlated with observations ($r > 0.70$), with the highest bias and error in April 2019 (*MB* = 5.1 %, *ME* = 11.3 %). Simulated wind speeds meet *MB* benchmark criteria in January 2019 and April 2019, while missing the correlation benchmark when the 4-months are averaged (i.e., annualized). Simulated wind speeds are biased low in each season. Our lowest WRF performance is shown by WD, which only meets suggested *MB* criteria for April 2019. Model simulated WD and station measurements are highly correlated ($r > 0.5$), except for August 2018 ($r = 0.3$). The simulation is wetter (*MB* < 6%) and warmer (*MB* < 0.8 °C) than observations for all months, except for August 2018, where RH is biased low (*MB* = -5.7%).

Within Chicago city limits, there is one NCDC meteorological station, located ~16 km inland from Lake Michigan at O'Hare International Airport on the northwestern edge of the city (denoted by a triangle in Figure 1b). Model performance in comparison to the O'Hare meteorological station is shown in Table S3. Similar to the full 1.3 km domain comparison, meteorology in the model grid cell that contains O'Hare shows high correlations with NCDC observations for T2, WS, and RH. Also, like the full domain comparison, WRF-simulated T2 has the highest correlation with observations, while WD correlations are lowest. The ME and RMSE

258 are higher for the O'Hare grid cell for WS and RH than for the full 1.3 km domain comparison,
259 but lower for T2 and WD. Unlike the full 1.3 km domain, simulated T2 is cooler than
260 observations at O'Hare (-0.2 °C), though the RH biases are similar to the domain average (Table
261 S3).

262 **Table 1** Comparison of two-way coupled WRF-CMAQ simulated hourly meteorological
263 variables with NCDC observations for 1.3 km (d03) simulations. The average observed value is
264 noted as μ_d, while the predicted value is noted as μ_p.
265

Var	Month	μ_d	μ_p	*MB*	*ME*	*RMSE*	*r*
T2 (°C)	08/2018	23.2	24.0	0.8*	1.8	2.4	0.9
	10/2018	10.8	11.1	0.3	1.8	2.3	0.9
	01/2019	-5.8	-5.0	0.8*	1.9	2.5	1.0
	04/2019	9.2	9.2	0.0	1.9	2.5	0.9
	Average	9.4	9.8	0.5	1.9	2.4	0.9
RH (%)	08/2018	76.6	71.0	-5.7	11.1	14.4	0.7
	10/2018	74.8	76.6	1.9	12.2	15.4	0.7
	01/2019	74.6	78.6	4.0	9.6	11.7	0.7
	04/2019	66.1	71.9	5.8	12.4	16.1	0.8
	Average	73.0	74.5	1.5	11.3	14.4	0.7
WS (m/s)	08/2018	6.9	6.1	-0.7*	2.9	3.7*	0.6
	10/2018	8.8	7.3	-1.5*	3.2	4.1*	0.7
	01/2019	10.2	10.1	-0.1	3.4	4.7*	0.7
	04/2019	10.6	8.3	-2.3*	3.8	4.8*	0.7
	Average	9.1	8.0	6.7	22.3	35.4	0.6
WD (°)	08/2018	166.0	196.7	30.7*	78.8*	127.8	0.3
	10/2018	190.6	207.7	17.1*	53.6*	102.2	0.5
	01/2019	192.8	208.5	15.7*	41.4*	89.5	0.6
	04/2019	166.8	171.0	4.3	52.2*	99.4	0.6
	Average	179.1	196.0	17.0	56.5	104.7	0.5

266 *Indicates performance outside of Emery et al. (2001) suggested benchmarks (Table S2).

3.2 WRF-CMAQ Pollutant Simulation Performance

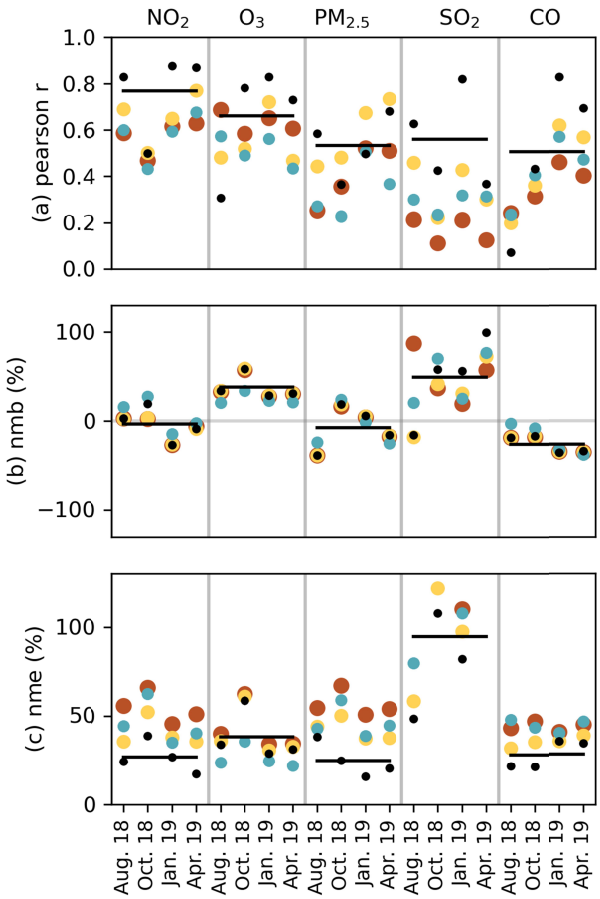


Figure 2 Model performance relative to observations of 5 pollutants over 5 different time-averaging periods: hourly (red), daily (yellow), daily maximum (blue), and monthly (black). The average of the four simulation months, each drawn from a different meteorological quarter, i.e., the annualized performance, is depicted as a black horizontal line.

To assess WRF-CMAQ pollutant simulation performance, we compare model simulated criteria pollutant concentrations to measurements of NO_2, O_3, $PM_{2.5}$, SO_2 and CO from EPA Air Quality System (AQS) monitoring stations. Model-observation comparisons occur where WRF-CMAQ grid cells contain EPA AQS stations (Figure 1c). The number of EPA monitoring stations changes within the 1.3 km domain depending on the season with 120 total stations in August 2018, 120 in October 2018, 78 in January 2019, and 125 in April 2019. Over our simulation period, there were a maximum of 15 NO_2, 67 O_3, 25 $PM_{2.5}$, 8 SO_2, and 5 CO monitors, though some stations occasionally drop offline during our simulation period. Notably, the number of O_3 monitors drops from 67 to 23 in January 2019. Due to erroneously low values, we exclude SO_2 measurements from from Ingham and La Salle County for August 2018, October 2018, and April 2019. We clean EPA station data by removing observations with negative concentration values.

We compare model simulated pollutants to AQS observations over several different time scales, including 4-monthly or annualized mean, monthly mean, daily mean, hourly mean, and daily

maximum (Figure 2). We calculate the annualized mean by averaging across our simulated months of August 2018, October 2018, January 2019, and April 2019. In the following, we provide quantitative assessments of each criteria pollutant across different temporal periods to provide context for model performance on both fine (hourly, daily, and daily maximum) and coarse scales (monthly and annual). In general, the model and observations have better agreement with longer time-averaging slices, e.g., lower biases and errors. To provide greater context for our model performance, we follow EPA recommendations (Dennis et al., 2010) and compare the performance of WRF-CMAQ over our 1.3 km domain to previously published CTM studies (Table S4). Our comparisons use fine-scale, domain-agnostic studies (<4 km horizontal resolution), or coarser-scale (>4 km horizontal resolution) studies focused on Chicago, the Midwest, or the Great Lakes region, i.e., studies with similar model domains. We select studies that use the same statistical metrics as in Table S1, simulate time periods after the year 2000, and integrate a similar CTM (WRF-CMAQ or WRF-Chem). We do not focus on other benchmark studies that use coarser and/or older versions of CTMs or emissions models (Simon et al., 2012; Emery et al., 2016).

Table 2 WRF-CMAQ performance metrics for hourly simulated 1.3 km (d03) pollutants as compared to EPA AQS station observations. The average observed value is noted as μ_d, while the predicted value is noted as μ_p.

		1.3 km Domain Performance				
Var	**Month**	μ_d	μ_p	***NMB%***	***NME%***	***r***
NO_2	08/18	10.38	10.69	2.98	55.76	0.59
	10/18	10.76	11.05	2.64	66.08	0.47
	01/19	13.13	9.59	-26.95	45.45	0.62
	04/19	11.19	10.57	-5.56	51.21	0.63
	Average	11.37	10.48	-6.72	54.62	0.57
O_3	08/18	30.28	40.25	32.92	40.04	0.69
	10/18	20.38	32.12	57.66	62.55	0.58
	01/19	24.61	31.47	27.88	34.10	0.65
	04/19	36.33	47.35	30.35	34.46	0.61
	Average	27.90	37.80	37.20	42.79	0.63
$PM_{2.5}$	08/18	12.12	7.49	-38.21	54.61	0.25
	10/18	6.78	7.89	16.43	67.18	0.35
	01/19	9.42	9.83	4.39	50.87	0.52
	04/19	7.60	6.26	-17.62	53.94	0.51
	Average	8.98	7.87	-8.76	56.65	0.41
SO_2	08/18	0.76	1.41	87.23	169.66	0.21
	10/18	0.83	1.13	37.14	139.52	0.11
	01/19	0.99	1.18	19.68	110.15	0.21
	04/19	0.79	1.25	57.57	152.94	0.12

	Average	0.76	0.96	25.33	116.39	0.18
CO	08/18	250.72	204.53	-18.42	43.25	0.24
	10/18	229.61	188.81	-17.77	46.99	0.31
	01/19	284.08	186.90	-34.21	40.55	0.46
	04/19	281.43	183.51	-34.79	44.70	0.40
	Average	261.46	190.94	-26.30	43.87	0.35

Of all pollutants, we find that WRF-CMAQ-simulated NO_2 is closest to the observations with low *NMB* and high correlations across months (Figure 2). The annualized average hourly correlation of NO_2 is high (r = 0.57; Figure 2, Table 2), while its bias is low (*NMB* < -7%; Figure 2, Table 2). NO_2 model-observation correlations are generally greater than 0.6 regardless of temporal assessment scale, except for October 2018 (Figure 2). We find slight high biases in model simulated hourly NO_2 in August (*NMB* = 3%) and October 2018 (*NMB* = 3%) and low biases in January (*NMB* = -27%) and April 2019 (*NMB* = -6%) (Table 2). When compared to previously published WRF-CMAQ studies with different domains/resolutions, our NO_2 simulation performance exceeds *NMBs* and correlations reported by (Bickford et al., 2014; Harkey et al., 2015; Vijayaraghavan et al., 2009).

Model-simulated O_3 is high relative to observations, with limited variation across seasons (Figure 2 and Table 2). We find that the annualized average correlation of simulated hourly O_3 is high (r = 0.6), but that the annualized *NMB* (38%) and *NME* (42%) are high. The highest *NME* for O_3 occurs in our October 2018 simulation (58%), which corresponds with the highest *NME* for NO2 (64%). The lowest *NMB* and *NME* are found in January 2019 (27%, 34%), which has the lowest concentrations of O_3. Compared to other studies in the Great Lakes region, our biases and errors are higher than those of Bickford et al. (2013), who ran WRF-CMAQ without two-way coupling, and Abdi-Ouskouei et al. (2020), who used WRF-Chem. Other similar CTMs studies report O_3 biases similar to those reported here (Abdi-Oskouei et al., 2020; Odman et al., 2019; Pan et al., 2017; Qin et al., 2019; Travis et al., 2016; Zhang et al., 2014a). Our high O_3 bias is mainly driven by an over prediction of simulated O_3 concentrations during periods of low observed O_3, particularly at night (Figure S1). During warm "ozone season" months when observed O_3 is high (O_3 > 60 ppb), our *NMB* is negative (-5.4% and -7.2% for August and April) and *NME* are less than 17% (Table S5). When our model performance evaluation is limited to hours when observed O_3 concentrations are greater than the 50%ile value, average annualized *NMB* is reduced to ~25% (Table S5). Lastly, and further confirming WRF-CMAQ's challenges with capturing low O_3 concentrations, the *NMB* in our model-simulated daily maximum 8-hr running average O_3 ($MDAO_3$) is ~27% when annualized, ~25% in O_3 season months, and only ±2% when $MDAO_3$ is greater than 60 ppb (Table S5).

Unlike model simulated O_3, our simulated hourly $PM_{2.5}$ concentrations have low biases and low correlations (Figure 2 and Table 2). The annualized average correlation of hourly $PM_{2.5}$ is 0.4, with *NMB* of -10% and *NME* of 56%. August 2018 hourly $PM_{2.5}$ has the largest bias (-38%) and lowest correlation (r = 0.25), while the highest *NME* (67%) and highest positive *NMB* (16%) are found in October 2018. Within the Great Lakes region, we find that our model-observation

agreement for $PM_{2.5}$ has higher correlations and similar *NME*s to Bickford et al., 2013. Compared to other WRF-CMAQ studies within the continental U.S., our $PM_{2.5}$ *NME* is lower than (Hogrefe et al., 2015), but higher than (Liu et al., 2010) and (F. Wang et al., 2021). Our $PM_{2.5}$ *NMB* and MB are similar to Hogrefe et al. (2015) and Wang et al. (2021), but lower than Liu et al. 2010 and (Torres-Vazquez et al., 2022).

The agreement of our model simulated SO_2 and CO compared to the AQS observations were the lowest of the 5 criteria pollutants (Figure 2 and Table 2). Annualized average correlation of SO_2 is 0.18, with *NMB* = 25% and *NME* = 116%. Annual average correlation of CO is 0.35, with *NMB* = -26% and *NME* = 44%. Few previous WRF-CMAQ studies report their performance of SO_2 and CO. Compared to those that do, our simulation of SO_2 had lower *NMB*s, higher *NME*s, and lower correlations (Bickford et al., 2014; Campbell et al., 2019).

3.3 Domain-wide Characterization of WRF-CMAQ Simulated Pollutants

In Figure 3, we show monthly-average simulated NO_2, O_3, and $PM_{2.5}$ concentrations over the 1.3 km domain (d03) for each season. Simulating pollutants at a 1.3 km spatial resolution facilitates the characterization of distinct urban-rural patterns, the influence of Lake Michigan on regional O_3 distributions, pollutant hotspots over highway corridors, stationary emitting sources, and urban centers. In the following, we individually discuss domain-wide analyses of each pollutant and then highlight the model's characterization of pollutants within the city of Chicago.

The simulated NO_2 concentrations largely track high-population areas and highway corridors (Figure 3a-e). In all seasons, the interstate highway system that connects population centers is highlighted by the spiraling web of roadways with elevated NO_2 concentrations. The lowest NO_2 concentrations in our domain are simulated over northernmost and easternmost portions of Lake Michigan, in areas distant from emissions sources. Likewise, rural areas distant from roadways have low NO_2 concentrations. We find that the average annual urban concentration of NO_2 in our domain is simulated to be 3.5 ppb (59.8%) higher than average concentrations in rural portions of our domain (Table S6.0 & Figure 3). Across seasons, domain-wide NO_2 concentrations tend to correspond to changes in simulated NOx emissions (Figure S3). In January 2019, domain average NO_2 concentrations are highest (μ = 3.2 ppb; Figure 3c), which corresponds to our highest simulated NOx emissions (Figure S3). The lowest domain average NO_2 concentrations occur in April 2019 (μ = 2.4 ppb; Figure 3d), which co-occurs with low NOx emissions (Figure S3) and the highest domain average O_3 concentrations (Figure 3i).

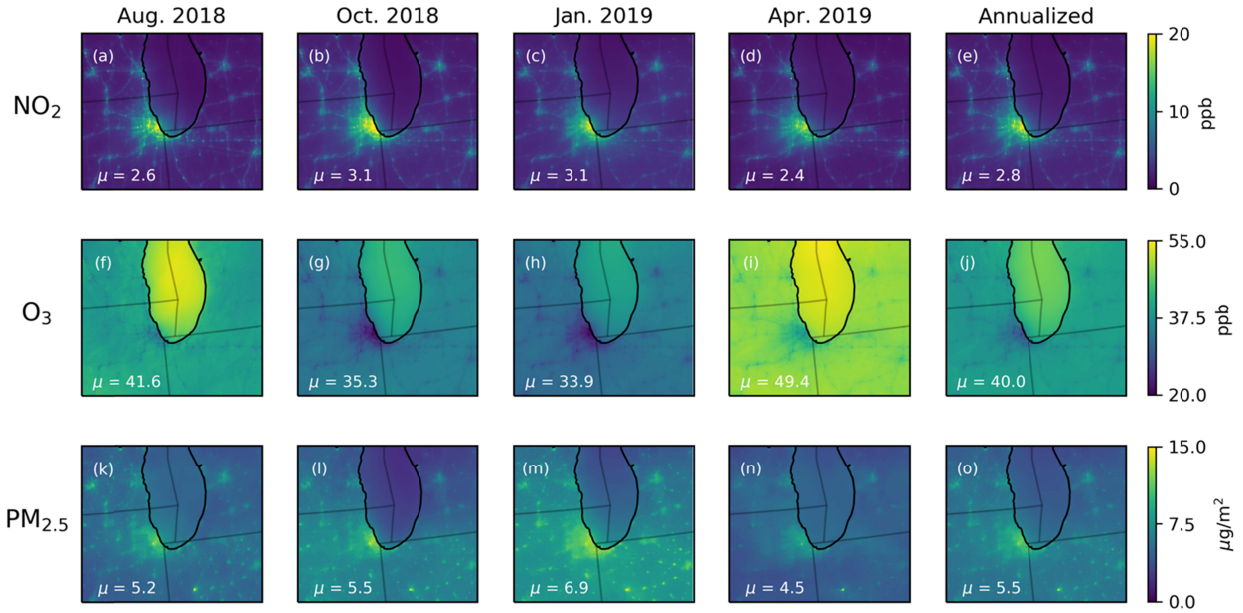


Figure 3 Monthly and annualized average WRF-CMAQ simulated NO_2 (a-e), O_3 (f-j) and $PM_{2.5}$ (k-o) for the 1.3 km simulation domain (d03). From left-to-right each column presents August 2018, October 2018, January 2019, and April 2019, with the domain-average (μ) concentrations annotated in the lower left of each panel. The right-most column provides the annualized average.

In contrast to NO_2, O_3 concentrations are simulated to be relatively low over urban areas and roadways, with individual highway corridors apparent (Figure 3f-j). Across seasons, the spatial distribution of O_3 concentrations is relatively consistent, however the magnitude of O_3 concentrations varies by season. Of the four months that we simulate, we find that the highest O_3 concentrations occur in April 2019, (μ = 49.4 ppb; Figure 3i), while the lowest concentrations O_3 simulated in January 2019 (μ = 33.9 ppb; Figure 3h). Over simulated warm season months (i.e., April 2019 and August 2018), we find that domain-wide O_3 concentrations are ~1.5 times higher than cool-season concentrations. Comparing concentrations across the urban-rural divide (Figure S2), we find that O_3 over urban areas is simulated to be ~3.3 ppb (9.4%) lower than over rural areas, with the greatest urban-rural difference in cool season months (Table 3).

Across all seasons, the highest simulated concentration of O_3 in the 1.3 km domain occurs over Lake Michigan (Figure 3f-j). The simulation of elevated over-lake O_3 concentrations is consistent with previous Lake Michigan observation and modeling campaigns (Doak et al., 2021; Dye et al., 1995; Foley et al., 2011), and similar to other studies focused on inland bodies of water (e.g., Chesapeake Bay, (Goldberg et al., 2014)). Elevated O_3 over Lake Michigan is thought to be dependent on the circulation of primary pollutants from land to lake via the lake-breeze. Elevated O_3 over the lake occurs through the following idealized sequence of events: (1) In the morning land-based emissions (O_3 precursors) are transported over the lake by a land-breeze, which combine with shipping emissions, and are trapped below a shallow boundary layer. (2) As the day warms, the land-lake temperature gradient weakens, and the land breeze dwindles. Fewer NO_x emissions are transported to the lake. (3) As sunlight increases photochemical production of O_3 over the lake is enhanced. Due to the lack of depositional

pathways over the lake, O_3 accumulates. (4) On days where a lake breeze forms, O_3 is advected inland, often to areas not responsible for the original precursor emissions. Lake breeze effects are primarily a warm season phenomenon, however, the lack of over-lake depositional pathways also contributes to elevated cool season O_3 concentrations (Figure 3f-j) (Doak et al., 2021; Dye et al., 1995). In our simulations land-lake O_3 concentration differences are greatest in August 2018, as the average concentration of O_3 over land (39.8 ppb) is 11 ppb lower than the average O_3 concentration over the lake (50.4 ppb).

Compared to NO_2 concentrations, domain-wide simulated $PM_{2.5}$ concentrations show greater spatial homogeneity in that elevated hot spots have a larger diffusive footprint (Figure 3k-o). Across months, the spatial pattern of simulated $PM_{2.5}$ concentrations is relatively consistent and largely tied to the location of emission sources. Despite consistent spatial patterns across months, the relative magnitude of $PM_{2.5}$ concentrations is influenced by meteorological conditions (e.g., boundary layer height and wind speeds), the magnitude of seasonal primary PM emissions, and secondary PM pollutant formation reactions. The domain-wide average concentration of $PM_{2.5}$ peaks in January 2019 (μ = 6.9 μ/m^3) and is lowest in April 2019 (μ = 4.5 μ/m^3), which mirrors the pattern of emissions of PM and its precursors (Figure S2). Both stationary and mobile sources of $PM_{2.5}$ typically co-emit NO_x emissions, as such simulated $PM_{2.5}$ hotspots tend to co-occur with NO_2 hotspots over urban areas, highways, and stationary sources (Figure 3a-e and k-o). However, compared to NO_2, the $PM_{2.5}$ hotspots are more spatially diffuse, likely due to longer $PM_{2.5}$ lifespans, secondary formation of $PM_{2.5}$, and the influence of meteorology, which disperses $PM_{2.5}$ concentrations from point sources. Despite the large number of sources within the domain of $PM_{2.5}$, the concentrations in urban areas are simulated to be 22% higher than rural areas (Table 3). Simulated grid cells with the highest concentration of $PM_{2.5}$ occur outside of urban areas and are primarily associated with emissions from industrial and manufacturing point sources.

3.4 Domain-wide Characterization of WRF-CMAQ Simulated Pollutants

Our domain-wide analysis demonstrates the ability of the 1.3 km WRF-CMAQ simulations to characterize differences in urban-rural regimes and identify pollutant hotspots, however it does not highlight the ability of the model to resolve and characterize neighborhood-scale air quality. To demonstrate this ability, we provide an in-depth analysis of a sub-region of the 1.3 km modeling domain, i.e., the city of Chicago (Figure 4a). Chicago sits close to the center of our 1.3 km domain, and in Figure 3 is identifiable as both an NO_2 and $PM_{2.5}$ hotspot at the southwest corner of Lake Michigan. Chicago has a population of 2.7 million people that are divided amongst 77 named community areas (Figure 4a). Major sources of emissions within Chicago include transportation, industry, and buildings. The city has 5 major interstate highways (I-290, I-294, I-90, I-94, I-55, I-57) that loosely outline the City's lakeside central business district or "Loop". There are two airports within City limits, O'Hare in the northwest and Midway in the south central. Most industrial activities occur on the west and southwest sides of the city.

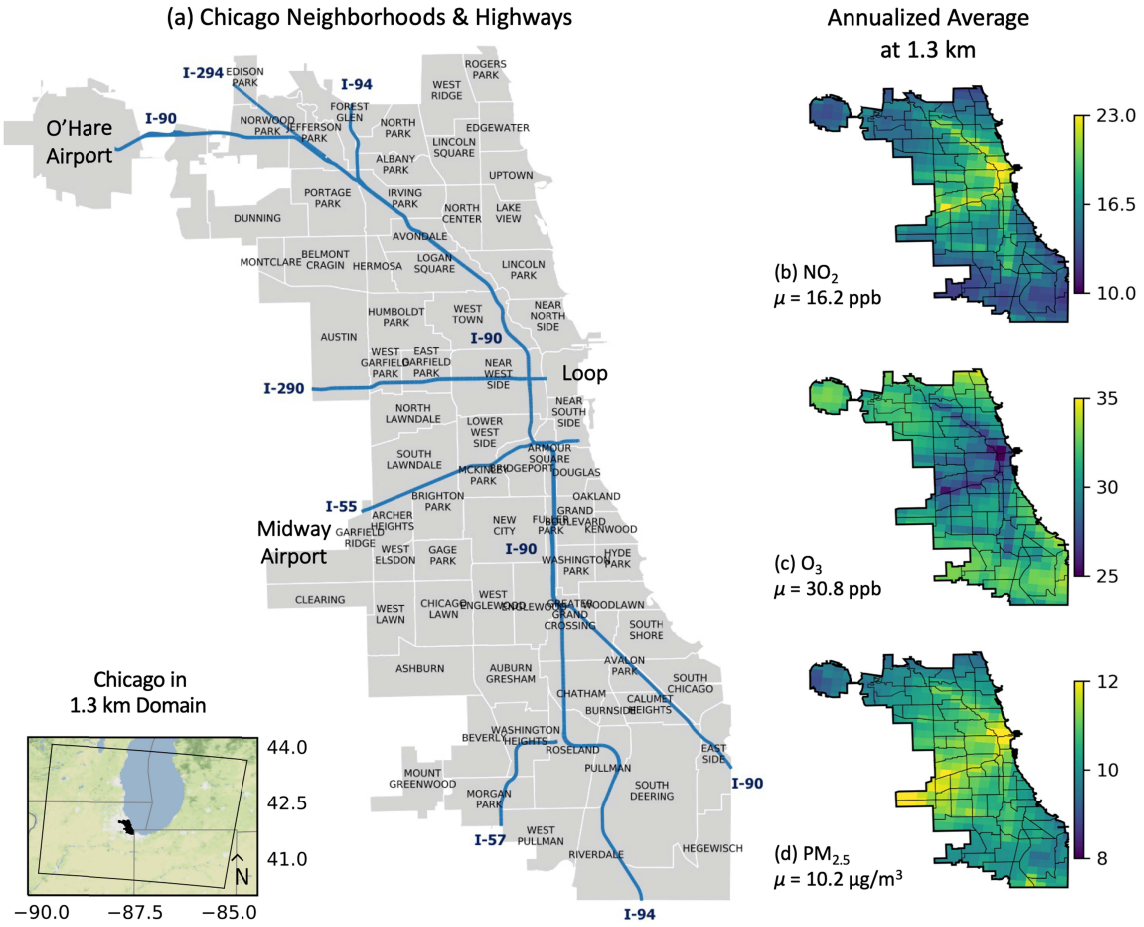


Figure 4 Chicago geography and simulated pollutants. (a) Neighborhoods and highways within the city of Chicago. Inset at lower left depicts location of Chicago (black) within the 1.3 km domain. In the right column, we provide annualized WRF-CMAQ simulated (b) NO_2, (c) O_3, and (d) $PM_{2.5}$ concentrations over Chicago and average concentrations within city limits (μ).

In Figure 4b-d, we show the city's average annualized concentrations of NO_2, O_3, and $PM_{2.5}$. Simulated concentrations of criteria pollutants are higher within Chicago than the domain average, so Figure 4 has a different color bar than that used in Figure 3. For all pollutants, the 1.3 km WRF-CMAQ simulations reveal robust spatial heterogeneity amongst neighborhoods. Spatial gradients are particularly substantial for simulated NO_2, with concentrations in some neighborhoods double those in others. The predominant spatial pattern of simulated pollutants strongly corresponds with the interstate highway system (Figure 4a), although the highest concentrations of NO_2 and $PM_{2.5}$ are simulated where the lowest concentrations of O_3 are simulated. O_3 concentrations are elevated in neighborhoods that abut Lake Michigan, consistent with the influence of the lake breeze, and in neighborhoods without interstate highways. In the following paragraphs we discuss annualized pollutant pattern details across Chicago, as well as patterns of individual months.

455 The highest annualized concentrations of NO_2 (where $\mu > 19.5$ ppb) are simulated on the West
456 side of Chicago and in the Loop, where highways are prevalent (e.g., I-90, I-290, and I-55) and
457 simulated NO_x emissions are high (Figure S3). The lowest NO_2 concentrations (where $\mu_{annual} <$
458 11.3 ppb) are simulated in the lake-front neighborhoods, with the exception of those in the Loop.
459 Lakefront neighborhoods are east of the main interstate highways, where lower NO_x emissions
460 and ventilation contribute to the relatively low simulated NO_2 (Figure S3). Across seasons, NO_2
461 concentrations within Chicago remain highest over the 3 most-trafficked inter-state highways: I-
462 290, I-90, and I-94 (Figure 4b). Simulated NO_2 concentrations are lowest in April 2019 (μ =
463 14.2 ppb) and highest in October 2018 (μ = 18.7 ppb), although the greatest NO_2 bias was also
464 found in October 2018. Compared to the full model domain (μ = 2.4-3.2 ppb; Figure 3a-d),
465 average NO_2 over Chicago is nearly 5 times higher (μ = 14.2-18.7 ppb) across seasons (Figure
466 5a-d). The differences in average NO_2 concentrations over Chicago are reflected in the
467 emissions, which are highest in October 2018 and are lowest in April 2019 (Figure S3).

468 Simulated annualized and individual month O_3 concentrations within Chicago tend to be the
469 spatial inverse of simulated NO_2 concentrations (Figures 4 & 5). The lowest concentrations
470 ($\mu_{annual} < 28.5$ ppb) of O_3 are simulated on the West side of the city, near the interstates. These
471 locations are also simulated to have the highest NO_2 concentrations, i.e., O_3 is suppressed via
472 titration by NO. O_3 concentrations are highest in the warm months, with August 2018 ($\mu = 35.7$
473 ppb) and April 2019 concentrations ($\mu = 41.1$ ppb) nearly double October 2018 ($\mu = 23.3$ ppb)
474 and January 2019 concentrations ($\mu = 22.7$ ppb). Average annualized O_3 concentrations in
475 Chicago ($\mu = 30.7$ ppb) are simulated to be significantly lower than the domain average because
476 of the lake reservoir of O_3 ($\mu = 40.0$ ppb). Even when land-only O_3 concentrations are isolated,
477 Chicago has concentrations that are slightly lower than the rest of the full model domain (μ =
478 38.8 ppb). Warm-season O_3 is highest near Northern lake-front neighborhoods, which are distant
479 from the major interstates, have low NO_x emissions (Figure S3), and subject to lake breeze
480 advection of the reservoir of O_3 over Lake Michigan. In the cool months, O_3 concentrations are
481 simulated to be highest on the western edges of the city. However, the cooler months have a
482 lower range of O_3 concentrations (±8.9 ppb) than warmer months (±15.5 ppb).

483 Annualized $PM_{2.5}$ concentrations in Chicago correspond well with the spatial patterns of the
484 interstate system, though the $PM_{2.5}$ footprint is spatially more extensive than that of NO_2 (Figure
485 4). $PM_{2.5}$ concentrations in Chicago are simulated to be 2 times higher than the average
486 concentration of the full model domain ($\mu_{domain,annual} = 5.5$ $\mu g/m^3$, $\mu_{Chicago,annual} = 10.2$ $\mu g/m^3$).
487 $PM_{2.5}$ concentrations peak on the west side of Chicago near Midway airport and the intersection
488 of I-290 and I-55 with I-90 (μ = 12 – 13 $\mu g/m^3$). The lowest concentrations of $PM_{2.5}$ occur on
489 the lakefront (μ = 8 – 10 $\mu g/m^3$). Similar to O_3, Chicago $PM_{2.5}$ levels show strong seasonal
490 variations, though the simulated concentrations of $PM_{2.5}$ are highest in the seasons when O_3 is
491 lowest. As such, simulated $PM_{2.5}$ peaks in the cooler months (Figure 5j-k) and is lowest in April
492 2019 (8 $\mu g/m^3$). Areas of high $PM_{2.5}$ in Chicago are consistent across seasons, in particular on the
493 west side of the city and within the Loop.
494

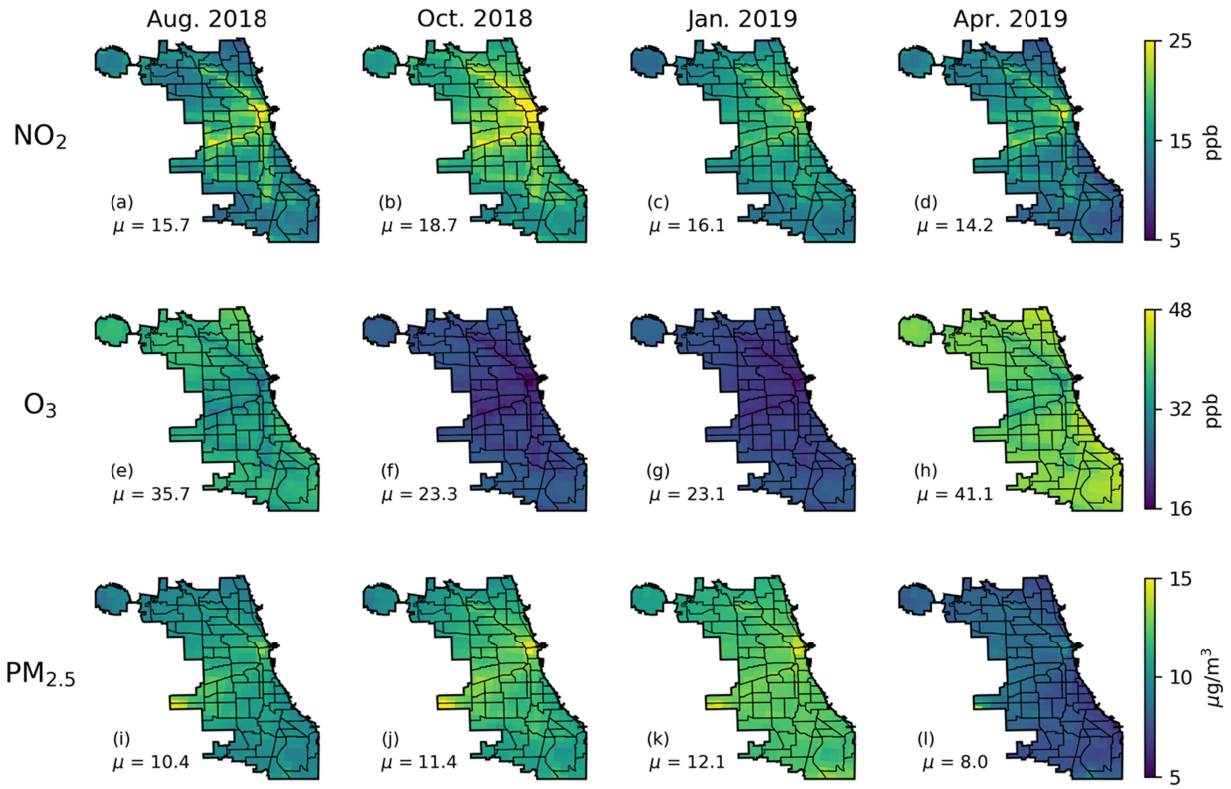


Figure 5 Monthly average NO_2 (a-d), O_3 (e-h) and $PM_{2.5}$ (i-l) concentrations over Chicago, as simulated in the 1.3 km domain. Columns depict August 2018, October 2018, January 2019, and April 2019 simulations. Chicago-average concentrations are annotated (μ).

3.5 Benefits and Disbenefits of Increased Model Resolution

Our nested modeling framework facilitates assessment of the potential benefits and/or disbenefits of attempting to resolve neighborhood scale meteorology and air pollutants with a CTM. That is, given that our methodology simulates atmospheric chemistry and meteorology at both 4 km and 1.3 km resolutions, an assessment of increased spatial resolution on model performance is possible. However, our chosen methodology does not provide a pure spatial resolution sensitivity analysis, i.e., while the underlying emissions data of each simulation is the same (NEI 2016v1), the meteorologically informed emissions are slightly different due to differences in the 4 km and 1.3 km WRF simulation. Further, we compare stations to grid-cells for each domain, so the 1.3 km grid cell covers a smaller area than the 4 km grid cell. Despite this imperfect sensitivity analysis, in the following we compare model performance at both resolutions. We restrict our comparison to model performance over the 1.3 km domain. As such, we use the same 10 NCDC meteorological observing stations and the ~125 EPA AQS stations shown in Figure 1 and discussed in the WRF-CMAQ simulation comparison. After this performance comparison, we investigate spatial changes in simulated pollutants over Chicago at the different resolutions.

For the simulated meteorology, we find that the increase in resolution benefits the performance of WS and WD, has no influence on RH, and has disbenefits for T2 (Table 1 & Table S7). Higher resolution provides the greatest benefit to simulated WD as *MB* and *ME* are lower at 1.3

km and correlations are higher. In contrast, we find that observed T2 is better captured in the 4 km simulation, as lower bias and error and higher correlations are found compared to the 1.3 km simulation. Within Chicago (with just one LCD station) the 1.3 and 4 km simulations perform similarly to their domain-wide performance, with WD showing slightly lower biases in the 4 km domain (Table S8).

For simulated pollutant concentrations we find higher model-observation correlations in the higher resolution (1.3 km) simulation, though this improvement is coupled with increased *ME* (Table S9). The 1.3 km simulation showed higher correlations than then 4 km simulation for all criteria pollutants, though this increase was marginal ($\Delta r < 0.1$, Table S9). The 1.3 km simulation of NO_2 has a closer agreement to the EPA stations, but this comes with slightly higher normalized errors (+0.25%; Table S9). On average, the 1.3 km simulation *NMB* was lower than that of the 4 km model simulation for each season, which came at a trade-off, as *NME* was only lower in the 1.3 km simulation in April 2019. The correlation between AQS observation and model outputs for NO_2 were similar for the 1.3 km and 4 km resolution simulations. The 1.3 km simulation lowered the *NMB* by 8% in August 2018 and January 2019, with marginal bias improvement in April 2019 and January 2019. In contrast, the 1.3 km simulation of O_3 showed higher *NMB* and *NME* than the 4 km simulation. For $PM_{2.5}$, we find that the 1.3 km resolution simulation has a lower *NMB* than the 4 km simulation for 3 out of 4 seasons, but the *NME* is marginally higher (0.9%) in the higher resolution simulation. Simulated SO_2 showed the largest improvement with finer model resolution ($\Delta NMB = 10\%$, $\Delta r = 0.03$), however this was also the pollutant with the lowest performance in both the 1.3 km and 4 km domains. CO had slightly better performance in the 4 km domain, as the *NME* and *NMB* were 0.5% to 1% higher, respectively, in 1.3 km simulation.

These meteorological and pollutant performance analyses are limited due to the low number of sensors relative to the number of grid cells simulated in our modeling domain (125 EPA stations and 10 NCDC stations vs 90,720 grid cells). In addition, the finer resolution creates opportunities for local-scale meteorological processes to influence agreement, particularly for pollutants which are not well mixed in the atmosphere (Zhang et al., 2014a). The measurement-prediction relationship can be greatly influenced by model grid cell size, plumes, and wind speeds and directions. For example, while the relative amount of SO_2 simulated in a plume may be correct, due to the increase in the number of grid cells in a higher resolution simulation, the probability that an erroneously simulated wind direction will adversely influence model grid cell-observation fidelity increases.

Given that model performance when assessed against limited station observations is similar between the 1.3 km and 4 km simulations, we now turn our focus to the primary advantage of the higher resolution simulation, i.e., the ability to characterize neighborhood-scale air quality. As Chicago is the densest metropolitan area in our modeling domain, we focus our analysis on air quality differences between the 1.3 and 4 km simulations within city limits. We find that over Chicago the 1.3 km simulation has higher average NO_2 and $PM_{2.5}$ concentrations and lower O_3 concentrations (Figure 6) than the 4 km simulation. Differences in pollutant concentrations at different model resolutions can be caused by several factors, including differences in the underlying emissions data, the ability to resolve fine-scale processes, and the nuances of grid cell-geography-chemistry feedbacks. In the following, we provide examples of each.

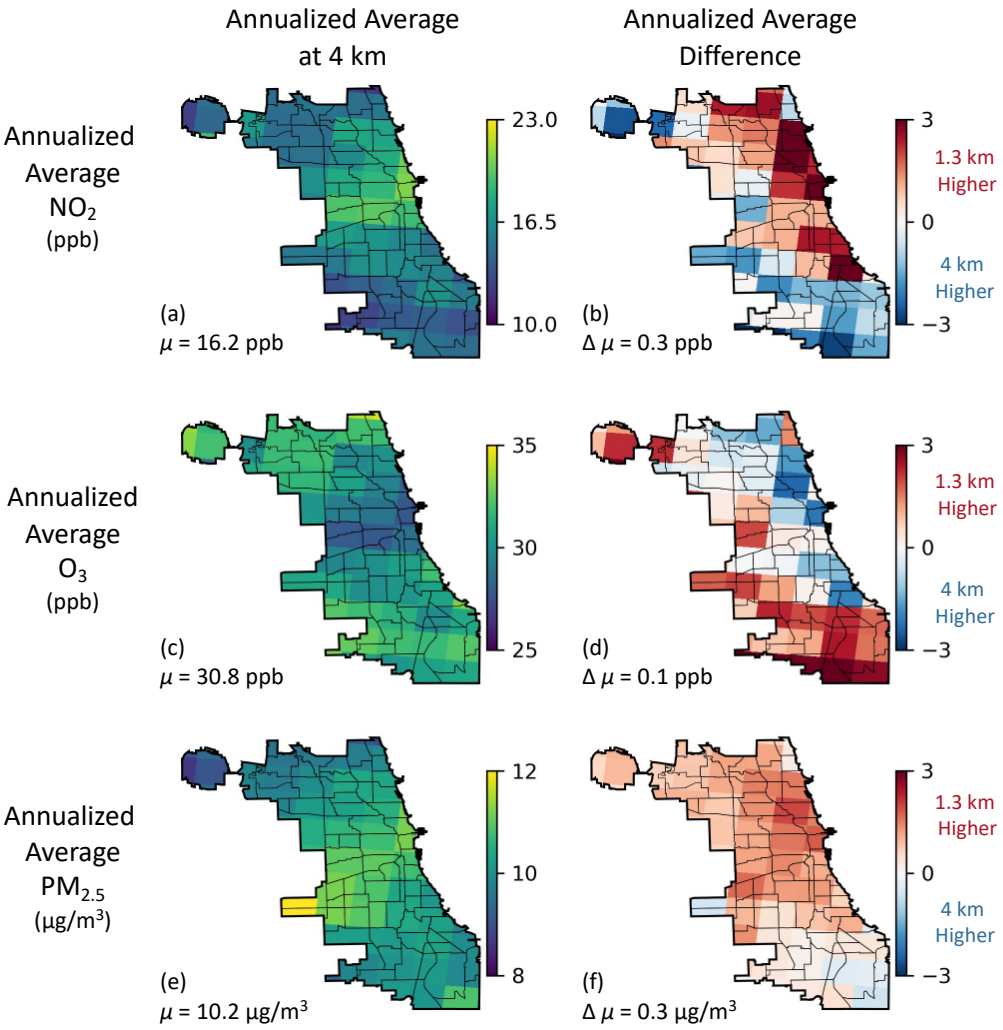


Figure 6 Average annualized NO_2 (a), O_3 (c), and $PM_{2.5}$ (e) within Chicago city limits as simulated in the 4 km simulation. (b,d,f). Relative pollutant differences between 1.3 km and 4 km resolution simulations. Average Chicago concentrations (μ) are annotated, with average differences also noted.

The overall higher concentrations of NO_2 and $PM_{2.5}$ simulated over Chicago at 1.3 km are in part due to differences in the underlying meteorologically-informed emissions. In the 1.3 km simulation, NO_2 concentrations are ~0.3 ppb higher than in the 4 km simulation. In the 1.3 km simulation, the rate of NO_2 emissions over Chicago is 2.5% higher than the in the 4 km simulation (Figure S4a). It is potentially these higher NO_x emission rates that result in slightly lower over-Chicago O_3 concentrations (-0.07 ppb) at 1.3 km, though VOC emissions are higher at 1.3 km as well (Figure S4b). Similarly, for $PM_{2.5}$, organic and elemental carbon emissions over Chicago are 1% higher in the 1.3 km emissions than in the 4 km emissions (Figure S4c). For $PM_{2.5}$ pollution, this difference leads to most (59/73) of Chicago's neighborhoods having higher $PM_{2.5}$ concentrations when simulated at the higher spatial resolution.

In our simulations, we also note that the higher resolution simulation better captures fine-scale processes. This ability is best illustrated by examining emissions and pollutants over highways (compare Figures 4 and 6). While the 4 km simulation shows elevated NO_2 and $PM_{2.5}$ (and depressed O_3) concentrations over the center of the city, the larger grid cells smooth the underlying emissions (Figure S4a, Figure S5) which leads to lower concentrations of NO_2 and $PM_{2.5}$ over highway pollutant corridors, features that are critical to resolve when assessing neighborhood-scale exposure. These finer-scale processes captured in the 1.3 km simulation may have profound implications for environmental justice-focused analyses.

4 Discussion

In the above we present the first neighborhood-scale (1.3 km) two-way coupled WRF-CMAQ simulations focused on the southern Lake Michigan-Chicago region. We perform hindcast simulations of individual months from each season and assess the model's performance against meteorological and pollutant station observations, as well as against coarser resolution (4 km) simulations. Below we summarize our results and discuss notable findings and experimental caveats.

In our WRF-CMAQ simulations, we show that the WRF-simulated meteorological variables WD, WS, and T2 meet performance criteria suggested by Emery et al. (2001). The lowest performing simulated meteorological variable is WD, a variable that models have historically struggled to reproduce with high fidelity and which has previously been shown to be sensitive to model resolution, boundary layer parameterization, and land cover schemes (Carvalho et al., 2012). In our simulations, we find that WRF-CMAQ best-captures observed WD in January 2019 and October 2018, but struggles in August 2018 and April 2019, likely due to more diffuse warm season winds, similar to findings presented in Zhang et al. (2014). In addition, recent WRF simulations have demonstrated the influence of different lake temperature datasets on meteorology, particularly air temperatures and convection, in domains near to Lake Michigan (J. Wang et al., 2022). Future work should assess the role of lake temperatures, and uncertainties therein, on the simulation of pollutants.

In the CMAQ portion of our two-way coupled WRF-CMAQ simulation, NO_2 concentrations show the highest agreement amongst EPA criteria pollutant AQS station observations. We find a similar correlation between observed and simulated O_3 concentrations, which reflects the strong anticorrelated relationship between O_3 and NO_2. However, we find that O_3 is biased high for all 4 months of simulation. Previous studies have noted that high O_3 biases in CTMs can be due to

overestimated NO_x emissions (Huang et al., 2018), excessive boundary layer mixing (Travis et al., 2016), long-range transport and boundary conditions (Sharma et al., 2017), and low night time titration by NO (Sharma et al., 2017). Compared to NO_2 and O_3, simulated $PM_{2.5}$ has lower correlation with EPA AQS station observations, though our model performance is comparable to results reported by many previous studies (e.g., Wang et al. 2021 and Torres-Vazquez et al., 2022). Given that SO_2 and CO both contribute to secondary PM formation, the relatively poor performance of WRF-CMAQ with respect to the simulation of these constituents likely also influences the $PM_{2.5}$ model-observation agreement. Previous studies have reported that model-station agreement of $PM_{2.5}$ can be strongly influenced by wind direction, wind speed, transport, and emissions inventories (Hughes et al., 2021; Zhang et al., 2014a) and it is likely that these meteorological factors also play a role here. For instance, in the results presented here, we employ MOVES2014a which does not account for emissions from off-road idling of heavy-duty vehicles. MOVES3, released in 2020 (EPA, 2020) does include these processes, which may be critical for more accurate simulation of PM, particularly in high-density warehouse environments common within urban settings.

Since the two-way coupled WRF-CMAQ methodology employs nested domains of increasing spatial resolution, we take the opportunity to discuss differences, advantages, and disadvantages of neighborhood-scale (1.3 km) simulations versus those performed at coarser resolutions (4 km). By and large, when model results are assessed against meteorological and pollutant station observations, we find comparable performance between the two resolutions. We note a few occurrences of slightly degraded model-observation fidelity at higher-resolution (e.g., T2 and O_3), but primarily find that higher resolution simulations marginally improve hindcast simulations of both meteorology and atmospheric chemistry, similar to previous thematically-similar studies (Torres-Vazquez et al., 2022). We note that our 1.3 km to 4 km simulation comparison is not a pure resolution-focused sensitivity experiment. MOVES emissions processing influences on-road sector emissions, and due to differences between the 1.3 km and 4 km WRF-simulated meteorology, on road emissions differ over roadways (Figure S3). However, total emissions are the same within the 1.3 km domain subset of the 4 km domain. Despite emissions differences of 1-6% over Chicago (Figure S3), pollutant concentrations differ by only 1-2% between the 1.3 km and 4 km simulations (Figure 6). Ultimately, we find that the most valuable feature of increasing model resolution comes from the finer characterization of emission sources and subsequent pollutant concentrations. The ability to resolve air pollution at neighborhood-scale resolutions, using a physics- and chemistry-based numerical model, represents a step change in air quality research, and continued efforts should be made to both improve model performance and apply these tools to fundamental research queries in the fields of health, policy, and environmental justice.

Characterizing neighborhood-scale spatial heterogeneities in pollutant concentrations over urban settings, such as Chicago, is critical for better understanding health impacts and constraining the contribution of pollutants to inequitable impacts across population subgroups. In our simulations, we find that Chicago has 2 to 5 times higher NO_2 and $PM_{2.5}$ concentrations than neighboring rural areas (Fig. 2), and within city limits annualized pollutant concentrations between neighborhoods can vary by a factor of 1.8 (Figure 4). Results such as these suggest that summarizing city-wide air quality using limited observations or coarse-scale model simulations could be problematic. To highlight the utility of high resolution spatially resolved model simulations, we analyze zonally averaged annualized pollutant concentrations from Chicago's

western suburbs to Lake Michigan in the east (Figure 7). Pollutants over this zonal swath display a distinct west-to-east profile, with NO_2 and $PM_{2.5}$ peaking over the core of the city, and relatively high O_3 concentrations over the lake and in the western suburbs. In general, zonal patterns of O_3 concentrations are the inverse of simulated NO_2 concentrations. This inverse pattern is replicated when comparing $PM_{2.5}$ to O_3, albeit with lesser fidelity. In addition to elevated NO_2 over the city center, a western peak is simulated near O'Hare International Airport (Figure 7a). Despite the inverse NO_2-$PM_{2.5}$-O_3 zonal pattern, some sections of the city do see relatively high co-occurring concentrations of NO_2, O_3, and $PM_{2.5}$. For example, Chicago's west side, near -87.8°W, has relatively high concentrations of each pollutant (Figure 7a), likely due to the confluence of highways and industrial areas. Our finding of elevated west side pollutants is consistent with previous health-focused work that has identified elevated clusters of air quality related diseases on Chicago's west side (Gupta et al., 2008).

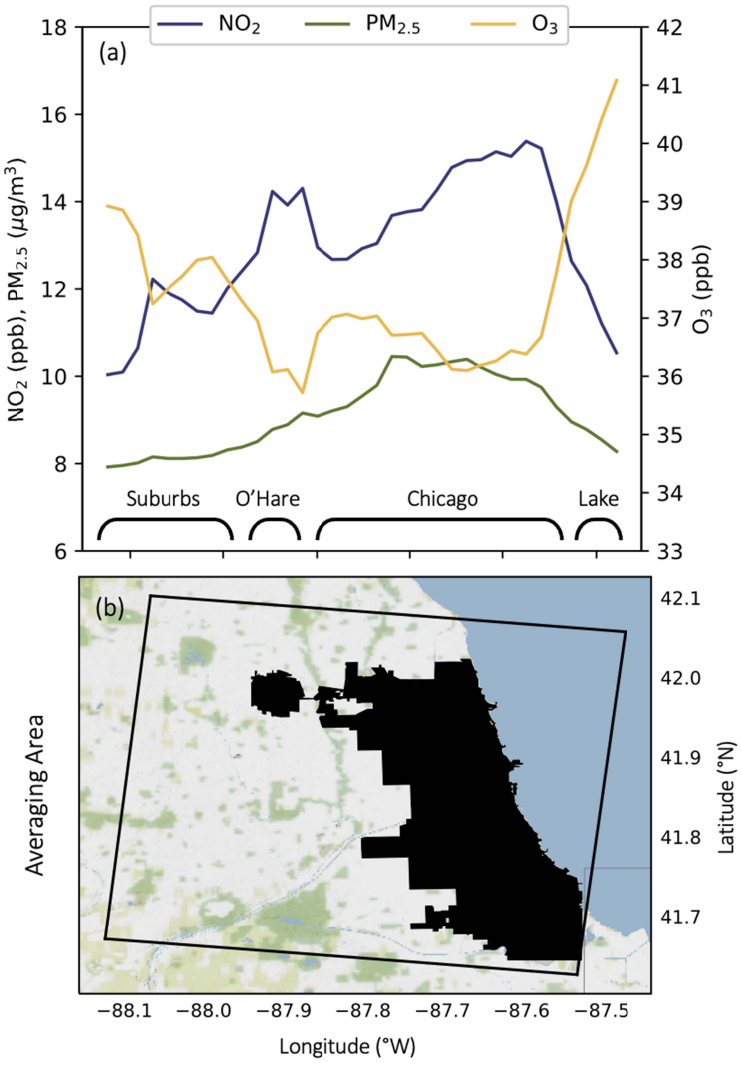
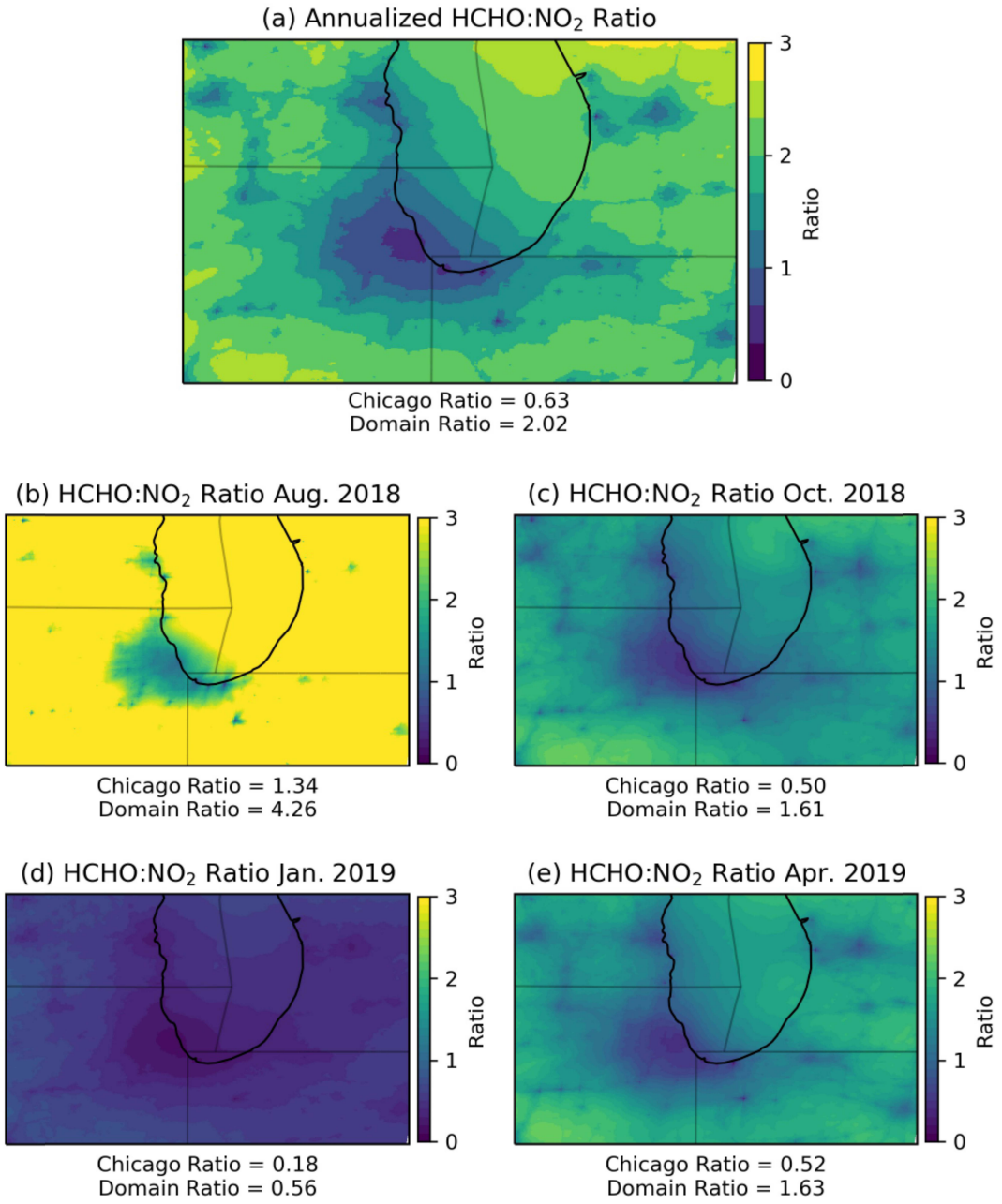


Figure 7 (a) Zonal average of pollutants over Chicago region, from suburbs to city center to lake. Mean pollutant concentrations of NO_2, $PM_{2.5}$, and O_3 are provided in and (b) the footprint

of the averaging domain is depicted around Chicago. Concentrations are plotted across longitudes and stretch from the western suburbs to Lake Michigan in the east.

Neighborhood-scale simulations may also be critical for designing effective mitigation and abatement strategies. For example, given Chicago's current EPA O_3 non-attainment status, designing strategies that effectively target O_3 precursors requires foreknowledge of the City and region's chemical regime. Localized formation of tropospheric O_3 is a nonlinear process that depends on the relative abundances of precursor emissions, the transport of O_3 and other precursor emissions from upwind areas, and the scale and magnitude of local sinks. Despite this complexity, O_3 production environments are often simplified as either NOx-limited or VOC-limited regimes (Sillman et al., 1990; Kleinman, 1994). An area is considered "NOx-limited" when VOCs are more available than NO_2, and as such, O_3 production is limited by the radical termination of NO_2 by OH. O_3 production is "VOC-limited" when NO_x is abundant, and O_3 production is limited by the availability of peroxy radicals from VOC oxidation (Schroeder et al., 2017). To determine if areas are NOx- or VOC-limited, the ratio of HCHO to NO_2 can be analyzed. This ratio serves as a proxy to describe the chemical loss of $HO_2 + RO_2$ (LROx) over the chemical loss of NO_x (LNOx) (Shroeder et al., 2017). While there are not universally agreed upon HCHO:NO_2 ratio values to delineate if an area is NO_x- or VOC-limited, it is accepted that very low ratios (e.g., < 1) of HCHO:NO_2 indicate an area is VOC-limited, very high ratio values (e.g., > 2) indicate that an area is NOx-limited, and values between the high and low range are considered "transitional" (Jin et al., 2017).

Several previous studies using a variety of methods have attempted to quantify the HCHO:NO_2 ratio of Chicago. These studies have arrived at different conclusions over the years, including some that have found Chicago to be NO_x-limited (Laughner & Cohen, 2019), VOC-limited (Blanchard et al., 2008; Koplitz et al., 2022), or in a transitional state (Jin et al., 2020; Jing & Goldberg, 2022). In Table S10 and Figure 8, we provide the HCHO:NO_2 ratios from our WRF-CMAQ simulations. We compute column average HCHO:NO_2 ratios for each individual simulated month and the annualized mean for TropOMI's overpass time (1 – 3 PM) and the daytime average (7 am – 7 pm), though the ratio values computed for Chicago do not change substantially between overpass and daytime average (Table S10). We find robust heterogeneity in column average HCHO:NO_2 ratios in both space and time. HCHO:NO_2 ratios vary across seasons (Table S10, Figure 8), with the lowest Chicago-average ratio simulated in January 2019 (0.20, VOC-limited) and highest in August 2018 (1.57, transitional). Chicago is simulated to be VOC-limited for most months except August 2018, where it is in a transitional regime. The HCHO:NO_2 ratio is higher over the full 1.3 km domain than over Chicago, indicating the full domain is more NO_x limited than the city itself. In the annualized mean plot (Figure 8), we find a robust spatial gradient in the HCHO:NO_2 ratio, suggesting that Chicago's chemical regime may change over relatively short distances, and in modeling studies, may be resolution dependent. The consequence of this finding, and its effect on policy design for O_3 precursor control alludes to the complexity of the system and the benefits of resolving atmospheric chemistry and pollutants at the neighborhood-scale.



709

710 **Figure 8** Daytime average HCHO:NO_2 ratios over 1.3 km domain for the (a) annualized and
711 monthly (b-e) simulations.

712 Despite the promising neighborhood-scale results reported here, there are several caveats to bear
713 in mind when considering our two-way coupled WRF-CMAQ results. Chief amongst these
714 considerations is the use of four individual months and their annualized means to characterize the
715 region's air quality and atmospheric chemistry regime. Neighborhood-scale CTM simulations
716 are computationally expensive, which has limited our ability to simulate full seasons or multiple
717 years. Previous studies have demonstrated that internal meteorological variability can have

718 profound consequences on pollutant concentrations (Garcia-Mendez et al., 2017), and this facet
719 should be remembered when considering our results. A key example/consequence from our
720 study is the anomalously high O_3 concentration in our April 2019 simulation. Typically, O_3 in
721 this region peaks in July, however April of 2019 (our chosen simulation month) had higher O_3
722 concentrations than the typical summer O_3 season. A second key consideration of our study is
723 that the EPA air quality monitoring system was not designed with CTM-validation in mind. AQS
724 sensors are relatively sparse and very often not within urban settings. As such, we use EPA data
725 here, but would advocate for the use of hyper-local observing networks to operationally monitor
726 neighborhood-scale air quality and perform model validation.

727 **5 Conclusions**

728 In the above, we present the first neighborhood-scale two-way coupled WRF-CMAQ simulations
729 to be performed over a Chicago-centric southern Lake Michigan domain. Both the
730 meteorological and chemical components of our model largely perform at or above
731 recommended standards. We note that our 1.3 km simulation outperforms our 4 km simulation
732 with respect to most air quality-relevant meteorological variables. In terms of chemical
733 performance, we observe that the 1.3 km simulation outperforms the 4 km simulation with
734 respect to grid cell-to-observation station comparisons for NO_2, O_3, $PM_{2.5}$, and CO
735 concentrations. SO_2 is the only pollutant that showed higher model-observation fidelity at the
736 coarser model resolution, but this was also the chemical with the lowest model-station agreement
737 at both the 1.3 km and 4 km resolutions. Consideration of these performance assessments should
738 be tempered by knowledge that both meteorological and pollutant observing networks allow for
739 model-to-observation comparisons at a maximum of 0.1% of simulated grid cells.

740 Neighborhood-scale, 1.3 km simulations, are made possible by spatial surrogates curated for the
741 region by LADCO. These surrogates facilitate the simulation of fine-scale features and
742 processes, none more evident than the effect of resolving on-road emissions within urban
743 settings, where we simulate anomalously high roadway-adjacent NO_2 and $PM_{2.5}$ concentrations,
744 and anomalously low O_3 concentrations. Over our full simulation domain, we find that the
745 highest concentrations of O_3 are found over Lake Michigan during warm season months, where
746 concentrations are simulated to be a full 30% higher than the domain average. In the largest
747 urban area simulated in our domain, Chicago, IL, we find that concentrations of NO_2 are five
748 times higher than the domain average, $PM_{2.5}$ three times higher, and O_3 slightly lower. We also
749 note spatiotemporal O_3 regime variability within the full model domain, where simulated column
750 average HCHO:NO_2 ratios differ substantially by season and location. Over Chicago, conditions
751 are simulated to be VOC-limited, except during the summer, during which time conditions are
752 considered transitional. Likewise in Chicago, our higher resolution simulations show higher
753 average concentrations of NO_2 and $PM_{2.5}$ than our coarser model simulations, suggesting that
754 coarser models may underestimate exposure to these pollutants and their associated health
755 impacts. Lastly, within Chicago city limits, we find that pollutants can vary by a factor of ~2
756 between neighborhoods, a finding potentially corroborated by observed inequitable health
757 outcomes.

758 **Acknowledgments**

- We declare there are no financial or conflicts of interest for the authors.
- We acknowledge support from the U.S. National Science Foundation Grant CBET-1848683 to DEH; McCormick Center for Engineering Sustainability and Resilience seed grant to DEH; The Ubben Program for Carbon and Climate Science postdoctoral fellowship to JLS; and The Data Science fellowship from the National Science Foundation Research Traineeship and Northwestern Integrated Data-Driven Discovery in Earth and Astrophysical Sciences to AM.

Open Research

In this study, we use the WRF-CMAQ model (v5.2; Byun & Schere, 2006), which is available from US EPA Office of Research and Development via https://doi.org/10.5281/zenodo.1079898. The WRFv3.4 model can be downloaded from: https://www2.mmm.ucar.edu/wrf/src/WRFV3.4.TAR.gz. We use the SMOKE 2016 Beta modeling platform (B.H. Baek & Seppanen, 2018) to process the input emissions for WRF-CMAQ which is available via ftp://newftp.epa.gov/air/emismod/2016/beta/. We use the 4-km spatial surrogates from CMAS (Eyth et al., 2019) which are available with the 2016 Beta Platform and 1.3 km spatial surrogates from LADCO (LADCO, 2022) which are available by request. Due to the large output file sizes and storage, model output can only be provided by request. For model performance analysis, we use measurements from the National Climate Data Center (NCDC; https://www.ncei.noaa.gov/data/local-climatological-data/) and the EPA Air Quality System Data Mart (EPA AQS; https://aqs.epa.gov/aqsweb/airdata/download_files.html). The data processing and visualization was performed in Python (v3) using the following packages which are available through the Anaconda repository: NetCDF4, Numpy, Pandas, Geopandas, Cartopy, Shapely, and Matplotlib. The shapefiles for Chicago are available from the Chicago Open Data Portal via https://data.cityofchicago.org/. The analysis and visualization Python scripts are available on Github the BSD-3-Clause License, available here: https://github.com/NU-CCRG/Montgomery-et-al-2022_WRFCMAQ-LakeMichChicago-Baseline .

References

Abdi-Oskouei, M., Carmichael, G., Christiansen, M., Ferrada, G., Roozitalab, B., Sobhani, N., et al. (2020). Sensitivity of Meteorological Skill to Selection of WRF-Chem Physical Parameterizations and Impact on Ozone Prediction During the Lake Michigan Ozone Study (LMOS). *Journal of Geophysical Research: Atmospheres*, *125*(5). https://doi.org/10.1029/2019JD031971

Alexeeff, S. E., Roy, A., Shan, J., Liu, X., Messier, K., Apte, J. S., et al. (2018). High-resolution mapping of traffic related air pollution with Google street view cars and incidence of cardiovascular events within neighborhoods in Oakland, CA. *Environmental Health*, *17*(1), 38. https://doi.org/10.1186/s12940-018-0382-1

B.H. Baek, & Seppanen, C. (2018, July 25). Bokhaeng/Smoke: Smoke V4.5 Public Release (April 2017). Zenodo. https://doi.org/10.5281/ZENODO.1321280

Bickford, E., Holloway, T., Karambelas, A., Johnston, M., Adams, T., Janssen, M., & Moberg, C. (2014). Emissions and Air Quality Impacts of Truck-to-Rail Freight Modal Shifts in the Midwestern United States. *Environmental Science & Technology*, *48*(1), 446–454. https://doi.org/10.1021/es4016102

Blanchard, C. L., Tanenbaum, S., & Lawson, D. R. (2008). Differences between Weekday and Weekend Air Pollutant Levels in Atlanta; Baltimore; Chicago; Dallas–Fort Worth; Denver; Houston; New York; Phoenix; Washington, DC; and Surrounding Areas. *Journal of the Air & Waste Management Association*, *58*(12), 1598–1615. https://doi.org/10.3155/1047-3289.58.12.1598

Byun, D., & Schere, K. L. (2006). Review of the Governing Equations, Computational Algorithms, and Other Components of the Models-3 Community Multiscale Air Quality

(CMAQ) Modeling System. *Applied Mechanics Reviews*, *59*(2), 51. https://doi.org/10.1115/1.2128636

Campbell, P. C., Bash, J. O., & Spero, T. L. (2019). Updates to the Noah Land Surface Model in WRF-CMAQ to Improve Simulated Meteorology, Air Quality, and Deposition. *Journal of Advances in Modeling Earth Systems*, *11*(1), 231–256. https://doi.org/10.1029/2018MS001422

Clark, L. P., Harris, M. H., Apte, J. S., & Marshall, J. D. (2022). National and Intraurban Air Pollution Exposure Disparity Estimates in the United States: Impact of Data-Aggregation Spatial Scale. *Environmental Science & Technology Letters*, acs.estlett.2c00403. https://doi.org/10.1021/acs.estlett.2c00403

Clough, S. A., Shephard, M. W., Mlawer, E. J., Delamere, J. S., Iacono, M. J., Cady-Pereira, K., et al. (2005). Atmospheric radiative transfer modeling: a summary of the AER codes. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *91*(2), 233–244. https://doi.org/10.1016/j.jqsrt.2004.05.058

Dennis, R., Fox, T., Fuentes, M., Gilliland, A., Hanna, S., Hogrefe, C., et al. (2010). A framework for evaluating regional-scale numerical photochemical modeling systems. *Environmental Fluid Mechanics*, *10*(4), 471–489. https://doi.org/10.1007/s10652-009-9163-2

Dewitz, J. (2021). National Land Cover Database (NLCD) 2019 Products [Data set]. U.S. Geological Survey. https://doi.org/10.5066/P9KZCM54

Doak, A. G., Christiansen, M. B., Alwe, H. D., Bertram, T. H., Carmichael, G., Cleary, P., et al. (2021). Characterization of ground-based atmospheric pollution and meteorology

sampling stations during the Lake Michigan Ozone Study 2017. *Journal of the Air & Waste Management Association*, 1–24. https://doi.org/10.1080/10962247.2021.1900000

Dye, T. S., Roberts, P. T., & Korc, M. E. (1995). Observations of Transport Processes for Ozone and Ozone Precursors during the 1991 Lake Michigan Ozone Study. *Journal of Applied Meteorology*, *34*(8), 1877–1889. https://doi.org/10.1175/1520-0450(1995)034<1877:OOTPFO>2.0.CO;2

Emmons, L. K., Schwantes, R. H., Orlando, J. J., Tyndall, G., Kinnison, D., Lamarque, J., et al. (2020). The Chemistry Mechanism in the Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(4). https://doi.org/10.1029/2019MS001882

Eyth, A., Vukovich, J., Farkas, C., & Strum, M. (2019, September). Technical Support Document (TSD): Preparation of Emissions Inventories for the Version 7.2 2016 North American Emissions Modeling Platform. U.S. Environmental Protection Agency. Retrieved from https://www.epa.gov/sites/default/files/2019-09/documents/2016v7.2_regionalhaze_emismod_tsd_508.pdf

Foley, T., Betterton, E. A., Robert Jacko, P. E., & Hillery, J. (2011). Lake Michigan air quality: The 1994-2003 LADCO Aircraft Project (LAP). *Atmospheric EnviroNMEnt*. https://doi.org/10.1016/j.atmosenv.2011.02.033

Fountoukis, C., Koraj, Dh., Denier van der Gon, H. A. C., Charalampidis, P. E., Pilinis, C., & Pandis, S. N. (2013). Impact of grid resolution on the predicted fine PM by a regional 3-D chemical transport model. *Atmospheric EnviroNMEnt*, *68*, 24–32. https://doi.org/10.1016/j.atmosenv.2012.11.008

851 Gan, C.-M., Hogrefe, C., Mathur, R., Pleim, J., Xing, J., Wong, D., et al. (2016). Assessment of
852 the effects of horizontal grid resolution on long-term air quality trends using coupled
853 WRF-CMAQ simulations. *Atmospheric EnviroNMEnt*, *132*, 207–216.
854 https://doi.org/10.1016/j.atmosenv.2016.02.036
855 Goldberg, D. L., Loughner, C. P., Tzortziou, M., Stehr, J. W., Pickering, K. E., Marufu, L. T., &
856 Dickerson, R. R. (2014). Higher surface ozone concentrations over the Chesapeake Bay
857 than over the adjacent land: Observations and models from the DISCOVER-AQ and
858 CBODAQ campaigns. *Atmospheric EnviroNMEnt*, *84*, 9–19.
859 https://doi.org/10.1016/j.atmosenv.2013.11.008
860 Goldberg, D. L., Anenberg, S. C., Kerr, G. H., Mohegh, A., Lu, Z., & Streets, D. G. (2021).
861 TROPOMI NO 2 in the United States: A detailed look at the annual averages, weekly
862 cycles, effects of temperature, and correlation with surface NO 2 concentrations . *Earth's*
863 *Future*, (2), 1–16. https://doi.org/10.1029/2020ef001665
864 Goodkind, A. L., Tessum, C. W., Coggins, J. S., Hill, J. D., & Marshall, J. D. (2019). Fine-scale
865 damage estimates of particulate matter air pollution reveal opportunities for location-
866 specific mitigation of emissions. *Proceedings of the National Academy of Sciences*,
867 *116*(18), 8775 LP – 8780. https://doi.org/10.1073/pnas.1816102116
868 Gupta, R. S., Zhang, X., Sharp, L. K., Shannon, J. J., & Weiss, K. B. (2008). Geographic
869 variability in childhood asthma prevalence in Chicago. *Journal of Allergy and Clinical*
870 *Immunology*, *121*(3), 639-645.e1. https://doi.org/10.1016/j.jaci.2007.11.036
871 Harkey, M., Holloway, T., Oberman, J., & Scotty, E. (2015). An evaluation of CMAQ NO_2
872 using observed chemistry-meteorology correlations. *Journal of Geophysical Research:*
873 *Atmospheres*, *120*(22), 11,775-11,797. https://doi.org/10.1002/2015JD023316

Hogrefe, C., Pouliot, G., Wong, D., Torian, A., Roselle, S., Pleim, J., & Mathur, R. (2015). Annual application and evaluation of the online coupled WRF–CMAQ system over North America under AQMEII phase 2. *Atmospheric EnviroNMEnt*, *115*, 683–694. https://doi.org/10.1016/j.atmosenv.2014.12.034

Hu, J., Ostro, B., Zhang, H., Ying, Q., & Kleeman, M. J. (2019). Using Chemical Transport Model Predictions To Improve Exposure Assessment of PM $_{2.5}$ Constituents. *Environmental Science & Technology Letters*, *6*(8), 456–461. https://doi.org/10.1021/acs.estlett.9b00396

Huang, R., Zhai, X., Ivey, C. E., Friberg, M. D., Hu, X., Liu, Y., et al. (2018). Air pollutant exposure field modeling using air quality model-data fusion methods and comparison with satellite AOD-derived fields: application over North Carolina, USA. *Air Quality, Atmosphere & Health*, *11*(1), 11–22. https://doi.org/10.1007/s11869-017-0511-y

Hughes, D. D., Christiansen, M. B., Milani, A., Vermeuel, M. P., Novak, G. A., Alwe, H. D., et al. (2021). PM2.5 chemistry, organosulfates, and secondary organic aerosol during the 2017 Lake Michigan Ozone Study. *Atmospheric EnviroNMEnt*, *244*, 117939. https://doi.org/10.1016/j.atmosenv.2020.117939

Jbaily, A., Zhou, X., Liu, J., Lee, T.-H., Kamareddine, L., Verguet, S., & Dominici, F. (2022). Air pollution exposure disparities across US population and income groups. *Nature*, *601*(7892), 228–233. https://doi.org/10.1038/s41586-021-04190-y

Jin, X., Fiore, A. M., Murray, L. T., Valin, L. C., Lamsal, L. N., Duncan, B., et al. (2017). Evaluating a Space-Based Indicator of Surface Ozone-NOx-VOC Sensitivity Over Midlatitude Source Regions and Application to Decadal Trends. *Journal of Geophysical Research: Atmospheres*, *122*(19), 10439–10461. https://doi.org/10.1002/2017JD026720

Jin, X., Fiore, A., Boersma, K. F., Smedt, I. D., & Valin, L. (2020). Inferring Changes in Summertime Surface Ozone–NO $_x$ –VOC Chemistry over U.S. Urban Areas from Two Decades of Satellite and Ground-Based Observations. *Environmental Science & Technology*, *54*(11), 6518–6529. https://doi.org/10.1021/acs.est.9b07785

Jing, P., & Goldberg, D. L. (2022). Influence of conducive weather on ozone in the presence of reduced NOx emissions: A case study in Chicago during the 2020 lockdowns. *Atmospheric Pollution Research*, *13*(2), 101313. https://doi.org/10.1016/j.apr.2021.101313

Jing, P., Lu, Z., Xing, J., Streets, D. G., Tan, Q., O'Brien, T., & Kamberos, J. (2014). Response of the summertime ground-level ozone trend in the Chicago area to emission controls and temperature changes, 2005–2013. *Atmospheric EnviroNMEnt*, *99*, 630–640. https://doi.org/10.1016/J.ATMOSENV.2014.10.035

Jing, P., Lu, Z., & Steiner, A. L. (2017). The ozone-climate penalty in the Midwestern U.S. *Atmospheric EnviroNMEnt*, *170*, 130–142. https://doi.org/10.1016/j.atmosenv.2017.09.038

Kain, J. S. (2004). The Kain–Fritsch Convective Parameterization: An Update. *Journal of Applied Meteorology*, *43*(1), 170–181. https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2

Koplitz, S., Simon, H., Henderson, B., Liljegren, J., Tonnesen, G., Whitehill, A., & Wells, B. (2022). Changes in Ozone Chemical Sensitivity in the United States from 2007 to 2016. *ACS Environmental Au*, *2*(3), 206–222. https://doi.org/10.1021/acsenvironau.1c00029

LADCO. (2022, July 27). Lake Michigan Air Directors Consortium (LADCO) Technical Support Document. Retrieved from https://www.ladco.org/technical/ladco-internal/ladco-projects/ladco-2015-o3-naaqs-moderate-area-sip-technical-support-document/

Lamsal, L. N., Duncan, B. N., Yoshida, Y., Krotkov, N. A., Pickering, K. E., Streets, D. G., & Lu, Z. (2015). U.S. NO<inf>2</inf> trends (2005-2013): EPA Air Quality System (AQS) data versus improved observations from the Ozone Monitoring Instrument (OMI). *Atmospheric EnviroNMEnt*, *110*, 130–143. https://doi.org/10.1016/j.atmosenv.2015.03.055

Laughner, J. L., & Cohen, R. C. (2019). Direct observation of changing NO_x lifetime in North American cities. *Science*, *366*(6466), 723–727. https://doi.org/10.1126/science.aax6832

Lin, J.-T., Wuebbles, D. J., Huang, H.-C., Tao, Z., Caughey, M., Liang, X.-Z., et al. (2010). Potential effects of climate and emissions changes on surface ozone in the Chicago area. *Journal of Great Lakes Research*, *36*, 59–64. https://doi.org/10.1016/J.JGLR.2009.09.004

Liu, X.-H., Zhang, Y., Olsen, K. M., Wang, W.-X., Do, B. A., & Bridgers, G. M. (2010). Responses of future air quality to emission controls over North Carolina, Part I: Model evaluation for current-year simulations. *Atmospheric EnviroNMEnt*, *44*(20), 2443–2456. https://doi.org/10.1016/j.atmosenv.2010.04.002

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., et al. (2006). North American Regional Reanalysis. *Bulletin of the American Meteorological Society*, *87*(3), 343–360. https://doi.org/10.1175/BAMS-87-3-343

Morrison, H., Thompson, G., & Tatarskii, V. (2009). Impact of Cloud Microphysics on the Development of Trailing Stratiform Precipitation in a Simulated Squall Line: Comparison

of One- and Two-Moment Schemes. *Monthly Weather Review*, *137*(3), 991–1007. https://doi.org/10.1175/2008MWR2556.1

Odman, M. T., White, A. T., Doty, K., McNider, R. T., Pour-Biazar, A., Qin, M., et al. (2019). Examination of Nudging Schemes in the Simulation of Meteorology for Use in Air Quality Experiments: Application in the Great Lakes Region. *Journal of Applied Meteorology and Climatology*, *58*(11), 2421–2436. https://doi.org/10.1175/JAMC-D-18-0206.1

O'Leary, B. F., & Lemke, L. D. (2014). Modeling spatiotemporal variability of intra-urban air pollutants in Detroit: A pragmatic approach. *Atmospheric EnviroNMEnt*, *94*, 417–427. https://doi.org/10.1016/j.atmosenv.2014.05.010

Pan, S., Choi, Y., Roy, A., & Jeon, W. (2017). Allocating emissions to 4 km and 1 km horizontal spatial resolutions and its impact on simulated NOx and O3 in Houston, TX. *Atmospheric EnviroNMEnt*, *164*, 398–415. https://doi.org/10.1016/j.atmosenv.2017.06.026

Pleim, J. E. (2007). A Combined Local and Nonlocal Closure Model for the Atmospheric Boundary Layer. Part I: Model Description and Testing. *Journal of Applied Meteorology and Climatology*, *46*(9), 1383–1395. https://doi.org/10.1175/JAM2539.1

Pleim, J. E., & Gilliam, R. (2009). An Indirect Data Assimilation Scheme for Deep Soil Temperature in the Pleim–Xiu Land Surface Model. *Journal of Applied Meteorology and Climatology*, *48*(7), 1362–1376. https://doi.org/10.1175/2009JAMC2053.1

Pleim, J. E., & Xiu, A. (2003). Development of a Land Surface Model. Part II: Data Assimilation. *Journal of Applied Meteorology*, *42*(12), 1811–1822. https://doi.org/10.1175/1520-0450(2003)042<1811:DOALSM>2.0.CO;2

962 Qin, M., Yu, H., Hu, Y., Russell, A. G., Odman, M. T., Doty, K., et al. (2019). Improving ozone
963 simulations in the Great Lakes Region: The role of emissions, chemistry, and dry
964 deposition. *Atmospheric EnviroNMEnt*, *202*, 167–179.
965 https://doi.org/10.1016/j.atmosenv.2019.01.025
966 Schnell, J. L., & Prather, M. J. (2017). Co-occurrence of extremes in surface ozone, particulate
967 matter, and temperature over eastern North America. *Proceedings of the National*
968 *Academy of Sciences*, *114*(11), 2854–2859. https://doi.org/10.1073/pnas.1614453114
969 Sharma, S., Sharma, P., & Khare, M. (2017). Photo-chemical transport modelling of tropospheric
970 ozone: A review. *Atmospheric EnviroNMEnt*, *159*, 34–54.
971 https://doi.org/10.1016/j.atmosenv.2017.03.047
972 Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Wang, W., et al. (2008). *A*
973 *Description of the Advanced Research WRF Version 3* [Application/pdf] (p. 1002 KB).
974 UCAR/NCAR. https://doi.org/10.5065/D68S4MVH
975 Southerland, V. A., Anenberg, S. C., Harris, M., Apte, J., Hystad, P., van Donkelaar, A., et al.
976 (2021). Assessing the Distribution of Air Pollution Health Risks within Cities: A
977 Neighborhood-Scale Analysis Leveraging High-Resolution Data Sets in the Bay Area,
978 California. *Environmental Health Perspectives*, *129*(3), EHP7679, 037006.
979 https://doi.org/10.1289/EHP7679
980 Tai, A. P. K., Mickley, L. J., & Jacob, D. J. (2010). Correlations between fine particulate matter
981 (PM2.5) and meteorological variables in the United States: Implications for the
982 sensitivity of PM2.5 to climate change. *Atmospheric EnviroNMEnt*, *44*(32), 3976–3984.
983 https://doi.org/10.1016/j.atmosenv.2010.06.060

Tessum, C. W., Paolella, D. A., Chambliss, S. E., Apte, J. S., Hill, J. D., & Marshall, J. D. (2021). PM $_{2.5}$ polluters disproportionately and systemically affect people of color in the United States. *Science Advances*, *7*(18), eabf4491. https://doi.org/10.1126/sciadv.abf4491

The CESM2 Development Team. (2019). CESM2.1/CAM-chem Instantaneous Output for Boundary Conditions [Data set]. UCAR/NCAR - Atmospheric Chemistry Observations and Modeling Laboratory. https://doi.org/10.5065/NMP7-EP60

Torres-Vazquez, A., Pleim, J., Gilliam, R., & Pouliot, G. (2022). Performance Evaluation of the Meteorology and Air Quality Conditions From Multiscale WRF-CMAQ Simulations for the Long Island Sound Tropospheric Ozone Study (LISTOS). *Journal of Geophysical Research: Atmospheres*, *127*(5). https://doi.org/10.1029/2021JD035890

Tran, T., Tran, H., Mansfield, M., Lyman, S., & Crosman, E. (2018). Four dimensional data assimilation (FDDA) impacts on WRF performance in simulating inversion layer structure and distributions of CMAQ-simulated winter ozone concentrations in Uintah Basin. *Atmospheric EnviroNMEnt*, *177*, 75–92. https://doi.org/10.1016/j.atmosenv.2018.01.012

Travis, K. R., Jacob, D. J., Fisher, J. A., Kim, P. S., Marais, E. A., Zhu, L., et al. (2016). Why do models overestimate surface ozone in the Southeast United States? *Atmospheric Chemistry and Physics*, *16*(21), 13561–13577. https://doi.org/10.5194/acp-16-13561-2016

Vijayaraghavan, K., Zhang, Y., Seigneur, C., Karamchandani, P., & Snell, H. E. (2009). Export of reactive nitrogen from coal-fired power plants in the U.S.: Estimates from a plume-in-grid modeling study. *Journal of Geophysical Research*, *114*(D4), D04308. https://doi.org/10.1029/2008JD010432

Wang, F., Li, Q., & Wang, Y. (2021). Lake-atmosphere exchange impacts ozone simulation around a large shallow lake with large cities. *Atmospheric EnviroNMEnt*, *246*, 118086. https://doi.org/10.1016/j.atmosenv.2020.118086

Wang, J., Xue, P., Pringle, W., Yang, Z., & Qian, Y. (2022). Impacts of Lake Surface Temperature on the Summer Climate Over the Great Lakes Region. *Journal of Geophysical Research: Atmospheres*, *127*(11). https://doi.org/10.1029/2021JD036231

Wong, D. C., Pleim, J., Mathur, R., Binkowski, F., Otte, T., Gilliam, R., et al. (2012). WRF-CMAQ two-way coupled system with aerosol feedback: software development and preliminary results. *Geoscientific Model Development*, *5*(2), 299–312. https://doi.org/10.5194/gmd-5-299-2012

Xiu, A., & Pleim, J. E. (2001). Development of a Land Surface Model. Part I: Application in a Mesoscale Meteorological Model. *Journal of Applied Meteorology*, *40*(2), 192–209. https://doi.org/10.1175/1520-0450(2001)040<0192:DOALSM>2.0.CO;2

Zhang, H., Chen, G., Hu, J., Chen, S.-H., Wiedinmyer, C., Kleeman, M., & Ying, Q. (2014a). Evaluation of a seven-year air quality simulation using the Weather Research and Forecasting (WRF)/Community Multiscale Air Quality (CMAQ) models in the eastern United States. *Science of The Total EnviroNMEnt*, *473–474*, 275–285. https://doi.org/10.1016/j.scitotenv.2013.11.121

Zhang, H., Chen, G., Hu, J., Chen, S. H., Wiedinmyer, C., Kleeman, M., & Ying, Q. (2014b). Evaluation of a seven-year air quality simulation using the Weather Research and Forecasting (WRF)/Community Multiscale Air Quality (CMAQ) models in the eastern United States. *Science of the Total EnviroNMEnt*, *473–474*, 275–285. https://doi.org/10.1016/j.scitotenv.2013.11.121