

High-Resolution Air Quality Modeling over Santiago, Chile: A Comparative Assessment of EDGAR and INEMA Anthropogenic Emission Inventories

Mohsen Pourkiaei

spourkiaei@ethz.ch

ETH Zurich

Jing Wang

ETH Zurich

Article

Keywords: High-Resolution Air Quality Modeling, INEMA, EDGAR, WRF, SMOKE, CMAQ, ISAM, HDDM

Posted Date: October 16th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10259950/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

High-Resolution Air Quality Modeling over Santiago, Chile: A Comparative Assessment of EDGAR and INEMA Anthropogenic Emission Inventories

Mohsen Pourkiaei^{1*}, Jing Wang^{1,2}

¹ Dept. of Civil, Environmental and Geomatic Engineering (D-BAUG), Institute of Environmental Engineering (IfU), Air Quality and Particle Technology, ETH Zürich, Zürich, Switzerland

² Empa, Laboratory for Building Energy Materials and Components, Überlandstrasse 129, 8600 Dübendorf, Switzerland

*Corresponding author, email: spourkiaei@ethz.ch

Abstract

Emission inventories are a major source of uncertainty in urban air quality modeling because differences in emission magnitude, spatiotemporal allocation, source classification, and chemical speciation propagate into simulated concentrations, source apportionment, and sensitivities. This study compares the global EDGAR.v8.1(2022) and regional INEMA.v2(2020) anthropogenic inventories within the CMAQ (Community Multiscale Air Quality) photochemical modeling framework over Santiago, Chile, using three one-way nested domains at 9, 3, and 1 km resolution. Inventory comparison showed substantial pollutant- and sector-dependent differences: EDGAR reported higher domain-aggregated emissions for most species, whereas aggregated fine particle (PM_{2.5}) emissions were similar despite markedly different sectoral distributions. EDGAR generally reproduced daytime ozone (O₃) variability more closely, while INEMA agreed better for PM; neither inventory consistently outperformed the other across all pollutants. Emission scaling constrained by mean observed concentrations established calibrated atmospheric states for source-apportionment and sensitivity analyses. CMAQ source apportionment differed between calibrated EDGAR- and INEMA-based configurations, with greater industrial-combustion influence under EDGAR and transport influence under INEMA, while energy use in buildings remained important in both. Decoupled Direct Method (DDM) sensitivity simulations revealed distinct sensitivity structures in the calibrated configurations. Under EDGAR, first-order sensitivities to NO_x were distributed mainly across industrial combustion, transport, and power industry, whereas under INEMA, NO, O₃, and PM_{2.5} showed their largest first-order sensitivities to transport NO_x emissions. Both configurations showed negative O₃ sensitivity to NO_x and positive sensitivity to volatile organic compounds (VOC) emissions, indicating a VOC-limited photochemical response during the simulated period. Second-order PM_{2.5} sensitivities to NO_x also showed substantial sector-specific nonlinearity.

Keywords: High-Resolution Air Quality Modeling, INEMA, EDGAR, WRF, SMOKE, CMAQ, ISAM, HDDM.

Introduction

Air pollution remains an important environmental and public-health challenge in Santiago, Chile, where dense anthropogenic emissions interact with complex basin topography and seasonally restricted atmospheric ventilation. The Santiago Metropolitan Region lies between the Andes Mountains and the Coastal Range, and stable conditions and thermal inversions can limit vertical mixing and favor pollutant accumulation¹. Road transport, residential combustion, industrial activity, and energy production contribute to urban emissions of nitrogen oxides (NO_x), particulate matter (PM), volatile organic compounds (VOCs), and other pollutants, while their effects on ambient concentrations vary strongly with meteorology, chemical transformation, and spatiotemporal patterns^{1,2}.

Chemical transport models (CTMs) can help resolve these interacting processes and provide a framework for evaluating concentration fields, source contributions, and responses to changes in precursor emissions. At kilometer-scale resolution, however, model performance is strongly conditioned by the emissions supplied to the model. Differences and uncertainties in emission magnitude, spatial allocation, temporal distribution, sector classification, or chemical speciation can propagate through atmospheric chemistry and transport and thereby affect not only simulated concentrations but also subsequent source-apportionment and sensitivity analyses³⁻⁵.

Global and regional emission inventories address different scientific needs and consequently differ in their construction. The Emissions Database for Global Atmospheric Research (EDGAR) provides globally consistent gridded emissions using harmonized activity data, emission factors, and spatial proxies, facilitating comparisons across countries and regions^{6,7}. Such consistency is valuable for large-scale applications but cannot reproduce every locally specific feature relevant to urban modeling^{5,8}. In contrast, Chile's INEMA (Inventario Nacional de Emisiones Antropogénicas) inventory incorporates nationally derived activity information and source categories at substantially finer native spatial resolution⁹. Regional inventories nevertheless have their own uncertainties and can differ from global inventories in reference year, source classification, spatial allocation, temporal assumptions, and pollutant definitions. Recent global inverse modeling has likewise identified substantial uncertainties in bottom-up ammonia (NH₃) emissions, particularly in less-constrained regions including South America¹⁰. Direct quantitative comparison is therefore necessary before differences in Community Multiscale Air Quality (CMAQ) model performance can be attributed to either inventory.

The CMAQ model provides a three-dimensional framework for simulating atmospheric transport, gas-phase chemistry, aerosol processes, deposition, and cloud interactions^{11,12}. Together with meteorological fields generated by the Weather Research and Forecasting (WRF) model and emissions processed through the Sparse Matrix Operator Kernel Emissions (SMOKE) system, CMAQ can be applied at nested hemispheric-to-urban scales to investigate spatiotemporal variations in pollutant levels and chemical transformation under complex meteorological conditions.

Diagnostic extensions of CMAQ can further determine how emission differences affect modeled source apportionment and chemical response. The Integrated Source Apportionment Method (ISAM) tracks the contribution of specified emission sectors, initial conditions, and boundary conditions to modeled pollutant concentrations¹³⁻¹⁵. The Decoupled Direct Method

in Three Dimensions (DDM-3D) and its Higher-order implementation (HDDM) calculate first- and second-order sensitivities of pollutant concentrations to perturbations in precursor emissions¹⁶⁻¹⁹. These diagnostics are inherently conditional on the atmospheric state within which they are calculated; consequently, their quantitative interpretation is strengthened when the underlying simulation provides reasonable agreement with observations.

Previous modeling studies in Chile have demonstrated the value of high-resolution atmospheric modeling while also identifying important uncertainties associated with emissions and meteorology. For example, WRF-SMOKE-CMAQ simulations for the Quintero-Puchuncaví-Concón industrial region reproduced important features of particulate matter (PM) and ozone (O₃), but the sensitivity to local emission inputs was not examined¹. Independent development of the INEMA inventory has likewise shown substantial differences between globally derived and Chile-specific emission estimates and sector distributions⁹. These findings motivate a controlled comparison in which different anthropogenic inventories are processed within a systematic modeling framework.

In this paper, we examine how differences arising from anthropogenic emission inventory selection propagate through high-resolution air quality modeling over Santiago during April 2025. April was selected as a challenging autumn transition period, combining shortening daylight and evolving meteorological conditions with the change from daylight-saving time to standard time. The choice of 2025 aligns with the broader objectives of the MASSEV project, which include developing predictive high-resolution CTM applications for periods beyond the temporal coverage of available emission inventories. EDGAR v8.1 (2022) and INEMA v2 (2020) were processed using novel inventory-specific preprocessing workflows and then passed to the WRF-SMOKE-CMAQ photochemical modeling framework with nested 9-, 3-, and 1-km domains. The emissions implementation was restricted to processing and comparing the two anthropogenic inventories; a separate biogenic emission component was not included. This scope restriction does not imply that biogenic emissions are negligible in Santiago. Their omission limits the representation of the atmospheric VOC budget and the interpretation of O₃ and secondary organic aerosol formation, source apportionment, and precursor sensitivities. These results are conditional on the adopted model configurations and, for the calibrated diagnostics, on the emission adjustments. We first quantify differences in pollutant totals and harmonized source sectors and evaluate the corresponding CMAQ simulations against observations from Las Condes, Parque O'Higgins, and Pudahuel. Pollutant-specific Detailed Emissions Scaling, Isolation, and Diagnostic (DESID) adjustments are then derived using pollutant-specific observational targets to construct calibrated model-ready emissions. Details of the DESID application are presented in the Methods Section under CMAQ DESID Scaling. These calibrated configurations are subsequently used for ISAM source apportionment and DDM/HDDM sensitivity analysis. The central objectives are to determine (i) how strongly the two inventories differ in emission magnitude and sectoral distribution, (ii) whether those differences produce consistent or pollutant-dependent changes in model performance, and (iii) how calibrated EDGAR- and INEMA-based configurations differ in modeled source contributions and first- and higher-order emission sensitivities. To our knowledge, this is the first CMAQ implementation and comparison of EDGAR v8.1 and INEMA v2 over Santiago

that evaluates inventory differences jointly through model performance, source apportionment, and higher-order concentration sensitivity.

Results

Emission Inventory Comparison

Before evaluating how the choice of emission inventory affects CMAQ concentration predictions, the two anthropogenic inventories were compared over the common Santiago-centered 9-km parent domain. Figures 1 and 2 show the absolute and relative sectoral contributions per pollutant of EDGAR v8.1 for April 2022 (most recent year) and INEMA v2 for April 2020 (most recent year), respectively. Since the two inventories use different native source classifications, INEMA sectors were mapped onto eight integrated EDGAR sectors to support a consistent sector-by-sector comparison: agriculture (AGR), energy for buildings (ENB), fuel exploitation (FUE), industrial combustion (INC), industrial processes (INP), power industry (POI), transport (TRA), and waste (WAS). The detail of this sector mapping is presented in the Methods Section; Emission Inventory Preprocessing.

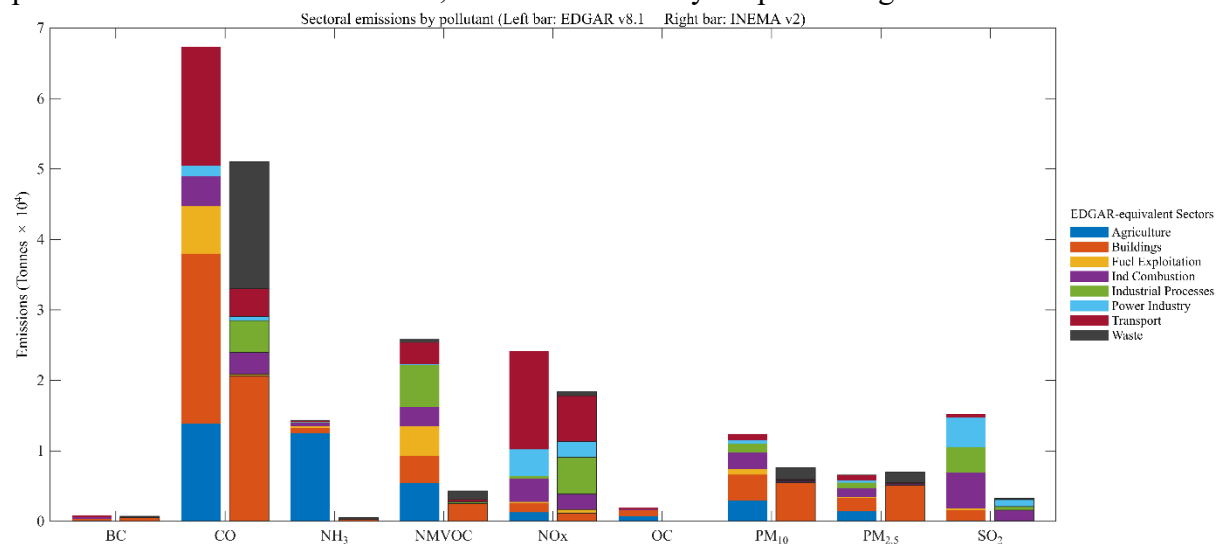


Figure 1. Sectoral composition of absolute emissions (Tonnes × 10⁴) of EDGAR v8.1 (April 2022) and INEMA v2 (April 2020) over the Santiago-centered 9-km parent domain. The eight harmonized comparison sectors are agriculture (blue), energy for buildings (orange), fuel exploitation (yellow), industrial combustion (purple), industrial processes (green), power industry (cyan), transport (maroon), and waste (dark gray). EDGAR and INEMA sector definitions, and their corresponding mapping are provided in the Methods section under Emission Inventory Preprocessing. Abbreviations: EDGAR = Emissions Database for Global Atmospheric Research; INEMA = Inventario Nacional de Emisiones Antropogénicas; BC = black carbon; CO = carbon monoxide; NH₃ = ammonia; NMVOC = non-methane volatile organic compounds; NO_x = nitrogen oxides; OC = organic carbon; PM_{2.5} = particulate matter with aerodynamic diameter ≤ 2.5 μm; PM₁₀ = particulate matter with aerodynamic diameter ≤ 10 μm; SO₂ = sulfur dioxide.

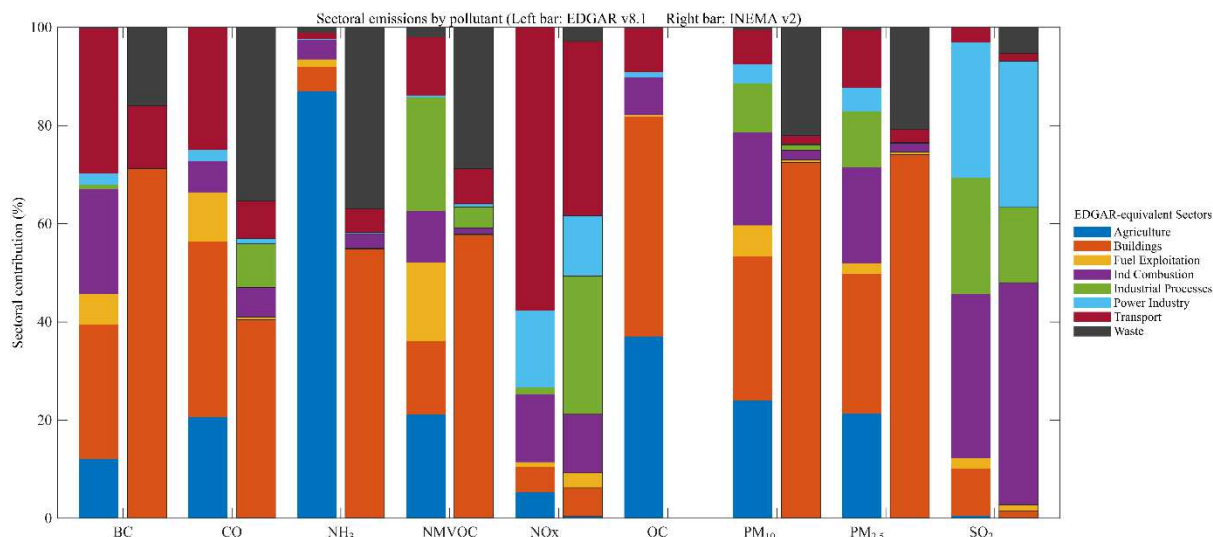


Figure 2. Sectoral composition of relative emission contributions (%) from EDGAR v8.1 (April 2022) and INEMA v2 (April 2020) over the Santiago-centered 9-km parent domain. The eight harmonized comparison sectors are agriculture (blue), energy for buildings (orange), fuel exploitation (yellow), industrial combustion (purple), industrial processes (green), power industry (cyan), transport (maroon), and waste (dark gray). EDGAR and INEMA sector definitions, and their corresponding mapping are provided in the Methods section under Emission Inventory Preprocessing. Abbreviations: EDGAR = Emissions Database for Global Atmospheric Research; INEMA = Inventario Nacional de Emisiones Antropogénicas; BC = black carbon; CO = carbon monoxide; NH₃ = ammonia; NMVOC = non-methane volatile organic compounds; NO_x = nitrogen oxides; OC = organic carbon; PM_{2.5} = particulate matter with aerodynamic diameter ≤ 2.5 μm; PM₁₀ = particulate matter with aerodynamic diameter ≤ 10 μm; SO₂ = sulfur dioxide.

Table 1 and Table 2 summarize domain-aggregated and sector-resolved emissions, respectively, for each pollutant across all source sectors within the Santiago-centered 9-km parent domain, comparing EDGAR v8.1 for April 2022 with INEMA v2 for April 2020.

Table 1. Domain-aggregated anthropogenic emissions for EDGAR v8.1 (April 2022) and INEMA v2 (April 2020) over the Santiago-centered 9-km parent domain

Pollutant	EDGAR v8.1 (t)	INEMA v2 (t)	Relative Difference
CO	71,938.7	51,008.8	EDGAR +41%
NO _x	24,167.0	18,362.7	EDGAR +32%
NMVOC / VOC	25,857.4	4,317.3	EDGAR +499%
SO ₂	15,215.1	3,306.2	EDGAR +360%
PM _{2.5}	6,619.9	6,951.7	INEMA +5%
PM ₁₀	12,418.7	7,626.3	EDGAR +63%
NH ₃	14,404.6	541.6	EDGAR +2,560%
BC	849.5	722.0	EDGAR +18%

* The underlying source definitions are not equivalent despite the harmonized eight-sector structure. Further details on the sector mapping between EDGAR v8.1 and INEMA v2 inventories are provided in the Methods section, Emission Inventory Preprocessing.

** Relative differences are expressed using total emissions (E) as: $100(E_{EDGAR} - E_{INEMA})/E_{INEMA}$.

While these harmonized categories facilitate comparison, they are not identical in their underlying activity definitions.

The inventories differ substantially in total emitted mass for most pollutants. Relative to INEMA, EDGAR contains approximately 41% more carbon monoxide (CO), 32% more NO_x, 63% more coarse particles (PM₁₀), 360% more sulfur dioxide (SO₂), and 499% more NMVOC (non-methane volatile organic compounds) over the parent domain. The largest relative

difference occurs for NH₃, for which the EDGAR total is more than an order of magnitude larger. Fine particles (PM_{2.5}) are the principal exception, with similar total emissions between the inventories; INEMA is approximately 5% higher than EDGAR.

Table 2. Sector-resolved EDGAR v8.1 (April 2022) and INEMA v2 (April 2020) anthropogenic emissions over the Santiago-centered 9-km parent domain (t-month⁻¹).

Pollutant Sector	Inventory	CO	NO _x	NMVOC	SO ₂	PM _{2.5}	PM ₁₀	NH ₃
AGR	EDGAR	13,902.9	1,290.2	5,468.1	73.5	1,415.2	2,979.6	12,536.7
	INEMA	47.5	66.6	1.4	1.8	1.9	2.1	0.2
ENB	EDGAR	24,078.7	1,245.8	3,856.9	1,469.8	1,884.2	3,643.0	702.4
	INEMA	20,604.2	1,072.8	2,487.3	49.3	5,144.9	5,525.1	297.1
FUE	EDGAR	6,728.8	228.3	4,138.7	319.4	135.7	787.9	223.5
	INEMA	253.9	557.2	10.0	39.2	38.3	38.5	0.9
INC	EDGAR	4,616.6	3,332.3	2,739.5	5,080.8	1,296.4	2,354.3	573.2
	INEMA	3,097.1	2,191.3	54.4	1,496.5	126.4	152.6	15.8
INP	EDGAR	4,261.0	337.1	5,972.6	3,611.0	757.0	1,227.3	12.5
	INEMA	4,496.1	5,178.3	185.2	506.7	0.4	85.0	0.03
POI	EDGAR	1,527.4	3,785.6	101.7	4,189.6	320.9	490.7	20.0
	INEMA	542.2	2,244.4	25.9	985.5	12.1	12.2	1.2
TRA	EDGAR	16,820.3	13,939.7	3,079.5	463.3	771.1	874.9	186.4
	INEMA	3,955.8	6,523.2	311.9	53.4	188.2	129.3	26.3
WAS	EDGAR	3.1	8.0	500.3	7.6	39.6	61.0	150.0
	INEMA	18,012.0	528.8	1,241.2	173.6	1,439.5	1,681.5	200.1
Total	EDGAR	71,938.7	24,167.0	25,857.4	15,215.1	6,619.9	12,418.7	14,404.6
	INEMA	51,008.8	18,362.7	4,317.3	3,306.2	6,951.7	7,626.3	541.6

* Note: The eight sectors are harmonized comparison categories. The underlying EDGAR and INEMA activity definitions are not identical; consequently, differences between corresponding sectors may reflect both differences in emission magnitude and differences in inventory classification. The complete mappings are provided in Tables 7 and 11.

The reported VOC/NMVOC-to-NO_x mass ratio is approximately 1.070 for EDGAR and 0.235 for INEMA across all sectors, and 0.221 and 0.048, respectively, for their transport groups. These pre-DESID ratios highlight substantial differences in the represented precursor balance, but do not independently establish the accuracy of either inventory. The transport comparison also involves different source coverage: EDGAR includes road, rail, shipping, and aviation, whereas the INEMA group comprises road transport and transport non-exhaust emissions. Sectoral allocation differs as strongly as the total emission magnitude. For NO_x, EDGAR assigns approximately 13.9 kt-month⁻¹ to transport compared with 6.52 kt-month⁻¹ in INEMA, whereas INEMA assigns substantially more NO_x to industrial processes than EDGAR. PM_{2.5} provides an especially informative contrast. Despite similar domain-aggregated emissions, INEMA concentrates much of its PM_{2.5} in energy use in buildings and the harmonized waste/biomass-burning category, whereas EDGAR distributes PM_{2.5} more broadly among buildings, agriculture, industrial combustion, transport, and other sectors. Supplementary Information (Supplementary Notes) presents the long-term evolution (1975–2020) and current sectoral composition of key emissions in Chile according to the EDGAR v8.1.

CMAQ Base Model Performance

Herein, the base model refers to CMAQ simulations that use the raw emission inventories as provided (EDGAR v8.1, 2022, and INEMA v2, 2020), with no subsequent *quantitative* updates or adjustments.

CMAQ base model performance was evaluated against observations from three SINCA (Sistema de Información Nacional de Calidad del Aire) monitoring stations spanning distinct parts of Santiago: Las Condes in the east, Parque O'Higgins in the central urban area, and Pudahuel in the west (Figure 3). Supplementary Information (Supplementary Results, Figures S2–S4) presents the April 2025 CMAQ base-model results (1-km resolution) at these three locations. SO₂ observations were available only at the La Florida station and were used as a representative reference for model evaluation.

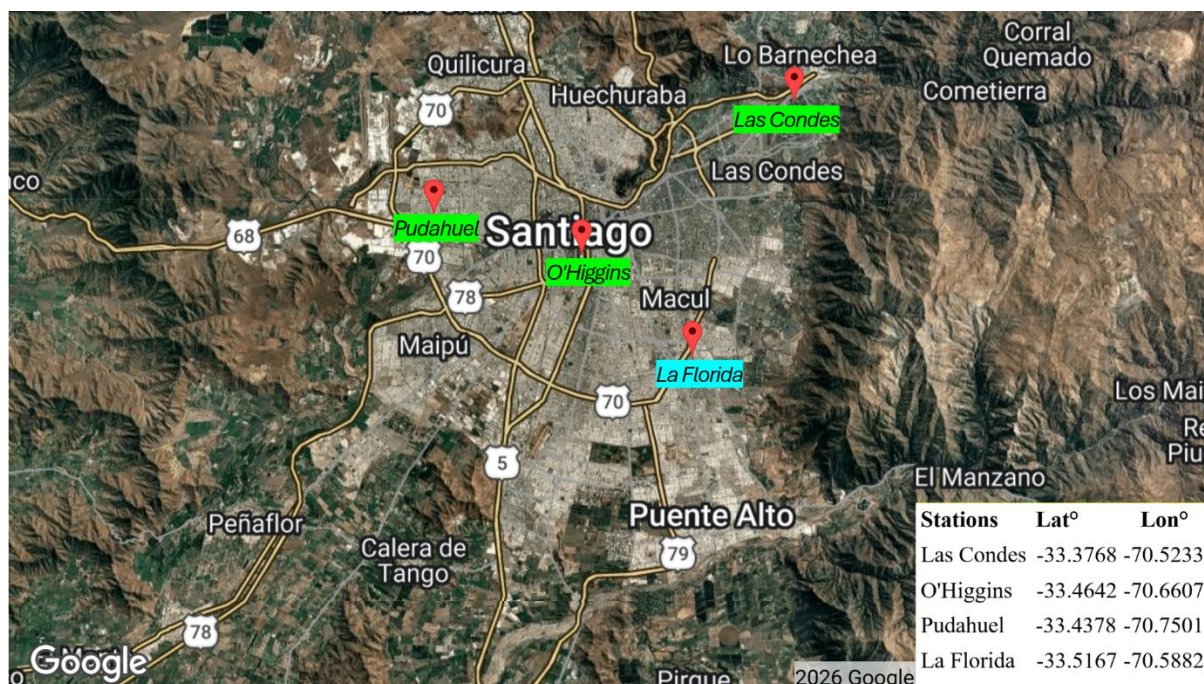


Figure 3. SINCA air quality monitoring stations used for CMAQ model evaluation and calibration in Santiago, Chile. Background map: Google Maps. Map data © Google and applicable data providers; original attribution retained in the figure. The map shows Las Condes in eastern Santiago, Parque O'Higgins in the central urban area, Pudahuel in western Santiago, and La Florida in the southeastern urban area. Las Condes, Parque O'Higgins, and Pudahuel were used for the principal model–observation evaluation and for construction of the three-site observational calibration targets. Representative SO₂ observations used in this study were obtained from La Florida, the only SINCA station recording SO₂ during the study period. Station labels indicate the locations from which observational data were extracted for comparison with the corresponding 1-km CMAQ grid cells. Abbreviations: SINCA = Sistema de Información Nacional de Calidad del Aire; CMAQ = Community Multiscale Air Quality model; SO₂ = sulfur dioxide.

The two inventories produce distinct base-model responses for NO_x and O₃ (Figures S2–S4). The EDGAR simulation generally reproduces the observed nitric oxide (NO) and nitrogen dioxide (NO₂) concentration range more closely than INEMA, although short-duration observed NO peaks are frequently underestimated. Contrarily, INEMA produces substantially larger NO concentrations and recurrent peak overprediction, particularly at Parque O'Higgins and Pudahuel. Both inventories underestimate daytime O₃ maxima, but the EDGAR simulation generally produces higher daytime O₃ and follows the observed diurnal cycle more closely. The higher daytime O₃ under EDGAR is consistent with weaker local O₃ titration (NO + O₃ → NO₂ + O₂) associated with its lower simulated local NO concentrations relative to INEMA.

CO performance varies among the three sites and does not identify a consistently superior inventory. On the other hand, PM shows clearer inventory dependencies. The EDGAR simulation retains positive biases in PM_{2.5} and especially PM₁₀, whereas INEMA generally produces concentrations closer to the observed ranges and captures a greater fraction of their temporal variability. Finally, in the representative SO₂ evaluation, EDGAR produces substantially higher SO₂ concentrations than INEMA, consistent with its considerably larger domain-aggregated SO₂ emissions. Supplementary Information (Supplementary Results) provides complementary model-only diagnostics of size-dependent aerosol mass and particle-number concentrations to characterize differences between the two base simulations beyond the PM_{2.5} and PM₁₀ observational comparisons.

DESID Adjustment (Calibrated CMAQ Performance)

The base-model evaluation revealed substantial discrepancies between simulated and observed concentrations for both emission inventories. Large biases in precursor concentrations can influence the quantitative interpretation of secondary pollutants and their responses to emission perturbations. An empirical emission-calibration step was therefore introduced to obtain model configurations that more closely reproduced the observed concentration levels before subsequent quantitative interpretation.

Scaling factors are empirical coefficients for the present modeling configuration and period. Applying a factor directly changes the magnitude of the corresponding emissions supplied to CMAQ. Because a single factor was applied uniformly across all anthropogenic sectors for each emission quantity, its relative sectoral allocation was retained, but the calibrated emissions no longer preserve the original inventory in absolute magnitude. The factors are therefore interpreted as model-adjustment coefficients rather than independently derived corrections to true emissions. For NO, NO₂, CO, PM_{2.5}, and PM₁₀, calibration targets were defined as the arithmetic mean of the April mean concentrations at Las Condes, Parque O'Higgins, and Pudahuel. SO₂ observations were available only at La Florida, and the measured concentrations were used as representative values for Santiago in this study. NMVOC scaling was constrained indirectly using observed O₃ rather than direct NMVOC measurements. Table 3 summarizes the resulting emission scaling factors, and the Methods describe the corresponding observational targets and modeled quantities used in calibration.

Table 3. Pollutant-specific DESID scaling factors used to construct calibrated emission configurations for April 2025, over Santiago's selected air quality monitoring stations.

	NO	NO ₂	CO	SO ₂	NMVOC	PM _{2.5}	PM ₁₀
EDGAR v8.1	1.548	0.851	1.127	0.026	6.742	0.040	0.061
INEMA v2	0.12	0.55	0.42	0.374	19.353	0.167	0.667

The magnitude of several factors, particularly those applied to VOC emissions, indicates that the adjustment is not an inventory correction in a strict inverse-modeling sense. The calibrated coefficients can compensate simultaneously for uncertainties in emission magnitude, spatial and temporal allocation, chemical speciation, unresolved sub-grid sources, and other model processes. For VOCs in particular, omission of natural biogenic emissions means that the empirical scaling may partly compensate for missing atmospheric VOC reactivity.

Figures 4–6 present the DESID-adjusted 1-km CMAQ simulations alongside the April 2025 observations, and Tables 4–6 summarize the corresponding statistical agreement. Because the observational records used to construct the calibration targets were also used in these comparisons, the results characterize calibration performance and remaining model–observation discrepancies rather than independent out-of-sample performance. Any improvement relative to the corresponding unadjusted base simulations is interpreted as an outcome of empirical fitting. The statistical metrics presented in these tables are N (Number of paired CMAQ–Observation data points), R (Pearson Correlation), $RMSE$ (Root Mean Square Error), MB (Mean Bias), NMB (Normalized Mean Bias), and NME (Normalized Mean Error).

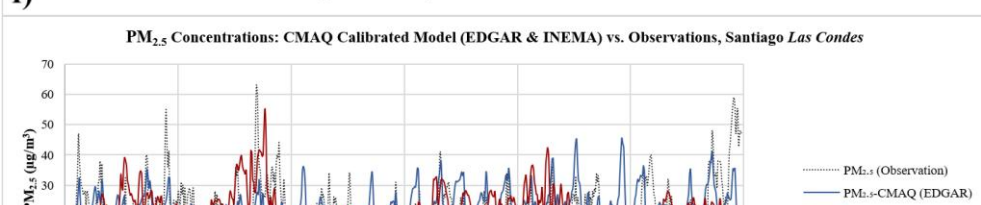
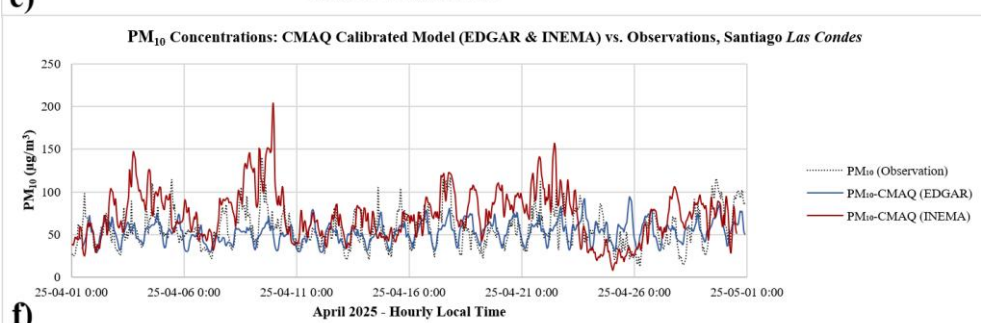
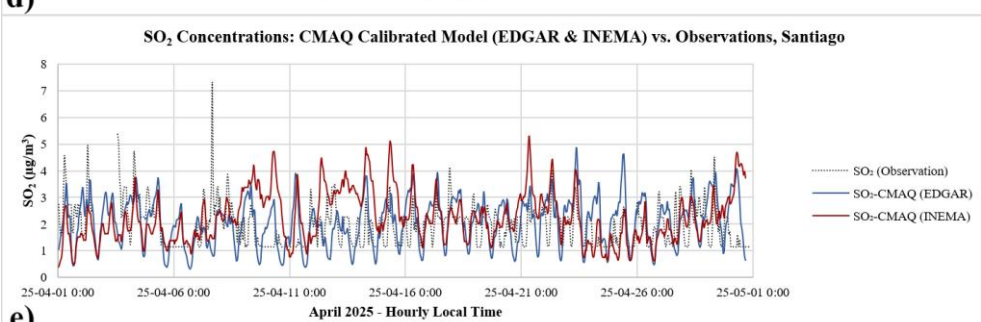
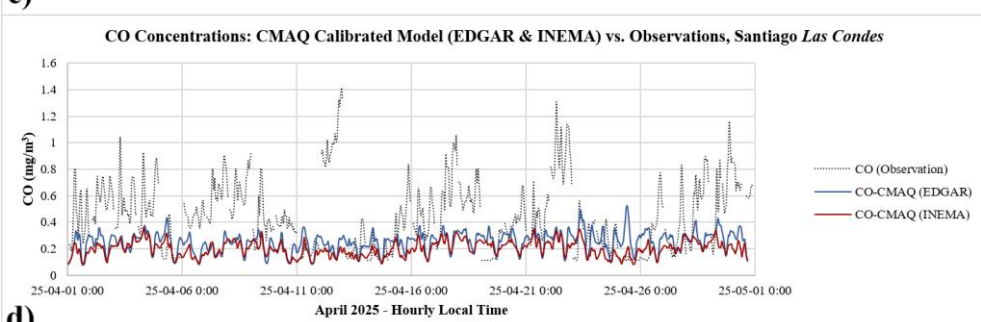
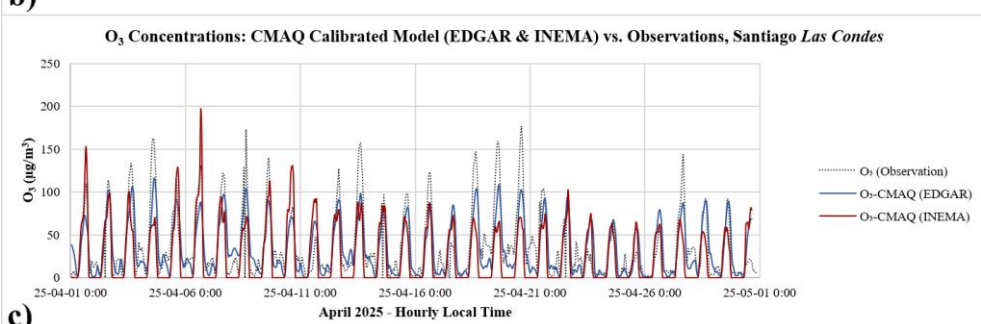
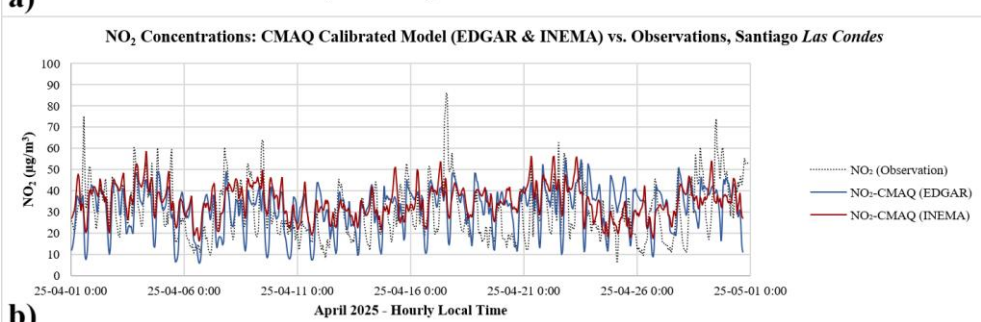
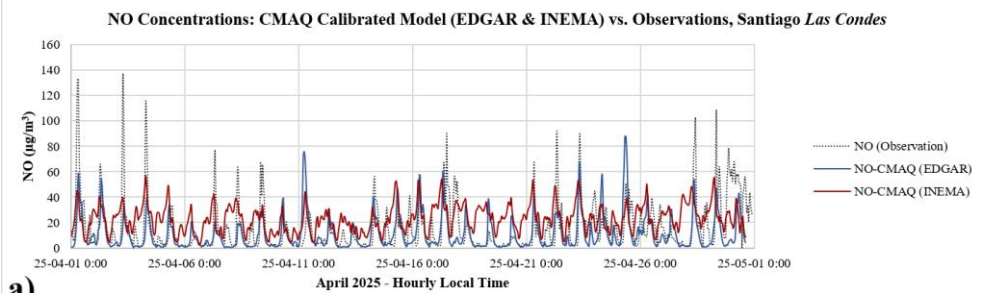


Figure 4. Hourly observed and calibrated CMAQ concentrations at Las Condes during April 2025.

CMAQ simulations driven by DESID-adjusted EDGAR v8.1 (2022) and INEMA v2 (2020) emissions are compared with SINCA observations for (a) NO, (b) NO₂, (c) O₃, (d) CO, (e) PM_{2.5}, (f) PM₁₀, and (g) SO₂. Model values correspond to the 1-km grid cell containing the monitoring station. SO₂ observations are provided from Santiago's representative dataset (La Florida station). Each panel shows hourly concentrations, with solid blue and red lines representing CMAQ outputs driven by EDGAR and INEMA emissions, respectively, and a dashed black line representing observations. All timestamps are reported in Santiago local time. Concentrations are reported in $\mu\text{g}\cdot\text{m}^{-3}$, except CO, which is reported in $\text{mg}\cdot\text{m}^{-3}$. Abbreviations: CMAQ = Community Multiscale Air Quality model; EDGAR = Emissions Database for Global Atmospheric Research; INEMA = Inventario Nacional de Emisiones Antropogénicas; NO = nitric oxide; NO₂ = nitrogen dioxide; O₃ = ozone; CO = carbon monoxide; PM_{2.5} = particulate matter with aerodynamic diameter $\leq 2.5 \mu\text{m}$; PM₁₀ = particulate matter with aerodynamic diameter $\leq 10 \mu\text{m}$; SO₂ = sulfur dioxide.

Table 4. Statistical performance of the calibrated 1-km EDGAR and INEMA CMAQ simulations at Las Condes during April 2025.

Pollutant	Inventory	N	R	RMSE	MB	NMB (%)	NME (%)
O ₃ ($\mu\text{g}\cdot\text{m}^{-3}$)	EDGAR	712	0.824	22.00	-3.62	-9.9	43.5
	INEMA	712	0.748	28.78	-12.63	-34.4	57.3
NO ₂ ($\mu\text{g}\cdot\text{m}^{-3}$)	EDGAR	712	0.321	13.78	+1.31	+4.3	36.4
	INEMA	712	0.343	13.41	+4.58	+15.1	36.1
NO ($\mu\text{g}\cdot\text{m}^{-3}$)	EDGAR	711	0.531	18.00	-4.81	-30.1	64.8
	INEMA	711	0.307	20.88	+6.97	+43.5	102.4
CO ($\text{mg}\cdot\text{m}^{-3}$)	EDGAR	678	0.141	0.331	-0.204	-44.8	56.1
	INEMA	678	0.272	0.358	-0.254	-56.1	60.4
SO ₂ ($\mu\text{g}\cdot\text{m}^{-3}$)	EDGAR	687	0.230	1.05	+0.14	+7.2	42.1
	INEMA	687	0.125	1.18	+0.31	+15.8	45.7
PM _{2.5} ($\mu\text{g}\cdot\text{m}^{-3}$)	EDGAR	713	0.194	11.22	-0.59	-2.9	42.8
	INEMA	706	0.233	10.36	-0.70	-3.5	40.6
PM ₁₀ ($\mu\text{g}\cdot\text{m}^{-3}$)	EDGAR	709	0.463	19.34	-3.66	-6.5	25.8
	INEMA	709	0.366	33.63	+17.23	+30.9	47.6
PM _{2.5} ($\mu\text{g}\cdot\text{m}^{-3}$) daily means (≥ 18 paired hours)	EDGAR	29	0.052	6.63	-0.30	-1.5	24.3
	INEMA	29	0.295	6.66	-0.50	-2.5	27.8

Note: SO₂ observations were available only at La Florida, and the measured concentrations were used as representative values for Santiago in this study.

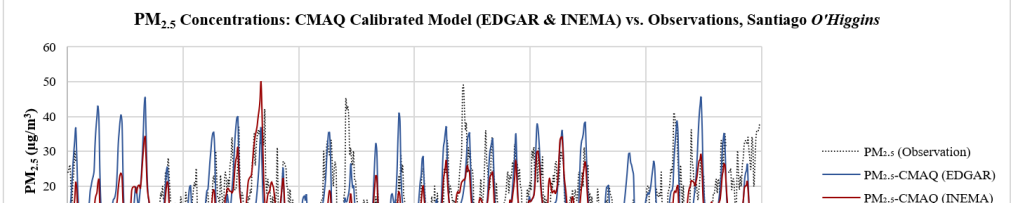
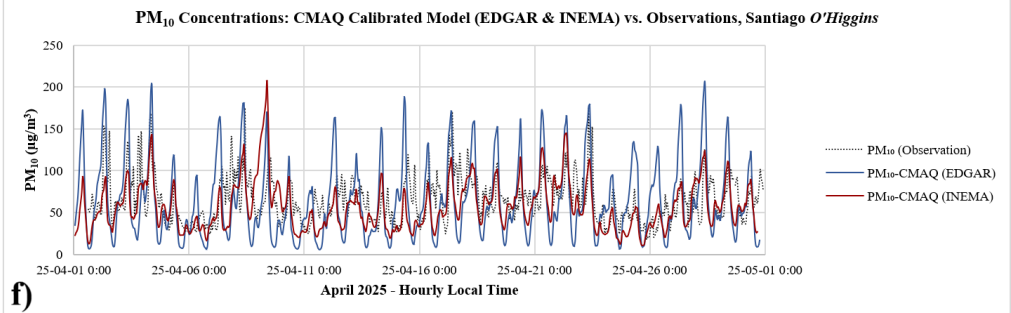
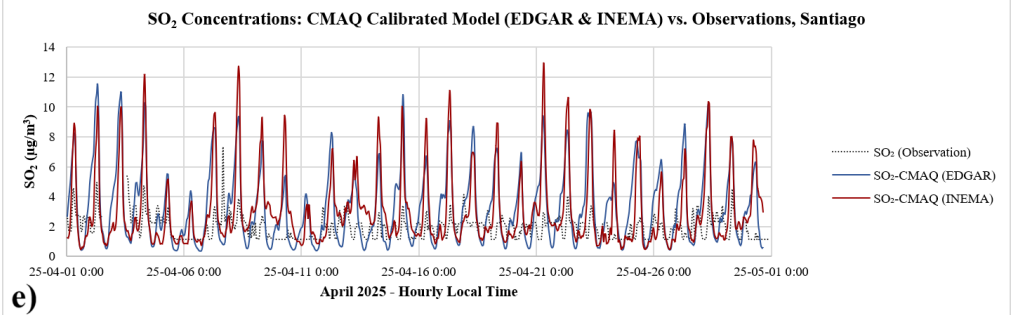
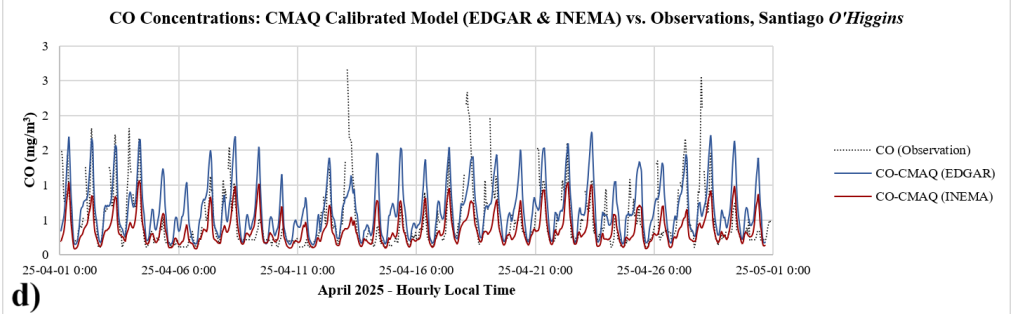
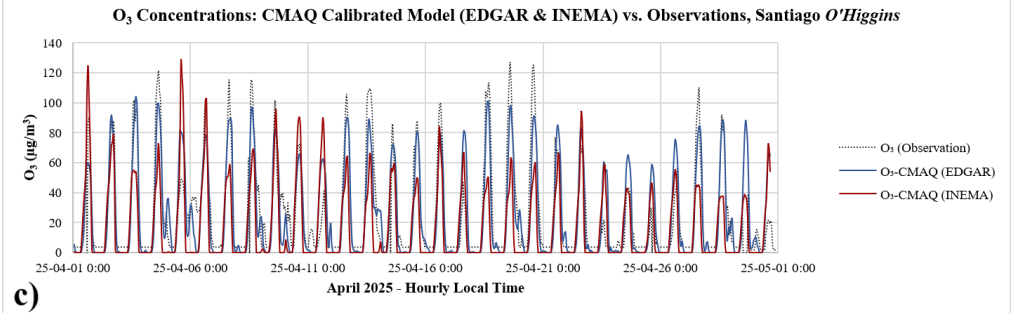
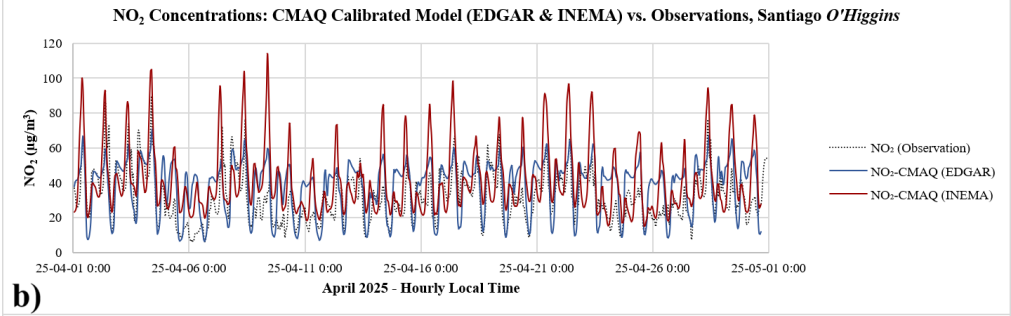
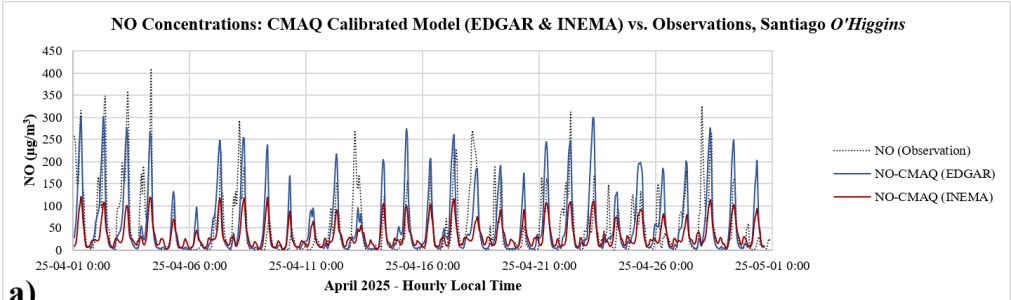


Figure 5. Hourly observed and calibrated CMAQ concentrations at O’Higgins during April 2025. CMAQ simulations driven by DESID-adjusted EDGAR v8.1 (2022) and INEMA v2 (2020) emissions are compared with SINCA observations for (a) NO, (b) NO₂, (c) O₃, (d) CO, (e) PM_{2.5}, (f) PM₁₀, and (g) SO₂. Model values correspond to the 1-km grid cell containing the monitoring station. SO₂ observations are provided from Santiago’s representative dataset (La Florida station). Each panel shows hourly concentrations, with solid blue and red lines representing CMAQ outputs driven by EDGAR and INEMA emissions, respectively, and a dashed black line representing observations. All timestamps are reported in Santiago local time. Concentrations are reported in $\mu\text{g}\cdot\text{m}^{-3}$, except CO, which is reported in $\text{mg}\cdot\text{m}^{-3}$. Abbreviations: CMAQ = Community Multiscale Air Quality model; EDGAR = Emissions Database for Global Atmospheric Research; INEMA = Inventario Nacional de Emisiones Antropogénicas; NO = nitric oxide; NO₂ = nitrogen dioxide; O₃ = ozone; CO = carbon monoxide; PM_{2.5} = particulate matter with aerodynamic diameter $\leq 2.5\ \mu\text{m}$; PM₁₀ = particulate matter with aerodynamic diameter $\leq 10\ \mu\text{m}$; SO₂ = sulfur dioxide.

Table 5. Statistical performance of the calibrated 1-km EDGAR and INEMA CMAQ simulations at Parque O’Higgins during April 2025.

Pollutant	Inventory	N	R	RMSE	MB	NMB (%)	NME (%)
O ₃ ($\mu\text{g}\cdot\text{m}^{-3}$)	EDGAR	712	0.851	16.73	-0.82	-3.2	43.3
	INEMA	712	0.736	22.95	-8.80	-33.9	58.6
NO ₂ ($\mu\text{g}\cdot\text{m}^{-3}$)	EDGAR	712	0.619	14.54	+7.03	+22.4	37.5
	INEMA	712	0.554	18.69	+9.70	+30.9	45.8
NO ($\mu\text{g}\cdot\text{m}^{-3}$)	EDGAR	711	0.641	59.52	+1.04	+2.0	66.4
	INEMA	711	0.560	62.90	-20.70	-38.7	69.7
CO ($\text{mg}\cdot\text{m}^{-3}$)	EDGAR	682	0.589	0.396	+0.165	+33.7	58.4
	INEMA	682	0.555	0.366	-0.128	-26.0	43.4
SO ₂ ($\mu\text{g}\cdot\text{m}^{-3}$)	EDGAR	687	0.284	2.68	+1.29	+66.1	99.8
	INEMA	687	0.290	2.58	+1.12	+57.1	88.8
PM _{2.5} ($\mu\text{g}\cdot\text{m}^{-3}$)	EDGAR	632	0.486	9.74	-3.82	-21.1	43.3
	INEMA	630	0.615	8.34	-4.89	-27.1	35.4
PM ₁₀ ($\mu\text{g}\cdot\text{m}^{-3}$)	EDGAR	699	0.338	44.30	-1.83	-2.8	55.5
	INEMA	697	0.478	29.75	-8.00	-12.3	34.7
PM _{2.5} ($\mu\text{g}\cdot\text{m}^{-3}$) daily means (≥ 18 paired hours)	EDGAR	26	0.416	6.03	-3.98	-21.9	25.7
	INEMA	25	0.813	5.42	-4.59	-25.8	26.0

Note: SO₂ observations were available only at La Florida, and the measured concentrations were used as representative values for Santiago in this study.

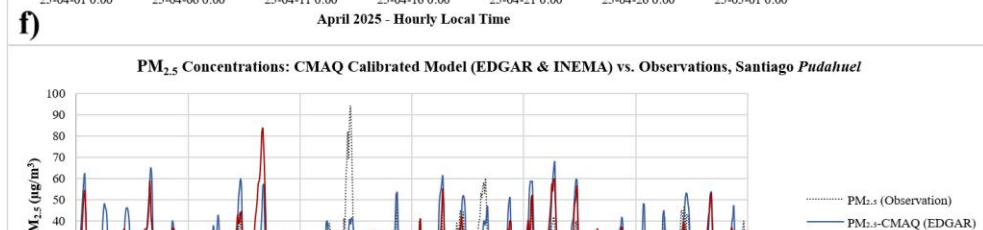
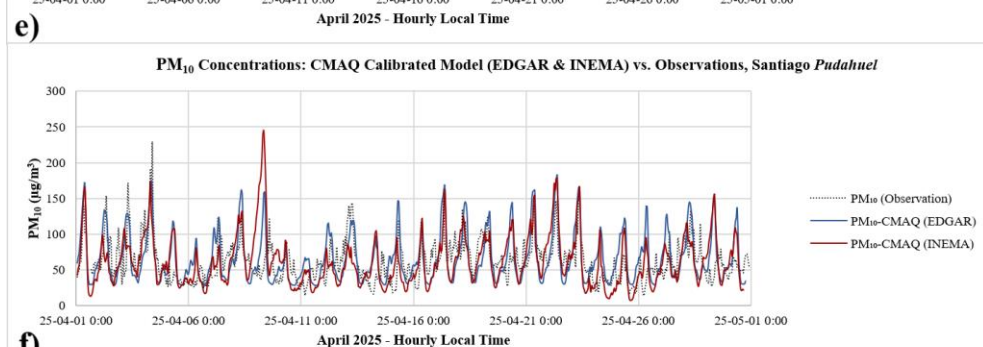
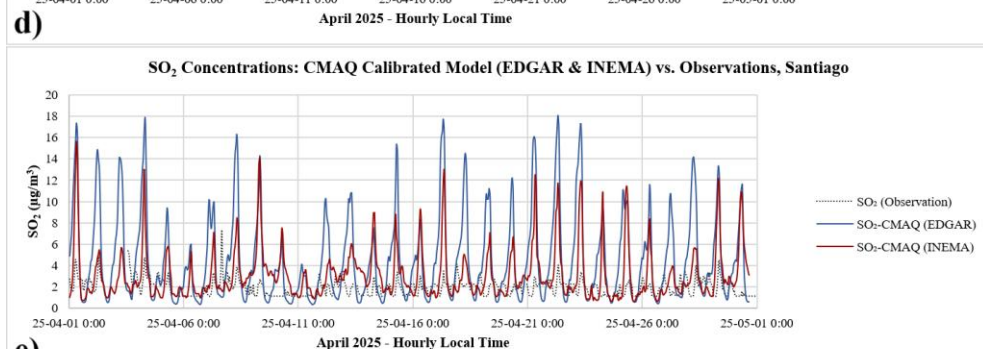
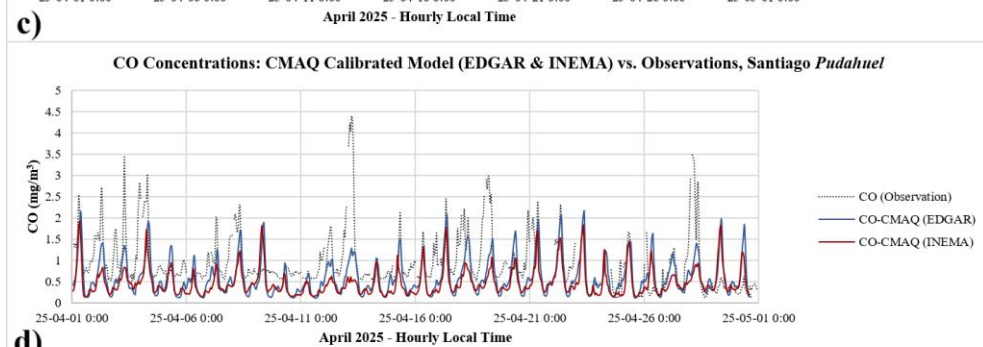
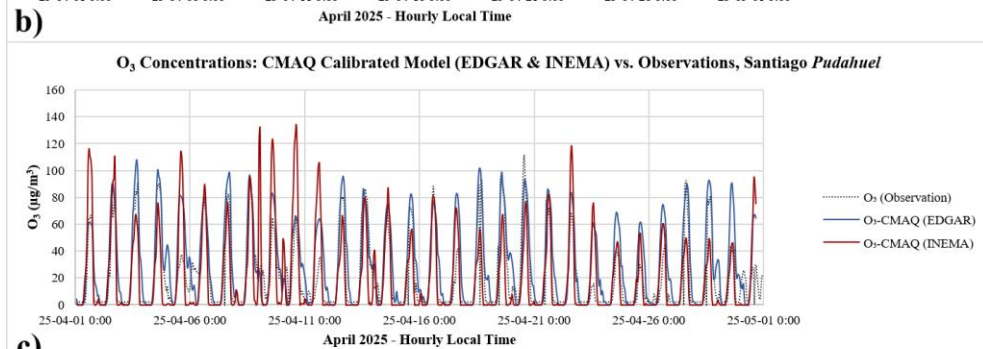
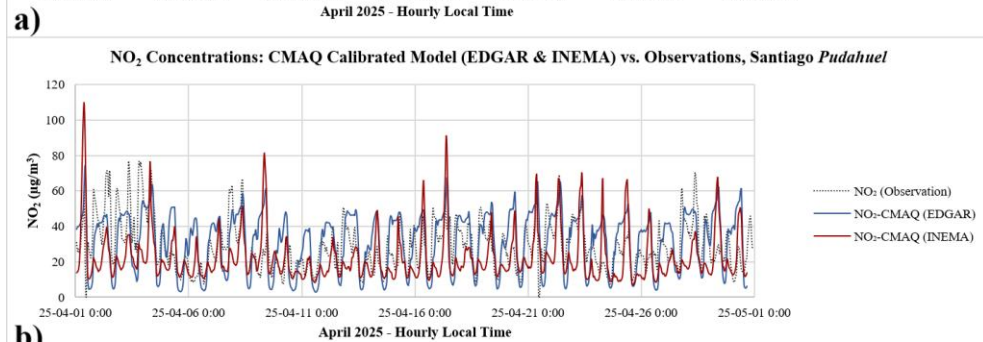
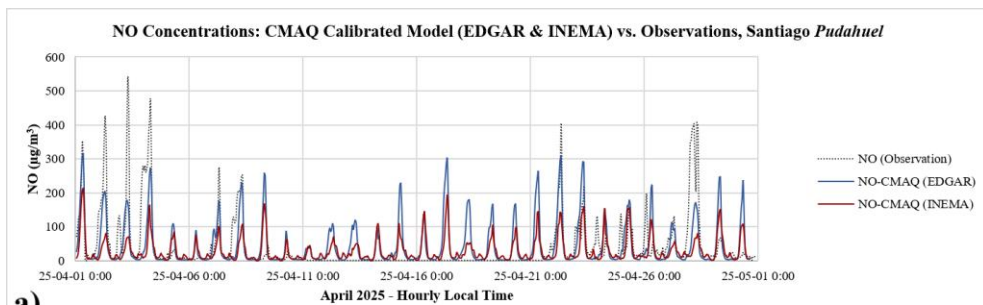


Figure 6. Hourly observed and calibrated CMAQ concentrations at Pudahuel during April 2025. CMAQ simulations driven by DESID-adjusted EDGAR v8.1 (2022) and INEMA v2 (2020) emissions are compared with SINCA observations for (a) NO, (b) NO₂, (c) O₃, (d) CO, (e) PM_{2.5}, (f) PM₁₀, and (g) SO₂. Model values correspond to the 1-km grid cell containing the monitoring station. SO₂ observations are provided from Santiago's representative dataset (La Florida station). Each panel shows hourly concentrations, with solid blue and red lines representing CMAQ outputs driven by EDGAR and INEMA emissions, respectively, and a dashed black line representing observations. All timestamps are reported in Santiago local time. Concentrations are reported in $\mu\text{g}\cdot\text{m}^{-3}$, except CO, which is reported in $\text{mg}\cdot\text{m}^{-3}$. Abbreviations: CMAQ = Community Multiscale Air Quality model; EDGAR = Emissions Database for Global Atmospheric Research; INEMA = Inventario Nacional de Emisiones Antropogénicas; NO = nitric oxide; NO₂ = nitrogen dioxide; O₃ = ozone; CO = carbon monoxide; PM_{2.5} = particulate matter with aerodynamic diameter $\leq 2.5 \mu\text{m}$; PM₁₀ = particulate matter with aerodynamic diameter $\leq 10 \mu\text{m}$; SO₂ = sulfur dioxide.

Table 6. Statistical performance of the calibrated 1-km EDGAR and INEMA CMAQ simulations at Pudahuel during April 2025.

Pollutant	Inventory	N	R	RMSE	MB	NMB (%)	NME (%)
O ₃ ($\mu\text{g}\cdot\text{m}^{-3}$)	EDGAR	712	0.847	18.29	+8.46	+39.5	58.1
	INEMA	712	0.724	21.18	-1.46	-6.8	60.6
NO ₂ ($\mu\text{g}\cdot\text{m}^{-3}$)	EDGAR	712	0.455	16.38	+3.37	+11.7	45.8
	INEMA	712	0.374	16.79	-6.14	-21.4	44.3
NO ($\mu\text{g}\cdot\text{m}^{-3}$)	EDGAR	711	0.511	74.53	+4.76	+12.0	104.4
	INEMA	711	0.376	76.15	-8.89	-22.5	102.1
CO ($\text{mg}\cdot\text{m}^{-3}$)	EDGAR	628	0.485	0.679	-0.340	-35.5	51.2
	INEMA	627	0.395	0.766	-0.466	-48.6	56.9
SO ₂ ($\mu\text{g}\cdot\text{m}^{-3}$)	EDGAR	687	0.263	4.77	+2.60	+133.0	162.3
	INEMA	687	0.265	2.69	+1.22	+62.2	90.1
PM _{2.5} ($\mu\text{g}\cdot\text{m}^{-3}$)	EDGAR	626	0.572	12.56	-1.54	-7.9	49.4
	INEMA	624	0.505	12.03	+0.68	+3.5	40.4
PM ₁₀ ($\mu\text{g}\cdot\text{m}^{-3}$)	EDGAR	703	0.493	32.69	+5.30	+8.6	41.6
	INEMA	701	0.437	34.63	+0.80	+1.3	41.6
PM _{2.5} ($\mu\text{g}\cdot\text{m}^{-3}$) daily means (≥ 18 paired hours)	EDGAR	26	0.645	5.27	-1.55	-7.9	20.0
	INEMA	25	0.517	6.81	+0.75	+3.8	25.8

Note: SO₂ observations were available only at La Florida, and the measured concentrations were used as representative values for Santiago in this study.

To place the calibrated model performance in quantitative context, the O₃ and PM_{2.5} results were compared with the historical photochemical-model benchmarks proposed by Emery and colleagues²⁰. These benchmarks summarize performance achieved in previous modeling applications and are intended as goals and criteria for contextual evaluation rather than strict pass-fail thresholds. For each metric, the first and second values denote the performance goal and criterion, respectively. For 1-h O₃, the “goal/criterion” values are |NMB|<5/15%, NME<15/25%, and R>0.75/0.50. EDGAR meets the O₃ correlation goal at all three stations (R=0.824–0.851), whereas INEMA meets the correlation criterion at all three stations (R=0.724–0.748) but remains below the goal. For O₃ NMB, EDGAR meets the criterion at Las Condes (-9.9%) and the goal at Parque O'Higgins (-3.2%) but exceeds the criterion at Pudahuel (+39.5%). INEMA exceeds the NMB criterion at Las Condes (-34.4%) and Parque O'Higgins (-33.9%) but meets the criterion at Pudahuel (-6.8%). Ozone NME exceeds the

25% criterion for both inventories at all three stations (43.3–60.6%). Consequently, none of the six site–inventory combinations satisfies all three O₃ criteria simultaneously, despite the relatively strong temporal correlations, particularly for EDGAR.

For 24-h PM_{2.5}, the corresponding goal/criterion values are |NMB|<10/30%, NME<35/50%, and R>0.70/0.40. At Las Condes, both calibrated configurations satisfy the PM_{2.5} NMB and NME goals but do not meet the correlation criterion. At Parque O’Higgins, both configurations satisfy all three PM_{2.5} criteria, with INEMA additionally meeting the correlation goal. At Pudahuel, both configurations also satisfy all three criteria; their NMB and NME meet the goals, whereas their correlations meet the criterion but not the goal. Thus, four of the six site–inventory combinations satisfy all three Emery et al. criteria for 24-h PM_{2.5}, although none satisfies all three goals simultaneously²⁰.

CMAQ-ISAM

The ISAM simulations were carried out using the calibrated EDGAR- and INEMA-based emission configurations. For each emission quantity scaled by a common coefficient across all contributing anthropogenic sectors, its relative sectoral allocation in the processed emission inputs is retained. This does not imply unchanged contributions to modeled concentrations, because the pollutant-specific adjustments also change the emission mixture and the resulting atmospheric chemical state. The ISAM results here characterize source apportionment within the respective calibrated configurations, rather than isolating the effects of the original inventories. Figure 7 presents the source apportionment of key pollutants in central Santiago during April 2025, obtained from CMAQ-ISAM simulations with DESID-adjusted EDGAR v8.1-2022 and INEMA v2-2020 emissions. The stacked horizontal bar chart illustrates the relative contributions of eight tagged emission sectors, together with relevant initial condition (ICO), boundary condition (BCO), and residual contributions (OTH). The modeled species are: NH₃ (ammonia), NO₃⁻ (particulate nitrate), NO_x, NO_z (total oxidized nitrogen compounds), EC (elemental carbon), NVPOA (non-volatile primary organic aerosol), O₃, total organic aerosol (TOA), total PM, SO₄²⁻ (particulate sulfate), SO₂, and VOCs. Figure 8 presents the corresponding comparative heatmaps.

Source apportionment in Santiago, April 2025 (CMAQ-ISAM)

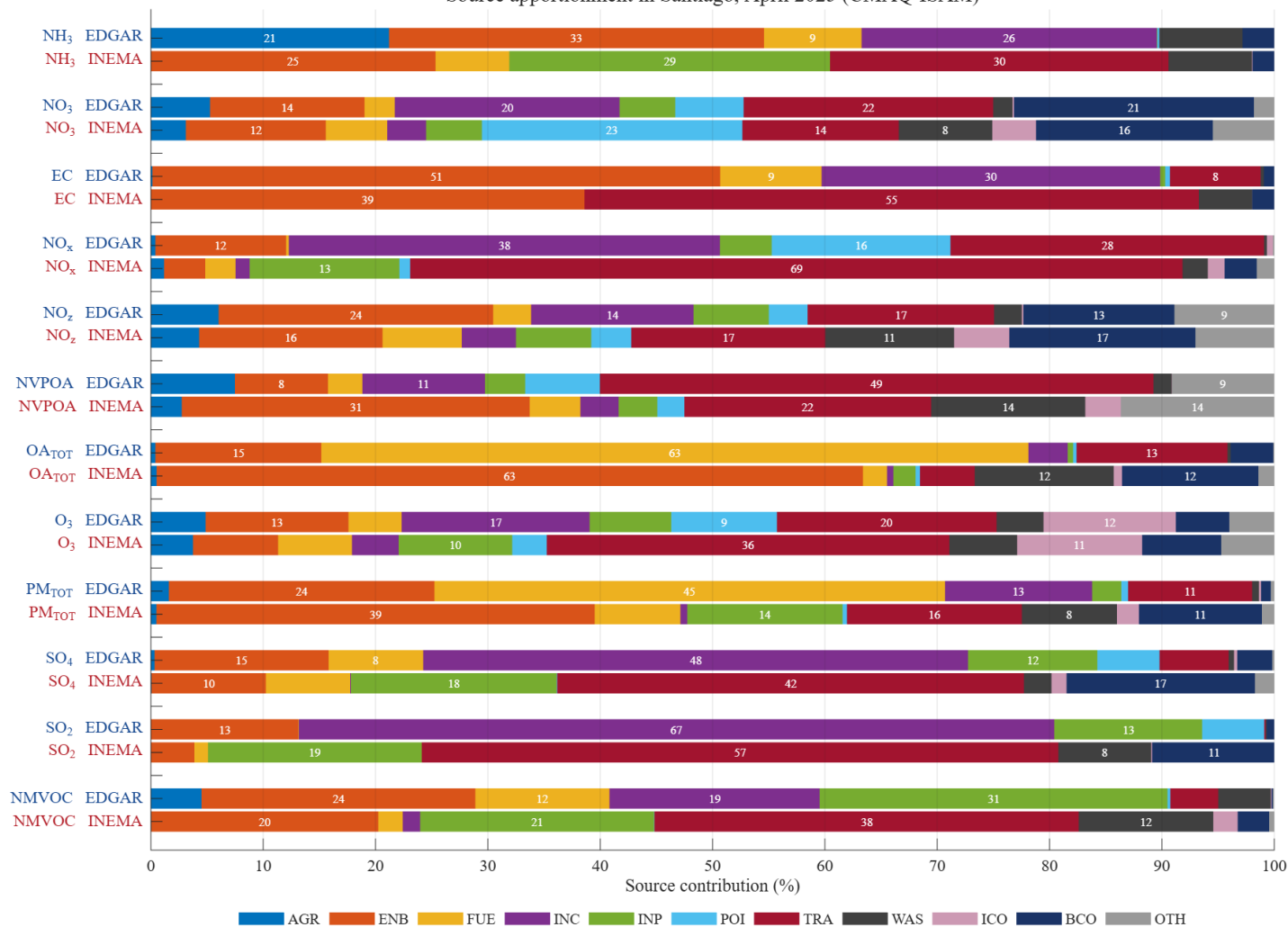


Figure 7. Inventory dependence of CMAQ-ISAM source apportionment in Santiago during April 2025.

Stacked horizontal bars show the percentage contribution of tagged sources to modeled concentrations in calibrated simulations driven by EDGAR v8.1 and INEMA v2 emissions. For each modeled species, separate EDGAR and INEMA bars show contributions from the eight harmonized anthropogenic sectors together with initial conditions (ICO), boundary conditions (BCO), and residual/other contributions (OTH). Numbers within bar segments indicate percentage contributions; each bar represents the total modeled concentration for that species, subject to rounding. The anthropogenic sectors are agriculture (AGR), energy for buildings (ENB), fuel exploitation (FUE), industrial combustion (INC), industrial processes (INP), power industry (POI), transport (TRA), and waste (WAS). Abbreviations: CMAQ = Community Multiscale Air Quality model; ISAM = Integrated Source Apportionment Method; EDGAR = Emissions Database for Global Atmospheric Research; INEMA = Inventario Nacional de Emisiones Antropogénicas; NH₃ = ammonia; NO₃ = particulate nitrate; NO_x = nitrogen oxides; NO₂ = total oxidized nitrogen compounds; EC = elemental carbon; NVPOA = non-volatile primary organic aerosol; TOA = total organic aerosol; O₃ = ozone; PM_{TOT} = total particulate matter; SO₄ = particulate sulfate; SO₂ = sulfur dioxide; VOC = volatile organic compounds.

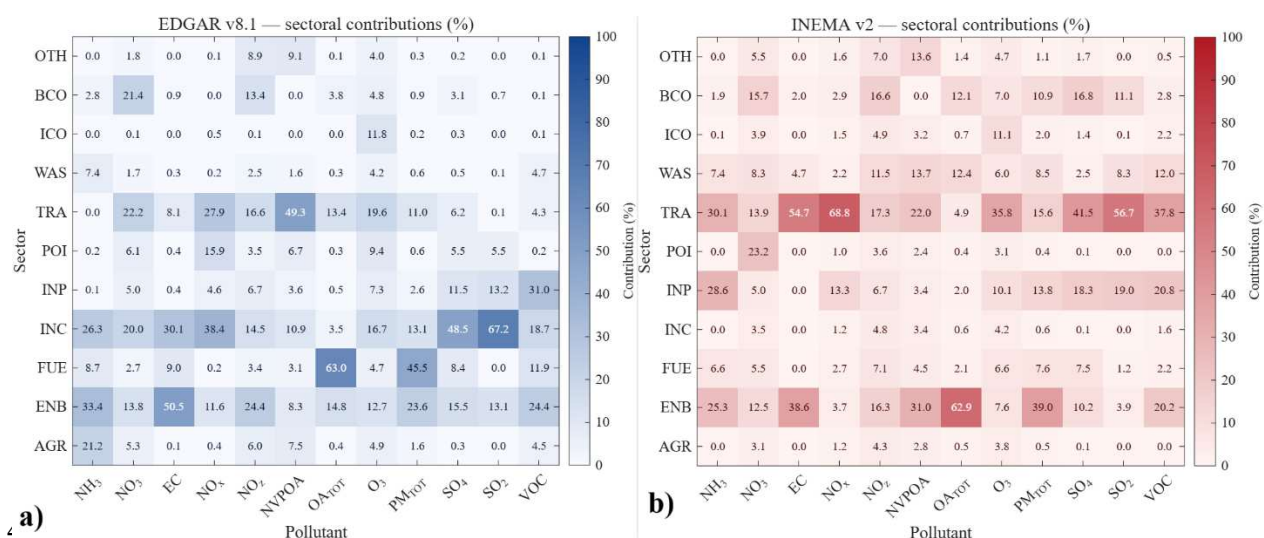


Figure 8. Heatmap representation of CMAQ-ISAM source apportionment in Santiago during April 2025 for calibrated simulations driven by (a) EDGAR v8.1 and (b) INEMA v2 emissions. Columns represent modeled species and rows represent the eight harmonized anthropogenic source sectors together with initial conditions (ICO), boundary conditions (BCO), and residual/other contributions (OTH). Each cell gives the percentage of the modeled concentration attributed to the corresponding source category, with increasing color intensity indicating a larger contribution. Numbers within cells indicate percentage contributions, subject to rounding. The anthropogenic sectors are agriculture (AGR), energy for buildings (ENB), fuel exploitation (FUE), industrial combustion (INC), industrial processes (INP), power industry (POI), transport (TRA), and waste (WAS). Abbreviations: CMAQ = Community Multiscale Air Quality model; ISAM = Integrated Source Apportionment Method; EDGAR = Emissions Database for Global Atmospheric Research; INEMA = Inventario Nacional de Emisiones Antropogénicas; NH₃ = ammonia; NO₃ = particulate nitrate; NO_x = nitrogen oxides; NO₂ = total oxidized nitrogen compounds; EC = elemental carbon; NVPOA = non-volatile primary organic aerosol; TOA = total organic aerosol; O₃ = ozone; PM_{TOT} = total particulate matter; SO₄ = particulate sulfate; SO₂ = sulfur dioxide; VOC = volatile organic compounds.

NVPOA requires separate interpretation due to specific regional emission events and application of dynamic boundary conditions (for dynamic boundary conditions see section Methods, CMAQ Processing). The April-mean NVPOA source contribution was strongly influenced by a short-lived episode dominated by BCO contributions. Here, NVPOA is evaluated after excluding this episode to provide a more representative characterization of April; detailed diagnostics are provided in Supplementary Information (Supplementary Results).

The EDGAR source apportionment distributes source influence across several combustion-related sectors. Industrial combustion (INC) appears repeatedly among the leading anthropogenic contributors and is especially important for sulfur-containing species, while energy use in buildings (ENB) contributes substantially to carbonaceous aerosols and several other pollutants. Transport (TRA) remains an important contributor to nitrogen-containing species. Overall, the EDGAR source structure is comparatively distributed among industrial combustion, buildings, transport, and other sectors rather than being dominated by a single anthropogenic category.

The INEMA source apportionment presents a different pattern, with a stronger relative influence of transport (TRA) across several pollutants. Transport is among the leading contributors to nitrogen-containing and combustion-related species, while energy use in

buildings (ENB) remains an important contributor to carbonaceous aerosol and total particulate matter. Industrial processes (INP) also contribute substantially to selected sulfur- and VOC-related species. Relative to EDGAR, INEMA shifts the modeled source hierarchy toward transport for several pollutants while retaining a significant role for buildings.

The contrast between EDGAR and INEMA is largest for several combustion-related species (NO_x , EC, and SO_2), for which INEMA attributes roughly twice the contribution to TRA compared with EDGAR. These differences are consistent with the distinct construction, sector definitions, and spatial allocation of the two inventories. ICO and BCO also make non-negligible contributions to selected secondary species, showing that the apportionment includes both locally generated and transported components.

Industrial combustion appears most frequently in the EDGAR source apportionment, whereas transport appears most frequently for INEMA. Energy use in buildings remains prominent in both. To summarize the recurring anthropogenic source hierarchy across species, Figure 9 counts how often each sector appears among the top three major ISAM contributors. This count is used only as a qualitative summary and does not assign equal atmospheric or policy importance to different pollutants.

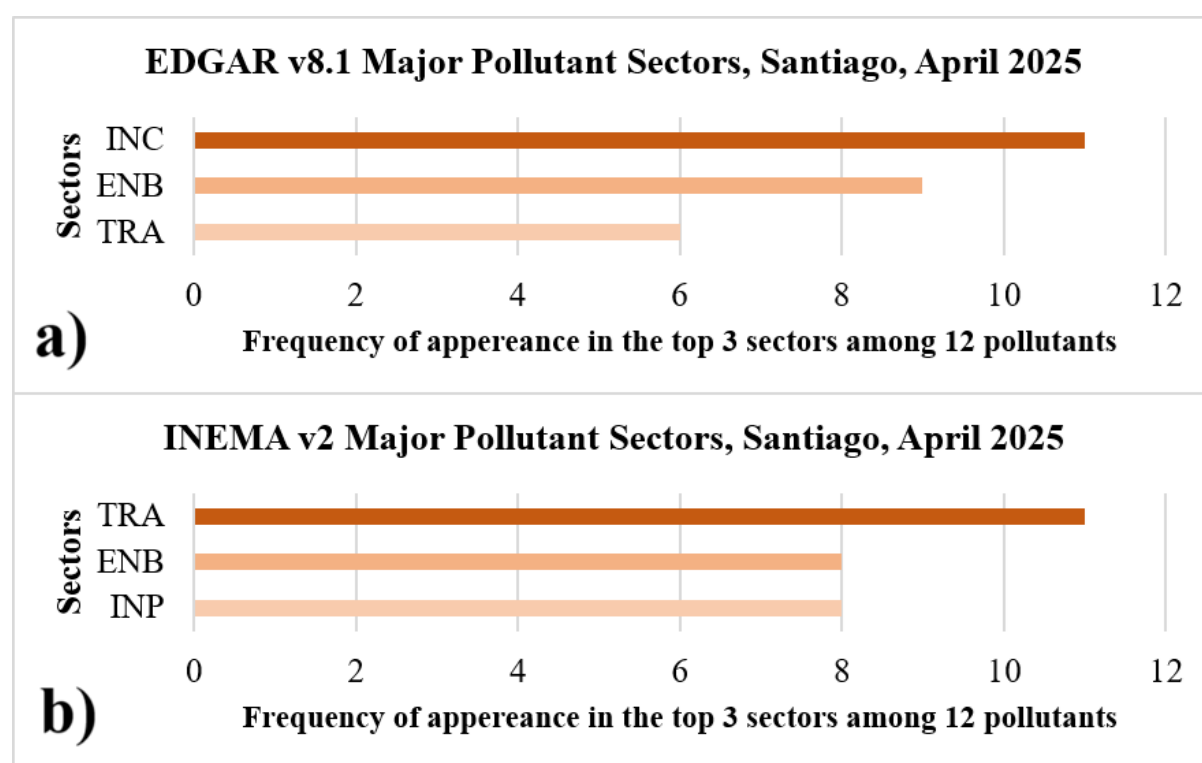


Figure 9. Ranking of emission sectors based on their frequency of appearance among the three leading CMAQ-ISAM contributors across 12 selected species (NH_3 , NO_3 , EC, NO_x , NO_z , NVPOA, TOA, O_3 , PM_TOT, SO_2 , SO_4 , and VOC) in Santiago during April 2025. Abbreviations: CMAQ = Community Multiscale Air Quality model; ISAM = Integrated Source Apportionment Method; NH_3 = ammonia; NO_3 = particulate nitrate; EC = elemental carbon; NO_x = nitrogen oxides; NO_z = total oxidized nitrogen compounds; NVPOA = non-volatile primary organic aerosol; TOA = total organic aerosol; O_3 = ozone; PM_TOT = total particulate matter; SO_2 = sulfur dioxide; SO_4 = particulate sulfate; VOC = volatile organic compounds.

As a qualitative summary of recurring top-three appearances across the twelve species, clear inventory-dependent hierarchies emerge. Under the EDGAR v8.1 inventory, Industrial

Combustion (INC) has the highest appearance frequency, appearing in the top three for 11 of the 12 pollutants, followed by Energy for Buildings (ENB) with nine appearances and Transport (TRA) with six. In contrast, the INEMA v2 inventory identifies TRA as the highest appearance frequency, ranking in the top three for 11 pollutants, while ENB and Industrial Processes (INP) tie for second place with eight appearances each. Figure 9 summarizes these frequencies without weighting the atmospheric importance of the individual species.

CMAQ-DDM/HDDM-3D

DDM-3D and HDDM-3D were applied to the calibrated EDGAR and INEMA emissions to quantify sector-resolved first- and second-order sensitivities of modeled pollutant concentrations to precursor-emission perturbations. First-order DDM coefficients describe the local linear response of concentration to an emission perturbation, whereas second-order HDDM coefficients describe the curvature of that response and therefore the change in first-order sensitivity as the emission perturbation increases^{16,19}.

The uncalibrated DDM/HDDM coefficients are mathematically valid local sensitivities of the base simulation but are not appropriate for quantitative extrapolation to the observationally constrained state. For this reason, the sensitivity analysis presented here is conducted after DESID calibration and is interpreted relative to the calibrated atmospheric state. Because the DESID adjustments differ between inventories, the calibrated sensitivity contrasts reflect both differences in the processed emission inputs and the effects of their respective adjustments on the emission baselines and chemical states. These contributions are not quantified separately in the present analysis. The DDM/HDDM sensitivities of the uncalibrated base model are provided in Supplementary Information (Supplementary Results).

Sector-by-sector sensitivities to NO_x, VOC, NH₃, SO₂, and directly emitted primary PM (PPM) emissions were quantified using CMAQ DDM/HDDM-3D for April 2025 over Santiago. The PPM species considered in the model include primary elemental carbon, primary nitrate, primary organic carbon, primary sulfate, primary ammonium, and primary non-carbon organic matter. First-order (DDM) sensitivities capture the linear response of pollutants (NO, NO₂, O₃, PM_{2.5}) to emission changes from each sector in the EDGAR v8.1 and INEMA v2 inventories. Second-order sensitivities (denoted here as “2NO_x” and “2VOC”, obtained via higher-order DDM/HDDM-3D) were also computed to assess nonlinearity. Hourly sensitivity values were derived for each sector over April 2025 and are expressed in concentration units (ppmv for gases; µg·m⁻³ for PM_{2.5}). Only one region, corresponding to the full extent of the 1-km domain, was considered in the DDM/HDDM analysis.

Supplementary Information (Supplementary Results) presents the sector-resolved sensitivity of key air pollutants to precursor emissions in Santiago during April 2025, derived from CMAQ-DDM-3D (first-order) and HDDM (second-order) simulations using the DESID-adjusted EDGAR v8.1 and INEMA v2 emission inventories, respectively. Figure 10 and Figure 11 present the corresponding heatmaps of median sensitivities for EDGAR and INEMA, respectively.

In the calibrated EDGAR configuration, NO and NO₂ sensitivities to NO_x are largest for industrial combustion (INC) and transport (TRA), with power industry (POI) also contributing

appreciably. Ozone sensitivity to NO_x is mostly negative, with the largest negative response associated with TRA and additional responses from INC and POI, whereas O_3 sensitivity to VOC emissions is positive and is strongest for energy use in buildings (ENB). The corresponding O_3 second-order NO_x sensitivities are predominantly positive, particularly for TRA, indicating curvature that opposes the negative first-order NO_x response. $\text{PM}_{2.5}$ shows a similarly nonlinear NO_x response: first-order $\text{PM}_{2.5}$ sensitivity to NO_x is negative, particularly for TRA and INC, whereas the second-order $\text{PM}_{2.5}$ sensitivity to NO_x is positive and pronounced for these sectors. Positive first-order $\text{PM}_{2.5}$ sensitivity to VOC is strongest for ENB. $\text{PM}_{2.5}$ sensitivities to NH_3 are greatest for fuel exploitation (FUE) and agriculture (AGR), sensitivity to SO_2 is dominated by INC, and the direct primary-PM response is strongest for ENB, INC and FUE.

The calibrated INEMA configuration shows a more sector-concentrated sensitivity structure. Transport (TRA) overwhelmingly dominates the NO response to NO_x emissions, whereas ENB produces the largest NO_2 sensitivity to NO_x emissions. The negative first-order O_3 sensitivity to NO_x is also dominated by TRA, while positive first-order O_3 sensitivity to VOC is concentrated mainly in ENB, waste (WAS) and TRA. The second-order O_3 sensitivity to NO_x is positive and again dominated by TRA. For $\text{PM}_{2.5}$, TRA produces the strongest negative first-order NO_x sensitivity and a pronounced positive second-order NO_x sensitivity. $\text{PM}_{2.5}$ sensitivity to VOC emissions is positive and dominated by ENB, with smaller contributions from WAS and TRA, whereas the second-order VOC response is negative mainly for ENB. First-order $\text{PM}_{2.5}$ sensitivity to NH_3 is concentrated in FUE and WAS, first-order $\text{PM}_{2.5}$ sensitivity to SO_2 is largest for TRA and ENB, and direct primary-PM sensitivities are greatest for ENB, industrial processes (INP) and TRA.

Both calibrated configurations show negative first-order O_3 sensitivity to NO_x and positive sensitivity to VOC emissions, indicating a VOC-limited photochemical response during April 2025. The agreement in sensitivity sign is notable despite the different sectoral distributions of the two inventories. Because VOC emissions were calibrated with DESID before the DDM/HDDM simulations, this diagnosis applies to the calibrated modeled state and should be interpreted in light of the absence of explicit biogenic VOC emissions.

The second-order results show that nonlinearity differs strongly among pollutant–precursor relationships. Ozone second-order sensitivities are generally smaller than the corresponding first-order responses, although positive second-order O_3 sensitivities to NO_x , particularly for TRA, indicate measurable curvature of the NO_x – O_3 response. In contrast, second-order $\text{PM}_{2.5}$ sensitivities to NO_x are substantial in both inventories and are especially pronounced for TRA, with an additional strong INC response under EDGAR. Their positive sign contrasts with the negative first-order $\text{PM}_{2.5}$ sensitivity to NO_x , emphasizing curvature of the $\text{PM}_{2.5}$ response to NO_x emissions. The $\text{PM}_{2.5}$ –VOC nonlinearity is more inventory dependent, including a negative second-order response for ENB under INEMA.

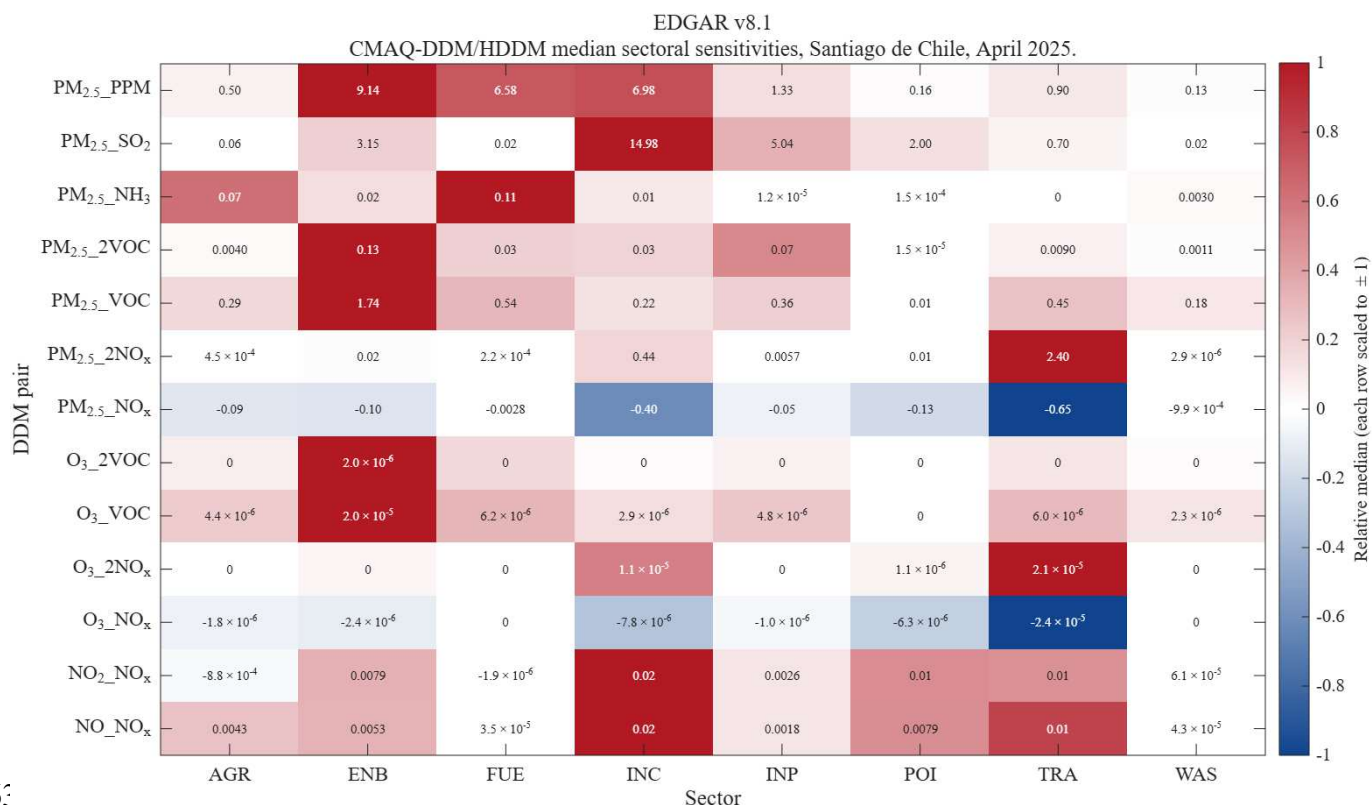


Figure 10. Sector-resolved median first- and second-order sensitivities from the calibrated EDGAR v8.1

CMAQ-DDM/HDDM simulation during April 2025. Each cell shows the median hourly sensitivity coefficient for a selected pollutant–precursor pair within one of the eight harmonized anthropogenic sectors.

Positive values indicate that an increase in the corresponding emission parameter increases the modeled pollutant concentration, whereas negative values indicate an inverse concentration response. Cell annotations report the median sensitivity coefficients, while the color scale is normalized independently by row to emphasize the relative sectoral pattern within each pollutant–precursor sensitivity pair. Values with $|x| < 10^{-6}$ are displayed as zero. From bottom to top, rows show sensitivities for NO_NO_x, NO₂_NO_x, O₃_NO_x, O₃_2NO_x, O₃_VOC, O₃_2VOC, PM_{2.5}_NO_x, PM_{2.5}_2NO_x, PM_{2.5}_VOC, PM_{2.5}_2VOC, PM_{2.5}_NH₃, PM_{2.5}_SO₂, and PM_{2.5}_PPM. Columns represent agriculture (AGR), energy for buildings (ENB), fuel exploitation (FUE), industrial combustion (INC), industrial processes (INP), power industry (POI), transport (TRA), and waste (WAS). Abbreviations: CMAQ = Community Multiscale Air Quality model; DDM = Decoupled Direct Method; HDDM = Higher-order Decoupled Direct Method; EDGAR = Emissions Database for Global Atmospheric Research; NO = nitric oxide; NO₂ = nitrogen dioxide; O₃ = ozone; NO_x = nitrogen oxides; PM_{2.5} = particulate matter with aerodynamic diameter ≤ 2.5 μm; VOC = volatile organic compounds; NH₃ = ammonia; SO₂ = sulfur dioxide; PPM = primary particulate matter; 2NO_x = second-order sensitivity to NO_x emissions; 2VOC = second-order sensitivity to VOC emissions.

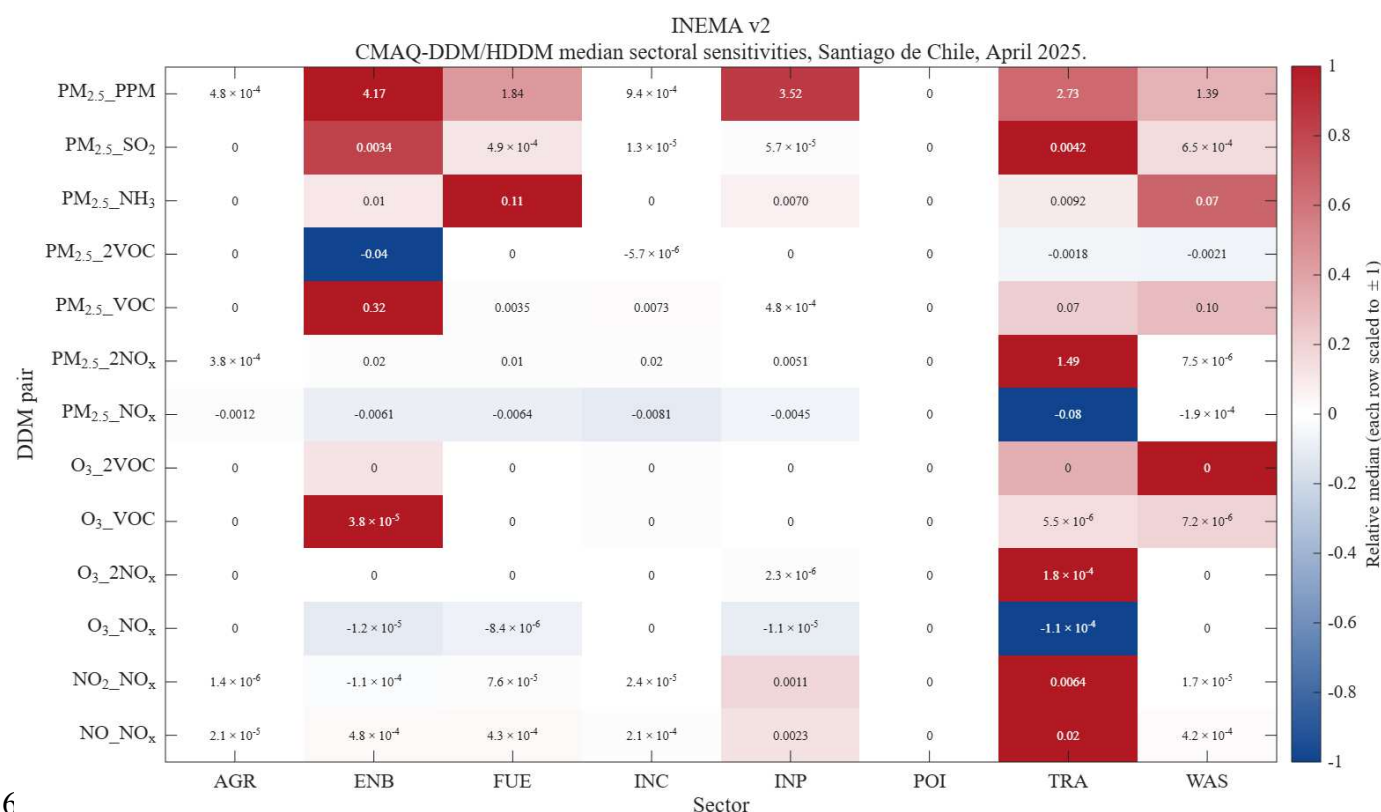


Figure 11. Sector-resolved median first- and second-order sensitivities from the calibrated INEMA v2 CMAQ-DDM/HDDM simulation during April 2025. Each cell shows the median hourly sensitivity

coefficient for a selected pollutant–precursor pair within one of the eight harmonized anthropogenic sectors.

Positive values indicate that an increase in the corresponding emission parameter increases the modeled pollutant concentration, whereas negative values indicate an inverse concentration response. Cell annotations report the median sensitivity coefficients, while the color scale is normalized independently by row to

emphasize the relative sectoral pattern within each pollutant–precursor sensitivity pair. Values with $|x| < 10^{-6}$ are

displayed as zero. From bottom to top, rows show sensitivities for NO_NO_x, NO₂_NO_x, O₃_NO_x, O₃_2NO_x,

O₃_VOC, O₃_2VOC, PM_{2.5}_NO_x, PM_{2.5}_2NO_x, PM_{2.5}_VOC, PM_{2.5}_2VOC, PM_{2.5}_NH₃, PM_{2.5}_SO₂, and

PM_{2.5}_PPM. Columns represent agriculture (AGR), energy for buildings (ENB), fuel exploitation (FUE),

industrial combustion (INC), industrial processes (INP), power industry (POI), transport (TRA), and waste

(WAS). Abbreviations: CMAQ = Community Multiscale Air Quality model; DDM = Decoupled Direct Method;

HDDM = Higher-order Decoupled Direct Method; INEMA = Inventario Nacional de Emisiones

Antropogénicas; NO = nitric oxide; NO₂ = nitrogen dioxide; O₃ = ozone; NO_x = nitrogen oxides; PM_{2.5} =

particulate matter with aerodynamic diameter $\leq 2.5 \mu\text{m}$; VOC = volatile organic compounds; NH₃ = ammonia;

SO₂ = sulfur dioxide; PPM = primary particulate matter; 2NO_x = second-order sensitivity to NO_x emissions;

2VOC = second-order sensitivity to VOC emissions.

Discussion

Emission inventory uncertainty enters CMAQ simulations through total emission magnitude, spatiotemporal patterns, chemical speciation, and sectoral distribution of emissions. The common meteorological and chemical-transport framework supports comparison of the two processed inventories before empirical scaling. The subsequent ISAM and DDM/HDDM analyses compare calibrated configurations and include the effects of inventory-specific DESID adjustments.

The mixed performance across pollutants does not support an overall preference for either inventory. EDGAR reproduces several characteristics of NO₂ and daytime O₃ more closely,

whereas INEMA performs more favorably for PM; CO performance is mixed. The lower reported VOC-to-NO_x ratio in INEMA and the large VOC adjustment required in the calibrated configuration warrant particular scrutiny of its precursor representation. However, because the VOC adjustment was constrained by O₃ rather than by direct VOC observations, its magnitude should be read as an O₃-informed estimate of the VOC scaling. Also, closer PM agreement does not establish that INEMA provides a more accurate representation of the complete emission mixture. Inventory suitability also depends on the sectoral distribution represented by the inventory, because different source activities have different temporal and chemical characteristics.

PM_{2.5} provides a clear example. Although the two inventories have similar domain-aggregated PM_{2.5} emissions, they differ substantially in their sectoral allocation and produce markedly different modeled concentrations. This shows that domain-aggregated emission magnitude alone does not determine inventory performance at the urban scale. The resulting concentration fields also depend on the spatial and temporal allocation of emissions, their distribution among source sectors, and their chemical speciation in the model-ready emissions.

The calibrated ISAM calculations show that the adjusted EDGAR- and INEMA-based configurations produce different modeled source hierarchies. These contrasts reflect both the inventory characteristics retained after preprocessing and the inventory-specific DESID adjustments. Although common scaling across sectors preserves the relative emission allocation of each individually scaled quantity, differing adjustments among pollutants change precursor ratios and can alter the resulting chemical states. Consequently, the ISAM and DDM/HDDM differences cannot be attributed exclusively to the original inventories. The present analysis does not quantify the separate contributions of inventory characteristics and calibration to these differences.

This distinction has practical consequences. Source-apportionment analyses are often used to prioritize emission controls, yet the apparent priority can change when the underlying inventory changes. If only the EDGAR attribution were considered, industrial combustion would receive greater relative emphasis for several species; if only the INEMA attribution were considered, transport would emerge more prominently. The disagreement should therefore be treated as information about inventory uncertainty rather than as evidence that one of these control strategies is necessarily correct.

Where both configurations identify the same sector as important, the conclusion is less sensitive to inventory choice. The recurring influence of energy use in buildings is one such example and supports continued investigation of residential-combustion mitigation using locally constrained emissions and activity data. Sector-to-sector comparison nevertheless remains approximate because the harmonized groups combine different native activities in the two inventories. For INEMA, temporal and chemical speciation profiles were assigned according to the underlying native activity rather than solely according to the harmonized group label. The remaining uncertainty concerns the representativeness of these activity-based cross-inventory profile matches, as well as differences in source definition and emission magnitude.

The calibrated EDGAR- and INEMA-based configurations differ in both the magnitude and sectoral pattern of their DDM/HDDM sensitivities, reflecting the combined influence of the processed emission inputs and their respective DESID adjustments. For NO, EDGAR distributes the dominant NO_x sensitivity mainly between industrial combustion and transport, whereas the INEMA response is strongly transport dominated. The NO₂ response differs further, with INC and TRA prominent under EDGAR but ENB dominating the INEMA sensitivity. These differences show that similar pollutant concentrations after calibration do not imply identical sector-resolved responses to subsequent emission changes.

Both calibrated configurations nevertheless give the same first-order directional response of O₃: negative sensitivity to NO_x and positive sensitivity to VOC emissions. This identifies a VOC-limited photochemical response in the calibrated April atmospheric states. Because the VOC abundance was adjusted before the sensitivity simulations, the chemical regime diagnosis reflects the resulting reactive atmospheric state, simulated without explicit biogenic emissions, rather than the original inventory source label assigned to the VOC emissions.

PM_{2.5} presents stronger nonlinear behavior. Negative first-order PM_{2.5} sensitivities to NO_x are accompanied by positive second-order NO_x sensitivities, particularly for TRA in both inventories and INC under EDGAR. Such nonlinear precursor responses are consistent with CMAQ scenario analyses showing that the relative PM_{2.5} response to NH₃, NO_x, VOC, and SO₂ controls can vary substantially with the magnitude of the emission perturbation²¹. Thus, the PM_{2.5} response to NO_x cannot be represented reliably by a simple linear extrapolation for large emission changes. The precursor structure also differs substantially between inventories. EDGAR shows particularly strong PM_{2.5} sensitivity to SO₂ from INC and to primary PM from ENB, INC and FUE, whereas INEMA shows much smaller SO₂ sensitivities and primary-PM sensitivity concentrated mainly in ENB, INP and TRA. The curvature of the PM_{2.5} response to VOC emissions also differs between inventories, including a negative second-order ENB response under INEMA.

To complement the full sector-by-sector DDM results, we examined median April 2025 sectoral sensitivities using heat-map summaries. Because the medians reduce the influence of short-lived extremes, they better reveal the persistent central tendency of the modeled chemical response. The EDGAR configuration shows a broader PM_{2.5} sensitivity structure, with strong median first-order responses to primary PM emissions from energy use in buildings, fuel exploitation, and industrial combustion, together with a pronounced sulfate-related response associated especially with industrial combustion. In contrast, the INEMA configuration presents a more concentrated structure, with the dominant PM_{2.5} sensitivities linked mainly to energy use in buildings, industrial processes, and transport, while sulfur-related median sensitivities are much weaker. In both inventories, PM_{2.5} displays negative first-order sensitivity to NO_x and positive second-order sensitivity to NO_x, most clearly for transport, indicating a nonlinear response of particulate matter to NO_x perturbations. The median O₃ sensitivities remain consistent with a VOC-limited regime, with positive first-order sensitivity to VOC emissions and negative first-order sensitivity to NO_x emissions, although the absolute median O₃ responses are small. Finally, the NO and NO₂ sensitivity patterns show that EDGAR

distributes the reactive-nitrogen response across several combustion sectors, whereas INEMA concentrates it much more strongly in transport.

The analysis covers April 2025 and therefore characterizes a single autumn period rather than the full seasonal cycle of Santiago. Emission activity, atmospheric stability, photochemical intensity, residential heating, and regional transport vary seasonally; source contributions and precursor sensitivities may consequently differ during winter stagnation episodes or summer O₃ events. These findings should therefore be interpreted as period-specific rather than annual climatological relationships²².

The inventory reference years also precede the modeled observational period. EDGAR v8.1 represents 2022 and INEMA v2 represents 2020, whereas the atmospheric simulation is evaluated against observations from 2025. Changes in traffic activity, vehicle technology, industrial production, fuel consumption, residential heating, and emission-control measures during the intervening years introduce uncertainty that is not independently separated from other inventory differences.

Explicit biogenic VOC emissions were not included. Instead, VOC emissions were scaled in DESID using observed O₃ as the target. The VOC-limited O₃ response therefore describes the photochemical regime of the calibrated state, not of the raw anthropogenic inventory. Ozone production depends on VOC abundance and reactivity rather than on the inventory source label, so the lack of an explicit biogenic stream does not, by itself, rule out a VOC-limited diagnosis. It does add uncertainty; scaling anthropogenic VOC can recover the domain-wide O₃ level without reproducing the composition, spatial pattern, or diurnal cycle of the true VOC mix, especially the more reactive biogenic fraction. Lightning NO_x and online windblown-dust emissions were also not activated and remain additional sources of uncertainty for NO_x and PM chemistry.

The DESID scaling factors are empirical model-adjustment coefficients rather than uniquely inferred corrections to the underlying inventories. A single pollutant-specific factor was applied across all corresponding anthropogenic sectors, so the calibration can compensate jointly for errors in inventory magnitude, spatiotemporal allocations, chemical speciation, missing source categories, and model processes.

The same three principal monitoring locations were used to construct the observational calibration targets and to evaluate the calibrated simulations. The resulting statistics therefore quantify post-adjustment agreement rather than independent out-of-sample validation. Closer agreement after scaling reflects compensation for underlying inventory and model deficiencies and does not, by itself, demonstrate that their causes have been identified or resolved. Future work should test the transferability of the derived coefficients against additional monitoring locations or periods not used in the adjustment procedure.

Observational representativeness also differs from model resolution. A 1-km CMAQ cell represents a spatially averaged concentration, whereas monitoring instruments respond to conditions at a specific location. Near road NO peaks, localized residential combustion, and industrial plumes can therefore produce short-duration concentrations that are not resolved by

the grid cell mean. The template-based conversion used to enable INEMA processing in SMOKE involved intermediate conservative regridding from $0.01^{\circ} \times 0.01^{\circ}$ to $0.1^{\circ} \times 0.1^{\circ}$. The influence of this intermediate spatial aggregation on model performance was not separately assessed.

Temporal allocation introduces an additional uncertainty. The use of a common temporal profile framework improves comparability between the two inventories, but sector-average monthly, weekly, and hourly profiles cannot reproduce every local activity pattern surrounding individual monitoring sites. Timing errors in emissions can interact with boundary layer evolution and thereby affect both primary pollutant peaks and secondary chemistry.

Meteorological uncertainty remains intertwined with the emission comparison. WRF dynamically calculates planetary-boundary layer evolution using the YSU scheme rather than prescribing a fixed mixing depth. Nevertheless, errors in wind speed, wind direction, nocturnal stability, or the timing of morning boundary layer breakup can alter pollutant accumulation and transport.

We reconstructed eight unavailable GEOS-CF boundary condition hours per day by linear interpolation between adjacent available fields. This provided continuous boundary forcing but can smooth short-lived transported features. The importance of resolving such episodic transport has also been demonstrated in WRF-CMAQ simulations, where cross-regional transport can dominate particulate pollution during individual episodes and depends strongly on boundary layer dynamics and chemical evolution²³. The episode-sensitive NVPOA response therefore reinforces the need to distinguish event-scale boundary effects from monthly-average source attribution.

The principal management implication of this work is that emission control priorities inferred from CTM analyses should be considered together with uncertainty in the underlying emission inventory.

Where the two inventory configurations converge, confidence in the qualitative priority is comparatively greater. Energy use in buildings remains important across multiple species under both EDGAR and INEMA, supporting continued evaluation of residential-combustion mitigation and cleaner household-energy technologies.

INEMA places stronger emphasis on transport for several key pollutants, supporting continued assessment of fleet modernization, high-emitter controls, low-emission transport policies, and public transport electrification. The magnitude of the expected air quality benefit, however, should be quantified using scenario simulations rather than inferred directly from source-apportionment percentages.

The stronger industrial contribution in EDGAR indicates that industrial source characterization also warrants independent evaluation, particularly where global and locally developed inventories disagree strongly. Facility-level activity data, stack information, fuel use, and local emission measurements could help determine whether the modeled discrepancy reflects inventory magnitude, spatial allocation, or source classification.

Energy for Buildings (ENB) and Transport (TRA) are the two sectors that rank among the three most frequently recurring major contributors in both inventories. This convergence is highly relevant for Santiago. Residential wood-stove replacement programs, such as the “Recambio de Calefactores” initiative, specifically target the ENB sector by replacing inefficient wood-burning stoves with cleaner technologies^{24,25}. Such measures could deliver substantial co-benefits for elemental carbon (EC), organic aerosol (OA), PM_{2.5}, and NH₃, regardless of which emission inventory is adopted. Industrial sources (INC/INP) follow closely, while agriculture, waste, and power generation have more localized influences.

For O₃ management, both calibrated DDM/HDDM configurations indicate a VOC-limited response during April 2025. Under this modeled state, reductions in VOC emissions act in the direction of decreasing O₃, whereas NO_x-only reductions can reduce titration and increase urban O₃. The nonlinear response indicated by HDDM means that the magnitude of larger precursor-control scenarios should be evaluated with explicit scenario simulations rather than by linear extrapolation of the first-order sensitivities.

The study therefore identifies several priorities for future inventory improvement in Santiago: updated activity years, locally constrained temporal profiles, improved vehicle-fleet and residential-heating information, enhanced industrial spatial allocation, mechanism-relevant VOC speciation, explicit natural VOC emissions, and routine mass-balance and spatiotemporal quality assessment of model-ready emissions.

More broadly, the study distinguishes differences between emission inventories before scaling from differences between their calibrated model configurations. Disagreement in the calibrated source-apportionment and sensitivity results highlights their dependence on both inventory representation and empirical adjustment. Separating these influences remains necessary before attributing the diagnostic contrasts to the original inventories alone.

Methods

The modeling framework was designed to investigate the influence of emission inventory selection on urban-scale pollutant distributions, source contributions, and chemical sensitivities during April 2025 in Santiago. Three one-way nested domains were configured with horizontal resolutions of 9, 3, and 1 km, centered over the Santiago Metropolitan Region. The nested-domain configuration enabled simultaneous representation of regional-scale transport processes and fine-scale spatiotemporal variations in urban pollutant levels associated with complex terrain, localized emissions, and boundary layer dynamics. Meteorological fields were generated using WRF v4.1.1 driven by ERA5 reanalysis data, while high-resolution geographical datasets derived from MODIS and Noah-modified IGBP classifications were employed to improve land-surface representation and urban meteorological processes.

Anthropogenic emissions were processed using SMOKE v5.2.1 for two distinct inventories: the global EDGAR v8.1 (2022) Air Pollution inventory and the regional high-resolution INEMA v2 (2020) inventory. Separate preprocessing chains were developed for each inventory to ensure compatibility with SMOKE/CMAQ and the nested-domain configuration. The comparative implementation of these inventories enabled evaluation of the sensitivity of

simulated pollutant concentrations to differences in emission magnitude and spatial allocation. Both inventories were processed within a common temporal-allocation and chemical-mechanism framework where feasible, while retaining the inventory-specific preprocessing required by their native structures.

In addition to standard chemical-transport simulations, ISAM and sector-resolved DDM-3D/HDDM-3D were executed with the calibrated emissions to quantify source contributions and first- and second-order pollutant sensitivities. ISAM was employed to evaluate the relative contributions of different source sectors to NO_x, NO_z, NH₃, NO₃, O₃, PM, SO₄, VOC, SO₂, OA, EC, and NVPOA concentrations. DDM/HDDM evaluated responses to perturbations in NO_x, VOC, NH₃, SO₂ and directly emitted primary PM (PPM), with first-order coefficients describing the local linear response and second-order coefficients describing curvature of the emission–concentration relationship.

The overall methodological workflow consists of: (1) harmonization of EDGAR and INEMA sectors and preprocessing of raw emission inputs for SMOKE; (2) meteorological simulations using WRF; (3) emissions processing using SMOKE; (4) CMAQ chemical-transport simulations and calibration; (5) source apportionment using CMAQ-ISAM; and (6) sensitivity analysis using CMAQ-DDM/HDDM-3D.

WRF Processing

Meteorological simulations were performed using WRF model v.4.1.1 to generate the atmospheric fields required for the air quality simulations²⁶. WRF was configured using three one-way square nested domains with horizontal resolutions of 9 km, 3 km, and 1 km, each with 121×121 cells (1080, 360, and 120 km side squares) centered over the Santiago Metropolitan Region. The spatial extent of the WRF modeling domains is shown in Figure 12. WRF domains were configured with 121 × 121 grid points, yielding 120 × 120 horizontal CMAQ/MCIP grid cells after processing. The ERA5 reanalysis datasets (0.25°×0.25°) of hourly data on pressure levels (37 levels) and on single levels were obtained for the month of April 2025²⁷. In addition, high-resolution geographical static data (modis_landuse_20class_30s_with_lakes) for the WRF Preprocessing System (WPS) were obtained from Mesoscale & Microscale Meteorology (MMM) laboratory of National Center for Atmospheric Research (NCAR)²⁸.



Figure 12. Modeling domain configuration of WRF over Santiago. The image was created by WRF Domain Wizard²⁹. The background map uses Esri basemap tiles with standard attribution (Leaflet, Esri and contributors). The figure shows a map of central Chile and western Argentina with three one-way nested square domains (three nested red square outlines) centered on Santiago. The outermost domain has 9 km horizontal resolution and covers approximately 1,080 km on each side; the middle domain has 3 km resolution and covers 360 km on each side; the innermost domain has 1 km resolution and covers 120 km on each side. A blue marker indicates the location of central Santiago. The domains are overlaid on a topographic and political map showing the Andes Mountains to the east and the Coastal range to the west.

The complete WRF configuration and meteorological evaluation are provided in Supplementary Information (Supplementary Materials).

Emission Inventory Preprocessing

EDGAR v8.1 and INEMA v2 are distributed with different spatial grids, temporal representations, source classifications, and units and therefore required separate preprocessing before SMOKE processing. The objective was to transform both inventories into technically compatible SMOKE-ready inputs for comparison within the adopted preprocessing workflow while retaining their differences in emission magnitude and source representation.

The EDGAR v8.1_AP air-pollutant dataset provides global $0.1^\circ \times 0.1^\circ$ emissions of SO_2 , NO_x , CO, NMVOC, NH_3 , PM_{10} , $\text{PM}_{2.5}$, BC (Black Carbon), OC (Organic Carbon) from 1970 to 2022 by sector. Annual time series and sector-specific annual and monthly grid maps are provided. The EDGAR dataset uses distinct sector classifications for the annual and monthly products. For the annual sector-specific grid maps, the sectors are detailed into 25 sectors (with originally 27 EDGAR codes). However, for the monthly sector-specific grid maps the sectors are categorized into eight major (integrated) classes of power industry, industrial combustion, buildings, transport, agriculture, fuel exploitation, processes, and waste. Table 7 presents the description of EDGAR's (v8.1) eight major sectors⁷. It should be noted that the monthly sector-specific gridded map files are provided in NetCDF format, where each individual file contains the complete set of monthly emissions data corresponding to all twelve months of a given year. The EDGAR files are available in two units of *emissions* (Tonnes) and *fluxes* ($\text{kg}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$), the latter of which is essential for SMOKE/CMAQ processing. To prepare the EDGAR v8.1 files for SMOKE processing in monthly simulations, the following steps were taken:

- a) Extract the data for the specific month from the raw Monthly file (the EDGAR v8.1 raw Monthly files contain data for all 12 months of the corresponding year).
- b) Convert the longitudes from [-180, 180] convention to [0, 360].
- c) Compute the PM_OTH files as: $[PM_{2.5} - (BC + OC)]$.
- d) Update the year of the files based on the year of running simulations (this must be carried out if no *Cntlmat* & *Grwinven* scenarios in SMOKE are considered, see section Methods, SMOKE processing).

EDGAR v8.1 also provides high-resolution temporal profiles which are essential for processing hourly sector-specific emissions. Although hourly emission time series and grid maps are not provided in NetCDF format, the monthly, weekly (seven days), and hourly profiles are sufficient for SMOKE processing requirements. Table 8 provides an explicit mapping between the 27 EDGAR v8.1 sector codes and the eight major EDGAR sectors^{30,31}.

Table 7. Description of EDGAR's (v8.1) eight major sectors

Integrated Sectors	Description
Agriculture	Agriculture - Agricultural soils, burning crop residues, enteric fermentation, manure management, indirect N ₂ O emissions from agriculture
Energy for Buildings	Buildings - Small scale non-industrial stationary combustion
Fuel Exploitation	Fuel exploitation - Production, transformation and refining of fuels
Industrial Combustion	Industrial combustion - Combustion for industrial manufacturing
Industrial Processes	Processes - Industrial processes (e.g. emissions from the production of cement, iron and steel, aluminum, chemicals, solvents, etc.)
Power Industry	Power Industry - Power and heat generation plants (public & auto producers)
Transport	Transport - Mobile combustion (road & rail & ship & aviation)
Waste	Waste - Solid waste disposal and wastewater treatment

Table 8. Mapping between the eight integrated EDGAR v8.1 sectors and the original EDGAR sector codes. Definitions of the original sector codes are provided in Table 5 of Crippa et al. (2020)³⁰.

EDGAR v8.1 Integrated Sectors	EDGAR Subsectors
Agriculture	AGS+AWB+ENF+MNM+N2O+IDE+BMB
Energy for Buildings	RCO
Fuel Exploitation	PRO+REF+TRF+FFF
Industrial Combustion	IND+NEU
Industrial Processes	CHE+FOO+SOL+IRO+NFE+NMM+PAP+PRU
Power Industry	ENE
Transport	TNR+TRO
Waste	SWD+WWT

The INEMA inventory provides emissions for vehicular, point sources (industrial, energy, etc.), residential, forest fires, and agricultural waste burning sectors estimated for 2015–2020 and spatially distributed onto a 0.01°×0.01° resolution, while extension boundaries of the covering domain are latitudes: [-66, -17], and longitudes: [-76, -56.3]. Pollutants included for all sectors are CO₂, NO_x, SO₂, CO, VOCs, NH₃, PM₁₀, and PM_{2.5}. Also, CH₄, N₂O, and BC are included for transport, forest fires, agricultural waste burning, and residential sources. Table 9 summarizes the sector structure of the INEMA emission inventory. Table 10 provides a detailed breakdown of the pollutants reported in each sector in INEMA v2 inventory.

961

Table 9. INEMA emission inventory sectors.

INEMA Sectors, V1	INEMA Sectors, V2
v1.1 Energy	V2.1 Energy Production
v1.2 Industry	V2.2 Industrial Energy Consumption
v1.3 Mining	V2.3 On Road Transport Energy Consumption
v1.4 Residential	V2.4 Commercial Energy Consumption
v1.5 Transport	V2.5 Residential Firewood Consumption
	V2.6 Agriculture Energy Consumption
	V2.7 Industrial Processes
	V2.8 Agriculture Waste Burning
	V2.9 Forest Fires
	V2.10 Transport Non-Exhaust Emissions

962

Table 10. INEMA v2 emission inventory pollutants per sector (per Table 9).

Pollutant Sector	BC	CH ₄	CO	CO ₂	N ₂ O	NO	NO ₂	NO _x	NH ₃	PM	PM _{2.5}	PM ₁₀	SO ₂	SO _x	VOC	NMVOC
v2.1			✓	✓				✓	✓		✓	✓	✓		✓	
v2.2			✓	✓				✓	✓		✓	✓	✓		✓	
v2.3	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓			✓	✓	✓
v2.4			✓	✓				✓	✓	✓	✓	✓	✓		✓	
v2.5	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓		✓	✓
v2.6			✓	✓				✓	✓		✓	✓	✓		✓	
v2.7			✓	✓				✓	✓		✓	✓	✓		✓	
v2.8		✓	✓	✓	✓			✓	✓		✓	✓		✓	✓	
v2.9	✓	✓	✓	✓	✓			✓	✓		✓	✓		✓	✓	✓
v2.10											✓	✓				

963 The INEMA emission data are available as annual NetCDF files in gigagrams per year (Gg·yr⁻¹).
 964 ¹). The native INEMA NetCDF files were not directly compatible with the SMOKE input
 965 workflow used in this study. EDGAR NetCDF files were used as templates to construct
 966 SMOKE-compatible inputs from the INEMA emissions. To prepare the raw INEMA files for
 967 monthly SMOKE simulations, the following steps were taken:

- 968 a) Calculate grid-cell areas and convert emissions from Gg·year⁻¹ to kg·m⁻²·s⁻¹.
- 969 b) Conservatively regrid the native 0.01° × 0.01° INEMA emission fields to the 0.1° ×
 970 0.1° grid of the EDGAR NetCDF templates used in this implementation.
- 971 c) Sort the latitudes and longitudes (from minimum negatives to maximum positives).
- 972 d) Convert the longitude from [-180, 180] convention to [0, 360].
- 973 e) Map INEMA sectors to eight comparable EDGAR-inspired sectors.
- 974 f) Compute the PM_{OTH} emissions: due to inconsistent availability of BC and OC, we
 975 approximate the PM_{OTH} by domain-aggregated EDGAR emissions (t) as (PM_{2.5}-BC-
 976 OC)/PM_{2.5}=(6619.91-849.52-1918.87)/6619.91=0.582, rounded to 0.6 [PM_{OTH}≈0.6
 977 × PM_{2.5}].
- 978 g) Update the year of the files based on the year of running simulations (this must be
 979 carried out if no *Cntlmat* & *Grwinven* scenarios in SMOKE are considered, see section
 980 Methods, SMOKE processing).

- h) Since the emission data files are only available in annual temporal format, monthly processing is handled solely by the SMOKE *Temporal* program.

Table 11. Correspondence between INEMA v2 source categories and the EDGAR-inspired eight processing groups used in this study.

Processing group	INEMA v2 source categories assigned to the group
Agriculture	Agriculture Energy Consumption
Energy for Buildings	Residential Firewood Consumption
Fuel Exploitation	Commercial Energy Consumption
Industrial Combustion	Industrial Energy Consumption
Industrial Processes	Industrial Processes
Power Industry	Energy Production
Transport	On Road Transport Energy Consumption + Transport Non-Exhaust Emissions
Waste	Agriculture Waste Burning + Forest Fires

Note: The harmonized sector mapping was used for inventory comparison and sector mapping. Temporal and chemical speciation profiles were assigned according to the underlying INEMA activity, as described in the corresponding SMOKE processing sections.

SMOKE Emissions Processing

Anthropogenic emissions were processed using the SMOKE modeling system version 5.2.1 to generate temporally allocated, chemically speciated, and gridded CMAQ-ready emissions. The SMOKE processing chain was implemented interactively to allow detailed control over inventory harmonization, sector mapping, and chemical speciation procedures required for both EDGAR v8.1 and INEMA v2 inventories. Core SMOKE modules including *smkinven*, *temporal*, *spcmat*, *grdmat*, and *smkmerge* were executed one after another to transform the raw inventory datasets into CMAQ-ready emissions fields.

Because the *smkmerge* module requires pre-existing control matrices generated through *cntlmat*, the *cntlmat* processor was executed using null-control configurations prior to emissions merging. This approach ensured compatibility with the SMOKE processing workflow while preserving the original emissions without applying control strategies or reduction scenarios during the preprocessing stage. The resulting emissions datasets were spatially allocated to the nested CMAQ/WRF domains, temporally distributed into hourly emissions, and chemically speciated to the selected CMAQ CB6R5 chemical mechanism. The complete EDGAR–SMOKE and INEMA–SMOKE processing chains were re-examined. Quality assurance included comparisons of inventory and model-ready emission totals, inspection of sector-resolved spatial fields, evaluation of hourly temporal behavior, and verification of unit, grid, and chemical-speciation transformations. Representative CMAQ-ready emission fields for EDGAR v8.1 and INEMA v2 are provided in the Supplementary Information (Supplementary Materials), illustrating the spatial and temporal variability across the eight harmonized sectors for selected pollutants. The following sections describe the SMOKE *temporal* and *spcmat* settings in more detail.

SMOKE High-Resolution Temporal Processing

Temporal allocation of anthropogenic emissions was performed within SMOKE to transform the inventory emissions into hourly emission fields compatible with the temporal resolution requirements of CMAQ. Since anthropogenic activities exhibit substantial temporal variability at seasonal, weekly, and diurnal scales, the application of representative temporal profiles is

essential for accurately reproducing the temporal evolution of atmospheric pollutant concentrations. CTMs require temporally resolved emissions to properly simulate photochemical processes, pollutant accumulation, and episodic air quality events.

In this study, high-resolution temporal profiles developed within the EDGAR framework were used and adapted for both EDGAR v8.1 and INEMA v2 emission inventories³⁰. For INEMA, temporal profiles were assigned according to the underlying native activity rather than solely according to the harmonized sector label. Commercial Energy Consumption, although mapped under FUE for sector comparison, was assigned the EDGAR Small Combustion temporal profile. Agricultural Waste Burning and Forest Fires, although mapped under WAS, were assigned the EDGAR Agricultural Waste Burning and Large-Scale Biomass Burning temporal profiles, respectively. The EDGAR temporal profile dataset provides country- and sector-specific temporal allocation factors, enabling the disaggregation of annual emissions into monthly, weekly, and hourly emissions. These profiles account for temporal variations associated with energy demand, transportation activity, industrial operations, and residential combustion, thereby improving the representation of real-world emission dynamics within CTM simulations.

For EDGAR v8.1, sector-specific monthly profiles, hourly profiles and day-of-week profiles (distinguishing all seven days) were extracted and harmonized through sector mapping of the 27 available EDGAR activity categories. April monthly emissions were supplied directly to SMOKE. Day-of-week and hourly profiles were then used to allocate those April totals within the month. The annual-to-month scaling step was left in place, as required for correct processing in SMOKE v5.2.1 (April monthly total was verified to be preserved during temporal allocation.).

WRF and CMAQ were run in UTC. Local temporal activity profiles were converted consistently to the corresponding model timestamps before being applied. The April 2025 daylight-saving transition in Santiago was explicitly considered during verification of the processed temporal profiles. Inspection of the final CMAQ-ready hourly emissions did not identify a persistent phase displacement.

SMOKE Chemical Speciation (PM & NMVOC)

Chemical speciation required additional processing because neither inventory provides a native mapping directly to the CMAQ CB6R5 mechanism. EDGAR provides sector-specific NMVOC information in the 25-group GEIA framework, whereas additional profiles were required for PM and for conversion of the inventory species into mechanism-compatible CMAQ emissions. Consequently, a customized speciation framework was developed to generate chemically consistent and SMOKE-compatible emissions for all modeled pollutants and precursor species.

To address this limitation, PM speciation profiles derived from the HTAP v3.1 framework were adapted and harmonized for compatibility with EDGAR v8.1 sector definitions and the CB6R5 mechanism³¹. Since several of the eight major EDGAR sectors did not directly correspond to the original HTAP source classification structure, additional preprocessing procedures were

implemented to calculate representative speciation factors (e.g., species-dependent split, divisor, and mass fraction parameters) for newly aggregated sectors.

Species-dependent factors were calculated using an emission-weighted averaging methodology based on pollutant-specific emission contributions from the original sectors to preserve both mass and mole consistency. For each pollutant–species pair, new mass fractions were computed as the emission-weighted average of the original sectoral mass fractions, while mole-based factors were derived from weighted combinations of split-to-divisor ratios from the original HTAP profiles. Table 12 presents the “EDGAR-equivalent” sector mapping between the EDGAR v8.1 inventory and the HTAP v3.1 speciation profiles used in this study.

Table 12. EDGAR-equivalent sector mapping for HTAP v3.1 speciation profiles.

EDGAR Sectors	“EDGAR-Equivalent” Sector Mapping of HTAP v3.1 Speciation
Agriculture	Agriculture
Energy for buildings	Residential
Fuel exploitation	Energy + Industry
Industrial combustion	Industry
Industrial processes	Industry
Power industry	Energy
Transport	Aircraft + Ships+ Transport
Waste	Waste

For the NMVOCs, EDGAR v8.1 provides globally distributed sector-specific speciated emissions based on the GEIA (Global Emissions Inventory Activity) 25-group classification framework, consisting of 25 individual NMVOC species and species groups³². These groups represent a broad range of chemically distinct compounds, including alkanes, alkenes, aromatics, alcohols, aldehydes, ketones, terpenes, and other reactive hydrocarbons relevant to ozone and secondary organic aerosol formation processes. Table S2 in Supplementary Information (Supplementary Materials) presents the 25 GEIA NMVOC species groups along with their representative molecular formulas. These speciation profiles were applied to disaggregate NMVOCs into species compatible with the CB6R5 chemical mechanism implemented in CMAQ. For INEMA, speciation profiles were assigned according to the underlying native activity rather than solely according to the harmonized sector label used for inventory comparison. Commercial Energy Consumption, although mapped under FUE in the harmonized sector structure, was speciated using the profile corresponding to Small Combustion. Agricultural Waste Burning and Forest Fires, although mapped under WAS, were speciated using the corresponding Agricultural Waste Burning and Large-Scale Biomass Burning profiles, respectively.

CMAQ Processing

The CMAQ modeling system version 5.5.0.3 was utilized to simulate the emission, transport, chemical transformation, and deposition of atmospheric pollutants over Santiago, Chile, during April 2025. CMAQ was configured using the *cb6r5_ae7_aq* chemical mechanism and aerosol module combination, enabling detailed representation of gas-phase photochemistry, secondary aerosol formation, heterogeneous chemistry, and cloud-related atmospheric processes. The model was applied over the three nested domains with horizontal resolutions of 9 km, 3 km, and 1 km, consistent with the WRF and SMOKE configurations. Supplementary Information

1090 (Supplementary Materials) presents the CMAQ Chemistry-Transport Model (CCTM)
1091 configurations in detail.

1092 Gridded uniform initial chemical conditions for CMAQ simulations on the 9-, 3-, and 1-km
1093 domains were generated using the standard vertically resolved concentration profiles
1094 distributed within the CMAQ GitHub repository provided by the U.S. Environmental
1095 Protection Agency:
1096 “*avprofile_cb6r3m_ae7_kmtbr_hemi2016_v53beta2_m3dry_col051_row068.csv*”³³. These
1097 files represent annual average concentrations at a grid cell over the Pacific derived from a
1098 simulation with the hemispheric version of CMAQ v5.3 beta2 for the year 2016. As such, these
1099 concentration profiles are reflective of conditions in a remote marine environment. The ICON
1100 (initial condition program of CMAQ) processor interpolates these predefined concentration
1101 profiles onto the user-defined CMAQ domain grid.

1102 A 48-h spin-up period was applied before the April analysis period. Initial chemical conditions
1103 were generated from vertically resolved CMAQ reference profiles and were allowed to adjust
1104 during the 48-h spin-up period. The generated ICON files were subsequently used as the initial
1105 chemical state for all CMAQ simulations, including the base, ISAM, and DDM/HDDM
1106 configurations.

1107 The gridded spatio-temporal chemical boundary conditions for CMAQ were generated using
1108 atmospheric composition fields derived from the NASA Global Earth Observing System
1109 Composition Forecast (GEOS-CF) system³⁴. The GEOS-CF dataset provides global
1110 reanalysis/assimilation-constrained three-dimensional atmospheric composition by integrating
1111 the GEOS Earth system model with detailed atmospheric composition simulations, thereby
1112 offering dynamically consistent global distributions of gaseous and aerosol species suitable for
1113 regional chemical transport modeling applications. Compared with the standard default CMAQ
1114 boundary profiles, GEOS-CF-based boundary conditions provide improved representation of
1115 large-scale background atmospheric composition, long-range pollutant transport, and upper-
1116 atmospheric chemical variability.

1117 The GEOS-CF boundary condition generation procedure followed the CMAQ-ready
1118 preprocessing workflow adapted from the publicly available *aqmbc* framework³⁵. Hourly
1119 GEOS-CF atmospheric composition files were initially downloaded and processed to generate
1120 CMAQ-compatible lateral boundary concentration fields consistent with the *cb6r5_ae7*
1121 chemical mechanism configuration employed in this study. The preprocessing workflow
1122 included species mapping, vertical interpolation, temporal interpolation, and conversion into
1123 CMAQ-ready BCON (boundary condition program of CMAQ) structures.

1124 The GEOS-CF files contained 16 of the 24-hour fields required for each day. The eight missing
1125 hourly fields were reconstructed by linear interpolation between adjacent available times.
1126 Following the correction of the temporal interpolation procedure and reconstruction of the
1127 missing hourly fields, monthly CMAQ-compatible BCON files were generated and
1128 subsequently applied to all simulations at the parent domain (9 km) boundary.

The resulting CMAQ-compatible BCON files provided temporally dynamic chemical boundary conditions for all gaseous and aerosol species throughout the simulation period (1 month). Compared with the use of static climatological boundary conditions, the GEOS-CF-driven BCON configuration allows improved representation of regional background atmospheric composition, hemispheric pollutant transport, and synoptic-scale variability influencing air quality over Santiago, Chile. The BCON files for the corresponding 3-km domain were extracted from the completed 9-km CMAQ simulation, and the 1-km BCON files were subsequently extracted from the completed 3-km simulation.

Total column O₃ data used within the CMAQ preprocessing framework was obtained from the TOMS/OMI (Total Ozone Mapping Spectrometer/Ozone Monitoring Instrument) ftp portal³⁶. A continuous 365-day dataset spanning from 1 July 2024 to 1 July 2025 was employed to provide temporally consistent total O₃ column fields for the required CMAQ OMI file preparation.

CMAQ ISAM

CMAQ-ISAM apportions modeled pollutant concentrations by tracking source-tagged contributions through transport, deposition, and chemical transformation. In gas-phase chemistry, source attribution is propagated from reactants to products using integrated reaction rates, product yields, and prescribed reactant-weighting rules. This provides source contributions within the simulated atmosphere without requiring separate simulations that remove individual emission sources.

For O₃ and its precursors, the dynamic attribution option was used¹⁴. This option determines the instantaneous O₃ formation regime using the ratio $R=P(H_2O_2)/P(HNO_3)$, where P denotes the gross chemical production rate. When $R>0.35$, the NO_x-limited attribution case transfers source identity preferentially through reactive nitrogen species; when $R\leq 0.35$, the VOC-limited attribution case transfers source identity through reactive VOCs and related organic species.

In this study, ISAM was applied to the calibrated EDGAR- and INEMA-based configurations over Santiago's 1-km domain during April 2025, using eight sector-specific emission tags and identical boundary and initial conditions. The resulting source fractions describe the partitioning of modeled concentrations within the respective calibrated configurations. Unlike DDM/HDDM sensitivities, ISAM contributions are not derivatives with respect to emissions and should not be interpreted directly as concentration changes following source reduction or removal. This distinction is particularly important for secondary pollutants, whose nonlinear formation makes the response to emission changes different from attribution within an unchanged emission scenario.

CMAQ DDM/HDDM

DDM/HDDM sensitivities are local derivatives evaluated about the model state. Their use to estimate the concentration response to a finite emission change therefore relies on a truncated Taylor expansion, $C(\epsilon_0+\Delta\epsilon)\approx C_0+\Delta\epsilon S^{(1)}+\frac{1}{2}\Delta\epsilon^2 S^{(2)}$, where the neglected third- and higher-order terms become increasingly important as the perturbation moves farther from the baseline state. In this expansion, $C_0=C(\epsilon_0)$ is the modeled concentration at the calibrated baseline emission state, ϵ_0 is the baseline value of the emission perturbation parameter, $\Delta\epsilon$ is its

perturbation, and $S^{(1)}$ and $S^{(2)}$ are the first- and second-order concentration sensitivities evaluated at ε_0 , respectively. First-order DDM has generally been shown to represent concentration responses well for relatively modest emission perturbations, whereas inclusion of second-order HDDM extends the useful projection range by representing curvature in the response surface^{16,37-39}. This limitation is important for the base (uncalibrated) simulations in the present study. The discrepancies in several of the base simulations require changes comparable to, or substantially larger than, the perturbation range over which a low-order Taylor expansion can reasonably be expected to approximate the full nonlinear CMAQ response. Moreover, because the sensitivity coefficients are evaluated about the biased base atmosphere, they describe the local chemical state of that simulation, including its oxidant levels, gas-particle partitioning and precursor regime, rather than the atmospheric state consistent with the observations.

Model Performance Statistics

CMAQ model performance was quantified from paired hourly observations and simulations using the number of valid pairs (N), Pearson correlation coefficient (R), root-mean-square error (RMSE), mean bias (MB), NMB (Normalized Mean Bias), and NME (Normalized Mean Error). The statistical metrics R, RMSE, MB, NMB, and NME were calculated using Equations 1–5, respectively.

$$R = \frac{\sum_{i=1}^N (M_i - \bar{M})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^N (M_i - \bar{M})^2} \sqrt{\sum_{i=1}^N (O_i - \bar{O})^2}} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (M_i - O_i)^2} \quad (2)$$

$$MB = \frac{1}{N} \sum_{i=1}^N (M_i - O_i) \quad (3)$$

$$NMB = 100 \frac{\sum_{i=1}^N (M_i - O_i)}{\sum_{i=1}^N O_i} \quad (4)$$

$$NME = 100 \frac{\sum_{i=1}^N |M_i - O_i|}{\sum_{i=1}^N O_i} \quad (5)$$

where M_i and O_i are the modeled and observed concentrations for the i -th paired hourly value, $\bar{M}$ and $\bar{O}$ are their respective means, and N is the number of valid model–observation pairs.

CMAQ DESID Scaling

Emission adjustment was implemented using the DESID capability within CMAQ⁴⁰. The calibration was pollutant specific rather than sector-specific; that is, a single scaling coefficient for a given pollutant was applied uniformly to that pollutant across all contributing emission sectors. Individual sectors were therefore not independently increased or reduced during the calibration procedure.

Rather than applying a predefined symmetric perturbation matrix, the scaling coefficients were determined using an iterative forward-model bracketing approach. Starting from the unscaled inventory, CMAQ concentrations were compared with the corresponding observations. For NO, NO₂, CO, PM_{2.5}, and PM₁₀, the observational calibration target for pollutant p was defined as the arithmetic mean of the April mean concentrations at Las Condes (LC), Parque O'Higgins (PO), and Pudahuel (PU), as presented in Equation 6:

$$O_p = \frac{1}{3} (\bar{O}_{p,LC} + \bar{O}_{p,PO} + \bar{O}_{p,PU}) \quad (6)$$

For these pollutants, the same three-site averaging was applied to the corresponding modeled April mean concentrations, as presented in Equation 7:

$$M_p(SF) = \frac{1}{3} (\bar{M}_{p,LC}(SF) + \bar{M}_{p,PO}(SF) + \bar{M}_{p,PU}(SF)) \quad (7)$$

SO₂ observations were available only at La Florida, and the measured concentrations were used as representative values for Santiago in this study. Thus, the April mean modeled concentrations at LC, PO, and PU were compared with the April mean La Florida observation used as the Santiago reference. For NMVOC adjustment, the corresponding observed O₃ concentrations at LC, PO, and PU were used as the calibration target rather than direct NMVOC measurements.

For pollutants exhibiting positive model bias, the relevant DESID Scaling Factor (SF) was progressively reduced in successive simulations; conversely, where an increase in the precursor emissions was required to improve model–observation agreement, the factor was progressively increased. Successive simulations were conducted until two neighboring scaling factors produced modeled responses that bracketed the observational target, with one case lying above (M_2) and the other below (M_1) the observational target (O_p). The final scaling coefficient was then estimated by linear interpolation between the two bounding simulations as presented in Equations 8 and 9:

$$SF_{final} = SF_1 + (SF_2 - SF_1) \frac{O_p - M_1}{M_2 - M_1}, \quad (8)$$

$$M_1 = M_p(SF_1), M_2 = M_p(SF_2) \quad (9)$$

where SF_1 and SF_2 are the two bounding DESID scaling factors. The interpolated coefficient was subsequently applied in an additional CMAQ simulation to assess the resulting model–observation agreement. This procedure was repeated independently for the selected pollutants and for the EDGAR and INEMA base model configurations. The procedure follows the general concept of iterative forward-model emission adjustment, in which emission scaling factors are progressively modified according to model–observation discrepancies and the atmospheric model is rerun until the observational target is bracketed or an acceptable level of agreement is reached^{41,42}.

Declarations

Data Availability

The datasets generated and analyzed during the current study are not publicly available due to the large data volume and substantial storage requirements associated with the high-resolution WRF–SMOKE–CMAQ simulations, but are available from the corresponding author upon reasonable request.

Code Availability

Custom Python scripts were developed and used for observational data processing, emission inventory preprocessing, spatial allocation, quality control, and preparation of model-ready inputs for the SMOKE–CMAQ modeling system. These scripts are not publicly available due to intellectual property considerations associated with ETH Zürich and the authors, but are available from the corresponding author upon reasonable request.

Acknowledgements

The first author gratefully acknowledges the use of the High-Performance Computing (HPC) infrastructure provided by the Air Quality and Particle Technology Group at ETH Zurich. We thank Dr. Monica Crippa for helpful correspondence and clarification regarding the EDGAR emission inventory.

Funding

This study was conducted as part of the MASSEV Project (*Monitoring, Mapping and Modelling Air Quality for Sustainable Environments in Cities and Communities*) with Swiss National Science Foundation grant no. 40CF40_221599. The funder had no role in the study design, data collection, analysis or interpretation of the results, preparation of the manuscript, or the decision to submit the work for publication.

Author Contributions

M.P. developed the methodology and software workflows, configured and executed the modelling system, curated and processed the data, performed the formal analyses, prepared the visualizations, interpreted the results, and wrote the original draft of the manuscript. J.W. conceptualized the study, acquired the funding, and administered the project. Both authors reviewed and edited the manuscript and approved the final version.

Competing Interests

The authors declare no competing financial or non-financial interests.

References

1. Pino-Cortés, E., et al. An evaluation of the photochemical air quality modeling using CMAQ in the industrial area of Quintero-Puchuncaví-Concón, Chile. *Atmos. Pollut. Res.* 13, 101336 (2022).
2. Nawaz, M. O. et al., Sources of air pollution health impacts and co-benefits of carbon neutrality in Santiago, Chile. *J. Geophys. Res. Atmos.* 128, e2023JD038808 (2023).
3. Cheng, X., Hao, Z., Zang, Z., Liu, Z., Xu, X., Wang, S., et al. A new inverse modeling approach for emission sources based on the DDM-3D and 3DVAR techniques: an application to air quality forecasts in the Beijing–Tianjin–Hebei region. *Atmos. Chem. Phys.* 21, 13747–13761 (2021).

4. Mousavinezhad, S., Choi, Y., Khorshidian, N., Ghahremanloo, M. & Momeni, M. Air quality and health co-benefits of vehicle electrification and emission controls in the most populated United States urban hubs: insights from New York, Los Angeles, Chicago, and Houston. *Sci. Total Environ.* 912, 169577 (2024).
5. Osses, M., Rojas, N., Ibarra, C., Valdebenito, V., Laengle, I., Pantoja, N., et al. High-resolution spatial-distribution maps of road transport exhaust emissions in Chile, 1990–2020. *Earth Syst. Sci. Data* 14, 1359–1376 (2022).
6. Crippa, M., Guizzardi, D., Pagani, F., Schiavina, M., Melchiorri, M., Pisoni, E., et al. Insights into the spatial distribution of global, national, and subnational greenhouse gas emissions in the Emissions Database for Global Atmospheric Research (EDGAR v8.0). *Earth Syst. Sci. Data* 16, 2811–2830 (2024).
7. Crippa, M., Guizzardi, D., Pagani, F., Banja, M., Muntean, M., Schaaf, E., et al. EDGAR v8.1 Global Air Pollutant Emissions [Dataset]. European Commission, Joint Research Centre (2026).
8. Castesana, P., Diaz Resquin, M., Huneus, N., Puliafito, E., Darras, S., Gómez, D., et al. PAPILA dataset: a regional emission inventory of reactive gases for South America based on the combination of local and global information. *Earth Syst. Sci. Data* 14, 271–293 (2022).
9. Álamos, N., Huneus, N., Opazo, M., Osses, M., Puja, S., Pantoja, N., et al. High-resolution inventory of atmospheric emissions from transport, industrial, energy, mining and residential activities in Chile. *Earth Syst. Sci. Data* 14, 361–379 (2022).
10. Chen, M., Zhang, W., Han, W., Xia, J. & Jin, J. Global ammonia emission inversion in 2022 via assimilating IASI observations. *npj Clean Air* 2, 35 (2026).
11. Byun, D. & Schere, K. L. Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system. *Appl. Mech. Rev.* 59, 51–77 (2006).
12. Appel, K. W., Bash, J. O., Fahey, K. M., Foley, K. M., Gilliam, R. C., Hogrefe, C., et al. The Community Multiscale Air Quality (CMAQ) model versions 5.3 and 5.3.1: system updates and evaluation. *Geosci. Model Dev.* 14, 2867–2897 (2021).
13. Kwok, R. H. F., Napelenok, S. L. & Baker, K. R. Implementation and evaluation of PM_{2.5} source contribution analysis in a photochemical model. *Atmos. Environ.* 80, 398–407 (2013).
14. Shu, Q., et al. Comparison of ozone formation attribution techniques in the northeastern United States. *Geosci. Model Dev.* 16, 2303–2322 (2023).
15. Xian, Y., et al. Source apportionment and formation of warm season ozone pollution in Chengdu based on CMAQ-ISAM. *Urban Clim.* 56, 102017 (2024).
16. Cohan, D. S. & Napelenok, S. L. Air quality response modeling for decision support. *Atmosphere* 2, 407–425 (2011).
17. Liu, C. & Shi, K. A review on methodology in O₃-NO_x-VOC sensitivity study. *Environ. Pollut.* 291, 118249 (2021).
18. Wang, Y., Yaluk, E. A., Chen, H., Jiang, S., Huang, L., Zhu, A., et al. The importance of NO_x control for peak ozone mitigation based on a sensitivity study using CMAQ-HDDM-3D model during a typical episode over the Yangtze River Delta region, China. *J. Geophys. Res. Atmos.* 127, e2022JD036555 (2022).
19. Yaluk, E. A., Wang, Y., Jiang, S., Huang, L., Lu, G., Zhu, A., et al. Changes in first- and second-order sensitivities of ozone concentration to its precursors over the Yangtze River Delta region of China due to COVID-19 lockdown: insights from CMAQ-HDDM modeling study. *Atmos. Environ.* 309, 119931 (2023).
20. Emery, C., et al. Recommendations on statistics and benchmarks to assess photochemical model performance. *J. Air Waste Manag. Assoc.* 67, 582–598 (2017).
21. Chen, G., Wang, Y., Tao, C., Zhang, Z., Zhou, M., Yan, R., et al. Nitrate-driven extreme winter PM_{2.5} pollution in Shanghai, China. *npj Clean Air* 1, 28 (2025).
22. Lapere, R., Menut, L., Mailler, S. & Huneus, N. Seasonal variation in atmospheric pollutants transport in central Chile: dynamics and consequences. *Atmos. Chem. Phys.* 21, 6431–6454 (2021).
23. Qu, K., Wang, X., Yan, Y., Jin, X., Cai, X., Shen, J., et al. Cross-regional PM_{2.5} sulfate transport to the Pearl River Delta: dynamics and chemistry. *npj Clean Air* 2, 14 (2026).

24. ChileAtiende. Programa de Recambio de Calefactores (2026).
25. Boso, À., Oltra, C. & Hofflinger, Á. Participation in a programme for assisted replacement of wood-burning stoves in Chile: the role of sociodemographic factors, evaluation of air quality and risk perception. *Energy Policy* 129, 1220–1226 (2019).
26. Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., et al. A Description of the Advanced Research WRF Model Version 4. NCAR Tech. Note NCAR/TN-556+STR (2019).
27. Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* 146, 1999–2049 (2020).
28. MMM, Laboratory. Geographical Data (Static Data) for the WRF Model. National Center for Atmospheric Research, https://www2.mmm.ucar.edu/wrf/site/access_code/geog_data.html, (2026).
29. Richter, J. WRF Domain Wizard [Computer software]. GitHub, <https://github.com/JiriRichter/WRFDomainWizard>, (2026).
30. Crippa, M., Solazzo, E., Huang, G., Guizzardi, D., Koffi, E., Muntean, M., et al. High resolution temporal profiles in the Emissions Database for Global Atmospheric Research. *Sci. Data* 7, 121 (2020).
31. Crippa, M., Guizzardi, D., Butler, T., Keating, T., Wu, R., Kaminski, J., et al. The HTAP_v3 emission mosaic: merging regional and global monthly emissions (2000–2018) to support air quality modelling and policies. *Earth Syst. Sci. Data* 15, 2667–2694 (2023).
32. Huang, G., Brook, R., Crippa, M., Janssens-Maenhout, G., Schieberle, C., Dore, C., et al. Speciation of anthropogenic emissions of non-methane volatile organic compounds: a global gridded data set for 1970–2012. *Atmos. Chem. Phys.* 17, 7683–7701 (2017).
33. CMAQ. CMAQ User’s Guide, Ch. 4: Model Inputs — Initial Conditions Processor (ICON). GitHub (2026).
34. Keller, C. A., Knowland, K. E., Duncan, B. N., Liu, J., Anderson, D. C., Das, S., et al. Description of the NASA GEOS composition forecast modeling system GEOS-CF v1.0. *J. Adv. Model. Earth Syst.* 13, e2020MS002413 (2021).
35. Barron, H. aqmbc: Air Quality Model Boundary Condition Processor (v1.x) [Computer software]. GitHub, <https://github.com/barronh/aqmbc>, (2026).
36. NASA, Goddard Space Flight Center, OMI Level 3e Ozone Data [Dataset], Atmospheric Chemistry and Dynamics Laboratory, <https://acd-ext.gsfc.nasa.gov/anonftp/toms/omi/data/Level3e/ozone/>, (2026).
37. Dunker, A. M., Yarwood, G., Ortmann, J. P. & Wilson, G. M. The decoupled direct method for sensitivity analysis in a three-dimensional air quality model: implementation, accuracy, and efficiency. *Environ. Sci. Technol.* 36, 2965–2976 (2002).
38. Hakami, A., Odman, M. T. & Russell, A. G. High-order, direct sensitivity analysis of multidimensional air quality models. *Environ. Sci. Technol.* 37, 2442–2452 (2003).
39. Hakami, A., Odman, M. T. & Russell, A. G. Nonlinearity in atmospheric response: a direct sensitivity analysis approach. *J. Geophys. Res. Atmos.* 109, D15303 (2004).
40. Murphy, B. N., Nolte, C. G., Sidi, F., Bash, J. O., Appel, K. W., Jang, C., et al. The detailed emissions scaling, isolation, and diagnostic (DESID) module in the Community Multiscale Air Quality (CMAQ) modeling system version 5.3.2. *Geosci. Model Dev.* 14, 3407–3420 (2021).
41. Mendoza-Dominguez, A. & Russell, A. G. Iterative inverse modeling and direct sensitivity analysis of a photochemical air quality model. *Environ. Sci. Technol.* 34, 4974–4981 (2000).
42. Mendoza-Dominguez, A. & Russell, A. G. Emission strength validation using four-dimensional data assimilation: application to primary aerosol and precursors to ozone and secondary aerosol. *J. Air Waste Manag. Assoc.* 51, 1538–1550 (2001).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [PourkiaeiSlv.R2c3.pdf](#)