PAPILA V2.0

Introduction

This work presents an actualization of PAPILA's inventory, the first AEI from anthropogenic sources covering the continental SA region, which combines local available information with a global database in a proper and rigorous way. For this purpose, global datasets were used as a basis, enriching it with locally developed inventories available in the literature until 2023 and selecting those with national coverage and with the availability of data for the period and species of interest. Since national territories are the common administrative entities that can be exchanged with the global inventory, the local information on emissions of cities from countries such as Colombia, Peru and Brazil was not included in this Inventory.

The main updates regarding the PAPILA 1 version are listed below:

1. **Adding new local inventories:**
    - High-resolution emission inventories from Mexico and Colombia are included in addition to those already considered for Chile and Argentina

    - National greenhouse inventories that include estimation for reactive gases are considered for Chile, Colombia and Ecuador

2. **Temporality**: the inventory was updated from the period 2014-2016 to 2014-2020
3. **Considered compounds:** in addition to reactive gases considered in the first version, CH_4 and CO_2 greenhouse gases are considered, and also particulate matter (PM10 and PM2.5) and its speciation in black carbon and organic carbon.
4. **Sectors**: emission from open biomass burning sector are included in addition to the sectors: Agriculture livestock, Agriculture soils; Fugitives, Industry, Refineries, Solvents, Inland Shipping, Non-road transportation, Power generation, Residential and other sectors; Road transportation, Ships and Solid waste disposal
5. **Base inventories considered:** CAMS-GLOB-ANT v6.1inventory was complemented with EDGAR v6.1 for particulate matter emissions and GFAS for open biomass burning emissions

Methodology and Data

The dataset presented in this work, hereinafter called PAPILA, focuses on emissions from: dioxide carbon (CO_2) and methane (CH_4) greenhouse gases; the group of species known as reactive gases, given their relevance in atmospheric chemistry as precursors of O_3 and PM2.5: carbon monoxide (CO), nitrogen oxides (NO_x), non-methane volatile organic compounds (NMVOCs), ammonia (NH_3), and sulfur dioxide (SO_2): and emission for particulate matter (PM10 and PM25) organic carbon, and black carbon (BC). The inventories are presented as netCDF4 files, one for each species gridded with a spatial resolution of $0.1^\circ \times 0.1^\circ$ covering the domain 32–120° W and 34° N–58° S. Each file contains 14 variables corresponding to the emissions in Tg yr^{-1} from the following categories, which are organized and denominated inspired mostly on the nomenclature given by CAMS: thermal power plants (ENE); road transportation (TRO); non-road transportation (TNR); fugitive emissions (FEF); industries, including fuel consumption in manufacturing industries and construction industrial processes, (IND); solvents (SLV); refineries (REF); agricultural soils (AGS); agricultural livestock (AGL); domestic and international navigation (SHP); solid waste disposal (including solid waste, wastewater, and incineration) (SWD); open biomass burning (OBB); residential , commercial and other sectors (RCO) and the sum of all categories (SUM) (table 1).

PAPILA description	PAPILA code	IPCC 2006 code	CAMS v6.1 code
Energy production	ENE	1A1a	ENE
industries, including fuel consumption in manufacturing industries and construction industrial processes	IND	1A2+2A+2B+2C+2H	IND
Reffineries	REF	1A1b+1A1ci+1A1cii+1A5biii+1B1b+1B2aiii6+1B2b	REF
Fugitive emissions from fuels	FEF	1B1A + 1B2 + 2D1+2D2+2D4	FEF
Road transport	TRO	1A3b	TRO
Non-road transport	TNR	1A3c_Rail	TNR
Shiping	SHP	1A3d	SHP
Residential commercial and others	RCO	1A4	COM + DOM
Solvents	SLV	2D3+2E+2F+2G	SLV
Agricultural soils	AGS	3C2+3C3+3C4+3C5+3C6+3C7	AGS
Agricultural livestock	AGL	3A	AGL
Open biomass burning	OBB	3C1	OBB
Solid Waste Disposal	SWD	4	SWD

Table 1: PAPILA emission categories.

Due to the availability of data in the local AEIs (**figure 1**) and the completeness of the sectors represented, the 2014–2020 period was selected for this update of the PAPILA dataset. As a basis, CAMS-GLOB-ANT v6.1 (Granier et al., 2019) inventory was used complemented with EDGAR v6.1 for particulate matter emissions and GFASv1.3 for open biomass burning emissions. The local emission inventories used for the mosaic include high-resolution local information from the continental areas of Argentina (Puliafito et al., 2021), Chile (Alamos et al., 2022), Mexico (Garcia, 2018), and Colombia (Rojas et al., 2023). In addition, information on National greenhouse inventories that includes completely, or partially reactive gasses and particles was considered for Chile, Ecuador and Colombia.

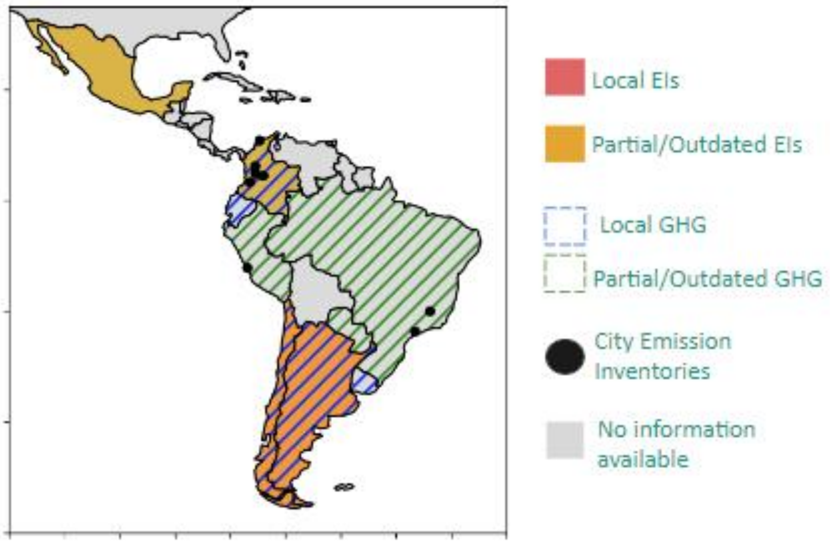

Figure 1: Summary of available information of local emission inventories in the South American region

In the process of creating the mosaic from each local inventory, only sectors that had emission estimates for all the specified species, throughout the entire study period, and covering the entire national territory were taken into consideration. In some cases, to meet this condition without compromising consistency, estimations for certain years and species were interpolated.

This condition represents an improvement over the first version of PAPILA's inventory, where categories without estimates for all species were considered, leading to inconsistencies in the database. These inconsistencies arose from different inventories and activity data being used in the estimation of different species within the same category for a particular country. For those categories with an absence of local data or with incomplete information in terms of species or years, the base inventory was used to fill the gaps.

The approach of combining the local emission inventories with the base inventory involved employing different methodologies based on the quality and spatial resolution of the datasets.

1. In cases where national greenhouse gas inventories were available, the emissions from the base inventory for each year and sector were scaled up to match the magnitudes reported in the local inventory. This scaling process allowed us to assign the spatial distribution of CAMS data to the total national emissions as reported in the GHG inventories.
2. For countries and sectors with high-resolution local inventory data available, the emissions from these inventories were directly incorporated into the mosaic, effectively replacing the emissions from the base data.

1.1. Base Inventory

The primary inventory used as a basis for this study was CAMS v6.1, which provided emissions data for CO2, CH4, CO, NOX, NMVOC, SO2, NH3, BC, and OC for the period 2014-2020. To incorporate the emissions of PM2.5 and PM10, we referred to the EDGAR v6.1 database. As CAMS is built upon the emissions data from EDGAR, there are no inconsistencies in combining the two inventories for the same sectors. However, it's important to note that the EDGAR database is only available until the year 2018. Therefore, we estimated EDGAR’s PM25 and PM10 emissions for years 2019 and 2020 by applying CAMS trajectories of the sum of BC and OC by sector and country. Before proceeding, we evaluated whether the relationship between OC, BC, and PM2.5 in the EDGAR database remained constant during the period 2014-2018 for each sector and country. The results were found to be satisfactory (see Figure 2).

It's important to note that emissions from the commercial (com) and residential (dom) sectors of CAMS were grouped into residential commercial and other sectors (RCO), to be consistent with the sectors provided in EDGAR. the sector RCO consists of sectors 1A4 and 1A5 of IPCC, which includes emission from energy uses in agriculture, forestry and livestock.

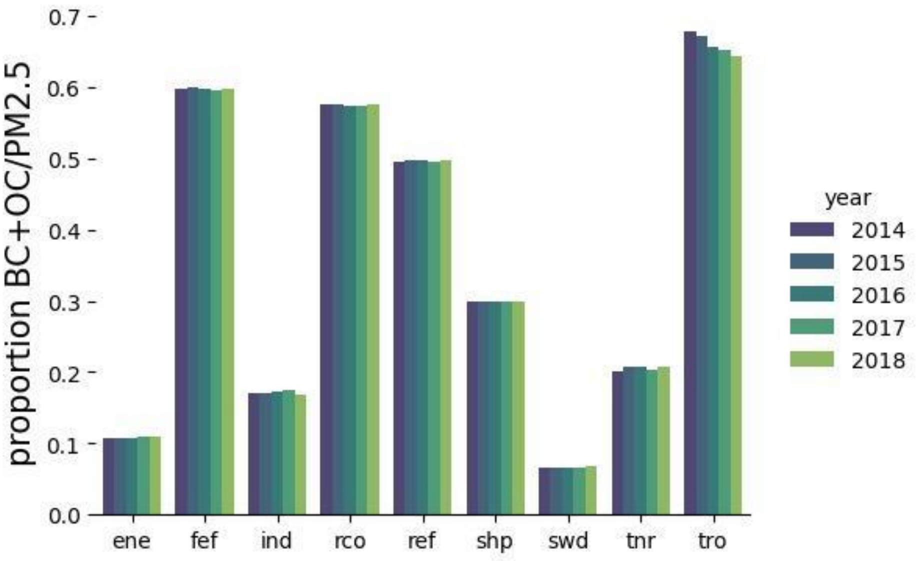

Figure 2: Proportion of BC + OC over PM2.5 for the whole domain in the period 2014-2018

For the Agriculture Waste Burning category, we did not consider the emissions from CAMS and EDGAR. Instead, we included the more comprehensive database of GFAS V1.3, which accounts for emissions from all open biomass burnings, including wildfires and agriculture waste burning. However, GFAS does not cover PM10. Consequently, the emissions estimation for the OBB sector lacks these specific species.
It is worth noting that the original EDGAR and GFAS databases had longitudes in the range of 0-360, which we rescaled to the range of -180 to 180 to ensure coherence with the CAMS grid. Additionally, we recalculated GFAS emission units from kg/m2/year to Tg/year.

Finally, we set the domain for our study to be within 32°–120° W and 34° N–58° S by applying a rectangular mask to the global database. Each grid pixel in the database was assigned to a specific country, aligning it with the GPW.v3 population product, which also utilizes the same grid as the other databases.

1.2. General Considerations for the inclusion of GHG inventories

National greenhouse gas (GHG) inventories, including estimates of reactive gases, are available for Chile, Colombia, Ecuador, Uruguay, Brazil, and Argentina. Brazil was excluded from the study due to significant incompleteness in species data for various sectors, and the inventory period was only extended until 2016. in the case of Uruguay, the database is grouped into IPCC categories bigger than CAMS (for example category 1A3 includes road transport, non-road transport, shipping, and aviation), which makes complicated to make an adequate match between the sectors of the GHG inventory and the PAPILA's categories.

Subsequently, we focused on GHG inventories for Chile, Colombia, and Ecuador. These inventories differed in terms of years, species, and sectors considered. To facilitate the integration, we first homogenized the databases to include information for all species and years present in the PAPILA inventory. Furthermore, we recategorized the IPCC sectors of each database to align with the sectors used in the PAPILA inventory.

For Ecuador and Colombia, which had emission estimates available until 2018, we estimated emissions for the years 2019 and 2020 by interpolating the CAMS trajectories for each species and sector-specific to each country and matching them with the GHG inventory data.

Notably, only the Colombian GHG inventory provided estimates for PM2.5, PM10, and BC emissions. To complement this data, we estimated emissions of these three species along with OC by utilizing the relationship between CO and these species in EDGAR for each sector and country. We rigorously assessed the consistency of this approach by examining the temporal stability of the relationship. In the case of Colombia, OC emissions were directly obtained from the relationship between PM2.5 and OC by sector and country, as derived from EDGAR.

Finally, none of these inventories included NH3 estimates, which cannot be derived from relationships with other species, such as particles. Consequently, NH3 emissions from CAMS were retained in this instance, recognizing the potential risk of introducing inconsistency in the utilization of different activity levels among species for the same sectors.

1.3. General Considerations for the inclusion of High-resolution local inventories

High-resolution emission inventories are available for several South American countries, including Chile, Argentina, Mexico, and Colombia, with the exception that emissions data for Colombia is limited to the transport sector. However, these inventories vary regarding years covered, species included, and sectors considered. Therefore, to facilitate a comprehensive integration, it becomes necessary to harmonize the databases, ensuring the incorporation of information for all species and years present in the PAPILA inventory. Additionally, the sectors in each database were reclassified to align with the sectors utilized in the PAPILA inventory.

None of the inventories provide emission estimates for OC species except for Argentina which has limited information restricted to specific sectors. To address this gap, we employed a methodology where OC estimates were derived by applying the OC-to-PM2.5 ratio from the EDGAR dataset for each sector, year, and country, based on the PM2.5 emissions available in each respective inventory. In instances where data for BC or PM10 were missing for certain sectors, a similar approach was adopted.

None of the databases provide complete sector coverage for the entire study period. Specifically, the Mexican inventory only pertains to the year 2016, while the Chilean inventory spans from 2015 to 2020. In the case of Argentina's inventory, there is a lack of information for the OBB sector in the year 2020.

Following a methodology similar to the one used for GHG inventories, missing years in the Chilean and Mexican datasets were estimated through interpolation, employing CAMS trajectories specific to each country, species, and sector. However, estimating emissions for the OBB sector in Argentina for 2020 proved to be more complex due to the spatial variation in fire distribution each year. Consequently, the emissions for the OBB sector in 2020 were estimated by scaling up GFAS 2020 emissions to match the level of the GEAAs inventory.

Grid Homogenization

Combining different local inventories into a common regional database poses the challenge of bringing them onto a single, uniform, and consistent grid. To achieve this, conflicts arising from overlapping cells shared by multiple countries or coastal areas needed to be resolved. To align with the base inventory used in this study, emissions from high-resolution inventories of Chile, Mexico, Argentina, and Colombia were transformed from their original grids to the CAMS 0.1-degree grid.

To accomplish this, each original grid was overlaid with the CAMS grid, and a matrix representing the proportion of each original grid cell corresponding to the CAMS grid cells was computed. This matrix was then used to estimate the emissions of each CAMS grid cell by taking a weighted average of all the original grid cells it covers. In this way, the emissions of each inventory are regridded to a harmonized grid in the same terms of CAMS.

1.4. Methods and Data by Country

1.4.1. Chile

In this study, the annual emissions for Chile were derived from two primary sources: the High-resolution INEMA inventory, available for the years 2015-2020 at a spatial resolution of 0.01 degree, covering sectors IND, ENE, REF, TRO, RCO, and OBB, and the GHG national inventory, which provided data for sectors AGS, FEF, and TNR. Emissions from SLV, AGL, and SHP were maintained at the same levels as those in the base data (for a detailed picture of the Chilean mosaic see Table 2).

However, certain sectors in the Chilean inventories did not have emissions data available for all species. To address this data gap, particle emissions were interpolated from the corresponding sectors of the GHG inventories. Additionally, for sectors IND, REF, and ENE, emissions of BC and OC were interpolated based on PM2.5 data. Furthermore, within these three sectors, NMVOC and CH4 emissions weren't disaggregated from total VOC emissions. To estimate emissions for CH4 and NMVOC within these sectors, proportions of these two species in CAMS data were considered for each sector and year.

Symbology	Base inventory			High-resolution INEMA				Interpolated emissions		GHG Inventory	
	GHG			REACTIVE GASES					PARTICLES			
Specie / Sector	CO2	CH4	N2O	CO	NOx	NMVOCs	NH3	SO2	PM10	PM2.5	BC	OC
ENE	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	Base inventory	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	Interpolated
RCO	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	Interpolated
TRO	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	Interpolated	High-resolution INEMA	High-resolution INEMA	Interpolated
TNR	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	Base inventory	GHG Inventory	Interpolated	Interpolated	Interpolated	Interpolated
FEF	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	Base inventory	Interpolated	Interpolated	Interpolated	Interpolated	Interpolated
IND	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	Base inventory	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	Interpolated
OBB	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA		High-resolution INEMA	High-resolution INEMA	Interpolated
AGS	GHG Inventory	GHG Inventory	GHG Inventory		GHG Inventory	GHG Inventory	Base inventory		Interpolated	Interpolated		Interpolated
AGL		Base inventory	Base inventory		Base inventory	Base inventory	Base inventory		Base inventory	Base inventory		Interpolated
SHP	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory
SWD	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory		Base inventory	Base inventory	Base inventory	Base inventory
REF	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	Base inventory	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	High-resolution INEMA	Interpolated
SLV	Base inventory					Base inventory	Base inventory					

Table 2: Database considered to estimate emissions for Chile for each sector and species

To account for the absence of emissions data in the INEMA inventory for the year 2014, interpolation was performed using CAMS emission trajectories specific to each sector and species for Chile.

Finally, Due to temporal inconsistencies in the magnitude of emissions associated with stationary sources in the REF, ENE, and IND sectors, a decision was made to scale the national total emissions by year, species, and sector to match the values estimated in the GHG national inventory while preserving the spatial distribution from INEMA. As a result, these sectors lack NH3 emissions data, owing to the absence of such information in the GHG inventory. NH3 emissions of the base inventory were used for these sectors.

1.4.2. Argentina.

Spatially disaggregated emissions for all the species and categories included in this work except SWD and SLV are available for Argentina in the GEAA inventory (table 3). Which consists of a high-resolution ($0.025^{\circ} \times 0.025^{\circ}$) inventory of 1990-2020 monthly emissions.

Symbology	Base inventory			High-resolution GEAA			Interpolated emissions					
	GHG			REACTIVE GASES					PARTICLES			
Specie / Sector	CO2	CH4	N2O	CO	NOx	NMVOCs	NH3	SO2	PM10	PM2.5	BC	OC
ENE												
RCO												
TRO												
TNR												
FEF												
IND												
OBB												
AGS												
AGL												
SHP												
SWD												
REF												
SLV												

Table 3: Database considered to estimate emissions for Argentina for each sector and species

In Table 3, a summary of the inventories used for each sector in Argentina is presented. For sectors RCO and FEF, where NH3 emissions were missing, the values were set to 0. Similarly, for the FEF sector, SO2 emissions were also assigned a value of 0. It is worth noting that these sectors collectively represent less than 1% of the total emissions of the species in the base inventory. The OC emissions were estimated through interpolation, utilizing the OC/PM25 ratio from the EDGAR dataset, as previously explained.

Furthermore, it is important to highlight that the TNR sector exclusively considers emissions from railways and does not account for emissions from other non-road transportation sources.
Regarding the Open Biomass Burning emissions, the GEAA inventory provides information only up to the year 2019. To extrapolate the emissions for the year 2020, we utilized the GFAS emissions data, adjusted based on the average proportion of species-specific emissions represented by GEAA relative to GFAS during the period 2014-2019. This approach ensures that the spatial representation of emissions from the GFAS satellite product for 2020 (that follows a similar methodology to the one used in GEAA), also maintains consistency in emission magnitudes with the GEAA time series."

1.4.3. Colombia

In the case of Colombia, a high-resolution database was available only for the road transport sector. for the rest of the sectors except SHP, SLV and OBB the emissions of the base data were scaled up to reach the magnitudes of the national Greenhouse inventory of Colombia. this inventory is the most complete of the region regarding species, considering that includes particles. For both, high-resolution transport inventory and national inventory, the emissions of OC were estimated by applying the OC/PM25 ratio from the EDGAR dataset, as previously explained.

Symbology	Base inventory			High-resolution inventory			Interpolated emissions		GHG Inventory			
	GHG			REACTIVE GASES					PARTICLES			
Specie / Sector	CO2	CH4	N2O	CO	NOx	NMVOCs	NH3	SO2	PM10	PM2.5	BC	OC
ENE	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	Base inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	Interpolated emissions
RCO	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	Base inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	Interpolated emissions
TRO	High-resolution inventory	High-resolution inventory	High-resolution inventory	High-resolution inventory	High-resolution inventory	High-resolution inventory	High-resolution inventory	High-resolution inventory	High-resolution inventory	High-resolution inventory	High-resolution inventory	Interpolated emissions
TNR	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	Base inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	Interpolated emissions
FEF	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	Base inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	Interpolated emissions
IND	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	Base inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	Interpolated emissions
OBB	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory
AGS		GHG Inventory	GHG Inventory		GHG Inventory	GHG Inventory	Base inventory		GHG Inventory	GHG Inventory		Interpolated emissions
AGL		GHG Inventory	GHG Inventory		GHG Inventory	GHG Inventory	Base inventory		GHG Inventory	GHG Inventory		Interpolated emissions
SHP	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory	Base inventory
SWD	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	Base inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	Interpolated emissions
REF	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	Base inventory	GHG Inventory	GHG Inventory	GHG Inventory	GHG Inventory	Interpolated emissions
SLV	Base inventory					Base inventory	Base inventory					

Table 4: Database considered to estimate emissions for Colombia for each sector and species

The composition of the mosaic in terms of the data used to estimate the emissions of each sector and species is summarized in Table 4. As explained before, for all sectors estimated from the national GHG inventory, the NH3 emissions of the base data remain unaltered.

1.4.4. Mexico

In the case of Mexico, a high-resolution inventory for the year 2016 available at a grid of 9x9 km was used. To account for the absence of emissions data in Mexico's inventory for the rest of the years of the period, interpolation was performed using CAMS emission trajectories specific to each sector and species for Mexico.

The sectors TNR, TRO, RCO, ENE, REF, IND, SWD, and AGL are directly estimated from the inventory itself. However, in the case of Mexico's inventory, there are missing data for the AGS and FEF sectors, particularly for species like NOx and CO2, which cannot be easily interpolated. These sectors represent more than 1% of the total emissions for the country, as per the base inventory.

Symbology	Base inventory		High resolution inventory			Interpolated emissions					
Specie / Sector	GHG		REACTIVE GASES					PARTICLES			
ENE											
RCO											
TRO											
TNR											
FEF											
IND											
OBB											
AGS											
AGL											
SHP											
SWD											
REF											
SLV											

Table 5: Database considered to estimate emissions for Mexico for each sector and species

Due to the lack of reliable data in Mexico's high-resolution inventory for these specific sectors, they are not considered for further analysis. Consequently, the emissions for these sectors from the high-resolution inventory are left unchanged from the data available on the base inventory.

1.4.5. Ecuador

Ecuador was the only country where only available information from the national GHG inventory. this inventory considers emissions for reactive gases and GHG from 1990 -2018. To ensure consistency with the rest of the database emissions of 2019 and 2020 were interpolated based on CAMS emission trajectories specific to each sector and species for Ecuador. Besides, as was explained before, emissions for particles were estimated based on the relationship between CO and each particle species for each sector and year in the EDDAR database for Ecuador.

Symbology	Base inventory			Interpolated emissions			GHG Inventory					
	GHG			REACTIVE GASES					PARTICLES			
Specie / Sector	CO2	CH4	N2O	CO	NOx	NMVOCs	NH3	SO2	PM10	PM2.5	BC	OC
ENE												
RCO												
TRO												
TNR												
FEF												
IND												
OBB												
AGS												
AGL												
SHP												
SWD												
REF												
SLV												

Table 6: Database considered to estimate emissions for Ecuador for each sector and species

Table 6 summarizes the composition of the Mosaic for Ecuador for each sector and species. the national GHG inventory was used to scale up the emissions of the base inventory for the sectors ENE, INF, REF, AGS, RCO, TNR and TRO. The FEF and AGL sectors weren't considered due to missing information on dame species or to big inconsistencies in the magnitudes reported between species.