



HBPG-CO₂: a global gridded monthly anthropogenic CO₂ emission dataset at 0.05° resolution for 1970–2025

Franz Pablo Antezana Lopez^{1,2}, Guanhua Zhou^{1*}, Kai Zhang^{3*}, Alejandro Casallas^{4,5}, Aamir Ali², Hongzhi Jiang¹, Jose Luis Alanoca Limachi²

- 5 1 School of Instrumentation Science and Optoelectronic Engineering, Beihang University, Beijing 100191, China.
2 Hangzhou International Innovation Institute, Beihang University, Hangzhou, China.
3 State Key Laboratory of Environmental Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, Beijing 100012, China.
4 Institute of Science and Technology Austria, Klosterneuburg 3400, Austria.
10 5 Atmospheric Chemistry Department, Max Planck Institute for Chemistry, 55128 Mainz, Germany.

Correspondence to: Guanhua Zhou (zhouguanhua@buaa.edu.cn), Franz Pablo Antezana Lopez (antezana@buaa.edu.cn)

Abstract. Carbon dioxide emissions, as geospatial information, are necessary to analyze seasonal and regional variability, as well as the changing distribution of anthropogenic activity. However, the availability of continuous data with high spatial and temporal resolution is limited, both sectorial and global. Therefore, in this study, we present the Hierarchical Bayesian Physics-Guided Carbon Dioxide (HBPG-CO₂) emission dataset, a global dataset of emissions from multiple sources for energy, industry, buildings, land transport, agriculture, waste, and fuel production. This dataset is based on multi-information harmonization and downscaling using geospatial information from multiple sources (nighttime lighting, roads, land use, population, settlements, meteorology, and point source data support temporal reconstruction and spatial refinement), various emissions inventories such as annual and monthly data from the Emissions Database for Global Atmospheric Research (EDGAR), as well as the Open source Data Inventory for Anthropogenic Carbon dioxide (ODIAC), Community Emissions Data System (CEDS), and daily estimates by country and sector. To integrate all the data, a hierarchical Bayesian component is applied to reconstruct monthly variability outside the directly observed period, while a physics-based component restricts the specific spatial allocation for each sector and preserves emissions data, while also allowing for area expansion based on decision trees. Furthermore, the integration of fossil fuel-related covariates that are underrepresented or absent in the original EDGAR fields helps reconstruct this dataset spatially and temporally. The resulting dataset contains 672 monthly layers per sector and global layers corresponding to the date range from 1970-01 to 2025-12. Comparisons with CEDS, ODIAC, EDGAR, and other studies show strong temporal and spatial consistency, while also preserving additional small-scale structure. Therefore, HBPG-CO₂ provides a long-term, high-resolution, sector-specific emissions record for atmospheric inversion, Earth system modeling, carbon cycle analysis, and regional emissions assessment. The complete HBPG-CO₂ dataset is openly accessible through Zenodo at <https://doi.org/10.5281/zenodo.21507026> (Antezana Lopez, 2026).



1 Introduction

Anthropogenic carbon dioxide (CO₂) emissions from power generation, industrial production, transportation, buildings, fuel production, agriculture, and waste are the primary direct cause of the long-term increase in atmospheric CO₂ and the associated climate forcing (Friedlingstein et al., 2025). Therefore, emission inventories are fundamental for atmospheric
 35 inversions, chemical transport modelling, Earth system modelling, regional carbon assessment, and the evaluation of mitigation pathways, while coordinated inventory development remains essential for ensuring consistency across scientific applications (Smith et al., 2022). These applications require more than national annual totals because the atmospheric signal depends not only on the magnitude of emissions but also on when and where source activity occurs. Existing spatially and temporally explicit information on fossil fuel CO₂ emissions has already demonstrated the importance of resolving this variability (Jones
 40 et al., 2021), particularly around cities, industrial belts, energy facilities, transport corridors, and rapidly changing land systems. Monthly information is needed to represent seasonal heating and cooling, agricultural cycles, mobility patterns, fuel consumption, and changes in economic activity. Moreover, a sufficiently fine grid is necessary to distinguish neighbouring source environments located within the same administrative unit. Long and spatially explicit monthly records are therefore required to connect national inventory accounting with regional and local analyses of the changing geography of human
 45 activity. Nonetheless, a persistent scale gap remains between national emission reporting and the high-resolution inputs required by atmospheric and Earth system models.

Several established datasets provide essential but complementary components of this information. The Emissions Database for Global Atmospheric Research (EDGAR) offers globally consistent annual sectoral estimates for 1970–2024 and monthly fields for January 2000 to December 2024 (Crippa et al., 2018; Janssens-Maenhout et al., 2019). The Open-source
 50 Data Inventory for Anthropogenic Carbon dioxide (ODIAC) provides monthly fossil fuel CO₂ emissions for 2000–2024 using power plant information and nighttime light observations to resolve the spatial distribution of nonpoint sources (Oda et al., 2018). Meanwhile, the Community Emissions Data System (CEDS) provides a long historical framework of country and sector emissions developed for climate and atmospheric applications (Hoesly et al., 2018; McDuffie et al., 2020). More recent emission mosaics have further combined regional and global inventory information to support atmospheric modelling (Crippa
 55 et al., 2023; Soulie et al., 2024), while regional inventories have demonstrated the value of integrating official statistics, activity information, remote sensing observations, and sector-specific spatial proxies. However, no product considered here combines global monthly coverage from 1970 to 2025, seven harmonised anthropogenic sectors, and a 0.05° grid within a single mass-conserving record. Long time series are commonly annual, whereas daily country and sector products are generally distributed as tabular estimates rather than continuous global grids. Monthly gridded products are often shorter, coarser, or focused mainly
 60 on fossil fuel combustion and cement production. Furthermore, although EDGAR provides the most consistent global sectoral basis, its native 0.10° spatial representation may not resolve secondary agricultural areas, minor road corridors, dispersed settlements, and evolving infrastructure that become relevant in regional and local analyses. Therefore, the remaining limitation



is not simply the scarcity of emission information, but the absence of a unified framework that combines temporal continuity, sectoral detail, fine spatial resolution, and accounting consistency.

65 Addressing this gap requires more than resampling an existing inventory. Near the equator, a 0.10° grid cell covers approximately 100 km^2 and may contain several roads, settlements, cultivated areas, industrial facilities, and other emission sources. Although dividing an inventory into finer grid elements increases the nominal spatial resolution, this operation alone neither determines the most plausible location of the conserved emission mass nor identifies strongly supported source areas that are absent from the original active domain. A coherent solution must therefore separate emission magnitude, monthly
 70 timing, and spatial allocation. In the Hierarchical Bayesian Physics Guided Carbon Dioxide framework, hereafter HBPG- CO_2 , EDGAR provides the principal annual and monthly mass anchor. The global daily CO_2 emission dataset developed by (Li et al., 2026) provides an additional temporal reference, reporting daily total and sectoral emissions from 1970 to 2024 at both global and individual country levels. CEDS supports historical consistency among harmonised sectors, whereas ODIAC provides an independent fine-scale spatial reference for fossil fuel activity. Multisource emission mapping has already
 75 demonstrated the value of combining inventory information, deep learning, and spatial covariates to improve regional source representation (Antezana López et al., 2025b). Within HBPG- CO_2 , the Bayesian component pools information across countries, sectors, years, and climatological profiles to reconstruct monthly variability where direct observations are incomplete. Subsequently, the physics-guided component applies process-relevant covariates and strict mass conservation constraints during spatial allocation. These covariates include HILDA+ land classes for agriculture (Winkler et al., 2021),
 80 global road information for transportation (Meijer et al., 2018), and spatial evidence from settlements, population, point sources, and nighttime lights for buildings, waste, and fossil fuel sectors (Elvidge et al., 2017; Urraca et al., 2024). Thus, the framework is designed to retain the accounting consistency of established inventories while improving spatial realism and temporal continuity.

HBPG- CO_2 is consequently presented as a global monthly anthropogenic CO_2 dataset at 0.05° resolution covering January
 85 1970 to December 2025. The product contains seven sectoral raster stacks and 672 monthly layers for agriculture, buildings, road transport, industry, power generation, waste, and fuel production, together with total emissions that also incorporate aviation and shipping components. Direct monthly inventory information is available for 300 months from January 2000 to December 2024 through EDGAR. In addition, the daily global and country level sectoral estimates of (Li et al., 2026) provide an independent temporal constraint for evaluating and harmonising monthly variability. The 360 months from 1970 to 1999
 90 are reconstructed from annual inventory anchors and learned temporal profiles, while the final 12 months constitute a documented extension through 2025. Spatial refinement redistributes each 0.10° source grid element across four nested 0.05° grid elements while preserving the prescribed mass at the source grid, sector, monthly, and global levels. For agriculture and road transport, a limited support expansion can activate high-confidence locations that are weakly represented or absent in the native EDGAR spatial support, while the redistributed fraction remains restricted to a small proportion of the corresponding
 95 monthly sector total. Previous satellite reconstruction studies have similarly demonstrated the value of combining complementary observations to recover spatially and temporally incomplete carbon information (Antezana López et al., 2022,



2025a). HBPG-CO₂ is therefore not intended as an unconstrained replacement for established inventories. Instead, it is a multisource reconstruction that extends temporal coverage, improves spatial representation, documents the degree of direct and reconstructed support, and provides a consistent high-resolution forcing record for atmospheric modelling, carbon cycle research, and regional emission assessment.

2 Methodology and data

Methodologically, HBPG-CO₂ separates temporal inference from spatial allocation, and its name reflects this division: Hierarchical Bayesian Physics Guided Carbon Dioxide. The Bayesian component operates on monthly timing. It uses an empirical hierarchical formulation in which directly supported country and sector profiles provide the observation-informed component, long-term climatologies and regularised trends provide prior structure, and their normalised partial pooling produces the monthly posterior shares used to distribute annual inventory mass. This formulation is designed for incomplete spatiotemporal observations (Alvera-Azcárate et al., 2005), while regularisation limits unstable coefficients (Friedman et al., 2010) and established temporal profiles provide an external basis for evaluating seasonality (Guevara et al., 2021). By contrast, the physics-guided component operates on spatial allocation. It applies sector process evidence, nonnegative weights, bounded support expansion, and exact mass closure at 0.05°. EDGAR establishes the mass and sector structure, ODIAC provides a fine-scale fossil fuel reference, and land, road, settlement, meteorological, vegetation, nighttime light, and point source information guide relative placement. Thus, the Bayesian component determines when the prescribed emissions occur, whereas the physics-guided component determines where they are allocated without creating unrestricted mass. Here, physics-guided denotes process constraints and conservation.

2.1 Input datasets and sector harmonisation

The main inputs and their functions are shown in Table 1. EDGAR's annual fields provide 55 source years from 1970 to 2024, while the documented 2025 extension yields the complete 56-year HBPG-CO₂ record. EDGAR's monthly stacks provide 300 directly backed months from January 2000 to December 2024 in their native 0.10° grid. In this study, "direct backing" refers to monthly temporal information from various emissions inventories at different temporal and spatial scales. ODIAC contributes the monthly fossil fuel structure, while CEDS provides an external historical comparison and supports sectoral harmonization. Daily activity data by country and sector limit month-to-month variability without replacing the bulk of the annual inventory. In addition, CAMS atmospheric fields and ERA5 meteorology provide diagnostic and environmental covariates (Hersbach et al., 2020; Inness et al., 2019). Settlements, population, roads, land use, vegetation, albedo, nighttime lights, and point sources provide spatial support for the sectors. Each dataset is aligned with a grid, a timeline, a country system, and a correspondence of seven common sectors prior to reconstruction.

Table 1: Principal datasets used in HBPG-CO₂, including their temporal coverage, native resolution, scientific role, provider, and official access location.



Dataset or variable	Role and provider	Official access or download URL
EDGAR_2025_GHG (monthly CO ₂)	EDGAR_2025_GHG — Emissions Database for Global Atmospheric Research; Joint Research Centre, European Commission	https://edgar.jrc.ec.europa.eu
EDGAR_2025_GHG (annual CO ₂)	EDGAR_2025_GHG — Emissions Database for Global Atmospheric Research; Joint Research Centre, European Commission	https://edgar.jrc.ec.europa.eu
ODIAC fossil-fuel CO ₂	ODIAC — Open-source Data Inventory for Anthropogenic CO ₂ ; National Institute for Environmental Studies (NIES)	https://www.odiac.org
CAMS EGG4 XCO ₂	CAMS global greenhouse-gas reanalysis column-averaged CO ₂ , used as an atmospheric consistency covariate rather than a direct emission constraint	https://ads.atmosphere.copernicus.eu
DMSP OLS / VIIRS night time lights	Harmonised DMSP OLS and VIIRS night-time light composites used as built environment and diffuse activity proxies; VIIRS products are provided by the Earth Observation Group.	https://eogdata.mines.edu/products/vnl/
GHS-SMOD R2023A	Global Human Settlement Layer — GHS-SMOD R2023A; European Commission Joint Research Centre	https://human-settlement.emergency.copernicus.eu
Harmonised NDVI composite	NOAA AVHRR/VIIRS normalized difference vegetation index climate data records, harmonised by the authors for long-term spatial support	https://www.ncei.noaa.gov/products/climate-data-records/normalized-difference-vegetation-index
GLASS BSA_SW albedo	GLASS Albedo Product — Global LAnd Surface Satellite; Beijing Normal University	https://glass-product.bnu.edu.cn
ERA5 total precipitation	ERA5 Monthly Averaged Reanalysis on Single Levels; Copernicus Climate Change Service / ECMWF	https://cds.climate.copernicus.eu
ERA5 2-m air temperature	ERA5 Monthly Averaged Reanalysis on Single Levels; Copernicus Climate Change Service / ECMWF	https://cds.climate.copernicus.eu
ERA5 10-m wind speed	ERA5 Monthly Averaged Reanalysis on Single Levels; Copernicus Climate Change Service / ECMWF	https://cds.climate.copernicus.eu
HILDA+ / ESA CCI–Copernicus C3S land cover	HILDA+ v2.0 land-use / land-cover dataset, Karlsruhe Institute of Technology; ESA CCI / Copernicus C3S land-cover extension	https://cds.climate.copernicus.eu
GRIP4 total road density	GRIP4 — Global Roads Inventory Project, Version 4; GLOBIO / PBL Netherlands Environmental Assessment Agency	https://www.globio.info
GRIP4 land-area support layer	GRIP4 — Global Roads Inventory Project, Version 4; companion land-area grid used in the road-density workflow	https://www.globio.info
gROADS v1 road network	gROADS v1 — Global Roads Open Access Data Set, Version 1; CIESIN / NASA SEDAC	https://sedac.ciesin.columbia.edu
CEDS country-sector CO ₂	Community Emissions Data System; independent annual country-sector benchmark and conservative calibration support	https://www.pnnl.gov/projects/ceds
Daily country and sector CO ₂ profiles	Global daily CO ₂ emissions from 1970 to 2024 for power, industry, residential, and transport sectors; (Li et al., 2026).	https://doi.org/10.6084/m9.figshare.28827932
Global Carbon Grid (GID)	Global 0.1° CO ₂ emission grid used as the common spatial source for major power and industrial point emissions.	https://gidmodel.org.cn/?page_id=1425

Accordingly, Fig. 1 distinguishes source coverage from reconstructed output. The annual EDGAR bridge contains 56 yearly fields for 1970–2024, and the 2025 extension completes the 56-year HBPg-CO₂ record. The directly supported monthly interval contains 300 source fields at 0.10°, whereas the final time axis contains 672 monthly fields at 0.05°. The importance of horizontal resolution for atmospheric CO₂ applications has been demonstrated in global modelling experiments (Agustí-Panareda et al., 2019). Spatially, the refinement creates four nested 0.05° sub cells within each 0.10° source grid cell and therefore increases the number of spatial elements by a factor of four. Satellite retrieval studies provide complementary context for interpreting the added spatial detail (Borsdorff et al., 2018; Crisp et al., 2017; Yokota et al., 2009), while global land surface



135 products support the environmental covariates (Liang et al., 2021). Thus, temporal continuity and spatial detail are increased without implying that every output month or fine grid element is directly observed.

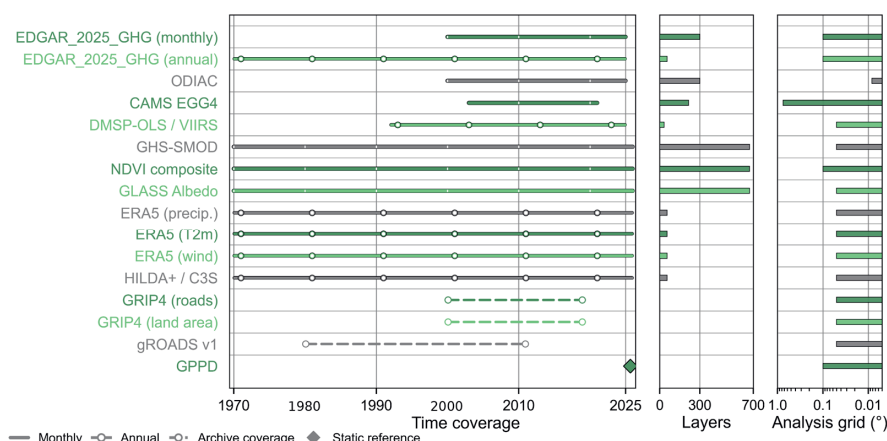


Figure 1: Temporal and spatial coverage of the principal inputs and the HBPG-CO₂ output. EDGAR contributes 56 annual source fields for 1970–2024 and 300 monthly source fields for 2000–2024; the documented 2025 extension completes a 56-year output containing 672 months. The spatial panel represents the conservative refinement from the 0.10° source grid to nested 0.05° grid elements.

Furthermore, the original source categories are assigned to seven common sectors: energy, industry, buildings, land transport, agriculture, waste, and fuel production. The same mapping is applied to all years and datasets. Therefore, changes in the resulting time series are not due to changes in sector labels. Aviation and international shipping are retained as complementary layers, as their geography and accounting differ from those of the seven land-based sectors.

Before the reconstruction, the available emissions data are placed within a common historical context, as shown in Figure 2. Global emissions from fossil fuels and processes increased from approximately 14 to 16 Gt CO₂ year⁻¹ in 1970 to around 35 to 38 Gt CO₂ year⁻¹ in more recent years. Long-term inventories document the same widespread growth in fossil fuel activity (Asefi-Najafabady et al., 2014; Oda and Maksyutov, 2011). However, regional trajectories differ. North America and Europe peak and then decline, while Asia shows the largest absolute increase. Similarly, high-resolution national mapping shows that aggregate trends exhibit strong subnational contrasts (Gurney et al., 2009). Africa and South America experience more gradual growth, while Australia remains relatively small. Furthermore, sectoral panels indicate that energy and industry dominate the long-term increase, while transport and construction contribute differently in each region. These comparisons are interpreted with due regard for the known uncertainty in the magnitude and spatial distribution of the global inventory (Solazzo et al., 2021).

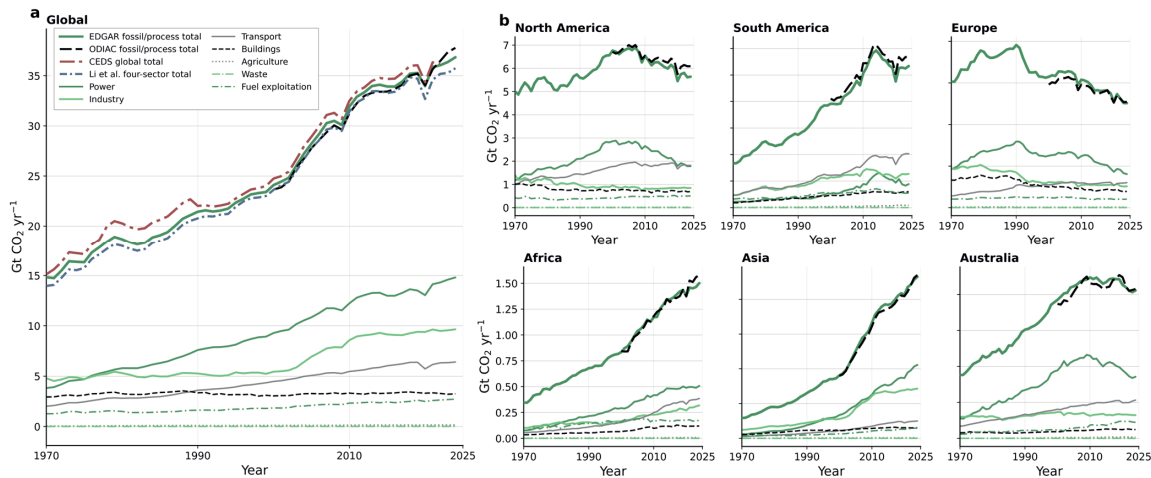


Figure 2: Historical context of the emission inventories and reference data used in HBPg-CO₂. (a) Global annual trajectories from EDGAR, ODIAC, CEDS, and daily data, together with the harmonised sector totals; the main global series rise from about 14 to 16 Gt CO₂ yr⁻¹ in 1970 to about 35 to 38 Gt CO₂ yr⁻¹ in recent years. (b to g) Regional trajectories for North America, South America, Europe, Africa, Asia, and Australia. The regional panels separate common long-term growth from contrasting peaks, declines, and sector transitions.

2.2 Covariate association and selection

To preserve sector specificity, the covariates were screened before spatial reconstruction. Power and industry are mainly supported by ODIAC, power point sources, nighttime lights, settlements, and road density. Moreover, transport is predicted primarily from GRIP4 and gROADS, whereas buildings are associated with nighttime lights, settlement structure, urban land, and 2 m air temperature. Agriculture is guided by HILDA land use, NDVI, precipitation, and albedo, while waste follows settlement and population proxies. Then, fuel production combines ODIAC, point sources, roads, and nighttime lights. The Mantel links classify these relationships from weak support, below 0.20, to strong support, at or above 0.60, following established guidance on interpreting spatial association tests (Legendre and Fortin, 2010). Hence, only sector covariates are retained, while satellite XCO₂ is also used to help the downscaling process.

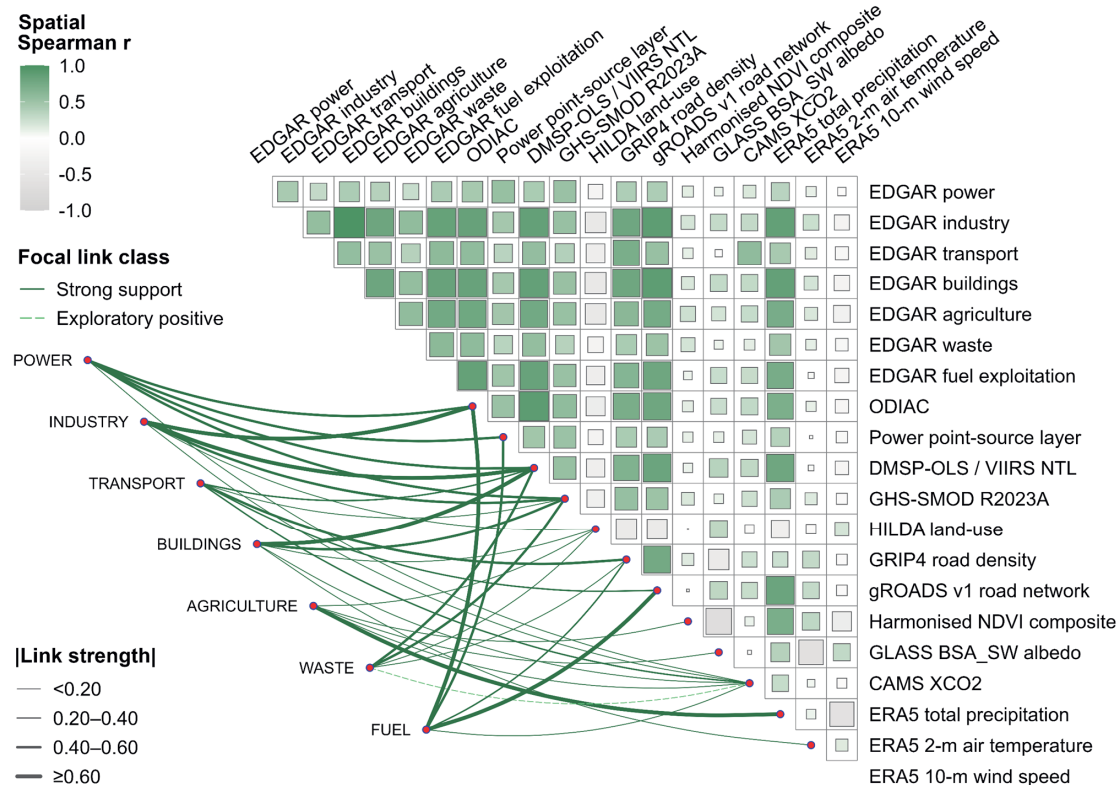


Figure 3: Spatial association analysis used to screen covariates for seven emission sectors across 20 candidate or reference fields. The upper triangular matrix reports spatial Spearman correlation among EDGAR sector fields and candidate datasets.

Consequently, the association structure supports different allocation rules by sector. Power and industry are most closely related to facility, settlement, night time light, and road information. Moreover, buildings are linked to population, settlement morphology, urban land, and temperature. By contrast, transport is controlled mainly by the two road products, while agriculture is more closely associated with land cover, vegetation, precipitation, and albedo. Long term land use products provide the historical context for these agricultural proxies (Chini et al., 2021; Klein Goldewijk et al., 2017). Fuel production combines point sources, night time lights, roads, and land information. Waste follows settlement and population support but remains the most weakly constrained diffuse sector. Biomass burning evidence is retained only where it is consistent with sector definitions (Akagi et al., 2011; van der Werf et al., 2017). Therefore, covariates are not transferred indiscriminately among sectors.

Accordingly, Fig. 4 summarises four stages with a single traceable flow from source data to validated output. Stage 1 harmonises inventories and spatial covariates to common units, sectors, dates, countries, and grids. Stage 2 estimates Bayesian monthly shares and sector specific spatial evidence separately, thereby preventing temporal information from being confused



with spatial support. Stage 3 combines these components through nonnegative allocation, bounded source support, and exact mass closure. Stage 4 exports the seven monthly sector fields together with support masks, uncertainty information, and independent validation diagnostics. Across the workflow, annual and monthly inventories define the admissible mass, while covariates only control its relative distribution.

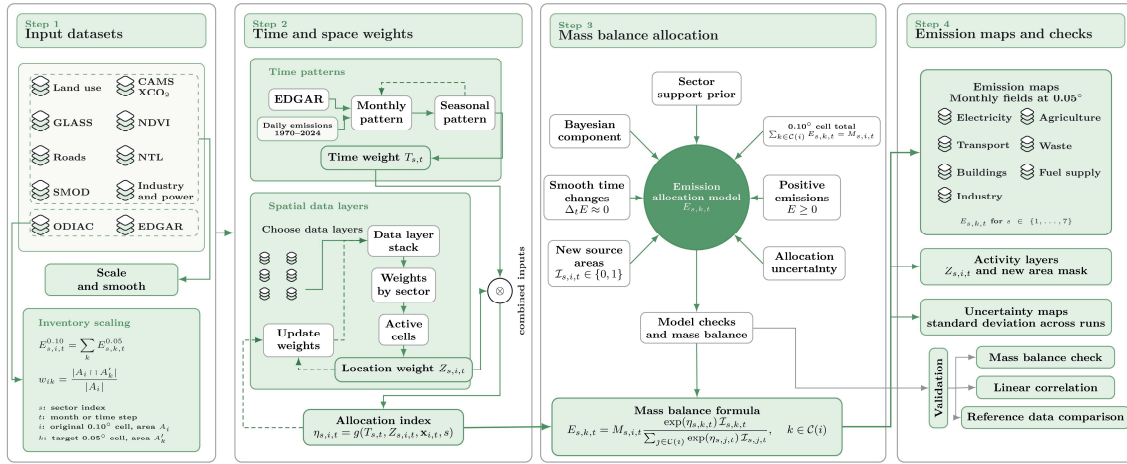


Figure 4: Traceable four stage workflow used to produce HBPG-CO₂.

2.3 Bayesian monthly temporal reconstruction

Within the temporal component, HBPG-CO₂ is implemented as an empirical hierarchical reconstruction of monthly shares rather than as a model of annual totals. In Eq. (1), c , s , y , and m index country, sector, year, and month; $E_{c,s,y,m}$ is the observed monthly inventory mass; k is the month summation index from 1 to 12; and $q_{c,s,y,m}^{obs}$ is the resulting positive dimensionless monthly share. The 12 shares sum to one for each valid country, sector, and year. Direct monthly observations provide the data supported profile, while weighted climatology, damped trends, and neighbouring temporal information provide hierarchical prior structure when observations are sparse. Consequently, the reconstructed profile changes only the within year distribution and cannot alter annual inventory mass.

$$q_{c,s,y,m}^{obs} = \frac{E_{c,s,y,m}}{\sum_{k=1}^{12} E_{c,s,y,k}} \quad (1)$$

Moreover, a representative monthly profile is estimated over the temporal training set $\mathcal{T} = 2000\text{--}2018$. This interval was selected because it contains continuous monthly EDGAR support while leaving the later years available for calibration and independent evaluation. In Eq. (2), $\mathcal{T}$ denotes the set of training years; $A_{(c,s,y)}$ is the annual country and sector mass used as a



nonnegative weight; $q_{c,s,y,m}^{obs}$ is the observed monthly share from Eq. (1); and $\bar{q}_{c,s,m}$ is the weighted climatological share for
 205 month m . Thus, years with greater activity contribute proportionally more information, and $\bar{q}_{c,s,m}$ acts as the prior mean. Well
 supported country and sector combinations remain close to their observations, whereas sparse combinations borrow strength
 from sectoral and regional seasonal structure.

$$\bar{q}_{c,s,m} = \frac{\sum_{y \in T} A_{c,s,y} \cdot q_{c,s,y,m}^{obs}}{\sum_{y \in T} A_{c,s,y}} \quad (2)$$

Furthermore, slow changes in seasonal timing are represented by a month specific trend in log share. In Eq. (3), $b_{c,s,m}$ is
 210 the regularised month specific trend coefficient, δ_τ is the damped and clipped year offset from the training reference period, ε
 is a small positive stabilising constant used before the logarithmic transformation, and k again indexes the 12 months in the
 normalising denominator. The resulting $\tilde{q}_{c,s,y,m}$ is positive and sums to one within each year. Extrapolation is limited to 12
 years on either side of the training endpoint so that weakly constrained trends cannot generate unrealistic historical or recent
 seasonality.

$$\tilde{q}_{c,s,y,m} = \frac{\exp\left[\log(\bar{q}_{c,s,m} + \varepsilon) + \delta_y b_{c,s,m}\right]}{\sum_{k=1}^{12} \exp\left[\log(\bar{q}_{c,s,k} + \varepsilon) + \delta_y b_{c,s,k}\right]} \quad (3)$$

At the local scale, country scale timing is combined with source grid climatology through confidence dependent partial
 pooling. In Eq. (4), p identifies a 0.10° source grid cell; $c(p)$ is the country containing p ; $q_{p,s,m}^{cell}$ is the local climatological
 monthly share; $\tilde{q}_{c(p),s,y,m}$ is the corresponding country and sector share; w_p is the local evidence weight; and N_m denotes
 normalisation over the 12 months. The pooled result $q_{p,s,y,m}^P$ therefore remains a valid annual composition. The weight w_p
 220 varies from 0.15 to 0.72 according to temporal support, allowing well supported cells to retain more local seasonality while
 incomplete or low activity cells borrow more strongly from the country profile.

$$q_{p,s,y,m}^P = N_m \left[w_p q_{p,s,m}^{cell} + (1 - w_p) \tilde{q}_{c(p),s,y,m} \right] \quad (4)$$

Consequently, temporal support differs across the 56-year record and is documented for every output band. Direct
 monthly timing at the EDGAR source resolution covers 300 of 672 months, or 44.6%, during 2000–2024. The 360 months
 225 from 1970–1999 represent 53.6% of the record and are reconstructed from annual constraints and learned profiles. The final
 12 months, or 1.8%, form the 2025 extension. Equation (5) defines that extension: $q_{p,s,2025,m}^P$ is the final monthly share,



$q_{p,s,2024,m}^{\text{obs}}$ is the last directly supported source grid profile, and $q_{p,s,2025,m}^{\text{model}}$ is the modelled 2025 profile. The fixed coefficients 0.65 and 0.35 retain continuity with 2024 while allowing the modelled seasonal evolution to contribute. The operator $\mathcal{N}_m$ again enforces a unit annual sum.

$$230 \quad q_{p,s,2025,m}^p = \mathcal{N}_m \left[0.65 q_{p,s,2024,m}^{\text{obs}} + 0.35 q_{p,s,2025,m}^{\text{model}} \right] \quad (5)$$

In addition, the daily country and sector series of (Li et al., 2026) is used as external temporal evidence rather than as a replacement for inventory mass. Daily values are aggregated to months and normalised within each country, sector, and year. Candidate blend strengths are estimated over 2019–2021, when both the daily reference and monthly EDGAR information are available, and are evaluated over the independent 2022–2024 period to prevent temporal information leakage. Only monthly
 235 shares are updated by the accepted blend. Therefore, the annual total remains controlled by the EDGAR based anchor, while the daily reference improves recent country and sector timing.

2.4 Physics guided spatial refinement and mass conservation

Subsequently, the physics guided component begins after the Bayesian monthly shares have been estimated. Here, physics guided refers to sector process knowledge, spatial support, positivity, bounded transfer, and conservation constraints rather
 240 than to atmospheric transport physics. In Eq. (6), $A_{p,s,y}$ is the annual mass assigned to source grid cell p , sector s , and year y ; $q_{p,s,y,m}^p$ is the pooled monthly share from Eq. (4) or Eq. (5); and $M_{p,s,y,m}$ is the resulting monthly source grid mass. For 2000–2024, the monthly shares use the certified EDGAR guide where available; for 1970–1999 and 2025, they use the Bayesian reconstruction. Recent EDGAR analyses provide the broader distribution framework (Crippa et al., 2024). A second support step is limited to agriculture and transport because these diffuse sectors are especially sensitive to incomplete land and
 245 road representation. High support inactive source grid cells are identified from process variables observed in active regions, while point source evidence remains anchored to independently mapped facilities (Nassar et al., 2017). The transferred pool is restricted to 2.2–4.0% of global agriculture mass and 1.0–2.0% of global transport mass per month; no more than 5.5% is withdrawn from the active source mass pool, and global sector mass is unchanged.

$$M_{p,s,y,m} = A_{p,s,y} q_{p,s,y,m}^p \quad (6)$$

250 At the fine grid stage, each accepted 0.10° source grid cell contains four nested 0.05° grid elements. In Eq. (7), j identifies a nested grid element, r identifies one of the R_s process variables available for sector s , $\mathcal{Z}_{j,r}$ is the standardised nonnegative value of variable r at j , $\omega_{s,r}$ is its sector specific weight, and $\mathcal{S}_{j,s}$ is the resulting baseline support score. The constraints

$\omega_{s,r} \geq 0$ and $\sum_{r=1}^{R_s} \omega_{s,r} = 1$ make the score interpretable as a convex combination. Agriculture uses HILDA+ cropland and



pasture, precipitation, NDVI, and albedo; transport uses the dynamic road product and gROADS with additional settlement,
 255 urban land, and nighttime light evidence. Buildings and waste use population, settlements, urban land, temperature, and
 nighttime lights, while power, industry, and fuel production use facility, road, settlement, and nighttime light evidence. These
 sector rules constrain redistribution according to plausible source processes.

$$S_{j,s} = \sum_{r=1}^{R_s} \omega_{s,r} z_{j,r}, \quad \omega_{s,r} \geq 0, \quad \sum_{r=1}^{R_s} \omega_{s,r} = 1 \quad (7)$$

Moreover, a regularised multivariable model learns within source grid contrasts not represented by the baseline score. In
 260 Eq. (8), $\mathbf{X}_s$ is the sector specific design matrix of centred feature contrasts, $\mathbf{y}_s$ is the reference contrast vector, λ_s is the ridge
 penalty, $\mathbf{P}$ is the penalty matrix with a zero entry for the unpenalised intercept, and $\hat{\boldsymbol{\beta}}_s$ is the fitted coefficient vector. The
 model uses 2000–2014 for fitting, 2015–2018 for selecting the refinement strength, and 2019–2024 for independent spatial
 evaluation. These intervals provide early, middle, and recent samples within the common EDGAR–ODIAC period and
 265 preserve a six-year holdout that is not used for fitting or parameter selection. ODIAC supplies the fine scale fossil fuel reference
 for power, industry, buildings, transport, and fuel production where its spatial support is informative, but it does not replace
 sector totals. A refinement is accepted only when it improves the held out spatial criterion and remains below the sector specific
 cap.

$$\hat{\boldsymbol{\beta}}_s = (\mathbf{X}_s^T \mathbf{X}_s + \lambda_s \mathbf{P})^{-1} \mathbf{X}_s^T \mathbf{y}_s \quad (8)$$

Consequently, Eqs. (9)–(11) convert the baseline support into a conservative fine grid allocation. In Eq. (9), p is the
 270 source grid cell, j and l index its nested 0.05° elements, $C(p)$ is their set, $b_{j|p,s}$ is the normalised physics guided baseline share,
 $\mathbf{x}_{j|p,s}$ is the centred feature contrast, $\hat{\boldsymbol{\beta}}_s$ is the fitted coefficient vector from Eq. (8), α_s is the accepted sector refinement
 strength, and $a_{j|p,s}$ is the final allocation share. Equation (10) defines $E_{j,s,y,m}^{HBPG}$ as the HBPG- CO_2 mass assigned to nested
 element j by multiplying $a_{j|p,s}$ by $M_{p,s,y,m}$. Equation (11) imposes exact source grid closure by requiring the sum of
 $E_{j,s,y,m}^{HBPG}$ over all j in $C(p)$ to equal $M_{p,s,y,m}$. Together with the global sector and month constraint, these equations allow
 275 supported locations to be resolved without an unconstrained increase in emissions.

$$a_{j|p,s} = \frac{b_{j|p,s} \exp(\alpha_s \mathbf{x}_{j|p,s}^T \hat{\boldsymbol{\beta}}_s)}{\sum_{\ell \in C(p)} b_{\ell|p,s} \exp(\alpha_s \mathbf{x}_{\ell|p,s}^T \hat{\boldsymbol{\beta}}_s)} \quad (9)$$

$$E_{j,s,y,m}^{HBPG} = M_{p,s,y,m} a_{j|p,s} \quad (10)$$



$$\sum_{j \in C(p)} E_{j,s,y,m}^{HBPG} = M_{p,s,y,m} \quad (11)$$

Therefore, the same sector rules are applied throughout the complete raster stack. Dynamic roads, population, land use, NDVI, albedo, and precipitation are read for their available dates, while static layers are used only as structural support. Once new agriculture or transport source-grid cells satisfy the global learned gate, their monthly mass follows the corresponding sector temporal profile and remains within the monthly expansion bounds. Therefore, the additional spatial support is propagated coherently through time rather than inserted as an isolated correction to a single map.

2.5 Evaluation strategy

For validation, four complementary tests are used. First, the temporal model is trained on 2000–2018, calibrated with the overlapping 2019–2021 reference period, and evaluated independently over 2022–2024. Second, the spatial model uses 2000–2014 for fitting, 2015–2018 for selecting the refinement strength, and 2019–2024 for evaluation. These chronological splits were chosen to represent early, intermediate, and recent conditions while avoiding the use of evaluation years during model fitting or parameter selection. Third, country and sector totals are compared with (Li et al., 2026) and CEDS using matched definitions and periods. Finally, signed residuals and source grid to nested grid closure quantify redistribution and numerical conservation. Because performance metrics emphasise different model properties (Willmott and Matsuura, 2005), no comparison inventory is treated as error free.

Moreover, the evaluation statistics include correlation, coefficient of determination, fitted slope, root mean square error, symmetric mean absolute percentage error, and the signed residual in Eq. (12). In that equation, $R_{j,s,y,m}$ is the signed residual at grid element j for sector s , year y , and month m ; $E_{j,s,y,m}^{HBPG}$ is the HBPG- CO_2 value; and $E_{j,s,y,m}^{EDGAR}$ is the EDGAR value represented on the comparison grid. The statistics are interpreted jointly: correlation describes ordering, the fitted slope describes multiplicative agreement, root mean square error describes absolute dispersion, symmetric mean absolute percentage error describes relative difference, and R identifies the direction of spatial redistribution.

$$R_{j,s,y,m} = E_{j,s,y,m}^{HBPG} - E_{j,s,y,m}^{EDGAR} \quad (12)$$

Consequently, a positive R value means that HBPG- CO_2 allocates more of the conserved mass to a grid element than the corresponding EDGAR representation, whereas a negative value means that mass is allocated away from that location. The residual therefore measures redistribution, not net emission creation. Positive and negative residuals must balance within each conserved source grid, sector, and month domain apart from numerical tolerance.



3 Results and dataset description

305 3.1 Data records and temporal coverage

In practical terms, the core release contains seven sector GeoTIFF raster stacks on a common WGS84 grid from 180° W to 180° E and from 60° S to 85° N. Each stack contains 7200 columns, 2900 rows, and 672 monthly Float32 layers from January 1970 through December 2025. One layer therefore contains 20,880,000 grid elements, one sector stack contains about 14.03 billion numerical values, and the seven core stacks contain approximately 98.22 billion numerical values in total.

310 Lossless GeoTIFF compression reduces storage size only; it does not alter the number of layers, grid elements, or numerical values. Overall, the product provides 4704 global sector and month fields.

Furthermore, the change from 0.10° to 0.05° increases the number of spatial elements by a factor of four. Within active source-grid cells, the increase is achieved by exact nested-grid allocation. Moreover, bounded support expansion allows a small number of well supported agriculture and transport cells to become active beyond the native source support. The
 315 expansion pool is limited to 2.2 to 4.0% of monthly global agriculture mass and 1.0 to 2.0% of monthly global transport mass, and no active source-grid cell contributes more than 5.5% of its original value.

3.2 Global and country scale emission patterns

At the global scale, the 2024 overview in Fig. 5 combines one global total field with six regional callouts for North America, South America, Europe, Africa, Asia, and Oceania. Each callout contains the total and six sector maps, producing
 320 42 regional panels in addition to the global map. Emission density spans approximately six orders of magnitude, from 10^{-2} to 10^4 t CO₂ km⁻² yr⁻¹. The strongest values occur in urban and industrial belts of eastern North America, Europe, eastern and southern Asia, and selected coastal regions. High resolution regional inventories report similarly concentrated urban, industrial, and transport structures (Kuenen et al., 2014). By contrast, agriculture and fuel production form broader and more diffuse patterns. Therefore, the 0.05° grid resolves both concentrated facilities and extended activity corridors within one global
 325 geometry, which is suitable for atmospheric transport modelling (Emmons et al., 2010).

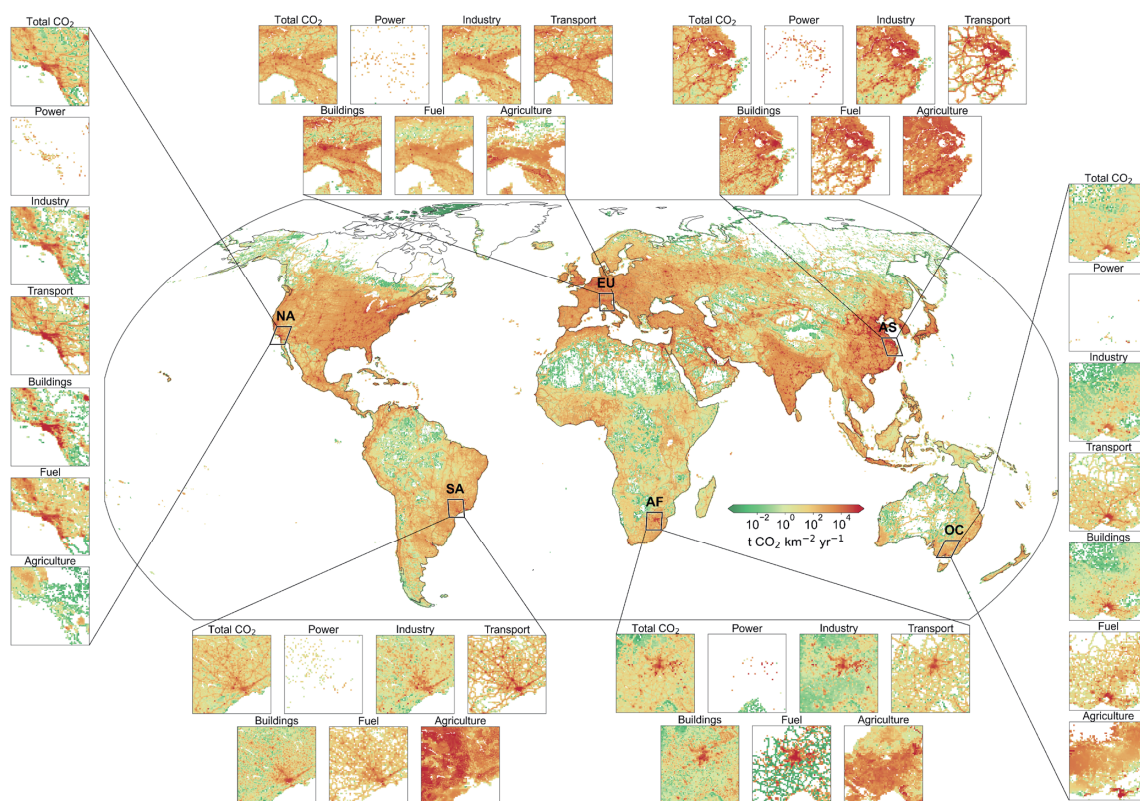
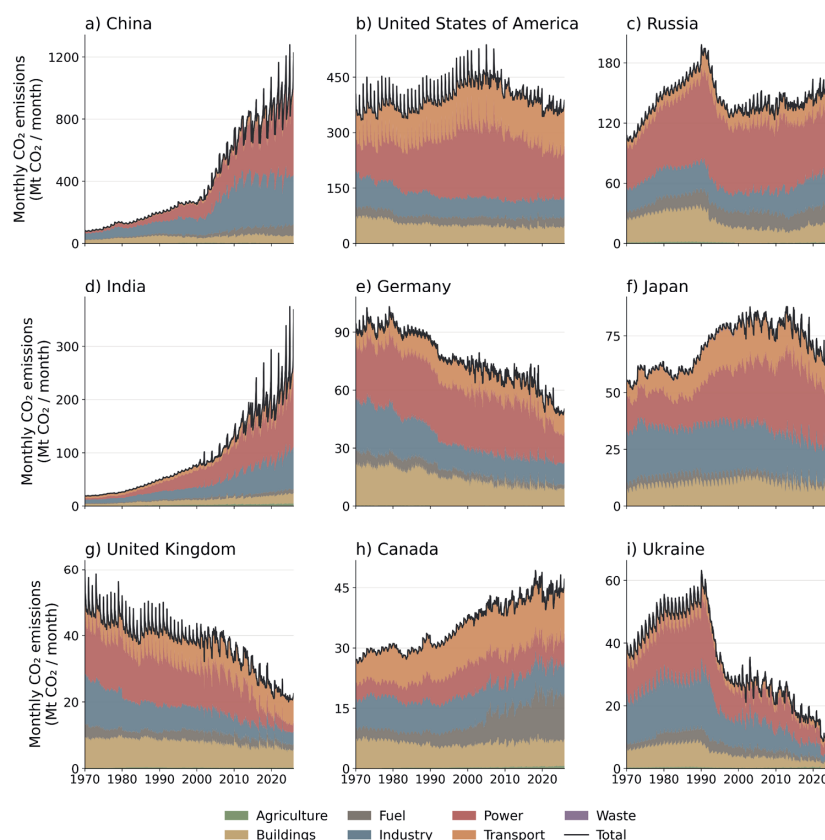


Figure 5: Global distribution of anthropogenic CO₂ emissions at 0.05° resolution, mean of 2024.

At the national scale, the revised monthly country series in Fig. 6 combines long-term change, seasonal variability, and sector composition over all 672 months. China increases from about 70 Mt CO₂ month⁻¹ in 1970 to recent seasonal peaks above 1200 Mt CO₂ month⁻¹, with the strongest growth arising from power and industry. India rises from roughly 15 Mt CO₂ month⁻¹ to peaks approaching 380 Mt CO₂ month⁻¹, particularly after 2000. The United States remains mostly between 350 and 450 Mt CO₂ month⁻¹, reaches short peaks above 500 in the 2000s, and then declines toward approximately 350 to 400. Russia grows from about 100 to nearly 200 Mt CO₂ month⁻¹ before the early 1990s contraction, falls to around 125, and subsequently recovers toward 170 to 180. Germany decreases from approximately 95 to 45 Mt CO₂ month⁻¹, the United Kingdom from about 50 to 20, and Ukraine from a peak above 60 to less than 10. Japan rises from about 55 to peaks near 85 before declining toward 60, whereas Canada increases from approximately 27 to recent peaks close to 50. The sector areas also distinguish the strong power and industry growth in China and India, the larger transport contribution in the United States and Canada, and the pronounced monthly building cycle in several temperate countries.



340 **Figure 6: Monthly HBPG-CO₂ emissions for nine countries from January 1970 to December 2025: (a) China, (b) the United States of America, (c) Russia, (d) India, (e) Germany, (f) Japan, (g) the United Kingdom, (h) Canada, and (i) Ukraine.**

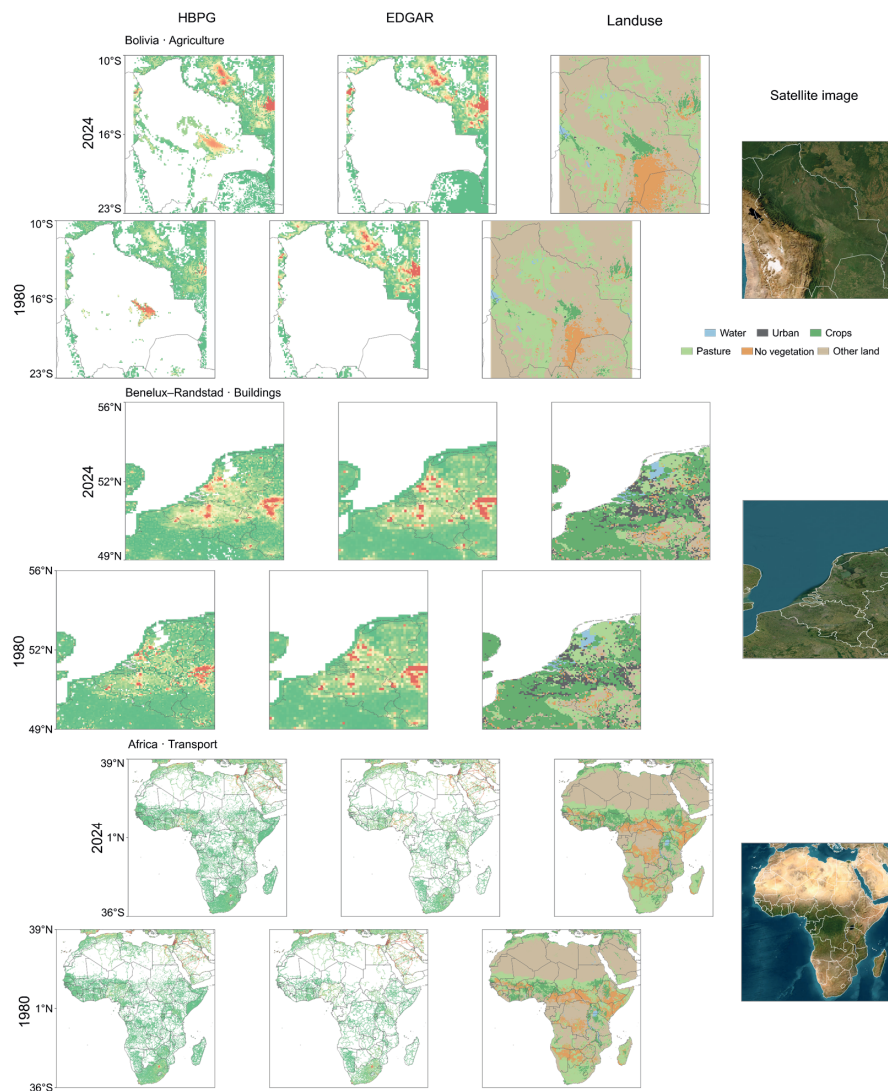
3.3 Spatial enhancement relative to EDGAR

Relative to EDGAR, HBPG-CO₂ is a spatially enhanced and temporally harmonised reconstruction of the EDGAR mass field, not a simple interpolation of the source grid. Reducing the cell size from 0.10° to 0.05° creates four possible nested grid
 345 locations within every source grid cell, while the multisource evidence determines their relative shares. Regional emission inventories show why sector specific proxies are required to recover local infrastructure and diffuse activity (Álamos et al., 2022; Kuenen et al., 2022). Moreover, the method can identify a limited number of strongly supported agriculture and transport locations that are not active in the native EDGAR support. Comparable national gridded inventories also demonstrate that remote sensing and statistical information can extend incomplete spatial records (Guo et al., 2025). Figure 7 therefore shows
 350 two related improvements: finer representation within existing source cells and conservative recovery of plausible activity areas that are not visible in EDGAR.



In the agricultural and urban examples, the Bolivia agriculture case shows that EDGAR represents the national mass but leaves several cultivated and pasture areas spatially unresolved. HILDA+ cropland and pasture classes, together with NDVI, precipitation, and albedo, provide independent land process evidence for these locations. Therefore, HBPG-CO₂ activates additional 0.05° cells that follow the observed agricultural landscape rather than spreading emissions uniformly across the coarse source-grid cell. The global agriculture expansion is nevertheless limited to 2.2 to 4.0% of monthly sector mass. In the Benelux and Randstad example, ODIAC fossil fuel structure, population, settlements, urban land, and night time lights sharpen buildings along the connected urban system. Moreover, the same ODIAC reference supports finer allocation in the fossil related areas and sectors for which it contains relevant information, including power, industry, buildings, transport, and fuel production, while the corresponding sector totals remain controlled by the inventory anchors.

Similarly, the Africa transport example shows the clearest improvement beyond the native EDGAR road support. The dynamic road product and gROADS recover mainland corridors, secondary connections, and routes around Madagascar that are weak or absent in the coarse inventory pattern. The accepted expansion remains limited to 1.0 to 2.0% of monthly global transport mass, and active-source withdrawal is capped at 5.5%, so the new road cells arise through redistribution rather than through added emissions. Across the complete framework, EDGAR provides the annual and monthly mass reference, (Li et al., 2026) improves recent monthly timing, CEDS constrains historical country and sector consistency, and ODIAC guides fine scale fossil fuel structure. Hence, HBPG-CO₂ improves both resolution and spatial support while remaining cross calibrated to established inventories and preserving the prescribed mass at sector, month, source-grid cell, and global levels.



370 **Figure 7: Examples of the HBPg-CO₂ spatial enhancement relative to the 0.10° EDGAR support. The rows show agriculture in**
 375 **Bolivia, buildings in the Benelux and Randstad region, and land transport across Africa for recent and historical periods.**

4 Technical validation

For technical validation, the framework combines staged temporal and spatial evaluation, cross inventory comparison, signed residual analysis, and mass accounting. Comparable atmospheric modelling datasets also use dedicated scenario and
 375 inventory checks to isolate methodological effects (Dolumbia et al., 2021). Each test addresses a different question. Held out



periods assess whether the learned temporal and spatial relationships extend beyond the fitting interval. On the other hand, (Li et al., 2026) evaluates monthly country and sector timing. CEDS evaluates annual country and sector consistency under a different inventory framework. Finally, residual maps and closure diagnostics show where mass is redistributed and whether the prescribed totals are conserved.

4.1 Monthly comparison with an external daily benchmark

In the monthly comparison, HBPg-CO₂ country and sector emissions were compared with (Li et al., 2026) after daily values were aggregated to months (Fig. 8). Each sector contains 7920 matched country and month observations, giving 31,680 comparisons in total. Rank correlation ranges from 0.96 for buildings to 0.99 for power and transport, while R² ranges from 0.94 to 1.00. The fitted slopes are 0.96 for power, 0.94 for transport, 1.10 for buildings, and 0.96 for industry. Symmetric mean absolute percentage error is lowest for transport at 7.1%, followed by industry at 16.1%, power at 19.4%, and buildings at 36.4%. Therefore, the reconstructed monthly timing is consistent for the four major sectors, although the broader relative error for buildings indicates greater sensitivity to seasonal and accounting differences.

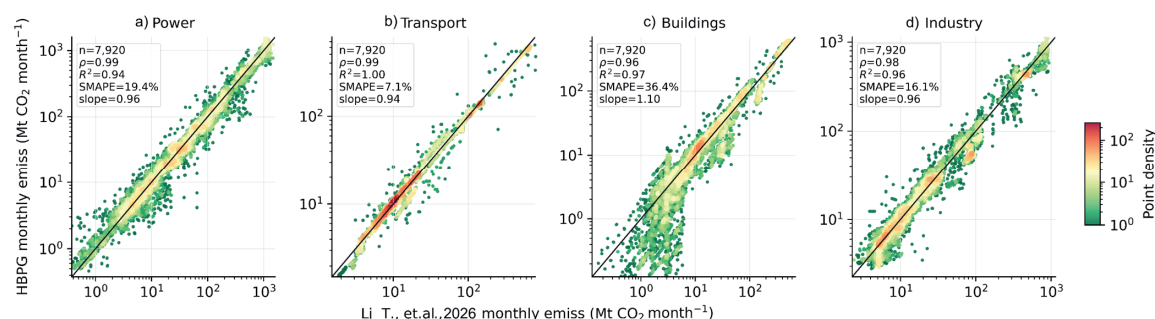


Figure 8: Monthly comparison of HBPg-CO₂ with (Li et al., 2026) after aggregation of the daily reference data. Panels show (a) power, (b) land transport, (c) buildings, and (d) industry.

4.2 Cross inventory comparison with CEDS

In the cross-inventory comparison, CEDS covers the 38-year period from 1985 to 2022 and contains 44,161 matched country, sector, and year observations (Fig. 9). The four concentrated fossil sectors retain strong country ranking: ρ is 0.98 for power, 0.97 for transport, 0.95 for buildings, and 0.98 for industry. Their R² values range from 0.97 to 1.00, and their fitted slopes range from 0.76 to 0.90. Fuel production also remains strongly ordered, with $\rho = 0.91$, R² = 0.94, and a slope of 0.77, although its symmetric mean absolute percentage error reaches 95.0%. By contrast, agriculture has $\rho = 0.77$, R² = 0.63, and an error of 104.2%, while waste has $\rho = 0.33$, R² = 0.10, an error of 140.6%, and a slope of 11.37. Therefore, CEDS supports the broad geographical ordering of the main energy sectors but also shows that diffuse sectors are highly sensitive to sector scope, activity data, and crosswalk definitions.

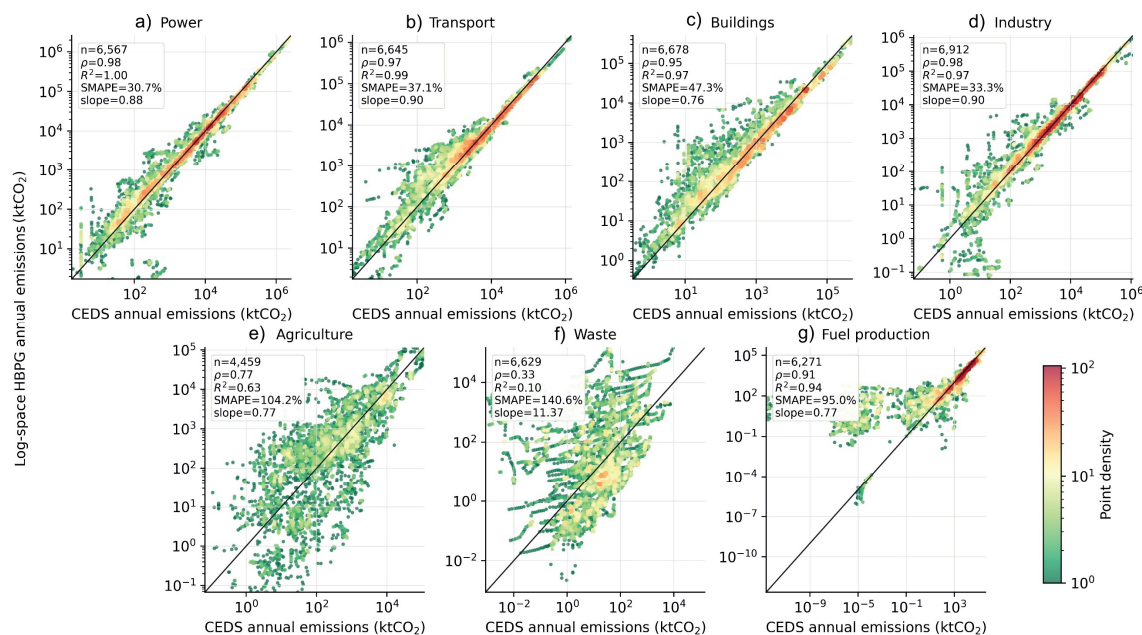


Figure 9: Annual country and sector comparison between HBPG-CO₂ and CEDS for 1985 to 2022. Panels show (a) power, (b) land transport, (c) buildings, (d) industry, (e) agriculture, (f) waste, and (g) fuel production.

4.3 Residual diagnostics and conservation checks

Spatially, the mean signed residuals in Fig. 10 show where HBPG-CO₂ redistributes mass relative to EDGAR for the total and seven sectors. Positive values identify cells that gain allocation, whereas negative values identify source locations contributing redistributed mass. The symmetric display range is approximately $\pm 1.8 \times 10^4$ kt CO₂ yr⁻¹ for the total. Sector ranges vary from ± 5.2 kt CO₂ yr⁻¹ for waste and $\pm 1.1 \times 10^2$ for agriculture to $\pm 1.4 \times 10^5$ for power. Fuel production, buildings, transport, and industry use ranges of approximately $\pm 1.2 \times 10^3$, $\pm 1.9 \times 10^3$, $\pm 2.6 \times 10^3$, and $\pm 4.0 \times 10^3$ kt CO₂ yr⁻¹, respectively. Moreover, the paired red and blue patterns around urban regions, road corridors, agricultural margins, and point sources show that spatial enhancement is achieved by relocation rather than by uniform amplification.

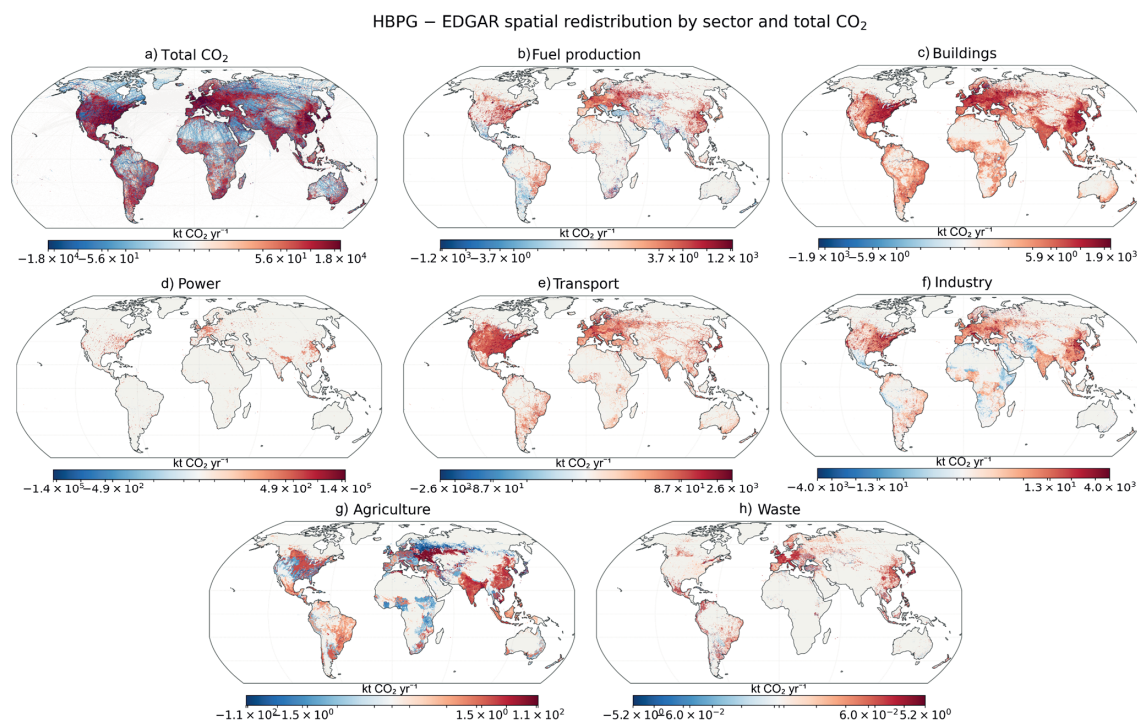


Figure 10: Mean signed spatial residuals between HBPG-CO₂ and EDGAR, defined as HBPG-CO₂ minus EDGAR, for (a) total CO₂, (b) fuel production, (c) buildings, (d) power, (e) transport, (f) industry, (g) agriculture, and (h) waste.

At the synthesis stage, the principal spatial and inventory comparisons are brought together in Fig. 11. Panel (a) shows the global HBPG-CO₂ total field in Robinson projection and preserves visible continental, urban, industrial, aviation, and shipping structure. Panel (b) summarises country level HBPG-CO₂ to reference ratios for seven sectors and the common total of power, industry, buildings, and land transport. The factor axis extends from one eighth to eight times the reference, and most central ranges remain between one half and twice the reference. EDGAR ratios cluster near one because it supplies the principal mass anchor. By contrast, CEDS shows the widest dispersion for agriculture and waste, while (Li et al., 2026) remains close to one for the four comparable sectors. Panel (c) compares the 15 largest common total emitters in 2020. China exceeds 10,000 Mt CO₂ yr⁻¹, the United States is near 4000, and India is close to 2000, while the remaining countries range from several hundred to about 1500 Mt CO₂ yr⁻¹. Therefore, the figure distinguishes agreement in country ranking from differences in sector scope and magnitude.

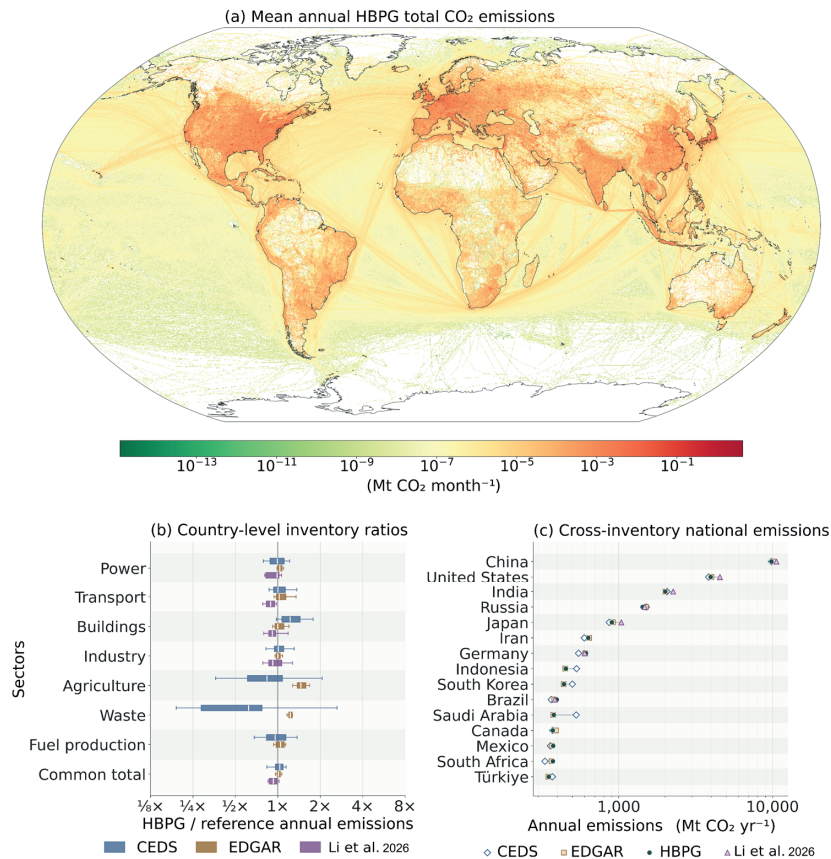


Figure 11: Synthesis of global spatial coverage and country level inventory consistency. (a) Mean HBPG-CO₂ total CO₂ field in Robinson projection, displayed logarithmically in Mt CO₂ month⁻¹. (b) Distributions of HBPG-CO₂ to reference annual emission ratios for power, transport, buildings, industry, agriculture, waste, fuel production, and the common total. Boxes compare CEDS, EDGAR, and (Li et al., 2026) on a factor axis from 1/8 to 8. (c) Annual emissions in 2020 for the 15 largest common total emitters, comparing CEDS, EDGAR, HBPG-CO₂, and daily emission data.

Finally, mass closure is verified at two levels. First, agriculture and transport support expansion conserves the global sector and month total before refinement. Second, the four nested 0.05° sub cells reproduce the resulting 0.10° source-grid mass according to Eqs. (10) and (11). The audit stores input mass, output mass, absolute difference, relative difference, and supported cell count for all 4704 core layers. Consequently, the increase in spatial detail does not alter the prescribed inventory totals apart from numerical precision.



435 5 Discussion, recommended use, and limitations

In terms of application, HBPG-CO₂ is intended for applications that require a continuous monthly, sector, and spatially anthropogenic CO₂ field. Moreover, these applications include atmospheric transport and inversion experiments, Earth system model forcing, regional carbon assessment, urban and infrastructure analysis, and studies of changing emission geography. The common grid and band index allow direct extraction of country, basin, latitude band, and user defined regional time series.

440 Moreover, conservative aggregation is recommended whenever the product is transferred to a coarser model grid.

Nevertheless, the added value and uncertainty differ by period and sector. The interval from 2000 to 2024 has direct monthly inventory support and therefore the strongest temporal basis. The interval from 1970 to 1999 preserves annual mass but reconstructs monthly timing from learned profiles. The 2025 fields form a documented extension. Spatial support is strongest where facilities, settlements, roads, land use, and point source information are available. By contrast, early decades
 445 and diffuse sectors depend more strongly on reconstructed timing and indirect spatial evidence. Therefore, every application should retain the temporal status flags and sector definitions supplied with the data.

Overall, the principal contribution of HBPG-CO₂ is not an independent replacement for EDGAR, CEDS, ODIAC, or national inventories. Instead, it converts complementary information into one monthly global record while preserving traceable mass constraints. Moreover, the comparisons show strong temporal and geographical consistency for the principal energy
 450 sectors, but they also reveal substantial differences for agriculture, waste, and fuel production. Cross inventory differences are expected when national histories and sector definitions are harmonised from distinct sources (Gütschow et al., 2021). These differences reflect inventory scope as well as reconstruction uncertainty, a limitation that also affects global budgets for diffuse emission sectors (Saunio et al., 2020). Nonetheless, the combination of 56-year continuity, a fixed 0.05° grid, seven harmonised sectors, explicit residual diagnostics, and exact accounting provides a transparent basis for modelling and
 455 comparative analysis. Fine-scale values should still be interpreted as support for informed allocations rather than direct measurements at individual cells.

6 Conclusions

In summary, HBPG-CO₂ provides 672 monthly anthropogenic CO₂ layers for each of seven harmonised sectors on a global 0.05° grid from 1970 through 2025. Annual inventory mass is retained throughout the reconstruction. Direct monthly
 460 timing constrains 300 months, or 44.6% of the record; 360 earlier months, or 53.6%, are reconstructed from annual anchors and learned profiles; and the final 12 months form a documented 1.8% extension. Each layer contains 20.88 million grid elements, while the complete release contains 4704 sector and month fields and approximately 98.2 billion numerical values. Lossless file compression changes storage size but not these counts. Moreover, every 0.10° source grid cell is allocated exactly among four nested 0.05° elements, so the added spatial detail does not change prescribed mass.

465 Moreover, the evaluation shows strong monthly agreement with Li et al. (2026), with correlations of 0.96 to 0.99 and a median symmetric mean absolute percentage error of 17.8 % across four major sectors. Furthermore, comparison with CEDS



also preserves strong country ranking for power, transport, buildings, industry, and fuel production, with correlations of 0.91 to 0.98. By contrast, agriculture and waste show larger differences because their definitions, activity data, and spatial support vary more strongly among inventories. The regional examples and signed residual maps further show that HBPG-CO₂ redistributes mass toward supported roads, settlements, agricultural areas, and point sources rather than creating new totals. Therefore, HBPG-CO₂ offers a coherent long-term monthly forcing field for atmospheric, carbon cycle, and regional studies. Nonetheless, early decades, diffuse sectors, and individual fine-scale cells remain more model dependent and should be interpreted with the supplied metadata and conservative aggregation rules.

Code and data availability

The complete HBPG-CO₂ data record, including the seven core sector and total rasterstacks in GeoTIFF format, figure source data, and evaluation summaries, is located here: <https://doi.org/10.5281/zenodo.21507026> (Antezana Lopez, 2026).

Author contributions

FPAL: conceptualisation, data curation, formal analysis, investigation, methodology, software, validation, visualisation, and writing – original draft. GZ: conceptualisation, funding acquisition, project administration, resources, and supervision. KZ, AC, AA, HJ, and JLAL: methodology, validation, interpretation, and writing – review and editing. All authors reviewed and approved the submitted manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Disclaimer

Copernicus Publications remains neutral about jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Acknowledgements

The authors acknowledge the institutions and data providers that made this work possible, including the Joint Research Centre of the European Commission for EDGAR and Global Human Settlement Layer products; the National Institute for Environmental Studies for ODIAC; the Copernicus Climate Change and Atmosphere Monitoring Services and the European Centre for Medium-Range Weather Forecasts for ERA5 and CAMS products; and the developers of the DMSP-OLS/VIIRS



night-time lights, GLASS albedo, HILDA+, GRIP, gROADS, and related open geospatial datasets. We also acknowledge
 495 Beihang University and the Hangzhou International Innovation Institute, Beihang University, for institutional support.

Financial support

This research was supported by the National Natural Science Foundation of China (Grant No. 42471425).

Review statement

The review statement will be added by Copernicus Publications and will list the handling editor and contributing referees
 500 according to their anonymous or identified status.

References

- Agustí-Panareda, A., Diamantakis, M., Massart, S., Chevallier, F., Muñoz-Sabater, J., Barré, J., Curcoll, R., Engelen, R., Langerock, B., et al.:
 Modelling CO₂ weather – why horizontal resolution matters, *Atmos. Chem. Phys.*, 19, 7347–7376, <https://doi.org/10.5194/acp-19-7347-2019>,
 2019.
- 505 Akagi, S. K., Yokelson, R. J., Wiedinmyer, C., Alvarado, M. J., Reid, J. S., Karl, T., Crounse, J. D., and Wennberg, P. O.: Emission factors for open
 and domestic biomass burning for use in atmospheric models, *Atmos. Chem. Phys.*, 11, 4039–4072, <https://doi.org/10.5194/acp-11-4039-2011>,
 2011.
- Álamos, N., Huneus, N., Opazo, M., Osses, M., Puja, S., Pantoja, N., Denier van der Gon, H., Schueftan, A., Reyes, R., and Calvo, R.: High-
 resolution inventory of atmospheric emissions from transport, industrial, energy, mining and residential activities in Chile, *Earth Syst. Sci. Data*,
 14, 361–379, <https://doi.org/10.5194/essd-14-361-2022>, 2022.
- 510 Alvera-Azcárate, A., Barth, A., Rixen, M., and Beckers, J.-M.: Reconstruction of incomplete oceanographic data sets using empirical orthogonal
 functions: application to the Adriatic Sea surface temperature, *Ocean Model.*, 9, 325–346, <https://doi.org/10.1016/j.ocemod.2004.08.001>, 2005.
- Antezana Lopez, F. P., Zhou, G., Jing, G., Zhang, K., and Tan, Y.: XCO₂ and XCH₄ reconstruction using GOSAT satellite data based on EOF-
 algorithm, *Remote Sens.*, 14, 2622, <https://doi.org/10.3390/rs14112622>, 2022.
- 515 Antezana Lopez, F. P., Zhou, G., Jing, G., Zhang, K., Chen, L., Chen, L., and Tan, Y.: Global Daily Column Average CO₂ at 0.1° × 0.1° Spatial
 Resolution Integrating OCO-3, GOSAT, CAMS with EOF and Deep Learning, *Sci. Data*, 12, 268, <https://doi.org/10.1038/s41597-024-04135-w>, 2025a.
- Antezana Lopez, F. P., Casallas, A., Zhou, G., Zhang, K., Jing, G., Ali, A., et al.: High-resolution anthropogenic emission inventories with deep
 learning in northern South America, *Remote Sens. Environ.*, 324, 114761, <https://doi.org/10.1016/j.rse.2025.114761>, 2025b.
- 520 Antezana Lopez, F. P.: HBPG-CO₂: a global gridded monthly anthropogenic CO₂ emission dataset at 0.05° resolution for January 1970–December
 2025, Zenodo [data set], <https://doi.org/10.5281/zenodo.21507026>, 2026.
- Asefi-Najafabady, S., Rayner, P. J., Gurney, K. R., McRobert, A., Song, Y., Coltin, K., Huang, J., Elvidge, C., and Baugh, K.: A multiyear, global
 gridded fossil fuel CO₂ emission data product: evaluation and analysis of results, *J. Geophys. Res.-Atmos.*, 119, 10213–10231,
<https://doi.org/10.1002/2013JD021296>, 2014.
- 525 Borsdorff, T., Aan de Brugh, J., Hu, H., Hasekamp, O., Susmann, R., and Rettinger, M.: Mapping carbon monoxide pollution from space down to
 city scales with daily global coverage, *Atmos. Meas. Tech.*, 11, 5507–5518, <https://doi.org/10.5194/amt-11-5507-2018>, 2018.
- Chini, L., Hurtt, G., Frolking, S., Sahajpal, R., and Goldewijk, K. K.: Land-use harmonization datasets for annual global carbon budgets, *Earth Syst.*
Sci. Data, 13, 4175–4189, <https://doi.org/10.5194/essd-13-4175-2021>, 2021.
- 530 Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., van Aardenne, J. A., Monni, S., Doering, U., Olivier, J. G. J., Pagliari, V., and
 Janssens-Maenhout, G.: Gridded emissions of air pollutants for the period 1970–2012 within EDGAR v4.3.2, *Earth Syst. Sci. Data*, 10, 1987–
 2013, <https://doi.org/10.5194/essd-10-1987-2018>, 2018.
- Crippa, M., Janssens-Maenhout, G., Guizzardi, D., Solazzo, E., Yan, F., Huo, D., Zheng, B., 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,
<https://doi.org/10.5194/essd-15-2667-2023>, 2023.
- 535 Crippa, M., Guizzardi, D., Pagani, F., Schiavina, M., Melchiorri, M., Pisoni, E., Graziosi, F., Muntean, M., Maes, J., Dijkstra, L., Van Damme, M.,
 Clarisse, L., and Coheur, P.: 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, <https://doi.org/10.5194/essd-16-2811-2024>,
 2024.



- 540 Crisp, D., Pollock, H. R., Rosenberg, R., Chapsky, L., Lee, R. A. M., Oyafuso, F. A., Frankenberg, C., O'Dell, C. W., Bruegge, C. J., Doran, G. B., et al.: The on-orbit performance of the Orbiting Carbon Observatory-2 (OCO-2) instrument and its radiometrically calibrated products, *Atmos. Meas. Tech.*, 10, 59–81, <https://doi.org/10.5194/amt-10-59-2017>, 2017.
- Doumbia, T., Granier, C., Elguindi, N., Bouarar, I., Darras, S., Brasseur, G., Gaubert, B., Liu, Y., Shi, X., Stavrakou, T., Tilmes, S., Lacey, F., Deroubaix, A., and Wang, T.: Changes in global air pollutant emissions during the COVID-19 pandemic: a dataset for atmospheric modeling, *Earth Syst. Sci. Data*, 13, 4191–4206, <https://doi.org/10.5194/essd-13-4191-2021>, 2021.
- 545 Eldering, A., Taylor, T. E., O'Dell, C. W., and Pavlick, R.: The OCO-3 mission: measurement objectives and expected performance based on 1 year of simulated data, *Atmos. Meas. Tech.*, 12, 2341–2370, <https://doi.org/10.5194/amt-12-2341-2019>, 2019.
- Elvidge, C. D., Baugh, K., Zhizhin, M., Hsu, F.-C., and Ghosh, T.: VIIRS night-time lights, *Int. J. Remote Sens.*, 38, 5860–5879, <https://doi.org/10.1080/01431161.2017.1342050>, 2017.
- 550 Emmons, L. K., Walters, S., Hess, P. G., Lamarque, J.-F., Pfister, G. G., Fillmore, D., Granier, C., Guenther, A., Kinnison, D., Laepple, T., et al.: Description and evaluation of the Model for Ozone and Related chemical Tracers, version 4 (MOZART-4), *Geosci. Model Dev.*, 3, 43–67, <https://doi.org/10.5194/gmd-3-43-2010>, 2010.
- Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J., Landschützer, P., Le Quéré, C., et al.: Global Carbon Budget 2024, *Earth Syst. Sci. Data*, 17, 965–1039, <https://doi.org/10.5194/essd-17-965-2025>, 2025.
- Friedman, J., Hastie, T., and Tibshirani, R.: Regularization paths for generalized linear models via coordinate descent, *J. Stat. Softw.*, 33, 1–22, <https://doi.org/10.18637/jss.v033.i01>, 2010.
- 555 Guevara, M., Jorba, O., Tena, C., Denier van der Gon, H., Kuenen, J., Elguindi, N., Darras, S., Granier, C., and Pérez García-Pando, C.: Copernicus Atmosphere Monitoring Service TEMPoral profiles (CAMS-TEMPO): global and European emission temporal profile maps for atmospheric chemistry modelling, *Earth Syst. Sci. Data*, 13, 367–404, <https://doi.org/10.5194/essd-13-367-2021>, 2021.
- Guo, F., Dai, F., Gong, P., and Zhou, Y.: CHN-CH₄: a gridded (0.1° × 0.1°) anthropogenic methane emission inventory of China from 1990 to 2020, *Earth Syst. Sci. Data*, 17, 4799–4819, <https://doi.org/10.5194/essd-17-4799-2025>, 2025.
- 560 Gurney, K. R., Mendoza, D. L., Zhou, Y., Fischer, M. L., Miller, C. C., Geethakumar, S., and de la Rue du Can, S.: High resolution fossil fuel combustion CO₂ emission fluxes for the United States, *Environ. Sci. Technol.*, 43, 5535–5541, <https://doi.org/10.1021/es900806c>, 2009.
- Gütschow, J., Günther, A., Pflüger, M., and Meinshausen, M.: The PRIMAP-hist national historical emissions time series (1750–2019) v2.3.1, *Earth Syst. Sci. Data*, 13, 5213–5252, <https://doi.org/10.5194/essd-13-5213-2021>, 2021.
- 565 Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al.: The ERA5 global reanalysis, *Q. J. Roy. Meteor. Soc.*, 146, 1999–2049, <https://doi.org/10.1002/qj.3803>, 2020.
- Hoesly, R. M., Smith, J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., et al.: Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), *Geosci. Model Dev.*, 11, 369–408, <https://doi.org/10.5194/gmd-11-369-2018>, 2018.
- 570 Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., et al.: The CAMS reanalysis of atmospheric composition, *Atmos. Chem. Phys.*, 19, 3515–3556, <https://doi.org/10.5194/acp-19-3515-2019>, 2019.
- Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., Bergamaschi, P., Pagliari, V., Olivier, J. G. J., Peters, J. A. H. W., et al.: EDGAR v4.3.2 Global Atlas of the three major greenhouse gas emissions for the period 1970–2012, *Earth Syst. Sci. Data*, 11, 959–1002, <https://doi.org/10.5194/essd-11-959-2019>, 2019.
- 575 Jones, M. W., Andrew, R. M., Peters, G. P., Janssens-Maenhout, G., De-Gol, A. J., Ciais, P., Patra, P. K., Chevallier, F., and Le Quéré, C.: Gridded fossil CO₂ emissions and related O₂ combustion consistent with national inventories 1959–2018, *Sci. Data*, 8, 2, <https://doi.org/10.1038/s41597-020-00779-6>, 2021.
- Klein Goldewijk, K., Beusen, A., Doelman, J., and Stehfest, E.: Anthropogenic land use estimates for the Holocene – HYDE 3.2, *Earth Syst. Sci. Data*, 9, 927–953, <https://doi.org/10.5194/essd-9-927-2017>, 2017.
- 580 Kuenen, J. J. P., Visschedijk, A. J. H., Jozwicka, M., and Denier van der Gon, H. A. C.: TNO-MACC-II emission inventory; a multi-year (2003–2009) consistent high-resolution European emission inventory for air quality modelling, *Atmos. Chem. Phys.*, 14, 10963–10976, <https://doi.org/10.5194/acp-14-10963-2014>, 2014.
- Kuenen, J., Dellaert, S., Visschedijk, A., Jalkanen, J.-P., Super, I., and Denier van der Gon, H.: CAMS-REG-v4: a state-of-the-art high-resolution European emission inventory for air quality modelling, *Earth Syst. Sci. Data*, 14, 491–515, <https://doi.org/10.5194/essd-14-491-2022>, 2022.
- 585 Legendre, P. and Fortin, M.-J.: Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data, *Mol. Ecol. Resour.*, 10, 831–844, <https://doi.org/10.1111/j.1755-0998.2010.02866.x>, 2010.
- Li, T., Wang, L., Qiu, Z., Ciais, P., Davis, S. J., Deng, Z., Zhao, Y., Peters, G. P., Ke, P., Jones, M. W., et al.: Global daily CO₂ emissions from 1970 to 2024, *Sci. Data*, 13, 605, <https://doi.org/10.1038/s41597-026-06621-9>, 2026.
- 590 Liang, S., Wang, D., He, T., and Yu, Y.: The Global LAnd Surface Satellite (GLASS) product suite, *B. Am. Meteorol. Soc.*, 102, E323–E337, <https://doi.org/10.1175/BAMS-D-18-0341.1>, 2021.



- McDuffie, E. E., Smith, S. J., O'Rourke, P., Tibrewal, K., Venkataraman, C., Marais, E. A., Zheng, B., Crippa, M., Brauer, M., and Martin, R. V.: A global anthropogenic emission inventory of atmospheric pollutants from sector- and fuel-specific sources (1970–2017): an application of the Community Emissions Data System (CEDS), *Earth Syst. Sci. Data*, 12, 3413–3442, <https://doi.org/10.5194/essd-12-3413-2020>, 2020.
- 595 Meijer, J. R., Huijbregts, M. A. J., Schotten, K. C. G. J., and Schipper, A. M.: Global patterns of current and future road infrastructure, *Environ. Res. Lett.*, 13, 064006, <https://doi.org/10.1088/1748-9326/aabd42>, 2018.
- Nassar, R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D. B. A., and Crisp, D.: Quantifying CO₂ emissions from individual power plants from space, *Geophys. Res. Lett.*, 44, 10045–10053, <https://doi.org/10.1002/2017GL074702>, 2017.
- Oda, T. and Maksyutov, S.: A very high-resolution global fossil fuel CO₂ emission inventory derived using a point source database and satellite observations of night-time lights, *Tellus B*, 63, 784–801, <https://doi.org/10.1111/j.1600-0889.2010.00510.x>, 2011.
- 600 Oda, T., Maksyutov, S., and Andres, R. J.: The Open-source Data Inventory for Anthropogenic CO₂, version 2016 (ODIAC2016): a global monthly fossil fuel CO₂ gridded emissions data product for tracer transport simulations and surface flux inversions, *Earth Syst. Sci. Data*, 10, 87–107, <https://doi.org/10.5194/essd-10-87-2018>, 2018.
- Saunio, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., Raymond, P. A., Dlugokencky, E. J., Houweling, S., Patra, P. K., et al.: The Global Methane Budget 2000–2017, *Earth Syst. Sci. Data*, 12, 1561–1623, <https://doi.org/10.5194/essd-12-1561-2020>, 2020.
- 605 Smith, S. J., Ahsan, H., Mott, A., McDuffie, E. E., Crippa, M., Klimont, Z., and Charles, M.: Opinion: Coordinated development of emission inventories for climate forcers and air pollutants, *Atmos. Chem. Phys.*, 22, 13201–13218, <https://doi.org/10.5194/acp-22-13201-2022>, 2022.
- Solazzo, E., Crippa, M., Guizzardi, D., Muntean, M., Choulga, M., and Janssens-Maenhout, G.: Uncertainties in the Emissions Database for Global Atmospheric Research (EDGAR) emission inventory of greenhouse gases, *Atmos. Chem. Phys.*, 21, 5655–5683, <https://doi.org/10.5194/acp-21-5655-2021>, 2021.
- 610 Soulie, A., Guevara, M., Jalkanen, J.-P., Keita, S., Denier van der Gon, H., Kuenen, J., Granier, C., et al.: Global anthropogenic emissions (CAMSGLOB-ANT) for the Copernicus Atmosphere Monitoring Service simulations of air quality forecasts and reanalyses, *Earth Syst. Sci. Data*, 16, 2261–2288, <https://doi.org/10.5194/essd-16-2261-2024>, 2024.
- Urraca, R., Oreggioni, G. D., Monforti-Ferrario, F., Crippa, M., Guizzardi, D., Janssens-Maenhout, G., et al.: CoCO₂-MOSAIC 1.0: a global mosaic of regional, gridded, fossil, and biofuel CO₂ emission inventories, *Earth Syst. Sci. Data*, 16, 501–523, <https://doi.org/10.5194/essd-16-501-2024>, 2024.
- 615 van der Werf, G. R., Randerson, J. T., Giglio, L., van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., van Marle, M. J. E., Morton, D. C., Collatz, G. J., et al.: Global fire emissions estimates during 1997–2016, *Earth Syst. Sci. Data*, 9, 697–720, <https://doi.org/10.5194/essd-9-697-2017>, 2017.
- Willmott, C. J. and Matsuura, K.: Advantages of the mean absolute error over the root mean square error in assessing average model performance, *Clim. Res.*, 30, 79–82, <https://doi.org/10.3354/cr030079>, 2005.
- Winkler, K., Fuchs, R., Rounsevell, M., and Herold, M.: Global land use changes are four times greater than previously estimated, *Nat. Commun.*, 12, 2501, <https://doi.org/10.1038/s41467-021-22702-2>, 2021.
- Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W. T., Sherlock, V., and Wennberg, P. O.: The Total Carbon Column Observing Network, *Philos. T. Roy. Soc. A*, 369, 2087–2112, <https://doi.org/10.1098/rsta.2010.0240>, 2011.
- 625 Yokota, T., Oguma, H., Morino, I., and Inoue, G.: A nadir looking SWIR sensor to monitor CO₂ column density for Japanese GOSAT project, *SOLA*, 5, 160–163, <https://doi.org/10.2151/sola.2009-041>, 2009.