



1 **A global, temporally continuous, hourly CO₂ concentration dataset at GAW monitoring**
2 **stations for 2000-2024: Enabling fine-scale carbon dynamics analysis**

3

4 Tan Mi¹, Xueting Pu², Xi Zheng¹, Guangcheng Nie¹, Xinyi Liu³, Han Zhang⁴, Michael L.
5 Grieneisen⁵, Yu Zhan¹, Maohua Wang⁶, and Fumo Yang¹

6

7 ¹College of Carbon Neutrality Future Technology, Sichuan University, Chengdu, Sichuan
8 610065, China

9 ²Department of Civil and Environmental Engineering, The Pennsylvania State University,
10 University Park, PA 16803, United States

11 ³Department of Environmental Science and Engineering, Sichuan University, Chengdu, Sichuan
12 610065, China

13 ⁴State Grid Sichuan Electric Power Research Institute, Chengdu, Sichuan, 610041, China

14 ⁵Department of Land, Air, and Water Resources, University of California, Davis, CA 95616,
15 United States

16 ⁶Institute for Disaster Management and Reconstruction, Sichuan University-The Hong Kong
17 Polytechnic University, Chengdu 610207, China

18

19 *Correspondence to:* Yu Zhan (yzhan@scu.edu.cn) and Maohua Wang (wmh@scu.edu.cn)



20 Abstract

21 Global Atmosphere Watch (GAW) carbon dioxide (CO₂) observations provide high-precision
22 constraints on surface CO₂ variability, but many station records remain temporally incomplete
23 because of flask sampling, instrument interruptions, and quality-control exclusions. Such gaps
24 can limit the representation of diurnal variability, bias daily, monthly, and annual mean CO₂
25 estimates, and reduce the ability to detect short-lived events. Here, we present a global,
26 temporally continuous hourly reconstruction dataset of GAW station CO₂ concentrations for
27 2000–2024. The dataset was generated using a two-stage reconstruction framework that
28 integrates GAW observations with Carbon Tracker (CT) and Copernicus Atmosphere Monitoring
29 Service (CAMS) CO₂ fields, together with anthropogenic, vegetation, meteorological, and
30 spatiotemporal predictors. In the intra-day gap filling stage, missing hours on days with at least
31 one valid observation were reconstructed using hour-pair models to produce complete daily
32 profiles. In the full-sequence reconstruction stage, the remaining periods with no valid
33 observations were reconstructed using a globally trained initial prediction model followed by
34 station-specific post-correction. The framework showed robust validation performance at the
35 hourly scale, with $R^2 = 0.96$ and RMSE = 3.65 ppm for intra-day gap filling and $R^2 = 0.93$ and
36 RMSE = 4.76 ppm for full-sequence reconstruction. Across all stations, the mean filling rate was
37 68.9 %, and the mean reconstruction uncertainty was 2.80 ppm. Applying the World Data Centre
38 for Greenhouse Gases (WDCGG) global analysis framework showed that reconstructed global
39 monthly mean CO₂ concentrations were, on average, 0.54 ± 0.21 ppm higher than those derived
40 from the original GAW records, indicating that incomplete temporal coverage can measurably
41 affect global monthly estimates. Compared with grid-cell CT and CAMS simulations, the
42 reconstructed dataset retained stronger station-scale diurnal variability and captured CO₂
43 enhancements during the 2023 Canadian wildfire period. This dataset can support long-term
44 assessments of global and regional surface CO₂ changes, fine-scale analyses of CO₂ variability,
45 and regional carbon emission inversion and evaluation. The dataset is available at
46 <https://doi.org/10.5281/zenodo.21306261> (Mi et al., 2026).



47 1. Introduction

48 Atmospheric carbon dioxide (CO₂) is the dominant anthropogenic greenhouse gas and a central
49 driver of contemporary climate change (Jin et al., 2024). Since the Industrial Revolution, its
50 global mean concentration has risen from about 280 ppm to more than 424 ppm by 2024, mainly
51 because of fossil fuel combustion, cement production, and land-use change (Cheng et al., 2023;
52 Tian et al., 2021; Lan et al., 2026). Although this long-term increase is well established,
53 atmospheric CO₂ also varies substantially across space and time due to anthropogenic emissions,
54 terrestrial biospheric exchange, ocean-atmosphere fluxes, and atmospheric transport
55 (Friedlingstein et al., 2025). This variability is especially pronounced near the surface, where
56 CO₂ concentrations can change rapidly in response to photosynthesis and respiration, local
57 emissions, convection, and boundary-layer dynamics (Mustafa et al., 2022; Te et al., 2023;
58 Roobaert et al., 2024). High-temporal resolution observations are therefore essential for
59 resolving fine-scale CO₂ variability, diagnosing carbon-cycle responses to environmental
60 perturbations, constraining and evaluating carbon-cycle models, assessing mitigation efforts, and
61 improving projections of future climate change.

62

63 Atmospheric CO₂ is monitored mainly through satellite remote sensing and ground-based
64 observation networks (He et al., 2022). Satellite missions such as Orbiting Carbon Observatory-2
65 (OCO-2) and Greenhouse Gases Observing Satellite (GOSAT) provide near-global observations
66 of column-averaged CO₂, but their retrievals are often limited by cloud cover, aerosols, and weak
67 solar illumination at high latitudes, leading to substantial data gaps and reduced sensitivity to
68 rapid fine-scale variability (Noël et al., 2022; Taylor et al., 2023; Hu et al., 2024). Ground-based
69 networks under the World Meteorological Organization (WMO) Global Atmosphere Watch
70 (GAW) program provide highly accurate in situ surface CO₂ measurements, which are archived
71 and distributed by the World Data Centre for Greenhouse Gases (WDCGG) operated by the
72 Japan Meteorological Agency (JMA) (Wu et al., 2024). However, these records are often
73 temporally incomplete. Many stations rely on flask sampling, in which air is collected



74 periodically and analyzed offline, typically at weekly or monthly intervals that cannot resolve
75 diurnal CO₂ variability (Lin et al., 2015). Even continuous in situ records can contain extended
76 gaps due to instrument failures, power interruptions, and quality-control filtering. Such temporal
77 incompleteness may introduce sampling bias, especially when missing observations occur
78 systematically at certain times of day. For example, the absence of afternoon measurements,
79 when photosynthetic uptake and boundary-layer mixing often lower near-surface CO₂ can bias
80 daily means upward. These gaps also affect estimates of long-term CO₂ trends and increase
81 uncertainty in global carbon-budget assessments and sources-sink attribution (Friedlingstein et
82 al., 2025).

83

84 Global atmospheric inversion and assimilation systems, such as Carbon Tracker (CT) and the
85 Copernicus Atmosphere Monitoring Service (CAMS), provide continuous CO₂ fields that can
86 help compensate for gaps in GAW measurements. These systems combine atmospheric CO₂
87 observations with prior surface fluxes and atmospheric transport models through inversion or
88 data-assimilation frameworks to estimate spatiotemporally complete CO₂ mole fractions (Van
89 Der Laan-Luijkx et al., 2017; Inness et al., 2019). However, these products cannot fully
90 substitute for high-frequency, station-level GAW observations. Their public outputs are typically
91 provided at intervals of 3 hours and coarse spatial resolutions, which limits their ability to
92 resolve rapid sub-daily variability, local emission signals, boundary-layer transitions, and sharp
93 gradients (Kenea et al., 2019; Chen et al., 2019; Wang et al., 2021). Their near-surface CO₂ fields
94 are also influenced by prior emission inventories and transport parameterizations, particularly in
95 regions with complex topography or strong synoptic variability (Kountouris et al., 2018). As a
96 result, simulated CO₂ fields often smooth local peaks and dampen diurnal and seasonal
97 amplitudes (Agustí-Panareda et al., 2023), which can obscure short-term emission or uptake
98 signals and limit their use for evaluating fine-scale carbon cycle processes and event-derived
99 CO₂ enhancements. Accordingly, CT and CAMS are better suited as physically informed
100 covariates than as direct replacements for GAW observations when reconstructing temporally



101 continuous hourly station-level CO₂ records.

102

103 These limitations point to a central data gap, namely the lack of long-term, globally distributed
 104 GAW CO₂ records that are continuous at hourly resolution. Current station-based assessments
 105 usually rely on monthly aggregation and statistical gap filling to estimate global or regional CO₂
 106 means and trends, which can suppress sub-daily variability and introduce uncertainty when
 107 missing values are inferred from nearby sites (Tsutsumi et al., 2009; Wu et al., 2024). Temporal
 108 gaps and uneven station availability may therefore bias estimates of mean CO₂ abundance,
 109 growth rates, and seasonal cycles, particularly at flask sampling stations that observe only during
 110 selected hours and cannot resolve diurnal variability associated with biospheric exchange and
 111 local transport. This limitation also affects atmospheric inversion systems such as CT, where
 112 observations are often assimilated under well-mixed conditions to reduce transport errors,
 113 leaving full diurnal constraints underused and potentially affecting annual source-sink estimates
 114 (Masarie et al., 2014; Munassar et al., 2025). Missing hourly records further limit the detection
 115 of short-lived CO₂ enhancements from sources such as wildfire plumes, whose timing,
 116 magnitude, and duration may be underestimated or missed entirely (Karandana Gamalathge et
 117 al., 2024). A temporally continuous hourly reconstruction of GAW CO₂ measurements is
 118 therefore needed to transform irregular station records into consistent time series for assessing
 119 global and regional mean CO₂ abundance, seasonal and diurnal variability, and rapid carbon-
 120 release events.

121

122 This study addresses that gap by reconstructing a global, temporally-continuous, hourly station-
 123 based GAW CO₂ dataset for 2000-2024. We combined WDCGG-archived in situ observations
 124 with simulated CO₂ fields from CT and the CAMS, together with anthropogenic, vegetation,
 125 meteorological, spatial, and temporal predictors. The reconstruction used a two-stage framework
 126 that first filled the missing hours within partially observed days and then estimated fully missing
 127 days, which preserved the observed diurnal information where available while extending records



through periods without observations. The resulting dataset complements gridded CO₂ products by retaining station-level variability that is often smoothed in global simulations. This enables more consistent analyses of global and regional CO₂ abundance, growth rates, seasonal cycles, and diurnal variability, and provides a basis for detecting short-lived anomalies such as wildfire-related CO₂ enhancements. By transforming irregular GAW records into a long-term, globally distributed hourly time series, this dataset offers a station-based benchmark for evaluating carbon-cycle models, assessing inversion products, and investigating fine-scale carbon dynamics over the past quarter century.

136

137 **2. Materials and Methods**

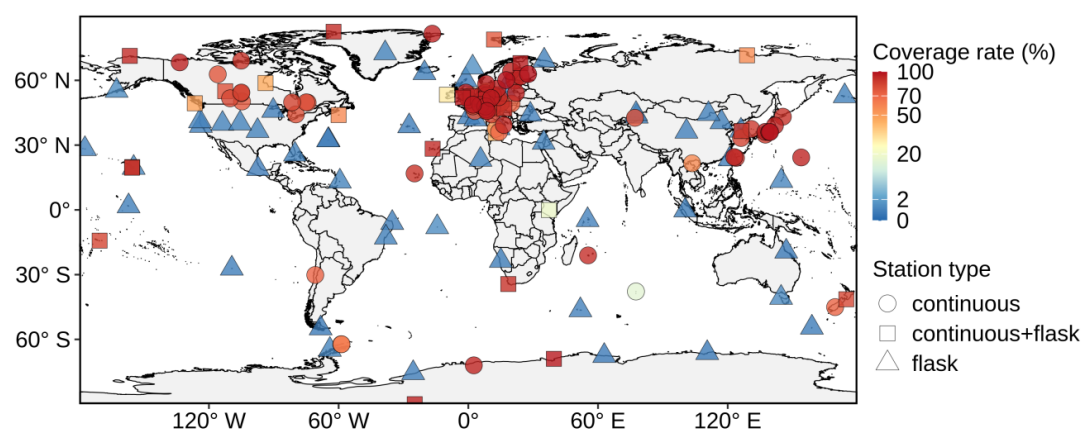
138 **2.1 Datasets**

139 **2.1.1 GAW CO₂ monitoring data**

GAW monitoring observations, including greenhouse gases, are collected, archived, and distributed by the WDCGG through contributions from participating institutions (Organization, 2025), and are publicly available at <https://gaw.kishou.go.jp/> (accessed on 27 June, 2025). We downloaded global CO₂ concentration records from all available stations from 2000 onward, including hourly and event records. After obtaining the initial data from 171 stations, we retained only valid measurements, corresponding to quality flags 1 and 2, and excluded invalid measurements with quality flag 3. To support reliable long-term analysis and representative seasonal coverage, we further removed stations with no significant interannual growth or with one or more months without observations in a single year (Fig. S1). This screening yielded 152 GAW stations for 2000-2024, comprising more than 10 million samples (Fig. 1). These stations are globally distributed, although 80.9 % are in the Northern Hemisphere, with particularly dense coverage in Europe. By observation methods, 73 stations provide continuous measurements (6,881,113 samples), while 53 rely on flask sampling only (43,826 samples), and 26 include both types (3,435,568 samples). For stations with both methods, continuous measurements were prioritized when both data types were available, while flask data were used as supplementary



155 records (Tsutsumi et al., 2009). Temporal coverage varies substantially among the stations, with
 156 a mean hourly coverage rate of 52.9 %. Continuous stations have much higher coverage than
 157 flask stations, at 84 % versus 0.5 %, reflecting the large temporal gaps inherent in flask
 158 sampling.
 159



160
 161 **Figure 1.** Global distribution of GAW CO₂ monitoring stations and their data coverage rates. The
 162 coverage rate for each station was calculated as the ratio of valid hourly samples to the total
 163 number of hours between the first and last observation timestamps. Points are colored according
 164 to coverage rate and shaped according to station type.

166 2.1.2 Model-simulated CO₂ data

167 Simulated near-surface CO₂ concentrations from CT and CAMS were used as key predictors for
 168 reconstructing the GAW CO₂ observation time series. We used two CT products: CT2022 for
 169 January 2000 to February 2021 and the near-real time CT-NRT v2025-1 for March 2021 to
 170 December 2024, both available at <https://gml.noaa.gov/ccgg/carbontracker/>. These products
 171 provide global three-dimensional CO₂ mole fraction fields at 3-hour resolution, with 34 vertical
 172 levels and a spatial resolution of 3°×2°. CO₂ concentrations from the near-surface layer,
 173 corresponding to level 1, were extracted for this study. We also used the CAMS greenhouse gas
 174 reanalysis product (EGG4), available at <https://ads.atmosphere.copernicus.eu/>, which covers



2003-2020 and provides global 3-hourly CO₂ fields at 0.75°×0.75° resolution with 137 vertical model levels. As with CT, near-surface CAMS CO₂ concentrations were extracted as a predictor for GAW CO₂ reconstruction. For each GAW station, CT and CAMS CO₂ concentrations were extracted from the model grid cell containing the station location and then temporally interpolated from 3-hour to hour resolution using a natural cubic spline (Perperoglou et al., 2019) (Supplement S1 and Fig. S2). For periods without CAMS data, namely 2000-2002 and 2021-2024, missing CAMS CO₂ values were estimated using a LightGBM model based on CT CO₂ data and the environmental variables described in Section 2.1.3 (Supplement S2 and Fig. S3).

2.1.3 Ancillary variables

In addition to the CT and CAMS simulated CO₂ fields, we incorporated ancillary variables to reconstruct temporally continuous hourly GAW CO₂ concentrations. These variables capture three major drivers of atmospheric CO₂ variability, including anthropogenic activities, vegetation indicators, and meteorological factors (Wang et al., 2025). Anthropogenic proxies characterize human-related emission patterns, including population, land use, nighttime lights, and emission inventories. Vegetation indicators represent biospheric carbon exchange, including the normalized difference vegetation index (NDVI), gross primary productivity (GPP), net primary productivity (NPP), fraction of photosynthetically active radiation (FPAR), and leaf area index (LAI). Meteorological factors reflect atmospheric transport and boundary-layer dynamics, including temperature, wind, pressure, precipitation, evaporation, radiation, and soil moisture. For each GAW station, all variables were extracted from the grid cell containing the station at their native spatial resolutions and used as station-specific predictors. For temporal matching, hourly variables were aligned with the corresponding station observation hour, whereas variables with coarser temporal resolutions were assigned to all hours within their native time interval. We also included static and proxy variables to better represent station characteristics and temporal patterns. Station elevation was obtained from the station metadata. Spatial proxy variables were generated by converting latitude and longitude into radians, providing a continuous



representation of geographic position and enabling the model to capture spatial heterogeneity among stations (Liu and Chen, 2022). Temporal indicators based on hour of day (HOD) and day of year (DOY) were transformed using sine and cosine functions to preserve their periodic continuity and avoid artificial discontinuities (Shao et al., 2026). Together, these predictors enhance the model's ability to capture complex hourly spatiotemporal variability in surface CO₂, with detailed data sources and descriptions provided in Supplement S3 and Table S1.

2.2 GAW CO₂ concentration record reconstruction

To reconstruct temporally continuous hourly CO₂ at GAW stations, we developed a station-based framework that first removes dominant large-scale CO₂ signals and then reconstructs missing residual variability using a two-stage machine learning strategy (Fig. 2). Because atmospheric CO₂ exhibits pronounced long-term growth and seasonal cycling, the original GAW records were adjusted using the CT-derived trend and seasonal components estimated with a Generalized Additive Model (GAM) decomposition (Simpson, 2018) (Supplement S4). CT products represent broad-scale atmospheric CO₂ variability, including interannual trends and seasonal patterns (Cui et al., 2021; Peters et al., 2007). The resulting GAW residual component (CO_{2r}), obtained by subtracting the CT-derived trend (CO_{2t}) and seasonal component (CO_{2s}), was then used as the reconstruction target in the subsequent machine learning procedures. A two-stage reconstruction strategy was employed to address both intra-day incompleteness and long-term missing segments. First, for days with at least one valid observation, missing hours were filled by learning hour-to-hour CO₂ residual differences from paired observations, thereby generating complete 24-hour profiles and a temporally balanced augmented dataset. Second, this augmented dataset was used to train a full-sequence model to estimate hourly CO₂ concentrations for entirely missing days. This design is particularly important for flask and semi-continuous stations, which often sample only once per day or week at fixed local times and thus provide temporally imbalanced training data that cannot resolve the full diurnal cycle. LightGBM was adopted as the core algorithm in both stages due to its efficiency in capturing nonlinear



relationships and discontinuous temporal dynamics in time series (Ke et al., 2017; Anderer and Li, 2022) (Supplement S5 and Table S2).

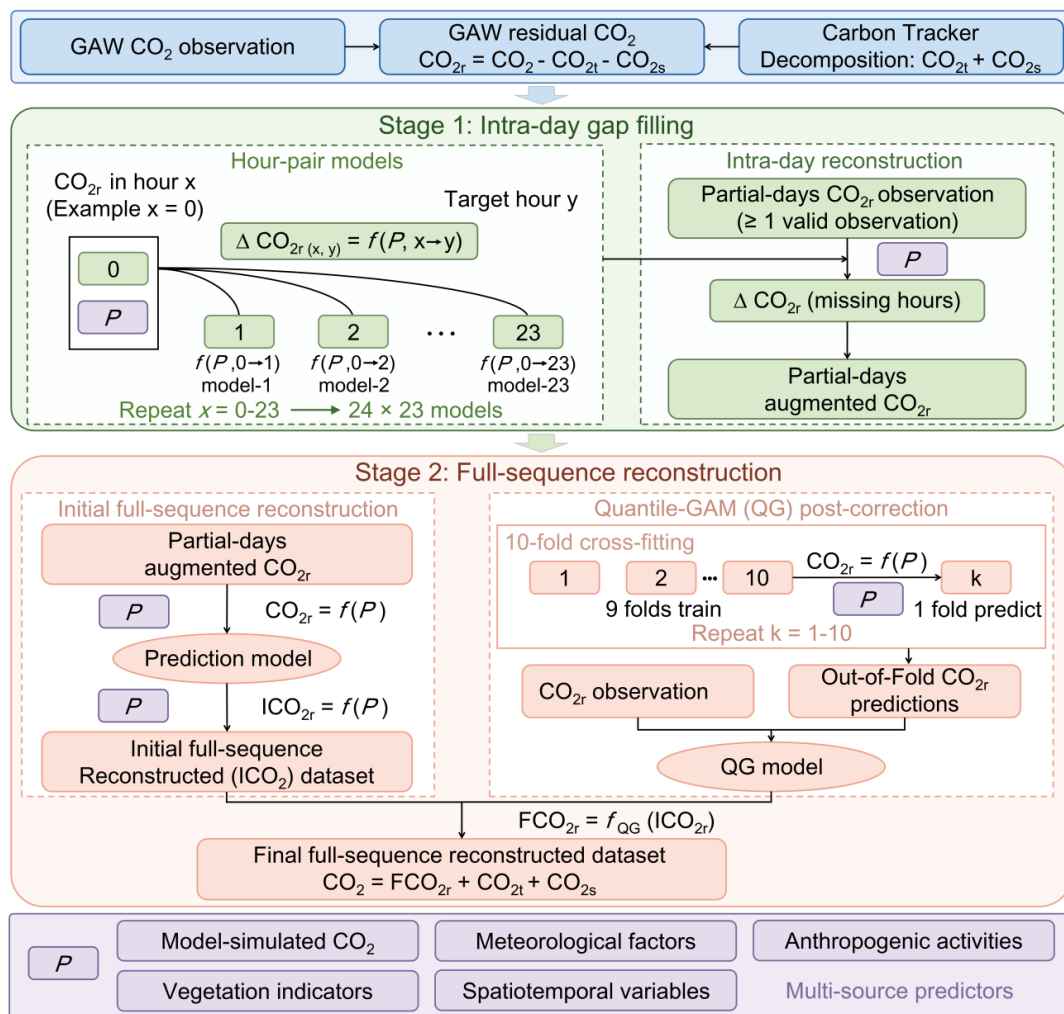


Figure 2. Flowchart of the reconstruction framework for hourly GAW CO₂ concentration records. Original GAW CO₂ observations were first converted into CO₂ residuals by removing the CT-derived long-term trend and seasonal components. The residual component was then reconstructed using a two-stage machine learning framework. In Stage 1, intra-day gap filling was performed for partially observed days with at least one valid observation. In Stage 2, the



augmented residual dataset from Stage 1 was used to reconstruct fully missing periods. The independent variable set included model-simulated CO₂, meteorological factors, anthropogenic activity indicators, vegetation indicators, and spatiotemporal variables.

241

2.2.1 Stage 1: Intra-day gap filling

In this stage, we developed an hour-pair modeling strategy to reconstruct missing hourly CO₂ concentrations within days that contained at least one valid observation. All 24 UTC hours were treated as ordered hour pairs (x, y) , where x denotes the monitoring hour with an available observation and y denotes the target hour to be filled (Figs. 2 and S4). Excluding self-pairs, this yielded 552 ordered hour-pair models. For each (x, y) pair, an independent LightGBM model was trained to predict the CO_{2r} difference defined by Eq. (1).

$$\Delta CO_{2r(x,y)} = CO_{2r,y} - CO_{2r,x} \quad (1)$$

Training samples were extracted from days with valid data at both x and y . GAW residual observations were combined with CT residual samples, calculated by subtracting the CT-derived trend and seasonal component from CT CO₂, to expand the training set, and samples from all stations were pooled to train a global model for each hour pair using Eq. (2).

$$\Delta CO_{2r(x,y)} = f_{(x,y)}(\Delta X_M, \Delta X_C, X_1, X_2, \dots, X_n) \quad (2)$$

where ΔX_M and ΔX_C are hour-to-hour differences in meteorological factors and CAMS CO₂ concentrations, respectively, and $X_1, X_2, \dots, X_n$ denote the anthropogenic, vegetation, static, and spatiotemporal proxy variables described in Section 2.1.3. The function $f_{(x,y)}$ denotes the model specific to the hour pair (x, y) . During reconstruction, this procedure was applied separately to each station. For each partially observed day, the corresponding hour-pair model $f_{(x,y)}$ was used to estimate $CO_{2r,y}$ from the available $CO_{2r,x}$. When multiple monitoring hours were available for the same target hour, their estimates were averaged. This procedure converted partially observed days into complete 24-hour records, providing a temporally balanced augmented dataset for the subsequent full-sequence reconstruction.

264



2.2.2 Stage 2: Full-sequence reconstruction

In this stage, we developed a full-sequence reconstruction strategy to fill time steps that remained completely missing after Stage 1, using LightGBM initial-prediction followed by quantile-GAM (QG) post-correction (Fig. 2 and S5). Using the augmented dataset generated from Stage 1, a LightGBM model was trained to predict the initial CO_{2r} for all missing time steps according to Eq. (3).

$$ICO_{2r} = f(X_{CTa}, X_M, X_C, X_1, X_2, \dots, X_n) \quad (3)$$

where ICO_{2r} is the initial predicted CO_{2r} , X_{CTa} represents the CT-derived CO_2 anomaly term, X_M and X_C denote meteorological factors and CAMS CO_2 , respectively, and $X_1, X_2, \dots, X_n$ represent the anthropogenic, vegetation, static, and spatiotemporal proxy variables described in Section 2.1.3. Because machine learning regressions can overestimate low values and underestimate high values, their predictions may be compressed toward the mean (Belitz and Stackelberg, 2021; Zhang and Lu, 2012). To reduce this bias, we applied a station-specific quantile-GAM post-correction based on the augmented dataset. Out-of-fold predictions were generated through 10-fold cross-fitting, in which calendar-month blocks for each station were randomly assigned to ten folds and iteratively used for training and prediction. These out-of-fold predictions were paired with the corresponding observed CO_{2r} values to estimate a station-specific quantile-GAM relationship f_{QG} between the predicted and observed residual distributions. The corrected CO_{2r} values were obtained using Eq. (4).

$$FCO_{2r} = f_{QG}(ICO_{2r}) \quad (4)$$

where FCO_{2r} denotes the bias-corrected CO_{2r} and f_{QG} represents the station-specific quantile-GAM correction function. Final CO_2 concentrations were obtained by adding back the CT-derived trend and seasonal components. Through this procedure, all remaining missing records were reconstructed, yielding continuous hourly CO_2 time series for each GAW station over 2000–2024.



292 2.2.3 Uncertainty assessment

293 To provide record-level uncertainty estimates for the reconstructed hourly CO₂ dataset, we used
 294 stage-specific validation-based error information to quantify the expected magnitude of the
 295 reconstruction error. For Stage 1, the validation root mean square error (RMSE) of each hour-
 296 pair model was used to represent the typical error in reconstructing target hour y from monitoring
 297 hour x (Fig. S4). RMSE is widely used to summarize prediction error magnitude in forecasting
 298 and model evaluation (Hyndman and Koehler, 2006). When multiple monitoring hours were
 299 available for the same target hour, their pair-specific RMSEs were combined using a quadratic
 300 mean to estimate U_y , as defined in Eq. (5).

$$301 \quad U_y = \sqrt{\frac{1}{n} \sum_{i=1}^n RMSE_{x_i,y}^2} \quad (5)$$

302 where n is the set of available monitoring hours used to reconstruct target hour y . For Stage 2,
 303 uncertainty was estimated from out-of-fold residual errors generated by the 10-fold cross-fitting
 304 procedure (Borjigen et al., 2026) (Fig. S5). Specifically, absolute residual errors were calculated
 305 between the reference CO_{2r} values in the Stage 1 augmented dataset and their corresponding out-
 306 of-fold predictions (Eq. (6)).

$$307 \quad E = |CO_{2r} - OOF_{2r}| \quad (6)$$

308 where OOF_{2r} denotes the out-of-fold predicted CO_{2r}. These absolute errors were then used as
 309 the response variable to train a LightGBM uncertainty model, which predicted the expected
 310 absolute reconstruction error for the Stage 2 missing time steps. In the final reconstructed
 311 dataset, each record was assigned an uncertainty value expressed as an absolute CO₂ error in
 312 ppm, which represents the expected magnitude of reconstruction error for that hourly estimate.

313

314 2.3 Model evaluation and interpretation

315 2.3.1 Model evaluation

316 We evaluated the two-stage reconstruction using validation designs tailored to each task,
 317 including intra-day prediction from observed hour pairs, full-sequence reconstruction under
 318 random and chronological station-month splits, and monthly-scale comparison with the WDCGG



319 benchmark. For intra-day gap filling, validation used 99 stations that had sufficient within-day
 320 observations, including continuous stations and stations with both continuous and flask
 321 sampling. On days with at least two valid hourly observations, each observed hour was in turn
 322 used as the monitoring hour, while the remaining observed hours were masked, reconstructed
 323 with the corresponding hour-pair models, and compared with the withheld observations. For full-
 324 sequence reconstruction, validation used the Stage 1 augmented dataset and was performed by
 325 station-month blocks to avoid information leakage. Under the random month-based split, station-
 326 month blocks within each station were assigned to training, validation, and testing subsets at
 327 proportions of 80 %, 10 %, and 10 %, respectively. Under the chronological split, the earliest
 328 80 % of months were used for training, the next 10 % for validation, and the most recent 10 %
 329 for testing. In both designs, the validation subset was used for model selection and performance
 330 monitoring during training, whereas the testing subset was held out for independent
 331 performance evaluation. As a monthly-scale benchmark, the WDCGG gap-filling procedure
 332 (Supplement S6) was applied to the same withheld station-month subsets, and its reconstructed
 333 monthly CO₂ values were compared with monthly means aggregated from the reconstructed
 334 hourly series (Tsutsumi et al., 2009).

335

336 **2.3.2 Model interpretation**

337 Predictor importance was evaluated using the split gain derived from the LightGBM models.
 338 Split gain represents the cumulative reduction in the training objective attributable to all
 339 decision-tree splits involving a given predictor. For each model, the gain associated with each
 340 predictor was divided by the total gain across all predictors and expressed as a percentage,
 341 hereafter referred to as normalized split gain (Saber et al., 2022). For intra-day gap filling,
 342 predictor importance was calculated separately for each of the 552 ordered hour-pair models.
 343 The normalized split gain of each predictor was first calculated within each model and then
 344 averaged across all hour-pair models to characterize its mean relative importance in a typical
 345 intra-day reconstruction model. For full-sequence reconstruction, predictor importance was



346 extracted from the final LightGBM residual model and expressed as the percentage of the total
 347 split gain attributable to each predictor. This metric summarizes the relative contribution of each
 348 predictor to reducing the training objectives when reconstructing CO_{2r} values during fully
 349 missing periods. To facilitate interpretation, individual predictors were further aggregated into
 350 six predictor groups of meteorology, vegetation, anthropogenic factors, location (including
 351 elevation and spatial proxies), time, and CT/CAMS. These gain-based importance measures
 352 indicate the relative extent to which the models relied on different predictors but do not imply
 353 causal relationships with CO_2 variability.

354

355 **2.4 Reconstructed CO_2 dataset analyses**

356 **2.4.1 Multi-scale evaluation of reconstructed CO_2 variability**

357 To evaluate the reconstructed GAW dataset, we compared it with the original GAW observations
 358 and with the CT and CAMS simulations at long-term, seasonal, and diurnal scales. The aim was
 359 to determine whether the reconstruction reliably captured CO_2 variability and provided added
 360 value relative to existing products. Long-term trends and mean seasonal cycles were analyzed
 361 using the WDCGG global analysis framework (Tsutsumi et al., 2009; Wu et al., 2024). This
 362 framework first selects stations suitable for large-scale analysis, derives station-specific trends
 363 and seasonal cycles from monthly mean CO_2 series, fills missing monthly values when needed,
 364 aligns all selected stations to a common monthly analysis period, and then calculates the zonal
 365 and global mean mole fractions, trends, and growth rates (Supplement S6). As the reconstructed
 366 dataset and model simulations were already continuous over their respective analysis periods,
 367 monthly gap filling and alignment to the common analysis period were applied only to the
 368 original GAW observations with incomplete monthly records. We also used this framework to
 369 benchmark monthly-scale performance by comparing monthly means derived from the
 370 reconstructed hourly dataset with those from a conventional monthly gap-filling method on the
 371 same test subsets. Hourly variability was assessed separately using climatological diurnal cycles.
 372 For each station, hourly CO_2 anomalies were calculated by subtracting a station-specific monthly



baseline, defined as the mean CO₂ concentration for each year-month, from each hourly value (Labzovskii et al., 2019). The anomalies were then grouped by local hour from 0 to 23 h to obtain the climatological diurnal cycle; and the diurnal amplitude was calculated as the difference between the maximum and minimum hourly anomalies. Larger amplitudes indicate greater mean intra-day CO₂ variability at a given site or region.

378

2.4.2 Wildfire-event responsiveness and interpretability analysis

To evaluate the response of the reconstructed hourly CO₂ dataset to extreme biomass-burning events, we focused on the 2023 Canadian wildfire season from May to September and selected 12 GAW stations in Canada and the surrounding regions. Hourly CO₂ anomalies from the original observations, reconstructed dataset, and CT simulation were compared between wildfire-affected periods and a non-fire baseline that was January-April and October-December 2023. This within-year baseline provides a reference for background CO₂ anomaly variability under non-fire-season conditions in the same year. Distributional differences were tested using the Wilcoxon rank-sum test, and mean diurnal cycles were calculated to examine changes in sub-daily variability. To support the interpretation of station-level CO₂ enhancements, we analyzed the anomalies with independent indicators of both fire activity and smoke transport, including Sentinel-5P aerosol absorbing index (AAI), MODIS/VIIRS active fire detections, ERA5 wind fields, and CT CO₂ fields. For selected wildfire episodes, these datasets were combined to assess whether the enhanced station-level CO₂ anomalies were spatially consistent with active fire regions and downwind plume transport. Model interpretation was further conducted using SHAP attribution and partial-dependence diagnostics. SHAP contributions were aggregated according to the predictor group definitions described in Section 2.3.2, and fire-period contributions were compared with the non-fire baseline to identify drivers associated with wildfire-related CO₂ anomaly changes.

398

399



400 3. Results and Discussion

401 3.1 Dataset description

402 The reconstructed hourly GAW dataset substantially improved the temporal completeness of
403 station CO₂ records while providing uncertainty estimates that characterized the reliability of
404 reconstructed values. Across the 152 stations, the mean filling rate, calculated as the number of
405 reconstructed records divided by the total number of hourly records in the complete station time
406 series, was 68.9 % (Fig. 3). This indicates that most stations required substantial reconstruction,
407 although the degree of filling varied strongly among sites because of differences in monitoring
408 type and record continuity. The average station-level uncertainty of reconstructed hourly CO₂
409 concentrations was 2.80 ppm. Reconstruction uncertainty varied across latitude bands, station
410 locations, and monitoring types. Stations in the Northern Hemisphere mid-latitudes had the
411 highest uncertainties, with the 30°N-60°N band reaching an average of 3.57 ppm. This higher
412 uncertainty likely reflects the combined effects of dense anthropogenic emissions, strong
413 biospheric seasonality, and heterogeneous atmospheric transport, which increase spatial and
414 temporal CO₂ variability (Liu et al., 2024). Uncertainty also differed among monitoring types.
415 Continuous stations had a higher average uncertainty of 3.32 ppm than flask stations, which
416 averaged 2.09 ppm. This difference may be partly related to gap structure, because some
417 continuous records contain longer consecutive missing periods that require greater temporal
418 extrapolation during reconstruction, whereas flask-based records are more often characterized by
419 shorter or more sparsely distributed gaps.

420

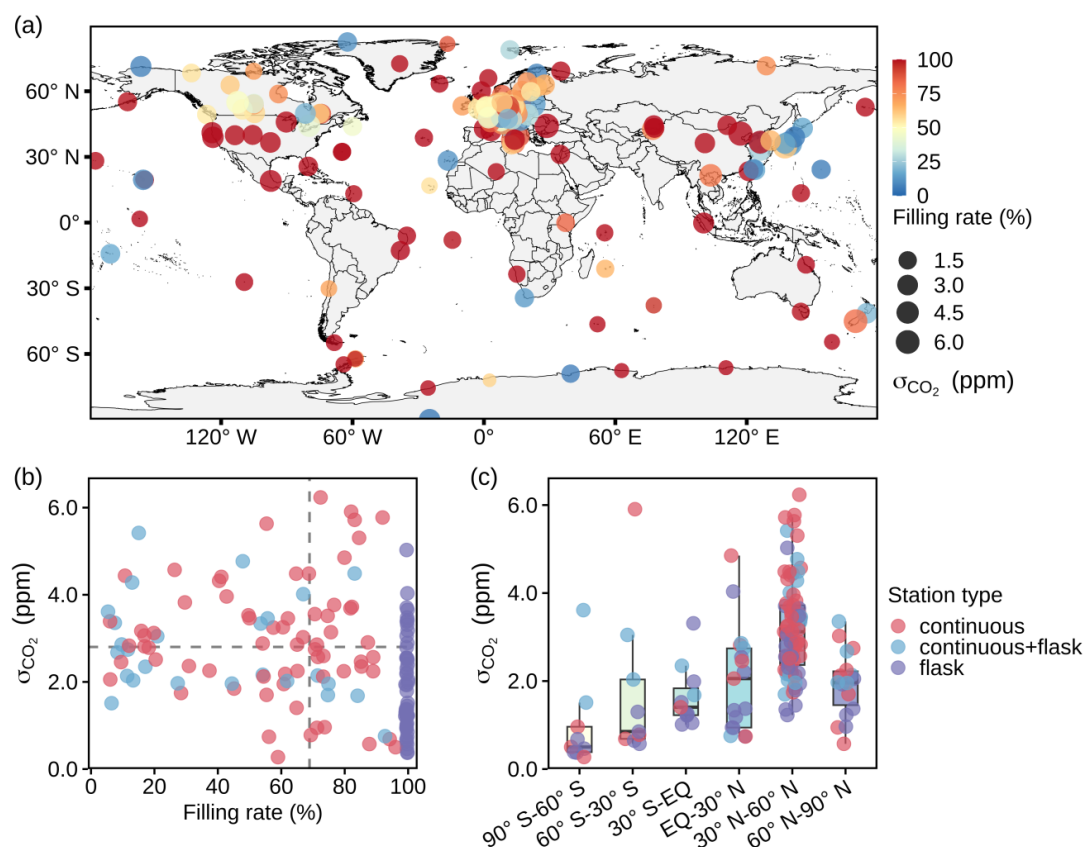


Figure 3. Dataset completeness and reconstruction uncertainty across global GAW CO₂ stations.

(a) Spatial distribution of station-level filling rates and mean reconstruction uncertainty. Point color indicates the filling rate, calculated as the proportion of reconstructed hourly values relative to the complete hourly time series at each station, and point size indicates the station-level mean uncertainty of the reconstructed CO₂ records. (b) Relationship between the station-level filling rate and mean reconstruction uncertainty. Dashed lines indicate the mean filling rate and mean uncertainty across all stations. (c) Distribution of station-level mean reconstruction uncertainty by latitude band. Boxes show the interquartile range and median, and points colored by station type indicate individual stations.

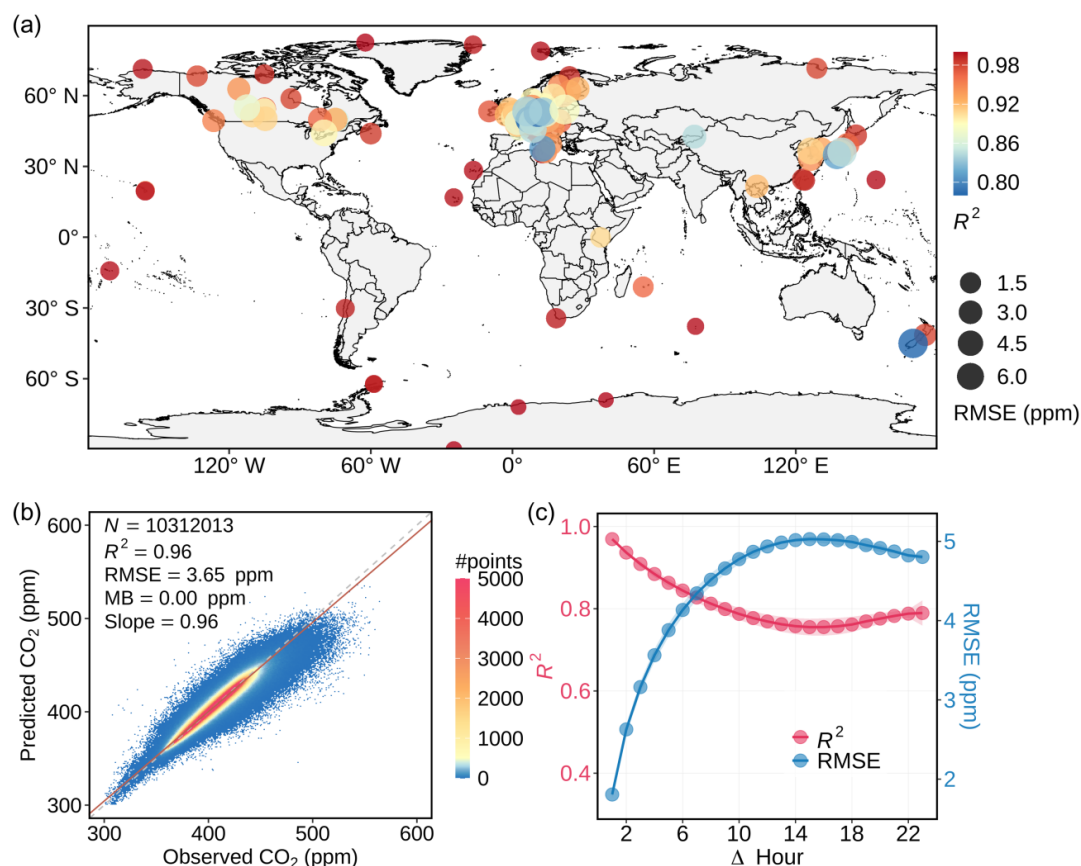


433 3.2 Model evaluation and interpretation

434 3.2.1 Evaluation of intra-day gap filling

435 The validation at 99 stations with continuous measurements, including stations with both
436 continuous and flask sampling, demonstrates the strong performance of the intra-day gap-filling
437 approach (Fig. 4). The hour-pair models perform well at most stations, with R^2 values generally
438 exceeding 0.90 and RMSE values generally below 4 ppm, indicating that the models can
439 effectively reproduce intra-day CO_2 variability across diverse station settings. When validation
440 samples from all stations are pooled, the reconstructed hourly CO_2 concentrations agree closely
441 with the observations, with an overall R^2 of 0.96 and an RMSE of 3.65 ppm. The mean bias is
442 close to 0 ppm, and the regression slope of 0.96 indicates little systematic bias, although with a
443 slight compression of variability. Performance varied with the absolute temporal lag between the
444 observed and target hours. Accuracy is highest for nearby hour pairs and generally decreases as
445 the lag increased, with lower R^2 and higher RMSE at larger separations. Overall, the hour-pair
446 models retain reasonable predictive skill across the full range of temporal lags. Representative
447 continuous observation sites in Asia (Cape Ochiishi, COI), Europe (Pallas, PAL), North America
448 (East Trout Lake, ETL), and at the remote marine background site Mauna Loa (MLO) also show
449 strong validation performance (Fig. S6), supporting the broad applicability of the intra-day gap-
450 filling method across different climatic and observational settings.

451



452

453 **Figure 4.** Validation of intra-day CO_2 reconstruction. (a) Spatial distribution of R^2 and RMSE for
 454 the intra-day gap-filling models at stations with continuous observations and stations with mixed
 455 continuous and flask sampling records, totaling 99 stations. (b) Observed versus predicted hourly
 456 CO_2 concentrations for all validation samples. (c) Model performance as a function of the
 457 absolute temporal lag (Δ Hour) between the observed hour and the target hour, used to assess
 458 how prediction accuracy changes with increasing separation between paired hours. Points
 459 represent mean R^2 and RMSE across hour-pair models with the same Δ Hour.

460

461

462

463



3.2.2 Evaluation of full-sequence reconstruction

The validation using both random and chronological month-based splitting strategies demonstrated strong predictive performance for the full-sequence reconstruction. Under the random month-based split, approximately 80 % of stations achieved R^2 values above 0.8, and about 80 % exhibited RMSE values below 5 ppm (Fig. 5a). Across all held-out samples, the framework achieved an overall R^2 of 0.93 and an RMSE of 4.76 ppm, with further improvement after daily aggregation ($R^2 = 0.97$, RMSE = 3.20 ppm; Fig. 5b, c). Under the more stringent chronological month-based split, the reconstruction remained robust, with an R^2 of 0.89 and an RMSE of 5.47 ppm (Fig. S7). Representative continuous and flask sampling sites across Asia (COI and Mt. Waliguan, WLG), Europe (PAL and Teriberka, TER), North America (ETL and Southern Great Plains, SGP), and remote background regions (MLO and Casey, CYA) showed consistent agreement between reconstructed CO_2 concentrations and held-out observations (Figs. S8 and S9). These results indicate that the framework can reconstruct hourly CO_2 time series across diverse climatic and observational settings. After QG correction, the station-level regression slopes moved closer to unity under both validation strategies, indicating reduced slope bias in the full-sequence reconstruction (Fig. S10a, b). Bias analysis across CO_2 concentration quantiles further showed that mean bias (MB) decreased after correction across most quantile bins, with especially pronounced improvements at the lower and upper tails of the CO_2 distribution (Fig. S10c, d). These results suggest that quantile-based correction alleviates the over-smoothing tendency of the regression model and improves the representation of extreme CO_2 values.

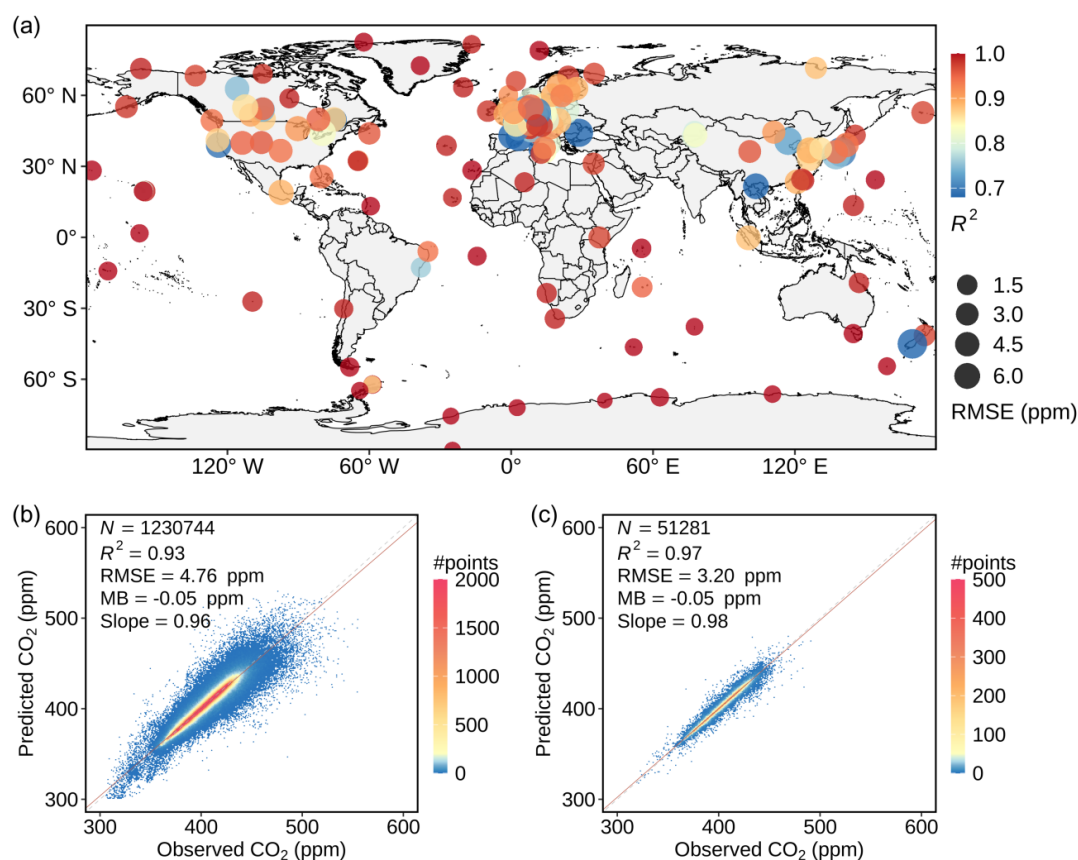


Figure 5. Validation of full-sequence CO_2 reconstruction using random month-based splitting.

(a) Spatial distribution of station-level validation performance. (b-c) Comparisons between observed and predicted CO_2 concentrations at the hourly and daily scales, respectively.

The full-sequence reconstruction framework also performed well after aggregation to monthly means, and it outperformed the WDCGG monthly gap-filling procedure on the same withheld station-month subsets (Supplement S6 and Fig. S11). Under the random month-based split, monthly means from the reconstructed hourly series agreed slightly better with observations than those from the WDCGG method, with $R^2 = 0.99$ and $\text{RMSE} = 1.71$ ppm compared to $R^2 = 0.98$ and $\text{RMSE} = 2.25$ ppm (Fig. S11a, c). Under the chronological split, the advantage was even larger, with $R^2 = 0.96$ and $\text{RMSE} = 2.45$ ppm for the proposed framework compared to $R^2 = 0.90$



and RMSE = 4.02 ppm for the WDCGG method (Fig. S11b, d). The weaker WDCGG performance under temporal extrapolation likely reflects its reliance on the fitted long-term trend, mean seasonal cycles, and zonal growth rates from long-term stations in the same latitude band (Tsutsumi et al., 2009), which assumes relatively stable inter-site relationships despite known temporal changes and gives limited representation to short-term variability (Wu et al., 2024; Masarie and Tans, 1995). By integrating simulated CO₂ fields, environmental predictors, and station-level observational constraints, the proposed framework better preserves month-to-month variability while maintaining strong performance at the native hourly scale.

3.2.3 Model interpretation

The gain-based predictor importance analysis showed that the two-stage reconstruction framework relied primarily on CT/CAMS atmospheric information, with local environmental and station-specific predictors further refining CO_{2r} variability (Fig. S12). In the intra-day gap-filling stage, CT provides the long-term trend and seasonal components used to define the CO_{2r} target, and CT-derived residual hour pairs were used to augment the training data. Among the predictors directly included in the 552 hour-pair models, CAMS was the most important predictor group, accounting for 38.5 % of the normalized split gain. Location-related, meteorological, anthropogenic, temporal, and vegetation predictors accounted for 18.2 %, 16.0 %, 13.3 %, 8.4 %, and 5.7 %, respectively. In the full-sequence reconstruction stage, which estimated CO_{2r} during fully missing periods, CT/CAMS remained the most important, accounting for 29.3 % of the normalized split gain. This contribution was driven mainly by the CT anomaly, the most important individual predictor in Stage 2, together with CAMS CO₂. Anthropogenic variables formed the second most important predictor group, contributing 23.3 %, followed by meteorology (16.3 %), location (12.1 %), time (10.2 %), and vegetation variables (8.9 %). These results are consistent with the design of the reconstruction framework, in which CT characterizes the background trend and seasonality, while CAMS and auxiliary covariates help resolve sub-daily and site-specific deviations from that background. Because split-gain



importance reflects the extent to which the models used each predictor to reduce the training objective, these results should not be interpreted as evidence of causal effects on atmospheric CO₂ variability.

3.3 CO₂ variability in the reconstructed GAW dataset

3.3.1 Improved sample representativeness after intra-day gap filling

The intra-day gap-filling procedure reduced sampling-time bias by expanding the sparsely sampled records into more complete hourly profiles. Across all stations, the number of hourly CO₂ samples increased from 10,360,507 to 11,981,280, corresponding to a 15.6 % increase. The largest relative gain occurred at flask-sampling stations, where the sample count increased from 43,826 to 1,014,792, reaching 23.2 times the original number. This effect is illustrated by two stations with contrasting sampling patterns. At WLG in western China, records increased from 1,026 to 24,192, and early-morning observations at 06:00-07:00 local time, which tend to capture elevated CO₂ under shallow and stable boundary-layer conditions, were supplemented with estimated lower-concentration values, which produced a more balanced distribution and more complete diurnal coverage in 2023 (Fig. S13a, b). In contrast, SGP in the central United States was mainly sampled at 13:00-14:00 local time, when stronger boundary-layer mixing and daytime biospheric uptake often lead to relatively lower near-surface CO₂ concentrations. After augmentation, the number of records increased from 1,017 to 24,408, and otherwise missing higher-concentration periods were estimated, broadening the distribution and producing continuous seasonal diurnal cycles with a weaker winter amplitude (Fig. S13c, d). Overall, intra-day gap filling improved the temporal representativeness of sparse observations and provided a more complete hourly basis for subsequent full-sequence reconstruction.

3.3.2 Effects of the reconstruction on CO₂ trends and seasonal estimates

The reconstructed GAW dataset provides more temporally complete station-based estimates of global and zonal CO₂ trends and seasonal cycles, while preserving the main variability in the



original observations. In the WDCGG global analysis framework comparison for 2001-2023, which was selected to limit filtering-related edge effects, the global monthly mean CO₂ concentrations derived from the reconstructed dataset were 0.54 ± 0.21 ppm higher than those derived from the original GAW records on average (Fig. 6). The reconstructed dataset also improves temporal continuity without substantially altering the growth-rate signal. Relative to the original GAW observations, the mean absolute difference in monthly growth rate was 0.10 ± 0.09 ppm yr⁻¹ for the reconstructed dataset. Similar patterns across most latitude zones further indicated that the reconstruction provides more complete station-based global and zonal estimates while maintaining consistency with the observed growth-rate variability (Fig. S14). In addition, the reconstructed dataset retained the observed latitudinal structure of the mean seasonal CO₂ cycle (Fig. S15). In both the original and reconstructed GAW records, seasonal amplitudes were largest in the Northern Hemisphere and became weaker toward the tropics and the ocean-dominated Southern Hemisphere. Differences between the reconstructed and original seasonal cycles were small in most latitude bands, with the largest mean absolute difference occurring at 30°-60°N (0.35 ± 0.21 ppm), which was still minor relative to the observed seasonal amplitude. These results indicate that reconstruction improves the temporal completeness of station records without substantially changing the observed growth-rate variability or the phase and amplitude of the mean seasonal cycle.

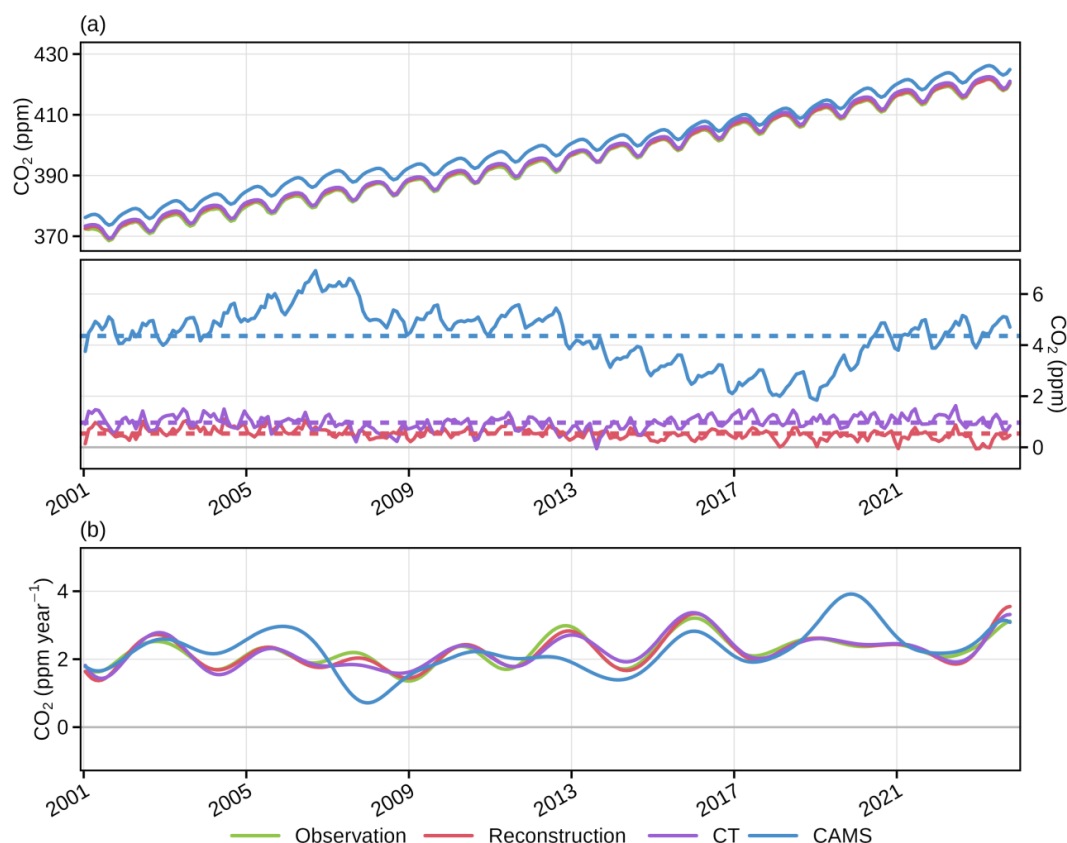


Figure 6. Global monthly mean CO₂ mole fractions and monthly CO₂ growth rates derived using the WDCGG global analysis framework. (a) Area-weighted global monthly mean CO₂ mole fractions derived from the original GAW observations, reconstructed GAW records, CT, and CAMS. The upper panel shows the global monthly mean CO₂ mole fractions, and the lower panel shows the differences between each reconstructed or model-based dataset and the original GAW observations. Dashed horizontal lines indicate the mean differences. (b) Global monthly CO₂ growth rates derived from the same four data sources. Global means were calculated by first aggregating station-level WDCGG products within 30° latitude bands and then applying latitude-band area weighting (Tsutsumi et al., 2009). CT and CAMS CO₂ values were extracted from the model grid cells containing each GAW station.



582 The offsets among the original GAW records, reconstruction, CT, and CAMS reflect differences
 583 in temporal sampling and model-simulated product observational constraints. The lower global
 584 monthly mean CO₂ derived from the original GAW records using the WDCGG global analysis
 585 framework likely results from the monthly means being calculated only from available hourly
 586 observations, making them susceptible to biases caused by missing data, uneven temporal
 587 coverage, and fixed-time sampling (Tsutsumi et al., 2009). In addition, the WDCGG global
 588 analysis framework is sensitive to station distribution, site selection, and data coverage within
 589 latitude bands, which can affect global mean surface CO₂ estimates (Wu et al., 2024). CT and
 590 CAMS were used here as model references by extracting the simulated CO₂ values at the GAW
 591 station locations, rather than as benchmarks for ranking product performance (Fig. 6). CT agreed
 592 well with the station observations and reconstruction, with a mean offset of 0.96 ± 0.28 ppm and
 593 a monthly growth-rate difference of 0.13 ± 0.09 ppm yr⁻¹ relative to the original GAW records,
 594 consistent with its assimilation of surface in situ CO₂ observations (Cui et al., 2021). CAMS
 595 showed larger differences from the GAW station records, with a mean offset of 4.36 ± 1.10 ppm
 596 and a growth-rate difference of 0.41 ± 0.38 ppm yr⁻¹, likely reflecting the differences in
 597 observational constraints, including XCO₂ retrievals, atmospheric transport, and prior flux
 598 information (Agusti-Panareda et al., 2023). Thus, the CT/CAMS comparison provides station-
 599 level context, whereas the main result is that hourly reconstruction reduces sampling-related
 600 differences in the global and zonal estimates derived from GAW station records.

601

602 Sampling-related changes were most evident at the flask-sampling stations, where observations
 603 are sparse and often concentrated at specific hours (Fig. S16). After reconstruction, the mean
 604 growth rate across these stations increased from 2.21 ± 0.14 ppm yr⁻¹ to 2.25 ± 0.11 ppm yr⁻¹.
 605 The largest increase occurred in the 30°-60° N band, where 23 flask-sampling stations were
 606 available and the mean growth rate increased by 0.08 ± 0.19 ppm yr⁻¹, indicating substantial
 607 station-to-station variability. Most other latitude bands showed smaller increases, whereas the
 608 60°-90° N band showed a slight decrease. Seasonal amplitude also changed after reconstruction,



with the mean amplitude across flask-sampling stations declining from 9.00 ± 5.88 ppm to 8.62 ± 5.51 ppm. This reduction was driven mainly by the 30° - 60° N band, where the mean amplitude dropped by 0.95 ± 2.26 ppm, although slight increases occurred in the northern tropics and polar bands. The stations with larger adjustments in this band are mostly located in Northern Hemisphere continental or coastal regions, where stronger biospheric seasonality, anthropogenic influences, and boundary-layer variability can amplify the sensitivity of monthly and seasonal CO_2 estimates to sparse or fixed-time sampling (Munassar et al., 2025). These patterns indicate that sparse, time-specific flask sampling can alter growth-rate and seasonal-cycle estimates derived directly from irregular observations, particularly in the Northern Hemisphere mid-latitudes. By providing more complete hourly records for monthly aggregation, the reconstruction reduces sampling-related differences while preserving the broad latitudinal structure of CO_2 trends and seasonal cycles.

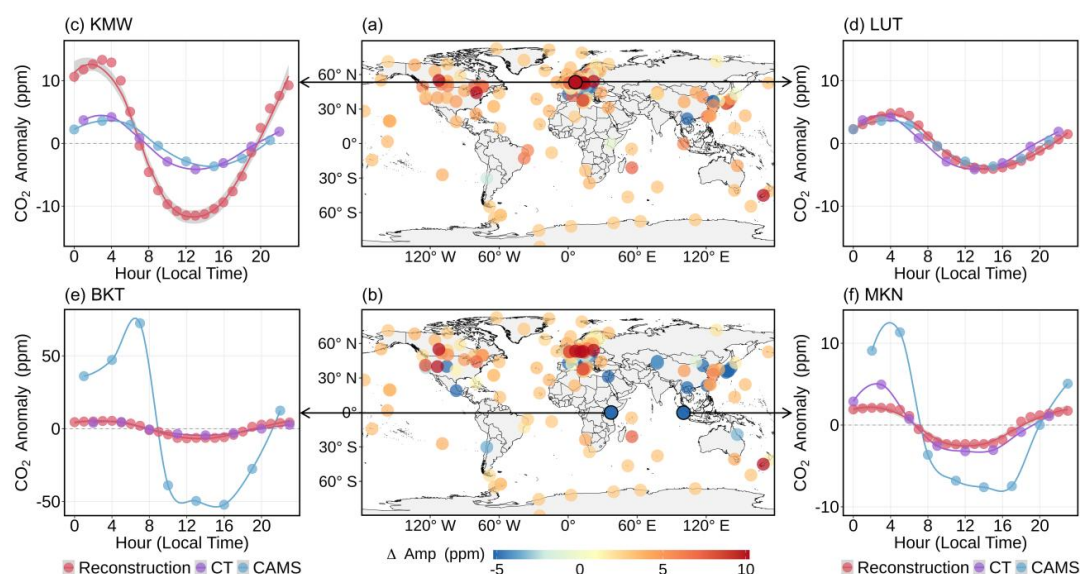
3.3.3 Representation of diurnal cycles

The reconstructed GAW hourly CO_2 dataset provides complete station-level diurnal profiles while retaining stronger local diurnal signals than the CT and CAMS grid-cell simulations. Flask-sampling stations are too sparsely sampled to allow the direct characterization of full diurnal cycles, whereas well-sampled continuous stations can capture long-term mean diurnal variability. Among continuous stations with observational completeness above 50 %, the reconstructed and original diurnal amplitudes differed by only 0.44 ± 0.53 ppm on average, indicating that the reconstruction largely preserves the observed mean diurnal structure. About 80 % of the stations showed larger diurnal amplitudes in the reconstructed dataset than in CT and CAMS (Fig. 7a, b). The differences were relatively small at high-latitude and remote equatorial or marine-influenced stations, where local diurnal CO_2 signals are generally weak (Yuan et al., 2018). The largest discrepancies occurred at terrestrial sites in the Northern Hemisphere mid-latitudes, especially within 30° - 60° N, where strong biospheric exchange and anthropogenic emissions generate pronounced diurnal CO_2 cycles (Tang et al., 2025). In model-simulated



products, these station-scale signals can be dampened by grid-cell averaging, transport mixing, and boundary-layer representation, leading to smaller apparent diurnal amplitudes (Ahmadv et al., 2009; Pillai et al., 2010). However, the reconstructed amplitudes at some stations were smaller than those from CT or CAMS, suggesting differences in representativeness between point-scale observations and grid-cell averages. The Kollumerwaard (KMW) and Lutjewad (LUT) stations in Europe illustrate this issue, as they fall within the same CT and CAMS grid cells but show distinct reconstructed diurnal amplitudes (Fig. 7c, d).

643



644

Figure 7. Station-scale diurnal CO₂ variability in the reconstructed GAW dataset compared with CT and CAMS sampled at GAW station locations. (a, b) Spatial distributions of differences in mean diurnal CO₂ amplitude (Δ Amp) between the reconstructed dataset and (a) CT or (b) CAMS. Δ Amp was calculated as the reconstructed diurnal amplitude minus the corresponding CT or CAMS amplitude. Positive values indicate stronger diurnal variability in the reconstructed dataset. CT and CAMS values were extracted from the model grid cells containing each GAW station. (c-f) Representative mean diurnal CO₂ anomaly cycles at Kollumerwaard (KMW), Lutjewad (LUT), Kototabang (BKT), and Mt. Kenya (MKN), respectively, comparing the reconstructed dataset with CT and CAMS in local time.



654

655 Global model-simulated CO₂ products are optimized to constrain large-scale atmospheric carbon
 656 variability, but they may dampen fine-scale signals in surface CO₂ records. The CAMS
 657 greenhouse gas reanalysis product EGG4 primarily assimilates XCO₂ rather than surface-based
 658 station observations, so its near-surface CO₂ fields depend strongly on prior emission inventories
 659 and atmospheric transport parameterizations (Agustí-Panareda et al., 2023). This may limit its
 660 ability to capture complex local surface diurnal cycles, as shown at the global background Bukit
 661 Kototabang (BKT) and Mt. Kenya (MKN) stations (Fig. 7e, f). CT assimilates station-based
 662 observations, but its filtering and sampling strategies emphasize regionally representative
 663 background conditions. For example, hourly observations are assimilated only during selected
 664 periods at some sites, and values strongly affected by local pollution or deviations from smooth
 665 synoptic-scale variability are excluded (Masarie et al., 2014; Jacobson, 2023). These criteria are
 666 appropriate for large-scale inversion constraints, but they can reduce the representation of all-
 667 hour station variability, including nocturnal CO₂ signals and diurnal transitions. In addition, the
 668 relatively coarse spatial resolution of CT limits its ability to resolve sub-grid processes, such as
 669 urban plumes and localized anthropogenic emission events. By providing temporally seamless
 670 hourly station-level CO₂ records, the reconstructed dataset can serve as a more complete
 671 benchmark for evaluating inversion models and may support the broader use of surface
 672 observations as temporally resolved atmospheric constraints.

673

674 **3.4 Capturing CO₂ responses to wildfire events**

675 The 2023 Canadian wildfire season provides a stringent event-scale test of the reconstructed
 676 dataset because it generated strong, short-lived, and spatially heterogeneous CO₂ enhancements
 677 while many regional GAW records were incomplete. All 12 GAW stations in Canada were
 678 included in the analysis (Table S3). Four stations had no valid observations during the wildfire
 679 period from May to September, and the remaining eight had an average data coverage of only
 680 64.5 %, meaning that the original records alone could not fully characterize the events. Under



these sparse observational conditions, the reconstructed dataset provided continuous hourly anomaly estimates across the regional station network and captured fire-period peaks that were consistent with the timing and magnitude of available observations (Fig. S17). Fire-period anomalies were clearly elevated relative to the non-fire baseline in both observations and the reconstruction, while CT showed a smaller and statistically less distinct shift (Fig. 8a). Diurnal composites reinforced this contrast, with the reconstruction closely matching the observed peak timing, minima, and amplitudes during wildfire periods, whereas CT underestimated both sub-daily variability and extreme anomalies (Fig. 8b). This pattern partly reflects the larger surface CO₂ diurnal variability expected during summer because of stronger biospheric exchange and boundary-layer cycling (Xueref-Remy et al., 2018). Wildfire emissions can further amplify this cycle, as shallow nighttime boundary layers favor near-surface accumulation of fire related CO₂, while daytime mixing reduces accumulation, but smoke stress may weaken ecosystem uptake, allowing the afternoon anomaly minima to remain above CT. The weaker CT response likely reflects its use of climatological fire-flux information in CT-NRT, together with the dilution of localized plume signals within coarse model grid cells (Chen et al., 2019; Jacobson, 2023).

The reconstructed wildfire anomalies were also spatially consistent with independent indicators of fire activity and smoke transport. On 14 July 2023 and 2 September 2023, elevated reconstructed station-level CO₂ concentrations aligned with S5P AAI, MODIS/VIIRS active fire detections and wind fields (Fig. 8c, d). Rather than being randomly distributed, high CO₂ concentrations occurred mainly at stations located downwind of active fire regions or near high-AAI smoke plumes. The CT surface CO₂ fields captured broad regional enhancement patterns, but their highest concentrations were less consistently collocated with the station-level smoke-transport signatures. On 14 July, widespread fire activity in western and central Canada coincided with plume transport along the prevailing winds, and LLB and EST showed elevated CO₂ concentrations near the transport pathway. On 2 September, fire detections and smoke plumes again indicated plume transport along the wind direction, with elevated reconstructed



concentrations at LLB, EST, and CDL. Because the raw GAW records were missing or temporally incomplete at several stations during these episodes, they could not continuously characterize the wildfire influence. By estimating missing hourly values while retaining constraints from available observations, the reconstruction provides a more complete station-based view of wildfire-related CO₂ enhancements and transport signatures.

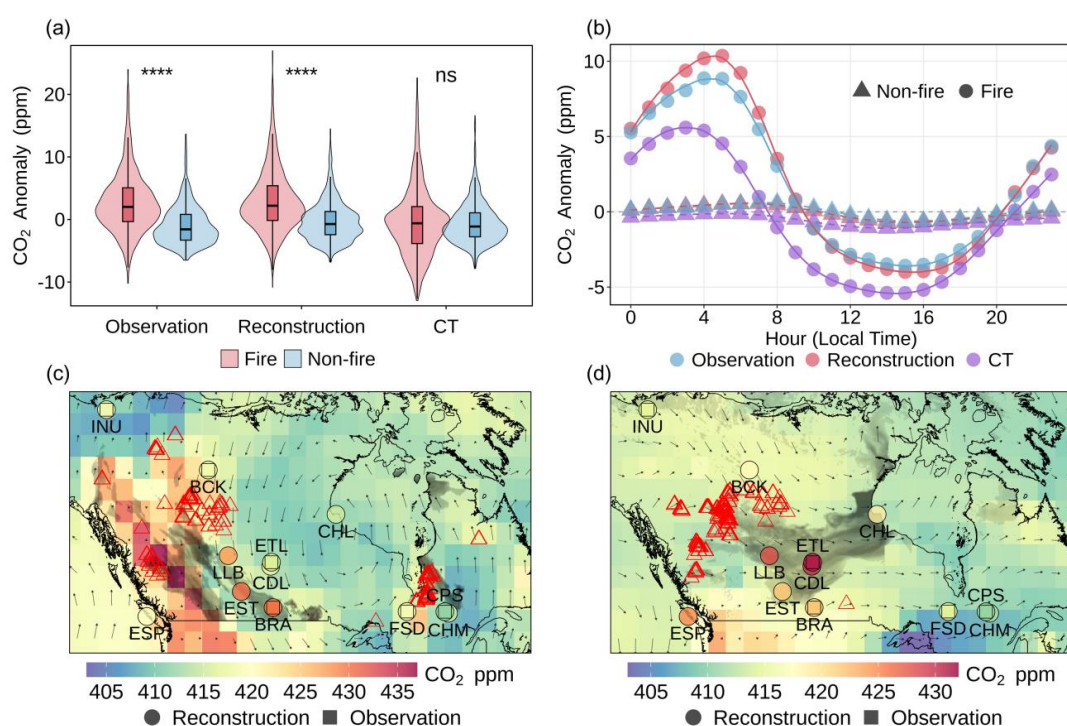


Figure 8. Comparisons of GAW hourly CO₂ observations, the reconstructed dataset, and CT surface CO₂ during major Canadian wildfire episodes in 2023. (a) Distribution of hourly CO₂ anomalies during the fire period, defined as May–September, and the non-fire period, defined as January–April and October–December. Horizontal markers indicate the statistical significance of differences in CO₂ anomalies between the fire and non-fire periods, determined using the Wilcoxon rank-sum test (**** $P < 0.0001$, *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$, and ns (nonsignificant) at $P \geq 0.05$). (b) Mean diurnal cycles of CO₂ anomalies. In (a) and (b), CT values were extracted from the model grid cells containing each GAW station. Reconstructed and



723 CT anomalies were calculated relative to their station-specific monthly means, whereas observed
 724 anomalies were calculated using the corresponding reconstructed monthly means as the baseline
 725 because the original observations were sparse. (c-d) Spatial distributions of daily mean surface
 726 CO₂ concentrations on 14 July and 2 September 2023, respectively, shown together with wildfire
 727 hotspots (indicated by red triangles), AAI smoke plumes, and ERA5 wind fields. (c, d) Absolute
 728 surface CO₂ concentrations rather than anomalies to illustrate regional concentration gradients
 729 and plume-related spatial patterns on specific wildfire days. Circles indicate reconstructed
 730 station-level daily means, and squares indicate stations with available observations.

731

732 Model diagnostics can further explain how the reconstruction represented these wildfire-related
 733 anomalies. Across the 12 Canadian GAW stations, SHAP attribution showed that meteorological
 734 and vegetation-related predictors generally increased the reconstructed CO₂ anomalies, whereas
 735 the CT/CAMS predictor group contributed weakly or negatively (Fig. S18). The largest positive
 736 changes were associated with temperature, FPAR/NDVI, and seasonal or spatial proxy variables
 737 (Fig. S19), indicating that the model used environmental conditions characteristic of the wildfire
 738 season to adjust station-scale residual variability. In contrast, the weak or negative CT/CAMS
 739 contributions suggest that these large-scale products primarily represent seasonal biospheric
 740 drawdown and smoothed background variability rather than localized fire-related CO₂
 741 enhancements, consistent with the climatological CT-NRT fire prior and coarse-resolution plume
 742 dilution (Jacobson, 2023; Brunner et al., 2023). Partial-dependence curves at representative
 743 stations also showed stronger predicted CO₂ anomalies under high temperature and FPAR
 744 conditions (Figs. S20 and S21). These diagnostics describe model behavior rather than direct
 745 physical causality. Together with the temporal, distributional, diurnal, and spatial evidence
 746 above, they support the interpretation that the reconstructed hourly GAW dataset can recover
 747 coherent station-scale biomass-burning signals even when individual station records are
 748 incomplete.

749



750 4. Conclusions

751 This study presents a temporally continuous, global hourly CO₂ dataset for GAW stations
 752 covering 2000-2024. By integrating original station observations with CT and CAMS CO₂ fields
 753 and meteorological, vegetation, anthropogenic, and spatiotemporal predictors, the proposed
 754 framework reconstructs missing observations to generate complete hourly time series while
 755 providing uncertainty estimates for the reconstructed values. The resulting data product
 756 addresses an important limitation of the original GAW records, whose use in many applications
 757 is constrained by irregular sampling and extended data gaps. The reconstructed dataset enables
 758 more consistent characterization of global and regional surface CO₂ variability, evaluation of the
 759 effects of incomplete sampling on aggregated statistics, and analysis of short-lived CO₂
 760 enhancements, including those associated with wildfire events. It therefore complements the
 761 original high-quality GAW observations by providing an observationally constrained, temporally
 762 complete hourly data product for carbon-cycle research. Potential applications include evaluating
 763 atmospheric transport and CO₂ simulations, quantifying surface CO₂ variability from diurnal to
 764 interannual timescales, investigating regional transport and emission events, and supporting
 765 atmospheric inversions and carbon-monitoring systems that require temporally continuous
 766 surface CO₂ information.

768 Code and data availability

769 The reconstructed GAW station CO₂ dataset and the source code are available on Zenodo via
 770 <https://doi.org/10.5281/zenodo.21306261> (Mi et al., 2026). The data package contains station-
 771 level NetCDF files global GAW CO₂ monitoring stations from 2000 to 2024, which provided
 772 data at hourly, daily, and monthly temporal resolutions. The hourly files include station
 773 information, CO₂ mole fractions, data-type labels (observation and reconstruction), and model-
 774 based uncertainty estimates. The daily and monthly files are aggregated from the hourly product
 775 and provide mean CO₂ mole fractions. The code used to generate the dataset is provided as
 776 Code.zip in the same repository. It contains two scripts implementing the two-stage GAW CO₂



777 reconstruction workflow. Detailed comments are included in the scripts to facilitate
778 understanding, reproduction, and further development of the workflow.

779

780 **Supplement**

781 This file provides supplementary tables and figures supporting the reconstruction of the global
782 hourly GAW CO₂ dataset. It also documents the ancillary variables and data sources, CT/CAMS
783 data processing, CO₂ residual decomposition, model parameter settings, WDCGG global
784 analysis framework, and procedures for model evaluation and interpretation.

785

786 **Team list**

787 No team authors are included in this manuscript.

788

789 **Author contributions**

790 Tan Mi: Data curation, Methodology, Visualization, Validation, Writing- Original draft
791 preparation. Xueting Pu: Visualization. Xi Zheng: Methodology. Guangcheng Nie: Validation.
792 Xinyi Liu: Investigation. Han Zhang: Investigation. Michael L. Grieneisen: Writing- Reviewing
793 and Editing. Yu Zhan: Investigation, Supervision, Writing- Reviewing and Editing, Software.
794 Maohua Wang: Investigation, Supervision. Fumo Yang: Supervision.

795

796 **Competing interests**

797 The contact author has declared that none of the authors has any competing interests.

798

799 **Disclaimer**

800 No additional disclaimer is required beyond the standard Copernicus Publications disclaimer.

801

802 **Acknowledgements**

803 This work was supported by the Science and Technology Project of State Grid Corporation of



804 China (No. 1400-202426279A-1-1-ZH). We thank the World Data Centre for Greenhouse Gases
 805 (WDCGG) and the contributing GAW institutions for providing the original CO₂ observation
 806 records. We also thank Xin Zheng from the College of Carbon Neutrality Future Technology,
 807 Sichuan University, for assistance with experimental detection and analysis.

808

809 **Financial support**

810 This work was supported by the Science and Technology Project of State Grid Corporation of
 811 China (No. 1400-202426279A-1-1-ZH).

812

813 **References**

- 814 Agustí-Panareda, A., Barré, J., Massart, S., Inness, A., Aben, I., Ades, M., Baier, B. C., Balsamo,
 815 G., Borsdorff, T., Bousserez, N., Boussetta, S., Buchwitz, M., Cantarello, L., Crevoisier, C.,
 816 Engelen, R., Eskes, H., Flemming, J., Garrigues, S., Hasekamp, O., Huijnen, V., Jones, L.,
 817 Kipling, Z., Langerock, B., McNorton, J., Meilhac, N., Noël, S., Parrington, M., Peuch, V. H.,
 818 Ramonet, M., Razinger, M., Reuter, M., Ribas, R., Suttie, M., Sweeney, C., Tarniewicz, J., and
 819 Wu, L.: Technical note: The CAMS greenhouse gas reanalysis from 2003 to 2020, *Atmos. Chem.*
 820 *Phys.*, 23, 3829-3859, <https://doi.org/10.5194/acp-23-3829-2023>, 2023.
- 821 Ahmadov, R., Gerbig, C., Kretschmer, R., Körner, S., Rödenbeck, C., Bousquet, P., and
 822 Ramonet, M.: Comparing high resolution WRF-VPRM simulations and two global CO₂ transport
 823 models with coastal tower measurements of CO₂, *Biogeosciences*, 6, 807-817,
 824 <https://doi.org/10.5194/bg-6-807-2009>, 2009.
- 825 Anderer, M. and Li, F.: Hierarchical forecasting with a top-down alignment of independent-level
 826 forecasts, *Int. J. Forecast.*, 38, 1405-1414, <https://doi.org/10.1016/j.ijforecast.2021.12.015>, 2022.
- 827 Belitz, K. and Stackelberg, P. E.: Evaluation of six methods for correcting bias in estimates from
 828 ensemble tree machine learning regression models, *Environ. Model. Softw.*, 139, 105006,
 829 <https://doi.org/10.1016/j.envsoft.2021.105006>, 2021.
- 830 Borjigen, S., Jin, X., Shao, M., Li, L., Nie, Z., Zha, Y., and Smith, W. L.: Generating seamless
 831 hourly XCO₂ maps over Asia: A CatBoost modeling approach integrating FY-4A GIIRS and
 832 OCO-2/3 (2021-2022), *J. Cleaner Prod.*, 563, 148493,
 833 <https://doi.org/10.1016/j.jclepro.2026.148493>, 2026.
- 834 Brunner, D., Kuhlmann, G., Henne, S., Koene, E., Kern, B., Wolff, S., Voigt, C., Jöckel, P.,
 835 Kiemle, C., Roiger, A., Fiehn, A., Krautwurst, S., Gerilowski, K., Bovensmann, H., Borchardt, J.,
 836 Galkowski, M., Gerbig, C., Marshall, J., Klonecki, A., Prunet, P., Hanfland, R., Pattantyús-
 837 Ábrahám, M., Wyszogrodzki, A., and Fix, A.: Evaluation of simulated CO₂ power plant plumes
 838 from six high-resolution atmospheric transport models, *Atmos. Chem. Phys.*, 23, 2699-2728,
 839 <https://doi.org/10.5194/acp-23-2699-2023>, 2023.



- 840 Chen, H. W., Zhang, L. N., Zhang, F., Davis, K. J., Lauvaux, T., Pal, S., Gaudet, B., and
- 841 DiGangi, J. P.: Evaluation of Regional CO₂ Mole Fractions in the ECMWF CAMS Real-Time
- 842 Atmospheric Analysis and NOAA CarbonTracker Near-Real-Time Reanalysis With Airborne
- 843 Observations From ACT-America Field Campaigns, *J. Geophys. Res. Atmos.*, 124, 8119–8133,
- 844 <https://doi.org/10.1029/2018JD029992>, 2019.
- 845 Cheng, D., Reiner, D. M., Yang, F., Cui, C., Meng, J., Shan, Y., Liu, Y., Tao, S., and Guan, D.:
- 846 Projecting future carbon emissions from cement production in developing countries, *Nat.*
- 847 *Commun.*, 14, 8213, <https://doi.org/10.1038/s41467-023-43660-x>, 2023.
- 848 Cui, Y. Y., Jacobson, A. R., Feng, S., Wesloh, D., Barkley, Z. R., Zhang, L., Gerken, T., Keller,
- 849 K., Baker, D., and Davis, K. J.: Evaluation of CarbonTracker's Inverse Estimates of North
- 850 American Net Ecosystem Exchange of CO₂ From Different Observing Systems Using ACT-
- 851 America Airborne Observations, *J. Geophys. Res. Atmos.*, 126, e2020JD034406,
- 852 <https://doi.org/10.1029/2020JD034406>, 2021.
- 853 Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J.,
- 854 Landschützer, P., Le Quéré, C., Li, H., Luijkx, I. T., Peters, G. P., Peters, W., Pongratz, J.,
- 855 Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Aas, K., Alin, S. R., Anthoni, P., Barbero,
- 856 L., Bates, N. R., Bellouin, N., Benoit-Cattin, A., Berghoff, C. F., Bernardello, R., Bopp, L.,
- 857 Brasika, I. B. M., Chamberlain, M. A., Chandra, N., Chevallier, F., Chini, L. P., Collier, N. O.,
- 858 Colligan, T. H., Cronin, M., Djeutchouang, L., Dou, X., Enright, M. P., Enyo, K., Erb, M., Evans,
- 859 W., Feely, R. A., Feng, L., Ford, D. J., Foster, A., Fransner, F., Gasser, T., Gehlen, M., Gkritzalis,
- 860 T., Goncalves De Souza, J., Grassi, G., Gregor, L., Gruber, N., Guenet, B., Gürses, Ö.,
- 861 Harrington, K., Harris, I., Heinke, J., Hurtt, G. C., Iida, Y., Ilyina, T., Ito, A., Jacobson, A. R.,
- 862 Jain, A. K., Jarníková, T., Jersild, A., Jiang, F., Jones, S. D., Kato, E., Keeling, R. F., Klein
- 863 Goldewijk, K., Knauer, J., Kong, Y., Korsbakken, J. I., Koven, C., Kunimitsu, T., Lan, X., Liu, J.,
- 864 Liu, Z., Liu, Z., Lo Monaco, C., Ma, L., Marland, G., McGuire, P. C., McKinley, G. A., Melton,
- 865 J., Monacci, N., Monier, E., Morgan, E. J., Munro, D. R., Müller, J. D., Nakaoka, S. I., Nayagam,
- 866 L. R., Niwa, Y., Nutzelt, T., Olsen, A., Omar, A. M., Pan, N., Pandey, S., Pierrot, D., Qin, Z.,
- 867 Regnier, P. A. G., Rehder, G., Resplandy, L., Roobaert, A., Rosan, T. M., Rödenbeck, C.,
- 868 Schwinger, J., Skjelvan, I., Smallman, T. L., Spada, V., Sreeush, M. G., Sun, Q., Sutton, A. J.,
- 869 Sweeney, C., Swingedouw, D., Séférian, R., Takao, S., Tatebe, H., Tian, H., Tian, X., Tilbrook,
- 870 B., Tsujino, H., Tubiello, F., van Ooijen, E., van der Werf, G., van de Velde, S. J., Walker, A.,
- 871 Wanninkhof, R., Yang, X., Yuan, W., Yue, X., and Zeng, J.: Global Carbon Budget 2025, *Earth*
- 872 *Syst. Sci. Data Discuss.*, 2025, 1–139, <https://doi.org/10.5194/essd-2025-659>, 2025.
- 873 He, C., Ji, M., Grieneisen, M. L., and Zhan, Y.: A review of datasets and methods for deriving
- 874 spatiotemporal distributions of atmospheric CO₂, *J. Environ. Manage.*, 322, 116101,
- 875 <https://doi.org/10.1016/j.jenvman.2022.116101>, 2022.
- 876 Hu, K., Zhang, Q., Gong, S., Zhang, F., Weng, L., Jiang, S., and Xia, M.: A review of
- 877 anthropogenic ground-level carbon emissions based on satellite data, *IEEE J. Sel. Top. Appl.*
- 878 *Earth Obs. Remote Sens.*, 1–22, <https://doi.org/10.1109/JSTARS.2024.3355549>, 2024.
- 879 Hyndman, R. J. and Koehler, A. B.: Another look at measures of forecast accuracy, *Int. J.*
- 880 *Forecast.*, 22, 679–688, <https://doi.org/10.1016/j.ijforecast.2006.03.001>, 2006.



- 881 Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A. M.,
882 Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z.,
883 Massart, S., Parrington, M., Peuch, V. H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.:
884 The CAMS reanalysis of atmospheric composition, *Atmos. Chem. Phys.*, 19, 3515-3556,
885 <https://doi.org/10.5194/acp-19-3515-2019>, 2019.
- 886 CarbonTracker Documentation CT2022 release:
887 <https://gml.noaa.gov/ccgg/carbontracker/documentation.php>, last access: 22 March 2026.
- 888 Jin, Z., Tian, X., Wang, Y., Zhang, H., Zhao, M., Wang, T., Ding, J., and Piao, S.: A global
889 surface CO₂ flux dataset (2015–2022) inferred from OCO-2 retrievals using the GONGGA
890 inversion system, *Earth Syst. Sci. Data*, 16, 2857-2876, [https://doi.org/10.5194/essd-16-2857-](https://doi.org/10.5194/essd-16-2857-2024)
891 2024, 2024.
- 892 Karandana Gamalathge, T. D., Tai, N., Jiang, X., Wang, X., Li, L., and Yung, Y. L.: Impacts of
893 California Wildfires on CO₂ and Other Trace Gases, *Geophys. Res. Lett.*, 51, e2024GL109352,
894 <https://doi.org/10.1029/2024GL109352>, 2024.
- 895 Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y.: LightGBM: A
896 Highly Efficient Gradient Boosting Decision Tree, *Adv. Neural Inf. Process. Syst.*, 30,
897 [https://proceedings.neurips.cc/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-](https://proceedings.neurips.cc/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html)
898 [Abstract.html](https://proceedings.neurips.cc/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html), 2017.
- 899 Kenea, S. T., Oh, Y.-S., Rhee, J.-S., Goo, T.-Y., Byun, Y.-H., Li, S., Labzovskii, L. D., Lee, H.,
900 and Banks, R. F.: Evaluation of Simulated CO₂ Concentrations from the CarbonTracker-Asia
901 Model Using In-situ Observations over East Asia for 2009–2013, *Adv. Atmos. Sci.*, 36, 603-613,
902 <https://doi.org/10.1007/s00376-019-8150-x>, 2019.
- 903 Kountouris, P., Gerbig, C., Rödenbeck, C., Karstens, U., Koch, T. F., and Heimann, M.:
904 Technical Note: Atmospheric CO₂ inversions on the mesoscale using data-driven prior
905 uncertainties: methodology and system evaluation, *Atmos. Chem. Phys.*, 18, 3027-3045,
906 <https://doi.org/10.5194/acp-18-3027-2018>, 2018.
- 907 Labzovskii, L. D., Jeong, S.-J., and Parazoo, N. C.: Working towards confident spaceborne
908 monitoring of carbon emissions from cities using Orbiting Carbon Observatory-2, *Remote Sens.*
909 *Environ.*, 233, 111359, <https://doi.org/10.1016/j.rse.2019.111359>, 2019.
- 910 Lan, X., Tans, P., and Thoning, K. W.: Trends in globally-averaged CO₂ determined from NOAA
911 Global Monitoring Laboratory measurements, Version 2026-02, [https://doi.org/10.15138/9N0H-](https://doi.org/10.15138/9N0H-ZH07)
912 [ZH07](https://doi.org/10.15138/9N0H-ZH07), 2026.
- 913 Lin, X., Indira, N. K., Ramonet, M., Delmotte, M., Ciais, P., Bhatt, B. C., Reddy, M. V.,
914 Angchuk, D., Balakrishnan, S., Jorphail, S., Dorjai, T., Mahey, T. T., Patnaik, S., Begum, M.,
915 Brenninkmeijer, C., Durairaj, S., Kirubakaran, R., Schmidt, M., Swathi, P. S., Vinithkumar, N. V.,
916 Yver Kwok, C., and Gaur, V. K.: Long-lived atmospheric trace gases measurements in flask
917 samples from three stations in India, *Atmos. Chem. Phys.*, 15, 9819-9849,
918 <https://doi.org/10.5194/acp-15-9819-2015>, 2015.
- 919 Liu, J. and Chen, W.: First satellite-based regional hourly NO₂ estimations using a space-time
920 ensemble learning model: A case study for Beijing-Tianjin-Hebei Region, China, *Sci. Total*
921 *Environ.*, 820, 153289, <https://doi.org/10.1016/j.scitotenv.2022.153289>, 2022.



- 922 Liu, Z., Rogers, B. M., Keppel-Aleks, G., Helbig, M., Ballantyne, A. P., Kimball, J. S.,
 923 Chatterjee, A., Foster, A., Kaushik, A., Virkkala, A.-M., Burrell, A. L., Schwalm, C., Sweeney,
 924 C., Schuur, E. A. G., Dean, J., Watts, J. D., Kim, J. E., Wang, J. A., Hu, L., Welp, L., Berner, L.
 925 T., Mauritz, M., Mack, M., Parazoo, N. C., Madani, N., Keeling, R., Commane, R., Goetz, S.,
 926 Piao, S., Natali, S. M., Wang, W., Buermann, W., Walker, X., Lin, X., Wang, X., Jin, Y., Yu, K.,
 927 and Zhang, Y.: Seasonal CO₂ amplitude in northern high latitudes, *Nat. Rev. Earth Env.*, 5, 802-
 928 817, <https://doi.org/10.1038/s43017-024-00600-7>, 2024.
- 929 Masarie, K. A. and Tans, P. P.: Extension and integration of atmospheric carbon dioxide data into
 930 a globally consistent measurement record, *J. Geophys. Res. Atmos.*, 100, 11593-11610,
 931 <https://doi.org/10.1029/95JD00859>, 1995.
- 932 Masarie, K. A., Peters, W., Jacobson, A. R., and Tans, P. P.: ObsPack: a framework for the
 933 preparation, delivery, and attribution of atmospheric greenhouse gas measurements, *Earth Syst.*
 934 *Sci. Data*, 6, 375-384, <https://doi.org/10.5194/essd-6-375-2014>, 2014.
- 935 Mi, T., Pu, X., Zheng, X., Nie, G., Liu, X., Zhang, H., Grieneisen, M. L., Zhan, Y., Wang, M.,
 936 and Yang, F.: A global, temporally continuous, hourly CO₂ concentration dataset at GAW
 937 monitoring stations for 2000-2024: Enabling fine-scale carbon dynamics analysis, Zenodo
 938 [dataset], <https://doi.org/10.5281/zenodo.21306261>, 2026.
- 939 Munassar, S., Rödenbeck, C., Gałkowski, M., Koch, F. T., Totsche, K. U., Botía, S., and Gerbig,
 940 C.: To what extent does the CO₂ diurnal cycle impact flux estimates derived from global and
 941 regional inversions?, *Atmos. Chem. Phys.*, 25, 639-656, [https://doi.org/10.5194/acp-25-639-](https://doi.org/10.5194/acp-25-639-2025)
 942 2025, 2025.
- 943 Mustafa, F., Bu, L., Wang, Q., Shahzaman, M., Bilal, M., Aslam, R. W., and Dong, C.:
 944 Spatiotemporal Investigation of Near-Surface CO₂ and Its Affecting Factors Over Asia, *IEEE*
 945 *Trans. Geosci. Remote Sens.*, 60, 1-16, <https://doi.org/10.1109/TGRS.2022.3178125>, 2022.
- 946 Noël, S., Reuter, M., Buchwitz, M., Borchardt, J., Hilker, M., Schneising, O., Bovensmann, H.,
 947 Burrows, J. P., Di Noia, A., Parker, R. J., Suto, H., Yoshida, Y., Buschmann, M., Deutscher, N.
 948 M., Feist, D. G., Griffith, D. W. T., Hase, F., Kivi, R., Liu, C., Morino, I., Notholt, J., Oh, Y. S.,
 949 Ohyama, H., Petri, C., Pollard, D. F., Rettinger, M., Roehl, C., Rousogonous, C., Sha, M. K.,
 950 Shiomi, K., Strong, K., Sussmann, R., Té, Y., Velasco, V. A., Vrekoussis, M., and Warneke, T.:
 951 Retrieval of greenhouse gases from GOSAT and GOSAT-2 using the FOCAL algorithm, *Atmos.*
 952 *Meas. Tech.*, 15, 3401-3437, <https://doi.org/10.5194/amt-15-3401-2022>, 2022.
- 953 Organization, W. M.: WMO Greenhouse Gas Bulletin: The State of Greenhouse Gases in the
 954 Atmosphere Based on Global Observations through 2024, Geneva, Switzerland,
 955 <https://wmo.int/publication-series/wmo-greenhouse-gas-bulletin>, 2025.
- 956 Perperoglou, A., Sauerbrei, W., Abrahamowicz, M., and Schmid, M.: A review of spline function
 957 procedures in R, *BMC Med. Res. Methodol.*, 19, 46, <https://doi.org/10.1186/s12874-019-0666-3>,
 958 2019.
- 959 Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J.
 960 B., Bruhwiler, L. M. P., Pétron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R.,
 961 Randerson, J. T., Wennberg, P. O., Krol, M. C., and Tans, P. P.: An atmospheric perspective on
 962 North American carbon dioxide exchange: CarbonTracker, *Proc. Natl. Acad. Sci. U S A*, 104,



- 18925-18930, <https://doi.org/10.1073/pnas.0708986104>, 2007.
- Pillai, D., Gerbig, C., Marshall, J., Ahmadov, R., Kretschmer, R., Koch, T., and Karstens, U.: High resolution modeling of CO₂ over Europe: implications for representation errors of satellite retrievals, *Atmos. Chem. Phys.*, 10, 83-94, <https://doi.org/10.5194/acp-10-83-2010>, 2010.
- Roobaert, A., Resplandy, L., Laruelle, G. G., Liao, E., and Regnier, P.: Unraveling the Physical and Biological Controls of the Global Coastal CO₂ Sink, *Glob. Biogeochem. Cycles*, 38, e2023GB007799, <https://doi.org/10.1029/2023GB007799>, 2024.
- Saberi, A. N., Belahcen, A., Sobra, J., and Vaimann, T.: LightGBM-Based Fault Diagnosis of Rotating Machinery Under Changing Working Conditions Using Modified Recursive Feature Elimination, *IEEE Access*, 10, 81910-81925, <https://doi.org/10.1109/ACCESS.2022.3195939>, 2022.
- Shao, W., Yue, T., Zhang, L., Tian, W., Wang, H., Zhou, H., Wu, C., and Zhang, L.: Mapping daily 1-km resolution XCO₂ in China using deep learning and multi-source data, *Resour. Conserv. Recycl.*, 227, 108755, <https://doi.org/10.1016/j.resconrec.2025.108755>, 2026.
- Simpson, G. L.: Modelling Palaeoecological Time Series Using Generalised Additive Models, *Front. Ecol. Evol.*, 6, 149, <https://doi.org/10.3389/fevo.2018.00149>, 2018.
- Tang, W., Zhang, F., Stoy, P. C., Scott, R. L., Tang, A. C. I., and Fu, Z.: Contribution of Carbon Dioxide Concentration to the Diurnal Variation in Land Surface Carbon Dioxide Uptake From the Atmosphere, *Plant Cell Environ.*, 48, 6722-6733, <https://doi.org/10.1111/pce.15638>, 2025.
- Taylor, T. E., O'Dell, C. W., Baker, D., Bruegge, C., Chang, A., Chapsky, L., Chatterjee, A., Cheng, C., Chevallier, F., Crisp, D., Dang, L., Drouin, B., Eldering, A., Feng, L., Fisher, B., Fu, D., Gunson, M., Haemmerle, V., Keller, G. R., Kiel, M., Kuai, L., Kurosu, T., Lambert, A., Laughner, J., Lee, R., Liu, J., Mandrake, L., Marchetti, Y., McGarragh, G., Merrelli, A., Nelson, R. R., Osterman, G., Oyafuso, F., Palmer, P. I., Payne, V. H., Rosenberg, R., Somkuti, P., Spiers, G., To, C., Weir, B., Wennberg, P. O., Yu, S., and Zong, J.: Evaluating the consistency between OCO-2 and OCO-3 XCO₂ estimates derived from the NASA ACOS version 10 retrieval algorithm, *Atmos. Meas. Tech.*, 16, 3173-3209, <https://doi.org/10.5194/amt-16-3173-2023>, 2023.
- Te, T., Bagan, H., Che, M., Hou, X., and Uudus, B.: Spatiotemporal variability of near-surface CO₂ and its affecting factors over Mongolia, *Environ. Res.*, 236, 116796, <https://doi.org/10.1016/j.envres.2023.116796>, 2023.
- Tian, S., Wang, S., Bai, X., Luo, G., Li, Q., Yang, Y., Hu, Z., Li, C., and Deng, Y.: Global patterns and changes of carbon emissions from land use during 1992–2015, *Environ. Sci. Ecotechnol.*, 7, 100108, <https://doi.org/10.1016/j.esec.2021.100108>, 2021.
- Tsutsumi, Y., Mori, K., Hirahara, T., Ikegami, M., and Conway, T. J.: Technical Report of Global Analysis Method for Major Greenhouse Gases by the World Data Center for Greenhouse Gases (WMO/TD-No. 1473), GAW Report No. 184, Geneva, WMO, 1-23, https://library.wmo.int/index.php?lvl=notice_display&id=12631, 2009.
- van der Laan-Luijkx, I. T., van der Velde, I. R., van der Veen, E., Tsuruta, A., Stanisławska, K., Babenhauserheide, A., Zhang, H. F., Liu, Y., He, W., Chen, H., Masarie, K. A., Krol, M. C., and Peters, W.: The CarbonTracker Data Assimilation Shell (CTDAS) v1.0: implementation and global carbon balance 2001–2015, *Geosci. Model Dev.*, 10, 2785-2800,



- 1004 <https://doi.org/10.5194/gmd-10-2785-2017>, 2017.
- 1005 Wang, Q., Mustafa, F., Bu, L., Zhu, S., Liu, J., and Chen, W.: Atmospheric carbon dioxide
- 1006 measurement from aircraft and comparison with OCO-2 and CarbonTracker model data, *Atmos.*
- 1007 *Meas. Tech.*, 14, 6601-6617, <https://doi.org/10.5194/amt-14-6601-2021>, 2021.
- 1008 Wang, Z., Zhang, C., Shi, K., Shangguan, Y., Hu, B., Chen, X., Wei, D., Chen, S., Atkinson, P.
- 1009 M., and Zhang, Q.: A full-coverage satellite-based global atmospheric CO₂ dataset at 0.05°
- 1010 resolution from 2015 to 2021 for exploring global carbon dynamics, *Earth Syst. Sci. Data*, 17,
- 1011 5355-5375, <https://doi.org/10.5194/essd-17-5355-2025>, 2025.
- 1012 Wu, Z., Vermeulen, A., Sawa, Y., Karstens, U., Peters, W., de Kok, R., Lan, X., Nagai, Y., Ogi,
- 1013 A., and Tarasova, O.: Investigating the differences in calculating global mean surface CO₂
- 1014 abundance: the impact of analysis methodologies and site selection, *Atmos. Chem. Phys.*, 24,
- 1015 1249-1264, <https://doi.org/10.5194/acp-24-1249-2024>, 2024.
- 1016 Xueref-Remy, I., Dieudonné, E., Vuillemin, C., Lopez, M., Lac, C., Schmidt, M., Delmotte, M.,
- 1017 Chevallier, F., Ravetta, F., Perrussel, O., Ciais, P., Bréon, F. M., Broquet, G., Ramonet, M.,
- 1018 Spain, T. G., and Ampe, C.: Diurnal, synoptic and seasonal variability of atmospheric CO₂ in the
- 1019 Paris megacity area, *Atmos. Chem. Phys.*, 18, 3335-3362, [https://doi.org/10.5194/acp-18-3335-](https://doi.org/10.5194/acp-18-3335-2018)
- 1020 2018, 2018.
- 1021 Yuan, Y., Ries, L., Petermeier, H., Steinbacher, M., Gómez-Peláez, A. J., Leuenberger, M. C.,
- 1022 Schumacher, M., Trickl, T., Couret, C., Meinhardt, F., and Menzel, A.: Adaptive selection of
- 1023 diurnal minimum variation: a statistical strategy to obtain representative atmospheric CO₂ data
- 1024 and its application to European elevated mountain stations, *Atmos. Meas. Tech.*, 11, 1501-1514,
- 1025 <https://doi.org/10.5194/amt-11-1501-2018>, 2018.
- 1026 Zhang, G. and Lu, Y.: Bias-corrected random forests in regression, *J. Appl. Stat.*, 39, 151-160,
- 1027 <https://doi.org/10.1080/02664763.2011.578621>, 2012.
- 1028