

# Interpretable Atmospheric Carbon Dioxide Growth Modeling with Hierarchical Station Effects and Transport-Aware Covariates

Austin Lubetkin

bocaaust@gmail.com

Florida Institute of Technology

---

## Article

## Keywords:

**Posted Date:** October 23rd, 2026

**DOI:** <https://doi.org/10.21203/rs.3.rs-11226520/v1>

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

[Read Full License](#)

**Additional Declarations:** No competing interests reported.

---

# Interpretable Atmospheric Carbon Dioxide Growth Modeling with Hierarchical Station Effects and Transport-Aware Covariates

Austin Lubetkin

Florida Institute of Technology, Melbourne, Florida, USA

Corresponding author: Austin Lubetkin (bocaaust@gmail.com)

## Abstract

Atmospheric carbon dioxide models serve distinct forecasting and process-inference goals, yet they are often compared on non-equivalent tasks. We evaluate a constrained corrected-exponential model, a nonlinear mixed-effects station hierarchy, and transport-aware CarbonTracker corrections using public observations. A pure exponential produced the lowest 2016–2026 Mauna Loa holdout error (RMSE 0.762 ppm), whereas the corrected exponential provided an interpretable acceleration estimate. Across 25,315 observations from 76 NOAA stations, partial pooling reduced 2016–2024 holdout RMSE from 3.985 to 1.850 ppm. On 12,257 matched station-months from 65 stations, component, meteorological and station corrections reduced 2016–2020 RMSE from 1.888 ppm for the CT2022 posterior simulation to 1.698 ppm. Because CT2022 assimilates observations, this is a time-split residual-correction evaluation rather than independent forecast validation. These results show that compact trend forecasting, station representation and transport-aware explanation require separate benchmarks and interpretations.

## Introduction

Atmospheric carbon dioxide concentration connects anthropogenic emissions with atmospheric transport, land and ocean uptake, wildfire activity and climate variability. The Mauna Loa record has provided a continuous observational anchor since 1958, and its long-term increase is among the clearest measurements of change in the Earth system.<sup>1;2</sup> A single smooth curve can summarize this aggregate trajectory, but it cannot by itself explain why stations differ or how process information modifies station-scale concentrations. Carbon-budget studies instead represent atmospheric growth through emissions and exchanges with land and ocean reservoirs,<sup>3;4</sup> while short-run growth anomalies are associated with modes of climate variability, including the El Nino–Southern Oscillation (ENSO).<sup>5</sup>

This distinction creates an evaluation problem. A compact trend model can be judged by forward error at one site, whereas a station-network model must represent persistent site offsets, heterogeneous trends and seasonal structure. A transport-aware model addresses a third target: whether physically informed simulated concentrations and process components can reduce residual station-scale error. Ranking these models by a single root-mean-square error (RMSE) would be misleading when the observations, time windows and scientific questions differ. The present study

therefore treats compact forecasting, interpretable growth estimation, station representation and process-side residual correction as related but non-equivalent tasks.

We first compare low-dimensional Mauna Loa trend models under a time-ordered 2016–2026 holdout. The proposed corrected-exponential form separates initial level, initial growth, nonlinear acceleration and two seasonal harmonics, while a pure exponential serves as a deliberately strong forecasting benchmark. We then extend the analysis to monthly NOAA Global Monitoring Laboratory (GML) surface-flask observations using a nonlinear mixed-effects hierarchy. The hierarchy profiles a shared curvature parameter while partially pooling station-specific level, trend and annual-seasonality deviations. Finally, we match NOAA observations to the CarbonTracker CT2022 posterior station operator and evaluate station-bias, component, meteorological and residual-state corrections on later observations. This final analysis is explicitly framed as residual correction because CT2022 is an assimilation-constrained posterior product rather than an independent forecast.

The main contribution is comparative. The pure exponential is retained when it performs best for compact Mauna Loa forecasting; the corrected exponential is evaluated for interpretable acceleration; the nonlinear hierarchy is tested for station-network representation; and the CarbonTracker correction models are compared only on their common matched panel. This design prevents model complexity from being treated as evidence by itself and makes the evidentiary boundary of each result explicit.

## Results

### Compact Mauna Loa benchmarks favor a pure exponential forecast

The full monthly Mauna Loa series contained 819 valid observations from March 1958 through May 2026. The corrected-exponential model fitted the full record with RMSE 0.7612 ppm (Fig. 1). Its inferred smooth growth rate increased from 0.878 ppm yr<sup>-1</sup> near the beginning of the record to 2.744 ppm yr<sup>-1</sup> at the end, with an end-of-sample acceleration of 0.0395 ppm yr<sup>-2</sup>. Twelve-month residual block bootstrap intervals were [2.6622, 2.8318] ppm yr<sup>-1</sup> for the end growth rate and [0.0327, 0.0477] ppm yr<sup>-2</sup> for acceleration (Supplementary Table S2).

Forward validation changed the model ranking. Models were fitted only to observations before 2016 and evaluated from 2016 through May 2026. The pure exponential produced the lowest test RMSE, 0.762 ppm, compared with 1.877 ppm for the corrected exponential, 2.198 ppm for the quadratic trend, 2.902 ppm for the cubic trend and 13.324 ppm for the linear trend (Table 1). The corrected exponential fitted the pre-2016 period more closely than the pure exponential but extrapolated less accurately. It is therefore retained as an interpretable acceleration estimator rather than presented as the best compact forecast equation.

### Climate covariates explain part of short-run growth variability

The corrected-exponential residuals retained structured variation, motivating distributed-lag diagnostics using standardized global temperature anomalies and the Multivariate ENSO Index version

**Table 1: Time-ordered Mauna Loa benchmark comparison.** Training models used observations before 2016; test metrics use observations from 2016 through May 2026. Lower RMSE is better. BIC is reported as an in-sample diagnostic and is not decisive because monthly residuals are serially correlated.

| Model                 | Train RMSE | Test RMSE    | Full RMSE    | Train BIC |
|-----------------------|------------|--------------|--------------|-----------|
| Pure exponential      | 0.800      | <b>0.762</b> | 0.771        | -264.1    |
| Corrected exponential | 0.747      | 1.877        | <b>0.761</b> | -352.4    |
| Quadratic trend       | 0.739      | 2.198        | 0.815        | -374.4    |
| Cubic trend           | 0.731      | 2.902        | 0.768        | -381.8    |
| Linear trend          | 3.123      | 13.324       | 4.753        | 1620.1    |

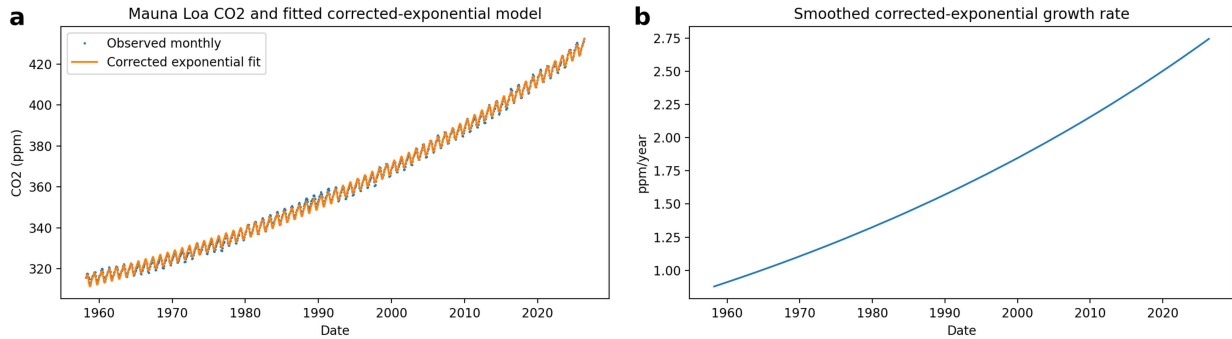


**Figure 1: Mauna Loa fit and inferred growth.** **a**, Observed monthly concentration and the full-sample corrected-exponential fit. **b**, Smooth growth rate implied by the fitted trend. Concentrations are parts per million (ppm); growth is ppm yr<sup>-1</sup>.

2 (MEI.v2). The residual regression had limited explanatory power ( $R^2 = 0.087$ ), and neither cumulative covariate effect was distinguishable from zero at the 5% level. The 12-month growth-anomaly regression was more informative ( $R^2 = 0.223$ ): the cumulative MEI.v2 effect over lags 0–6 was 0.167 ppm yr<sup>-1</sup> per one-standard-deviation increase (heteroskedasticity- and autocorrelation-consistent  $p = 0.013$ ). The cumulative temperature effect was 0.117 ppm yr<sup>-1</sup> ( $p = 0.100$ ). These associations are diagnostics of short-run variability, not causal partitioning of carbon-cycle fluxes (Supplementary Table S3 and Supplementary Figs. S1–S3).

An independent accounting check reconstructed atmospheric growth from the Global Carbon Budget 2025 components. When fossil emissions, land-use change, ocean uptake, land uptake, cement carbonation and budget imbalance were entered with their accounting signs, the identity closed to numerical precision, with maximum absolute residual  $3.35 \times 10^{-14}$  GtC yr<sup>-1</sup>. This result verifies data parsing and sign conventions rather than predictive skill. A separate climate-only annual regression for 1979–2024 explained  $R^2 = 0.595$  of annual atmospheric growth; however, it omits emissions and sinks and is consequently interpreted only as a descriptive climate diagnostic (Supplementary Tables S4–S5 and Supplementary Figs. S4–S5).

## Partial pooling improves station-network prediction

The NOAA GML surface-flask panel contained 26,534 valid station-month observations from 94 station codes before coverage filtering. Requiring at least 120 valid observations yielded 25,315 observations from 76 stations. The profiled nonlinear mixed-effects model estimated a shared curvature parameter  $\hat{\gamma} = 0.021584$  with a likelihood-ratio 95% interval [0.019668, 0.023517]. The likelihood profile had a clear interior optimum (Fig. 2a), although the curvature basis remained strongly correlated with time and its derivative. Interpretation therefore focuses on the profiled trend shape rather than assigning independent physical meaning to individual nonlinear coefficients.

Station heterogeneity was substantial. Estimated standard deviations were 2.481 ppm for station intercept deviations, 0.101 ppm yr<sup>-1</sup> for station slope deviations, 2.245 ppm for annual sine deviations and 2.476 ppm for annual cosine deviations. On the full panel, partial pooling reduced RMSE from 3.627 ppm for a fixed-only pooled curve to 1.563 ppm. In time-ordered evaluation, the model was estimated before 2016 and applied to 5,012 observations from 2016–2024 at stations already represented during training. RMSE decreased from 3.985 to 1.850 ppm (Table 2). These are forward predictions for previously observed stations, not predictions for entirely unseen locations.

Additional fixed-effect, variance-component, identifiability, coverage and training-profile diagnostics are reported in Supplementary Tables S6–S8 and Supplementary Figs. S6–S8.

**Table 2: Nonlinear hierarchical station-panel performance.** Conditional hierarchical predictions include empirical-Bayes station effects. The holdout contains later observations at stations represented before 2016.

| Sample                   | Observations | Fixed-only RMSE | Hierarchical RMSE |
|--------------------------|--------------|-----------------|-------------------|
| Full panel, 1968–2024    | 25,315       | 3.627           | 1.563             |
| Training panel, pre-2016 | 20,171       | 3.527           | 1.541             |
| Holdout panel, 2016–2024 | 5,012        | 3.985           | <b>1.850</b>      |

## Transport-aware corrections improve a common matched panel

The process-side analysis used CarbonTracker CT2022 gridded fluxes, station metadata, ERA5 validation fields and the CT2022 ObsPack posterior station simulation (Fig. 3). Annual CT2022 aggregation over 2000–2020 gave mean fossil-fuel flux of 8.714 Pg C yr<sup>-1</sup>, wildfire flux of 2.298 Pg C yr<sup>-1</sup>, ocean uptake of 2.916 Pg C yr<sup>-1</sup> and land-biosphere uptake of 3.544 Pg C yr<sup>-1</sup> (Fig. 4; Supplementary Table S9). The 2016 land-sink anomaly was the weakest uptake anomaly in this interval, while the largest positive wildfire anomaly occurred in 2003. These quantities define context for the correction models; they are not estimated causal effects.

The matched panel contained 12,257 station-month observations from 65 surface-flask stations spanning 2000–2020. Models were estimated on 9,343 observations before 2016 and evaluated on 2,914 observations from 2016–2020 at stations represented in training. Raw CT2022 posterior station concentrations achieved evaluation RMSE 1.888 ppm and  $R^2 = 0.893$ . Adding persistent station bias reduced RMSE to 1.754 ppm. The component-meteorology model, which used transport-resolved



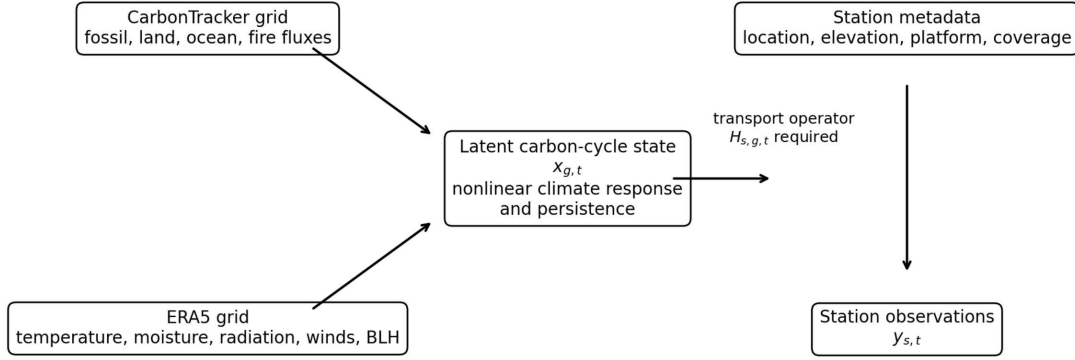


**Figure 3: Transport-aware modeling framework.** CarbonTracker flux components and ERA5 meteorology define process covariates; station metadata represent observation heterogeneity; and the CT2022 ObsPack posterior station simulations provide the transport-resolved mapping to observed station concentration. BLH denotes boundary-layer height.

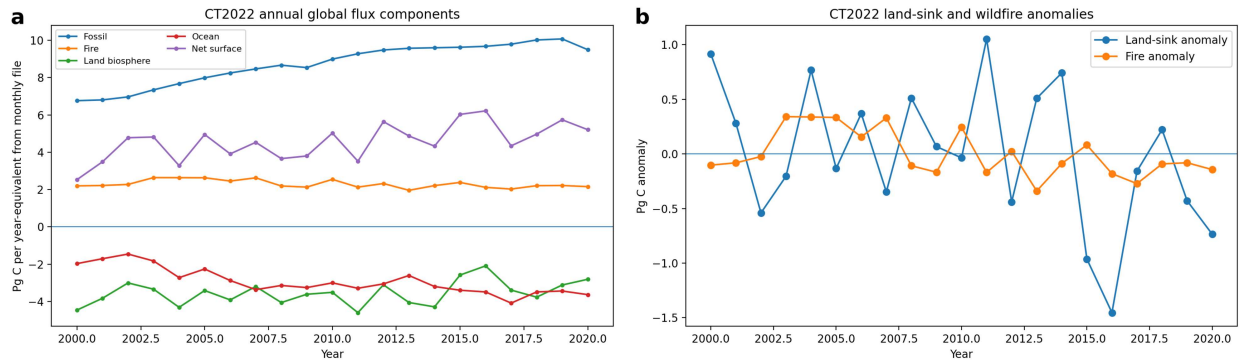


**Figure 4: CarbonTracker process diagnostics for 2000–2020.** **a**, Annual global fossil, wildfire, ocean and land-biosphere flux components; negative ocean and land values denote net uptake. **b**, Annual land-sink and wildfire anomalies. Fluxes and anomalies are  $\text{Pg C yr}^{-1}$ .

fossil, biosphere, ocean, wildfire and forecast components together with station-sampled meteorology and station effects, obtained the lowest evaluation RMSE, 1.698 ppm, and  $R^2 = 0.914$  (Table 3; Fig. 5). This was a 10.0% reduction in RMSE relative to raw CT2022, 3.2% relative to station-bias correction and 3.3% relative to the residual-state transition.

Station metadata and ERA5 decoding checks are shown in Supplementary Figs. S9–S10. A nearest-cell CarbonTracker join reduced descriptive station-panel RMSE from 3.050 to 2.785 ppm, but this diagnostic omits upwind transport and was not used as the final observation operator (Supplementary Table S10 and Supplementary Figs. S11–S12).

The residual-state transition remained competitive at 1.756 ppm, but its evaluation-period residual-state  $R^2$  was negative ( $-0.130$ ). Aggregate flux anomalies, station-sampled meteorology and nonlinear interaction terms therefore did not fully predict month-to-month residual-state variation outside the estimation interval. Moreover, component coefficients are conditional transport-space corrections and cannot be interpreted as causal source sensitivities (Supplementary Table S11 and Supplementary Fig. S13). Most importantly, CT2022 is a posterior assimilation product. The time split evaluates whether later matched residuals can be reduced after conditioning on that product; it does not constitute independent forecast validation of CarbonTracker.

**Table 3: Common-panel transport-aware evaluation.** All models use the same 65-station matched panel. Training uses observations before 2016; evaluation uses 2016–2020 observations. CT2022 values are posterior simulations, so metrics characterize residual correction rather than independent CarbonTracker forecasting.

| Model                                           | Train RMSE | Eval. RMSE   | Train $R^2$ | Eval. $R^2$  |
|-------------------------------------------------|------------|--------------|-------------|--------------|
| Raw CT2022 simulation                           | 2.191      | 1.888        | 0.957       | 0.893        |
| CT2022 plus station bias                        | 1.807      | 1.754        | 0.970       | 0.908        |
| Components plus meteorology and station effects | 1.658      | <b>1.698</b> | 0.975       | <b>0.914</b> |
| Residual-state transition plus station operator | 1.784      | 1.756        | 0.971       | 0.908        |

## Continuations clarify rather than replace scenario projections

The trend, hierarchy and process models were extended as empirical continuations to 2100 (Table 4; Fig. 6). Because the monthly series used for the full fit includes observations through May 2026, the 2026 entries are continuation nowcasts rather than ex ante forecasts. The pure and corrected exponential values, 429.86 and 429.52 ppm, fall within the Met Office 2026 annual Mauna Loa forecast of  $429.4 \pm 0.6$  ppm.<sup>6</sup> The hierarchy and transport values serve different station-conditioned targets and are not calibrated seasonal Mauna Loa forecasts.

By 2100, unconstrained statistical continuations ranged from 779.83 to 908.04 ppm. These values lie within the 393–1135 ppm span of the published Shared Socioeconomic Pathway (SSP) concentration trajectories,<sup>7</sup> but this numerical overlap does not make the models policy-conditioned projections. SSPs encode emissions, mitigation and carbon-cycle assumptions. The continuations instead show what fitted empirical forms imply if extended without scenario-specific forcing. Their long-horizon divergence is therefore evidence of extrapolation uncertainty and model-form

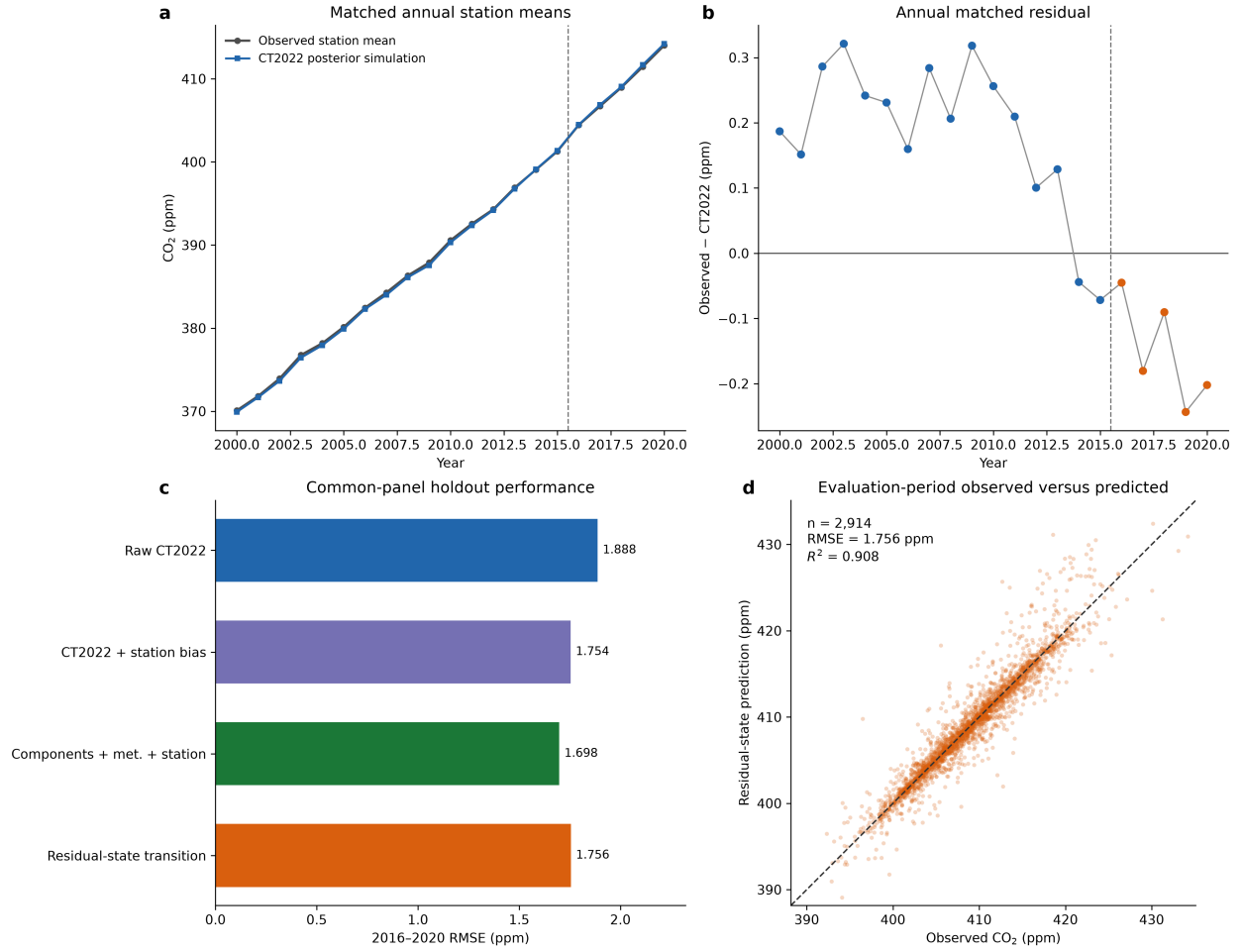


**Figure 5: Transport-aware matched-panel diagnostics.** **a**, Annual observed station mean and CT2022 posterior station simulation. **b**, Annual mean matched residual. **c**, Evaluation RMSE for the four common-panel models. **d**, Observed and predicted concentrations for the reduced residual-state transition during evaluation; the line denotes equality.

dependence, not a probabilistic forecast envelope. External-reference and complete model-to-task comparisons are provided in Supplementary Tables S12–S13.

**Table 4: Annual-mean carbon dioxide continuations.** Values are ppm. The 2026 entries are continuation nowcasts because the full monthly record includes observations through May 2026; later values are empirical continuations rather than policy-conditioned projections.

| Model                      | 2026   | 2030   | 2040   | 2050   | 2100   |
|----------------------------|--------|--------|--------|--------|--------|
| Pure exponential Mauna Loa | 429.86 | 441.50 | 474.15 | 512.56 | 833.53 |
| Corrected exponential      | 429.52 | 440.84 | 472.10 | 508.01 | 779.83 |
| Hierarchy, MLO conditional | 430.77 | 442.73 | 476.81 | 518.03 | 908.04 |
| Transport station mean     | 430.25 | 442.36 | 476.99 | 518.90 | 903.65 |

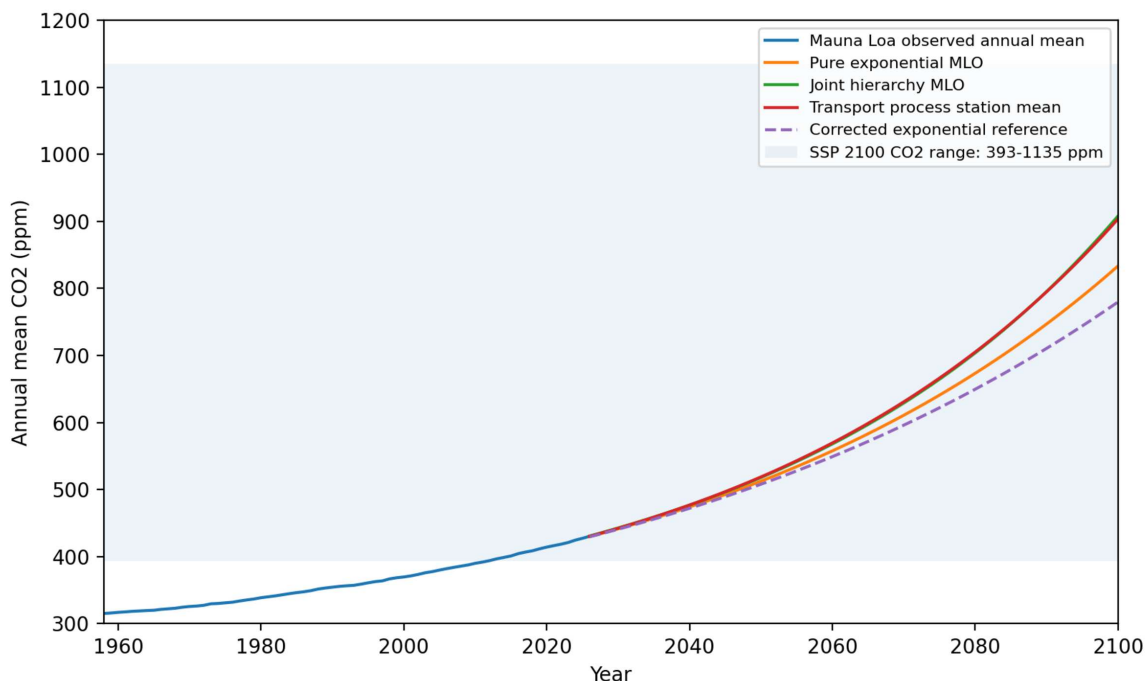


**Figure 6: Historical concentration and empirical continuations.** Observed Mauna Loa annual means and four model continuations to 2100. The shaded region marks the 393–1135 ppm SSP concentration range for 2100; it is a scenario span, not a confidence interval for the fitted models.

## Discussion

The results support a model-by-task interpretation rather than a universal ranking. The pure exponential was the strongest compact Mauna Loa forecaster on the 2016–2026 holdout. This finding prevents the corrected-exponential model from being oversold on predictive performance. Its value lies elsewhere: it provides a differentiable decomposition of initial growth, later acceleration and seasonality, while block-bootstrap intervals quantify uncertainty in derived acceleration quantities.

The distinction between forecast performance and interpretability is consequential because the model with the best in-sample fit did not provide the best forward extrapolation.

The station hierarchy provides the strongest observation-side result. Monthly concentrations differ systematically across location, record length and seasonal regime; a fixed-only pooled curve consequently leaves large station-structured errors. Partial pooling reduced holdout RMSE by 53.6% for later observations at stations present during training. The model does not establish performance at new stations, and the curvature estimate changed when the estimation period was shortened. Nevertheless, the forward result shows that explicitly representing station heterogeneity is essential when moving beyond a single aggregate record.

The process-side result is narrower but physically better aligned with station measurements. CarbonTracker maps gridded flux information through atmospheric transport to the sampling location and time. On the common CT2022 matched panel, component, meteorological and station corrections reduced later-period residual error beyond both the raw posterior simulation and a station-bias-only correction. The relatively small improvement over station bias, 3.2%, also matters: persistent location-specific offsets account for much of the correctable error, and additional process covariates provide incremental rather than transformative gains. The negative out-of-period  $R^2$  for the monthly residual-state transition further indicates that aggregate forcing variables do not capture all local representativeness and stochastic variability.

Several limitations define the evidence boundary. First, none of the statistical coefficients identifies causal carbon-cycle sensitivity. The Mauna Loa curve estimates measured acceleration, the hierarchy represents station heterogeneity, and the CT2022 coefficients condition on posterior transport simulations and station controls. Second, RMSE values from the Mauna Loa, station-network and CT2022 panels are not directly comparable because their observations and periods differ. Third, nonlinear trend parameters are coupled: profile likelihood and block bootstrap reduce, but do not eliminate, model-form and identifiability uncertainty. Fourth, CT2022 assimilates observations, so later matched residuals are not independent CarbonTracker forecasts. Finally, the 2100 continuations omit scenario-dependent emissions and future transport or meteorological forcing and should not be used in place of Earth-system or integrated-assessment projections.

Within these boundaries, the study offers a reproducible evaluation framework. Compact trend forecasting should be tested against simple extrapolative baselines; station representation should be evaluated under time-ordered partial pooling; and process-side corrections should be compared on a common transport-matched panel. The most credible next step is an independent forecast experiment using a CarbonTracker release or frozen state estimate unavailable to model fitting, paired with evaluation at both previously observed and entirely held-out stations. Such a design would determine whether the incremental process-side gains persist outside a posterior-assimilation setting.

## Methods

### Data sources and preprocessing

All analyses used public environmental data (Table 5). The Mauna Loa monthly series and NOAA annual growth series were downloaded from NOAA GML.<sup>8</sup> Surface-flask monthly station files were obtained from the NOAA Global Greenhouse Gas Reference Network.<sup>9</sup> MEI.v2 came from NOAA Physical Sciences Laboratory,<sup>10</sup> and global monthly temperature anomalies came from NASA GISTEMP v4.<sup>11</sup> Annual carbon-budget components came from the Global Carbon Budget 2025 data release.<sup>12</sup> CarbonTracker CT2022 supplied monthly gridded flux fields and posterior station simulations,<sup>13;14</sup> while ERA5 supplied reanalysis fields used to validate the intended meteorological forcing block.<sup>15;16</sup> Public aggregate environmental records were analyzed; no human participants, animals or identifiable personal data were involved, and ethics approval was therefore not required.

The Mauna Loa analysis excluded nonpositive concentration values. Monthly date was represented by elapsed years  $t$  since March 1958 and within-year phase  $p \in [0, 1]$ . The station panel harmonized NOAA station files by station code and month, then retained stations with at least 120 valid monthly observations for mixed-effects estimation. CT2022 posterior events were aggregated to station-month and matched to NOAA surface-flask observations by station code and month. Only stations represented before 2016 were included in later-period conditional evaluation.

**Table 5: Public datasets and analytical roles.** URLs and persistent identifiers are repeated in the Data availability statement. GML denotes Global Monitoring Laboratory; PSL, Physical Sciences Laboratory.

| Dataset                             | Period or frequency                | Analytical role                                           |
|-------------------------------------|------------------------------------|-----------------------------------------------------------|
| NOAA GML Mauna Loa CO <sub>2</sub>  | Monthly, March 1958–May 2026       | Single-site trends, benchmarks and growth anomalies       |
| NOAA GML surface-flask network      | Monthly, 1968–2024                 | Nonlinear hierarchical station panel                      |
| NOAA PSL MEI.v2 and NASA GISTEMP v4 | Monthly                            | Distributed-lag climate diagnostics                       |
| Global Carbon Budget 2025           | Annual, 1959–2024                  | Budget identity and annual climate diagnostics            |
| CarbonTracker CT2022 fluxes         | Monthly, 2000–2020                 | Fossil, ocean, land-biosphere and wildfire process fields |
| ERA5 monthly single levels          | Monthly                            | Meteorological forcing validation                         |
| CT2022 ObsPack station simulations  | Event and station-month, 2000–2020 | Transport-aware posterior station operator and components |

### Corrected-exponential trend and benchmark models

The corrected-exponential concentration model was

$$\widehat{C}(t, p) = \theta_0 + \theta_1 t + \theta_2 \{\exp(\theta_3 t) - 1 - \theta_3 t\} + S(p), \quad (1)$$

202 where

$$S(p) = \theta_4 \sin(2\pi p) + \theta_5 \cos(2\pi p) + \theta_6 \sin(4\pi p) + \theta_7 \cos(4\pi p). \quad (2)$$

203 Subtracting the constant and linear terms from the exponential basis preserves  $\theta_0$  as the initial  
 204 level and  $\theta_1$  as the initial growth component. Parameters minimized mean squared error subject to  
 205 box constraints and seasonal amplitude limits  $\sqrt{\theta_4^2 + \theta_5^2} \leq 8$  ppm and  $\sqrt{\theta_6^2 + \theta_7^2} \leq 3$  ppm. Both  
 206 amplitude constraints were inactive at the optimum. The smooth growth rate and acceleration were

$$\frac{d\hat{C}}{dt} = \theta_1 + \theta_2\theta_3\{\exp(\theta_3 t) - 1\}, \quad (3)$$

$$\frac{d^2\hat{C}}{dt^2} = \theta_2\theta_3^2 \exp(\theta_3 t). \quad (4)$$

207 Linear, quadratic, cubic and pure-exponential trends used the same annual and semiannual  
 208 seasonal harmonics. The time-ordered split was fixed before model comparison: observations before  
 209 2016 formed the training set and observations from 2016 onward formed the test set. Model ranking  
 210 used test RMSE. Training Bayesian information criterion was retained only as a rough in-sample  
 211 diagnostic because residual dependence violates an independent-error interpretation.

## 212 Uncertainty and residual climate diagnostics

213 Uncertainty for derived corrected-exponential quantities used 250 successful residual block-bootstrap  
 214 refits with 12-month blocks. The block length preserved short-range monthly dependence and the  
 215 seasonal cycle more appropriately than independent residual resampling. Local Gauss–Newton  
 216 standard errors were treated as diagnostic because the fitted residuals were serially correlated.

217 Monthly residuals  $r_m = C_m - \hat{C}_m$  and 12-month growth anomalies

$$a_m = (C_m - C_{m-12}) - \left. \frac{d\hat{C}}{dt} \right|_m \quad (5)$$

218 were regressed on standardized temperature and MEI.v2 at lags 0–6. Heteroskedasticity- and  
 219 autocorrelation-consistent covariance used a Newey–West estimator with 12 monthly lags. Cumula-  
 220 tive covariate effects were sums of the seven lag coefficients and were tested two-sided against zero.  
 221 The annual climate-only diagnostic used two annual lags in the covariance estimator. The Global  
 222 Carbon Budget identity was checked arithmetically and was not treated as an inferential regression  
 223 because its residual was numerical roundoff.

## 224 Nonlinear hierarchical station model

225 For station  $s$  and month  $m$ , the partially pooled model was

$$C_{s,m} = x_m(\gamma)^\top \beta + z_m^\top b_s + \epsilon_{s,m}, \quad \epsilon_{s,m} \sim N(0, \sigma^2), \quad (6)$$

$$b_s \sim N(0, D). \quad (7)$$

226 The fixed design  $x_m(\gamma)$  contained an intercept, time, the corrected-exponential basis  $q(t_m; \gamma)$  and  
 227 annual and semiannual sine and cosine terms. The random design  $z_m$  contained an intercept,  
 228 centered time in decades and annual sine and cosine terms. The shared nonlinear curvature was  
 229 estimated by profile maximum likelihood,

$$\hat{\gamma} = \arg \max_{\gamma} \max_{\beta, D, \sigma^2} \ell(\beta, D, \sigma^2, \gamma \mid C). \quad (8)$$

230 At each trial  $\gamma$ , the resulting linear mixed-effects model re-estimated fixed effects, the station  
 231 covariance matrix, residual variance and empirical-Bayes station effects. A likelihood-ratio interval  
 232 was obtained from the one-parameter profile. Conditional predictions include learned station  
 233 effects; evaluation was therefore restricted to stations represented during training. Identifiability  
 234 diagnostics included the fixed-design condition number, variance inflation for  $q(t; \hat{\gamma})$ , correlations  
 235 among nonlinear basis terms and a training-period profile.

## 236 **Transport-aware matched-panel models**

237 CarbonTracker flux fields were aggregated annually after area-weighted integration of monthly fossil,  
 238 ocean, land-biosphere and wildfire fields. A nearest-cell station join was used only as a descriptive  
 239 diagnostic because a local grid cell does not represent upwind transport. The inferential station  
 240 calculation instead used CT2022 ObsPack posterior values evaluated at the sampling location and  
 241 time. To match the NOAA monthly observation stream without duplicating laboratory products, we  
 242 paired each observed `co2_SITE_surface-flask_1_ccgg_month.txt` series with the corresponding  
 243 ObsPack `co2_SITE_surface-flask_1_representative.nc` product and retained one representative  
 244 dataset per site; alternative laboratory streams and duplicate products were excluded. The matched  
 245 product included posterior concentration, transport-resolved fossil, biosphere, ocean, wildfire and  
 246 forecast components, ensemble range, model-data mismatch and station-sampled meteorological  
 247 variables.

248 Four models were compared on identical observations: raw CT2022 posterior concentration;  
 249 CT2022 plus a persistent station bias; a component-meteorology regression with station effects; and  
 250 a residual-state transition added to the station operator. The residual observation model was

$$y_{s,t} = m_{s,t}^{\text{final}} + b_s + z_t + \epsilon_{s,t}, \quad (9)$$

251 where  $m_{s,t}^{\text{final}}$  is the CT2022 posterior station simulation,  $b_s$  is station bias and  $z_t$  is a shared monthly  
 252 residual state. The transition included lagged state, global fossil and wildfire fluxes, land-sink and  
 253 wildfire anomalies, station-averaged temperature, boundary-layer height, humidity, wind, ensemble  
 254 spread and prespecified fire-by-mixing and land-anomaly-by-temperature interactions. Continuous  
 255 covariates were standardized on the pre-2016 training period and the same transformations were  
 256 applied to evaluation data.

257 The primary process-side comparison used 9,343 observations before 2016 for estimation and

2,914 observations from 2016–2020 for evaluation. RMSE and coefficient of determination were computed on the same rows for every model. Because the baseline station simulation is posterior and assimilation constrained, the evaluation target was improvement in later matched residuals, not independent forecast skill of CT2022.

## Continuation calculations

Annual-mean continuations evaluated the fitted pure exponential and corrected-exponential Mauna Loa models, the conditional MLO station hierarchy and the station-mean transport model at future years. No future emissions, meteorology or transport pathways were supplied. The outputs are consequently deterministic model-form continuations. They were compared descriptively with the Met Office 2026 Mauna Loa forecast and the published SSP concentration range; no uncertainty interval or probability was assigned from that scenario comparison.

## Statistics and reproducibility

Sample sizes are reported as discrete station-month observations and station counts for each analysis. Measurements were repeated over time within stations; time-ordered rather than random validation was used to avoid leakage from later observations. RMSE was defined as  $\sqrt{n^{-1} \sum_i (y_i - \hat{y}_i)^2}$ . Reported hypothesis tests were two-sided. Monthly and annual regression covariance estimators accounted for serial dependence using prespecified lag lengths. No adjustment for multiple comparisons was applied to the exploratory distributed-lag coefficient profiles; interpretation emphasizes the prespecified cumulative covariate tests and treats the profiles as diagnostic. Profile-likelihood and residual block-bootstrap procedures were used where nonlinear parameter coupling or serial dependence made local asymptotic uncertainty inadequate. No observations were excluded after outcome inspection; exclusions were based on nonpositive concentration values, station coverage and availability of matched covariates.

## Data availability

The public source data are available as follows: NOAA GML Mauna Loa monthly concentration and annual growth data at <https://gml.noaa.gov/ccgg/trends/data.html>; NOAA GML surface-flask carbon dioxide files at [https://gml.noaa.gov/aftp/data/trace\\_gases/co2/flask/surface/txt/](https://gml.noaa.gov/aftp/data/trace_gases/co2/flask/surface/txt/); NOAA PSL MEI.v2 at <https://psl.noaa.gov/enso/mei/>; NASA GISTEMP v4 at <https://data.giss.nasa.gov/gistemp/>; Global Carbon Budget 2025 at <https://doi.org/10.18160/GCP-2025>; CarbonTracker CT2022 fluxes at <https://doi.org/10.25925/z1gj-3254>; CT2022 simulated-observation products through the NOAA ObsPack portal at <https://gml.noaa.gov/ccgg/obspack/>; and ERA5 monthly averaged single-level data at <https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels-monthly-means>. These sources were accessed during June 2026; NOAA trend data were refreshed through May 2026.

## References

- [1] Keeling, C. D. et al. Atmospheric carbon dioxide variations at Mauna Loa Observatory, Hawaii. *Tellus* **28**, 538–551 (1976).
- [2] Thoning, K. W., Tans, P. P. & Komhyr, W. D. Atmospheric carbon dioxide at Mauna Loa Observatory 2 Analysis of the NOAA GMCC data 1974–1985. *J. Geophys. Res.* **94**, 8549–8565 (1989).
- [3] Friedlingstein, P. et al. Global Carbon Budget 2025. *Earth Syst. Sci. Data* **18**, 3211–3288; <https://doi.org/10.5194/essd-18-3211-2026> (2026).
- [4] Bennedsen, M., Hillebrand, E. & Koopman, S. J. A regression-based approach to the CO<sub>2</sub> airborne fraction enhancing statistical precision and tackling zero emissions. *Nat. Commun.* **15**, 8333; <https://doi.org/10.1038/s41467-024-52728-1> (2024).
- [5] Wang, W. et al. Variations in atmospheric CO<sub>2</sub> growth rates coupled with tropical temperature. *Proc. Natl Acad. Sci. USA* **110**, 13061–13066; <https://doi.org/10.1073/pnas.1219683110> (2013).
- [6] Betts, R. A., Jones, C. D., Knight, J. R., Pope, J. O. & Sandford, C. Mauna Loa carbon dioxide forecast for 2026. Met Office <https://www.metoffice.gov.uk/research/climate/seasonal-to-decadal/seasonal-forecast/forecasts/co2-forecast> (2026).
- [7] Meinshausen, M. et al. The shared socio-economic pathway greenhouse gas concentrations and their extensions to 2500. *Geosci. Model Dev.* **13**, 3571–3605; <https://doi.org/10.5194/gmd-13-3571-2020> (2020).
- [8] NOAA Global Monitoring Laboratory. Trends in atmospheric carbon dioxide Mauna Loa data. <https://gml.noaa.gov/ccgg/trends/data.html> (2026).
- [9] NOAA Global Monitoring Laboratory. Surface flask carbon dioxide measurements from the Global Greenhouse Gas Reference Network. [https://gml.noaa.gov/aftp/data/trace\\_gases/co2/flask/surface/txt/](https://gml.noaa.gov/aftp/data/trace_gases/co2/flask/surface/txt/) (2025).
- [10] NOAA Physical Sciences Laboratory. Multivariate ENSO Index version 2. <https://psl.noaa.gov/enso/mei/> (2026).
- [11] GISTEMP Team. GISS Surface Temperature Analysis version 4. NASA Goddard Institute for Space Studies <https://data.giss.nasa.gov/gistemp/> (2026).
- [12] Global Carbon Project. Global Carbon Budget 2025 data release. ICOS Carbon Portal <https://doi.org/10.18160/GCP-2025> (2025).
- [13] Jacobson, A. R. et al. CarbonTracker CT2022. NOAA Global Monitoring Laboratory; <https://doi.org/10.25925/z1gj-3254> (2023).

- [14] NOAA Global Monitoring Laboratory. CarbonTracker CT2022 simulated observations, ObsPack product `obspack_co2_1_CARBONTRACKER_CT2022_2023-01-26`. <https://gml.noaa.gov/ccgg/obspack/> (2023).
- [15] Hersbach, H. et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049; <https://doi.org/10.1002/qj.3803> (2020).
- [16] Copernicus Climate Change Service. ERA5 monthly averaged data on single levels from 1940 to present. Climate Data Store <https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels-monthly-means> (2026).

## Acknowledgements

The author acknowledges the data producers and scientific teams responsible for NOAA GML observations, CarbonTracker, ERA5, GISTEMP, MEI.v2 and the Global Carbon Budget. CT2022 ObsPack identifies the contributing laboratories and investigators whose observations were simulated; the corresponding product citation is given above. Consistent with the ObsPack Fair Use Statement, the applicable data providers should be contacted before publication to confirm the appropriate acknowledgement or authorship. This study received no external funding. The data providers had no role in study design, analysis, interpretation or manuscript preparation.

## Author contributions

A.L. conceived the study; acquired and processed the data; developed the methodology and software; performed the formal analysis and validation; created the visualizations; and wrote, reviewed and edited the manuscript.

## Competing interests

The author declares no financial or non-financial competing interests.

## Supplementary Files

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

- [CO2SupplementaryInformation.zip](#)