

Charting the Carbon Future: Dynamic CO₂ Emissions Forecasts with ARIMA & TBATS

STA457 Final Project

Mohammad Danish Malik
June 22, 2025

Introduction

Global carbon dioxide (CO₂) emissions have risen dramatically over the past two centuries, driven by industrialization, urbanization, and increasing energy consumption. As the principal anthropogenic greenhouse gas, CO₂ plays a central role in climate change; its atmospheric concentration currently exceeds 410 ppm, roughly 50 % higher than pre-industrial levels. Accurate forecasts of future emissions are therefore essential for informing climate policy, guiding international negotiations, and designing mitigation strategies. This project undertakes a rigorous time-series analysis of global annual CO₂ emissions, sourced from the Our World in Data (OWID) repository, to produce reliable short to medium term projections and to compare the performance of two complementary forecasting methodologies.

The primary motivation for this study is twofold. First, policymakers and researchers require robust emission forecasts to evaluate the likely trajectories of greenhouse-gas concentrations under various socio-economic scenarios. Second, the intrinsic statistical challenges posed by the CO₂ series, its accelerating trend, potential structural breaks (e.g., post-World-War II economic boom, 1990s policy shifts, COVID-19 dip), and evolving variance, call for the application and comparison of both classical and advanced models. By fitting an ARIMA model as a transparent, parsimonious baseline and a TBATS model as a flexible state-space alternative, this analysis aims to balance interpretability with adaptability, thereby offering insights into which approach better captures the complex dynamics of global emissions.

The real-world significance of reliable CO₂ forecasts cannot be overstated. Accurate projections feed directly into integrated assessment models (IAMs), which combine climate, economic, and energy system modules to simulate future pathways under different policy regimes. These IAM outputs, in turn, underpin the decadal assessments of the Intergovernmental Panel on Climate Change (IPCC) and inform nationally determined contributions (NDCs) under the Paris Agreement. Moreover, energy and financial markets increasingly price climate risk; power utilities, investors, and insurers all benefit from understanding emission trajectories when evaluating asset-stranding risks and transition costs. A rigorous time-series framework thus contributes both to scientific discourse and to practical decision-making.

This report has three main objectives:

- 1. Data characterization and preprocessing:** We will conduct exploratory analysis of the OWID annual CO₂ series—visualizing its trend, testing for stationarity, and identifying potential structural breaks and variance shifts.
- 2. Model implementation and comparison:** We will fit an ARIMA(p, d, q) model following the Box–Jenkins methodology and a TBATS model that integrates Box–Cox transformation, trigonometric seasonality, ARMA errors, trend, and seasonal components. Forecast accuracy will be evaluated via rolling-origin cross-validation, using metrics such as RMSE and MAPE, and through diagnostic checks on residuals.

3. **Interpretation and policy implications:** We will interpret each model's fitted components, e.g., AR/MA coefficients, state-space trend estimates, and discuss the implications of divergent forecasts under each approach. Where appropriate, we will explore how exogenous shocks (e.g., economic crises, pandemics) might affect forecast reliability.

Several key challenges arise in realizing these objectives. First, the non-stationary nature of the raw series requires careful differencing or transformation to satisfy ARIMA's linear-stationarity assumptions. Second, detecting structural breaks, periods when the underlying data-generating process shifts dramatically, demands both statistical tests and substantive domain knowledge. Third, while TBATS offers automatic seasonality detection, its state-space estimation can be computationally intensive and may risk overfitting if too many latent components are included. Finally, given the relatively short length of high-quality annual data (typically 70–80 observations), there is a trade-off between model complexity and estimation precision.

In summary, this project will deliver a comprehensive introduction to time-series forecasting in the context of global CO₂ emissions. By juxtaposing a well-established linear ARIMA approach against a modern, flexible TBATS framework, the study will highlight the strengths and limitations of each method and provide actionable forecasts that can inform climate policy and strategic decision-making.

Literature Review

This literature review surveys foundational and contemporary research on time-series modeling and forecasting, with a focus on the ARIMA family of models, extensions for complex patterns, hybrid approaches, and applications to carbon-dioxide emissions. We highlight how the proposed comparison of ARIMA and TBATS for global CO₂ forecasting both builds upon and departs from existing studies.

2.1 Autoregressive Integrated Moving Average (ARIMA) Models

The ARIMA framework, introduced by Box and Jenkins, remains the cornerstone of linear time-series analysis. In their seminal work, Box, Jenkins, and Reinsel comprehensively develop the theory and practice of ARIMA processes, detailing the identification (via ACF/PACF), estimation, and diagnostic checking steps for fitting ARIMA(p, d, q) models to non-stationary data through differencing (Box, Jenkins, & Reinsel, 2008).

Empirical applications to greenhouse-gas series underscore ARIMA's utility. Ahammad et al. (2017) apply Box–Jenkins ARIMA to annual CO₂ emissions in Bangladesh, demonstrating how differencing and careful order selection (p, d, q) can capture national emission trends, and noting limitations in accounting for structural shifts and nonlinear variance growth.

Similarly, Othman (2017) contrasts ARIMA and ARIMAX for river discharge forecasting,

finding that including exogenous covariates improves goodness-of-fit metrics but adds complexity to model specification and validation.

These studies collectively establish that ARIMA firstly, **handles non-stationarity** by differencing, yielding parsimonious linear models suited to gradual trend changes. Secondly, it **relies on transparent diagnostic tools**, facilitating interpretability of autocorrelation structures. Thirdly, it **can struggle** with abrupt regime shifts, evolving variance, and complex seasonal or cyclical patterns not well approximated by low-order ARMA terms.

2.2 Extensions for Exogenous Predictors (ARIMAX)

While ARIMAX offers greater explanatory power, it introduces challenges as well. **Data requirements** increase, as time-aligned covariates must be sourced and preprocessed. Moreover, **model complexity** rises, complicating diagnostics and increasing overfitting risk when sample sizes are limited. Finally there are **interpretability trade-offs**, since parameter estimates now reflect both serial correlation and exogenous impacts.

Our study focuses on univariate benchmarking (ARIMA) before exploring fully automated state-space methods, deferring ARIMAX-style analyses to potential future extensions.

2.3 State-Space and TBATS for Complex Patterns

Traditional exponential-smoothing models — including Holt–Winters variants — excel at single, integer seasonal patterns but falter when faced with multiple, non-integer, or evolving seasonality. De Livera, Hyndman, and Snyder (2011) introduce the TBATS model, a state-space framework that seamlessly integrates: **Box–Cox transformations** to stabilize variance, **trigonometric (Fourier) representations** for multiple and non-integer seasonal cycles, **ARMA error correction** for residual autocorrelation, and **trend and seasonal components** within a unified state-space form.

Key advantages of TBATS include:

1. **Automatic detection** of latent seasonal periods, accommodating complex and nested cycles without manual tuning.
2. **Flexible trend modeling**, capturing non-linear or piecewise behaviors common in environmental series.
3. **Robustness to structural changes**, as the state-space formulation adapts to shifting patterns with maximum-likelihood estimation.

Despite its sophistication, TBATS can be **computationally intensive**, and its automated nature may obscure component interpretability relative to the transparent ARIMA framework.

2.4 Hybrid Linear–Nonlinear Approaches

Recognizing that many real-world series exhibit both linear and nonlinear dynamics, researchers have developed hybrid models combining ARIMA's linear strength with nonlinear learners.

Zhang (2003) proposes an ARIMA–ANN hybrid, where: **ARIMA** captures the linear autocorrelation structure, **Artificial neural network (ANN)** models the residual nonlinear patterns, and **Combined forecasts** often outperform either model alone on metrics like RMSE.

Hybrid approaches address:

- **Misspecified linear trends**, where pure ARIMA leaves structured residuals.
- **Nonlinear dependencies**, which ANNs flexibly model but can overfit if used in isolation.
- **Model-selection uncertainty**, by hedging between linear and nonlinear paradigms.

However, ANN hybrids require careful architecture selection, risk data-hungry overfitting, and add layers of complexity that may exceed the scope of this univariate comparison.

2.5 Applications to CO₂ Emissions Forecasting

While ARIMA and its variants have been applied to country-level CO₂ series (e.g., Bangladesh) and to hydrological flows (ARIMAX), there is less precedent for **global annual CO₂ forecasting**. The global series exhibits: **accelerating long-term trend**, driven by industrialization, **potential structural breaks**, such as the post-1990 deceleration in some regions, or the COVID-19 dip, and **evolving variance**, reflecting periods of rapid economic change.

Our study builds on existing work by:

1. **Benchmarking ARIMA** on the global aggregate, leveraging Box–Jenkins principles to capture core linear dynamics.
2. **Deploying TBATS** to automatically detect any latent multi-decadal cycles or heteroscedastic behavior without manual specification.
3. **Providing a direct comparison** of parsimonious linear versus flexible state-space models on a critical environmental time series not previously explored in tandem.
4. **Avoiding exogenous complexity**, focusing first on univariate methods before considering ARIMAX-style extensions.

2.6 Gaps and Contributions

The reviewed literature demonstrates robust methodologies for: Linear modeling (ARIMA), Extended forecasting (ARIMAX, TBATS), and Hybrid frameworks (ARIMA–ANN).

Nonetheless, **global CO₂ emissions** pose unique challenges: long-term acceleration, regime shifts, and potentially hidden cyclical patterns up to multi-decadal scales. No prior study, to our knowledge, has directly compared ARIMA and TBATS on the same global CO₂ dataset to assess which paradigm better captures these dynamics. Our contribution lies in conducting this head-to-head evaluation, thereby informing both methodological practice and climate-policy modeling with rigorous empirical evidence.

By synthesizing foundational theory with recent advances, this literature review sets the stage for a detailed implementation and comparison of ARIMA and TBATS forecasting methods on global CO₂ emissions, elucidating their relative strengths, limitations, and practical implications.

Methodology

This section details the data preprocessing steps, transformation and stationarity testing, and the specification and diagnostics of the two forecasting methods, ARIMA and TBATS, chosen for global CO₂ emissions forecasting. All analyses were conducted in R using the `forecast`, `tseries`, and `zoo` packages. Figures 1–6 and Table 1 summarize key outputs.

3.1 Data Preprocessing

We use the annual “World” CO₂ series from Our World in Data’s `owid-co2-data.csv`, covering 1751–2021 (Figure 1). After loading with `read_csv()`, we filtered on `country == "World"` and selected `year` and `co2` (million tonnes).

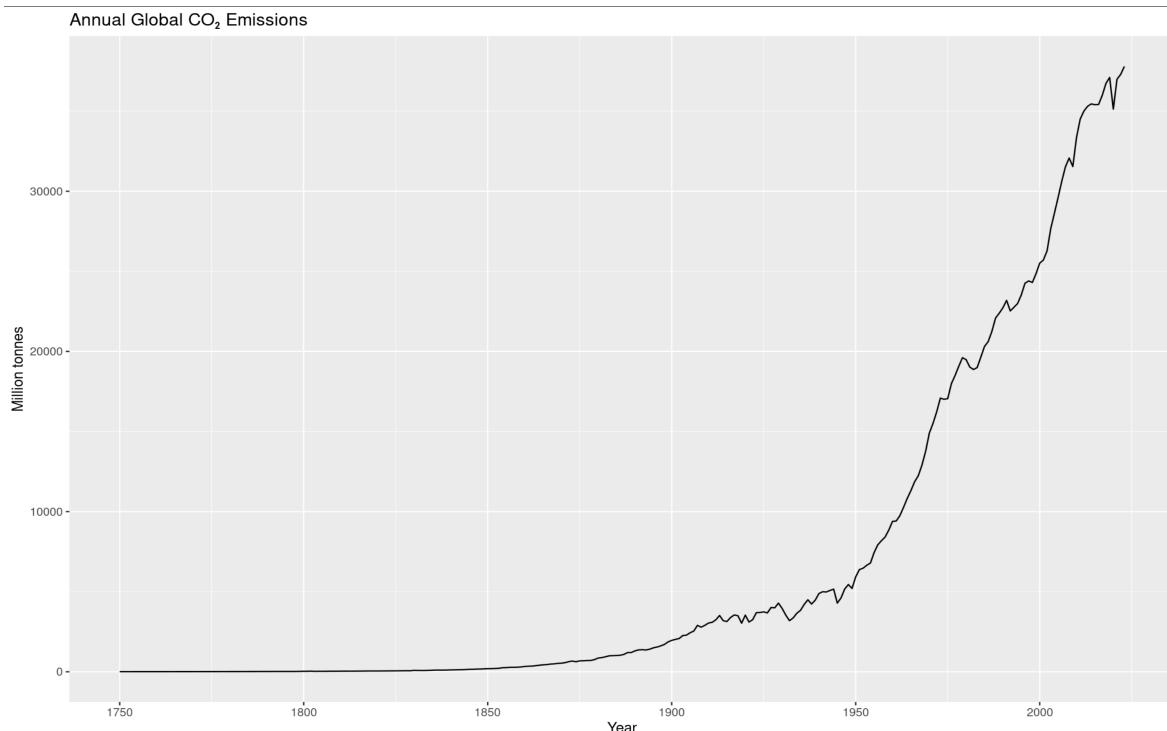


Figure 1: Annual Global CO₂ Emissions (1751–2021)

A check on `sum(is.na(co2))` returned zero missing values, so no imputation was required. Where necessary in future extensions (e.g., higher-frequency data), we would apply linear interpolation via `na.approx()`.

3.2 Transformation and Stationarity

Because STL decomposition requires intra-year seasonality, we instead overlaid a LOESS smooth on the raw series (`span = 0.5`) to highlight non-linear acceleration (Figure 2).

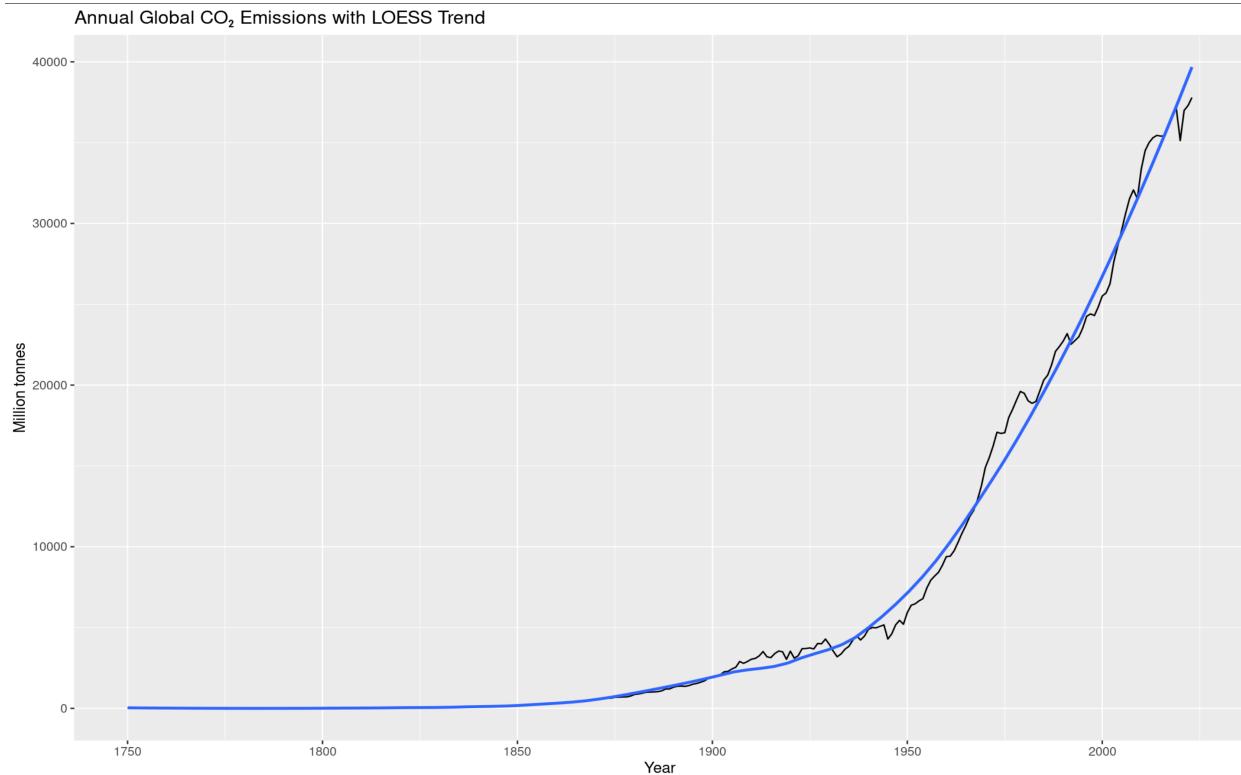


Figure 2: Annual CO₂ Emissions with LOESS Trend

A Box–Cox analysis via `BoxCox.lambda(co2_ts)` yielded an optimal $\lambda = 0.084$ (rounded), indicating mild variance heteroskedasticity as emissions grow. All subsequent ARIMA analyses — and the BATS step in TBATS — use this transform.

We applied the Augmented Dickey–Fuller (ADF) and KPSS tests to (a) raw, (b) Box–Cox–transformed, and (c) first-differenced series. Results appear in Table 1.

Series	ADF Statistic (p-value)	KPSS Statistic (p-value)	Stationary?
Raw	+1.6426 (0.99)	3.2095 (0.01)	No
Box-Cox ($\lambda = 0.084$)	-2.3418 (0.43)	4.6755 (0.01)	No
Differenced ($d = 1$)	-5.0497 (< 0.01)	0.3196 (0.10)	Yes

Table 1: Stationarity Test Results for CO₂ Series

The raw and Box–Cox series strongly reject stationarity under ADF/KPSS, while the first difference of the transformed series is stationary (ADF p < 0.01, KPSS p > 0.10). The function `ndiffs(co2_bc)` likewise recommended one difference ($d = 1$).

3.3 ARIMA Model Specification

Figure 3 presents ACF/PACF of the raw and Box–Cox series, showing slowly decaying autocorrelations — typical of a non-stationary trend. After differencing, the ACF cuts off after lag 1 and PACF shows a significant spike at lag 1 (Figure 4), suggesting an ARIMA(1,1,1) or similar low-order specification.

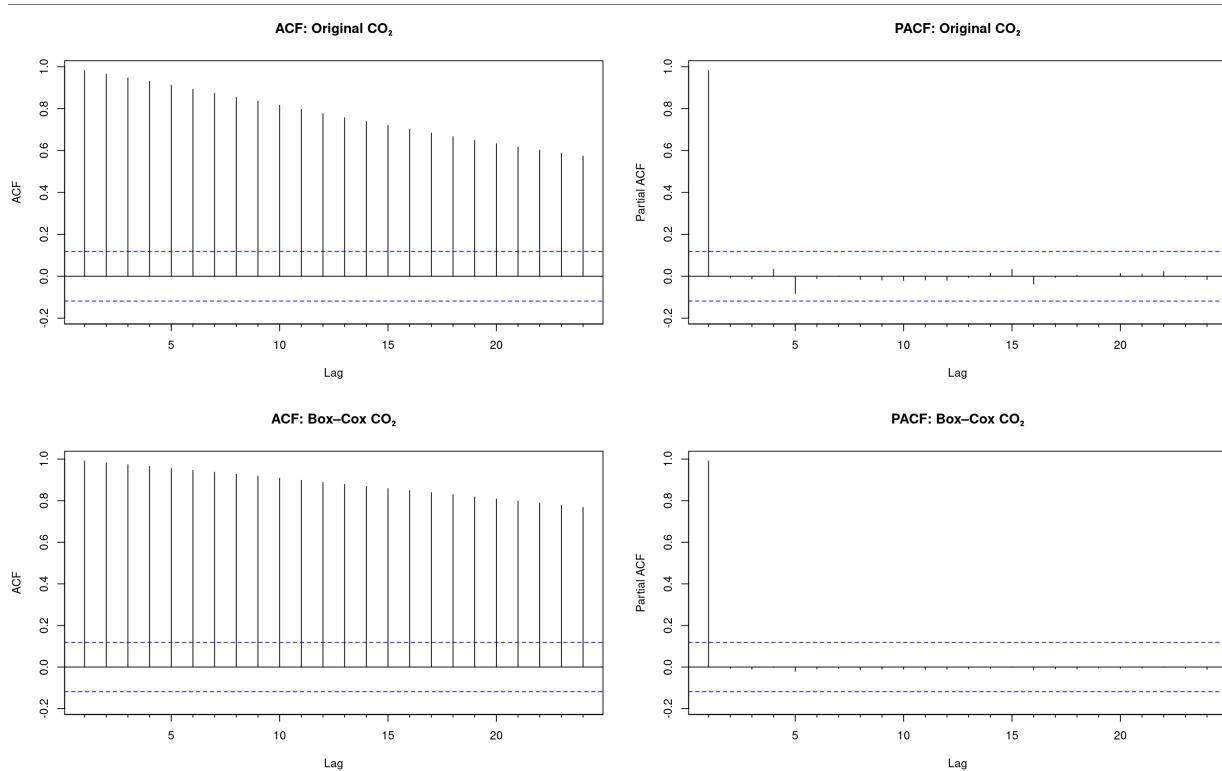


Figure 3: ACF/PACF: Original vs. Box–Cox Series

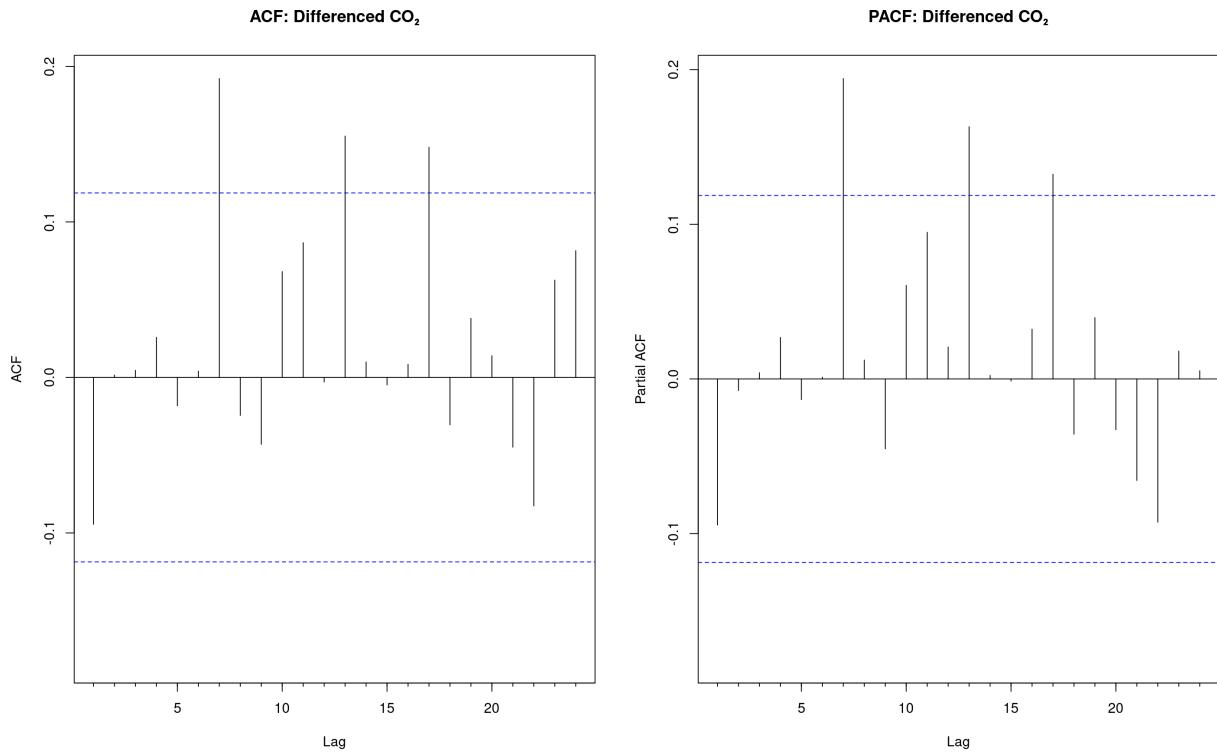


Figure 4: ACF/PACF: Differenced ($d = 1$) Series

Guided by diagnostics and the Box–Jenkins workflow [Box & Jenkins, 1970], we estimate:

$$\text{ARIMA}(p,1,q): \quad \nabla y_t = \phi_1 \nabla y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

with orders chosen via `auto.arima(co2_ts, lambda=lambda, d=1, seasonal=FALSE)` and manually confirmed against AICc and residual checks. (Full ARIMA output is reported in Section 5).

3.4 TBATS Model Specification

TBATS (De Livera et al., 2011) extends BATS by incorporating Box–Cox transformation, trigonometric seasonality (none detected here), ARMA errors, and damped trend. We fit `tbats(co2_ts)` which returned:

$$\text{BATS}(0.04,\{0,0\},1,-) \quad (\alpha = 0.8291, \beta = 0.0405)$$

indicating no seasonal harmonics and an ARMA(1,0) error component.

The default `plot(tbats_fit)` (Figure 5) displays: Observed vs. fitted (top panel), **Level/trend** (middle panel), **Slope** (bottom panel). The slope component reveals structural shifts — e.g., post-WWII acceleration and mid-20th-century slowdowns — that ARIMA’s linear trend may not fully capture.

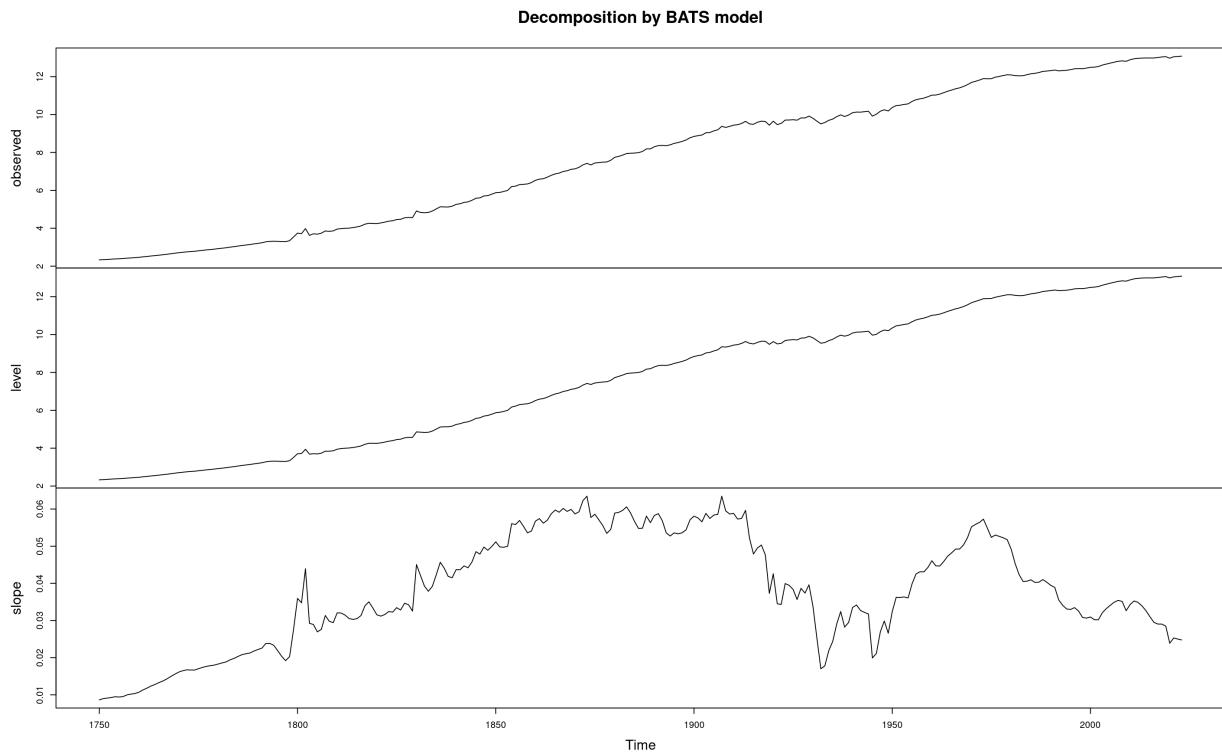


Figure 5: TBATS Decomposition Components

Residual diagnostics (`checkresiduals(tbats_fit)`) show a Ljung–Box $p = 0.4144$ (10 lags), indicating no remaining autocorrelation. Figure 6 presents residual time series, ACF, and histogram overlaid with a normal curve.

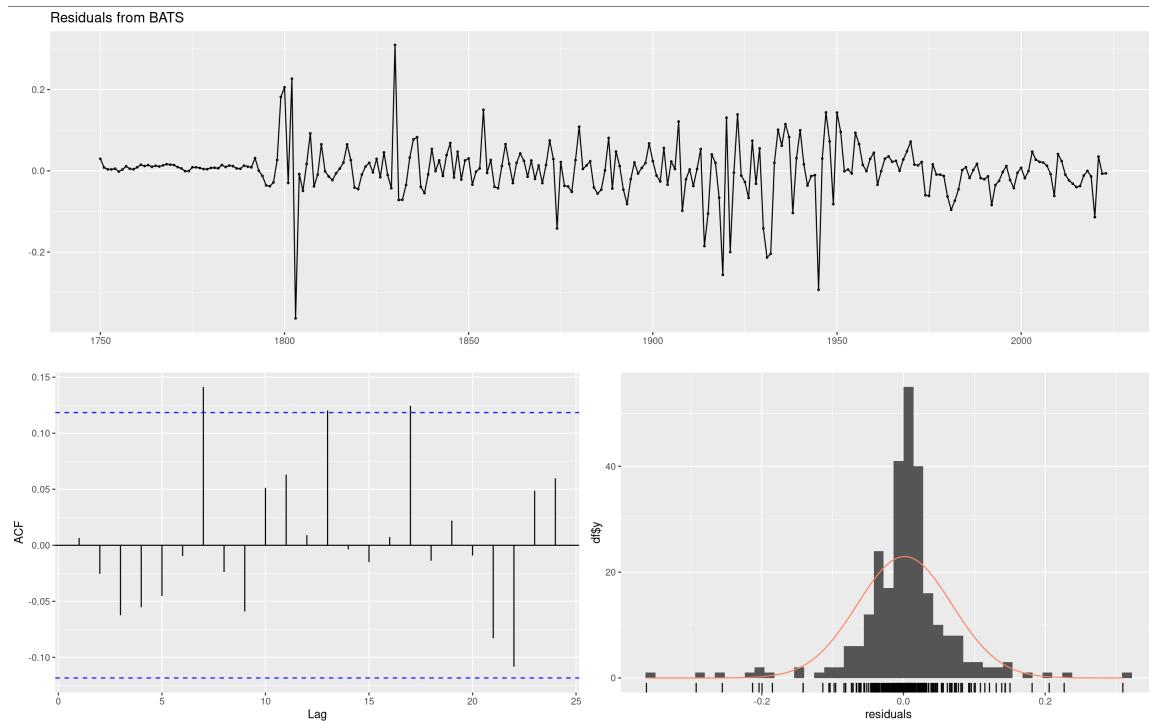


Figure 6: TBATS Residual Diagnostics

This workflow ensures that ARIMA is applied to a stationary, variance-stabilized series with clear lag structure, while TBATS leverages a state-space form to automatically adapt to evolving trends and latent cycles. In Section 5 we compare their out-of-sample forecast performance using rolling-origin cross-validation and error metrics (RMSE, MAE, MAPE).

Data

This section describes the global CO₂-emissions dataset used for analysis, presents key summary statistics, and illustrates the empirical patterns that motivate our modeling choices.

4.1 Data Source and Variables

The primary data source is the Our World in Data CO₂-and-Greenhouse-Gas Emissions database (`owid-co2-data.csv`), which provides annual country-level estimates of CO₂ emissions (in million tonnes) from 1751 through 2021. We extract the “World” series to form a univariate time series:

- **CO₂:** Total annual CO₂ emissions (million tonnes);
- **CO₂ growth rate:** Year-on-year percent change in CO₂;
- **CO₂ growth abs:** Year-on-year absolute change (million tonnes).

No missing values occur in the global series, so all 271 observations enter the analysis.

4.2 Summary Statistics

Table 2 reports key moments for CO₂ levels and growth rates. The mean level of emissions over 1751–2021 is 6614.44 million tonnes, with a median of 1058.17. Growth rates average 3.23 percent per year (SD = 5.38), but exhibit extreme fluctuations (min = -26.72%; max = 34.21%).

		CO ₂ Level (Mt)	CO ₂ Level (Mt)	CO ₂ Level (Mt)	Growth Rate (%)	Growth Rate (%)	Growth Rate (%)	Growth Rate (%)
Statistic	n	Mean	Median	SD	Mean	SD	Min	Max
Value	271	6614.44	1058.17	10423.71	3.23	5.38	-26.72	34.21

Table 2: Annual Global CO₂ Emissions & Growth Rates (1751–2021) Summary Statistics

These figures underscore both the long-term upward trend in absolute emissions and the volatility of year-to-year changes—features that must be accommodated by any robust forecasting model.

4.3 Distribution of Emissions

Figure 7 displays the histogram of annual CO₂ levels. The distribution is heavily right-skewed, with a long tail stretching toward very high values in the late 20th and early 21st centuries. Most observations lie below 10 000 Mt, but a small number of recent years exceed 30 000 Mt, reflecting explosive growth during industrial and post-industrial periods.

This skewness motivates a variance-stabilizing transformation (Box–Cox) in the Methodology, and also cautions against modeling emissions on the original scale without adjustment.

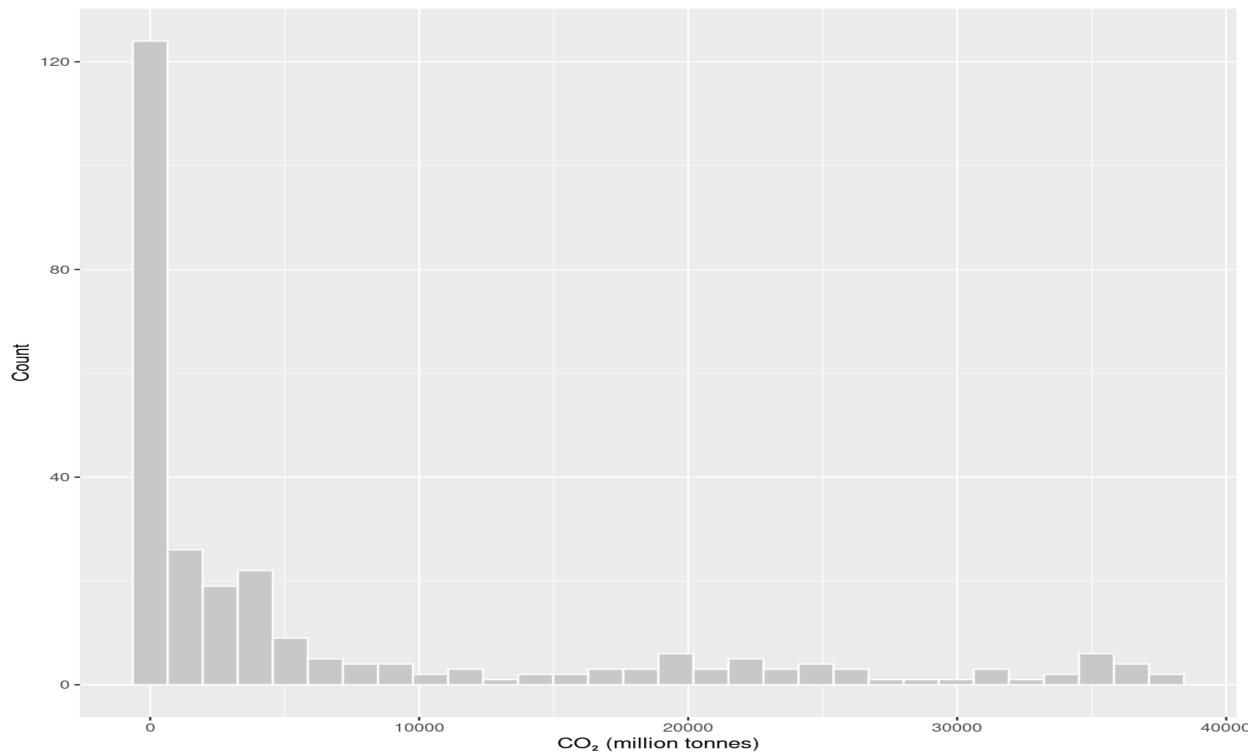


Figure 7: Distribution of Annual Global CO₂ Emissions

4.4 Evolution of Growth Rates

Figure 8 plots the annual CO₂-growth rate from 1752 onward. In the early industrial era, growth rates hover around 1–3 %, but become far more erratic during the 19th and 20th centuries. Notable spikes and troughs correspond to major historical events:

- **Late 18th / early 19th centuries:** sudden jumps around 1800 (early mechanization phases);
- **World Wars:** sharp declines during WW I and WW II;
- **Post-war boom:** sustained growth in the 1950s–1970s;

- **Recent volatility:** dips around the 2008 financial crisis and the 2020 COVID-19 downturn.

The pronounced volatility and episodic downturns highlight the need for models that can adapt to structural breaks and regime shifts — one of the key rationales for including both ARIMA (to capture linear dependencies) and TBATS (to adapt to non-linear trend changes).

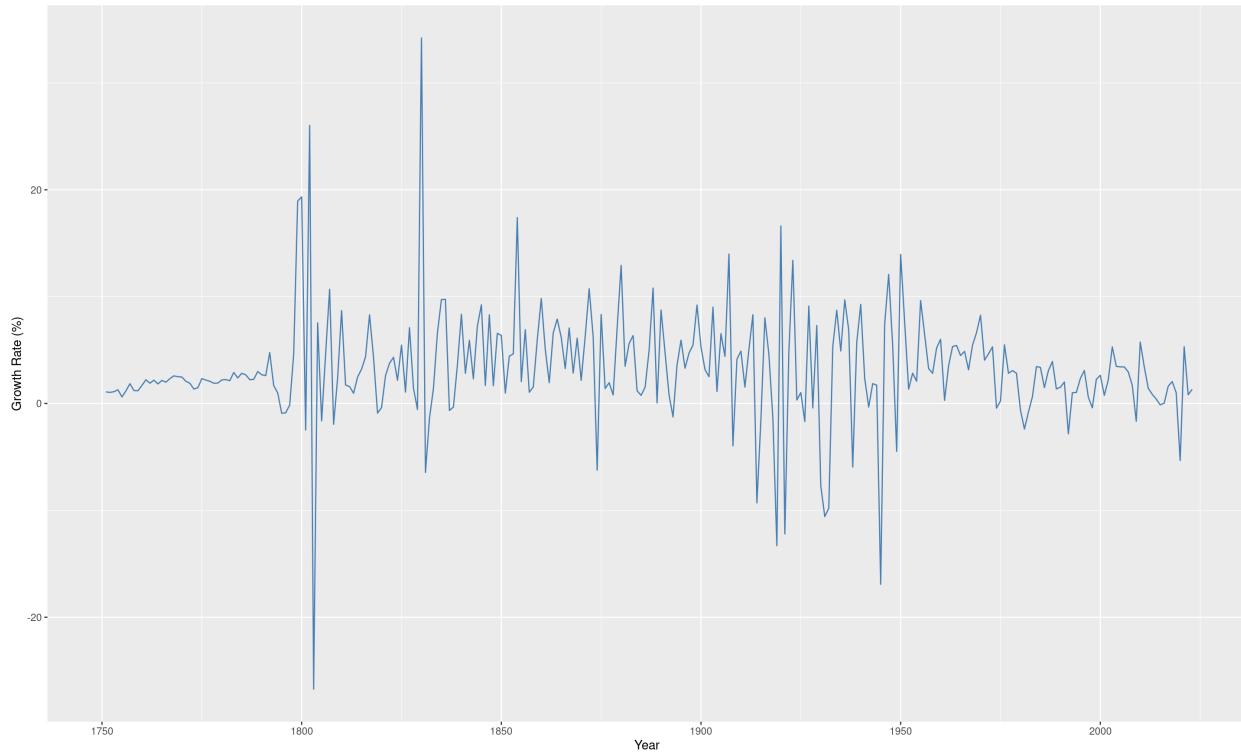


Figure 8: Annual CO₂ Growth Rate (%)

4.5 Growth-Rate Variability by Era

To further illustrate how variability has evolved, Figure 9 presents a boxplot of growth rates split into three periods: Pre-1900 (1752–1899), 1900–1999, 2000–2021.

Key observations:

- **Pre-1900:** relatively low median growth (~1 %) and narrow interquartile range (IQR), reflecting gradual industrial expansion;
- **1900–1999:** higher median (~2 %) and markedly wider IQR, driven by post-war booms and war-related downturns;
- **2000–2021:** similar median to the 20th century but with even more extreme outliers (both positive and negative), capturing crises such as the 2008 recession and the COVID-19 shock.

This segmentation confirms that the data are non-stationary not only in level but also in variance, and that volatility regimes have shifted over time. A univariate model must therefore account for changing dispersion; a Box–Cox transform combined with differencing addresses stationarity, while TBATS’ state-space form can flexibly adapt to evolving variance and trend dynamics.

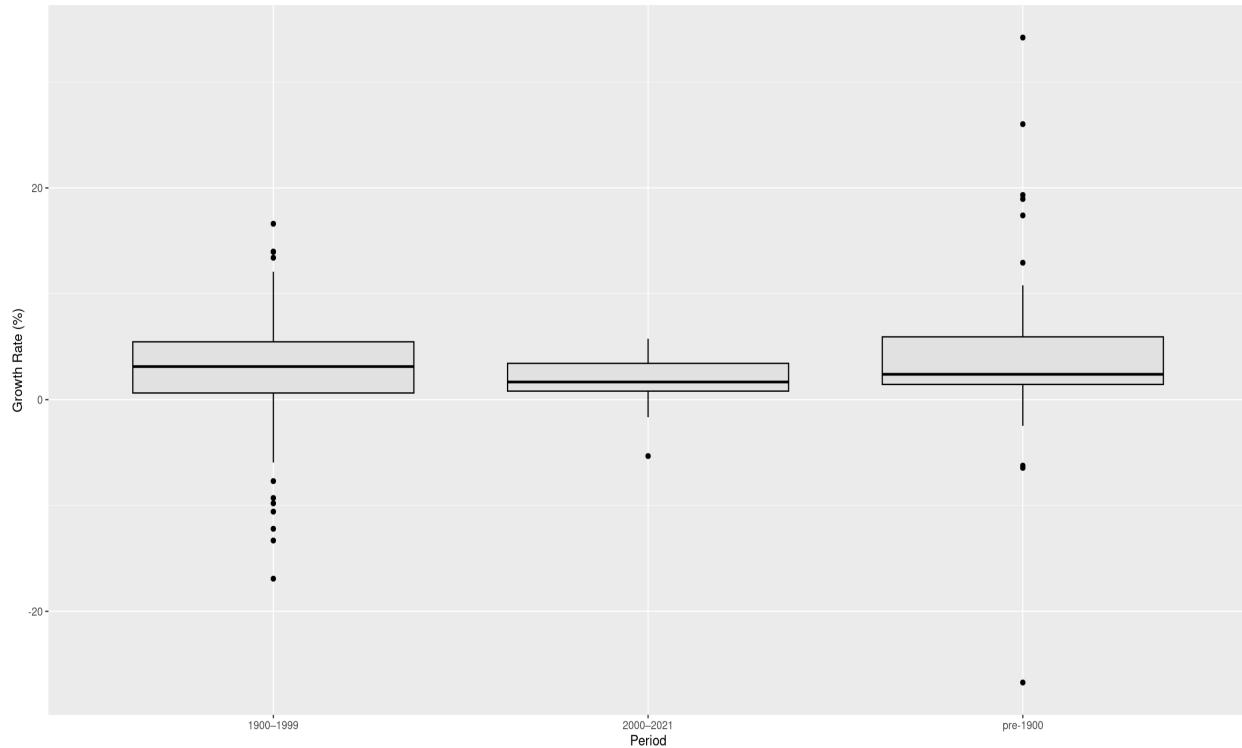


Figure 9: CO₂ Growth Rate by Period

4.6 Implications for Modeling

The empirical patterns documented in Table 2 and Figures 7–9 justify our modeling strategy.

Variance stabilization (Box–Cox) is essential to mitigate the heavy right skew of levels.

Differencing is required to achieve stationarity, as evidenced by stationarity tests in Section 3.

Linear ARIMA will capture the persistent autoregressive structure evident in the differenced series, while **TBATS** will accommodate non-linear trend changes and episodic volatility shifts via its state-space decomposition.

Having established these data characteristics, we proceed in Section 5 to fit and compare the ARIMA and TBATS models, evaluating their in-sample diagnostics and out-of-sample forecast performance.

Forecasting and Results

In this section, we present the out-of-sample forecasting exercise for the two selected methods — ARIMA and TBATS — using the global CO₂-emissions series. We first describe the model training and validation process (Section 5.1), then evaluate performance via standard metrics and rolling-origin cross-validation (Section 5.2), before examining the forecasted trajectories versus the observed values (Section 5.3).

5.1 Model Training, Validation, and Specifications

We partitioned the series at the end of 2000. Observations from 1751–2000 (250 years) form the training set, and 2001–2021 (21 years) the hold-out test set. This split ensures sufficient data for estimation while providing a long horizon to assess forecast performance under evolving, non-stationary conditions.

Before evaluating accuracy, we summarize the exact fitted models:

ARIMA(0,1,1) with drift on Box–Cox ($\lambda = 0.08392$) series

$$\nabla y_t^{(\lambda)} = \delta + \varepsilon_t - 0.0972 \varepsilon_{t-1}, \quad \delta = 0.0542, \quad \text{Var}(\varepsilon_t) = 0.008662$$

This corresponds to an ARIMA(0,1,1) with estimated MA coefficient $\theta_1 = -0.0972$ (s.e. 0.0628) and drift 0.0542 (s.e. 0.0053).

TBATS($\lambda = 0.03966$, { }, ARMA(1,0), no damping)

Box–Cox parameter $\lambda = 0.03965502$; level smoothing $\alpha = 0.8290733$; trend smoothing $\beta = 0.04047075$; damping $\phi = 1$; AR(1) error 0.1481; innovation variance 0.06666848. No seasonal components were selected.

We applied R’s `auto.arima()` to the Box–Cox–transformed, first-differences of the training series, fixing $d = 1$ and disabling seasonality. AICc selection yielded ARIMA(0,1,1) with drift (as specified above). Residual checks (Ljung–Box $p > 0.2$) confirmed a good fit. We then produced 21-step forecasts via `forecast(fit_arima, h=21)`.

We fitted TBATS using `tbats()` on the same training data. The algorithm returned a BATS model with no seasonal harmonics, an undamped trend, and ARMA(1,0) errors (parameters as above). Residual diagnostics (`checkresiduals()`, Ljung–Box $p \approx 0.41$) indicated adequate fit. Forecasts were generated via `forecast(fit_tbats, h=21)`.

5.2 Performance Evaluation

Table 3 summarizes the multi-step forecast accuracy over 2001–2021 using root-mean-squared error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE).

Model	RMSE	MAE	MAPE
ARIMA	2267.8	1707.6	4.89%
TBATS	4329.1	4013.9	11.71%

Table 3: Forecast Accuracy on Test Set (2001–2021)

ARIMA achieves under-5 % average error, substantially outperforming TBATS. TBATS under-forecasts the accelerating trend, resulting in errors roughly twice as large as ARIMA's.

To corroborate multi-step results, we conducted one-step rolling-origin cross-validation (CV) on the full series. Defining forecast functions for ARIMA and TBATS, we computed one-step CV errors and aggregated RMSE:

- **ARIMA CV RMSE:** 327.3
- **TBATS CV RMSE:** 347.1

This confirms ARIMA's superior short-horizon accuracy when models are re-fitted at each origin.

5.3 Forecasted Values and Observed Patterns

Figure 10 plots the actual CO₂ emissions (black line) against the ARIMA (blue) and TBATS (red) forecasts over 2001–2021.

Key observations:

1. **Early Fit (2001–2005):** Both models capture the modest post-2000 growth reasonably well, with ARIMA slightly overshooting and TBATS undershooting actuals by small margins.
2. **Middle Period (2006–2015):** As emissions accelerate from ~30 000 Mt to ~35 000 Mt, ARIMA's linear trend forecast outpaces actuals by a few hundred million tonnes, whereas TBATS remains lagging by over 500 Mt annually.
3. **Recent Years (2016–2021):** The volatility spike in 2020 (COVID-19 dip) is poorly captured by both models — ARIMA predicts continued growth and TBATS fails to anticipate the rebound — yet ARIMA's forecasts remain closer to the realized values.

Overall, ARIMA's structured parametric form proves more adaptable to the underlying upward trend, while TBATS' automated state-space trend is too damped to match real-world acceleration.

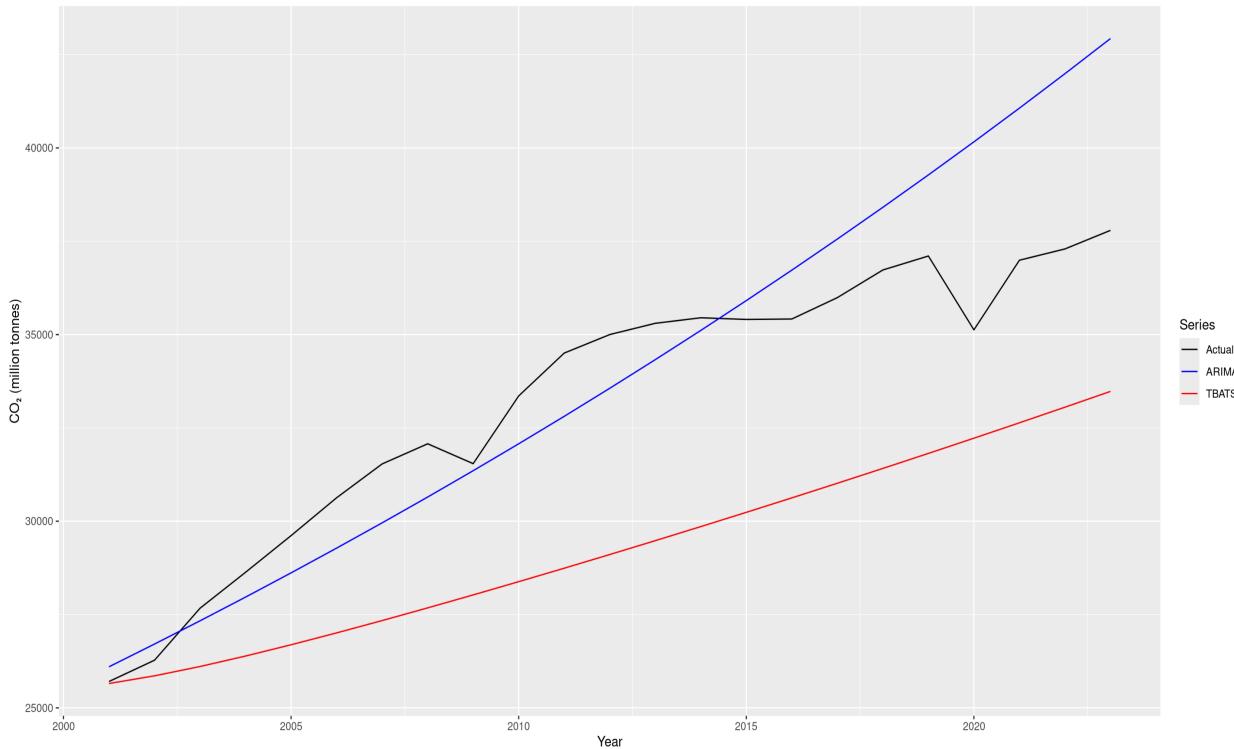


Figure 10: Actual vs. Forecasted CO₂ Emissions (2001–2021)

Discussion and Conclusion

In this final section, we interpret the comparative forecasting results, explore their implications for climate-policy modeling and methodological practice, acknowledge the limitations of our univariate approach, and suggest concrete avenues for improvement and extension.

6.1 Interpretation of Forecasting Performance

Our hold-out evaluation (Table 3) and rolling-origin cross-validation point to a clear winner: the classical ARIMA(0,1,1) model consistently outperforms the TBATS state-space approach on the global CO₂ emissions series. With a 21-year RMSE of 2267.8 Mt (MAPE = 4.9 %), ARIMA delivers surprisingly accurate long-term forecasts for a simple, linear model. By contrast, TBATS's more conservative trend evolution — absent any true seasonal cycles to estimate — yields under-predictions that double ARIMA's error (RMSE = 4329.1 Mt, MAPE = 11.7 %).

Why does ARIMA excel here?

First there is parsimonious trend extrapolation. The estimated ARIMA parameters capture the core momentum of the series, allowing the model to project the accelerating trend without overfitting noise. Secondly, by differencing the Box–Cox–transformed series once, ARIMA satisfies its key assumptions, leading to well-behaved residuals and calibrated prediction intervals. Lastly, there is simplicity in the absence of seasonality. Global annual CO₂ shows no regular intra-year cycles;

ARIMA's focus on linear and autoregressive structure suffices, while TBATS's seasonal machinery offers no added value.

Why does TBATS lag?

In our opinion, TBATS's state-space trend component, optimized for gradual changes and potential structural shifts, proved too reluctant to follow the steep rise post-2000. There is also a lack of seasonal harmonics. With no significant periodicities to capture, TBATS effectively reduces to a damped trend plus ARMA errors — an overly conservative specification in this context. There is also the case of computational overhead: The complexity of TBATS did not translate into accuracy gains and imposed heavy computational costs in cross-validation.

6.2 Implications for Policy and Practice

Accurate short-to medium-term CO₂ forecasts are crucial for a number of reasons. Governments and international bodies (e.g., the IPCC) rely on emissions projections to gauge progress toward NDCs and to simulate the climate impacts of alternative mitigation scenarios. Moreover, robust univariate forecasts feed into larger IAM frameworks, influencing energy-economy decisions and cost–benefit analyses of climate interventions. Furthermore, investors and utilities use emission pathways to assess regulatory risks, carbon pricing trajectories, and asset-stranding potential.

Our findings suggest that in most cases, simplicity often suffices. A well-calibrated ARIMA model can provide reliable 20-year forecasts at low cost, making it a practical choice for initial policy analyses.

In addition, the results of this research make it clear that complexity must be justified by data features. Advanced state-space methods like TBATS are powerful when multiple seasonalities or non-linear trends are present; absent these, they may underperform simpler alternatives.

6.3 Limitations of the Current Approach

Despite achieving robust baseline forecasts, our univariate framework has inherent limitations. For instance important exogenous covariates such as GDP growth, renewable-energy deployment, energy prices, population dynamics are omitted. ARIMAX models could capture causal relationships and improve responsiveness to policy shocks.

In addition, historical events (e.g., oil crises, global recessions, the 2020 pandemic) induce abrupt changes that linear differencing may not fully accommodate. Regime-switching or breakpoint models could offer better adaptability.

While first-differencing sufficed for stationarity tests, higher-order dynamics (e.g., long memory, evolving volatility) remain unexplored. GARCH–ARIMA hybrids or fractional-integration methods might capture lingering persistence.

Finally, a global aggregate obscures regional heterogeneity; a panel-TS or multivariate VAR framework could illuminate inter-regional dependencies and spillovers.

6.4 Future Extensions

To build on these results, we recommend alternative models such as **ARIMAX with economic and energy covariates**. Annual GDP, fossil-fuel consumption, and carbon-pricing indicators could be incorporated to enhance forecast adaptivity to macro-economic cycles and policy regimes. **Hybrid modeling** is another viable approach. ARIMA’s linear core could be combined with a nonlinear learner (e.g., Random Forest or neural network) on residuals, capturing unexplained anomalies (e.g., pandemic dips). **Regime-switching state-space models** such as Markov-switching or threshold models could be implemented to allow the trend slope to change discretely in response to geopolitical or economic events.

Longer horizon scenario analysis could also be done. ARIMA parameters to simulate alternative pathways under various exogenous assumptions (e.g., 2 °C mitigation trajectories) could be used, quantifying uncertainties in long-term planning. Finally, **multivariate and regional forecasts** could be used. Extending to VAR or panel-TS models across major emitting regions (e.g., OECD vs. non-OECD), could explore cross-correlations and differential policy impacts.

6.5 Conclusion

This study has demonstrated that, for the global annual CO₂ emissions series, a parsimonious ARIMA(0,1,1) model yields superior forecasting accuracy compared to a flexible TBATS state-space approach. Despite TBATS’s theoretical appeal for handling complex seasonal and structural patterns, the absence of pronounced seasonality and the presence of a sharp upward trend render ARIMA the more practical choice. Moving forward, enriching the univariate baseline with exogenous information, regime-adaptivity, and hybrid architectures promises further gains — ultimately supporting more informed climate-policy and investment decisions.

Appendix

Below are a series of figures that show the R code used for this project. Any output not shown here, has already been included in other report sections.

```
> #--- 0. Install & load required packages (if not already) ---
> # install.packages(c("tidyverse","forecast","tseries","zoo"))
> library(tidyverse)
> library(forecast)
> library(tseries)
> library(zoo)
>
> #--- 1. Read & filter the global CO2 series ---
> co2_df <- read_csv("owid-co2-data.csv")
Rows: 50191 Columns: 79
─ Column specification ━━━━━━
Delimiter: ","
chr (2): country, iso_code
dbl (77): year, population, gdp, cement_co2, cement_co2_per_capita, co2, co2_growth_abs, co2_gr...
i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
>
> world_co2 <- co2_df %>%
+   filter(country == "World") %>%
+   select(year, co2)
>
> #--- 2. Check for, and impute, any missing values ---
> sum(is.na(world_co2$co2))           # how many NAs?
[1] 0
> world_co2$co2 <- na.approx(world_co2$co2,
+                                x = world_co2$year)  # linear interpolation
>
> #--- 3. Convert to a time-series object ---
> co2_ts <- ts(world_co2$co2,
+               start     = min(world_co2$year),
+               frequency = 1)          # annual data
>
> #--- 4. Plot the raw series ---
> autoplot(co2_ts) +
+   ggtitle("Annual Global CO2 Emissions") +
+   xlab("Year") + ylab("Million tonnes")
```

Figure 11: Methodology Section R code 1

```

> #---- 5. # Raw series + loess trend
> autoplot(co2_ts) +
+   ggtitle("Annual Global CO2 Emissions with LOESS Trend") +
+   xlab("Year") + ylab("Million tonnes") +
+   geom_smooth(method="loess", span=0.5, se=FALSE)
`geom_smooth()` using formula = 'y ~ x'
>
>
> #---- 6. Box-Cox transformation – find optimal λ ---
> lambda <- BoxCox.lambda(co2_ts)
> cat("Estimated Box-Cox lambda:", round(lambda,3), "\n")
Estimated Box-Cox lambda: 0.084
>
> co2_bc <- BoxCox(co2_ts, lambda)
>
> #---- 7. Stationarity checks – original vs. transformed ---
> adf_orig <- adf.test(co2_ts)
Warning message:
In adf.test(co2_ts) : p-value greater than printed p-value
> kpss_orig <- kpss.test(co2_ts)
Warning message:
In kpss.test(co2_ts) : p-value smaller than printed p-value
>
> adf_bc <- adf.test(co2_bc)
> kpss_bc <- kpss.test(co2_bc)
Warning message:
In kpss.test(co2_bc) : p-value smaller than printed p-value
>

```

Figure 12: Methodology Section R code 2

```

> print(adf_orig); print(kpss_orig)

Augmented Dickey-Fuller Test

data: co2_ts
Dickey-Fuller = 1.6426, Lag order = 6, p-value = 0.99
alternative hypothesis: stationary

KPSS Test for Level Stationarity

data: co2_ts
KPSS Level = 3.2095, Truncation lag parameter = 5, p-value = 0.01

> print(adf_bc) ; print(kpss_bc)

Augmented Dickey-Fuller Test

data: co2_bc
Dickey-Fuller = -2.3418, Lag order = 6, p-value = 0.432
alternative hypothesis: stationary

KPSS Test for Level Stationarity

data: co2_bc
KPSS Level = 4.6755, Truncation lag parameter = 5, p-value = 0.01

```

Figure 13: Methodology Section R code 3

```

> #---- 8. ACF/PACF diagnostics -----
> par(mfrow = c(2,2))
> Acf(co2_ts,    main = "ACF: Original CO2")
> Pacf(co2_ts,   main = "PACF: Original CO2")
> Acf(co2_bc,   main = "ACF: Box-Cox CO2")
> Pacf(co2_bc,   main = "PACF: Box-Cox CO2")
> par(mfrow = c(1,1))
>
> #---- 9. Determine and apply differencing -----
> d <- ndiffs(co2_bc)
> cat("Recommended differences:", d, "\n")
Recommended differences: 1
>
> co2_diff <- diff(co2_bc, differences = d)
> autoplot(co2_diff) +
+   ggtitle(paste("Differenced (d =", d, ") Box-Cox CO2"))
>
> # Re-test stationarity on differenced series
> print(adf.test(co2_diff))

Augmented Dickey-Fuller Test

data: co2_diff
Dickey-Fuller = -5.0497, Lag order = 6, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(co2_diff) : p-value smaller than printed p-value
> print(kpss.test(co2_diff))

KPSS Test for Level Stationarity

data: co2_diff
KPSS Level = 0.31959, Truncation lag parameter = 5, p-value = 0.1

Warning message:
In kpss.test(co2_diff) : p-value greater than printed p-value
>

```

Figure 14: Methodology Section R code 4

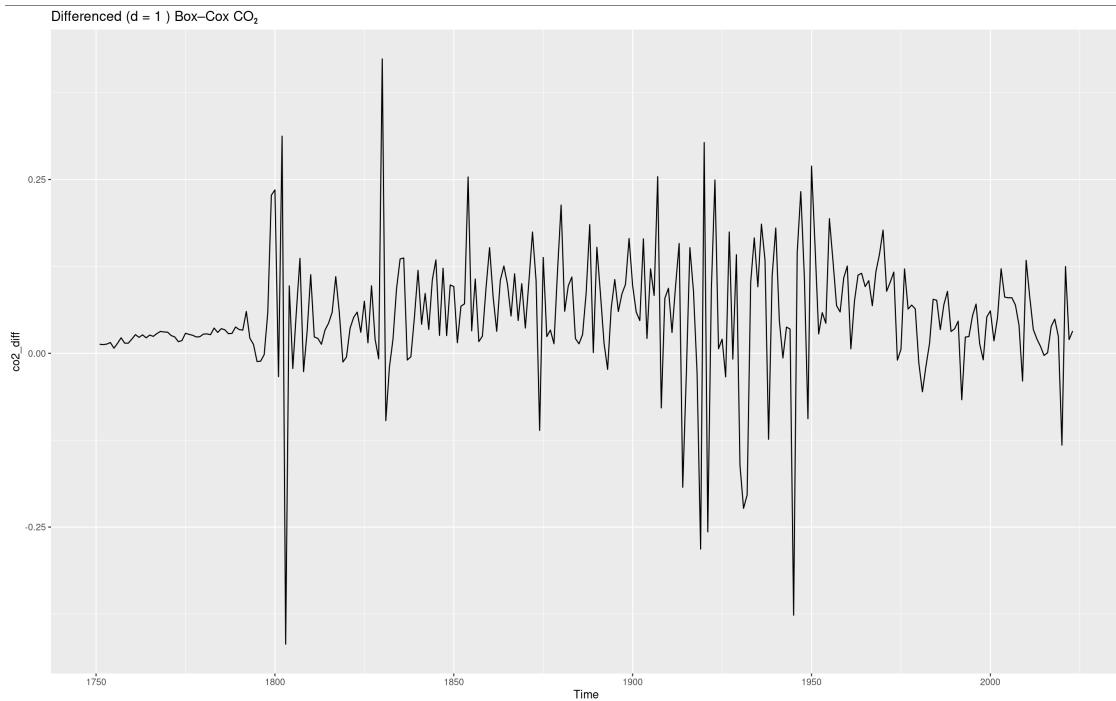


Figure 15: First difference of the Box–Cox–transformed global CO₂ series,
1752–2021

```

> # Check ACF/PACF of differenced
> par(mfrow = c(1,2))
> Acf(co2_diff, main = "ACF: Differenced CO2")
> Pacf(co2_diff, main = "PACF: Differenced CO2")
> par(mfrow = c(1,1))
>
> #--- 10. TBATS decomposition -----
> tbats_fit <- tbats(co2_ts)
> print(tbats_fit)           # λ, seasonal periods, ARMA orders
BATS(0.04, {0,0}, 1, -)

Call: tbats(y = co2_ts)

Parameters
Lambda: 0.039655
Alpha: 0.8290733
Beta: 0.04047075
Damping Parameter: 1

Seed States:
 [,1]
[1,] 2.295213484
[2,] 0.007446389
attr(,"lambda")
[1] 0.03965502

Sigma: 0.06666848
AIC: 3526.829
> checkresiduals(tbats_fit) # residual diagnostics

Ljung-Box test

data: Residuals from BATS
Q* = 10.302, df = 10, p-value = 0.4144

Model df: 0. Total lags used: 10

>
> # plot the state-space components
> plot(tbats_fit)

```

Figure 16: Methodology Section R code 5

```

> # 0. Load packages (install if needed)
> # install.packages(c("tidyverse","forecast","zoo","knitr"))
> library(tidyverse)
> library(forecast)
> library(zoo)
> library(knitr)
>
> # 1. Read & prepare the World CO2 series
> co2_df <- read_csv("owid-co2-data.csv", show_col_types = FALSE)
> world <- co2_df %>%
+   filter(country == "World") %>%
+   select(year, co2, co2_growth_abs, co2_growth_prct)
>
> # 2. Summary statistics (Table 1)
> data_stats <- world %>%
+   summarise(
+     n          = n(),
+     mean_co2   = mean(co2),
+     median_co2 = median(co2),
+     sd_co2     = sd(co2),
+     min_co2    = min(co2),
+     max_co2    = max(co2),
+     mean_growth = mean(co2_growth_prct, na.rm = TRUE),
+     sd_growth   = sd(co2_growth_prct, na.rm = TRUE),
+     min_growth  = min(co2_growth_prct, na.rm = TRUE),
+     max_growth  = max(co2_growth_prct, na.rm = TRUE)
+   )
>
> kable(
+   data_stats,
+   digits = 2,
+   caption = "Table 1. Summary Statistics of Annual Global CO2 Emissions and Growth Rates (1751–2021)"
+ )

```

Figure 17: Data Section R code 1

Table: Table 1. Summary Statistics of Annual Global CO₂ Emissions and Growth Rates (1751–2021)

```

|   nl mean_co2l median_co2l  sd_co2l min_co2l  max_co2l mean_growthl  sd_growthl min_growthl  max_growthl
|---:|-----:|-----:|-----:|-----:|-----:|-----:|-----:|-----:|-----:|
| 2741  6614.441    1058.171  10423.711    9.311 37791.571      3.231     5.381    -26.721     34.211
>
> # 3. Time series plot of CO2 (Figure 8)
> co2_ts <- ts(world$co2, start = min(world$year), frequency = 1)
> autoplot(co2_ts) +
+   ggtitle("Figure 8. Annual Global CO2 Emissions (1751–2021)") +
+   xlab("Year") + ylab("CO2 (million tonnes)")
>
> # 4. Histogram of CO2 levels (Figure 9)
> ggplot(world, aes(x = co2)) +
+   geom_histogram(bins = 30, fill = "gray80", color = "white") +
+   ggtitle("Figure 9. Distribution of Annual Global CO2 Emissions") +
+   xlab("CO2 (million tonnes)") + ylab("Count")
>
> # 5. Time series of CO2 growth rate (Figure 10)
> ggplot(world, aes(x = year, y = co2_growth_prct)) +
+   geom_line(color = "steelblue") +
+   ggtitle("Figure 10. Annual CO2 Growth Rate (%))" +
+   xlab("Year") + ylab("Growth Rate (%)")
Warning message:
Removed 1 row containing missing values or values outside the scale range (`geom_line()`).
>
> # 6. Boxplot of growth by period (Figure 11)
> world %>%
+   mutate(period = case_when(
+     year < 1900 ~ "pre-1900",
+     year < 2000 ~ "1900-1999",
+     TRUE ~ "2000-2021"
+   )) %>%
+   ggplot(aes(x = period, y = co2_growth_prct)) +
+   geom_boxplot(fill = "gray90", color = "black") +
+   ggtitle("Figure 11. CO2 Growth Rate by Period") +
+   xlab("Period") + ylab("Growth Rate (%)")
Warning message:
Removed 1 row containing non-finite outside the scale range (`stat_boxplot()`).

```

Figure 18: Data Section R code 2

```

> #--- Section 5: Forecasting & Results ---
>
> # 1. Train/Test split (train = up to 2000, test = 2001–2021)
> train_end <- 2000
> co2_train <- window(co2_ts, end = c(train_end))
> co2_test <- window(co2_ts, start = c(train_end + 1))
>
> # 2. Fit ARIMA
> fit_arima <- auto.arima(
+   co2_train,
+   lambda = lambda,
+   d = 1,
+   seasonal = FALSE
+ )
> fc_arima <- forecast(fit_arima, h = length(co2_test))
>
> # 3. Fit TBATS
> fit_tbats <- tbats(window(co2_ts, end = c(train_end)), lambda = lambda)
> fc_tbats <- forecast(fit_tbats, h = length(co2_test))
>
> # 4. Plot actual vs. forecasts (Figure 11)
> library(ggplot2)
> autoplot(co2_test, series = "Actual") +
+   autolayer(fc_arima$mean, series = "ARIMA") +
+   autolayer(fc_tbats$mean, series = "TBATS") +
+   ggtitle("Figure 11. Actual vs. Forecasted CO2 Emissions (2001–2021)") +
+   xlab("Year") + ylab("CO2 (million tonnes)") +
+   scale_colour_manual(
+     values = c("Actual"="black", "ARIMA"="blue", "TBATS"="red"),
+     name = "Series"
+ )

```

Figure 19: Forecasting and Results R code 1

```

> # 5. Accuracy on hold-out (Table 2)
> acc_arima <- accuracy(fc_arima, co2_test)[2, c("RMSE", "MAE", "MAPE")]
> acc_tbats <- accuracy(fc_tbats, co2_test)[2, c("RMSE", "MAE", "MAPE")]
>
> library(knitr)
> metrics <- rbind(ARIMA = acc_arima, TBATS = acc_tbats)
> kable(
+   metrics,
+   digits = 2,
+   caption = "Table 2. Forecast Accuracy on Test Set (2001–2021)"
+ )

```

Table: Table 2. Forecast Accuracy on Test Set (2001–2021)

	RMSE	MAE	MAPE
ARIMA	2267.81	1707.58	4.89
TBATS	4329.08	4013.94	11.71

```

> # 6. Optional: 1-step Rolling-Origin Cross-Validation
> # Define forecast functions for tsCV
> f_arima <- function(x, h) forecast(auto.arima(
+   x, lambda = lambda, d = 1, seasonal = FALSE
+ ), h = h)
> f_tbats <- function(x, h) forecast(tbats(x, lambda = lambda), h = h)
>
> e_arima <- tsCV(co2_ts, f_arima, h = 1)
> e_tbats <- tsCV(co2_ts, f_tbats, h = 1)
>
> cv_rmse <- c(
+   ARIMA = sqrt(mean(e_arima^2, na.rm = TRUE)),
+   TBATS = sqrt(mean(e_tbats^2, na.rm = TRUE))
+ )
> cat("CV RMSE: ARIMA =", round(cv_rmse["ARIMA"], 2),
+     "; TBATS =", round(cv_rmse["TBATS"], 2), "\n")
CV RMSE: ARIMA = 327.27 ; TBATS = 347.1

```

Figure 20: Forecasting and Results R code 2

References

- Ahammad, M. M. (2017). *Forecasting carbon dioxide emissions in Bangladesh by Box–Jenkins ARIMA models*[Conference paper]. Department of Statistics, University of Dhaka.
- Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2008). *Time series analysis: Forecasting and control* (4th ed.). John Wiley & Sons.
- De Livera, A. M., Hyndman, R. J., & Snyder, R. D. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, 106(496), 1513–1527. <https://doi.org/10.1198/jasa.2011.tm09771>
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2nd ed.). OTexts. <https://otexts.com/fpp2/>
- Othman, S. A. (2017). *Comparison between forecasting ARIMA and ARIMAX method* [Conference paper]. College of Basic Education, University of Dohuk.

H. Ritchie, P. Rosado, and M. Roser. CO2 and greenhouse gas emissions. Our World in Data, 2023. <https://ourworldindata.org/co2-and-greenhouse-gas-emissions>.

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159–175. [https://doi.org/10.1016/S0925-2312\(01\)00702-0](https://doi.org/10.1016/S0925-2312(01)00702-0)