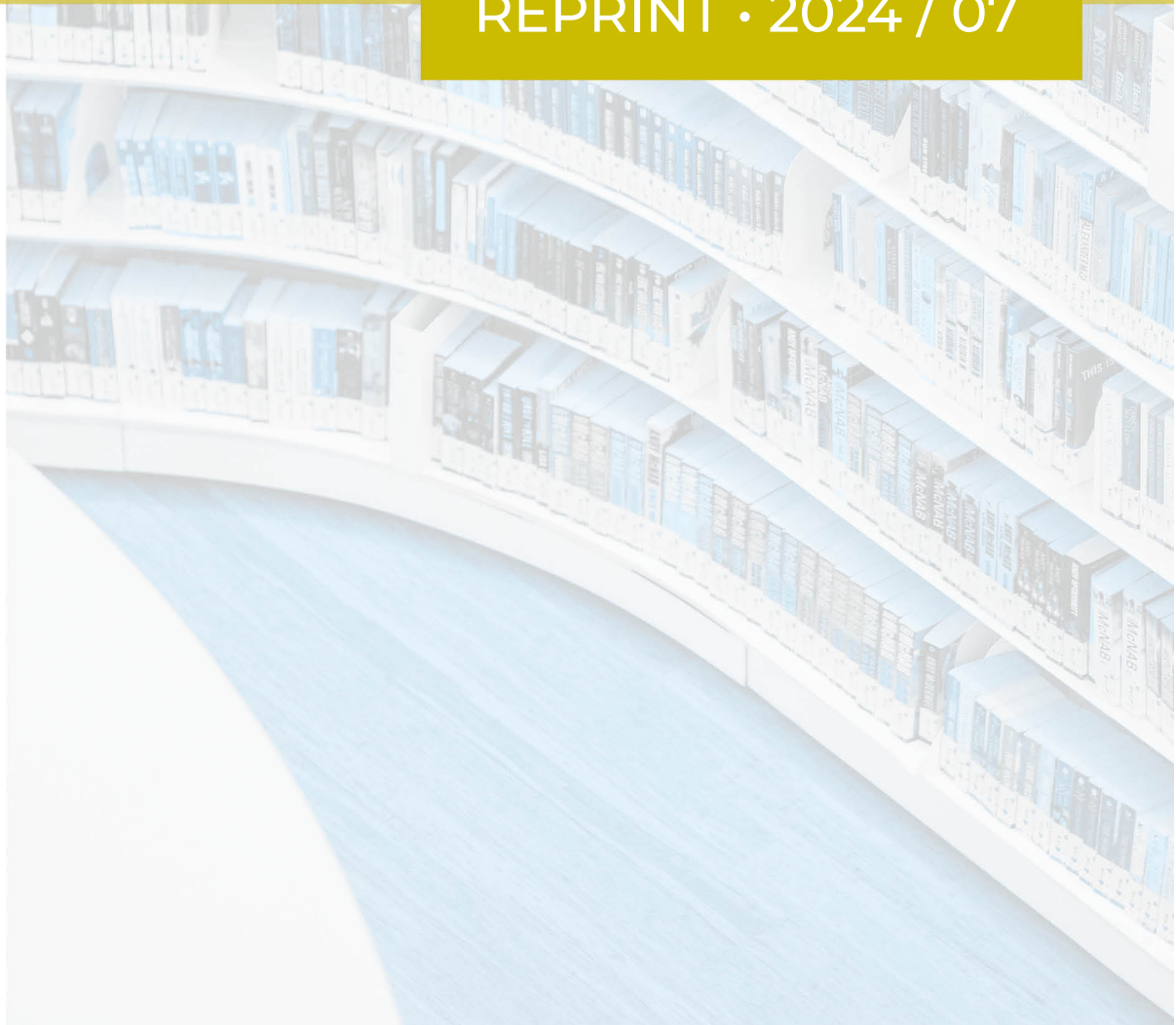


LOOKING AHEAD: FORECASTING TOTAL ENERGY CARBON DIOXIDE EMISSIONS

Bernardina Algieri, Leonardo Iania, Arturo
Leccadito

REPRINT • 2024 / 07



LFIN

Voie du Roman Pays 34, L1.03.01 B-1348 Louvain-la-Neuve

Tel (32 10) 47 43 04

Email: lidam-library@uclouvain.be

<https://uclouvain.be/en/research-institutes/lidam/lfin/publications.html>



Looking ahead: Forecasting total energy carbon dioxide emissions

Bernardina Algieri^{a,b,*}, Leonardo Iania^{c,d}, Arturo Leccadito^{a,c}

^a University of Calabria, Department of Economics, Statistics and Finance, Ponte Bucci, Rende, CS, 87036, Italy

^b Zentrum für Entwicklungsforschung, ZEF, Universität Bonn, Walter-Flex-Strasse 3, Bonn, 53113, Germany

^c CORE/LFIN, Université catholique de Louvain, Voie du Roman Pays 34, B-1348, Louvain-la-Neuve, Belgium

^d KU Leuven, Department Accounting, Finance and Insurance, Naamsestraat 69, 3001, Leuven, Belgium

ARTICLE INFO

JEL classification:

C53
C55
E71
Q47
Q53

Keywords:

CO₂ emissions
Forecasting models
Quantile factors
Economic
Sentiment and nature-related drivers

ABSTRACT

In recent years, the international community has been increasing its efforts to reduce the human footprint on air pollution and global warming. Total carbon dioxide (CO₂) releases are a crucial component of global greenhouse gas emissions, and as such, they are closely monitored at the national and supranational levels. This study presents different models to forecast energy CO₂ emissions for the US in the period 1972–2021, using quarterly observations. In an in-sample and out-of-sample analysis, the study assesses the accuracy of thirteen forecasting models (and their combinations), considering an extensive set of potential predictors (more than 260) that include macroeconomic, nature-related factors and different survey data and compares them to traditional benchmarks. To reduce the high-dimensionality of the potential predictors, the study uses a new class of factor models in addition to the classical principal component analysis. The results show that economic variables, market sentiment and nature-related indicators, especially drought and Antarctic wind indicators, help forecast short/medium-term CO₂ emissions. In addition, some combinations of models tend to improve out-of-sample predictions.

1. Introduction

Global warming and climate change are vital challenges that society is currently facing. Over the recent decades, economies worldwide have recognized (at various scales) the importance of implementing coordinated policies to reduce the anthropogenic influence on climate and preserve ecosystems, the environment, and human health. More than 190 parties have thus ratified the Paris Agreement and signed the Agenda 2030 to save the planet. The issue of pollutant reduction is very complex, but it requires a prompt response, even if the transition towards greener production is not immediate and needs time. Arguably, as most global economic activities are strongly carbon-based, carbon dioxide (CO₂) emissions are among the pivotal factors linking climate change and human activity (UNEP, 2022). A sharp rise in CO₂ emissions reinforces the greenhouse effect, strengthening the ongoing global average temperature increase. Greenhouse gas (GHG) emissions are expected to peak by 2025 (Intergovernmental Panel on Climate Change, IPCC, 2022); however, if policies beyond those implemented by the end of 2020 would not strengthen, the GHG concentrations will rise significantly above the current IPCC projections, causing a median global

warming of 3.2°C by 2100. Given the relevance of climate change policies and their connection with carbon emissions, empirical science is called upon to deliver theoretical and empirical models that monitor and forecast carbon dioxide releases, the main contributor to GHG emissions. Indeed, CO₂ emissions represent nearly 75% of total GHG (Ding et al., 2023) and marginal changes of CO₂ concentrations are crucial in determining policy interventions to mitigate the social cost of carbon dioxide, namely the dollar value of climate change damages following a marginal increase in CO₂ emissions (Rode et al., 2021).

We feed the ongoing debate by focusing on the forecasting of energy CO₂ emissions¹ for the US, the world's largest per-capita CO₂ emitter (UNEP, 2022) and address a problem that arises, for instance, in the annual update of the Global Carbon Budget, i.e. the need for statistical models able to capture the cumulative effects of economic activity on CO₂ emissions and that can reliably be used for forecasting purposes. Indeed, an accurate assessment of CO₂ concentrations and their redistribution among the atmosphere, ocean, and terrestrial biosphere in a changing climate is extremely important to better understand the global carbon cycle, support the development of climate policies, and project future climate change (Friedlingstein et al., 2022). We investigate to

* Corresponding author. University of Calabria, Department of Economics, Statistics and Finance, Ponte Bucci, Rende, CS, 87036, Italy.

E-mail addresses: b.algieri@unical.it (B. Algieri), leonardo.iania@uclouvain.be (L. Iania), arturo.leccadito@unical.it (A. Leccadito).

¹ Unless explicitly stated, for the rest of the study we use the terms 'energy CO₂ emissions' and 'CO₂ emissions' interchangeably.

what extent an ample group of potential drivers contribute to improving the predictive power of different forecasting models. For this purpose, we use three large datasets that include measures of anthropogenic activity (arising from various human activities) as potential determinants of CO₂ releases, namely (i) the [McCracken and Ng \(2020\)](#) database of macroeconomic variables; (ii) the University of Michigan's Surveys of Consumers and the Survey of Professional Forecasts series and (iii) energy-related indicators. We further augment the dataset with a rich group of nature-related data, the so-called local and global interannual variability (IAV) indicators. These indicators comprise information regarding, for instance, oceans/land surface temperature, local extreme weather events, and local droughts.

Previous studies have stressed the importance of macro variables or climate factors in explaining/forecasting the evolution of atmospheric CO₂ emissions (e.g. [Luo and Keenan, 2022](#); [Betts et al., 2016](#)). These types of variables have been considered separately. We propose to adopt a more comprehensive approach that evaluates the performance of a greater variety of data sources, including professional forecasts of economic activity and nature-related factors, in predicting energy CO₂ emissions. Our study fills the gap in the literature by using more than 260 anthropogenic and nature-related potential determinants and new methodological features thus providing important information to policymakers on what factors impact and could help predicting changes in CO₂ emissions.

In particular, for each group of variables, we first use dimensionality reduction techniques based on the Principal Component Analysis (PCA)² and the novel Quantile Factor (QF) Models introduced by [Chen et al. \(2021\)](#). These authors point out that, being a linear conditional mean model, PCA might be too restrictive in summarizing the information embedded in the joint distribution of the original variables. This might result in a loss of precious information in a forecasting or modelling setting. Instead, the extraction of the quantile factors, including those affecting the mean of the considered variables, could ameliorate the information traditionally provided by PCA. Further, Quantile Factor estimators present some robustness properties in the presence of outliers and heavy-tailed distributions, which would impair PCA. We build up thirteen models from our group of anthropogenic and local/global interannual variability predictors and then employ a series of point forecasting methods to assess their predictive power against the classical benchmarks. We consider each forecasting model separately and in combination, i.e. presenting model averaging prediction schemes. We perform both an in- and out-of-sample analysis using quarterly data for the period 1972Q1–2021Q3 and target different forecasting horizons ranging from one quarter to two years. Finally, we compare forecasting performance across models and methods and investigate if common factors, extracted with PCA and a novel technique from an extensive panel of macroeconomic variables, professional forecasts and nature-related factors, have predictive power over the rate of change in CO₂ emissions. This comprehensive approach is new in the empirical forecasting literature.

The remainder of the study is organised as follows. Section 2 reviews the literature on the topic. Section 3 presents our methodology and Section 4 describes the data. Section 5 shows and discusses the results of the in-sample and out-of-sample analyses. Section 6 concludes.

2. Review of the literature

The literature on predicting CO₂ emissions can be divided into different strands. The first strand has stressed the importance of anthropogenic variables in explaining, forecasting, and projecting the evolution of atmospheric CO₂ emissions. These anthropogenic factors

include, for instance, economic growth (e.g. [Filimonova et al., 2021](#); [Ozturk, 2010](#); [Raupach et al., 2007](#)), trade (e.g. [Olaizola and Diezgebacher, 2014](#); [Helm et al., 2012](#); [Munksgaard and Pedersen, 2001](#)), and tourism (e.g. [Wang and Wu, 2022](#); [İşik et al., 2017](#); [Peeters and Dubois, 2010](#)). Popular settings linking economic variables to CO₂ emissions range from the applied quantitative frameworks, such as the STIRPAT model and its extensions (e.g. [Algieri et al., 2022](#)), the system dynamics and the grey predictions (e.g. [Ren and Long, 2021](#); [Gao et al., 2021](#)), to the calibration-based settings, such as the integrated assessment models (e.g. [Wang et al., 2017](#)). Recently, [Fosten \(2019\)](#) and [Bennedsen et al. \(2021\)](#) have scrutinised the potential macroeconomic determinants of US CO₂ emissions via wavelet techniques combined with dynamic factor models and structural dynamic factor models, respectively. The former study found that economic activity and emissions have strong linkages from one to three years, while the latter highlights that CO₂ emissions are best explained by industrial production indices linked to manufacturing and residential utilities.

Another strand of the literature has highlighted how local and global climate variables (such as temperatures) affect CO₂ emission directly or indirectly via production, conversion, delivery, and energy use ([Ren and Long, 2021](#)). Building on the observation that terrestrial carbon sinks absorb up to 30% of anthropogenic CO₂ emissions, [Luo and Keenan \(2022\)](#) compare a battery of nine linear models to pinpoint the direct nexus between the global atmospheric CO₂ growth rate variability and extreme drought conditions in tropical areas. In relation to the latter channel, [Zhou et al. \(2013\)](#) assess the impact of climate change uncertainty on heating and cooling requirements by constructing an integrated assessment modelling framework. Coupling three different global climate model outputs with three scenarios of gridded population distribution, the study identifies to which extent the residential and building sector's energy consumption and CO₂ emissions of the US and China will respond to climate change.

A final strand of the literature has performed CO₂ emissions projections, using only the information included in the stochastic process underlying the baseline data, meaning that a univariate analysis has been carried out. In particular, [Belbute and Pereira \(2020\)](#) provide forecasts for CO₂ emissions in Portugal based on long memory Autoregressive Fractionally Integrated Moving Average (ARFIMA) approach and predict a reduction of 17.5% in carbon dioxide emissions by 2050 relative to the level recorded in 2010.

Our contribution to the current literature and policy debate is manifold. First, we highlight the importance of accounting for a broader range of anthropogenic and nature-related variables, both local and international, to improve the forecasting accuracy of CO₂ for different time horizons. Our results indicate that macro factors obtained from [McCracken and Ng \(2020\)](#), survey data that mirrors market sentiments and nature-related factors, notably drought indicators and Arctic anomalies, have strong in-sample and out-of-sample predicting power for future growth rates of CO₂ emissions. In some cases, superior predictions compared to the benchmarks can be obtained by combining forecasting models. These results might have policy implications as they indicate that, for example, policymakers should ponder among several sources of information when assessing and forecasting the evolution of CO₂ emissions, including weather anomalies like extreme drought episodes. To the best of our knowledge, this set of results is novel in the literature. [Fosten \(2019\)](#) and [Bennedsen et al. \(2021\)](#) look exclusively at macroeconomic determinants of US CO₂ emissions. [Luo and Keenan \(2022\)](#) link the global atmospheric CO₂ growth rate variability to extreme droughts conditions in tropical areas, while [Betts et al. \(2016\)](#) show that El Niño events have contributed to the rise in CO₂ emissions. Second, in contrast to previous methodological approaches, we use a direct forecasting setting with a new variable reduction technique, namely Quantile Factor Models, and we use a finer data frequency than the annual forecasting analysis conducted by [Bennedsen et al. \(2021\)](#). Finally, we point out that in monitoring and predicting CO₂ emissions, it might be advisable to strike a good balance between the type of IAV and

² The PCA is one of the oldest and most widely used methods to shrink the dimensionality of one or more datasets while preserving as much statistical information ('variability') as possible.

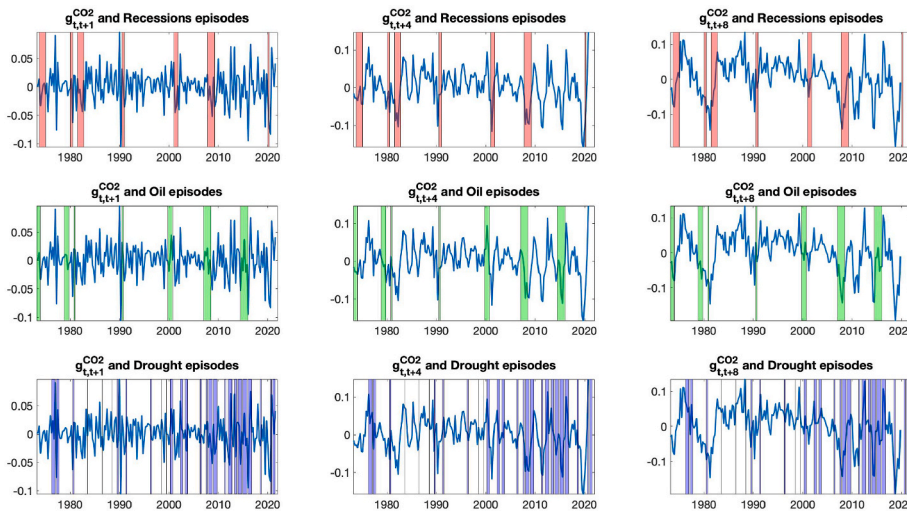


Fig. 1. Growth rates of total Energy CO_2 emissions **Fig. 1: Note:** CO_2 emissions are seasonally adjusted using the X-13ARIMA-SEATS filter. The Figure plots the growth of CO_2 emissions over one quarter (left column), one year (middle column) and two years (right column). The shaded areas refer to US recessions (top row), oil crises (middle row) and severe drought periods (bottom row). The US recession dates are from the National Bureau of Economics and Research. The oil episode periods are from Hamilton (2011), with the addition of the 2010s oil glut. The severe drought periods are those that, according to the National Centers for Environmental Information have caused more than 3 billion dollars in damages.

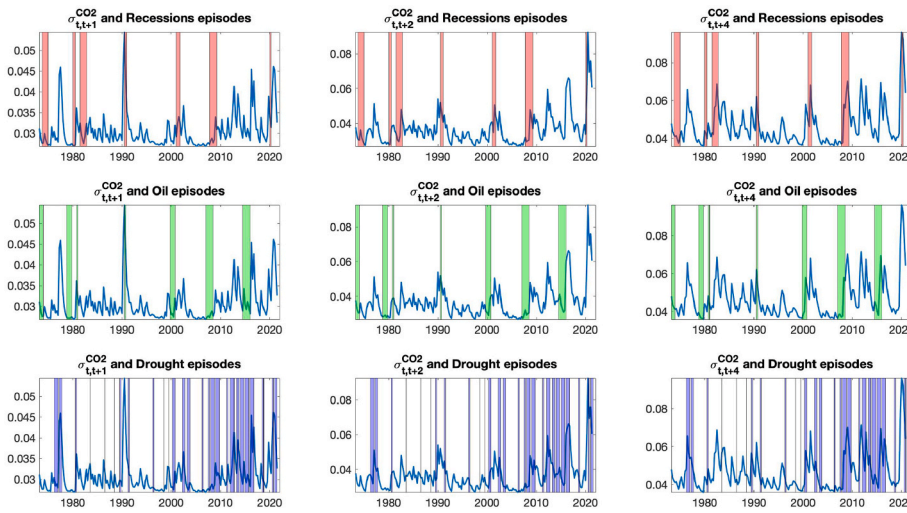


Fig. 2. Volatility of growth rates of total Energy CO_2 emissions

Fig. 2: Note: CO_2 emissions are seasonally adjusted using the X-13ARIMA-SEATS filter. The Figure plots the volatility of the growth of CO_2 emissions over one, (left column), (middle column) quarter, and one year (right column). Each volatility series has been extracted by a simple GARCH(1,1) model, see Bollerslev (1986). The shaded areas refer to US recessions (top row), oil crises (middle row) and severe drought periods (bottom row). The US recession dates are from the National Bureau of Economics and Research. The oil episode periods are from Hamilton (2011), with the addition of the 2010s oil glut. The severe drought periods are those that, according to the National Centers for Environmental Information, have caused more than 3 billion dollars in damages.

anthropogenic indicators and their combination. For example, we find that a combination of IAV and anthropogenic factors based on simple in-sample criteria such as the AIC/BIC criteria might deliver superior forecasting results.

3. Methodology

To set notation, we define the quarterly observation of the h -period growth rate of total CO_2 emissions as:

$$g_{t \rightarrow t+h}^{CO_2} = \log(E_{t+h}) - \log(E_t), \quad (1)$$

where $\{E_t\}_{t=1, \dots, T}$ is the level of carbon dioxide emissions.

With the objective to forecast CO_2 releases and assess the role of a group of potential drivers, we link $g_{t \rightarrow t+h}^{CO_2}$ to the last p quarter-on-quarter growth rates and a set of determinants at time t , X_t :

$$g_{t \rightarrow t+h}^{CO_2} = \beta^0 + \sum_{j=1}^p \beta_j g_{t-j \rightarrow t+1-j}^{CO_2} + \gamma' X_t + \varepsilon_t, \quad (2)$$

$$t = p+1, p+2, \dots, T-h.$$

where ε_t is the classical idiosyncratic disturbance at time t . The linear

model (2) is particularly suited for forecasting purposes, as it allows to directly forecast the growth rate of CO_2 emissions over h periods, i.e. from time T to time $T+h$, conditional on the information up to time T . The explanatory variables in (2) might be connected to anthropogenic determinants and nature-related factors.

In addition to the linear model (2), we follow Taylor (2007) and Lima et al. (2020) and implement robust forecast of CO_2 emissions growth rates based on linear combinations of quantiles. In particular, we consider the following quantile combinations.

● Trimean

$$\widehat{g_{T \rightarrow T+h}^{CO_2}} = 0.25Q_{0.25}(g_{T \rightarrow T+h}^{CO_2} | \mathcal{F}_T) + 0.5Q_{0.5}(g_{T \rightarrow T+h}^{CO_2} | \mathcal{F}_T) + 0.25Q_{0.75}(g_{T \rightarrow T+h}^{CO_2} | \mathcal{F}_T)$$

● Gastwirth:

$$\widehat{g_{T \rightarrow T+h}^{CO_2}} = 0.3Q_{1/3}(g_{T \rightarrow T+h}^{CO_2} | \mathcal{F}_T) + 0.4Q_{0.5}(g_{T \rightarrow T+h}^{CO_2} | \mathcal{F}_T) + 0.3Q_{2/3}(g_{T \rightarrow T+h}^{CO_2} | \mathcal{F}_T)$$

● 5-quantile:

$$\widehat{g_{T \rightarrow T+h}^{CO_2}} = 0.05Q_{0.1}(g_{T \rightarrow T+h}^{CO_2} | \mathcal{F}_T) + 0.25Q_{0.25}(g_{T \rightarrow T+h}^{CO_2} | \mathcal{F}_T) + 0.4Q_{0.5}(g_{T \rightarrow T+h}^{CO_2} | \mathcal{F}_T) + 0.25Q_{0.75}(g_{T \rightarrow T+h}^{CO_2} | \mathcal{F}_T) + 0.05Q_{0.9}(g_{T \rightarrow T+h}^{CO_2} | \mathcal{F}_T)$$

● 99-quantile:

$$\widehat{g_{T \rightarrow T+h}^{CO_2}} = \frac{1}{99} \sum_{i=1}^{99} Q_{0.01 \times i}(g_{T \rightarrow T+h}^{CO_2} | \mathcal{F}_T)$$

The conditional quantiles of $g_{t \rightarrow t+h}^{CO_2}$, appearing in the four linear combinations described above are obtained from the following quantile regression³ based on the same explanatory variables of the linear model (2):

$$Q_\tau(g_{t \rightarrow t+h}^{CO_2} | \mathcal{F}_t) = \beta_\tau^0 + \sum_{j=1}^p \beta_{j,\tau} g_{t-j \rightarrow t+1-j}^{CO_2} + \gamma_\tau X_t, \quad (3)$$

$$t = p+1, p+2, \dots, T-h$$

where $\tau \in (0, 1)$ is the probability level, $Q_\tau(g_{t \rightarrow t+h}^{CO_2} | \mathcal{F}_t)$ is the τ -quantile of $g_{t \rightarrow t+h}^{CO_2}$ and $\mathcal{F}_t$ denotes the information set available up to time t . Using a model like the above for conditional quantile of the growth rate over h periods is particularly useful since iterative forecasts (i.e. forecasts obtained from a one-period ahead model, iterated forward for h periods, as explained in Marcellino et al., 2006) are not available for quantiles. Indeed, quantiles cannot be easily temporally aggregated (Ghysels, 2014).

As a next step, we move from the five forecasting methods implemented separately (i.e. one linear and four quantile-based) to the combinations of forecasts by adopting the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) model averaging.⁴

Given $\mathcal{M}$ models and their predictions $\{\hat{y}\}_{m=1, \dots, \mathcal{M}}$ for the mean of CO_2 emissions growth rates, we obtain the predictions from model averaging based on information criteria as

$$\sum_{m=1}^{\mathcal{M}} w_m \hat{y}_m \quad (4)$$

where the weights are calculated as

$$w_m = \frac{\exp(-\frac{1}{2}IC_m)}{\sum_{j=1}^{\mathcal{M}} \exp(-\frac{1}{2}IC_j)}$$

and the Information Criterion (IC) can be either AIC or BIC.

For some of the considered models, we reduce the dimension of the explanatory variables adopting the Principal Component or the Quantile Factor Analysis (QFA) method. Both techniques allow us to extract the most relevant common factors from a sizeable set of variables. However, while PCA does not capture hidden factors that may shift characteristics (quantiles or moments) of the variable distributions other than their means, QFA seizes a larger amount of information, including those affecting the means and is robust to outliers and eventual heavy tails

(Chen et al., 2021). For the QFA method, we assume that for $\tau \in (0, 1)$ the observed explanatory variable $Z_{i,t}$, with $i = 1, \dots, N$ and $t = 1, \dots, T$ can be decomposed as

$$Z_{i,t} = \lambda'_i(\tau) f_t(\tau) + e_{i,t}(\tau)$$

where the error term $e_{i,t}(\tau)$ satisfies the quantile restriction

$$P(e_{i,t}(\tau) \leq 0 | f_t(\tau)) = \tau \quad \text{almost surely,}$$

and where for $r(\tau) \ll N$, $f_t(\tau)$ is an $r(\tau) \times 1$ vector of random common factors, and $\lambda_i(\tau)$ is an $r(\tau) \times 1$ vector of non-random factor loadings, that represent the contributions of the variable Z_i to the common factor (f_t) of dimension $r(\tau) \times 1$. Factors and loadings are estimated by minimizing

$$\frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \rho_\tau(Z_{i,t} - \lambda'_i(\tau) f_t(\tau)) \quad (5)$$

where $\rho_\tau(u) = u \cdot (\tau - I(u < 0))$. Eq. (5) is minimized under the restrictions

$$\frac{1}{T} \sum_{t=1}^T f_t(\tau) f'_t(\tau) = I_{r(\tau)}$$

$$\frac{1}{N} \sum_{i=1}^N \lambda_i(\tau) \lambda'_i(\tau) \quad \text{is diagonal with non-increasing diagonal elements.}$$

Here I_s denotes the identity matrix of dimension s . Contrary to PCA, however, it is not possible to obtain in closed form estimators of factors and loadings that minimize function (5). Hence, we follow the computationally efficient algorithm proposed by Chen et al. (2021), i.e. the Iterative Quantile Regression (IQR), that enables to efficiently identify the stationary points of the object function (5). The optimal number of factors, $r^*(\tau)$, is chosen by rank-minimization (see Section 3.2.1. of Chen et al., 2021). Once factors and loadings have been extracted from the set of observed explanatory variables $Z_{i,t}$ via IQR, we use the $r^*(\tau)$ factors $f_t(\tau)$ in equations (2) and (3) as explanatory variables, X_t . We follow a similar approach in the case of PCA.

4. Data

4.1. CO_2 emission rates

Our measure of CO_2 emission, E_t , is the seasonally-adjusted US Total Energy CO_2 Emissions (TETCEUS) retrieved from the US Energy Information Administration and spanning from Q1 1972 to Q3 2021. The seasonal adjustment is performed using the X-13ARIMA-SEATS method of the US Census Bureau.⁵

Fig. 1 sketches the CO_2 growth rate, $g_{t \rightarrow t+h}^{CO_2}$, over one quarter (left column), one year (middle column) and two years (right column). To gain some insight into the potential links between CO_2 emissions growth and natural and anthropogenic episodes, the figure also reports periods in which severe nature-related and economic events took place. In

³ Quantile regression is a valuable generalization of median regression (Koenker and Bassett, 1978) when the interest is not limited to the analysis of the average relationship between a specific variable and its predictors, but it involves the local behaviour of this linkage at different quantiles of the distribution.

⁴ See Moral-Benito (2015) for an overview of the topic.

⁵ See: <https://www.census.gov/data/software/x13as.html>.

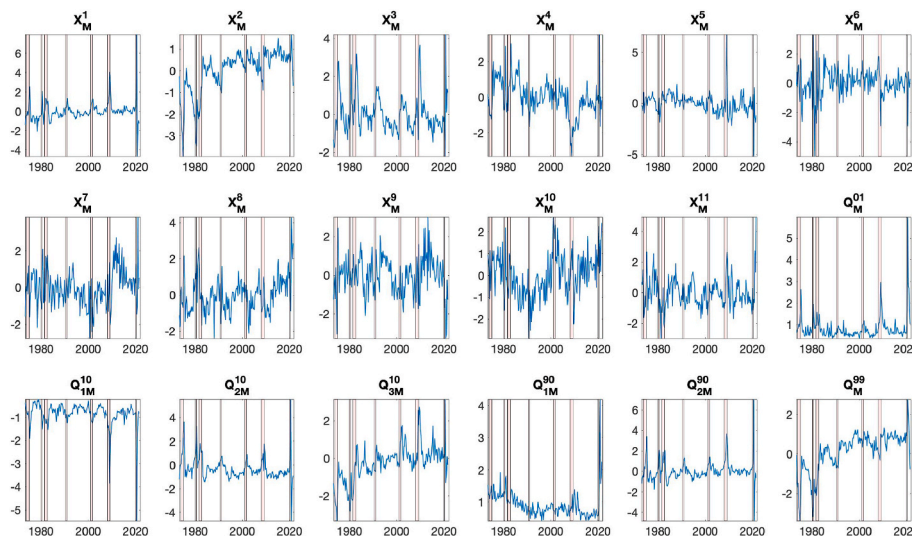


Fig. 3. Macro factors and quantiles.

Fig. 3: Note: Shaded areas refer to the US recession periods, as indicated by the National Bureau of Economics and Research. The Figure plots the first eleven principal components of the McCracken and Ng (2020) quarterly series, which explain 70% of the variation of the whole series. The rest of the series are the 1%, 10%, 90% and 99% quantile factors of the same dataset. The PCs and quantile factors are computed over the whole period, Q1 1972 to Q3 2021.

detail, the shaded areas in the top row of Fig. 1 highlight periods of severe drought, defined as episodes that have caused losses of more than \$3 billion, according to the Centers for Environmental Information.⁶ The middle row reports periods of oil crises, as detected by Hamilton (2011), plus the 2010s oil glut. Finally, the bottom row presents phases of economic recessions in the US as identified by the National Bureau of Economics and Research.⁷

The visual inspection of the top rows of Fig. 1 suggests that recession episodes are generally associated with drops in CO_2 emissions, both over short-term (one quarter) and medium-term (two years) horizons. For example, five out of the seven largest drops in yearly CO_2 emissions were associated with recession periods, particularly the early '80s, early '90s, early '00s, the period 2007–08 and the Covid-19 crisis. Turning to oil episodes, the subplots in the middle row of Fig. 1 show that oil crises (glut) were followed by a decrease (increase) in CO_2 emissions. The intuition behind this pattern might be that major oil crises were often followed by recessions, which caused reductions in anthropogenic activities and hence in CO_2 emissions.

Finally, the bottom-left panel of Fig. 1 seems to highlight that drought episodes were linked to an increase in one-quarter growth rate of CO_2 emissions. For example, CO_2 emissions increased during the short-term drought of the '70s, the Northern Plains drought of 1989, the drought in 1991, the drought and heat wave of 2000, and the drought and heat wave of 2012. The reason can be ascribed to the fact that when soils dry out due to droughts, plants breathe less and reduce their photosynthesis; hence, they can no longer seize carbon dioxide from the surrounding air, and more CO_2 pollutant stagnates (Humphrey et al., 2018; ETH Zurich, 2018; Betts et al., 2016). Thus, forests and trees have the ability to absorb and store carbon dioxide (Halsey, 2022; Ren and Long, 2021).

To gain further insights into the potential relationships between CO_2 emissions, droughts and economic episodes, we compute the conditional volatility⁸ of CO_2 growth rates and observe its dynamics over time. In particular, Fig. 2 reports the time-varying CO_2 growth volatility over one (left column), two (middle column) and four (right column) quarters. Focusing on the three plots on the bottom row, where there are severe drought periods, two points are worth noticing. First, volatility episodes

tend to spike in times of drought. Second, as drought frequency intensifies, so does the volatility of short-term growth CO_2 emissions. A similar nexus between CO_2 volatility and extreme tropical droughts has been recently pointed out by Luo and Keenan (2022). Overall, these simple graphical exercises suggest that the growth rate of CO_2 emissions is potentially linked to anthropogenic and natural determinants. The following sections lay out the dataset we use in our analysis.

4.2. Anthropogenic determinants: economic data, sentiment and energy factors

We trace the US-based historical evolution of economic activities that can influence the amount of CO_2 in the atmosphere by considering the large-scale dataset of McCracken and Ng (2020), available at the Federal Reserve of Saint Louis website. The dataset consists of 248 quarterly series classified into 14 groups ranging from financial and nominal variables, such as interest rates and prices, to output data, such as industrial production.⁹ To reduce the dimensionality of the dataset, we work with two types of factors: (i) $X_{M,t}^{70\%}$, i.e. the set of principal components that explain up to 70% of the variation of all the original series¹⁰, and (ii) $Q_{M,t}$, namely the set of 1%, 10%, 90% and 99% quantile factors of all the series.

Fig. 3 and Table 1 report the macro factors (in total eleven principal components plus seven quantile factors) extracted over the whole sample period and their correlation with the growth rate of CO_2 emission over 1, 2, 4, 6, and 8 quarters, respectively. Several quantile and macro factors are correlated with future CO_2 emissions evolution. Quarterly growth rates significantly correlate to the first principal component and quantile factors 1 and 2 (Table 1). In particular, the longer-term CO_2 emission growth, i.e. over 4 and 8 quarters, has the strongest link with macro factors 1 and 2 and quantile factor 1 at different τ levels. The shorter-term CO_2 emission growth, i.e. over 1 and 2 quarters, displays the strongest correlation with macro factor 1 and quantile factor 2 at 10% and 90% τ levels. Fig. 3 shows the dynamics of

⁹ We use the data and data-transformation proposed by McCracken and Ng (2020) at the Federal Reserve of Saint Louis website: <https://research.stlouisfed.org/econ/mccracken/fred-databases/> with two exceptions. First, from this database, we remove the series with ID: 32, 128, 129, 130, 194 as we included them in a larger energy-related database. Second, for nominal prices, we use the first differences of the log indexes instead of the second; using the proposed transformation by McCracken and Ng (2020) does not alter the results.

¹⁰ We follow the common practice of selecting principal components that explain 70% of the total variance (Jolliffe and Cadima, 2016).

⁶ See: <https://www.ncei.noaa.gov/access/billions/events/US/1980-2021>.

⁷ See: <https://www.nber.org/research/data/us-business-cycle-expansions-and-contractions>.

⁸ Each volatility series derive from a generalized autoregressive conditional heteroskedasticity model, GARCH(1,1), following Bollerslev (1986).

Table 1
Correlation table: macro factors.

	$h = 1$	$h = 2$	$h = 4$	$h = 6$	$h = 8$
X_M^1	-0.327***	0.428***	0.393***	-0.448***	0.408***
X_M^2	0.091	0.052	0.032	-0.043	-0.061
X_M^3	0.033	0.060	-0.261***	0.271***	0.339***
X_M^4	0.003	0.002	0.079	-0.054	0.090
X_M^5	0.051	0.038	0.033	0.177**	0.192***
X_M^6	-0.051	0.046	0.080	0.141 *	0.209***
X_M^7	-0.024	-0.034	-0.064	0.128 *	-0.049
X_M^8	0.006	-0.071	-0.178**	0.028	-0.084
X_M^9	-0.050	-0.024	0.196***	0.045	0.005
X_M^{10}	-0.042	0.091	0.026	-0.142*	-0.155**
X_M^{11}	0.123 *	0.042	0.045	-0.149**	0.093
Q_{1M}^{01}	-0.197***	0.358***	0.502***	-0.458***	0.460***
Q_{1M}^{10}	0.236***	-0.444***	-0.526***	0.493***	-0.471***
Q_{2M}^{10}	-0.312***	-0.359***	-0.309***	0.281***	0.237***
Q_{1M}^{90}	0.093	0.070	0.247***	-0.144**	0.100
Q_{2M}^{90}	-0.325***	0.395***	0.386***	-0.324***	-0.357***
Q_{3M}^{90}	0.033	-0.052	0.038	0.330***	0.256***
Q_{1M}^{99}	0.034	-0.254***	-0.395***	0.340***	-0.277***

Note: Table 1 reports the correlation between the macro factors/quantile and the growth rate of CO_2 emissions over 1, 2, 4, 6 and 8 quarters. h stands for the time horizon (quarters). ***, **, and * indicate 1, 5 and 10% significance levels, respectively.

the extracted factors.

In addition to the historical macroeconomic data, we consider (potential) future tendencies on economic activity and consumption by using a set of consumer spending and saving indicators retrieved from the University of Michigan's Surveys of Consumers,¹¹ from which we extracted $X_{MSc,t}^{70\%}$ via PCA. We further include the one-year ahead average forecast of the US real gross domestic product made by the Survey of Professional Forecasters (SPF), $X_{SPF,t}$. Data from surveys could be relevant because they provide 'soft' information, such as household perceptions and expectations about economic development, thereby reflecting market sentiments. These types of data have been exploited at the moment to a limited extent but could be very important to track human behaviour concerning decisions on saving, spending, prices, output, investment, finance, preferences and reactions to economic policies. Table 2 lists the data used and their sources.

Our last set of anthropogenic variables derives from the energy market. Table 3 shows the 15 energy-related variables we employ, which refer to (i) energy consumption, (ii) energy production, export, and import, (iii) energy prices, and (iv) the temperature-related measures of energy demand across the contiguous 48 States plus the District of Columbia. The latter indicator refers to the Residential Energy Demand Temperature Index (REDTI), a population-weighted heating and cooling degree days index tracing residential energy demand for air conditioning and heating. The index is correlated with daily temperatures: it increases when the number of heating and cooling degree days rises and decreases when the number of heating and cooling degree days falls (Heim et al., 2003). We summarize the information contained in these 15 variables using $X_{E,t}^{70\%}$, and $Q_{E,b}$ both obtained with the same procedure adopted for the macro variables.

Fig. 4 and Table 4 report $X_{MSC,t}^{70\%}$, $X_{SPF,b}$, $X_{Et}^{70\%}$ and $Q_{E,b}$ and their

¹¹ The Surveys of Consumers are nationally representative surveys based on repeated phone interviews with adults living in households in the conterminous (or contiguous) United States (i.e. 48 States plus the District of Columbia). The surveys capture changes in attitudes and behaviours at the aggregate and individual levels.

Table 2
Surveys of consumers and professional forecasters.

University of Michigan's Survey of Consumers
Current Financial Situation Compared with a Year Ago
Expected Change in Financial Situation in a Year
Annual Trend in Past and Expected Household Financial Situation
News Heard of Recent Changes in Business Conditions
Current Business Conditions Compared with a Year Ago
Expected Change in Business Conditions in a Year
Business Conditions Expected During the Next Year
Business Conditions Expected During the Next 5 Years
Expected Change in Unemployment During the Next Year
Expected Change in Interest Rates During the Next Year
Buying Conditions for Large Household Durables Vehicles
Buying Conditions for Vehicles

Note: The Surveys of Consumers are retrieved from the University of Michigan website, where a full description of each series is available: <https://data.sca.isr.umich.edu/tables.php><<https://data.sca.isr.umich.edu/tables.php>>. The surveys of professional forecasters are from the Philadelphia Federal Reserve website, where a full description of each series is available: <https://www.philadelphiafed.org/surveys-and-data/data-files>.<<https://www.philadelphiafed.org/surveys-and-data/data-files>>.

Table 3
Energy related variables.

Number	Variable name	Tranformation
1	Total Petroleum Consumed by the Residential Sector	1
2	Total Petroleum Consumed by the Commercial Sector	1
3	Total Petroleum Consumed by the Industrial Sector	1
4	Total Petroleum Consumed by the Transportation Sector	1
5	Fraction of gasoline on total for Transportation Sector	1
6	Total Petroleum Field Production	1
7	Total Renewable Energy Consumption	1
8	Petroleum Imports	1
9	Petroleum Exports	1
10	Petroleum Stock Change	0
11	IP: Consumer energy products (Index, 2012 = 100)	2
12	PPI by CFRPP: Natural Gas (1982 = 100)	2
13	PPI by CFRPP (Domestic Production) (1982 = 100)	2
14	Real Crude Oil Prices: WTI-CO (2012 \$/b), Core PCE deflated	2
15	REDTI: Residential Energy Demand Temperature Index	0

Note: **PPI by CFRPP** = Producer Price Index by Commodity for Fuels and Related Products and Power. **WTI-CO** = West Texas Intermediate (WTI) - Cushing, Oklahoma. **Transformation:** 0 = no transformation, 1 = year on year growth rate, 2 = quarter on quarter growth rate. **Sources:** series 1 to 10 are from <https://www.eia.gov/totalenergy/data/monthly/>, series 11 to 14 are from <https://research.stlouisfed.org/econ/mccracken/sel/>, and series 15 is available at <https://www.ncdc.noaa.gov/societal-impacts/redti/overview>.

correlation with the growth rate of CO_2 emissions over 1, 2, 4, 6, and 8 quarters, respectively.

Table 4 indicates that the first energy factor and the 10% quantile factor (i.e. $X_{E,t}^1$ and $Q_{E,t}^{10}$) have strong correlations with the growth of CO_2 emissions. In addition, there are also significant correlations between the growth of CO_2 releases and the second Consumer Survey factor, i.e. $X_{MSC,t}^2$.

4.3. Nature-related determinants: Interannual variability drivers

Our potential determinants of CO₂ emissions related to nature hinges

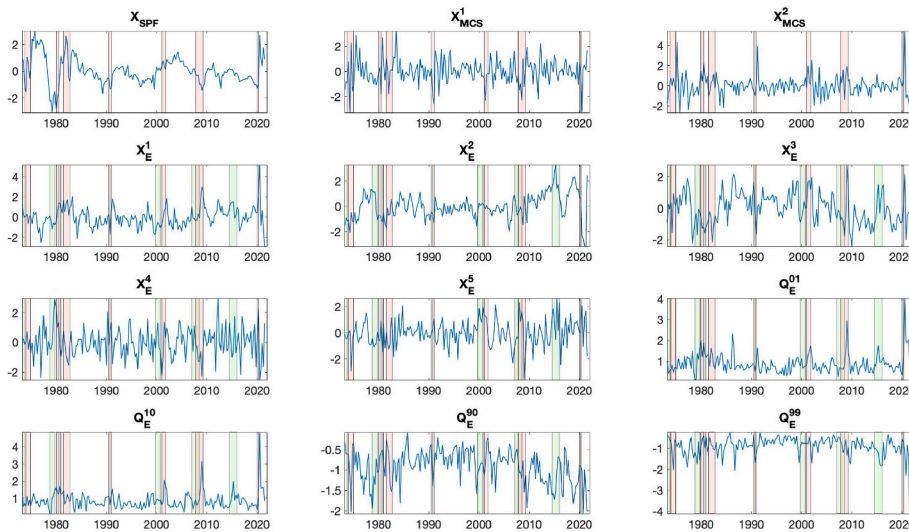


Fig. 4. Surveys and energy factors

Fig. 4: Note: Light-red shaded areas refer to the US recession periods, as indicated by the National Bureau of Economics and Research, while light-green shaded areas indicate oil episodes, which are from Hamilton (2011), with the addition of the 2010s oil glut. The Figure plots the SPF surveys, top-left panel, the first two principal components of the Michigan surveys, top-middle and top-right panels, the first five principal components of the energy dataset, all other panels except the bottom-right, and the first principal component of the 1%, 10%, 90% and 99% quantile factors of the energy dataset, bottom-right panel. All PCs are selected to explain 70% of the variation of the original series. The PCs and quantile factors are computed over the whole period, Q1 1972 to Q3 2021. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 4

Correlation table: surveys and energy factors.

Correlation with the growth of CO ₂ emissions					
	$h = 1$	$h = 2$	$h = 4$	$h = 6$	$h = 8$
SPF	0.034	0.002	0.029	0.045	−0.098
X^1_{MSC}	0.072	0.054	−0.044	−0.193***	0.173**
X^2_{MSC}	−0.288***	0.327***	−0.333***	−0.202***	−0.130*
X^1_E	−0.334***	0.532***	0.606***	−0.584***	0.513***
X^2_E	0.063	0.135*	0.126*	0.038	−0.007
X^3_E	−0.081	0.121*	0.274***	0.267***	0.225***
X^4_E	−0.240***	0.145**	0.070	−0.044	0.037
X^5_E	−0.166**	−0.099	−0.068	0.067	0.023
Q^{01}_E	−0.195***	0.156**	0.132*	0.109	0.242**
Q^{10}_E	−0.214***	0.448***	0.478***	−0.374***	0.336***
Q^{90}_E	−0.113	0.083	−0.023	0.043	−0.083
Q^{99}_E	−0.058	−0.079	−0.166**	0.110	−0.013

Note: Table 4 reports the correlation between the surveys/energy factors/quantile and the growth rate of CO₂ emissions over 1, 2, 4, 6 and 8 quarters. h stands for the time horizon (quarters). ***, **, and * indicate 1, 5 and 10% significance levels, respectively.

upon measures of temperature and precipitation variations, drought indicators, and climate indices, called Interannual Variability (IAV) determinants and presented in Table 5. The motivation for choosing this set of variables stems from previous studies. For instance, Matthews et al. (2021); Tokarska et al. (2020); Ray et al. (2020); Kushnir et al. (2019) have documented that atmospheric patterns and anomalies related to ocean and land surface conditions are crucial triggers of the natural fluctuations in the climate system, which subsequently, impact CO₂ emissions. Furthermore, Luo and Keenan (2022) and Ray et al. (2020) have highlighted that CO₂ emission cycles are linked to the vegetated land surface and soil conditions, which in turn are influenced by natural events, such as drought, temperature heat waves, and rain-falls. Along these lines, Algieri et al. (2022) have shown that extreme temperatures warm up and dry lands, which result in reduced carbon uptake by vegetation and higher carbon dioxide releases. The authors show that, conversely, precipitations reduce CO₂. Finally, Meangbua et al. (2019) have pointed out that temperature is the crucial factor influencing energy and CO₂ requirements, with a 1°C increase in temperature resulting in a 200% rise in consumption-based carbon dioxide emissions.

Starting from a rich nature-related dataset, we narrow the large number of IAV variables to a small range of factors common to these

Table 5

Nature-related determinants.

Label	Temperature
NHLTA	Northern hemisphere land temperature anomalies
NHOTA	Northern hemisphere ocean temperature anomalies
SHLTA	Southern hemisphere land temperature anomalies
SHOTA	Southern hemisphere ocean temperature anomalies
M-ARTIC	Average Arctic temperatures anomalies
RA-ARTIC	Cross sectional robust asymmetry measure of Arctic temperatures anomalies
M-US	Average US temperatures anomalies
RA-US	Cross sectional robust asymmetry measure of US temperatures anomalies
M-US-Extreme	Average US maximum and minimum temperature extremes
M-US-Anomalies	Average US maximum and minimum temperature anomalies
D-US-Anomalies	Difference of maximum and minimum US temperature anomalies
Drought and precipitations	
PDI	First principal component of the PDIs for US Lower 48 Divisions
EP	US extremes in 1-day precipitation
Oscillation indices	
AO	Arctic Oscillation index
PNA	Pacific-North America pattern
PDO	Pacific Decadal Oscillation
NAO	North Atlantic Oscillation index
SOI	Southern Oscillation Index
ONI	Oceanic Niño Index

Note: All the data on natural determinants are from the National Centers for Environmental Information: <https://www.ncei.noaa.gov/access/monitoring/products/>. The average and the robust asymmetry measures for the US and the Arctic regions were obtained by collecting data each 5 latitudes/longitudes.

potential nature-related predictors distinguished in (a) measures of temperature and precipitation variations, (b) drought indicators, and (c) climate indices.

In detail, as regards temperature/precipitation measures, we use three types of data: (i) The northern/southern hemisphere anomalies indexes for land and ocean, from which we extracted $X^{70\%}_{WAL,t}$. (ii) The year-to-date temperature anomalies in the Arctic region, i.e. the area within the 25th and 50th latitude. We proceeded in two steps. First, we gridded the area by varying the latitude and the longitude by one degree

and collected the year-to-date temperature anomalies of each gridded location. Afterwards, via standard principal component analysis and the new QFA, we extracted $X_{A,t}^{70\%}$ and $Q_{A,t}$. (iii) Measures of extreme weather conditions, namely the average of the maximum and minimum contiguous US temperature extremes, the average and the difference of the contiguous US maximum and minimum temperature anomalies, and the contiguous US extremes in 1-day precipitation. For these three sets of series, we use $X_{Extr,t}^{70\%}$.

Concerning the drought conditions' measures, we consider the Palmer Drought Index¹² (PDI) for each of the contiguous United States (the so-called 'Lower 48'). PDI indicators thus reflect local dryness conditions. We narrow down the 48 PDI indicators considering $X_{PDI,t}^{70\%}$ and $Q_{PDI,t}$.

Finally, as climate indices, we select measures covering vast global areas, namely: the Arctic Oscillation (AO) indicator, the Pacific-North America (PNA) pattern, the Pacific Decadal Oscillation (PDO) indicator, the North Atlantic Oscillation (NAO) index, the Southern Oscillation Index (SOI), and the Oceanic Niño Index¹³ (ONI). As with the other set of data, we use $X_{OSCI,t}^{70\%}$ for the six climate indices. All nature-related variables are reported in Table 5.

Fig. 5 and Table 6 depict the set of IAV determinants and their correlation with the growth rate of CO₂ emission over 1, 2, 4, 6, and 8 quarters, respectively. Table 6 points out that drought indicators, especially $X_{PDI,t}^4$ and Arctic temperature anomalies, $X_{A,t}^1$ and $Q_{A,t}^{99}$, exhibit the largest significant correlations with the growth rate of CO₂ emissions.

5. Results

5.1. In-sample analysis

5.1.1. Setting

For the in-sample analysis, we build thirteen competing models from our set of potentially relevant anthropogenic and nature-related predictors (Table 7) and compare them to the 'baseline' setting based on a direct forecasting framework with four lags of quarterly growth rates, $p = 4$:

$$g_{t \rightarrow t+h}^{CO_2} = \beta^0 + \sum_{j=1}^p \beta_j g_{t-j \rightarrow t+1-j}^{CO_2} + \varepsilon_t. \quad (6)$$

Equation (6) describes an autoregressive model with four lags (AR-4) and can be obtained as a particular case of eq. (2) by setting $\gamma = 0$. Our baseline has four lags for two reasons. First, controlling for past-year growth rates allows us to parsimoniously consider the potential medium-term correlation structure in multi-period CO₂ growth rates. Second, from a more formal standpoint, we compare the adjusted R-squared of four forecasting regressions based on equation (6) with $p = 1, 2, 3$, and 4. We estimate the four models for different time horizons, namely $h = 1, 2, 4, 6$, and 8 quarters, and over rolling samples of 25 years fixed-window.

Fig. 6 reports the evolution of the adjusted R-squared for each model

¹² The PDI index proposed by the meteorologist Wayne Palmer (1985) is computed using temperature data and a physical water balance model. As a result, the index is able to capture the primary effect of global warming on drought through changes in potential evapotranspiration.

¹³ The considered climate variables track the patterns of air pressure, winds and sea surface temperatures. For instance, the Southern Oscillation Index measures the large-scale fluctuations in air pressure between the western and eastern tropical Pacific. The Oceanic El Niño Index measures the anomalies in the ocean temperatures across the equatorial Pacific Ocean. The Arctic Oscillation Index monitors the circulation of the winds around the Arctic. Detailed information can be found at the National Centers for Environmental Information <https://www.ncei.noaa.gov/access/monitoring/products>.

and the five forecasting growth periods. With just a few exceptions, the adjusted R-squared of the baseline model with four lags (blue line) is consistently equal to or higher than the adjusted R-squared of the other AR models with one, two and three lags.

In this in-sample exercise, the baseline AR-4 model (6) is compared to each of our competing models, which augment the baseline with one set of the control factors listed in Table 7. Note that all the explanatory variables enter the thirteen models with a time delay so that current or existing data might help the predictions of future values.

We perform the in-sample evaluations over the whole sample (static setting) and via rolling estimations¹⁴ based on fixed windows of 25 years (dynamic setting). The choice of the window size averages out the window sizes of the macro-based study by Chen et al. (2021) and the IAV-based analysis performed by Luo and Keenan (2022), i.e. 30 and 20 years, respectively. We consider the two approaches since, with the rolling window setting, the group of indicators used to predict the CO₂ growth rate and their associated coefficients could vary over time. Conversely, with the static setting, we account for the possibility that a stable set of predictors may produce less volatile forecasts.

Fig. 7 reports the ratio between the adjusted coefficient of determination of the baseline AR-4 model over that of each competing augmented model for the whole sample estimation. Fig. 8 displays the mean ratio between the adjusted R-squared of the baseline model and the adjusted R-squared of each competing model for the rolling windows estimation. Therefore, models for which the ratio is less than one (red scale) have better in-sample goodness of fit than the baseline model. Vice-versa, models with ratio larger than one (green scale) underperform the baseline. For the rolling window analysis, we estimate equation (2) for the models described in Table 7 and test null $H_0: \gamma = 0$ on every window. For each model, we then employ the resulting p-values to derive the combined p-value using the method of Benjamini and Hochberg (1995) that controls for the family-wise error rate.

5.1.2. Discussion

Comparing Figs. 7 and 8, a few points stand out. First, models augmented with macro factors (M1 and M4), energy (M5), drought (M8), and Arctic factors (M11 and M12) have the strongest in-sample predicting power for future growth rates of CO₂ emissions in all forecasting horizons (mainly for time horizons above one quarter). For these models, the gains in the adjusted R-squared relative to the AR-4 baseline are mostly above 30%, and the null hypothesis on the coefficient γ is rejected with a p-value lower than 1%. These results hold both in the static and rolling setting, indicating that the in-sample predicting power is potentially robust to time variation and structural breaks. To a lesser extent, i.e. with lower gains in the adjusted R-squared relative to the benchmark, these findings hold also for models augmented with the University of Michigan's Surveys of Consumers (M3) in both static and rolling settings and for models enhanced with the Surveys of Professional Forecasters (M2) and World Anomalies Indexes (M10) in the static setting. Second, all models augmented with quantile factors have higher predictive power than the baseline within a dynamic setting. Instead, models enhanced with quantile factors linked to energy and drought (M6 and M13) show lower predictive performance in a static framework. Finally, models augmented with extreme events (M7) improve short-run predictions compared to the baseline, while the climate-oscillation variables (M9) outperform the benchmark only for horizons longer than six months.

To the best of our knowledge, this set of results is new in the literature. Differently from Fosten (2019) and Bennedsen et al. (2021) that point out the exclusive role of macroeconomic determinants in predicting US CO₂ emissions, we also stress the importance of considering a broader range of factors when forecasting carbon dioxide emissions. Those factors include nature-related drivers, such as local drought

¹⁴ See Luo and Keenan (2022) for a similar approach.

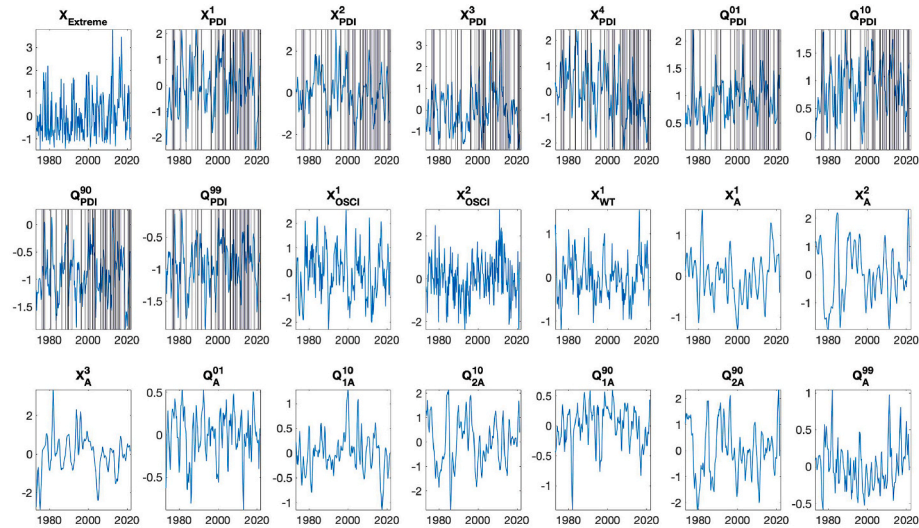


Fig. 5. IAV factors

Fig. 5: Note: Shaded areas refer to US drought periods.

Table 6

Correlation table: nature-related factors, IAV.

	Correlation with the growth of CO ₂ emissions				
	<i>h</i> = 1	<i>h</i> = 2	<i>h</i> = 4	<i>h</i> = 6	<i>h</i> = 8
$X_{Extreme}$	0.078	-0.025	-0.012	0.120	-0.120
X_{PDI}^1	0.071	0.122*	0.145**	0.033	0.016
X_{PDI}^2	-0.065	-0.108	-0.089	0.114	-0.098
X_{PDI}^3	-0.073	-0.064	0.014	-0.037	-0.023
X_{PDI}^4	0.129*	-0.189***	0.276***	-0.258***	-0.140*
Q_{PDI}^{01}	0.020	0.073	0.037	0.135*	0.134*
Q_{PDI}^{10}	0.022	0.080	-0.060	0.004	-0.026
Q_{PDI}^{90}	0.108	-0.139*	0.167**	-0.079	-0.086
Q_{PDI}^{99}	0.075	0.029	0.041	0.012	-0.022
X_{OSCI}^1	0.020	-0.004	0.105	0.200***	0.183**
X_{OSCI}^2	0.003	-0.118	-0.195***	0.088	-0.063
X_{WT}^1	-0.052	-0.110	-0.185**	-0.177**	-0.188**
X_A^1	-0.069	-0.120*	-0.213***	-0.193***	-0.187**
X_A^2	0.060	0.075	0.028	0.029	0.031
X_A^3	0.015	-0.035	0.068	-0.074	-0.061
Q_A^{01}	-0.053	-0.106	-0.172**	-0.105	-0.035
Q_A^{10}	0.054	0.116	0.209***	0.156**	0.131*
Q_A^{90}	0.033	0.034	-0.010	-0.023	0.091
Q_A^{99}	0.062	0.104	0.203***	0.199***	0.194***
Q_{2A}^{90}	-0.004	-0.013	-0.072	-0.034	0.045
Q_A^{99}	-0.073	-0.143**	-0.208***	-0.180**	-0.160**

Note: Table 6 reports the correlation between the IAV factors/quantile and the growth rate of CO₂ emissions over 1, 2, 4, 6 and 8 quarters. *h* stands for the time horizon (quarters).***, **, and * indicate 1, 5 and 10% significance levels, respectively.

episodes and global Arctic anomalies, and consumer and professional forecaster surveys that reflect market sentiment. In particular, consumer surveys are helpful for predictions in static and dynamic settings, while surveys of professional forecasters are valuable in a static environment. Drought and Antarctic indices are important factors for all forecasting horizons. Oscillation indicators and northern/southern hemisphere anomalies are useful for predictions in the long term but not in the short term. Finally, different types of quantile factors might contain relevant information for predicting CO₂ emissions, especially within a rolling-dynamic setting. Thus, selecting a more comprehensive group of variables is crucial to improve the accuracy of CO₂ forecasts since sharper predictions can be used to evaluate the policy effort necessary to achieve emissions goals to net zero by 2050 and give the world an even

Table 7

In-sample analysis: List of models.

Benchmark name	Direct forecasting equation	
Baseline	$\beta^0 + \sum_{j=1}^p \beta_j g_{t-j+1-j}^{CO_2}$	
Model name	Direct forecasting equation	Type of control variables
M1	Baseline + $\gamma X_{M,t}^{70\%}$	All Macro, except Oil
M2	Baseline + $\gamma X_{SPF,t}$	Survey of professional forecasters
M3	Baseline + $\gamma X_{MSC,t}^{70\%}$	GDP growth
M4	Baseline + $\gamma Q_{M,t}$	Michigan surveys
M5	Baseline + $\gamma X_{EI,t}^{70\%}$	Quantiles factors of all Macro, except Oil
M6	Baseline + $\gamma Q_{E,t}$	Energy series
M7	Baseline + $\gamma X_{Ext,t}^{70\%}$	Quantiles factors of Energy series
M8	Baseline + $\gamma X_{PDI,t}^{70\%}$	US extreme events
M9	Baseline + $\gamma X_{OSCI,t}^{70\%}$	Palmer Drought Index
M10	Baseline + $\gamma X_{WAL,t}^{70\%}$	OSCI variables
M11	Baseline + $\gamma X_{A,t}^{70\%}$	World Anomalies Indexes
M12	Baseline + $\gamma Q_{A,t}$	Arctic temperatures
M13	Baseline + $\gamma Q_{PDI,t}$	Quantiles factors of Arctic temperatures
		Quantiles factors of Palmer Drought Index

Note: $X_V^{70\%}$ denotes the set of principal components that explain up to 70% of the variation of the series *V*. Q_V is the set of 1%, 10%, 90% and 99% quantile factors of the series *V*.

chance of limiting the global temperature rise to 1.5°C.

5.2. Out-of-sample analysis

5.2.1. Setting

In the out-of-sample analysis, we split the sample into two equal parts: the training sample, spanning from 1972Q1 to 1997Q4, and the testing sample ranging from 1998Q1 to 2021Q3. For each forecasting horizon, *h* = 1, 2, 4, 6, and 8 quarters, we consider a rolling window of 25 years and make predictions with (i) the previous baseline AR-4 model, eq. (6); (ii) a naïve benchmark model, i.e. the random walk which predicts that $g_{T \rightarrow T+h}^{CO_2} = g_{T-h \rightarrow T}^{CO_2}$; (iii) the models with macroeconomic, survey data and nature-related variables (see Table 7) and (iv) a

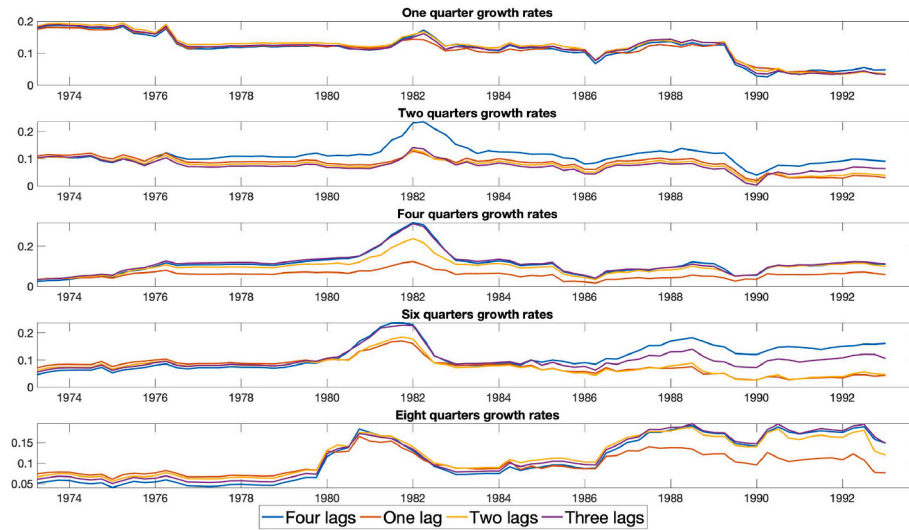


Fig. 6. Rolling estimates of Adjusted R-squared

Fig. 6: Note: Comparison of adjusted R-squared. The dates on the horizontal axis represent the beginning of the sample period of the 25 years window used to obtain the adjusted R-squared.

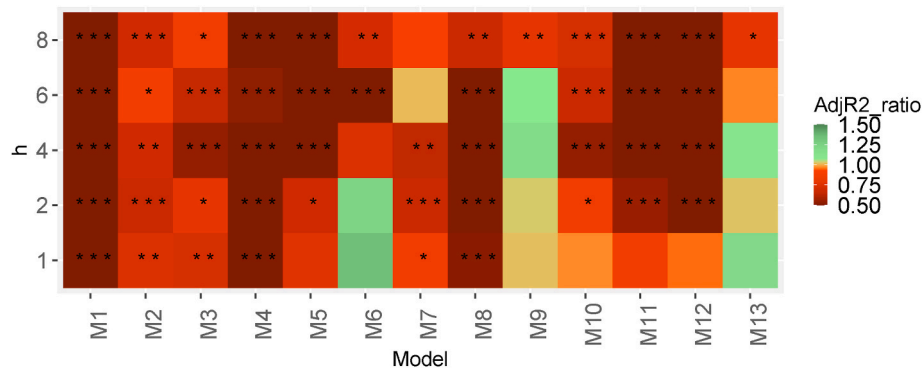


Fig. 7. In-sample Analysis: Heatmap for the ratio $\text{AdjustedR}^2(\text{Baseline})/\text{AdjustedR}^2(i)$, whole sample (static setting).

Fig. 7: Note: $\text{AdjustedR}^2(\text{Baseline})$ is the Adjusted R^2 of model (6). $\text{AdjustedR}^2(i)$ is the Adjusted R^2 of model i . The stars in each cell are related to the p-values for the test with null $H_0 : \gamma = 0$ in eq. (2). ***, **, and *, denote significance at the 1%, 5% and 10% levels, respectively.

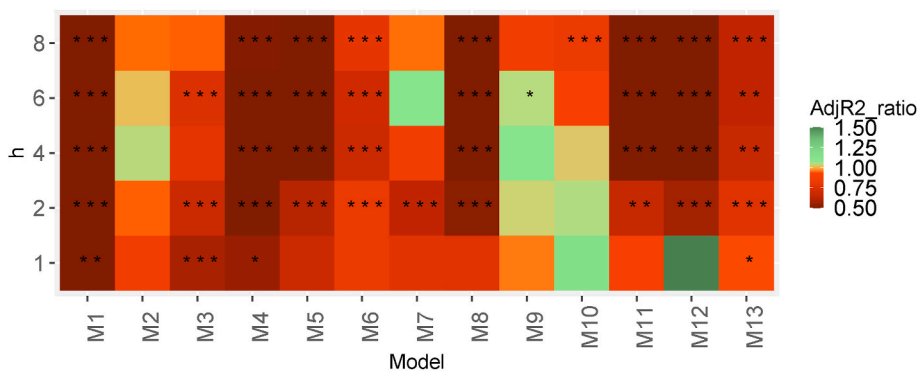


Fig. 8. In-sample Analysis: Heatmap for the mean of $\text{AdjustedR}^2(\text{Baseline})/\text{AdjustedR}^2(i)$, rolling windows (dynamic setting).

Fig. 8: Note: $\text{AdjustedR}^2(\text{Baseline})$ is the Adjusted R^2 of model (6). $\text{AdjustedR}^2(i)$ is the Adjusted R^2 of model i . The stars in each cell are related to the p-value for the combined test (Benjamini and Hochberg, 1995) of the null hypothesis that no p-value is significant. The combined p-values are for the test with null $H_0 : \gamma = 0$ in eq. (2) in all the possible windows. ***, **, and *, denote significance of γ at the 1%, 5% and 10% levels, respectively.

linear combination of forecasts resulting from those latter models, which group all (or a part of) the predictions of the original thirteen models using weights obtained via the AIC or the BIC criteria as outlined in Section 3. The models with combined forecasts are labelled MAICX and MBICX, respectively, where $X = 1, \dots, 7$ refer to seven combinations of the original models (Table 8). The rationale for presenting combinations (or averages) of the models is that they could display better forecasting performance than individual models by smoothing out the impact of

outlier observations.

Model comparison is carried out using the average loss function based on the Root Mean Square Error (RMSE) for mean predictions. Each loss function is divided by the loss function of the benchmark model to obtain the Score Ratio (SR)

Table 8

Out-of-sample Analysis: List of models.

Benchmark name	Direct forecasting equation
Baseline	$\rho^0 + \sum_{j=1}^p \beta_j g_{t-j \rightarrow t+1-j}^{CO_2}$
Naïve	$g_{T-h \rightarrow T}^{CO_2}$
Model name	Description
M1-M13	Models in Table 7
MXYZ1	Combination of all factors
MXYZ2	Combination of all factors less quantiles
MXYZ3	Combination of macro factors (M1 to M6)
MXYZ4	Combination of macro factors less quantiles
MXYZ5	Combination of IAV factors (M7 to M13)
MXYZ6	Combination of IAV factors less quantiles
MXYZ7	Combination of all quantile factors

Note: The string XYZ can be: i) ‘AIC’, for the case when the combination is via AIC; or ii) ‘BIC’, for the case when the combination is via BIC.

$$SR_j^h = \frac{\sqrt{\sum_{T=1}^{T_h} (\hat{y}_{T+h,j} - y_{T+h})^2}}{\sqrt{\sum_{T=1}^{T_h} (\hat{y}_{T+h,benchmark} - y_{T+h})^2}} \quad (7)$$

Here $\hat{y}_{T+h,j}$ is the CO_2 growth rate from T to $T+h$ predicted by model j , y_{T+h} the realized value, and T_h the number of out-of-sample predictions. The measure (7) can be used to test if a model outperforms (underperforms) its benchmark, i.e. SR_i 1 (SR_i 1), via the statistics proposed by Clark and West (2007). The Clark and West test is run under the null hypothesis of equal predictive accuracy between two competing forecasts. The online appendix reports the detailed results. Figs. 9 and 10 instead present the heatmaps based on the SR against the two benchmark models. In each figure/heatmap cell, we also report the number of stars based on the significance of the Clark and West (2007) testing procedure. In particular, Fig. 9 indicates that all the considered models outperform the naïve benchmark.

We have finally carried out a more robust model comparison by implementing the multiple testing proposed by White (2000). White’s test has the objective of mitigating spurious out-of-sample forecasting performance substantially. Indeed, White (2000) argued that when a ‘good’ forecasting model derives from an extensive specification search, there is always the possibility that good performances could come not from its actual forecasting ability but just by chance. Therefore, he proposed the Reality Check test that permits avoiding data snooping (or mining) with some degree of confidence so that one does not mistake results generated by chance for genuinely good results. The null hypothesis is formulated as a multiple hypothesis, i.e. the intersection of l one-sided hypotheses, where l is the number of the considered models and/or methods of producing forecasts. The null hypothesis of interest is that the best model has no predictive superiority over the benchmark. Table 9 reports the bootstrap Reality Check p-values for comparing l models and/or methods with the AR-4 and naïve benchmarks. In each of the two panels of Table 9, we identify the best model based on RMSE, using predictions from models M1 to M13 (see Table 7) and from models MXYZ1 to MXYZ7 (see Table 8). We consider, for each model, predictions obtained using the linear and the four quantile-based methods, meaning that the multiple hypothesis we test involves $l = 135 (= (13 + 7 \times 2) \times 5)$ different types of predictions. Indeed, in the case of combined forecasts we jointly consider the five different methods together with in-sample combinations (XYZ = AIC or XYZ = BIC).

5.2.2. Discussion

The findings from the out-of-sample heatmaps (Figs. 9 and 10) corroborate the in-sample results with a few exceptions. In particular, compared to the naïve model (Fig. 9), the forecasting performance of the

augmented models always improves out-of-sample predictions. The reason can stem from the fact that the current values of CO_2 growth series have little information about their future evolution, as explained by Elliott and Timmermann (2016) and as visible from Fig. 1. Compared to the AR-4 baseline (Fig. 10), the predicting power of the augmented models with macro (M1), energy (M5), drought (M8) and Arctic (M11) factors found out in the previous in-sample analysis also hold for the out-of-sample forecasts. The quantile factors are instead less powerful in the out-of-sample analysis. All the results are robust on how the point forecast is obtained; see Section 3.

It is worth noting that the forecasting power of macroeconomic factors is in line with the results of Bennedsen et al. (2021); however, in contrast with their findings, we show that the predictive ability of macroeconomic covariates is valid for longer forecasting horizons. Moreover, while Bennedsen et al. (2021) compare their results to a constant growth model, whereby the point forecast is the historical mean of the CO_2 growth rate up until the time of the forecast, our approach is a more conservative one as we use a more performing AR-4 baseline model.

The out-of-sample performance of energy factors and local and global IAV (PDI index and Arctic) is, to the best of our knowledge, a novel result. It indicates that augmenting the forecasting model with these types of indicators might result in a substantial gain in forecasting power. For short-term forecasting horizons (below one year), our results suggest employing local IAV factors, while for longer horizons (over one year), it is important to account for ‘teleconnection patterns’¹⁵ among vast global geographical areas. Thus, our findings, differently from Belbute and Pereira (2020), indicate that it is not only useful to consider the information included in the stochastic process underlying the baseline data but also a richer set of other information that changes over time.

Furthermore, Fig. 9 shows that the combination of models/factors always beats the naïve model and Fig. 10 highlights that there might be forecasting power gains by combining together macro, sentiment and nature-related factors (see models from MAIC1 to MBIC2). Superior predictive ability is found for the combinations of forecasts MBIC1, 99-Quantile that outperform the AR-4 benchmark for $h = 2$ (see online appendix). In some cases, the forecasting accuracy of combined predictions is superior to the accuracy of predictions coming from single models. For instance, for the six-quarters horizon, the MAIC7 combination of quantile predictions based on the 99-quantile method has a score ratio of 0.987 (and a p-value of 0.049 for the CW test, see online appendix). This value is smaller (and thus better) than the score ratios of the models M4 (containing macro quantile factors), M6 (including energy quantile factors), M12 (containing Arctic temperature quantile factors), and M13 (with Palmer drought quantile factors) on which the MAIC7 is based. Indeed, the four models have all score ratios larger than one in the case of the 99-quantile method. Our results indicate that adding to macroeconomic drivers, nature-related and sentiment factors extracted from surveys is a promising route to significantly improve CO_2 forecasting accuracy. This is relevant since accurate forecasting has always been at the forefront of decision-making and planning.

Finally, the results of White’s test (Table 9) indicate that the best of all the models under consideration always has superior predictive ability over the naïve Random Walk benchmark. Compared to the AR-4 baseline (Table 9), the best of the M1-M13 models (namely the linear M1 with macro factors, the 99-quantile M8 with drought, the 99-quantile M4 with quantile macro factors and the linear M5 with energy factors) outperforms this benchmark for longer time horizons (from $h = 2$ to $h = 8$ quarters). All in all, these results suggest that there has been no data snooping for the models enhanced with macro factors and models augmented with macro, sentiment and nature-related covariates.

¹⁵ Teleconnection patterns refer to large-scale changes in the atmospheric wave and jet stream dynamics.

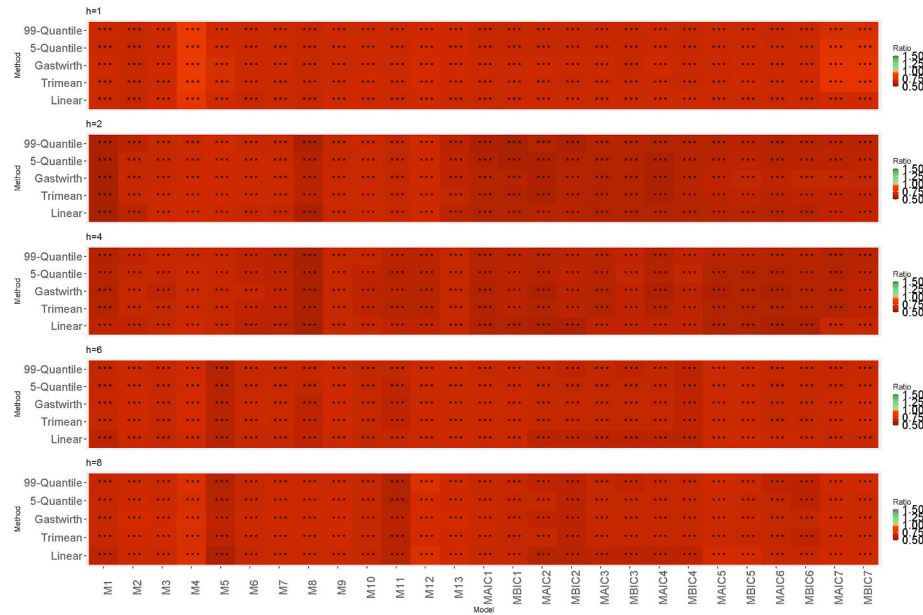


Fig. 9. Out-of-sample Analysis: Heatmap for SR, eq. (7). Naïve benchmark.

Fig. 9: Note: SR is the ratio between the RMSE of a given model and the one of the benchmark model (Random Walk). The stars in each cell are related to the p-values for the null hypothesis of equal predictability (test of Clark and West, 2007). ***, **, and *, denote significance at the 1%, 5% and 10% levels, respectively.

6. Conclusions

Forecasting has always been a key element in decision-making and planning. In this study, we perform point predictions of short- and medium-term, i.e. up to 2-year, US energy CO_2 emissions using different forecasting methods for the period 1972Q1–2021Q3. We consider, as potential predictors, large-scale series of anthropogenic and local (i.e. 48 States in the US plus the District of Columbia) and global interannual variability indicators and extract a small number of factors common to this set of variables using the classical Principal Component Analysis and the novel Quantile Factor Modelling Analysis. The extracted factors are then used as explanatory variables to predict CO_2 growth rates over different forecast horizons. We assess the in-sample and out-of-sample forecasting power of the resulting models and compare them to a naïve (Random Walk) and baseline (with only past quarterly CO_2 growth

rates) benchmark. The results show that models augmented with anthropogenic and nature-related indicators outperform the naïve and baseline settings, thus revealing their importance in predicting carbon dioxide emissions. In particular, factors extracted from macroeconomics data, energy series, drought indexes and Arctic temperature anomalies improve in-sample and out-of-sample forecasting power. In addition, market sentiment, especially the ‘soft’ information coming from the University of Michigan’s Surveys of Consumers, helps to improve in-sample forecasts. Furthermore, we document that model combinations tend to ameliorate out-of-sample forecasts particularly, over a two month forecasting horizon. These results have important policy implications. First, accurate emissions projections could derive not only from the information included in the stochastic process underlying the baseline data but also from a broader set of economic, nature-related and sentiment information that vary over time. Therefore

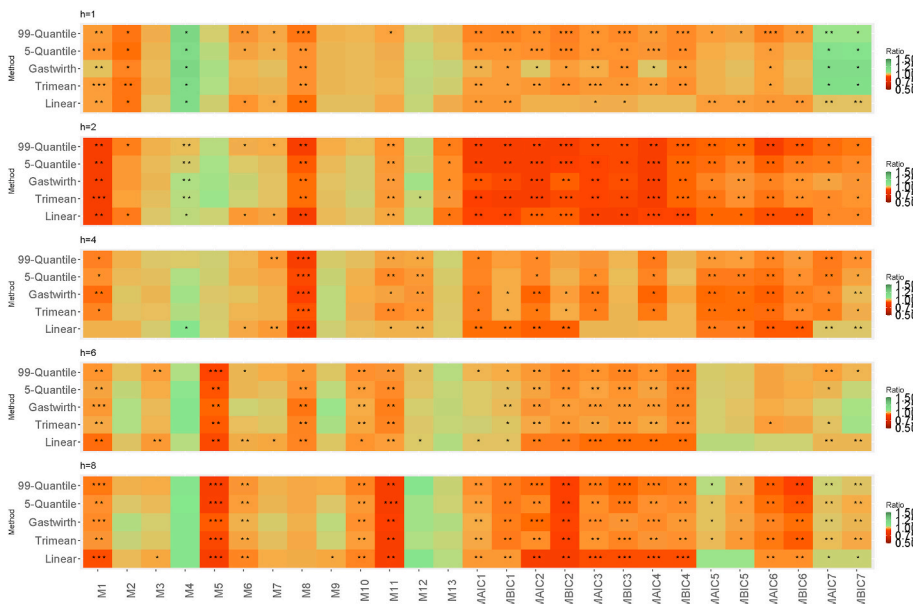


Fig. 10. Out-of-sample Analysis: Heatmap for SR, eq. (7). Benchmark: baseline model.

Fig. 10: Note: SR is the ratio between the RMSE of a given model and the one of the benchmark model (baseline model). The baseline model is given in eq. (6). The stars in each cell are related to the p-values for the null hypothesis of equal predictability (test of Clark and West, 2007). ***, **, and *, denote significance at the 1%, 5% and 10% levels, respectively.

Table 9
Results of White's multiple test.

h	Benchmark: Naïve		
	Best Model	CW pval	White pval
1	M2,5-Quantile	7.440E-07	1.000E-04
2	M1,Linear	1.945E-04	0.004
4	M8,99-Quantile	4.315E-04	0.003
6	M5,99-Quantile	2.996E-06	1.000E-04
8	M5,Linear	2.604E-04	0.004

h	Benchmark: Baseline		
	Best Model	CW pval	White pval
1	M2,5-Quantile	0.052	0.392
2	M1,Linear	0.014	0.022
4	M8,99-Quantile	0.006	0.042
6	M4,99-Quantile	0.010	0.014
8	M5,Linear	0.001	0.057

Note: The table reports, for each forecasting horizon h , the best model (based on RMSE), the p-value for the test Clark and West (2007) and the p-value for the Reality Check procedure of White (2000). The latter are bootstrap p-values based on 10,000 replicas for the null hypothesis that the best model (based on RMSE) has no predictive superiority over the benchmark model.

policymakers could ponder among several sources of information when assessing and forecasting the evolution of CO_2 emissions, including climate extremes and severe drought episodes. Second, short-medium-term forecasts represent a relevant tool to gauge the precise nature of the current CO_2 threats and to evaluate the policy effort involved in reaching the intermediate and final targets defined by IPCC for achieving carbon neutrality in 2050. Third, models such as the one proposed could provide information essential for further climate policy planning and subsequent negotiation agreements. Given the importance of monitoring pollutant emissions, future work could build on our approach complemented with models to forecast the probability of tail events, as in Algieri and Leccadito (2019), for GHG releases at a global level or to predict other pollutant components such as N_2O and SO_2 .

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The authors wish to thank the Editor, Ferda Halicioglu and two anonymous reviewers for their valuable comments and suggestions. Leonardo Iania acknowledges the support of the following research grant: ARC 18/23-089. The remaining errors are the authors' only responsibility.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cesys.2023.100112>.

References

Algieri, B., Füg, O., Lombardo, R., 2022. The Italian journey: carbon dioxide emissions, the role of tourism and other economic and climate drivers. *J. Clean. Prod.* 375, 134144.

Algieri, B., Leccadito, A., 2019. Ask CARL: forecasting tail probabilities for energy commodities. *Energy Econ.* 84 (C).

Belbute, J.M., Pereira, A.M., 2020. Reference forecasts for CO_2 emissions from fossil-fuel combustion and cement production in Portugal. *Energy Pol.* 144, 111642.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. B* 57 (1), 289–300.

Bennedsen, M., Hillebrand, E., Koopman, S.J., 2021. Modeling, forecasting, and nowcasting U.S. CO_2 emissions using many macroeconomic predictors. *Energy Econ.* 96, 105118.

Betts, R.A., Jones, C.D., Knight, J.R., Keeling, R.F., Kennedy, J.J., 2016. El Niño and a record CO_2 rise. *September Nat. Clim. Change* 6 (9), 806–810.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. *J. Econom.* 31 (3), 307–327.

Chen, L., Dolado, J.J., Gonzalo, J., 2021. Quantile factor models. *Econometrica* 89 (2), 875–910.

Clark, T.E., West, K.D., 2007. Approximately normal tests for equal predictive accuracy in nested models. *J. Econom.* 138 (1), 291–311.

Ding, Q., Xiao, X., Kong, D., 2023. Estimating energy-related CO_2 emissions using a novel multivariable fuzzy grey model with time-delay and interaction effect characteristics. *Energy* 263, 126005.

Elliott, G., Timmermann, A., 2016. *Economic Forecasting*. Princeton University Press.

ETH Zurich, 2018. Drought increases CO_2 concentration in the air. *ScienceDaily*, 29 August 2018 Technical Report. www.sciencedaily.com/releases/2018/08/180829133214.htm.

Filimonova, I., Nemov, V., Komarova, A., Mishenin, M., Kozhevnikov, V., 2021. Relationship of renewable energy consumption to economic, environmental and institutional factors in Europe (Technologies and Materials for Renewable Energy, Environment and Sustainability). *Energy Rep.* 7, 358–365.

Fosten, J., 2019. CO_2 emissions and economic activity: a short-to-medium run perspective. *Energy Econ.* 83, 415–429.

Friedlingstein, P., Jones, M.W., O'Sullivan, M., Andrew, R.M., Bakker, D.C.E., Hauck, J., Le Quéré, C., Peters, G.P., Peters, W., Pongratz, J., Sitch, S., Canadell, J.G., Ciais, P., Jackson, R.B., Alin, S.R., Anthoni, P., Bates, N.R., Becker, M., Bellouin, N., Bopp, L., Chau, T.T.T., Chevallier, F., Chini, L.P., Cronin, M., Currie, K.I., Decharme, B., Djutchouang, L.M., Dou, X., Evans, W., Feely, R.A., Feng, L., Gasser, T., Gilfillan, D., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, O., Harris, I., Houghton, R.A., Hurtt, G.C., Iida, Y., Ilyina, T., Luijckx, I.T., Jain, A., Jones, S.D., Kato, E., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J.I., Körtzinger, A., Landschützer, P., Lauvset, S.K., Lefèvre, N., Lienert, S., Liu, J., Marland, G., McGuire, P.C., Melton, J.R., Munro, D.R., Nabel, J.E.M.S., Nakaoka, S.-I., Niwa, Y., Ono, T., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Rosan, T.M., Schwinger, J., Schwingshackl, C., Séférian, R., Sutton, A.J., Sweeney, C., Tanhua, T., Tans, P.P., Tian, H., Tilbrook, B., Tubiello, F., van der Werf, G.R., Vuichard, N., Wada, C., Wanninkhof, R., Watson, A.J., Willis, D., Wiltshire, A.J., Yuan, W., Yue, C., Yue, X., Zaehle, S., Zeng, J., 2022. Global carbon budget 2021. *Earth Syst. Sci. Data* 14 (4), 1917–2005.

Gao, M., Yang, H., Xiao, Q., Goh, M., 2021. A novel fractional grey riccati model for carbon emission prediction. *J. Clean. Prod.* 282, 124471.

Ghysels, E., 2014. Conditional skewness with quantile regression models: SoFIE presidential address and a tribute to Hal White. *J. Financ. Econom.* 12 (4), 620–644.

Halsey, M., 2022. *Establishing Policies and Procedures for IT Use*. Apress, Berkeley, CA, pp. 117–136.

Hamilton, J.D., 2011. Historical oil shocks. *NBER Working Papers* (16790), 1–52.

Heim, R.R., Lawrimore, J.H., Wuertz, D.B., Waple, A.M., Wallis, T.W.R., 2003. The RDTI and MSI: two new national climate impact indices, 10 *J. Appl. Meteorol.* 42 (10), 1435–1442.

Helm, D., Hepburn, C., Ruta, G., 2012. Trade, climate change, and the political game theory of border carbon adjustments. *Oxf. Rev. Econ. Pol.* 28 (2), 368–394.

Humphrey, V., Zscheischler, J., Ciais, P., Gudmundsson, L., Sitch, S., Seneviratne, S., 2018. Sensitivity of atmospheric CO_2 growth rate to observed changes in terrestrial water storage. *Nature* 560 (7720), 628–631.

Isik, C., Kasimati, E., Ongan, S., 2017. Analyzing the causalities between economic growth, financial development, international trade, tourism expenditure and/on the CO_2 emissions in Greece. *Energy Sources B Energy Econ. Plann.* 12 (7), 665–673.

Intergovernmental Panel on Climate Change, IPCC, 2022. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.

Jolliffe, I.T., Cadima, J., 2016. Principal component analysis: a review and recent developments. *Philos. Trans. A Math. Phys. Eng. Sci.* 374 (4), 2–16.

Koenker, R.W., Bassett, G., 1978. Regression quantiles. *Econometrica* 46 (1), 33–50.

Kushnir, Y., Scaife, A.A., Arritt, R., Balsamo, G., Boer, G., Doblas-Reyes, F., Hawkins, E., Kimoto, M., Kolli, R.K., Kumar, A., Matei, D., Matthes, K., Müller, W.A., O'Kane, T., Perlwitz, J., Power, S., Raphael, M., Shimp, A., Smith, D., Tuma, M., Wu, B., 2019. Towards operational predictions of the near-term climate. *Nat. Clim. Change* 9, 94–101.

Lima, L.R., Meng, F., Godeiro, L., 2020. Quantile forecasting with mixed-frequency data. *Int. J. Forecast.* 36 (3), 1149–1162.

Luo, X., Keenan, T., 2022. Tropical extreme droughts drive long-term increase in atmospheric CO_2 growth rate variability. *Nat. Commun.* 13.

Marcellino, M., Stock, J.H., Watson, M.W., 2006. A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. *J. Econom.* 135 (1), 499–526.

Matthews, H.D., Tokarska, K.B., Rogelj, J., Smith, C.J., MacDougall, A.H., Karsten Haustein, N.M., Sippel, S., Forster, P.M., Knutti, R., 2021. An integrated approach to quantifying uncertainties in the remaining carbon budget. *Nature Communications Earth Environment* 2 (7).

- McCracken, M.W., Ng, S., 2020. Fred-qd: A Quarterly Database for Macroeconomic Research, 005. Federal Reserve Bank of St. Louis Working Paper, pp. 1–52.
- Meangbua, O., Dhakal, S., Kuwornu, J.K., 2019. Factors influencing energy requirements and CO₂ emissions of households in Thailand: a panel data analysis. *Energy Pol.* 129, 521–531.
- Moral-Benito, E., 2015. Model averaging in economics: an overview. *J. Econ. Surv.* 29 (1), 46–75.
- Munksgaard, J., Pedersen, K.A., 2001. CO₂ accounts for open economies: producer or consumer responsibility? *Energy Pol.* 29 (4), 327–334.
- Olaizola, A., Dietzenbacher, E., 2014. Drivers of the growth in global greenhouse gas emissions. *Environ. Sci. Technol.* 48 (10), 5388–5394.
- Ozturk, I., 2010. A literature survey on energy–growth nexus. *Energy Pol.* 38 (1), 340–349.
- Palmer, W.C., 1985. Meteorological Drought, vol. 45. U.S. Weather Bureau Research Paper, pp. 1–65.
- Peeters, P., Dubois, G., 2010. Tourism travel under climate change mitigation constraints. *J. Transport Geogr.* 18 (3), 447–457 (Tourism and climate change).
- Raupach, M.R., Marland, G., Ciais, P., Quéré, C.L., Canadell, J.G., Klepper, G., Field, C.B., 2007. Global and regional drivers of accelerating CO₂ emissions. *Proc. Natl. Acad. Sci. USA* 104 (24), 10288–10293.
- Ray, R., Griffin, R., Fares, Elhassan, A.A., Awal, R., Woldeesenbet, S., Risc, E., 2020. Soil CO₂ emission in response to organic amendments, temperature, and rainfall. *Nature scientific reports* 10 (5849).
- Ren, F., Long, D., 2021. Carbon emission forecasting and scenario analysis in guangdong province based on optimized fast learning network. *J. Clean. Prod.* 317, 128408.
- Rode, A., Carleton, T., Delgado, M., Greenstone, M., Houser, T., Hsiang, S., Hultgren, A., Jina, A., Kopp, R.E., McCusker, K.E., Nath, I., Rising, J., Yuan, J., 2021. Estimating a social cost of carbon for global energy consumption. *Nature* 598, 308–314.
- Taylor, J.W., 2007. Forecasting daily supermarket sales using exponentially weighted quantile regression. *Eur. J. Oper. Res.* 178 (1), 154–167.
- Tokarska, K.B., Arora, V.K., Gillett, N.P., Lehner, F., Rogelj, J., Schleussner, C.-F., Seferian, R., Knutti, R., 2020. Uncertainty in carbon budget estimates due to internal climate variability. *Environ. Res. Lett.* 15 (10).
- UNEP, 2022. Emissions Gap Report 2022: the Closing Window — Climate Crisis Calls for Rapid Transformation of Societies. Technical report. United Nations Environment Programme, 27 October.
- Wang, C.-M., Wu, T.-P., 2022. Does tourism promote or reduce environmental pollution? evidence from major tourist arrival countries. *Environ. Dev. Sustain.: A Multidisciplinary Approach to the Theory and Practice of Sustainable Development* 24 (3), 3334–3355.
- Wang, Z., Wu, J., Liu, C., Gu, G., 2017. Integrated Assessment Models of Climate Change Economics. Springer Singapore.
- White, H., 2000. A reality check for data snooping. *Econometrica* 68 (5), 1097–1126.
- Zhou, Y., Eom, J., Clarke, L., 2013. The effect of global climate change, population distribution, and climate mitigation on building energy use in the U.S. and China. *Climatic Change* 119, 979–992.