Environmental Impact Assessment Review 115 (2025) 107987

Contents lists available at ScienceDirect

Environmental Impact Assessment Review

journal homepage: www.elsevier.com/locate/eiar

Disentangling and hedging global warming risk: A machine learning approach

Shusheng Ding[a], Tianxiang Cui[b], Anna Min Du[c,*], John W. Goodell[d], Nanjiang Du[b]

[a] Business School, Ningbo University, China
[b] School of Computer Science, University of Nottingham Ningbo, China
[c] The Business School, Edinburgh Napier University, United Kingdom
[d] College of Business, The University of Akron, Akron, OH, United States of America

ARTICLE INFO

JEL code:
G17

Keywords:
Climate finance
Extreme gradient boosting
Global warming risk
Greenhouse gas emission

ABSTRACT

As global warming provokes increasing attention from investors, this study disentangles global warming risk (GWR) for investors by leveraging energy futures volatilities. This study derives GWR from energy futures using an extreme gradient boosting (XGB)-genetic programming (GP) framework. Our XGB-GP framework develops volatility forecasting models for GWR from selected energy futures markets identified by XGB as key contributors to global warming, surpassing traditional models in forecasting accuracy. The originality of the study rests on the pioneering integration of the XGB-GP framework in predicting climate risk, linking energy futures markets with climate risk management and enabling feasible climate-featured portfolio hedging. Our study also sheds new insights for policymakers to design carbon trading systems and carbon pricing mechanisms, as they can use relevant energy futures prices as a basis for carbon trading calibration.

1. Introduction

Global warming has become one of the major concerning issues facing by human community at the moment (Amicarelli et al., 2021; Wu et al., 2024). The growing greenhouse gas (GHG) emissions resulted in notable changes toward our climate, putting human society and environment into potential catastrophe (Agnew, 2024; Zhang et al., 2023). GHG emissions, on the other side, also have broadly impacted economic development and firm values. As noted by Nishitani and Kokubu (2012), reducing firm GHG emissions is a corporate social responsibility (CSR) of firms, and CSR success could reinforce firms' reputational capital. Further, Cooper et al. (2018) reveal that the negative effect of GHG emissions on a company's value is undermined by the damage it causes to the company's reputation for social responsibility. Consequently, scholars in climate finance are investigating how investors exercise shareholder disciplines over firms with regard to companies facilitating a global transformation from an energy-intensive economy to a sustainable economy through climate risk management (Hong et al., 2020; Stroebel and Wurgler, 2021). On these bases, the study of climate risk impacting on company value analysis attracts the attention of scholars (Bartram et al., 2022; Bose et al., 2024; Pham et al., 2024). There has been a heightened awareness among investors regarding the implications of climate risk, demanding to a greater transparency in the reporting such risk for companies (Huang et al., 2018). As a result, analyzing and hedging climate risks, such as global warming risk (GWR), are key ingredients of risk management for firms.

Nevertheless, one of the pivotal obstacles in assessing the effects of climate change on firm reputations, values, and behaviors is the complicated process of determining the explicit mechanism through which companies and financial markets are impacted. These concerns further extend to whether or not financial markets can help to evaluate and price those risks and potentially aid the risk management of climate change for investors and financial institutions (Sautner et al., 2023). Commodity futures markets, especially energy futures markets have significantly impacted on the climate risks and sustainable development (Gong et al., 2023; Hoque et al., 2023). As a result, the research question how financial market dynamics, particularly energy futures volatilities, can be leveraged to estimate and predict global warming risk (GWR) through greenhouse gas (GHG) emission fluctuations has been brought into being. This research theme is pivotal as GHG emissions may jeopardize economic development and firm values (Cruz and Rossi-Hansberg, 2024).

* Corresponding author.
E-mail addresses: dingshusheng@nbu.edu.cn (S. Ding), tianxiang.cui@nottingham.edu.cn (T. Cui), a.du@napier.ac.uk (A.M. Du), johngoo@uakron.edu (J.W. Goodell), scynd1@nottingham.edu.cn (N. Du).

https://doi.org/10.1016/j.eiar.2025.107987
Received 11 February 2025; Received in revised form 29 April 2025; Accepted 11 May 2025
Available online 14 May 2025
0195-9255/

While prior studies explored the climate risks' economic consequences, few actionable hedging mechanisms have been proposed against such risk. Therefore, our paper intends to provide new understanding of global warming risk (GWR) in terms of GHG emission volatility by employing a machine learning approach from a financial market perspective. On this basis, this study further proposes a hedging portfolio from energy futures markets to neutralize the negative effect of GWR. Overall, we could furnish the research gap by offering the deep understanding regarding how financial markets price GWR and how markets can hedge against such risk.

In fact, global warming mainly stems from GHG emissions, and thus, this study uses GHG emission volatility to approximate the GWR. The main types of GHGs included in this paper are carbon dioxide (CO_2), nitrous oxide (N_2O), and sulfur hexafluoride (SF_6). CO_2, N_2O and SF_6 are the major contributors to global warming (Allen et al., 2009; Battaglia and Joos, 2018; Li et al., 2018). CO_2 is the most prevalent greenhouse gas emitted by human activities. It absorbs and re-emits infrared radiation, frustrating the heat release from Earth's atmosphere, resulting in the greenhouse effect. Increased levels of CO_2 amplify this effect, leading to increased global temperatures (Franta, 2018; Matthews et al., 2009). Although N_2O is emitted in smaller quantities than CO_2, it is a much more potent greenhouse gas.

Existing literature has identified the GHG emissions as key driver of global warming, generating profound effects on ecosystems and economies (Geiger et al., 2021). The transparency and accuracy of climate risk is increasingly demanding, yet existing climate risk studies prevalently rest their forecasting models on environmental data, neglecting the interconnectedness between energy markets and GHG volatilities. Current risk forecasting models, such as GARCH-family models, including GARCH-MIDAS model and HAR-RV models can be only applied in the area of energy futures market risk forecasting, few models can be used to predict GWR since the time series data of GWR is scarce. Our remarkable contribution thereby is proposing a cutting-edge methodology to scrutinize GWR in terms of GHG emission volatility from a novel angle. We synthesize the Extreme Gradient Boosting (XGB) with Genetic Programming (GP) method to create an XGB-GP methodology to examine the GWR. This framework challenges the "black box" paradigm of machine learning by providing interpretable GWR forecasting model, which considerably reduces forecasting errors compared to traditional methods. By employing the energy future market data, our framework further decomposes GWR using energy futures, which delivers practical hedging strategies operationalize climate risk management.

Since analyzing the GWR from the firm values and energy markets angle is crucial, our paper contribute to the existing literature by filling the research gap in several ways. Firstly, existing studies scrutinize the climate risk through the lens of climate models or environmental indicators, which may not fully capture the dynamic interconnection between financial markets and climate risk (Bovari et al., 2018; Lamperti et al., 2019). There is a burgeoning number of literature suggests that financial market factors like energy market factors, are sound predictors of greenhouse gas emissions (Brehm, 2019; Guo et al., 2021). However, current literature has not scrutinized multiple factors in a comprehensive framework by including financial market factors to forecast the GWR. Our paper can fulfill this research gap by putting GWR into the economic context by revealing how energy futures respond to GWR and, more importantly, which energy markets contribute more to GWR than other markets. This study enriches the global warming research area from an economic perspective other than the existing studies, which are based on environmental information. This scarcity of linking energy futures market volatilities directly to GWR highly motivates our research. Our study thereby furnishes the research gap by establishing a predictive model, which uses machine learning approach to learn from financial market data as a basis for forecasting GWR, providing a new perspective on how financial activities in the energy sector can influence environmental outcomes.

Without such vigorous climate risk forecasting model, current literature leaves the vital research question how to hedge and alleviate climate risk still open, given aforementioned massive climate risk effects on financial investments. Lacking of hedging strategies toward GWR substantiates the research motivation of our paper. Additionally, since the factors related to human activities cannot be used to hedge, our research lens of this paper is to focus on the GWR related to the financial markets and how the financial markets can be used to hedge such risk. The energy futures markets and the financial instruments that can be used for hedging are the key players of our paper. Our study thereby challenges this conventional paradigm by linking energy market volatilities, to the volatilities of GHG emissions, including CO_2, N_2O, and SF_6. By demonstrating that energy futures markets can serve as leading indicators of GHG emission risks, this study attempts to provide a deeper understanding of the global warming risk structure.

In addition, our XGB-GP framework can also shed new insights into GWR deconstruction. The key advantage of our XGB-GP framework is that the final model is a mathematical expression that can be interpreted and analyzed. This is a notable contrast to most machine learning models, which are often seen as black boxes (Chen et al., 2024). In addition, by using XGBoost's information gain for feature selection, we can identify the most relevant features from the dataset, reducing the dimensionality and complexity of the problem. This makes the subsequent GP process more efficient and focused on the most important relationships in the data. Subsequently, GP can automatically discover complex feature interactions and non-linear relationships, which might not be easily captured by traditional statistical methods or even other machine learning techniques (Jin et al., 2024).

Furthermore, our XGB-GP framework unravels that key contributors to CO_2 emission volatility are the volatility of Brent oil futures, the volatility of coking coal futures, the volatility of gas futures, and the volatility of gasoline futures. The key contributors to N_2O emission volatility are the volatility of coking coal futures, the volatility of gasoline futures, the volatility of thermal coal futures, and the volatility of WTI oil futures. Finally, the futures that have the highest prediction power for SF_6 emission volatility are the volatility of Brent oil futures, the volatility of coke futures, the volatility of heating oil futures, and the volatility of Rotterdam Coal futures. Based on the key contributors of the three GHG volatilities, we developed three GHG volatility forecasting models (see Eqs. (17)–(19)). By employing those three volatility forecasting models, our results indicate that our XGB-GP framework exhibits a larger prediction accuracy than the traditional MIDAS model, with an overall forecasting error of approximately 2 %. Based on our key contributors of GHG emission volatilities and volatility forecasting models, we propose a hedging portfolio for neutralizing three GHG emission volatilities (see Table 4), which could be extremely useful for firms inside related industries.

The structure of this paper can be outlined as follows. In Section 2, we provide an overview of the methodology applied in this paper as well as the sample data and variable measures regarding the GHG volatilities and energy futures market volatilities. Section 3 presents the empirical results and the model performance of GWR forecasting. In Section 4, we further discussed the implications of our empirical findings. Section 5 covers the conclusion summary of the study and highlights relevant research implications.

2. Data and methodology

2.1. Data and variables

In order to investigate the predictability of futures market volatilities on the GWR, we proxy GWR as the volatility of three GHG emissions, with these three GHGs the main contributors to global warming (Lashof and Ahuja, 1990; Shine et al., 2005; Zhang et al., 2021a). Specifically, the three main types of GHGs we include in this study are carbon dioxide (CO_2), nitrous oxide (N_2O), and sulfur hexafluoride (SF_6). Further, as the emission of GHG mainly stems from the burning of fossil fuels, including coal, gas and crude oil (Gillingham and Stock, 2018; Lichtfouse et al., 2003), we use different energy futures markets to represent the volatility of fossil fuel within the financial markets as well as the hedging instruments. We define t_i as the conditional volatility of a particular time series i for time t. Therefore, the conditional volatilities of the three gases are tCO_2, tN_2O, tSF_6, respectively. Conditional volatility is estimated with GARCH modeling. We use the percentage change for all variables in this paper to fit the GARCH model.

In particular, we include 12 major energy futures volatilities and three GHG emission volatilities, defining two vectors as described below:

$$\overrightarrow{x} = \sigma_t^{bo}, \sigma_t^{cc}, \sigma_t^{cf}, \sigma_t^{do}, \sigma_t^{fo}, \sigma_t^{gas}, \sigma_t^{go}, \sigma_t^{ho}, \sigma_t^{tc}, \sigma_t^{rbc}, \sigma_t^{rc}, \sigma_t^{wo}, \tag{1}$$

where σ_t^{bo} the volatility of Brent oil futures, $\sigma_t^{cc}$is the volatility of coking coal futures, σ_t^{cf} is the volatility of coke futures, σ_t^{do} is the volatility of diesel oil futures, $\sigma_t^{fo}$is the volatility of fuel oil futures, $\sigma_t^{gas}$is the volatility of gas futures, σ_t^{go} is the volatility of gasoline futures, σ_t^{ho} is the volatility of heating oil futures, σ_t^{tc} is the volatility of Richards Bay Coal futures, σ_t^{rbc} is the volatility of Rotterdam Coal futures, σ_t^{rc} is the volatility of thermal coal futures, and σ_t^{wo} is the volatility of WTI crude oil futures. Further,

$$\overrightarrow{y} = \sigma_t^{CO2}, \sigma_t^{N2O}, \sigma_t^{SF6}, \tag{2}$$

where σ_t^{CO2} is the volatility of CO_2 emissions, σ_t^{N2O} is the volatility of N_2O emissions, and $\sigma_t^{SF6}$is the volatility of SF_6 emissions.

Additionally, we employ an environmental-related index from the stock market in China, which includes the environment protection sector (EP), hydropower development sector (HY), new energy vehicle sector (NEV), nuclear power development sector (NP), and solar photovoltaic sector (SP). These indices reflect the firm share price in those industries that are highly related to environmental protection and global warming issue.

The sample period covers 1 January 2016 to 1 January 2023. We select this period because the Paris Agreement, signed in December 2015 as part of the United Nations Framework Convention on Climate Change, brought the global warming issue into a legal form (Guiot and Cramer, 2016). Therefore, we use the GHG index as a main explanatory variable to represent the GWR highlighted by the Paris Agreement. GHG index data are provided by the Global Monitoring Laboratory (GML) of the National Oceanic and Atmospheric Administration (NOAA), which is freely available to the public (Lan et al., 2023). Future market data on coal, gas, crude oil, and environmental-related stock indices are collected from WIND. The GHG index is monthly data, whereas the futures market data are daily data. Therefore, we use a high-frequency to low-frequency mapping technique for GWR prediction. We further use the MIDAS method as the benchmark model to evaluate our mapping technique. We describe our GARCH modeling and MIDAS methodology in the following two subsections.

2.2. GARCH model

GARCH modeling is widely used for volatility estimation and forecasting, especially for time-varying volatility in the finance area (Ding et al., 2019; Zhang et al., 2024). Consequently, we use GARCH modeling to estimate the volatility of GHG emissions, the volatility of the stock market subindex, and the volatility of futures prices.

The widely used GARCH (1,1) model takes the following format:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \sigma_{t-1}^2 + \alpha_2 \varepsilon_{t-1}^2 \tag{3}$$

where σ_t is the volatility of the time series employed (namely, the percentage change of GHG, stock market subindex and futures prices) and ε_t is the residual term taken from the conditional mean equation, which is:

$$r_t = \varnothing + \varepsilon_t, \tag{4}$$

Ø presents the conditional mean and $\varepsilon_t \sim N(0, \sigma_t^2)$.

2.3. Mixed data sampling

The dynamic conditional volatility for GHG emissions is on a monthly basis, and so there is a relatively limited data sample. To further forecast GHG emissions with a fruitful data sample, we employ the data from the energy market conditional volatility, which is on a daily basis and can enhance the forecasting sample capacity. As those two datasets have different data frequencies, the application of the mixed data sampling (MIDAS) regression method is essential, as it can serve as the benchmark model to compare with our XGB-GP method.

Ghysels et al. (2007) developed the MIDAS model to handle mixed-frequency data in one regression model, especially when the explanatory variables are at a high frequency. The MDAS model can be formulated as follows:

$$\sigma_{t+1}^M = \alpha + \beta \sum_{i=0}^{q^D-1} \sum_{j=0}^{N^D-1} \omega_{i+j^*N^D}\left(\varphi^D\right)\sigma^D_{N^D-j,t-i} + \varepsilon_{t+1}, \tag{5}$$

where σ_{t+1}^M is the low-frequency data with the frequency M (the monthly GHG emission volatility at time t + 1, q^D is the number of lags for the daily frequency data (the daily energy market conditional volatility), N^D is the number days in one month (21 trading days in one month), ω is the weighting function regarding the energy market conditional volatility toward the GHG emission volatility, and φ^D represents the parameters in the Almon lag polynomial function since we use the Almon lag weighting (i.e., polynomial distributed lag) method for the MIDAS regression with the residual term ε_{t+1}.

2.4. Extreme gradient boosting

Extreme gradient boosting (XGBoost) (Chen and Guestrin, 2016) is an advanced implementation of the gradient boosting algorithm. Gradient boosting is a method that goes through cycles to iteratively add models into an ensemble. It is a sequential technique that combines weak learners to create a strong learner by assigning accurate predictions on the basis of a cumulative error function. XGBoost was originally designed for speed and performance. It has several key advantages that have made it a popular choice among data scientists. First, it employs a highly efficient and powerful algorithm known for its speed and performance. Second, it includes built-in regularization parameters that impede overfitting. This means that it not only improves model performance but also controls the model's complexity, making it less likely to overfit to the training data. Third, it offers flexibility by supporting various objective functions, including regression, classification, and ranking. Additionally, XGBoost can handle missing values internally, reducing the need for external preprocessing. Finally, it supports parallel processing, making it highly scalable, and has built-in cross- validation capabilities at each iteration, making model selection easier.

The XGBoost algorithm follows a boosting framework where models are trained sequentially, and each subsequent model aims to correct the errors of the previous one. The general procedure can be broken down into the following steps:

(1) *Model Initialization*

XGBoost begins by initializing a model with a constant value. This value is chosen to minimize the loss function *L*, which measures the difference between the predicted values and the actual target values based on the problem type. For example, in a regression problem with mean squared error (MSE) as the loss function, the model is initialized with the mean of the target values:

$$f_0(\vec{x}) = arg_\gamma min \sum_{i=1}^{n} L(y_i, \gamma) \tag{6}$$

where:

- n is the number of data instances,
- $\vec{x}$ represents the features of the i-th instance,
- y_i is the target value for the i-th instance,
- γ is the constant value that minimizes the loss function during initialization (e.g., the mean of target values for regression).

The loss function L varies depending on the task:

- For regression problems, a common choice is the Mean Squared Error (MSE):

$$L(y_i, \hat{y}_i) = (y_i - \hat{y}_i)^2,$$

where $\hat{y}_i$ is the predicted value.

- For binary classification problems, a common loss function is Logistic Loss (Log Loss):

$$L(y_i, \hat{y}_i) = -[y_i\, log(\hat{y}_i) + (1 - y_i,)log(1 - \hat{y}_i)].$$

- For multiclass classification problems, a Softmax loss is often used to account for multiple class probabilities.

In this step, the loss function L is minimized to initialize the model with a constant value. It's important to note that here is not the scaling factor used for the base learners; that is introduced in the boosting rounds.

(2) *Boosting Rounds (form m = 1 to M)*

XGBoost runs for M boosting rounds (iterations), where each round builds a new base learner that is added to the model from the previous round.

(a) Compute Pseudo-Residuals

In each round m, the pseudo-residuals r_{im} are calculated as the negative gradient of the loss function L, evaluated at the current prediction $f_{m-1}(\vec{x})$:

$$r_{im} = -\left[\frac{\partial L\left(y_i, f\left(\vec{x_i}\right)\right)}{\partial f\left(\vec{x_i}\right)}\right]_{f=f_{m-1}} \quad \text{for } i = 1, \dots, n. \tag{7}$$

These pseudo-residuals represent the errors made by the current model and guide the construction of the next base learner.

(b) Fit a Base Learner to the Pseudo-Residuals

A base learner (typically a decision tree) $h_m(\vec{x})$is then fitted to predict the pseudo-residuals. This involves constructing the tree by splitting the data into partitions that minimize the sum of residuals within each partition.

(c) Compute the Optimal Multiplier

After fitting the base learner, the algorithm finds an optimal multiplier m that minimizes the loss when the new base learner is added to the current model:

$$\gamma_m = arg_\gamma min \sum_{i=1}^{n} L\left(y_i, f_{m-1}\left(\vec{x_i}\right) + \gamma h_m\left(\vec{x_i}\right)\right).$$

This step ensures that the new base learner is appropriately scaled to maximize its contribution to improving the model's performance.

(d) Update the Model

The current model is updated by adding the newly scaled base learner to the previous model:

$$f_m(\vec{x}) = f_{m-1}(\vec{x}) + \gamma_m h_m(\vec{x}).$$

This process repeats for each boosting round, gradually refining the model to reduce the error in the predictions.

(3) *Final Model Output*

After M boosting rounds, the final model is the sum of the initial model and all the subsequent base learners:

$$\hat{f}(\vec{x}) = f_M(\vec{x}). \tag{8}$$

This final model is used for making predictions on new data.

(4) *Regularization*

XGBoost includes regularization terms to prevent overfitting and improve generalization. This is done by adding a regularization term $\Omega(h_m)$ to the objective function when computing m:

$$\gamma_m = arg_\gamma min\left[\sum_{i=1}^{n} L\left(y_i, f_{m-1}\left(\vec{x_i}\right) + \gamma h_m\left(\vec{x_i}\right)\right) + \Omega(h_m)\right]. \tag{9}$$

The regularization function $\Omega(h_m)$ typically takes the form:

$$\Omega(h_m) = \alpha T + \frac{1}{2}\lambda\|\omega\|^2,$$

where:

- T is the number of leaves in the decision tree,
- ω is the vector of leaf weights,
- α and λ are regularization parameters that control the model's complexity and prevent overfitting.

This regularization improves the robustness of the model, allowing it to generalize better on unseen data.

2.5. *Genetic programming*

Genetic Programming (GP) (Poli et al., 2008) is an evolutionary computation technique that automatically solves problems without the user having to know or specify the form or structure of the solution in advance. At the heart of GP is the idea of evolving computer programs using principles of evolution and natural selection, such as mutation, crossover (recombination), and survival of the fittest.

The general procedure of GP is as follows:

(1) Population Initialization: A population of randomly generated computer programs (individuals) is created. Each individual is a potential solution to the problem at hand. We denote the initial population as P_0 where $P_0 = \{p_1, p_2, \dots, p_n\}$, and n is the size of the population.
(2) Fitness Evaluation: Each individual in the population is evaluated for its fitness, i.e., how good of a solution it is to the problem at hand. This can be represented by a fitness function f: $P \rightarrow R$,

where P is the population and R is the real number. The fitness of an individual program pi in the population is given by f (pi).

(3) Selection: individuals are probabilistically selected from the population for reproduction (mutation and crossover) based on their fitness. The fitter the individual, the higher the chance of being selected. This can be represented by a selection functions f: $P \times F \rightarrow P'$, where F is the fitness value of the population, and P' is the selected population. $P' = s\ (P,\ F)$.

(4) Reproduction: The selected individuals undergo genetic operations such as crossover and mutation to produce offspring. Crossover can be represented as c: $P' \times P' \rightarrow P''$, where P'' is the new population after crossover. Mutation can be represented as m: $P'' \rightarrow P''$', where P''' is the final new population after mutation.

(5) Replacement: The newly created offspring replace some or all of the individuals in the current population. This completes one generation.

(6) Termination: The above steps are repeated for many generations or until some termination conditions are met, such as a solution that satisfies minimum criteria, fixed number of generations reached, or allotted computation time reached. The final best solution is given by $p_{best} = argmax_{p \in P} f(p)$.

2.6. Neural networks

In this study, we consider one simple form of the neural network, namely, the single-layer perceptron. It consists of an input layer and an output layer that are fully connected. Generally, it works as follows:

We denote the input as a vector $\vec{x} = [x_1, x_2, \dots, x_n]^T$, where n is the dimension of the input, and the weights if the neural network as a vector $\vec{w} = [w_1, w_2, \dots, w_n]^T$.

(1) Linear Transformation: The first step in the neural network is a linear transformation of the input, which can be represented as a dot product between the input and the weight vectors:

$$z = \vec{w}^T \vec{x} + b, \tag{10}$$

where b is a bias term.

(2) Activation Function: The result of the linear transformation is then passed through an activation function. The choice of activation function can vary. In this work, we choose the sigmoid function, which can be represented as follows:

$$\sigma(z) = \frac{1}{1 + e^{-z}}. \tag{11}$$

Therefore, the output of the neural network, given an input $\vec{x}$, is:

$$\hat{y} = \sigma(\vec{w}^T \vec{x} + b) \tag{12}$$

The weights $\vec{w}$ and bias b are learned from the data by minimizing a loss function, which measures the difference between the network's predictions and the true values. For a binary classification problem, a common choice of loss function is the binary cross entropy loss:

$$L(y, \hat{y}) = -y log(\hat{y}) - (1 - y) log(1 - \hat{y}), \tag{13}$$

where y is the true value and $\hat{y}$ is the predicted value. The weights and bias are updated by performing gradient descent on this loss function.

2.7. The proposed XGB-GP framework

As the prediction of GWR has no unanimous answer over the existing literature, following the vein of energy market risk prediction, the theoretical foundation of climate risk forecasting is inherently tied to synthesizing factors such as macroeconomic variables, and spanning historical volatility patterns, extracting available information from the historical data (Fu et al., 2024; Liu et al., 2018). In order to acquire information from the historical data, a large number of models have been applied in the volatility forecasting field, especially energy market risk prediction. Early GARCH-family models, leveraged low-frequency data but constrained in capturing intraday volatility dynamics (Escobar-Anel et al., 2025; Naysary and Shrestha, 2024), provoking innovations like HAR-RV models that integrate high-frequency data to encapsulate multi-scale risk persistence (Li et al., 2025). Additionally, popularized by artificial intelligentized models, machine learning techniques, such as LASSO and gradient-boosted decision trees (GBDT), have differentiated themselves for their ability to handle high-dimensional data and incorporate nonlinear relationships, outperforming traditional linear models volatility forecasts (Lee et al., 2022; Zhang et al., 2021b).

Our XGB-GP framework further extended the existing method to create interpretable climate risk forecasting models, whereas machine learning models usually produce black-box outputs without sufficient interpretability. The overall structure of our proposed XGB-GP framework is illustrated in Fig. 1. We trained the XGB model on the dataset and compute feature importance scores using information gain, which is used to measure the effectiveness of a feature in splitting the data to create more homogeneous subsets. It is based on the concept of entropy, which quantifies the randomness or impurity of a dataset. In this work, information gain is used to determine the best split points in the tree structure. By extracting feature importance, we can identify the most relevant features in the dataset. Next, we select a subset of the top 4 most important features for GP processing.

We then use GP to evolve mathematical expressions that describe the relationship between these top 4 features and the target variable. We initialize a population of random symbolic expressions containing the selected features and evolve the population using GP's standard processes, such as selection, crossover, and mutation. The fitness of an individual expression can be evaluated based on its prediction accuracy. We continue the GP process for a fixed number of generations.

3. Results

3.1. Baseline results

Before we further analyze the GWR, it is essential to demonstrate the significance of GHG volatility on the financial investments, such as impacting on the stock market. Therefore, we scrutinize the influence of GHG volatility on the volatility of stock subindices, demonstrating the nexus of GWR with financial markets and firm value. In particular, we employ five environmental-related stock subindices in China: the environmental protection sector index, hydropower development sector

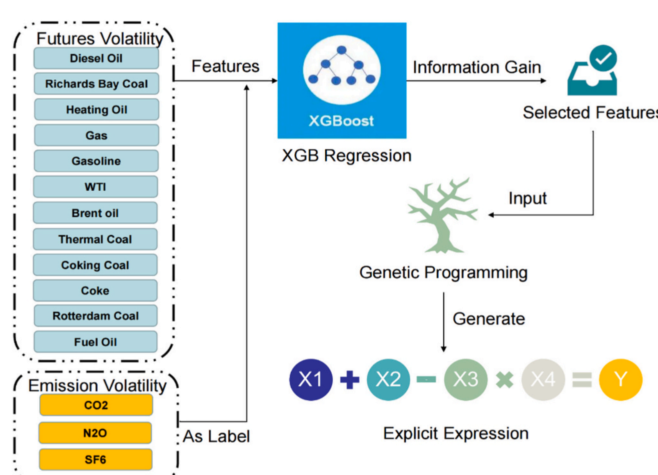

Fig. 1. The proposed XGB-GP framework.

index, new energy vehicle sector index, nuclear power development sector index and solar photovoltaic sector index. Those five stock subindices are constituted by the share price of listed companies that are significantly involved in the area of environmental protection and emission reduction.

Table 1 presents the baseline result of our paper, which delivers empirical data analysis of the five stock subindices volatilities regarding the impact of three GHG volatilities based on Eq. (14). It is observable that the CO_2 emission volatility generates considerable impact on the stock subindices of the hydropower development sector, new energy vehicle (NEV) sector, nuclear power development sector and solar photovoltaic sector. In particular, CO_2 emission volatility has the most significant impact on new energy vehicles according to the coefficients and statistical significance. Since CO_2 emissions play key roles in global warming, governments worldwide have implemented policies to truncate CO_2 emissions, including promoting NEV utilization. In countries where there are strict regulations on CO_2 emissions to achieve carbon neutrality, the NEV industry has experienced substantial growth. As in China, the government has implemented policies to promote NEVs, including subsidies and tax exemptions, which have led to a surge in NEV sales.

However, on the other hand, if there is a decrease in CO_2 emissions due to a decrease in economic activity, the demand for NEVs may decrease as well. This is because NEVs tend to be more expensive than traditional vehicles, and consumers may be less willing to pay the premium for a vehicle that is perceived as less necessary. Therefore, the GWR in terms of CO_2 emission volatility could shape NEV sales, which in turn can affect the listed companies inside the NEV industry, reflecting the NEV stock market subindex. The hydropower development, nuclear power development sector and solar photovoltaic sectors, which are highly correlated with the reduction of CO_2 emissions, could also be affected by the CO_2 emission volatility.

From Table 1, N_2O volatility also exhibits crucial effects on the stock subindices of the environmental protection sector, hydropower development sector, and nuclear power development sector. In fact, N_2O is a potent GHG that has a tremendous global warming potential vastly greater than that of CO_2. Therefore, any increase in N_2O emissions can significantly contribute to global warming and climate change. It can create severe consequences for those sectors that concentrate GHG emissions reduction. In addition, the N_2O emission variation may deteriorate the air quality. N_2O is a precursor to ozone, which is a lethal air pollutant that can cause respiratory problems and other health issues. Moreover, N_2O emissions can also erode the water quality. N_2O can be converted to nitrate, which is a water pollutant that can cause eutrophication and other environmental problems. Therefore, an increase in N_2O emissions can lead to a degradation in water quality, which can have significant impacts on aquatic ecosystems and human health. Therefore, the GWR in terms of N_2O emission volatility could impact the environment protection sector and nuclear power development sector (global warming and air quality), as well as the hydropower development sector (global warming and water quality).

It can also be observed in Table 1 that SF_6 volatility has a significant impact on the stock subindices of the environmental protection sector, hydropower development sector, and solar photovoltaic sector. Currently, most medium- and high-voltage gas-insulated switchgear, including those deployed in solar photovoltaic (PV) powering systems, still uses sulfur hexafluoride (SF_6) gas as the insulating medium. Therefore, the impact of SF_6 emission reduction on the solar PV industry is increased costs. SF_6 is a relatively inexpensive insulating gas; however, if the use of SF_6 is eliminated because of global warming issues, alternative insulating gases may need to be used, which could elevate the production cost for the whole solar PV industry. Additionally, SF_6 emission reduction in the solar PV industry could also transform the design and operation of current solar PV powering systems. If SF_6 is no longer used in switchgear, alternative technologies may need to be developed, which could impact the performance and cost of these systems and thus the whole industry. This could also be true for the hydropower development listed companies, which may also depend on the use of SF_6. Therefore, the GWR in terms of SF_6 emission volatility could generate a considerable impact on the listed companies' insider environment protection, hydropower development, and solar photovoltaic sectors.

$$\sigma_t^j = \alpha_0 + \alpha_1\sigma_t^{CO2} + \alpha_2\sigma_t^{N2O} + \alpha_3\sigma_t^{SF6} \tag{14}$$

where σ_t^j is the conditional volatility of a particular time series of stock market sector index j for time t, j includes the sector index mentioned in Section 2.1, namely, EP, HY, NEV, NP and SP.

Table 1
The impact of GHG volatility on environment protection sector index volatility, hydropower development sector index volatility, new energy vehicle sector index volatility, nuclear power development sector index volatility, solar photovoltaic sector index volatility, respectively (see Eq. 14).

GHG volatility stock subindex	σ_t^{EP}	σ_t^{HY}	σ_t^{NEV}	σ_t^{NP}	σ_t^{SP}
σ_t^{CO2}	−45.41 (101.57)	16.12** (7.89)	98.32*** (43.52)	28.24** (13.56)	31.73* (18.2
σ_t^{N2O}	40.47*** (8.69)	144.15* (85.51)	28.61 (35.44)	22.73** (11.09)	32.74 (39.26)
σ_t^{SF6}	28.73* (16.34)	30.75** (16.91)	75.94 (70.66)	40.04 (83.09)	16.32*** (7.76)
Constant	0.04 (0.016)	0.021 (0.017)	0.01 (0.006)	0.095 (0.085)	0.007 (0.007)

Notes: The robust standard errors are in parentheses, with ***, ** and *, denoting significance at 1 %, 5 % and 10 %, respectively.

3.2. Global warming risk decomposition

By using the stock market data, we have demonstrated the importance of GHG volatility on the financial markets in the previous session. In this subsection, we use the information from energy futures markets to investigate the impact of energy futures volatilities on GWR proxied by GHG volatility. We aim to constitute the GWR by four energy futures volatilities that have the most prediction power toward GWR among 12 major energy futures markets.

The primary rationale for adopting machine learning methods in our research stems from the complexity of GWR, represented by GHG emission volatility. Traditional statistical methods often depend on restrictive assumptions, such as presumed data distributions, which may not adequately capture the intricate relationships and interactions among multiple influencing factors. In contrast, machine learning techniques, such as XGB, expertise in encapsulating complex interactions and nonlinear relations among variables into the risk prediction models, making them particularly suitable for analyzing GWR through financial market data. Additionally, GP complements this approach by providing interpretable symbolic expressions, thus overcoming the common "black-box" limitations associated with various machine learning models.

Our proposed XGB-GP framework refines existing methodologies by significantly enhancing predictive accuracy, interpretability, and practical applicability. Specifically, the integration of XGB's strong feature selection capability and GP's symbolic regression approach results in improved forecasting performance, as demonstrated by the notably reduced prediction errors compared to traditional benchmark models like MIDAS. Furthermore, this combined approach provides explicit mathematical equations that clearly illustrate the relationships between energy futures market volatilities and GHG emission volatility, enhancing transparency and facilitating informed decision-making for stakeholders and policymakers. By effectively identifying the most influential energy futures markets through XGB's feature selection, our method also reduces computational complexity and improves modeling efficiency.

Therefore, our proposed XGB-GP framework offers several notable advantages over existing methods. First, it significantly improves prediction accuracy by effectively capturing complex nonlinear relationships and interactions among predictors, outperforming traditional statistical models such as MIDAS and linear regression-based approaches. Second, it provides enhanced interpretability compared to purely black-box machine learning techniques (e.g., neural networks, random forests). By integrating GP, our model generates explicit symbolic equations that clearly illustrate the relationship between energy market volatilities and CO_2 emission volatility. This transparency is particularly valuable for policymakers and practitioners. Third, the built-in feature selection capability of XGB identifies the most influential predictors, thereby reducing model complexity, improving computational efficiency, and enabling targeted monitoring of critical markets.

However, our proposed XGB-GP framework also has certain limitations. One potential disadvantage is the increased computational complexity compared to simpler statistical models, as the GP component involves evolutionary processes that require more computational resources and time. Despite these limitations, the proposed XGB-GP method demonstrates clear advantages in predictive accuracy, interpretability, and feature relevance compared to existing methods, making it a valuable tool for modeling and forecasting CO_2 emission volatility.

Therefore, we adopt the XGBoost method to aid our understanding of GWR since XGBoost can be used to identify the significance of different energy futures volatility in predicting GWR. Typically, energy futures volatility importance is measured by the total information gain and the total number of splits in the decision tree. The total information gain refers to the amount of information gained about the GWR from all energy futures volatilities, while the total split number represents how many times one energy futures volatility is split when creating the decision tree. When anode is split into two leaves in XGBoost, the gain can be defined as:

$$Gain = \frac{1}{2}\left[\frac{G_L^2}{H_L+\lambda}+\frac{G_R^2}{H_R+\lambda}-\frac{(G_L+G_R)}{H_L+H_R+\lambda}\right]-\gamma, \tag{15}$$

where G_L and G_R are the sums of gradient statistics for the left and right nodes after the split, respectively. H_L and H_R are the sums of the Hessian statistics for the left and right nodes after the split, respectively. λ is the regularization term for the leaf weights, and γ is the regularization term that controls the complexity of the tree (i.e., the number of leaves in the tree).

The total information gain across all splits of a feature used in the tree for the CO_2 volatility prediction model, N_2O volatility prediction model, and SF_6 volatility prediction model can be found in Figs. 2, 3 and 4, respectively.

In Figs. 2, 3 and 4, the total information gain is calculated based on the XGBoost algorithm. Specifically, when constructing decision trees within the XGBoost model, each node split generates a "gain" value, quantifying how effectively the split reduces the prediction error (i.e., impurity or randomness) of the model. The total information gain for each feature (energy futures volatility) is then obtained by summing the gain values across all splits where this specific feature is used. A higher total information gain implies that the feature contributes more significantly to reducing prediction errors and, consequently, plays a more important role in predicting the target variable. In this specific context, the target variable is the volatility of GHG emissions, including CO_2, N_2O, and SF_6. Therefore, features with higher total information gains are deemed more influential in explaining or predicting the fluctuations in GHG emissions.

Within our proposed XGB-GP framework, the total information gain plays a critical role in feature selection. Specifically, after training the XGBoost model, we rank the energy futures volatilities based on their total information gain to identify the most influential predictors of GHG emission volatilities. Subsequently, we select the top four energy futures volatilities with the highest information gains to serve as inputs for the GP algorithm. GP then evolves mathematical expressions that accurately capture the relationship between these selected energy futures volatilities and the targeted GHG emission volatilities.

In particular, the feature importance results depicted in Figs. 3 and 4 reveal the relative contributions of various energy futures market volatilities in predicting GHG emission volatility. Specifically, the computed importance indicates that certain energy futures markets, such as crude oil, natural gas, and coal, exhibit significantly higher predictive power compared to other markets. These results align with expectations, as these commodities are closely linked to energy production and consumption patterns, directly influencing GHG emissions. The high volatility of these markets often reflects uncertainty and rapid shifts in global energy supply and demand, making them particularly informative predictors for GHG emission volatility.

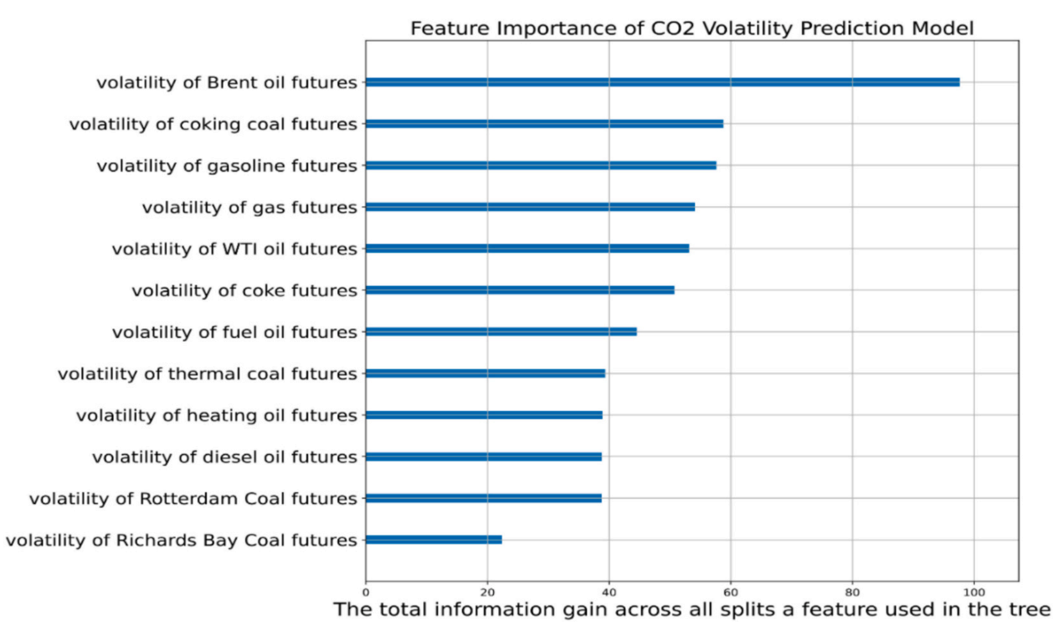

Fig. 2. Feature importance in the CO_2 volatility prediction model measured by information gain.

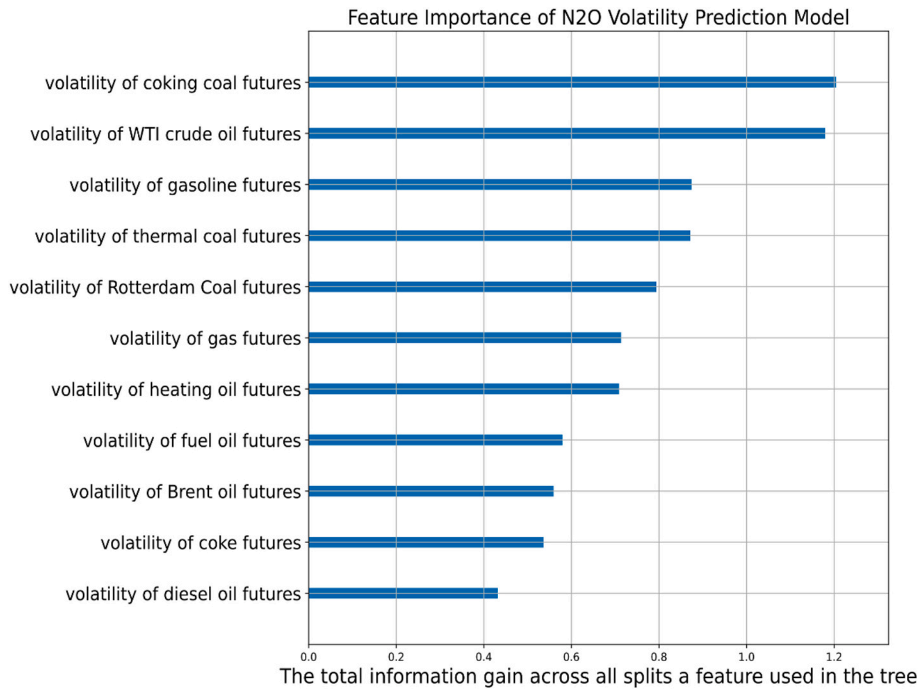

Fig. 3. Feature importance in the N_2O volatility prediction model measured by information gain.

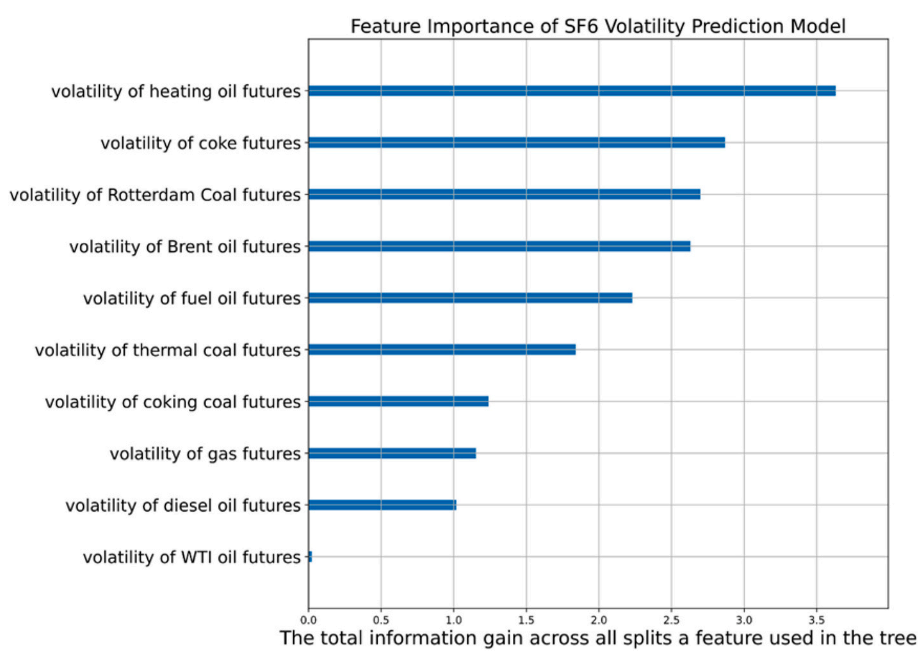

Fig. 4. Feature importance in the SF_6 volatility prediction model measured by information gain.

These computed feature importance results are valuable for subsequent predictive work in several ways. Firstly, they facilitate targeted feature selection by identifying the most influential predictors, allowing for more streamlined and efficient forecasting models. By focusing on features with higher predictive power, future forecasting models can achieve improved accuracy and reduced complexity. Secondly, understanding the relative importance of different energy futures markets provides critical insights for policymakers and industry stakeholders, enabling them to prioritize monitoring efforts and formulate targeted risk management strategies. For example, heightened attention to markets with higher predictive importance (e.g., crude oil and natural gas) can help stakeholders anticipate potential spikes in GHG emissions and proactively implement mitigation measures.

According to the aforementioned total information gain, the top four features that affect GHG volatility are presented in Table 2. Specifically, the most influential futures that can predict CO_2 emission volatility are the volatility of Brent oil futures, the volatility of coking coal futures, the volatility of gas futures, and the volatility of gasoline futures. Additionally, the most influential futures that can predict N_2O emission volatility are the volatility of coking coal futures, the volatility of gasoline futures, the volatility of thermal coal futures, and the volatility of WTI oil futures. Finally, the futures that have the highest prediction power for SF_6 emission volatility are the volatility of Brent oil futures, volatility of coke futures, volatility of heating oil futures and volatility of

Table 2
Global warming risk (i.e., CO_2, N_2O and SF_6 emission volatility) decomposition.

GWR	Features			
σ_t^{CO2}	σ_t^{bo}	σ_t^{cc}	σ_t^{gas}	σ_t^{go}
σ_t^{N20}	σ_t^{cc}	σ_t^{go}	σ_t^{tc}	σ_t^{wo}
σ_t^{SF6}	σ_t^{bo}	σ_t^{cf}	σ_t^{ho}	σ_t^{rc}

Notes: σ_t^{bo} is the volatility of Brent oil futures, σ_t^{cc} is the volatility of coking coal futures, σ_t^{cf} is the volatility of coke futures, σ_t^{gas} is the volatility of gas futures, σ_t^{go} is the volatility of gasoline futures, σ_t^{ho} is the volatility of heating oil futures, σ_t^{rc} is the volatility of Rotterdam Coal futures, σ_t^{tc} is the volatility of thermal coal futures, σ_t^{wo} is the volatility of WTI crude oil futures.

Rotterdam Coal futures. Based on the top four energy futures volatilities selected by XGBoost for three GHG XGBoost, we can construct our GWR prediction model by employing the GP method in the next subsection.

3.3. GP-based model specification

For our GP model development, we evolve the following Eq. (16) based on the top 4 features selected by XGBoost, which are volatilities from different energy futures markets:

$$f(x_1, x_2, x_3, x_4) = y_i, y_i \in \vec{y}, i \in [1, 3]. \tag{16}$$

Specifically, our GP approach consists of the following parts:

- Terminal Set: x_1, x_2, x_3, x_4 (i.e., top 4 selected features).
- Function Set: $+$, $-$, $\times$.
- Fitness measure: the error between the value of the individual function and the corresponding desired output.
- GP parameters: population = 10,000, the maximum length of the program = 1000 (i.e., up to 1000 subitems within one polynomial function), probability of crossover operation = 0.8 (i.e., 80 % of population functions will be mixed with other functions to generate new functions) and probability of mutation operation = 0.1 (i.e., 10 % of population functions will be mutated to generate new functions).
- Termination criterion: the system runs up to 100 generations.

With the settings stated in the previous section, the best functions for the three GHGs obtained by GP are:

$$\sigma_t^{CO2} = \left(\sigma_t^{bo} - \sigma_t^{go}\right) * \frac{\sigma_t^{cc} * \sigma_t^{go}}{\sigma_t^{gas}} + \left(\sigma_t^{gas}\right)^2 * \frac{\sigma_t^{bo} + \sigma_t^{cc}}{\sigma_t^{gas} + \sigma_t^{go} - \sigma_t^{cc}}, \tag{17}$$

$$\sigma_t^{N2O} = \left(\sigma_t^{cc} - \sigma_t^{go}\right) * \frac{\sigma_t^{cc} * \sigma_t^{wo}}{\sigma_t^{tc}} + \sigma_t^{tc} * \frac{\sigma_t^{go} + \left(\sigma_t^{wo}\right)^2}{\sigma_t^{cc}}, \tag{18}$$

$$\sigma_t^{SF6} = \sigma_t^{bo} * \frac{\sigma_t^{rc}}{\sigma_t^{ho} - \sigma_t^{cf}} + \left(\sigma_t^{cf}\right)^2 * \frac{\sigma_t^{ho} + \sigma_t^{rc}}{\sigma_t^{bo}}. \tag{19}$$

Therefore, our XGB-GP framework has created three GHG emission volatility forecasting models, namely, Eqs. (17)–(19) for CO_2 emission volatility, N_2O emission volatility, and SF_6 emission volatility, respectively. It is observable for all models that the GP has evolved into three models with two major terms. The first term is the volatility information extracted from vital energy futures markets selected by XGB. The second term is the variance information, which can serve as the second level of magnitude since the variance term (volatility squared term) is considerably smaller than volatility. This variance information can be used to adjust the first volatility term in a subtle way to capture the infinitesimal movement of GHG emission volatility. In the next subsection, we evaluate the model performance compared with the MIDAS method using mean absolute error (MAE) and Root Mean Square Error (RMSE) loss functions.

3.4. Model performance evaluation with benchmark model

In order to evaluate our XGB-GP model performance, mean absolute error (MAE) and Root Mean Square Error (RMSE) were employed to obtain the volatility forecasting error generated by the models (Bollerslev et al., 2016; Somu et al., 2020). We find the periodic averaged MAE as follows:

$$MAE_T = \frac{1}{T}\sum_{t=1}^{T} |Observed_t - Predicted_t|,$$

where T is total observations during the forecasting period, $Observed_t$ is the observed energy futures volatility obtained from the corresponding energy futures markets and $Predicted_t$ is the energy futures volatility predicted from the corresponding models.

Furthermore, we obtain the periodic averaged RMSE as follows:

$$RMSE_T = \sqrt{\frac{1}{T}\sum_{t=1}^{T} (Observed_t - Predicted_t)^2}.$$

Lower MAE and RMSE imply higher forecasting accuracy with more prediction power toward GWR, and we use MIDAS as our benchmark model. As exhibited in Table 3, the errors in predicting using our GP method are significantly lower than those predicted by the MIDAS method. Using the MAE loss function, our XGB-GP method has an overall prediction error of 0.49 %, 2.21 %, and 0.82 % for CO_2, N_2O, and SF_6 volatility forecasting, respectively. In contrast, the traditional MIDAS method has overall prediction errors of 2.74 %, 4.75 %, and 3.34 %. It is clear that our XGB-GP method overwhelmingly surpasses the MIDAS method in terms of prediction accuracy. On the other hand, for the RMSE loss function, our XGB-GP method's prediction errors are 0.65 %, 3.08 %, and 1.02 % for CO_2, N_2O, and SF_6 volatility forecasting, respectively.

The prediction accuracy for our XGB-GP method is also considerably higher than the MIDAS method exhibited in Table 3 regarding the RMSE loss function. Therefore, our proposed XGB-GP method provides a more accurate prediction of GWR and thus can be adopted as the fundamental model for constructing the portfolio to neutralize the GWR.

As a result, our empirical results clearly demonstrate that the proposed XGB-GP model significantly outperforms the traditional MIDAS method in forecasting GHG emission volatility regarding the forecasting accuracy. Specifically, our XGB-GP method achieves substantially lower forecasting errors, as indicated by both MAE and RMSE, compared to the MIDAS benchmark. These findings underscore the superior predictive capability of our methodology, which effectively captures the nonlinear and interactive relationships among energy futures volatilities and GHG emissions.

In terms of model stability, our proposed framework leverages the robustness of XGB, which incorporates built-in regularization mechanisms to mitigate overfitting, thus enhancing stability and generalizability. Additionally, GP contributes to model stability by evolving symbolic expressions through multiple generations, systematically refining solutions to achieve stable and interpretable models. Moreover, compared with other commonly used forecasting methods, such as traditional regression-based models and neural network-based approaches, our XGB-GP framework not only achieves higher predictive

Table 3
Model prediction accuracy for XGB-GP and MIDAS comparison using MAE and RMSE.

Forecasting error	MAE		RMSE	
	XGB-GP	MIDAS	XGB-GP	MIDAS
CO_2 volatility	0.49 %	2.74 %	0.65 %	3.17 %
N_2O volatility	2.21 %	4.75 %	3.08 %	5.91 %
SF_6 volatility	0.82 %	3.34 %	1.02 %	4.86 %

accuracy but also offers greater transparency and interpretability through explicit mathematical expressions. This interpretability is particularly valuable for policymakers and industry stakeholders, as it provides clear insights into the underlying factors driving GHG emission volatility.

3.5. Hedging portfolio construction

Based on the results from previous subsections, we construct hedging portfolios to neutralize the GWR in terms of GHG volatilities in this subsection. According to Section 2.6, we employ the neural network to identify the weights for each futures to hedge against the GWR extracted from Table 2. Furthermore, in Table 4, we outline three GHG volatility hedging portfolios. In particular, for hedging one unit of CO_2 volatility, we need to shorten 1.381 units of Brent oil futures, 0.272 units of gas futures, 0.138 units of gasoline futures and 0.791 units of coking coal futures. Likewise, to hedge one unit of N_2O volatility, we need to shorten coking coal futures by 0.671 units, gasoline futures by 0.971 units, thermal coal futures by 0.798 units and WTI crude oil futures by 1.44 units. Finally, in particular, for hedging one unit of SF_6 volatility, we need to short 0.0551 units of Brent oil futures and 1.06 units of coke futures and long 0.101 heating oil futures and 0.0141 units of Rotterdam Coal futures. All hedging portfolios can be used to neutralize GHG volatilities, relieving firms from GWR.

4. Discussions

Compared with prior studies that prevailing associate environmental factors with climate risk (Park and Lee, 2020; Ren et al., 2022), our study highlights the use of energy futures data to deconstruct GWR, emphasizing the dependencies of GWR on financial market fluctuations. Unlike traditional environmental risk assessments, our study reveals energy futures volatilities as leading factors influencing GHG emission fluctuations, leveraging forward-looking pricing mechanisms that integrate geopolitical, economic, and policy-driven expectations (Yalew et al., 2020). This contrasts with studies attributing emission volatility solely to consumption patterns (Mideksa and Kallbekken, 2010), rather, we position energy market volatilities as critical mediators. Furthermore, the proposed hedging strategy extends beyond traditional sectoral analyses by operationalizing energy futures to manage climate risk exposure, supplementing existing static ESG frameworks with dynamic analysis of climate risk (Eskantar et al., 2024). Methodically, compared to existing machine learning models (Guo et al., 2023; Wang et al., 2022), which focus on general predictive accuracy, our XGB-GP method specifically targets on GWR forecasting with explicit model formulation, suggesting actionable insights for regulatory compliance and low-carbon transitions.

Linking with existing literature of XGB and big data analytics studies, which verify that the XGB method can provide more accurate forecasting models (Guo et al., 2023; Wang et al., 2022). Our XGB-GP method also proposes a more accurate GWR prediction model in terms of GHG emission volatility. Accurate forecasting is valuable for firms to respond to the negative impact of global warming. As the global warming issue has created major concerns for policymakers in the global context, governments worldwide are adopting stringent climate policies and regulations to curb greenhouse gas emissions and promote a low-carbon economy. Our GHG emission volatility forecasting model can enable firms to anticipate regulatory changes based on precise GHG emission volatility and anticipate adapting to policy changes. This facilitates compliance, minimizing costs of noncompliance such as carbon emission penalties.

Unlike the current research focusing on the environmental factors to predict climate risk (Park and Lee, 2020; Ren et al., 2022), our study leverages energy futures market data to predict climate risk, which illustrates the intricate relationships between climate risk and financial factors. Our results thereby show that energy futures volatilities can be used as prevailing predictors of GHG emission volatilities. It can be attributed that energy futures prices encapsulate expectations about future energy supply and demand since they are forward-looking future prices for energies. These expectations are shaped by a variety of factors, including economic outlook, geopolitical events and environmental policies (Yalew et al., 2020). Remarkably, concerns about global warming and the potential for increased GHG emissions are also embedded in these expectations. Consequently, energy futures prices and their volatilities can signal the market's anticipation of these climate-related risks and the potential for increased GHG emissions. More importantly, changes in energy prices can have profound effects on energy consumption patterns (Mideksa and Kallbekken, 2010). As energy prices fluctuate due to supply and demand imbalances, which are influenced by climate change impacts, consumers and industries may adjust their behavior to be more energy-efficient or to switch to alternative energy sources. These adjustments in energy consumption patterns can, in turn, feedback into GHG emissions, contributing to their volatilities as a result.

Based on energy market selections by XGB, our paper intends to formulate portfolios to hedge against GWR for all three GHG emission volatilities. Our proposed portfolio to hedge against GWR could reduce the financial risks associated with climate change. As global warming continues, major weather extremes will become more frequent and intense. These events could result in supply chain disruptions and increased insurance premiums. Our portfolio can hedge GWR according to GHG emission volatilities, and when GHG emissions become volatile, the probability of weather extremes occurrence increases. As a result, firms can thereby use our portfolio to hedge against high GHG emission volatilities.

On this basis, we reveal that CO_2 emission volatility has a significant impact on the values of those firms inside the hydropower development sector, new energy vehicle sector, nuclear power development sector, and solar photovoltaic sector. N_2O emission volatility has a significant impact on the values of firms inside the environmental protection sector, hydropower development sector, and nuclear power development sector. The SF_6 emission volatility has a significant impact on the values of firms within the environmental protection sector, hydropower development sector, and solar photovoltaic sector. As a result, we demonstrate that GHG emission volatility has a close nexus with firm values in relevant industries. Therefore, our paper has several salient implications. First, it is evident that GWR has a substantial impact on stock markets and listed company values. The impact of global warming is reflected in the financial performance of firms. The negative impact of global warming on listed companies' earnings has been well documented (Huang et al., 2018). Shareholders, investors, and other stakeholders of listed companies are increasingly demanding that companies take responsibility for their environmental impact. Firms that understand and hedge global warming risks demonstrate their commitment to sustainability, bolstering their reputation and strengthening stakeholder relations. Therefore, companies and investors need to learn how to manage their exposure to GWR stemming from climate change as

Table 4
CO_2, N_2O and SF_6 emission volatility hedging coefficients.

σ_t^{CO2}	HR	σ_t^{N20}	HR	σ_t^{SF6}	HR
σ_t^{bo}	−1.381	σ_t^{cc}	−0.671	σ_t^{bo}	−0.0551
σ_t^{cc}	0.791	σ_t^{go}	−0.971	σ_t^{cf}	−1.060
σ_t^{gas}	−0.272	σ_t^{tc}	−0.798	σ_t^{ho}	0.101
σ_t^{go}	−0.138	σ_t^{wo}	1.440	σ_t^{rc}	0.0141

Notes: σ_t^{bo} is the volatility of Brent oil futures, σ_t^{cc} is the volatility of coking coal futures, σ_t^{co} is the volatility of coke futures, σ_t^{gas} is the volatility of gas futures, σ_t^{go} is the volatility of gasoline futures, σ_t^{ho} is the volatility of heating oil futures, σ_t^{rc} is the volatility of Rotterdam Coal futures, σ_t^{tc} is the volatility of thermal coal futures, and σ_t^{wo} is the volatility of WTI crude oil futures. HR is the hedging ratio.

financial management increasingly focuses on ESG (environmental, social, and governance) considerations.

We further unveil the response of energy futures markets toward GWR. Energy markets have substantial carbon footprints, while commodity futures markets are closely related to respective industries. The energy industry, facing decarbonization, is expected to experience significant changes in the coming years, with major implications for energy futures markets. Global demand for energy continues to rise, and an increasing share of that energy demand is being supplied by renewable sources such as solar and hydropower.

Intuitively, future non-renewable energy scarcity is already being priced, with the cost of renewables (fossil fuels) expected to fall (rise). Investors will be expected to become increasingly cautious of high-carbon-use assets, diversifying portfolios across lower-carbon-intensity investments. Our study decomposes GWR from the energy futures market to highlight a hedging portfolio within the energy futures market that has flexible investment opportunities.

Further, we disentangle the global warming risk from the energy market perspective toward creating an accurate GWR prediction model. Our decomposition of GWR from energy markets enables firms to make informed decisions about investments, resource allocation, and business strategies. The energy market, being a primary source of GHG emissions, is inherently linked to GWR. The volatility in energy futures can signify potential shifts in the market's response to climate change policies and changing environmental conditions. By decomposing GWR from energy markets, we unravel the underlying factors contributing to these volatilities, thereby enabling firms to make more informed decisions regarding investments, resource allocation, and business strategies. It could thereby empower companies to identify new opportunities related to the GWR, mitigate potential threats created by GWR, and allocate resources more effectively.

5. Conclusions

Considering that global warming is causing increasing concerns among investors and policy-makers worldwide, we disentangle and forecast GWR with energy futures volatilities. We first claim that GWR has a considerable effect on listed company values by impacting stock markets. We subsequently derive the GWR forecasting model from energy futures using an extreme gradient boosting (XGB)-genetic programming (GP) framework. The XGB-GP framework creates volatility forecasting models for GWR proxied by GHG volatilities by selecting the major contributors to global warming from energy futures markets.

XGB methodology accommodates the main constituents of CO_2, N_2O, and SF_6 emission volatilities within 12 major energy futures. Creating a GP algorithm for these three volatility forecasts, we demonstrate that our GP models are significantly more accurate than the MIDAS method in predicting greenhouse gas emission volatilities. Additionally, based on a decomposition of global warming risk by XGB-GP, we propose risk hedging portfolios using energy futures to mitigate risk to investors.

As our paper indicates, the use of energy futures volatilities as predictors of GHG emission volatilities emphasizes the role of energy markets in GWR assessment. Compared with existing studies, we demonstrate that the energy markets are closely intertwined with greenhouse gas emissions, and thus understanding these volatilities can provide further deconstruction of GWR evolving in the long-run. The long-term and dynamic analysis offers a more comprehensive framework on GWR, which can aid the understanding in identifying trends and moving patterns of GWR. Further, by putting different energy markets within a single prediction framework of GWR, our paper also delivers the understanding of cross-market connection among different energy markets.

Moreover, another implication of our study is its potential to help investors to design hedging strategies in energy futures markets. Accurate forecasts of GHG emission volatilities are essential for designing effective hedging strategies in energy markets, especially for those investors increasingly exposed to the risks associated with climate changes. They can thereby apply our XGB-GP model to develop more effective hedging strategies against GWR. Given the growing exposure of investors to climate related risks, accurate forecasts of GHG emission volatilities play vital role in designing effective hedging strategies. Our XGB - GP model, provides superior model performance in GWR forecasting, which can be further leveraged by investors. They can employ the improved forecasts to develop hedging strategies in energy futures markets. These strategies can be helpful to mitigate the risks related to GHG emission volatilities by taking positions in energy futures. This paper thereby bridges a critical gap in climate risk management, as existing studies often underestimate the impact of GWR as there are lacking of effective hedging tools.

Additionally, our study also illuminates the interactions between energy markets and climate change in policy formulation. Our study has emphasized the necessity for hedging GWR and other climate related risk, which sheds the insights for proposing financial instrument regarding the GWR. Policymakers could thereby collaborate with financial institutions to develop standardized financial instruments for investors to hedge the GWR and other climate related risk. Because developing standardized hedging instruments relies on the predictability of GWR and other climate related risk. Our developed model that has improved the predictability of GWR largely facilitates such instrument development and thus delivers the policy feasibility. On the basis, creating standardized hedging financial instruments based on energy futures under our framework, policymakers can provide a more structured and accessible way for investors and companies to manage climate-related risks more effectively.

Finally, the findings of this study offer valuable insights for the design of emissions trading systems (ETS), which are steadily being used by governments to reduce GHG emissions. In an ETS, companies can acquire a predetermined quantity of emission allowances, which can be traded in the open market. The price of these allowances can be influenced by energy prices and GWR. Understanding the energy futures volatilities with emission volatilities, this study provides policymakers with a valuable tool for forecasting the future price of emission allowances and adjusting the supply of allowances as the futures markets move.

Although our study adopts the novel machine learning approach to scrutinize the global warming risk by employing energy futures data, it still suffers from research limitations. The main constraint of our paper is considering the futures markets as perfect markets where the market liquidity risk and transaction costs have been overlooked. Market liquidity risk can be manifested during episodes of geopolitical events, such as Russian-Ukraine conflict. Another limitation is that we focus on the financial perspectives of the climate risks where other policy related issues have not been taken into account, such as the climate policy uncertainty.

As a result, future research could extend this paper from those two aspects. First, it is essential to explore the role of futures market trading frictions such as transaction costs and liquidity risk in shaping hedging effectiveness for investors. Moreover, encompassing the impact of climate policy related data such as climate policy uncertainty index to could further strengthen GWR analysis and prediction. Future research that incorporates market trading friction factors as well as the climate policy uncertainty, could strengthen the effectiveness of hedging strategies for investors and enlarge policymakers' capacity to design resilient ETS architecture and risk management framework accommodating to both market shocks and evolving climate policies.

Funding information

This paper is supported by The Philosophy and Social Science Planned Research Project of Zhejiang Province, "Study of Intelligentized Prediction and Management of Small and Medium-sized Enterprises' Credit Risk in Zhejiang Province" (grant number: 24NDJC240YBM), The

Ningbo Natural Science Foundation supported project (Project ID: 2023J194), and The Ningbo Natural Science Foundation supported project (Project ID: 2024J126).

CRediT authorship contribution statement

Shusheng Ding: Conceptualization, Writing – original draft. **Tianxiang Cui:** Data curation, Software, Methodology, Writing – original draft. **Anna Min Du:** Investigation, Project administration, Supervision, Validation, Writing – review & editing. **John W. Goodell:** Investigation, Writing – review & editing. **Nanjiang Du:** Methodology, Formal analysis.

Declaration of competing interest

The authors report no conflicts of interest in this work. The sponsors had no role in the design of the study, in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

Data availability

Data will be made available on request.

References

Agnew, D.C., 2024. A global timekeeping problem postponed by global warming. Nature 628 (8007), 333–336.

Allen, M.R., Frame, D.J., Huntingford, C., Jones, C.D., Lowe, J.A., Meinshausen, M., Meinshausen, N., 2009. Warming caused by cumulative carbon emissions towards the trillionth tonne. Nature 458 (7242), 1163–1166.

Amicarelli, V., Lagioia, G., Bux, C., 2021. Global warming potential of food waste through the life cycle assessment: An analytical review. Environmental Impact Assessment Review 91, 106677.

Bartram, S.M., Hou, K., Kim, S., 2022. Real effects of climate policy: financial constraints and spillovers. J. Financ. Econ. 143 (2), 668–696.

Battaglia, G., Joos, F., 2018. Marine N_2O emissions from nitrification and denitrification constrained by modern observations and projected in multimillennial global warming simulations. Glob. Biogeochem. Cycles 32 (1), 92–121.

Bollerslev, T., Patton, A.J., Quaedvlieg, R., 2016. Exploiting the errors: a simple approach for improved volatility forecasting. J. Econ. 192 (1), 1–18.

Bose, S., Shams, S., Ali, S., AlMamun, A., Chang, M., 2024. Economic policy uncertainty, carbon emissions and firm valuation: international evidence. Br. Account. Rev. 16 (6), 101453.

Bovari, E., Giraud, G., Mc Isaac, F., 2018. Coping with collapse: a stock-flow consistent monetary macrodynamics of global warming. Ecol. Econ. 147, 383–398.

Brehm, P., 2019. Natural gas prices, electric generation investment, and greenhouse gas emissions. Resour. Energy Econ. 58, 101106.

Chen, T., Guestrin, C., 2016. Xgboost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794.

Chen, L., Zhong, X., Li, H., Wu, J., Lu, B., Chen, D., Xie, S.-P., Wu, L., Chao, Q., Lin, C., et al., 2024. A machine learning model that outperforms conventional global subseasonal forecast models. Nat. Commun. 15 (1), 6425.

Cooper, S.A., Raman, K., Yin, J., 2018. Halo effect or fallen angel effect? Firm value consequences of greenhouse gas emissions and reputation for corporate social responsibility. J. Account. Public Policy 37 (3), 226–240.

Cruz, J.L., Rossi-Hansberg, E., 2024. The economic geography of global warming. Rev. Econ. Stud. 91 (2), 899–939.

Ding, S., Zhang, Y., Duygun, M., 2019. Modeling price volatility based on a genetic programming approach. Br. J. Manag. 30 (2), 328–340.

Escobar-Anel, M., Hou, Y., Stentoft, L., 2025. The shifted GARCH model with affine variance: applications in pricing. Financ. Res. Lett. 71, 106371.

Eskantar, M., Zopounidis, C., Doumpos, M., Galariotis, E., Guesmi, K., 2024. Navigating ESG complexity: an in-depth analysis of sustainability criteria, frameworks, and impact assessment. Int. Rev. Financ. Anal. 95, 103380.

Franta, B., 2018. Early oil industry knowledge of CO_2 and global warming. Nat. Clim. Chang. 8 (12), 1024–1025.

Fu, T., Huang, D., Feng, L., Tang, X., 2024. More is better? The impact of predictor choice on the INE oil futures volatility forecasting. Energy Econ. 134, 107540.

Geiger, T., Gütschow, J., Bresch, D.N., Emanuel, K., Frieler, K., 2021. Double benefit of limiting global warming for tropical cyclone exposure. Nat. Clim. Chang. 11 (10), 861–866.

Ghysels, E., Sinko, A., Valkanov, R., 2007. Midas regressions: further results and new directions. Econ. Rev. 26 (1), 53–90.

Gillingham, K., Stock, J.H., 2018. The cost of reducing greenhouse gas emissions. J. Econ. Perspect. 32 (4), 53–72.

Gong, X., Ye, X., Zhang, W., Zhang, Y., 2023. Predicting energy futures high-frequency volatility using technical indicators: the role of interaction. Energy Econ. 119, 106533.

Guiot, J., Cramer, W., 2016. Climate change: the 2015 Paris agreement thresholds and mediterranean basin ecosystems. Science 354 (6311), 465–468.

Guo, L.-N., She, C., Kong, D.-B., Yan, S.-L., Xu, Y.-P., Khayatnezhad, M., Gholinia, F., 2021. Prediction of the effects of climate change on hydroelectric generation, electricity demand, and emissions of greenhouse gases under climatic scenarios and optimized ANN model. Energy Rep. 7, 5431–5445.

Guo, X., Gui, X., Xiong, H., Hu, X., Li, Y., Cui, H., Qiu, Y., Ma, C., 2023. Critical role of climate factors for groundwater potential mapping in arid regions: insights from random forest, XGBoost, and LightGBM algorithms. J. Hydrol. 621, 129599.

Hong, H., Karolyi, G.A., Scheinkman, J.A., 2020. Climate finance. Rev. Financ. Stud. 33 (3), 1011–1023.

Hoque, M.E., Soo-Wah, L., Billah, M., 2023. Time-frequency connectedness and spillover among carbon, climate, and energy futures: determinants and portfolio risk management implications. Energy Econ. 127, 107034.

Huang, H.H., Kerstein, J., Wang, C., 2018. The impact of climate risk on firm performance and financing choices: an international comparison. J. Int. Bus. Stud. 49, 633–656.

Jin, J., Cui, T., Bai, R., Qu, R., 2024. Container port truck dispatching optimization using real2sim based deep reinforcement learning. Eur. J. Oper. Res. 315 (1), 161–175.

Lamperti, F., Bosetti, V., Roventini, A., Tavoni, M., 2019. The public costs of climate-induced financial instability. Nat. Clim. Chang. 9 (11), 829–833.

Lan, X., Thoning, K.W., Dlugokencky, E.J., 2023. Trends in Globally-Averaged Ch4, N_2O, and SF_6 Determined from Noaa Global Monitoring Laboratory Measurements. NOAA Global Monitoring Laboratory. Accessed: 2023-04.

Lashof, D.A., Ahuja, D.R., 1990. Relative contributions of greenhouse gas emissions to global warming. Nature 344 (6266), 529–531.

Lee, J.H., Shi, Z., Gao, Z., 2022. On LASSO for predictive regression. J. Econ. 229 (2), 322–349.

Li, Y., Zhang, X., Xiao, S., Chen, Q., Wang, D., 2018. Decomposition characteristics of c5f10o/air mixture as substitutes for SF_6 to reduce global warming. J. Fluor. Chem. 208, 65–72.

Li, H., Huang, X., Luo, F., Zhou, D., Cao, A., Guo, L., 2025. Revolutionizing agricultural stock volatility forecasting: a comparative study of machine learning and HAR-RV models. J. Appl. Econ. 28 (1), 2454081.

Lichtfouse, E., Lichtfouse, M., Jaffrézic, A., 2003. $\delta^{13}C$ values of grasses as a novel indicator of pollution by fossil-fuel-derived greenhouse gas CO_2 in urban areas. Environ. Sci. Technol. 37 (1), 87–89.

Liu, J., Ma, F., Yang, K., Zhang, Y., 2018. Forecasting the oil futures price volatility: large jumps and small jumps. Energy Econ. 72, 321–330.

Matthews, H.D., Gillett, N.P., Stott, P.A., Zickfeld, K., 2009. The proportionality of global warming to cumulative carbon emissions. Nature 459 (7248), 829–832.

Mideksa, T.K., Kallbekken, S., 2010. The impact of climate change on the electricity market: a review. Energy Policy 38 (7), 3579–3585.

Naysary, B., Shrestha, K., 2024. Financial technology and ESG market: a wavelet-DCC GARCH approach. Res. Int. Bus. Financ. 71, 102466.

Nishitani, K., Kokubu, K., 2012. Why does the reduction of greenhouse gas emissions enhance firm value? The case of japanese manufacturing firms. Bus. Strateg. Environ. 21 (8), 517–529.

Park, S.-J., Lee, D.-K., 2020. Prediction of coastal flooding risk under climate change impacts in South Korea using machine learning algorithms. Environ. Res. Lett. 15 (9), 094052.

Pham, L., Hay, D., Miihkinen, A., Myllymäki, E.-R., Niemi, L., Sihvonen, J., 2024. Climate Risk Disclosures and Auditor Expertise. Available at SSRN 4766071.

Poli, R., Langdon, B., McPhee, N., 2008. A Field Guide to Genetic Programming. Lulu Enterprises, UK Ltd.

Ren, X., Li, Y., Shahbaz, M., Dong, K., Lu, Z., 2022. Climate risk and corporate environmental performance: empirical evidence from China. Sustain. Prod. Consum. 30, 467–477.

Sautner, Z., Van Lent, L., Vilkov, G., Zhang, R., 2023. Firm-level climate change exposure. J. Financ. 78 (3), 1449–1498.

Shine, K.P., Fuglestvedt, J.S., Hailemariam, K., Stuber, N., 2005. Alternatives to the global warming potential for comparing climate impacts of emissions of greenhouse gases. Clim. Chang. 68 (3), 281–302.

Somu, N., MR, G. R, Ramamritham, K., 2020. A hybrid model for building energy consumption forecasting using long short term memory networks. Appl. Energy 261, 114131.

Stroebel, J., Wurgler, J., 2021. What do you think about climate finance? J. Financ. Econ. 142 (2), 487–498.

Wang, C.-C., Kuo, P.-H., Chen, G.-Y., 2022. Machine learning prediction of turning precision using optimized xgboost model. Appl. Sci. 12 (15), 7739.

Wu, W., Zhang, N., Li, A., Chen, Y., 2024. The path to global climate justice: from the perspective of regional discrepancy in embodied carbon emissions. Environ. Impact Assess. Rev. 105, 107410.

Yalew, S.G., van Vliet, M.T., Gernaat, D.E., Ludwig, F., Miara, A., Park, C., Byers, E., De Cian, E., Piontek, F., Iyer, G., et al., 2020. Impacts of climate change on energy systems in global and regional scenarios. Nat. Energy 5 (10), 794–802.

Zhang, J., Guo, Y., Han, J., Ji, Y., Zhang, L., 2021a. Greenhouse gas emissions and net global warming potential of vineyards under different fertilizer and water managements in North China. Agric. Water Manag. 243, 106521.

Zhang, W., Yu, J., Zhao, A., Zhou, X., 2021b. Predictive model of cooling load for ice storage air-conditioning system by using GBDT. Energy Rep. 7, 1588–1597.
Zhang, J., Shen, J., Xu, L., Zhang, Q., 2023. The CO_2 emission reduction path towards carbon neutrality in the Chinese steel industry: a review. Environ. Impact Assess. Rev. 99, 107017.
Zhang, L., Chen, Y., Bouri, E., 2024. Time-varying jump intensity and volatility forecasting of crude oil returns. Energy Econ. 129, 107236.