Computational Economics (2023) 61:775–805
https://doi.org/10.1007/s10614-021-10231-5

CO_2 Emission Allowances Risk Prediction with GAS and GARCH Models

Nader Trabelsi[1,2] · **Aviral Kumar Tiwari**[3]

Accepted: 27 December 2021 / Published online: 16 January 2022

Abstract

We analyse the predictive and the forecasting ability of various Generalized Autoregressive Score (GAS) and GARCH frameworks for European Union Allowances (EUAs) daily returns (EUAs returns) for the period 22/04/2005–28/02/2019. We further examine the impact of different distributional assumptions on risk prediction. The Model Confidence Set (MCS) is employed to compare and select a superior predictive model of Value-at-Risk (VaR) thresholds. We find that GAS under skewed t-student error distribution and gjr-GARCH under general error distribution deliver excellent results for the Value-at-Risk (VaR) prediction for EUA at 1% and 5% levels, respectively. These results are robust with respect to three back-testing procedures (i.e., Unconditional Coverage, Conditional Coverage, and Dynamic Quantile tests). These results are of particular importance for the development of EUA pricing policies and risk management strategies.

Keywords CO_2 emission allowances · GARCH · GAS · MCS procedure · VaR

1 Introduction

The European Union (EU) Emissions Trading System (ETS) is the world's largest organized financial market for CO_2 Emission Allowances. In ETS, polluters can trade enough emission rights with another one who may have a surplus of them. Much like any market, the EUA price, at any given point in time, will result in supply and

✉ Nader Trabelsi
nhtrabelsi@imamu.edu.sa

Aviral Kumar Tiwari
aviral.eco@gmail.com

1 Imam Mohammad Ibn Saud Islamic University (IMSIU), PO Box 5701, Riyadh, Saudi Arabia

2 LARTIGE ASTURIMA, University of Kairouan, Kairouan, Tunisia

3 Indian Institute of Management Bodh Gaya, Bodh Gaya, India

demand conditions. As this market works on the 'cap and trade' programs, the supply or the cap is pre-determined for 31 countries, but the allowance demand from energy-intensive installations across the EU (i.e., combustion plants, coke ovens, oil refineries, iron and steel plants, and factories making cement, brick, ceramics, glass, lime, pulp and paper) is rather flexible and can hardly be predicted. Thus, although the carbon market has been growing extremely fast, there is a consensus in the literature that there is an ambiguous picture of the EUA price formation.[1] Precisely, the EUA price determination and their "forecastability" are still open questions that continue to attract increasing interest from researchers. Indeed, a better understanding of the price formation of EUAs will help companies, among other things, to assess their business future risk, and policymakers to evaluate the performance of each national allocation plan to achieve compliance with their commitments under the Kyoto Protocol. In this sense, Benz and Trück (2006) describe the EUAs as a new class of assets with distinct characteristics, which demand efficient techniques to produce accuracy price forecasts. According to Mansanet-Bataller et al. (2007), Alberola et al., (2008a, 2008b), Chevallier (2009), Hintermann (2010), Hitzemann and Uhrig-Homburg (2013), among others, the long-term EUAs spot prices depend on several determinants such as the price of traditional energy products, weather conditions, and economic fluctuations. For Nazifi and Milunivich (2010), there are only significant short-run linkages between the price of EUAs and traditional energy prices (i.e., coal, oil, natural gas, and electricity). On the other hand, Wang et al. (2018) find significant time-varying spillover effects between returns and volatility of energy and EUA future markets. Moreover, Carnero et al. (2018) show that EUAs price changes can be explained by the dynamics of fuel prices.[2]

While there are several studies performing multiple regression analysis to find how the EUAs prices are influenced by various independent variables, there are few other studies that go beyond the financial econometric analysis (Daskalakis et al., 2009; Benz & Trück, 2009; Hitzemann and Uhrig-Homburg, 2013; Gil-Alana, 2016; Müller et al., 1997). Their results also demonstrate the capability of models that capture the dynamic features of the financial series, such as volatility clustering, skewness and excess kurtosis. However, Paolella and Taschini (2008) believe we cannot predict carbon emission permits under EU ETS accurately. This paper adds to this literature by providing a statistical procedure that delivers the "best" model to forecast EUA's price change. This issue is of paramount importance because a "poor" model would significantly affect the risk management strategies of a range of industrial companies, being direct CO_2 emitters or consumers, as well as players in financial market including hedgers and portfolio managers. In fact, the participants in the EUA market have increased in number, and previous studies provide evidence for its utility for hedging portfolio risk (Zhang & Sun, 2016, among others).

[1] At the moment, the EU ETS is in Phase 3 (2013–2020), which is significantly different from Phases 1 (2005–2007) and II (2008–2012).

[2] The factors that can influence the EUA prices are: (i) Changes in regulations as they can essentially change the supply of the EUAs. (ii) The influence of other energy commodity prices. (iii) Macroeconomic and financial markets indicators. (iv) The financial markets as when there is a negative shock in the financial markets, the financial situations of individual firms will be affected; meanwhile, the economy may deteriorate, which in turn shrinks the emissions demand.

Clearly, this study offers support for the following idea: GAS model of Creal et al. (2013) and Harvey (2013) is the most appropriate model to use when one has to evaluate the volatility of EUA returns. The appropriateness of this model can be seen through its mathematical flexibility in solving the complicated dynamics characteristics of times series (i.e. asymmetry, long memory, etc.). Particularly, this requires to computing a score among different scales rather than solving a number of moment equations.

The second contribution of this study is to evaluate the impact of different error distributional approximations on the accuracy of the predictive and forecasting results. Precisely, we assume that innovations are normal (norm), std (t-student), sstd (skewed t-student), Asymmetric *t*-Student with two decay parameters (ast) or Asymmetric *t*-Student with a left-tail decay parameter (ast1). To the best of the authors' knowledge, this will be the first study that checks the appropriateness of GAS model and these distribution assumptions to predict EUA's price changes.

The third contribution of this study is to focus on the comparison of GAS model with multiple General Autoregressive Conditional Heteroskedasticity (GARCH) models. For GARCH models, innovations are assumed to follow: Normal (Norm), t-Student (St), Skewed-t-Student (Sts), Generalized Error Distribution (GED) or Johnson's Reparametrized SU (JSU) distributions. Such models are useful as they demonstrate their ability to capture several stylised facts of the financial time series. Moreover, as they have been applied successfully in several areas (i.e., default and credit risk modelling, stock volatility, and correlation modelling, modelling time-varying dependence structures, CDS spread modelling and questions relating to financial stability and systemic risk, modelling high-frequency data, etc.), It's worth noting that by scaling the score function appropriately, standard observation-driven models such as the GARCH models can be recovered. For that, only non-Gaussian-GARCH models are discussed in our analysis. Comparison between general GARCH models and GAS model is made so that researchers and policy makers are in better position to use evaluate the difference when traditional models are used as opposed to the advanced models. In addition to that to choose the best model we used three back-testing procedures namely Unconditional Coverage, Conditional Coverage, and Dynamic Quantile tests.

In order to choose the best model among GARCH and GAS models, we implement several model selection criteria. More precisely, the best-fitted model with maximum likelihood (ML) will be validated by the minimum AIC (Akaike Information Criteria) and BIC (Bayesian Information Criteria) criterion.

Our research is also centred on the prediction of the EUA extreme volatility using the Value-at-Risk (VaR) measure. To this end, we use the MCS approach to perform a comparison of a number of VaR forecasts. The MCS approach consists of a sequence of tests that permits to construct of a set of "superior" models. In other terms, GAS and GARCH models are compared in terms of their predictive ability, so that models that produce better forecasts of EUA's VaR will have the highest ranks. Specifically, we assess whether GAS model leads to more accurate volatility forecasts of EUA return than any other alternative models. Contrary to other researchers, we do not use distinct phases (i.e., phase I, II and III). Note that the CO_2 allowance trading periods of the EU ETS are usually referred to as "phases". Our empirical results will be, however, identified from a vast amount of data (i.e., from 2005 to nowadays), which might make this study rather unique in the environmental literature.

Results show the GAS model under skewed t-distribution has better volatility forecasts of EUA price changes compared to any other GARCH models. This model can detect common facts about conditional volatility, such as fat-tails, excess kurtosis, persistence of volatility, asymmetry and leverage effects. We find also the superiority of some GARCH types to find VaR predictions. Thus, we conclude that it is feasible to discriminate between the estimation methods based on an analysis of the VaR forecast accuracy.

The remainder of the paper is organized as follows. Section 2 gives an overview of the regulatory design of the EU ETS. Section 3 presents a brief literature review on the formation of EUA prices and their determinants. Section 4 describes the data and the econometric models used in our analysis. Section 5 presents the main results, while Sect. 6 concludes.

2 Background on EUAs

The EU ETS remains the largest scheme to trade greenhouse gas emissions allowances in the world. As part of the EU future climate policies to enhance the energy renewable shares and reduce carbon emission by 50% in 2050, the EU EST has been entered into force in 2005 to attain these goals. This is a "cap and trade" European design that is adopted by 31 countries (28 EU countries, Iceland, Lichtenstein, and Norway). Despite its novelty compared to some rather mature markets (e.g. bond or stock markets), the CO_2 emission permit market has been growing faster than predicted. Up to now, there are 3 trading periods of CO_2 credits, the first of which was from 2005 to 2007, the second from 2008 to 2012, and the third from 2013 to 2020. The trading periods are usually referred to as "phases". In the future there may be more phases coming up; phase 4 has already been planned.[3]

There are at least six trading platforms for the trading of EUAs spot and future/forward contracts under the EU ETS: EEX (the German electricity market),[4] ECX (the European Climate Exchange), Nordpool (the Nordic power market), Powernext (the French power exchange), EXAA (the Austrian energy exchange) and Climex (an alliance formed between the Amsterdam Power Exchange-APX, the UK Power Exchange UKPX, the Spanish CO2 Exchange-SENDECO2 and a number of European companies that provide consulting and trading services in the environmental markets). As well, EUAs future/forwards contracts are listed only in Nordpool, the EEX, and the ECX while Powernext provides a continuous trading platform of EUAs spot contracts. Similarly, Australia, China, and Japan have promoted the development of carbon emission derivatives markets. The United States did not become part to the Kyoto Protocol.

[3] Three phases can be defined: (i) Phase 1 was a learning-by-doing process, during which the system was built up and experiences were accumulated for further improvement of the market. EUAs were allocated freely by the participating countries. (ii) Phase 2 expanded the scope of the scheme significantly. Norway, Iceland, and Liechtenstein joined the scheme as non-EU members, and CDM and JI credits were introduced into the trading system and were formalized by the Linking Directive. (iii) Phase 3 saw some further changes e.g. auctioning became the major way to release EUAs to the market.

[4] This is one of the largest power markets in Europe.

Participants in the EU ETS can be access international credits (i.e., emission reduction units (ERUs) and certified emission reductions (CERs)) to fulfil part of their obligations, in regard of qualitative and quantitative restrictions. However, since Phase 3, these international credits are no longer compliance units and must be exchanged for EU allowances. By 30 April of each year, installations must surrender a EUA for each ton of CO2 emitted in the previous year. During the first trading phase, most allowances in all countries can be either given away for free or sold (i.e., also referred to as grandfathering). However, during the third phase, 57% of the total amount of permits is auctioned off.

In principle, anyone can trade in the carbon market, but the main traders are energy companies, industrial companies and financial intermediaries such as banks. EUAs spot trading began in early 2005, while in August of the same year it started providing clearing services for the OTC trading of EUAs. Each EUA gives the right to its owner to emit one metric tonne of carbon dioxide during the preliminary period (years 2005–2007) and is quoted in Euros per ton of CO_2 up to two decimal places. All EUAs are kept in the national registers that are in charge of issuing or deleting them after the actual emissions have taken place. The national registers set up accounts for the interested parties (e.g. power plants, power exchanges, and others) to manage the EUAs inventories more efficiently. EEX holds such an account of the German Emissions Trading Authority (DEHSt) that is used as a collective and the central depository account.

In 2001, the exchange began offering futures contracts on the electricity index. After three years, the trading of options on these futures began, while 2005 saw the lunch of futures contracts on physical electricity and European Carbon Futures (ECFs).

Finally, the settlement price is established by the exchange following the current market price of the daily contract. At maturity, the final settlement price will automatically be the one set on the last trading day. This price is generally the base price received by clients (i.e., in the form of settlement note) to generate payments and delivery. The last business day of trading is the penultimate trading day of November with the actual delivery, taking place two days after the last scheduled trading day.

3 Literature Review

In general, financial forecasts are done in different ways. Among them, the time series forecast is a popular quantitative forecasting technique that involves collecting data during a certain period in order to identify trends. Concerning the forecasting of EUA prices, it has become a hot filed and attracted many researchers' attention. The review of previous frameworks underlines two main strand of forecasting techniques: the first one uses the linear regression technique to explain the allowance prices by different economic and financial determinants with differences in the data analysed (i.e., phases I, II and II). These varying ways to explain the allowance prices can elucidate in part the different and, sometimes, contradictory results. In this sense, we find that the judgment of the authors can be related to the trading phase selected (Chevallier, 2012; Creti et al., 2012; Rickels et al., 2015, among others) and the factors which appear to drive the carbon price as well as the key determinants of the price of EUAs (Fezzi &

Bunn, 2009; Hintermann, 2010; Maydybura, 2011; Bredin & Muckley, 2011, among others).[5] Being aware of the problems of assuming the EUA price formation in terms of a set of exogenous determinants, an alternative line of research (i.e. second strand) has emerged using more complex models and techniques susceptible to incorporate the stochastic properties of EUA data. Seifert et al. (2008), Daskalakis et al. (2009), Benz and Trück (2009) and Hitzemann and Uhrig-Homburg (2013) which focus on the stochastic properties of daily price data and provide, amongst other things, evidence for conditional heteroskedasticity. Paolella and Taschini (2008) propose mixed GARCH models which allowed for the unconditional tail behaviour and heteroskedasticity in the EUA price series. According to their results, they have reported the validity of these models in capturing the complex price volatility of EUAs at the end of Phase I. Seifert et al. (2008) use a stochastic equilibrium model to analyse the dynamics of EUA spot prices. Their main conclusion is that EUAs pricing model exhibits a time- and price volatility structure. Daskalakis et al. (2009) model the effects of abolishing banking on futures prices during Phase I and develop a framework for pricing and hedging of intra-phase and inter-phase futures and options on futures. The study covers data of three main markets for emission allowances within the EUETS: Powernext, Nord Poolandh European Climate Exchange (ECX). Their analysis suggests that the prohibition of banking of emission allowances between distinct phases of the EUETS has significant implications in terms of EUA futures pricing. Besides, the non-mean reverting models proposed by Merton (1976) are the more appropriate process. Benz and Trück (2006) advocate the use of Markov switching and GARCH classes' models to depict the volatility of the EUA spot prices in Phase I accurately. Their results support the strength of both models to emphasis the specific characteristics of the EUA time series, such as cyclic phases, volatility clustering, skewness, and excess kurtosis. Bao (2013) analyses the EUA end-of-day spot price and real-time price using the change point analysis. A key result is that the EUAs spot price can be decomposed into two parts: a diffusion part which resembles white noise, and a jump part which can be linked to influential political news in the market. It's important to note that these studies addressing the stochastic properties of EUAs prices are limited to data from Phase I. Moreover, it is possible that the results of the first phase cannot be fully generalizable to other phases. In this context, Benschopa et al. (2014) support the performance of Markov Switching GARCH models to predict EUA log-returns during the second trading phase. Recall that these models are developed to capture some characteristics of data such as the volatility clustering, breaks in the volatility process and heavy-tailed distributions. Recently, Gil-Alana (2016) re-examines the behaviour of persistence in carbon emission allowance prices. For this purpose, they use daily data for the period between 2007 and 2014 and techniques based on the concept of long memory accounting for structural breaks and non-linearities in the data. Results indicate that, while there is no evidence of non-linearity, when allowing for structural breaks, the persistence of shocks to the carbon emission allowance is markedly reduced, with the same being transitory for recent sub-samples. Similarly, Mölar et al. (2017) study structural breaks in the emission allowance price process of

[5] For additional reviews of the research during phases 1 and 2 see Zhang (2016) and Hintermann et al. (2016). For phase 3, you can cite Angeles Carnero et al. (2018).

the European Union Trading System but during Phase Π and Phase III. There is indeed a structural break between Phase II and Phase III. However, there are several regimes within each of these phases. Moreover, the findings suggest that the high-volatility regimes are usually the regimes with negative average returns, whereas low-volatility regimes usually exhibit zero or positive average returns.

Yang et al. (2016) is the first study to introduce the jump effects in modelling CO_2 emission allowance prices. In particular, they demonstrate that the dynamic jump ARMA–GARCH model can provide more accurate valuations of the CO_2 emission allowance options on futures than other models in terms of a pricing error. This result has been also showed by Song and Liang (2018) and Kostrzewski (2019).

It is important to note that this second strand has revealed the efficacity of some complex models and techniques to reflect the properties of EUA data. Our contribution to this literature is to propose a simple and sophisticated model (i.e., GAS model) to better predict EUA price volatility.

4 Data and Econometric Models

4.1 Data and Descriptive Statistics

We use daily data spanning from 22/04/2005 to 28/02/2019, comprising a total of 3566 observations. The log-return transformation to the raw data is defined as $r_t = \log\left(\frac{P_t}{P_{t-1}}\right)$, where P_t is the EUA price at time t. Table 1 and Figs. 1 and 2 show the main characteristics of EUA's data.

It is clear from Fig. 1 that EUA's datasets are characterized by extreme events during 2008. On the other hand, the Jarque–Bera test presented in Table 1 and the QQ plot reject the null hypothesis of normality distributed EUA prices and returns. Moreover, the unconditional distributions exhibit high kurtosis and positive skew. These indicate that EUA prices and returns distributions tend to have asymmetric and heavy tails. The KPSS p-value is 0.01 which rejects the null hypothesis of stationarity for both data (i.e., price and return). Since the p-value of the Ljungbox test is below 0.05, this agrees that there we cannot reject the null hypothesis of no-autocorrelation. Concerning the test for conditional heteroscedasticity on EUA price and return time series, LiMcLeod shows p-values less than 0.05 for all lags. This means the reject of the null hypothesis of no heteroscedasticity.

In sum, we find that the EUA series can involve the main characteristics of several financial markets (i.e., fat tails and excess kurtosis, clustering, long memory, and leverage effects), which justify the use of GAS and non-gaussian GARCH models in our analysis.

The top panel of Table 2 presents the GAS models, proposed in this paper to estimate and predict the EUA price changes, with different error distributions: normal (norm), t-student (std), skewed-t-student (stsd), asymmetric t-student with two decay parameters (ast), and asymmetric t-student with a left-tail decay parameter (ast1).

Table 1 Descriptive statistics

	Last price	Log change
Mean	10.59	6.95E-05
Standard deviation	6.94	0.13
Kurtosis	2.64	2183.38
Skewness	0.64	40.94
Minimum	0.01	− 1.38
Maximum	29.8	7.02
ARCH test	3527.1 [2.2e-16]	0.32 [1.000]
Chi-squared Df	12	12
J.B	266.55 [0.000]	7.07e08 [0.000]
KPSS	4.8441 [0.010]	0.04 [0.010]
Ljunbox		
Q(5)	17,442.92 [0.000]	23.12 [3.1e-04]
Q(10)	34,170.53 [0.000]	24.45 [6.4e-03]
Q(20)	65,557.52 [0.000]	28.52 [9.7e-02]
Q(30)	94,579.10 [0.000]	74.16 [1.3e-05]
LiMcLeod		
Q(5)	17,418.57 [0.000]	23.10 [3.2e-04]
Q(10)	34,099.39 [0.000]	24.43 [6.5e-03]
Q(20)	65,333.12 [0.000]	28.53 [9.7e-02]
Q(30)	94,131.71 [0.000]	73.93 [1.4e-05]
Sample	3567	3566

The descriptive statistics of EUA last price and log change for the period spans from April 22, 2005, to February 28, 2019. Ljunbox is the Ljung and Box portmanteau test applied on the residuals of the fitted model at lag values $m = 5$, 10, 20 and 30. *P*-values are calculated LiMcLeod is Li and McLeod portmanteau test. J-B is the Jaque-Bera test, while KPSS is the Kwiatkowski-Philips-Schmidt-Shin test or stationary test (the null hypothesis is that the data is stationary)

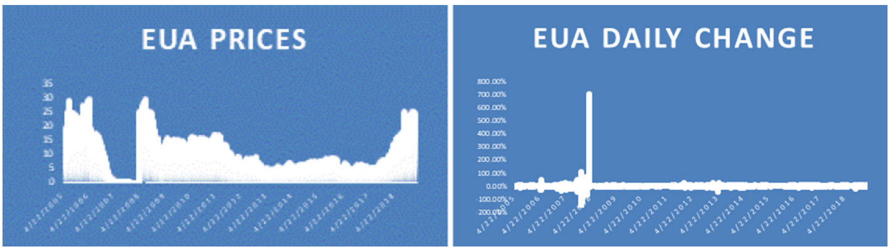

Fig. 1 Evolution of EUA price and. *Note* Data of EUA last price and log change for the period spans from April 22, 2005 to February 28, 2019

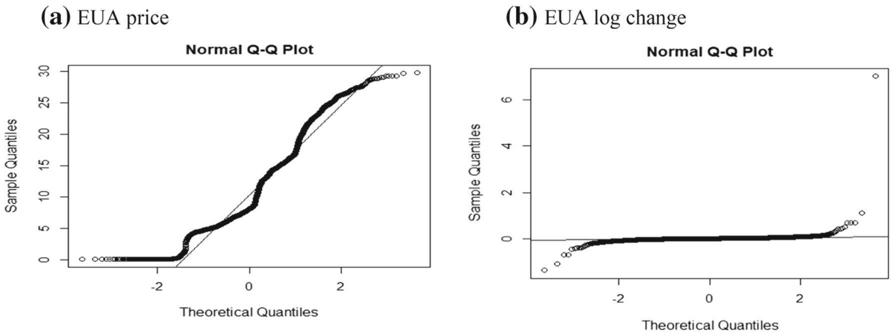

Fig. 2 QQ plot of EUA price and log change. *Note* Quantile—quantile plot between the sample of EUA prices (or return) and the normal distribution. The period spans from April 22, 2005 to February 28, 2019

Panel 2 is reserved to present the well-known GARCH models (i.e. SGARCH, IGARCH, GJRGARCH, APARCH, CSGARCH, NGARCH, NAGARCH and ALLGARCH), under different innovation distribution hypotheses such as the t-Student distribution (std), Skew-t-Student's (ged) distribution Barndorff-Nielsen (1977, 1978) and Johnson's SU (jsu) distribution (Johnson (1949)).[6]

4.2 GAS and GARCH Models

Let Φ_{t-1} be the past information set of r_t up to t–1, $p(r_t; \theta_t)$ be the conditional distribution of the returns, $r_t/\Phi_{t-1} \sim p(r_t; \theta_t)$ and $\theta_t \in \Psi \subseteq N$ be a vector of time-varying parameters that completely identifies $p(\bullet)$. We specify GAS models as follows:

$$\theta_{t+1} = \kappa + \Delta_1 s_t + \Delta_2 \theta_t \tag{1}$$

$$s_t = S_t(\theta_t) \frac{\partial logp(r_t; \theta_t)}{\partial \theta_t} \tag{2}$$

where κ, Δ_1 and Δ_2 are coefficient matrices, s_t is a vector of scaled-score steps, and $S_t(\theta_t)$ is a positive-definite scaling matrix that adjusts the shape of the score, for instance:

$$S_t(\theta_t) = E_{t-1}\left[\frac{\partial \log p(r_t; \theta_t)}{\partial \theta_t} \frac{\partial \log p(r_t; \theta_t)'}{\partial \theta_t}\right]^{-1} \tag{3}$$

For GARCH family models, the mean equation of EUA log return (r_t) can be expressed as:

$$r_t = \mu_t + \sigma_t Z_t \tag{4}$$

6 For more details about error distributions see Appendix 1.

Table 2 Estimate model specifications

	Norm	Std	Sstd	AST	AST1
Panel 1. GAS models					
GAS	AR(1)-GAS-norm	AR(1)-GAS-t-Student	AR(1)-GAS-Skewed t-Student	AR(1)-GAS-AST	AR(1)-GAS-AST1

GARCH models	Error distribution hypothesis			
	Std	Ged	sstd	jsu
Panel 2. Non-Gaussian GARCH family models				
NGARCH	NGARCH-std	NGARCH-ged	NGARCH-sstd	NGARCH-jsu
NAGARCH	NAGARCH-std	NAGARCH-ged	NAGARCH-sstd	NAGARCH-jsu
ALLGARCH	ALLGARCH-std	ALLGARCH-ged	ALLGARCH-sstd	ALLGARCH-jsu
sGARCH	sGARCH-std	sGARCH-ged	sGARCH-sstd	sGARCH-jsu
csGARCH	csGARCH-std	csGARCH-ged	csGARCH-sstd	csGARCH-jsu
apARCH	apARCH-std	apARCH-ged	apARCH-sstd	apARCH-jsu
gjrGARCH	gjrGARCH-std	gjrGARCH-ged	gjrGARCH-sstd	gjrGARCH-jsu
iGARCH	AR(1)-iGARCH-std	AR(1)-iGARCH-ged	AR(1)-iGARCH-sstd	AR(1)-iGARCH-jsu

where μ_t denotes the conditional mean, σ_t denotes a volatility process and Z_t is error terms.

In the standard specifications of GARCH (1,1) model of Bollerslev (1986), denoted by SGARCH (1, 1), the second conditional variance equation is given by:

$$\sigma_t^2 = \omega + \alpha_1 Z_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{5}$$

For $\alpha_1 > 0$, $\beta_1 > 0$ and $\omega > 0$. The constraint of the volatility clustering in the data is captured by $\alpha_1 + \beta_1 > 1$.

The IGARCH(1,1) model generates the standard model by allowing $\alpha_1 + \beta_1 = 1$ (the null hypothesis of a unit root), so that Engle and Bollerslev (1986) rewritten the conditional variance equation as:

$$\sigma_t^2 = \omega + \sigma_{t-1}^2[(1 - \alpha_1) + \alpha_1 Z_{t-1}^2] \tag{6}$$

The Glosten-Jagannathan-Runkle GARCH (GJR-GARCH) model assumes a specific parametric form for the conditional heteroskedasticity, this is an asymmetric version of the SGARCH model. More specifically, Glosten et al. (1993) express the conditional variance equation as:

$$\sigma_t^2 = \omega + \alpha_1 Z_{t-1}^2 + \gamma_1 I_{t-1} Z_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{7}$$

$$I_{t-1} = \begin{cases} 1 & \text{if} Z_{t-1} \leq 0 \\ 0 & \text{otherwise} \end{cases}$$

For $\alpha_1 > 0$, $\beta_1 > 0$, $\gamma_1 > 0$ and $\omega > 0$.

The ALL-GARCH (1, 1) of Hentschel (1995) represents a general class of models that use the Box and Cox (1964) transformation formation. This can be written as:

$$\sigma_t^\delta = \omega + \alpha_1 \sigma_{t-1}^\delta \left[|Z_{t-1} - \eta_1| + \gamma_1 (Z_{t-1} - \eta_1)\right]^\delta + \beta_1 \sigma_{t-1}^\delta \tag{8}$$

For $\delta > 0$, $\alpha_1 \geq 0$, $\beta_1 \geq 0$, $-1 \leq \gamma_1 < 1$, $-\infty \leq \eta_1 < \infty$ and $\omega > 0.\delta$ is a parameter for the Box-Cox transformation. The persistence parameter is equal to $\alpha_1 + \beta_1 + \gamma_1 \in_1$, where $\in_1$ is the expected value of the standardized residuals under the Box-Cox transformation of the absolute value of the asymmetry term.

The ALLGARCH(1,1) model can contain the following as particular variants: the NAGARCH (1, 1) model of Engle & Ng (1993) for $\delta = 2$ and $\gamma_1 = 0$; the GARCH (1, 1) model of Bollerslev (1986) for $\delta = 2$ and $\gamma_1 = \eta_1 = 0$ and $g1 = h1 = 0$; the GJR-GARCH (1, 1) model of Glosten et al. (1993) for $\delta = 2$ and $\eta_1 = 0$; the NGARCH (1, 1) model of Higgins and Bera (1992a) for $\gamma_1 = \eta_1 = 0$; the TGARCH (1, 1) model of Zakoian (1994) for $\delta = 1$ and $\eta_1 = 0$; the APARCH (1, 1) model of Ding et al. (1993) for $\eta_1 = 0$.

Two additional classes of models are also proposed:

The asymmetric power ARCH model (Ding et al., 1993), denoted by APARCH (1, 1):

$$\sigma_t^\delta = \omega + \alpha_1(|Z_{t-1}| - \gamma_1 Z_{t-1})^\delta + \beta_1\sigma_{t-1}^\delta \tag{9}$$

For $\delta > 0, \alpha_1 \geq 0, \beta_1 \geq 0, -1 \leq \gamma_1 < 1$ and $\omega > 0.\delta$ is a parameter for the Box-Cox transformation and γ_1 is a leverage parameter.

The component standard GARCH model (Engle et al., 1999), denoted by CSGARCH (1, 1):

$$\sigma_t^2 = q_t + \alpha_1\left(Z_{t-1}^2 - q_{t-1}\right) + \beta_1\left(\sigma_t^2 - q_{t-1}\right) \tag{10}$$

where

$q_t = \omega + \rho q_{t-1} + \phi\left(Z_{t-1}^2 - \sigma_{t-1}^2\right)$.

For $\delta > 0, \alpha_1 \geq 0, \beta_1 \geq 0, \phi > 0$ and $\omega > 0$. Weak stationarity holds if $\alpha_1 + \beta_1 < 1$ and $\rho < 1$. CSGARCH (1, 1) decomposes the conditional variance into permanent and transitory components to investigate the long- and short-run movements of volatility. Compared to SGARCH (1, 1), the intercept parameter is now a time-varying first-order autoregressive process.

To choose the best alternative that model and predict the EUA daily returns, we use ML, AIC, and BIC as fit criteria. The model which gives the minimum AIC or BIC is selected as the best.[7]

4.3 Comparison of Value-at-Risk Forecasts Using MCS Procedure

In this paper, we will also evaluate how EUA VaR returns fit. For that, we compare the 1-day ahead forecasts of q%-VaR forecast models delivered by alternative model specifications above introduced (i.e., GAS and GARCH models), using the MCS procedure. To this end, models are compared in terms of their predictive ability, so that models that produce better forecasts of EUA VaR price are preferred.

Formally, let Φ denotes the cumulative density function (CDF) and the conditional VaR denoted as $VaR_{t+1}^\alpha \backslash F_t = \alpha$. For GAS and GARCH models, the VaR forecasts are obtained by inverting the corresponding conditional cumulative density function . If the density of y belongs to the location-scale family, it may be estimated from:

$$VaR_{t+1}^\alpha = \mu_{t+1}(\theta_t) + \Phi_{t+1}^{-1}(\alpha)\sigma_{t+1}(\theta_t), \tag{11}$$

where Φ_{t+1}. is the forecast cumulative distribution of the standardized return, $\mu_{t+1}(\theta_t) = E(y_{t+1}/F_t)$ is the conditional mean forecast of the return, and $\sigma_{t+1}^2(\theta_t) = E\left(\varepsilon_{t+1}^2/F_t\right)$ is the conditional variance forecast based on the volatility models of previous sections, and θ_t is the parameter vector estimated by using the information up to time t.

[7] It's interesting to note that we use three criterions because AST density does not satisfy the usual regularity conditions for maximum likelihood estimation.

Among the multiple-testing procedures available in the forecasting literature (see White (2000), Romano and Wolf (2005), Hansen (2005), Giacomini and White (2006) and Hansen et al. (2011), among others), we conduct the MCS procedure of Hansen et al. (2011) to evaluate EUA VaR forecasts. This procedure is based on an equivalence test, and an elimination rule.

(i) Equivalence test

This consists of selecting a "Superior Set of Models" (SSM) after a sequence of tests under the null hypothesis of equal predictive ability (EPA) and given an arbitrary loss function. Formally, let M_0 denotes the start number of predictive models and the smaller set of superior models by $\hat{M}^*_{1-\alpha}$. The MCS procedure consists of performing multiple pairwise comparisons and eliminates at each iteration the inadequate models. At the final step, the MCS procedure generates a smaller set of acceptable models. The loss function, $\ell_{i,t}$, associated with the ith model is defined as:

$$\ell_{i,t} = \ell\left(r_t, \hat{r}_{i,t}\right) \tag{12}$$

where r_t denote the observation of EUA daily returns at time t and $\widehat{r}_{i,t}$ be the output of model indexed by i = 1,…,m. This function measures the difference between the output $\hat{r}_{i,t}$, and the "posteriori" at time t. Formally, let $d_{i,j,t}$ denote the loss differential between models i and j at time t:

$$d_{ij,t} = \ell_{i,t} - \ell_{j,t},\ \text{i, j} = 1, \ldots, \text{m},\ \text{t} = 1, \ldots, \text{n} \tag{13}$$

We rank models in terms of expected loss, so that model "i" is preferred to model "j" if $E(d_{ij,t}) < 0$.

The EPA hypotheses for a given set of models "M" take the form:

$$H_{0,M} : E(d_{ij,t}) = 0$$

$$H_{1,M} : E(d_{ij,t}) \neq 0$$

To test $H_{0,M}$, Hansen et al. (2011) propose the following test statistics:

$$T_{R,M} = \max_{i,j \in M} \left|t_{ij}\right| \tag{14}$$

$$T_{max,M} = \max_{i \in M} t_{i,} \tag{15}$$

where t_{ij} and $t_{i.}$ are two statistics defined as follow:

$$t_{ij} = \frac{\overline{d}_{ij}}{\sqrt{\widehat{\text{var}}\left(\overline{d}_{ij}\right)}} \tag{16}$$

$$t_i = \frac{\overline{d}_i}{\sqrt{\widehat{\text{var}}\left(\overline{d}_{i.}\right)}} \tag{17}$$

$\overline{d}_i = (m-1)^{-1} \sum_{j \in M \setminus \{i\}} \overline{d}_{ij}$ is the average loss of ith model relative to the average losses across the models belonging to the set M, and $\overline{d}_{ij} = n^{-1} \sum_{jt=1} d_{ij,t}$ measures the relative average loss function between models i and j. The variances $\widehat{\text{var}}(\overline{d}_{i.})$ and $\widehat{\text{var}}(\overline{d}_{ij})$ are bootstrapped estimates of $\widehat{\text{var}}(\overline{d}_{i.})$ and $\widehat{\text{var}}(\overline{d}_{ij})$, respectively. (for more details see White (2000)).

(ii) Elimination rule

After the equivalence test, the MCS procedure eliminates the worst model using an elimination rule that is coherent with the statistic tests defined in Eqs. (14)–(15). For the test statistic $T_{\max,M}$, the natural elimination rule is:

$$e_{\max,M} = \arg\max_{i \in M} \frac{\overset{=}{d}_{i,.}}{\sqrt{\widehat{\text{var}}(\overline{d}_{i,.})}} \tag{18}$$

With the other statistic test, we have:

$$e_{R,M} = \arg\max_{i} \left\{ \sup_{j \in M} \frac{\overset{=}{d}_{ij}}{\sqrt{\widehat{\text{var}}(\overline{d}_{ij})}} \right\} \tag{19}$$

5 Main Results

It is interesting to note that, in the literature, we can find two main types of models regarding directly or indirectly the CO_2 emission allowance price: Models in which the CO_2 emission allowance is a determinant influencing the broadly understood energy market, and models where the main goal is to analyze the CO_2 emission allowance price and the determinants that influence this price. Our study seems different because it considers the prediction and modeling of CO_2 emission allowance price volatility.

We first estimate the parameters of model fits, and then to assess the performance, we use the maximum Likelihood or the minimum AIC and BIC criterion as model selection measures. For that, the baseline GARCH model of Bollerslev (1986) and a list of popular extensions, are compared to GAS models, which are promising in modelling highly nonlinear volatility dynamics (Creal et al., 2013 and Harvey, 2013).

Comparison results are presented in Table 3.

The top panel "a" of Table 3 reports the results related to GAS models. We show AR(1)-GAS-ast the best fit model using ML measure, whereas AR(1)-GAS-std achieved the minimum BIC and AIC.

The bottom panel "b" reports the results of ARCH- type models. Here, we consider these not only because of their popularity, but also because of their ability to account for the main stylised facts about EUA daily returns. According to ML measure, the APARCH model concerning skewed t-student density for the error term appears the

Table 3 Estimation of EUAs daily returns

GAS model	ML	AIC	BIC	np
Panel a: GAS models				
AR(1)− GAS− norm	− 8603.76	17,219.51	17,256.07	6
AR(1)-GAS-std	− 8184.62	**16,383.24**	**16,425.89**	7
AR(1)-GAS-sstd	− 8187.16	16,390.31	16,439.06	8
AR(1)-GAS- ast	**− 8180.77**	16,387.53	16,466.74	13
AR(1)-GAS-ast1	− 8210.26	16,440.51	16,501.44	10

	ML	AIC	BIC	$Q^2(10)$	ARCH(5)
Panel b: NON-GAUSSIAN GARCH modelsPanel b. NON-GAUSSIAN GARCH models					
AR(1) NGARCH-std	− 8192.74	5.013597	5.026635	37.010*	5.620123
AR(1)-NGARCH-ged	− 8246.74	5.046615	5.059654	28.478*	4.900852
AR(1)-NGARCH-sstd	− 8191.23	5.013288	5.02819	37.947*	4.716558
AR(1)-NGARCH-jsu	− 8195.95	5.016169	5.031071	42.607*	5.934971
AR(1)-NAGARCH-std	− 8219.88	5.030192	5.043231	5.56129	0.029062
AR(1)-NAGARCH-ged	− 8265.08	5.057827	5.070866	5.68901	0.029941
AR(1)-NAGARCH-sstd	− 8217.14	5.029127	5.044029	4.88864	0.025436
AR(1)-NAGARCH-jsu	− 8220.50	5.031184	5.046085	5.33412	0.028871
AR(1)-ALLGARCH-std	− 8180.95	5.007612	5.024376	30.332*	1.811083
AR(1)-ALLGARCH-ged	− 8232.35	5.039039	5.055803	18.24**	1.177289
AR(1)-ALLGARCH-sstd	− 8177.59	5.006170	5.024797	15.9***	0.512621
AR(1)-ALLGARCH-jsu	− 8180.94	5.008219	5.026846	22.406*	1.188631
AR(1)-sGARCH-std	− 8227.76	5.034402	5.045578	5.91244	0.052401
AR(1)-sGARCH-ged	− 8273.8	5.062552	5.073728	5.93874	0.049222
AR(1)-sGARCH-sstd	− 8227.49	5.034848	5.047887	5.81846	0.047560
AR(1)-sGARCH-jsu	− 8229.36	5.035990	5.049028	6.14506	0.053701
AR(1)-csGARCH-std	− 8230.92	5.037553	5.052454	6.14401	0.054813
AR(1)-csGARCH-ged	− 8278.34	5.066546	5.081448	6.26576	0.050135
AR(1)-csGARCH-sstd	− 8233.95	5.040018	5.056782	5.91532	0.056092
AR(1)-csGARCH-jsu	− 8234.75	5.040508	5.057272	7.13552	0.061508
AR(1)-apARCH-std	− 8178.36	5.005415	5.020316	25.086*	1.287079
AR(1)-apARCH-ged	− 8232.27	5.038382	5.053284	18.553*	1.316956
AR(1)-apARCH-sstd	**− 8177.09**	**5.005250**	**5.022015**	**17.30****	0.682402
AR(1)-apARCH-jsu	− 8181.49	5.007943	5.024707	19.280*	1.111426
AR(1)-gjrGARCH-std	− 8216.86	5.028347	5.041386	4.67388	0.021750
AR(1)-gjrGARCH-ged	− 8261.43	5.055599	5.068638	5.19307	0.023212
AR(1)-gjrGARCH-sstd	− 8214.80	5.027697	5.042598	5.21030	0.024869
AR(1)-gjrGARCH-jsu	− 8218.00	5.029653	5.044555	5.35531	0.025851

Table 3 (continued)

	ML	AIC	BIC	$Q^2(10)$	ARCH(5)
AR(1)-iGARCH-std	− 8227.93	5.033894	5.043208	4.93764	0.041133
AR(1)-iGARCH-ged	− 8273.71	5.061881	5.071195	6.57911	0.060453
AR(1)-iGARCH-sstd	− 8226.01	5.033329	5.044505	5.46085	0.045751
AR(1)-iGARCH-jsu	− 8229.10	5.035217	5.046393	6.17584	0.055874

We report ML, AIC, BIC and number of parameters (np) of different GAS specifications. The error distributions: Normal (norm), t-student (std), skewed t-student distribution (sstd), Asymmetric t-Student with two decay parameters (ast) and Asymmetric t-Student with a left-tail decay parameter (ast1), Johson's Reparametrized SU (jsu)or generalized error distribution (ged)

The best models are in boldface. The model which gives the maximum LL or the minimum AIC or BIC is selected as the best

LL, AIC and BIC denote Log-likelihood, Akaike Information criterion and Bayesian Information criterion, respectively. *, ** and *** denote p-value of accepted H0. $Q^2(10)$ is the Ljung-Box Q-statistic of order 10 computed on the squared standardized residuals. The sample spans from April 22, 2005 to February 28, 2019

more suitable model for the EUA's daily data. The performance of this model is also confirmed by AIC and BIC criteria (i.e., they are smaller than the values under other models). This empirical finding documents that the standard GARCH model should not be adopted to assessing EUA's daily data. For Ding (2011), the APARCH model can forecast and capture common facts about conditional volatility, such as fat-tails, the persistence of volatility, asymmetry and leverage effect. For the skewed t-student's distribution that describes the leverage effects of data precisely, generates the larger ML.

These features are also captured by our descriptive statistics (see Table 1). Precisely, we find that the EUA market can involve the main characteristics of several financial markets (i.e., fat tails and excess kurtosis, clustering, long memory, and leverage effects), which justify the use of GAS and non-gaussian GARCH models in our analysis. In other words, the EUAs market and the main financial markets (e.g., Standard & Poor 500 stock market daily closing price index and MSCI EUROPE INDEX, see Ding (2011)) can share the same characteristics.[8]

Considering now all results of both panels, the models under skewed and heavy tail distributions (i.e. ast, and sstd) outperform the models under gaussian innovation distribution. This empirical finding is consistent with the greater impact that the GFC and Europe sovereign debt had in the European financial markets as compared to the non–Euro areas. Hence, the related distribution to these features can produce accurate VaR forecasts for the EUA market.

To examine the performance of models used to forecast VaRs, the complete dataset of EUA daily returns is divided into two samples: an out-sample period from April 22, 2005, to April 15, 2015, for a total of 2567 observations, and a forecast or validation period, containing the remaining 1000 observations. A fixed rolling window approach is also implemented to produce 1-day ahead forecasts VaRs at quantile levels 1% and

[8] Estimated coefficients for each model are not reported to save space, but they are available upon request.

5%. This means that we consider the VaR measure that represents the quantiles 1% and 5% of the profit-loss distribution. Table 4 presents the main results.

In this table, the MCS procedure of Hansen et al. (2011) is applied to obtain the set of models with superior predictive ability in terms of the supplied VaR forecasts.[9] After that, the selected models are ranked according to their $T_{R,M}$ and $T_{max,M}$ statistics. The *p*-values of the $T_{R,M}$ and $T_{max,M}$ statistics, are reported in the third and sixth columns of Table 4, respectively. For the final column, it is, however, reserved for the expected loss given a VaR violation (i.e., the difference between realized and evaluated VaR measures).

The top panel shows an SSM of 18 models. Other models are, however, removed by the elimination rule. This is the aim advantage of the MCS approach as it is possible to reduce the set of competing models when is large, to a smaller set of models.

It is worth noting that the models belonging to SSM can be used for different purposes. For example, they can be used to forecast future volatility levels, to predict future observations, conditional to the past information, see, e.g., Gneiting (2011), or to deliver future VaR estimates, as argued by Bernardi et al. (2015). In this study, the SSM set is used to find the true model that provides more precise EUA's VaR forecasts.

Regarding the rank of the accepted models, the GJR-GARCH-std is the best model to predict the 1%-VaR of the EUA daily change. These results are related to the $T_{max,M}$ statistic and the expected loss. Whereas $T_{R,M}$ statistic considers the GJR-GARCH-ged followed by the GJR-GARCH-std as the best models to predict 1%-VaR of EUAs returns. Concerning the GAS models, they are among the models surviving the test. Although these models are not ranked in the three top models (GAS-sstd has ranked 9), they have similar forecasting ability of GARCH family models to forecast EUA VaRs. As Hansen et al. (2011), if the final SSM contains a big portion of the starting M^0 set, then the competing model are statistically equivalent in term of their forecast ability of future VaR levels.

These empirical findings are interesting because they prove some previous results of Paolella and Taschini (2008), Benz and Trück (2006), Benschopa et al. (2014) that support the performance of GARCH models to predict EUA log-returns but for other sample periods and separate phases.

For 5% VaR forecasting results, the MCS procedure accepts all considered models of GARCH and GAS families. Naturally, the higher the number of eliminated models, the higher the heterogeneity of the competing forecasts. On the contrary, if the final set of models contains a big portion of the starting set, then the competing model are statistically equivalent in term of their forecast ability of future VaR levels.

Based on $T_{R,\,M}$ and $T_{max,M}$ statistics, we show that both GAS-sstd followed by GAS-std models consistently outperform all other models to forecast the 5% VaR of EUA daily returns.

Overall the skewed Student's t distribution and Generalized error distributions fits well EUA VaRs. The implication of these results is that, the volatility of CO_2 emission allocations may not be affected by some stylized facts such as the leverage effect or by complex nonlinear conditional volatility dynamics. Nonetheless, it exhibits excess

[9] It is worth noting that the estimated coefficients for each model, over the out-sample period, are not reported to save space, but they are available upon request to the first author.

Table 4 Comparison of Value-at-Risk models using the MCS approach

	Rank_M	t_{ij}	$p-\text{value}_{R,m}$	$\text{Rank}_{\max,M}$	t_i	$p-\text{value}_{\max,M}$	Loss
Panel a: For quantile = 1%							
GAS_norm	12	0.802377	0.9855	14	1.837407	0.8505	0.000942
GAS_std	10	0.028623	1	8	1.010690	0.9995	0.000923
GAS_sstd	9	− 0.260190	1	7	0.985736	1	0.000920
GAS_ast	14	1.000301	0.937	15	1.840773	0.8480	0.000933
GAS_ast1	15	1.964728	0.303	18	3.274984	0.068	0.000947
ALLGARCH-std	16	1.019026	0.929	12	1.578733	0.9505	9.37E−04
ALLGARCH-ged	11	0.689162	0.9945	11	1.436428	0.975	9.30E−04
sGARCH-std	3	− 0.858690	1	3	0.853273	1	9.16E−04
sGARCH-ged	4	− 0.808700	1	6	0.921187	1	9.16E−04
csGARCH-std	8	− 0.419820	1	9	1.181648	0.996	9.19E−04
csGARCH-ged	7	− 0.444830	1	10	1.222359	0.9935	9.20E−04
apARCH-std	17	1.497575	0.64	16	2.106949	0.659	9.41E−04
apARCH-ged	13	0.986871	0.9415	13	1.823774	0.855	9.31E−04
apARCH-sstd	18	2.174969	0.2055	17	2.992392	0.1355	9.51E−04
gjrGARCH-std	2	− 1.59956	1	**1**	**0.753278**	1	**9.11E−04**
gjrGARCH-ged	**1**	**− 1.64036**	1	4	0.873460	1	9.13E−04
iGARCH-std	5	− 0.78079	1	2	0.851362	1	0.000916
iGARCH-ged	6	− 0.76444	1	5	0.911884	1	0.000916
Panel b: For quantile = 5%							
GAS_norm	37	1.668076	0.4915	27	2.119677	0.715	0.003199

Table 4 (continued)

	Rank_M	t_{ij}	$p-\text{value}_{R,m}$	$\text{Rank}_{\max,M}$	t_i	$p-\text{value}_{\max,M}$	Loss
GAS_std	2	− 1.86113	1	2	0.364946	1	0.003088
GAS_sstd	**1**	**− 2.04919**	1	**1**	**− 0.364950**	**1**	**0.003086**
GAS_ast	8	− 1.00816	1	14	0.962580	1	0.003103
GAS_ast1	7	− 1.13764	1	15	1.119941	1	0.003103
NGARCH-std	3	− 1.65661	1	3	0.574086	1	0.003096
NGARCH-ged	9	− 0.90948	1	13	0.990018	1	0.003108
NGARCH-sstd	4	− 1.46743	1	9	0.875740	1	3.10E−03
NGARCH− jsu	5	− 1.30687	1	11	0.900223	1	3.10E−03
NAGARCH-std	15	− 0.10137	1	17	1.205714	0.9985	3.13E−03
NAGARCH-ged	29	0.772616	0.9835	28	2.185286	0.6745	3.14E−03
NAGARCH-sstd	25	0.359407	1	26	1.931693	0.8365	3.14E−03
NAGARCH-jsu	27	0.59119	0.9965	29	2.229389	0.6345	3.14E−03
ALLGARCH-std	18	− 0.0179	1	12	0.961043	1	3.13E−03
ALLGARCH-ged	20	0.191984	1	16	1.136352	1	3.14E−03
ALLGARCH-sstd	21	0.252826	1	21	1.567884	0.965	3.14E−03
ALLGARCH-jsu	22	0.281895	1	20	1.545705	0.9705	3.14E−03
sGARCH-std	23	0.340878	1	23	1.839962	0.884	3.14E−03
sGARCH-ged	35	1.445344	0.6505	36	2.891081	0.2145	3.15E−03
sGARCH-sstd	30	0.824666	0.9755	31	2.358263	0.547	3.14E−03
sGARCH-jsu	33	1.193821	0.8255	34	2.64082	0.3435	3.15E−03
csGARCH-std	19	0.167509	1	22	1.603001	0.965	3.13E−03

Table 4 (continued)

	Rank_M	t_{ij}	$p-\text{value}_{R,m}$	$\text{Rank}_{\max,M}$	t_i	$p-\text{value}_{\max,M}$	Loss
csGARCH-ged	28	0.656483	0.994	33	2.417426	0.5035	3.14E−03
csGARCH-sstd	24	0.359401	1	25	1.894551	0.866	3.14E−03
csGARCH-jsu	31	0.858112	0.97	30	2.30132	0.6345	3.15E−03
apARCH-std	13	− 0.33766	1	5	0.688239	1	3.12E−03
apARCH-ged	14	− 0.19229	1	10	0.877377	1	3.12E−03
apARCH-sstd	16	− 0.07619	1	19	1.27765	0.9955	3.13E−03
apARCH-jsu	17	− 0.05652	1	18	1.186712	0.9985	3.13E−03
gjrGARCH-std	6	− 1.17531	1	4	0.576688	1	3.11E−03
gjrGARCH-ged	12	− 0.56995	1	8	0.845795	1	3.12E−03
gjrGARCH-sstd	10	− 0.86383	1	6	0.708671	1	3.11E−03
gjrGARCH-jsu	11	− 0.73515	1	7	0.752732	1	3.11E−03
iGARCH-std	26	0.38096	0.9995	24	1.87841	0.866	3.14E−03
iGARCH-ged	36	1.474622	0.632	37	2.901935	0.2115	3.15E−03
iGARCH-sstd	32	0.883886	0.9655	32	2.405433	0.5035	0.003146
iGARCH-jsu	34	1.260072	0.788	35	2.681756	0.32	0.00315

Better models are in boldface. The model with minimum $\boldsymbol{t_{ij}}$, $\boldsymbol{t_i}$ and /or expected loss are are the best

The table reports the results of the Model Confidence Set (MCS) procedure for the one-day ahead 5% and 1% VaR. This procedure is based on equivalence tests and elimination rules. For the equivalence test, Hansen et al. (2011) propose two measures based on $\boldsymbol{t_{ij}}$ and $\boldsymbol{t_i}$ statistics. For the elimination rules, panel "a" presents a reduced set of 18 models, the other models are removed

Table 5 Back-testing Measures using the unconditional coverage (UC)

	$VaR_t^{1\%}$			$VaR_t^{5\%}$		
	AE	AD.mean	AD.max	AE	AD.mean	AD.max
GAS_norm	0.9	2.399880	**10.20778**	0.64	1.945316	**12.93999**
GAS_std	0.5	3.006804	7.774074	0.82	1.764254	12.14513
GAS_sstd	0.6	**2.602399**	8.122245	0.84	1.766260	12.30644
GAS_ast	0.5	3.232374	**8.951641**	0.8	1.766848	**12.92184**
GAS_ast1	0.4	3.876954	**8.903665**	0.78	1.759065	**12.91503**
NGARCH-std	0.5	3.094111	8.099698	0.9	**1.633248**	12.50680
NGARCH-ged	0.7	**2.263126**	7.959888	0.7	1.834309	12.11649
NGARCH-sstd	0.5	3.110867	8.258092	0.84	1.701959	12.57223
NGARCH-jsu	0.5	3.073640	8.169831	0.84	**1.650869**	12.48818
NAGARCH-std	0.6	2.691131	8.015050	0.88	1.823896	12.44474
NAGARCH-ged	0.6	2.689661	7.866236	0.84	1.685707	12.04748
NAGARCH-sstd	0.6	2.692132	8.188470	0.86	1.806257	12.51640
NAGARCH-jsu	0.6	2.649587	8.117098	0.84	1.802402	12.43869
ALLGARCH-std	0.7	2.439187	8.218095	0.86	1.814818	12.57081
ALLGARCH-ged	0.7	2.364060	8.102379	0.76	1.777042	12.19834
ALLGARCH-sstd	0.6	2.786668	8.364188	0.84	1.795091	12.63284
ALLGARCH-jsu	0.5	3.304144	8.297453	0.82	1.798167	12.56067
sGARCH-std	0.6	2.556116	7.895462	0.88	1.839871	12.37799
sGARCH-ged	0.6	2.597669	7.746672	0.84	1.719813	11.97669
sGARCH-sstd	0.5	3.060238	8.061028	0.86	1.816225	12.44593
sGARCH-jsu	0.5	3.009582	7.973912	0.86	1.764801	12.36161
csGARCH-std	0.7	2.383590	8.392860	0.88	1.878146	12.60322
csGARCH-ged	0.8	**2.092109**	8.238233	0.84	1.748327	12.27117
csGARCH-sstd	0.7	2.352876	8.504578	0.86	1.838409	12.65151
csGARCH-jsu	0.6	2.658127	8.397718	0.88	1.761595	12.57820
apARCH-std	0.7	2.424832	8.523484	0.84	1.803384	12.76541
apARCH-ged	0.6	2.727751	8.33826	0.74	1.774940	12.35600
apARCH-sstd	0.6	2.774371	8.692636	0.82	1.797793	12.83615
apARCH-jsu	0.6	2.737018	8.610819	0.8	1.789233	12.75470
gjrGARCH-std	0.5	3.055815	8.157655	0.88	1.770077	12.53406
gjrGARCH-ged	0.5	3.100562	8.01478	0.82	**1.664333**	12.14563
gjrGARCH-sstd	0.5	3.058460	8.334611	0.88	1.700729	12.60735
gjrGARCH-jsu	0.5	3.018121	8.265416	0.86	1.689946	12.53204

Table 5 (continued)

	$\text{VaR}_t^{1\%}$			$\text{VaR}_t^{5\%}$		
	AE	AD.mean	AD.max	AE	AD.mean	AD.max
iGARCH-std	0.6	2.526113	7.836081	0.88	1.831694	12.34950
iGARCH-ged	0.6	2.575656	7.697123	0.84	1.711237	11.94614
iGARCH-sstd	0.5	3.029973	8.014794	0.86	1.808706	12.42383
iGARCH-jsu	0.5	2.971440	7.903930	0.86	1.754506	12.32524

Better models are in boldface. The model with lower mean and/or maximum ADs are preferred
The table reports the results of the unconditional coverage (UC) test using the Quantile loss (QL) function of Gonzalez-Rivera et al. (2004) for the one-day ahead 5% and 1% VaR. AE is the ratio between the realized q%- VaR exceedances and their "a priori" expected values. ADmean and ADmax are the mean and maximum Absolute Deviation (AD). Results are related to 1000 daily rolling out-of-sample forecasts

kurtosis and heavy-tailed distribution (i.e. std or sstd). These features are probably due to the higher impact of the Global Financial Crisis (see Fig. 1). Hence, our empirical analyses suggest that making inferences without considering the distribution of error could severely impact the EUA risk assessment procedure.

5.1 Robustness Check

To examine whether evaluation results of the MCS procedure are robust and to investigate which model has the best forecasting performance of the EUA extreme volatility, we perform the following three backtesting procedures: the unconditional coverage (UC) test of Kupiec (1995), the conditional coverage (CC) test of Christoffersen (1998) and the dynamic quantile (DQ) test of Engle and Manganelli (2004):

5.1.1 Unconditional Coverage (UC) Test

As it is one of the most used loss functions in the VaR context, we propose the use in this test the Quantile loss (QL) function of Gonzalez-Rivera et al. (2004) for selecting the best fitting models of EUA VaR forecasts. This is an asymmetric loss function that imposes a higher penalty with (1-q), for a risk level q, to the observation where returns display VaR exceedance. Formally, it can be defined as:

$$QL_{t+1}(q\%) = (q\% - d_{t+1})(r_{t+1} - VaR_{t+1}(q\%)) \tag{22}$$

where $\text{VaR}_{t+1}(q\%)$ is a q%-VaR "a priori" expected value at t + 1 and $d_{t+1} = 1(r_{t+1} - \text{VaR}_{t+1}(q\%))$ is a realized VaR exceedance. The UC test calculates the Actual over Expected Ratio (AE) backtesting measure over a given time horizon, which is the ratio between the realized q%- VaR exceedances and their "a priori" expected values:

$$AE = \frac{\sum_{j=1}^{P} d_{t+j}}{q\% P} \tag{23}$$

We perform 1000 daily rolling out-of-sample forecasts (i.e. $p = 1000$). The VaR forecast series that has an AE ratio closer to the unity is preferred, while values higher or lower than the unity indicate that the models overestimate or underestimate risk, respectively.

Besides AE ratio, we suggest two other measures such as the mean and maximum Absolute Deviation (ADmean and ADmax). The AD provides a measure of the expected loss given VaR violation models and expressed as:

$$AD_t = \left| r_t - \left(-(q\% VaR)_t \right) \right| \tag{24}$$

The mean and maximum AD are calculated here to compare the competing q% VaR models. In general, models with lower mean and/or maximum ADs are preferred. Table 5 reports the results

With q = 1%, the results show that AE ratios are statistically less than 1. All GAS and ARCH-type models under-estimate EUA VaRs. The risk violation is then important with these models. As Basel Accords, models that over-estimate risk are preferable to those that under-estimate risk levels.

The AD backtesting measure may not also support a specific model. For instance, the results related to 1%-VaR show that GAS-norm model ranks best on ADmax, while csGARCH-ged and NGARCH-std rank in top for ADmean. In sum, the UC test gives unsatisfying information about the best model for 1% and 5% EUA VaR forecasts. For this reason, we propose two other tests for the robustness of our previous results.

5.1.2 Conditional Coverage (CC) and Dynamic Quantile (DQ) Tests

Christoffersen (1998) recommends the use of the conditional coverage test, which is a joint test of unconditional coverage and independence. Precisely, the UC test verifies the correct coverage at the left-tail of the marginal distribution of the returns, while the CC test analyses the conditional density of the returns. In this case, VaR exceedances should be independently distributed over time.

Engle and Manganelli (2004) have proposed a general testing procedure for dynamic quantile models. The dynamic quantile test (DQ) jointly tests for UC and CC and has more power than previous alternatives under some form of model misspecification. Under the null hypothesis, we accept the correct model specification.

The results of both tests are reported in Table 6.

Table 6 presents the p-values of the CC and DQ tests for 1% and 5%-VaR forecasts using different GAS and GARCH models' specifications. The worst models appear in bold. For DQ test, one can see the presence of larger p-values for all VaR predictions with different model specifications. Similar results are also found for CC test at 5% significance level, except GAS-norm model for quantile 5% and GAS-ast1 for 1%

Table 6 Back-testing measures using the dynamic quantile (DQ) test and the conditional coverage (CC) test

	$VaR_t^{1\%}$		$VaR_t^{5\%}$	
	DQ p.value	*CC p*.value	*DQ p*.value	*CC p*.value
GAS_norm	0.993803788	0.746470877	0.113462460	**0.005292895**
GAS_std	0.920508525	0.078594056	0.529989746	0.178266318
GAS_sstd	0.969802863	0.169627481	0.611023780	0.233157393
GAS_ast	0.908442445	0.078594056	0.473859128	0.133320187
GAS_ast1	0.813564517	**0.030058131**	0.388086401	0.097443577
NAGARCH-std	0.973883693	0.169627481	0.498955353	0.374560586
NAGARCH-ged	0.973878411	0.169627481	0.411836075	0.233157393
NAGARCH-sstd	0.973219987	0.169627481	0.465545869	0.298544544
NAGARCH-jsu	0.973083764	0.169627481	0.421015492	0.233157393
ALLGARCH-std	0.974403702	0.313557231	0.729848644	0.298544544
ALLGARCH-ged	0.954378197	0.313557231	0.301027425	0.069544257
ALLGARCH-sstd	0.946105074	0.169627481	0.655471473	0.233157393
ALLGARCH-jsu	0.863984916	0.078594056	0.664664525	0.178266318
sGARCH-std	0.970971173	0.169627481	0.280698530	0.374560586
sGARCH-ged	0.971456856	0.169627481	0.430119659	0.233157393
sGARCH-sstd	0.919705985	0.078594056	0.47932248	0.298544544
sGARCH-jsu	0.920309596	0.078594056	0.480611008	0.298544544
csGARCH-std	0.988760696	0.313557231	0.276585386	0.374560586
csGARCH-ged	0.995774603	0.510159049	0.413080995	0.233157393
csGARCH-sstd	0.991342439	0.313557231	0.463718361	0.298544544
csGARCH-jsu	0.968752084	0.169627481	0.471645743	0.374560586
apARCH-std	0.978126738	0.313557231	0.454559787	0.233157393
apARCH-ged	0.946446922	0.169627481	0.153497711	0.048421047
apARCH-sstd	0.950309417	0.169627481	0.390367981	0.178266318
apARCH-jsu	0.950498412	0.169627481	0.418800694	0.133320187
gjrGARCH-std	0.914422326	0.078594056	0.38111328	0.374560586
gjrGARCH-ged	0.915627211	0.078594056	5.58E-01	1.78E-01
gjrGARCH-sstd	0.919034551	0.078594056	0.383051305	0.374560586
gjrGARCH-jsu	0.919527733	0.078594056	0.343094941	0.298544544
iGARCH-std	9.71E-01	1.70E-01	2.81E-01	0.374560586
iGARCH-ged	0.971302735	0.169627481	0.430213305	0.233157393
iGARCH-sstd	0.91947743	0.078594056	0.480011167	0.298544544
iGARCH-jsu	0.919998682	0.078594056	0.480975992	0.298544544

The table reports the *p*–values of the conditional coverage (CC) test of Christoffersen (1998) and the dynamic quantile test (DQ) for the one-day ahead 5% and 1% VaR. Models rejected for the null hypothesis and under the significance level 5% are in boldface quantile. Overall, the results of these three Back-testing tests have confirmed the results of MCS procedure that support the accuracy of GAS-sstd and gjr-GARCH-ged to forecast EUA VaRs at 1% and 5% quantiles, respectively. Our research seems in line with some previous that support the precision of some GARCH types to predict and forecast EUA VaR results (Paolella & Taschini, 2008; Seifert et al., 2008; Daskalakis et al., 2009; Benz & Trück, 2009; Hitzemann and Uhrig-Homburg, 2013; Gil-Alana, 2016; Müller et al., 1997). On the other hand, our results appear more sophisticated than results of previous studies because they are not directed by phase data.

6 Conclusions

This study investigates the volatility of European Union Allowances (EUAs) or CO_2 emission permits for the period from 22/04/2005 to 28/02/2019, which has importance in the evolution of energy markets and for the advance of new risk management approaches. Particularly, we focus on the predictive and forecastability of various GARCH and GAS models for emission permit daily returns under different error distributions. We have also used the MCS procedure to obtain the set of models with superior predictive ability of emission permit VaR forecasts. For the estimation of daily returns, we show that the APARCH model with skewed-t-distribution outperforms all other models. This model is capable of capturing the fat tails, excess kurtosis and leverage effects characterizing the UEAs data. However, to forecast a 1% and 5% -VaR we find that a simple model such as the gjr-GARCH and the GAS-sstd performs excellent results, respectively. These results are also confirmed by three backtesting tests (i.e., Unconditional coverage, Conditional coverage, and dynamic quantile tests). Thus, we conclude that it is feasible to discriminate between the estimation methods based on an analysis of the VaR forecast accuracy.

Since modelling volatility is linked to portfolio optimization, hedging decisions and derivative pricing models, our findings can be helpful to investors and policymakers while developing the risk management strategies. Specifically, investors could forecast future EUA volatility more precisely, as our results suggest more sophisticated models under different error distributions. In addition, investors holding positions in the emission market could make proper asset-allocation and hedging decisions by understanding the effect of extreme events on the emission permit VaR measure. Policymakers often need to formulate effective hedging strategies to avoid the adverse impact of extreme events. A more accurate estimate of volatility is essential for implementing such strategies. The results of this study may help in this regard. Finally, it is important to note that the shift from fossil fuels to green power will take decades. As mentioned earlier, the EU Commission plans to achieve a zero net carbon footprint by 2050. So, in the nearest future, it is very important for the CO_2 emitters or EUA buyers to do the pricing of the emission allowances as part of their strategies. As Ciesielska-Maciągowska et al. (2021), the accurate forecast of the CO_2 prices is one of the most important things in large scale utilities that have exposure to CO_2 prices.

Although many types of GARCH and heavy-tailed distributions have been captured in the study, future research could also compare the performance of an increasing line of research based on Complex Networks and Machine learning regression models

(Hamrani et al., 2020; Xu et al., 2020, among others) over already used models, and see if such models are better in explaining the predictive ability.

Acknowledgements The authors thank the anonymous reviewers for their valuable comments, which helped us to considerably improve the content, quality and presentation of this paper.

Funding No funds, grants, or other support was received.

Declarations

Conflicts of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

Consent to Participate All the authors listed have approved the manuscript that is enclosed.

Consent for Publication The manuscript is approved by all authors for publication.

Appendix 1

Error Term Assumptions

Let $x_1, \ldots, x_n$ be independent and identically distributed as $N(\mu, \sigma^2)$, i.e. this is a sample of size n from a normally distributed population with expected mean value μ and varianceσ^2. The table below regroup the probability density function for each category:

		The probability density function	Parameters
Normal (N)	Symmetric (mesokurtic)	$f_x\left(x;\mu,\sigma^2\right)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	Where μ is the mean and σ^2 the variance
Student t- (std)	Symmetric	$f_x(x;\upsilon)=\frac{\Gamma\left(\frac{\upsilon+1}{2}\right)}{\sqrt{\pi(\upsilon-2)}\Gamma\left(\frac{\upsilon}{2}\right)}\left(1+\frac{x^2}{(\upsilon-2)}\right)^{\frac{\upsilon+1}{2}}$	With $\upsilon>2$ degrees of freedom, unit variance and Γ gamma function
Skewed-t-Student (Sstd)	Asymmetric	$f_x(x;\upsilon,\lambda;a,b,c)=$ $\begin{cases} bc\left(1+\frac{1}{\upsilon-2}\left(\frac{bx+a}{1-\lambda}\right)^2\right)^{-\frac{\upsilon+1}{2}}, & x<-a/b \\ c\left(1+\frac{1}{\upsilon-2}\left(\frac{bx+a}{1+\lambda}\right)^2\right)^{-\frac{\upsilon+1}{2}} & ,bx\geq -a/b \end{cases}$	With mean zero and unit variance $2<\upsilon<\infty$ and $-1<\lambda<1$. The constants a, b and c are defined as: $a=4\lambda c\left(\frac{\nu-2}{\nu-1}\right)$, $b^2=1+3\lambda^2-a^2$, and $c=\frac{\Gamma(\frac{\upsilon+1}{2})}{\sqrt{\pi(\upsilon-2)}\Gamma\left(\frac{\upsilon}{2}\right)}$.
Johnson's Reparametrized SU (jsu)		$f(t)=\frac{\delta}{\lambda\sqrt{1+\left(\frac{t-\zeta}{\lambda}\right)^2}}\phi\left[\gamma+\delta\sinh^{-1}\left(\frac{t-\zeta}{\lambda}\right)\right]$	Where ϕ is the density function of N(0,1), ζ and λ both are positive as well as location-scale parameters respectively, γ and δ can be interpreted as a skewness and kurtosis parameters
Generalized Error Distribution (ged)		$f_x(x;\upsilon,\lambda)=\frac{\upsilon}{\lambda 2^{\left(\frac{\upsilon+1}{\upsilon}\right)}\Gamma\left(\frac{1}{\upsilon}\right)}e^{-\frac{1}{2}\left\lvert\frac{x}{\lambda}\right\rvert^{\upsilon}}$	Where $\lambda=\left[\frac{2^{-\frac{2}{\upsilon}}\Gamma\left(\frac{1}{\upsilon}\right)}{\Gamma\left(\frac{1}{\upsilon}\right)}\right]^{\frac{1}{2}}$ If $\upsilon=2$, the GED is a standard normal distribution, whereas $\upsilon<2$ yields a leptokurtic distribution and $\upsilon>2$ yields a platykurtic distribution

	The probability density function	Parameters
AST distribution	$f_{SST}(y;\alpha,\upsilon,\mu,\sigma) = \begin{cases} \frac{1}{\sigma}K(\upsilon)\left[1+\frac{1}{\upsilon}\left(\frac{y-\mu}{2\alpha\sigma}\right)^2\right]^{-(\upsilon+1)/2}, & y \le \mu \\ \frac{1}{\sigma}K(\upsilon)\left[1+\frac{1}{\upsilon}\left(\frac{y-\mu}{2(1-\alpha)\sigma}\right)^2\right]^{-(\upsilon+1)/2}, & y > \mu \end{cases}$	With $\alpha = 1/\left(1+\gamma^2\right)$ and $\sigma = \frac{\frac{\gamma+1}{\gamma\sigma'}}{2}$, SST will become that of Fernandez and Steel (1998); letting $\alpha = \frac{1-\lambda}{2}, \sigma = \frac{1}{b}\sqrt{(\upsilon - 2/\upsilon)}$ And $\mu = -a/b$, the SST will be that of Hansen 1994). With $\alpha = 1/2$, the SST reduces to the general form of Student-t distribution. The SST density is skewed to the right for $\alpha < 1/2$ and to the left for $\alpha > \frac{1}{2}$.

References

Alberola, E., Chevallier, J., & Chèze, B. (2008a). Price drivers and structural breaks in European carbon prices 2005–07. *Energy Policy, 36*(2), 787–797.

Alberola, E., Chevallier, J., & Chèze, B. (2008b). The EU emissions trading scheme: the effects of industrial production and CO2 emissions on European carbon prices. *International Economics, 116*, 93–126.

Angeles Carner, M., Olmo, J., & Pascual, L. (2018). Modelling the dynamics of fuel and eu allowance prices during phase 3 of the EU ETS. *Energies, 11*, 3148. https://doi.org/10.3390/en11113148

Bao, H. (2013). *Analysis of carbon market pricebehaviour*. Universiteit van Amsterdam.

Barndorff-Nielsen, O. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society of London, 353*, 401–419.

Barndorff-Nielsen, O. E. (1978). *Information and exponential families in statistical theory. Wiley series in probability and mathematical statistics*. New York: Wiley.

Benschopa, T., & LÃ³pez Cabrera, B. (2014). Volatility modelling of CO2 emission allowance spot prices with regime-switching GARCH models, No SFB649DP2014-050, SFB 649 discussion papers, Humboldt University, Collaborative Research Center 649, https://EconPapers.repec.org/RePEc:hum:wpaper:sfb649dp2014-050.

Benz, E., & Trück, S. (2006). CO_2 emission allowances trading in Europe-specifying a new class of assets. *Problems and Perspectives in Management, 3*(3), 30–40.

Benz, E., & Trück, S. (2009). Modelling the price dynamics of CO2 emission allowances. *Energy Economics, 31*, 4–15. https://doi.org/10.1016/j.eneco.2008.07.003.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, Elsevier, 31*(3), 307–327.

Bredin, D., & Muckley, C. (2011). The price forming process in energy markets. In A. Dorsman, W. Westerman, M. Karan, & Ö. Arslan (Eds.), *Financial aspects in energySpringer* (pp. 85–107). New York.

Carnero, M. A., Olmo, J., & Pascual, L. (2018). Modelling the dynamics of fuel and EU allowance prices during phase 3 of the EU ETS. *Energies*, *11*(11), 3148. https://doi.org/10.3390/en11113148.

Chevallier, J. (2009). Carbon futures and macroeconomic risk factors: A view from the EU ETS. *Energy Economics, 31*(4), 614–625.

Chevallier, J. (2012). CO_2 price fundamentals. In Econometric analysis of carbon markets. Springer

Christoffersen, P. (1998). Evaluating interval forecasts. *International Economic Review, 39*(4), 841–862.

Ciesielska-Maciągowska, D., Klimczak, D., & Skrzek-Lubasińska, M. (2021). Central and eastern European CO2 market—challenges of emissions trading for energy companies. *Energies, 14*, 1051. https://doi.org/10.3390/en14041051

Creti, A., Jouvet, P. A., & Mignon, V. (2012). Carbon price drivers: Phase 1 versus phase 2 equilibrium? *Energy Economics., 34*, 327–334.

Daskalakis, G., Psychoyios, D., & Markellos, R. N. (2009). Modelling CO_2 emission allowance prices and derivatives : Evidence from the European trading scheme. *Journal of Banking and Finance, 33*(7), 1230–1241.

Ding, Z., Granger, C. W. J., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance, 1*(1), 83–106.

Engle, R. F., & Bollerslev, T. (1986). Modelling the persistence of conditional variance. *Econometric Reviews, 5*, 1–50.

Engle, R. F., & Lee, G. G. J. (1999). A permanent and transitory component model of stock return volatility. In R. F. Engle & H. White (Eds.), *Cointegration, causality, and forecasting: A festschrift in Honor of Clive W.J. Granger.* Oxford University Press.

Engle, R. F., & Manganelli, S. (2004). CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. *Journal of Business & Economics Statistics, 22*(4), 367–381.

Engle, R. F., & Ng, V. K. (1993). Measuring and Testing the Impact of News on Volatility. *Journal of Finance, 48*, 1749–1778. https://doi.org/10.1111/j.1540-6261.1993.tb05127.x

Fezzi, C., & Bunn, D. W. (2009). Structural interactions of European carbon trading and energy prices. *Journal of Energy Markets., 2*, 53–69.

Giacomini, R., & White, H. (2006). Tests of conditional predictive ability. *Econometrica, 74*, 1545–1578.

Gil-Alana, L. A., Gupta, R., & Perez De Gracia, F. (2016). Modelling persistence of carbon emissions allowance prices. *Renewable and Sustainable Energy Reviews*, *55*, 221–226.

Glosten, L., Jagannathan, R., & Runkle, D. (1993). Relationship between the expected value and volatility of the nominal excess returns on stocks. *Journal of Finance, 48*, 1779–1802.

Gneiting, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association, 106*, 746–762.

Gonzalez-Rivera, G., Lee, T. H., & Mishra, S. (2004). Forecasting volatility: A reality check based on option pricing, utility function, value-at-risk and predictive likelihood. *International Journal of Forecasting, 20*, 629–645.

Hamrani, A., Akbarzadeh, A., & Madramootoo, C. A. (2020). Machine learning for predicting greenhouse gas emissions from agricultural soils. *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2020.140338 PMID: 32610233.

Hansen, P. R. (2005). A test for superior predictive ability. *Journal of Business & Economic Statistics, 23*, 365–380.

Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The model confidence set. *Econometrica, 79*(2), 453–497.

Hansen, P. R. (2003). Asymptotic tests of composite hypotheses. http://www.stanford.edu/people/peter.hansen.

Harvey, A. C. (2013). *Dynamic models for volatility and heavy tails*. Cambridge University Press.

Hentschel, H. (1995). All in the family: Nesting symmetric and asymmetric GARCH models. *Journal of Financial Economics, 39*, 71–104.

Higgins, M. L., & Bera, A. K. (1992a). A class of nonlinear ARCH models. *International Economic Review, 33*(1), 137–158.

Hintermann, B. (2010). Allowance price drivers in the first phase of the EU ETS. *Journal of Environmental Economics and Management, 59*, 43–56.

Hintermann, B., Peterson, S., & Rickels, W. (2016). Price and market behavior in phase II of the EU ETS: a review of the literature. *Review of Environment Economic Policy, 10*(1), 108–128.

Hitzemann, S., Uhrig-Homburg, M. (1993). Empirical performance of reduced-form models for emission permit prices. In Winner of the energy & finance conference best paper Award 2013).

Johnson, N. L. (1949). System of frequency curves generated by translation. *Biometrika, 36*, 149–176.

Kostrzewski, M. (2019). The Bayesian Methods of Jump Detection: The Example of Gas and EUA Contract Prices. *Central European Journal of Economic Modelling and Econometrics*, 107–131.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives, 3*, 73–84. https://doi.org/10.3905/jod.1995.407942

Mansanet-Bataller, M., Pardo, A., & Valor, E. (2007). CO2 prices, energy and weather. *The Energy Journal, 28*, 73–92.

Maydybura, A., & Andrew, B. (2011). A study of the determinants of emissions unit allowance price in the European union emissions trading scheme. *Australasian Accounting Business and Finance Journal, 5*(4), 123–142.

McAleer, M., & Hafner, C. M. (2014). A one line derivation of EGARCH. *Econometrics, 2*, 92–97.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics, 3*, 125–144. https://doi.org/10.1016/0304-405X(76)90022-2

Müller, U. A., Dacorogna, M. M., Davé, R. D., Olsen, R. B., Pictet, O. V., & Weizsäcker, J. V. (1997). Volatilities of different time resolutions—analyzing the dynamics of market components. *Journal of Empirical Finance, 4*(2/3), 213–239.

Nazifi, F., & Milunovich, G. (2010). Measuring the impact of carbon allowance trading on energy prices. *Energy Environ, 21*(5), 367–383.

Paolella, M. S., & Taschini, L. (2008). An econometric analysis of emission allowance prices. *Journal of Banking and Finance, 32*(10), 2022–2032.

Rickels, W., Gorlich, D., & Peterson, S. (2015). Explaining European emission allowance price dynamics: Evidence from phase II. *Germany Economic Review, 16*, 181–202.

Romano, J. P., & Wolf, M. (2005). Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association, 100*(469), 94–108.

Seifert, J., Uhrig-Homburg, M., & Wagner, M. (2008). Dynamic behavior of CO2 spot prices. *Journal of Environmental Economics and Management*, *56*(2), 180–194.

Song, Y., & Liang, D. (2018). *How to discriminate market conditions based on its price jump information? A result in EU ETS* (pp. 77–81). IEEE.

Wang, Y., & Guo, Z. (2018). The dynamic spillover between carbon and energy markets: New evidence. *Energy, 149*, 24–33. https://doi.org/10.1016/j.energy.2018.01.145

White, H. (2000). A reality check for data snooping. *Econometrica, 68*, 1097–1127.

Xu, H., Wang, M., Jiang, S., & Yang, W. (2020). Carbon price forecasting with complex network and extreme learning machine. *Physica a: Statistical Mechanics and Its Applications*. https://doi.org/10.1016/j.physa.2019.122830
Yang, L., Li, F., & Zhang, X. (2016). Chinese companies' awareness and perceptions of the Emissions Trading Scheme (ETS): Evidence from a national survey in China. *Energy Policy*, *98*, 254–265. https://doi.org/10.1016/j.enpol.2016.08.039.
Zakoian, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control.*, *18*(5), 931–955.
Zhang, Y. (2016). Research on carbon emission trading mechanisms: Current status and future possibilities. *International Journal Global Energy Issues, 39*, 89–107.
Zhang, Y. J., & Sun, Y. F. (2016). The dynamic volatility spillover between European carbon trading market and fossil energy market. *Journal of Cleaner Production, 112*, 2654–2663.