Energy Economics 121 (2023) 106625

Contents lists available at ScienceDirect

Energy Economics

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

Evaluation of Cost-at-Risk related to the procurement of resources in the ancillary services market. The case of the Italian electricity market

Francesco Lisi [a,b,*], Luigi Grossi [a,b], Federico Quaglia [c]

[a] Department of Statistical Sciences, University of Padua, Italy
[b] Interdepartmental Centre "Giorgio Levi Cases" for Energy Economics and Technology, University of Padua, Italy
[c] Terna Rete Italia S.p.A., Italy

ARTICLE INFO

JEL classification:
C01
C14
C22
D81
G32
Q4

Keywords:
Cost-at-risk (caR)
Electricity market
Ancillary services
GAM models
Q-GAM models
GARCH models
Quantile regression

ABSTRACT

Measuring the risk exposure of TSOs on the dispatching market is a crucial task for the correct management of liberalized electricity markets. To fill a gap in the literature, the notion of Cost-at-Risk (CaR) is defined in the context of the dispatching market. Moreover, we propose a set of semi-parametric and non-parametric models for the estimation of the Cost at Risk (CaR) for the Italian TSO (Terna) and evaluate the corresponding out-of-sample forecasting performance. The empirical analysis relies on a rich hourly dataset provided by Terna, including several costs' drivers. The results, in terms of 1-day and 30-day ahead predictions, suggest that the model with the globally best performance is the semi-parametric GAM-GARCH model.

1. Introduction

The dispatching market (in Italian Mercato dei Servizi di Dispacciamento, MSD) is the place where Terna, the Italian Transmission System Operator (TSO), procures the resources needed to manage and control the system, generically called "ancillary services" (Kaushal and Van Hertem, 2019; Lobato Miguélez et al., 2008). To obtain the needed resources, Terna accepts sell and buy offers submitted by market participants, acting as central counterpart. Costs borne while getting ancillary services are relevant (Li and Ho, 2022; Graf et al., 2021; Ghiani et al., 2020; Graf et al., 2020; Liu and Wu, 2007) and are charged onto the end-consumers.

In the last years, the increasing penetration of renewable energy resources has led to a greater uncertainty about the level of power production and, hence, to a growing need of flexibility and of ancillary services (Lamadrid and Mount, 2012; Godoy-González et al., 2020). In turn, this has led to an increasing cost risk when operating in the MSD market.

Thus, measuring the risk exposure connected to the costs paid on the MSD market becomes crucial (Falvo et al., 2022) for TSOs. Identifying in advance potential risks in terms of incurred costs allows the TSO to activate in advance proper countermeasures (if available) to mitigate the risk in the short term. Risk models, measuring the effect of each single variable on the expected costs, support also the TSO in setting the priority list of mid-term actions and to perform sensitivity/scenario analyses. Despite its importance, the issue of risk assessment for ancillary services costs has not been adequately considered in the literature and only very few works have been produced with the aim of predicting expected costs/prices incurred in spot ancillary service markets (Hadzic and Bisanovic, 2019). This may be also related to the heterogeneous design adopted for these markets across the world. To fill the gap, in this work we are going to investigate how to model costs dynamics and evaluate the cost risk related to a market-based ancillary services procurement by an electricity TSO.

The first contribution of this paper is the extension of the idea of the Value at Risk (VaR), a well-known risk measure used in the literature on financial markets, to the risk analysis connected to costs borne on the MSD. To this aim, we re-formulate the idea of VaR focusing on costs. The VaR, introduced in the early 1990s, is typically used to quantify the

* Corresponding author at: Department of Statistical Sciences, University of Padua, Italy.
E-mail address: francesco.lisi@unipd.it (F. Lisi).

https://doi.org/10.1016/j.eneco.2023.106625
Received 17 October 2022; Received in revised form 1 March 2023; Accepted 7 March 2023
Available online 21 March 2023
0140-9883/

risk related to volatility and fluctuations of prices. It has been widely used in finance (Jorion, 2006; Alexander, 2009) and in energy finance, the area of energy markets analysis looking at energy products/markets from a financial perspective (Sadeghi and Shavvalpour, 2006; Fiorenzani, 2006; Chiu et al., 2010; Laporta et al., 2018). In analogy with the VaR, we define the Cost-at-Risk ($CaR_{t,\alpha,h}$) as the maximum amount of money expected to be borne over a given horizon (h), with a given probability (α), under normal market conditions. For example, if for a one-day ahead horizon we have $CaR_{t,\alpha,1} = 1$ million, costs in the next day will not exceed 1 million with a $(1-\alpha)$ probability. CaR can also be defined as the $\alpha-$th percentile of the $h-$day costs distribution.

Costs borne in the MSD may depend on the kind of services bought, as well as on several other factors, including the renewable power production, the available level of the ramping capability of thermal power plants, possible calendar effects, the price of gas, etc.

As a further contribution, we propose a set of models able to capture the main empirical characteristics of costs and to select among them the best specification. The dynamics of costs and corresponding CaR are described in function of a vector of covariates $\boldsymbol{X}_t$. In particular, CaR is defined as the quantile of the costs C_t conditioned to the covariates $\boldsymbol{X}_t$. We will focus on h = 1-day and h = 30-day Car at level $\alpha = 10\%$. This probability level has been set in agreement with the TSO, who considers it adequate because the interest is centered on high, but not extreme, level of costs.

The distinctive feature of the procedure we are going to introduce is that, differently from the usual context where the VaR is applied, costs cannot be assumed to be generated by a null conditional mean process, as in the case of financial returns. This implies that the conditional mean equation of costs time series must be modeled and all its characteristics have to be accounted for.

In the following sections, we consider five classes of CaR models and analyze their performances in the conditional CaR estimation. Four models are based on a two-step procedure involving a non-parametric GAM model for the conditional mean of costs and four different approaches for the conditional quantile of the residuals and, hence, of CaR. Conditional quantiles of residuals are estimated using (i) non-parametric kernel density, estimated under the assumption of homoscedasticity; (ii) a semi-parametric GARCH model; (iii) a linear parametric[1] quantile regression model and (iv) a the parametric CAViaR model. The fifth model describes the dynamics of CaR using a single-step approach based on a Q-GAM model and is fully nonparametric.

Models are estimated and compared using the time series of the costs incurred in by Terna during the time interval 2017–2021. A final empirical contribution is connected to the introduction of relevant exogenous regressors that can be reasonably considered as possible costs' drivers. The set of exogenous regressors contains: actual consumptions, levels of reserve requirements, network nodal and zonal constraints, as well as prices of the main commodities affecting the Italian electricity market (namely, natural gas and carbon dioxide). Calendar variables are also included to account for periodic and other deterministic effects.

The rest of the paper is organized as follows: Section 2 contains a short introduction to the Italian electricity market, a description of the dataset and some descriptive features of the data; in Section 3, we outline the general form of CaR models used in the work; Section 4 is devoted to the in-sample models estimation and out-of-sample forecasting performance in the case of 1-day-ahead CaR; results referring to the 30-day-ahead CaR are described in Section 5. Section 6 concludes.

[1] Note that, even if the quantile representation is parametric, no distributional assumptions are required on the error terms.

2. The Italian market of ancillary services and data description

The Italian electricity market is the place where transactions involving electricity are conducted and consists of two main segments: the spot market and the forward market. In the former, products are traded for immediate or day-ahead delivery while, in the latter, future deliveries and withdrawals are negotiated. In turn, the spot market consists of an energy market and a dispatching market. The energy market includes a day-ahead market (in Italian, Mercato del Giorno Prima, MGP) and an intra-day market (ID) where producers, wholesalers and end customers buy and sell wholesale quantities of electricity for the next day. Although the injections and withdrawals schedules accepted in the MGP account for interconnection capacity among market zones, they do not consider intra-zonal congestions or any other network security constraint which could make the physical delivery of energy unfeasible. Moreover, schedules defined in the MGP represent only a first schedule of actual injections and withdrawals occurring the next day, which can be affected by demand and generation forecast uncertainties and unplanned outages (e.g. fault) of generation units. More specifically, demand and generation forecast is mainly related to weather conditions which, in turn, can impact on the demand level as well as on intermittent renewable energy sources, such as wind and solar energy. As mentioned in the introduction, power systems must comply with several technical constraints in order to work correctly and safely; to this purpose, law provides for an institutional subject called TSO (in Italy Terna), who operates in the MSD and takes care that the balance between injections and withdrawals, as well as all other security criteria, are always satisfied. Terna (Caprabianca et al., 2020) operates within the MSD in order to (i)balance injections and withdrawals in real-time keeping the system frequency within the security ranges, accepting balancing energy offers; (ii) procure reserve capacity, ensuring the availability of balancing resources; (iii) relieve congestions, ensuring that each element of the power system operates inside its security limits.

Both MGP and MSD generate massive datasets that could help operators to understand the underlying processes. In this paper, we collect a rich dataset in order to estimate reliable CaR models. Originally, data have hourly frequency but, since we are interested in $h-$day ahead CaR, we move to daily data by (algebraically) summing them over the 24 h.

Data cover the period January 1, 2017–September 20, 2021 and, for this period, the following time series are available for each day t ($t = 1, \dots, 1826$):

- C_t: total daily costs incurred in by Terna in the dispatching market;
- D_t: daily national demand for energy (in MWh);
- $Wind_t$, PV_t, $Hydro_t$: time series of daily wind, photo-voltaic, hydro (excluding pumping units) energy production (in MWh);
- $aFRR_t$: national needs of automatic frequency restoration reserve (in MWh), also called secondary reserve. The secondary control is automatic and based on the secondary reserve provided by generators connected to the grid which have to vary their power supply in order to restore the nominal value of the frequency after any deviation. The service must be completely delivered within 180 s, so only some generators can provide it. In Europe, the secondary reserve is called "automatic Frequency Restoration Reserve" (aFRR) and is defined as the active power reserves, available to re-establish the frequency to the rated value and, for synchronous areas comprising more than one Load-Frequency Control area, to bring back the power balance to the scheduled value.
- RR_t: sum of the "manual frequency restoration reserve" and of the "replacement reserve" (in MWh). This sum is also called tertiary reserve. It is the active power reserve available to restore and support the required level of FRR and it must be prepared for further system imbalances and unexpected events, including generation reserve. There are two types of tertiary control reserve: (i)

Spinning tertiary control reserve, fully delivered within 15 min, in order to restore the secondary reserve. It can be activated manually and it is the same as the European "manual Frequency Restoration Reserve" (mFRR).
(ii) Replacement tertiary control reserve, fully delivered within 120 min and necessary to restore the tertiary reserve against shifts in demand, injection from renewable sources, long-lasting faults of power plants. It corresponds to the European "Replacement Reserve" (RR).
- VRI_t: number of generation units with must run constraints to provide system services in order to ensure voltage control and stability of the power system.
- TTF_t: price of gas at the TTF index (in Euros);
- ETS_t: cost of the green certificates (in Euros) according to the EU Emissions Trading System (ETS).

In addition, for each t, ($t = 1, \dots, n$), the following calendar variables are used:

- T_t: it represents the "trend" variable or the long-run dynamics of the costs at day t;
- DY_t: day of the year; it represents the yearly periodicity of the data and is described by a vector repeating the sequence $1, 2, \dots, 365$ (366for leap years). For 2021 the sequence stops at $t = 273$;
- DW_t: day of the week; it represents the weekly periodicity of the data and is described by repeating the periodic sequence $1, \dots, 7$;
- $bank_t$: a dummy variable accounting for bank holidays. It assumes value 1 if day t is an Italian bank holiday, 0 otherwise.

In this work, C_t is the response variable we want to model, while all other variables define the set $\boldsymbol{X}_t$ (or, equivalently, $\boldsymbol{I}_t = \{\boldsymbol{X}_s,\ s \le t\}$), describing the information available at time t, with respect to the predictions to which they are conditioned.

Fig. 1, panel in position (1,1), shows the series of costs in million euros and some calendar effects present in its dynamics. Real costs are described by negative values which mean an overall payment from the consumers to the producers, while positive values represent an overall payment from the generators to the consumers. Indeed, in the ancillary services market, generators are paid to increase their grid injection infeed, while they pay back money for decreasing their infeed from the scheduled value (since they are saving, at least, their short-run marginal cost of production). In the observed period the daily average cost is around 5,392,000 Euros with a standard deviation of 3,000,000 Euros. Average daily costs for each month (panel (1,2)) and for each day of the week (panel (2,1)) clearly suggest the presence of an yearly periodic component and a weekly periodic component. Note that, in the weekend, costs are higher than in the working days. The explanation is that average daily costs for bank/no-bank holidays define the impact of a bank holiday. These findings result in the need to include in the model some calendar variables to model the structural components.

It is important to take note that true costs assume negative values, so that smaller (negative) values denote higher costs. Table 1 lists the main descriptive statistics of all variables. The daily average amount of costs borne by Terna for ancillary services is around five million euros and this explains the importance of controlling such source of risk.

3. CaR models

This section is devoted to find a model that suitably describes the dynamics of CaR_t, defined as the quantile of the costs C_t conditionally to a set of covariates $\boldsymbol{X}_t$

$$CaR_{t,\alpha,h} = q^C_{t,\alpha,h}(\mathbf{X}_t).$$

To this end, we are going to identify and estimate five different competitive models and compare their performance with respect to CaR.[2]

The first four models are semi-parametric and are based on a two-step procedure: first, the conditional mean is estimated, and then, using the residuals from the mean, the (dynamic) quantile of residuals is estimated and added to the conditional mean to produce the CaR_t. The fifth model, otherwise, is fully nonparametric and estimates CaR in a single step.

The general specification of the two-step models is given by

$$C_t = \mu\left(\boldsymbol{X}_t\right) + \varepsilon_t \tag{1}$$

where $\mu\left(\boldsymbol{X}_t\right) \equiv \mu_t = E(C_t|\boldsymbol{X}_t)$ is the conditional mean of costs with respect to a vector of covariates (exogenous or not) and ε_t is a random variable whose characteristics depend on the approach and will be specified case-by-case.

The four two-step models share the same specification of the conditional mean, which follows a GAM model (Hastie and Tibshirani, 1986):

$$\mu\left(\boldsymbol{X}_t\right) = \mu + \sum_{i=1}^{p} f_i(X_{i,t}). \tag{2}$$

Functions f_j are smoothers describing the expectation of the response variable, C_t, conditionally to all other regressors. They do not have a specific functional form, allowing for both linear and non-linear specifications, but they are required to be smooth, i.e. continuous together with their first and second derivatives f'_i and f''_i. To avoid problems of identifiability, it is usually assumed that $E(f_i(X_i)) = 0$, for $i = 1, 2, \dots, p$. Parameter μ is the unconditional mean.

There are different ways to represent the non-linear function $f_j(X_j)$ but in this work spline functions are used (Hastie and Tibshirani, 1986; Wood, 2006). The smooth components of the model can be estimated using the back-fitting algorithm, according to the (Hastie and Tibshirani, 1986) approach, or by penalized likelihood maximization, in which the model (negative log) likelihood is modified adding a penalty factor for each smooth function, penalizing its 'wiggliness' (Wood, 2006). In this paper we follow the second approach.

For the error term $\varepsilon_t = C_t - \mu\left(\mathbf{X}_t\right)$, we consider four different cases, based on proper assumptions:

1. homoscedasticity, i.e. a sequence of i.i.d random variables, $\varepsilon_t \sim D(0, \sigma^2)$. In this case the expression of the CaR is given by:
$$\begin{aligned} CaR_{t,\alpha,h} &= \mu\left(\boldsymbol{X}_t\right) + q^\varepsilon_\alpha \\ &= \mu + \sum_{i=1}^{p} f_i(X_{i,t}) + q^\varepsilon_\alpha \end{aligned} \tag{3}$$
where q^ε_α is the α−th quantile of ε_t.
We estimate the distribution D and, hence, q^ε_α, using nonparametric kernel methods (Silverman, 1986). For this reason, we refer to this model as GAM-K.
2. heteroscedasticity and GARCH dynamics. In this case we can write $\varepsilon_t = \sigma_t\, z_t$, where $z_t \sim iid(0, 1)$, $\varepsilon_t|\boldsymbol{I}_t \sim D(0, \sigma^2_t)$ and
$$\sigma^2_t = \gamma_0 + \gamma_1 \varepsilon^2_{t-1} + \gamma_2 \sigma^2_{t-1}.$$
The conditional distribution D depends on the distribution of z_t. Parameters γ_i can be estimated using Maximum Likelihood (ML) or Quasi-Maximum Likelihood (QML) methods.
In this approach, the conditional quantile of ε_t is time-varying and is a function of the conditional variance. The CaR for C_t is given by:
$$\begin{aligned} CaR_{t,\alpha,h} &= \mu\left(\mathbf{X}_t\right) + q^\varepsilon_{t,\alpha} \\ &= \mu + \sum_{i=1}^{p} f_i(X_{i,t}) + q^z_\alpha \sigma_t. \end{aligned} \tag{4}$$
The quantile q^z_α can be estimated assuming for z_t a parametric or a non parametric distribution without distributional assumptions.
As ε_t follows a GARCH(1,1) process, we denote the model as GAM-GARCH.

[2] Henceforth, when we refer to costs the superscript C will be omitted.

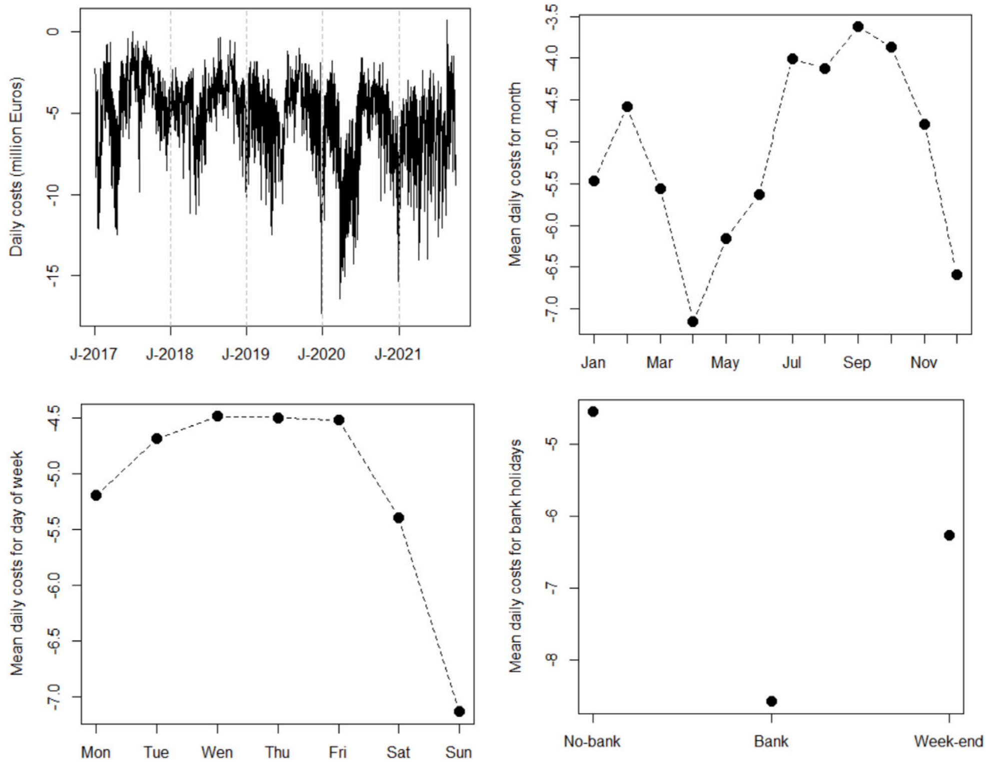

Fig. 1. Time series of daily costs in million euros (panel 1,1); average daily costs for each month (panel 1,2); average daily costs for each day of the week (panel 2,1); average daily cost for bank/no-bank holidays (panel 2,2).

Table 1
Descriptive statistics for the considered variables. Q1 and Q3 are the first and third quartiles, Sk and $Kurt$ are skewness and kurtosis coefficients, respectively. Costs are expressed in million euros, while Demand, Wind, PhotoVoltaic, Hydro, aFFR and RR in thousands of MWh.

Variable	Q1	Mean	Median	Q3	St.Dev.	Sk	$Kurt$
C_t	−6.421	−5.132	−4.564	−3.315	2.596	−1.087	4.362
D_t	695.113	786.347	813.917	872.484	117.248	−0.393	2.414
$Wind_t$	22.862	50.821	44.152	71.662	33.395	0.788	2.828
PV_t	31.693	53.502	56.729	74.318	23.132	−0.232	1.794
$Hydro_t$	88.087	119.158	114.370	149.060	41.611	0.304	2.515
$aFFR_t$	13.679	15.134	14.982	16.402	2.374	0.553	3.592
RR_t	82.312	87.918	87.185	93.158	7.835	0.363	2.953
Vri_t	515	606	584	677	127.285	0.849	3.882
TTF_t	12.112	18.043	16.750	21.575	9.618	2.376	12.865
ETS_t	11.100	22.640	22.905	26.688	13.958	0.890	3.539

3. $\varepsilon_t = \beta_{0,\alpha} + \sum_{i=1}^{m} \beta_{i,\alpha} Y_{i,t} + z_t$, where, in the general case, regressors $Y_{i,t}$ can be lagged values of ε_t or other exogenous variables. Under the further assumption[3] that $q_\alpha^z(\boldsymbol{Y}_t) = 0$, the conditional quantile of ε_t is given by

$$q_\alpha^\varepsilon(\boldsymbol{Y}_t) = \beta_{0,\alpha} + \sum_{i=1}^{m} \beta_{i,\alpha} Y_{i,t}$$

This approach corresponds to a quantile regression (QR) on ε_t and does not require distributional assumptions on z_t (Koenker, 2005). In the context of the quantile regression, estimates of parameters β_i can be obtained by solving a problem of least absolute deviation (LAD). This can be achieved using the simplex algorithm where the initial LAD problem is reformulated as a linear programming problem (Koenker and Bassett, 1978).

The related CaR for C_t is given by:

$$\begin{aligned} CaR_{t,\alpha,h} &= \mu\left(\mathbf{X}_t\right) + q_\alpha^\varepsilon(\boldsymbol{Y}_t) \\ &= \mu + \sum_{i=1}^{p} f_i(X_{i,t}) + \beta_{0,\alpha} + \sum_{i=1}^{m} \beta_{i,\alpha} Y_{i,t} \end{aligned} \quad (5)$$

We call this model GAM-QR.

4. A variant of the previous model assumes that the conditional quantile of ε_t can be written as:

$$q_{t,\alpha}^\varepsilon = \beta_{0,\alpha} + \beta_{1,\alpha} q_{\alpha,t-1}^\varepsilon + \beta_{2,\alpha}\varepsilon_{t-1}^{+} + \beta_{3,\alpha}\varepsilon_{t-1}^{-} \quad (6)$$

where ε_t^+ and ε_t^- denote positive and negative values of ε_t, respectively. The unknown parameters are estimated using Koenker and Bassett (1978) regression quantile framework. Eq. (6) corresponds to the CAViaR model proposed by Engle and Manganelli (2004): thus, the whole model is called GAM-CAViaR. For the GAM-CAViaR model, the time-varying CaR for C_t is given by

$$\begin{aligned} CaR_{t,\alpha,h} &= \mu\left(\mathbf{X}_t\right) + q_{t,\alpha}^\varepsilon \\ &= \mu + \sum_{i=1}^{p} f_i(X_{i,t}) + \beta_{0,\alpha} + \beta_{1,\alpha} q_{\alpha,t-1}^\varepsilon + \beta_{2,\alpha}\varepsilon_{t-1}^{+} + \beta_{3,\alpha}\varepsilon_{t-1}^{-}. \end{aligned}$$

[3] This is a standard assumption in the quantile regression and is the equivalent to require that the expectation of the error term of a regression model is zero.

Finally, in the fifth approach, we model the conditional quantile of costs (the CaR) in just one step. It is a generalization of the GAM model (1), introduced by Fasiolo et al. (2021), which directly refers to conditional quantiles, rather than to the conditional mean and it is known as Quantile-GAM (QGAM).
In this approach a nonparametric quantile regression is applied to C_t by assuming that

$$C_t = \beta_{0,\alpha} + \sum_{i=1}^{p} f_{i,\alpha}(X_{i,t}) + \epsilon_t. \tag{7}$$

Under the assumption that $q_{\alpha}^{\epsilon_t}(X_t) = 0$, the α−th quantile of C_t, conditionally to X_t, and coinciding with the CaR, is:

$$CaR_{t,\alpha,h} \equiv q_{t,\alpha}^{C} = \beta_{0,\alpha} + \sum_{i=1}^{p} f_{i,\alpha}(X_{i,t}) \tag{8}$$

where functions $f_{i,\alpha}$ have the same meaning as in model (1), but in this case, they (their parameters) depend on the quantile's level (Fasiolo et al., 2021).
This approach will be referred to as Q-GAM. In this model, unlike the GAM-QR models, the conditional quantile of C_t is directly modeled, while in the GAM-QR the quantile regression is applied to the residuals of the conditional mean.

4. 1-day-ahead CaR

In this section the previous general specifications are applied to our dataset for 1-day-ahead CaR computation. Their performance is evaluated and compared both in-sample and out-of-sample in order to choose the best model.

To that end, the whole dataset, ranging from January 1, 2017 to September 30, 2021 was divided into two periods: an in-sample period covering the interval January 1, 2017–September 30, 2020 and an out-of-sample period covering the interval October 1, 2020–September 30, 2021, that is the last 365 days. The in-sample set was further divided in a training set, from January 1, 2017 to September 30, 2019, used to identify and estimate models, and a validation set, from October 1, 2019 to September 30, 2020, used to evaluate the CaR performance in a dataset different from the one used for estimation.

In all analyses, we use lagged values only for costs, while we use the actual values for the exogenous variables. For calendar variables this is not a problem, because they are known at any time. For the other variables, this choice is motivated by the consideration that the Italian TSO has access to some kind of predictions for the variables used as regressors up to 2-day-ahead but, for privacy issues, they are not publicly available. The important point, however, is that the comparison is done for all models under the same conditions.

To calibrate the models, we consider a mixed strategy based on parameters significance, analysis of residuals, R^2, value of MAE in the training set and on the comparison of real versus nominal coverage in the validation set. With respect to the last point, in particular, we use several tests: (i) the Kupiec (1995) test of correct unconditional coverage; (ii) the Christoffersen (1998) test of correct conditional coverage, which jointly tests if CaR violations appear independently, and the right unconditional coverage; (iii) the Engle and Manganelli (2004) test, which considers the dynamic quantile and can be interpreted as an overall goodness-of-fit test for the estimated CaR process.

4.1. In sample analyses and results

The GAM model which best fits the in-sample data includes several calendar variables, lagged costs, hydro and wind energy production, secondary and tertiary reserve requirements and the number of units needed to ensure control and stability of the power system. The expression of the conditional mean is:

$$\mu_t = \beta_0 + f_1(T_t) + f_2(DY_t) + f_3(DW_t) + bank_t + f_4(C_{t-1}) + f_5(C_{t-7}) + f_6(Hydro_t) + f_7(Wind_t) + f_8(aFRR_t) + f_9(RR_t) + f_{10}(vri_t) \tag{9}$$

For T_t and DY_t we consider adaptive basis splines, which use a weighted penalty matrix, where the weights are allowed to vary smoothly over the range of the covariate. For the other variables, otherwise, thin plate regression basis splines are used. Fig. 2 shows the estimated effects, and their variability bands, of hydro and wind energy production, of the reserve requirements (aFRR and RR) and of the number of "must-run" units (VRI). For any specific level of a variable, a positive value means that it reduces costs, given the values of all other variables, while a negative value suggests an increasing of costs.

Results show that increasing volumes of hydro from river basins and wind energy production are associated to an increase (more negative values of) costs. The relationship with the hydro power can be explained by a seasonal effect: hydro is higher in the spring season (due to the snow melting process) when the load is lower and system constraints are typically higher.

As far as wind is concerned, interpretation is less straightforward. High wind scenarios could imply higher costs due to higher reserve requirements to cope with its uncertainty, possible congestions in the power system to be solved and a lower amount of conventional generation units online. This, in turn, could require some re-dispatching actions from the TSO in order to activate a minimum amount of units to ensure the stability of the power system.

For the analysis of the impact of renewable sources penetration on electricity market volatility see Bigerna et al. (2017). With respect to the reserve requirements, the effect of the secondary reserve is to increase the costs' expectation up to around 16 000 MWh and, after this threshold, to decrease them. As in almost 77% of days aFRR is below 16 000 MWh, most of times an increase in the secondary reserve leads to an increase in costs. Also regarding the tertiary reserve needs, we observe a parabolic relationship but with a downward concavity. The interpretation is that the need for RR and electricity demand are linked by a direct relation (more demand, more RR needs). Thus, it is reasonable to assume that low values of RR are associated to low levels of demand and, at the same time, also the number of production units active after market results is low, so that Terna has to look for other units in order to build a reserve. Finally, as expected, the relation with VRI is positive so that an increase in the number of units needed to ensure control and stability of the power system leads to an increase in costs. The curve of the relation becomes more sloped approximately beyond 650 units.[4]

Eq. (9) allows to compute $\hat{\mu}_t$ and, thus, the residuals $\hat{\epsilon} = C_t - \hat{\mu}_t$, on which second-step-models are estimated.

For the GAM-K model the marginal distribution of $\hat{\epsilon}$ is estimated by means of nonparametric gaussian kernel methods. The quantile is then computed using numerical integration.

In the GAM-GARCH approach we estimate a zero-mean GARCH(1,1) model $\hat{\epsilon}_t = \sigma_t z_t$ whose estimated conditional variance is

$$\hat{\sigma}_t^2 = (2565.5 \cdot 10^8) + 0.118\ \hat{\epsilon}_{t-1}^2 + 0.728\ \hat{\sigma}_{t-1}^2$$

The estimation is performed by QML and, to avoid distributional assumptions, the conditional quantile of $\hat{\epsilon}$ is based on the empirical quantile of $\hat{z}_t$.

The best results for the GAM-QR model are obtained by considering the estimated conditional quantile given by

$$\hat{q}_{(\hat{\epsilon},0.10,t)} = -1.27 \cdot 10^6 + 0.197\ \hat{\epsilon}_{t-1} + 0.186\ \hat{\epsilon}_{t-7} - 2.25 \cdot 10^{-7}\ mm\hat{\epsilon}_{t-1}^2,$$

where $mm\hat{\epsilon}_t^2$ is a 7-day moving average of $\hat{\epsilon}_t^2$.

Parameters have been estimated using a modified version of the Barrodale and Roberts' algorithm for l1-regression, described in detail in Koenker and d'Orey (1994).

[4] Hourly values of the VRI requirements are summed up and some VRI constraints overlap.

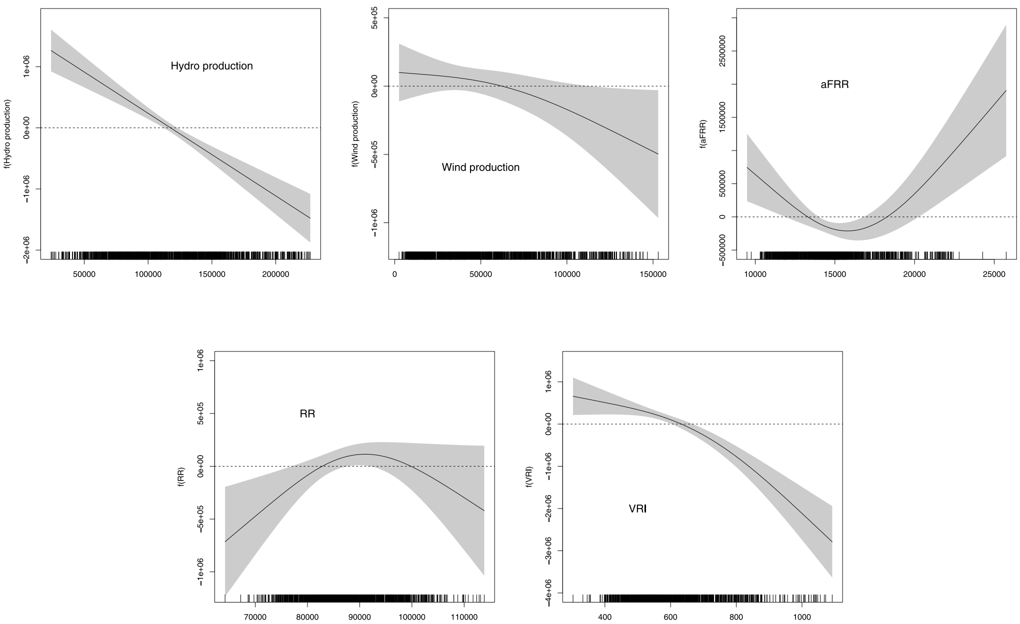

Fig. 2. Estimated effects of different variables on the conditional mean of costs. Panel (1,1): hydro energy production; panel (1,2): wind energy production; panel (1,3): secondary reserve; panel (2,1): tertiary reserve; panel (2,2): number of units in service.

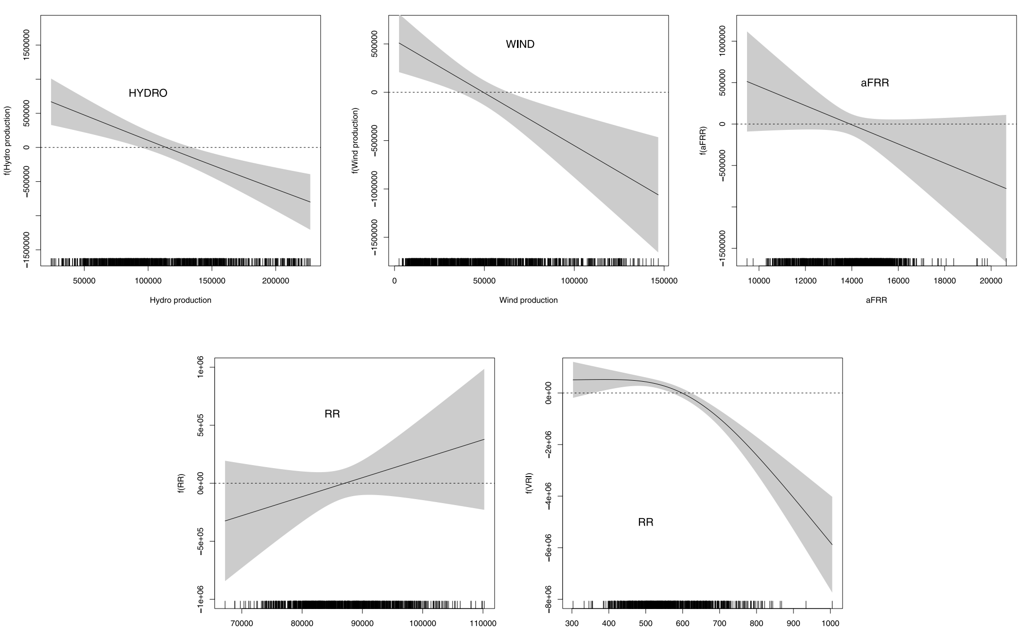

Fig. 3. Estimated effects of different variables on the 10% quantile of costs. Panel (1,1): hydro energy production; panel (1,2): wind energy production; panel (1,3): secondary reserve; panel (2,1): tertiary reserve; panel (2,2): number of units in service.

The dynamics of the conditional quantile of $\hat{\varepsilon}_t$ estimated within the GAM-CAViaR model is

$$\hat{q}_{(\hat{\varepsilon},0.10,t)} = -0.197 + 0.877\hat{q}_{(\hat{\varepsilon},0.10,t-1)} - 0.140\hat{\varepsilon}^{+}_{t-1} + 0.260\hat{\varepsilon}^{-}_{t-1},$$

where $\hat{\varepsilon}^{+}_t$ and $\hat{\varepsilon}^{-}_t$ denote positive and negative values of $\hat{\varepsilon}_t$, respectively.

All variables entering these models are significant at 5% level. Fig. 3 shows the results related to a Q-GAM specification including the same variables entering the model for the conditional mean. Even if the direction of the impact of the variables is the same as for the conditional mean, it is clear that the impact itself on the 10% conditional quantile is not the same. In particular, the graphs show that, in the Q-GAM model, aFRR and RR are not significant and the wind production is significant only for high levels of production. This implies that variables significantly impacting on the conditional mean are always significant also for a specific quantile. Furthermore, the impact of the number of activated units (VRI) is much stronger for the 10% quantile, than for the mean. The model including only the significant variables (therefore denoted by Q-GAM$_{sig}$) is

$$\begin{aligned}\hat{q}^{C}_{0.10,t} &= \beta_0 + f_1(DY_t) + f_2(DW_t) + bank_t + f_3(C_{t-1}) + f_4(C_{t-7}) \\ &\quad + f_5(Hydro_t) + f_6(Wind_t) + f_7(vri_t)\end{aligned}$$

However, it turned out that for this model the observed coverage is larger than the expected one with an observed level[5] of 6.3% versus

[5] The same results are obtained including aFRR and RR.

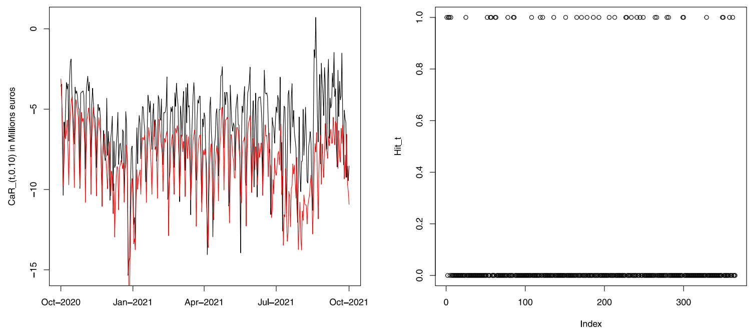

Fig. 4. Left panel: time series of daily costs (in million euros) October 1, 2020–September 30, 2021 (in black) and related CaR series (in red). Right panel: hits sequence in the out-of-sample period. $Hit_t = 1$: a CaR violation occurred at time t; $Hit_t = 0$, otherwise. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Observed level (α_{oss}) versus nominal level of 10% in the validation set (365 daily observations) and p-values of the corresponding tests for CaR models: Kupiec test (Kup), Christoffersen test (Chris) and Dynamic Quantile test (DQ).

Model	α_{oss}	Kup	Chris	DQ
GAM-K	0.0739	0.0836	0.0290	0.978
GAM-GARCH	0.1062	0.665	0.906	0.980
GAM-QR	0.0876	0.423	0.622	0.768
GAM-CAViaR	0.0876	0.423	0.305	0.977
QGAM$_{sig}$	0.0630	0.012	0.036	0.257
QGAM$_{cov}$	0.1051	0.794	0.739	0.997

Table 3
Observed level (α_{oss}) versus nominal level of 10% in the out-of-sample period (365 daily observations) and p-values of the corresponding tests for CaR models. Kupiec test (Kup), Christoffersen test (Chris) and dynamic quantile test (DQ).

Model	α_{oss}	Kup	Chris	DQ
GAM-K	0.159	<0.001	<0.001	0.331
GAM-GARCH	**0.117**	**0.347**	**0.041**	**0.998**
GAM-QR	0.139	0.024	0.007	0.989
GAM-CAViaR	0.148	0.004	<0.001	0.739
QGAM$_{sig}$	0.151	0.002	0.001	0.684
QGAM$_{cov}$	0.153	0.001	<0.001	0.772

a nominal level of 10% (see Table 2). This suggests that the model is too conservative. Thus, we looked for the Q-GAM model performing at best with respect to the in-sample CaR, at the cost of excluding some significant variable, and we found that the simple model,

$$\hat{q}^{C}_{0.10,t} = \beta_0 + f_1(DY_t) + f_2(DW_t) + bank_t + f_3(C_{t-1}) + f_4(C_{t-7}) \qquad (10)$$

denoted by Q-GAM$_{cov}$ (cov stands for coverage) leads to an in-sample observed level of 10.5%, which looks acceptable.

When these six models have been used to estimate CaR$_t$ in the validation set, they have led to the results showed in Table 2, which lists the observed level and the p-values of the three considered tests. We can see that, apart from the GAM-K and the Q-GAM$_{sig}$ models, the tests never reject the null hypothesis although most of the models tend to be too conservative. The best empirical coverages are observed for the GAM-GARCH and Q-GAM$_{cov}$ models, fitting the quantile very well.

Since excluding significant variables is an unusual strategy, we temporarily conclude that the first choice is the GAM-GARCH model, while the Q-GAM$_{cov}$ model is the second best.

4.2. Out-of-sample results

The analyses shown in the previous section refer to the in-sample period. Now we extend them to the out-of-sample period keeping fixed the models estimated in-sample. Ideally, we should consider only the in-sample "winner" model, that is the GAM-GARCH model. Nevertheless, as a further check, in Table 3 observed levels are listed for all models. Results point out that all methods, including the too conservative ones, show a number of CaR exceedances larger than the expected level. In particular, the observed coverage for the QGAM$_{cov}$ model is only 84.5% in the face of a nominal coverage of 90%. This confirms the doubts stated in Section 4.2 about this model.

The best in-sample model, the GAM-GARCH, performs quite well even out-of-sample, with an observed level of 11.7%, implying a coverage of 88.3%. *A posteriori*, among all models, it is also the one giving the most stable results. The p-value of the Christoffersen test for this model is 0.041, which could suggest possible problems connected with the independence of the hits. However, the p-value is not too far from the 5% threshold and, in addition, the CaR hits sequence (see Fig. 4) shows that hits distribute almost uniformly in the out-of-sample period. Thus, we can conclude that the GAM-GARCH model works quite well also out-of-sample.

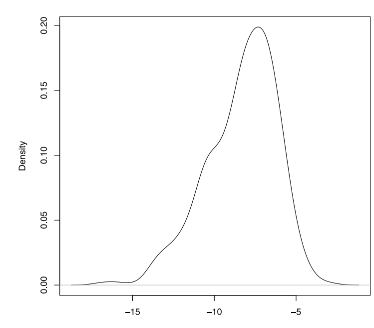

Fig. 5. Estimated distribution of $CaR_{t,0.10}$.

The above tests and criteria only consider whether the violations occur independently and in the right proportion, but do not give any piece of information about the magnitude and the variability of the CaR forecasts. To assess these features Table 4 lists the average and median CaR in the out-of-sample period, as well as the first and the third quartile of the CaR distribution, given in Fig. 5. The same table contains the value of the expected shortfall (ES), i.e. the average cost conditionally to a CaR exceedance.

5. 30-day CaR

In this section we cope with the 1-month CaR, assuming that a month is composed of 30 days. Thus, the problem is equivalent to finding the 30-day-ahead CaR.

This problem is much more complex than the 1-day-ahead CaR because, if we use cumulated costs over rolling windows of thirty consecutive days, we induce a very strong autocorrelation due to the overlapping of two of the windows over which costs are cumulated. On

Table 4
Descriptive statistics for 1-day CaR in the out-of-sample period (million euros) using the GAM-GARCH model.

Average $CaR_{t,0.10}$	Median $CaR_{t,0.10}$	$Q_{0.25}(CaR_{t,0.10})$	$Q_{0.75}(CaR_{t,0.10})$	ES
−8.176	−7.916	−9.124	−6.819	−9.752

Table 5
Observed level (α_{oss}) versus a nominal level of 10% in the out-of-sample period (336 observations) and p-values of the corresponding tests for CaR models. Kupiec test (Kup), Christoffersen test (Chris) and dynamic quantile test (DQ).

Model	α_{oss}	Kup	Chris	DQ
GAM-GARCH	0.108	0.665	<0.001	0.817

the other hand, if we consider non-overlapping windows, the number of observations scales as 30^{-1}, producing a too short time series. To bypass this problem, and since the best model for 1-day-ahead was the GAM-GARCH, we adapt the method of the filtered historical simulation (Barone-Adesi, 2015) to our context.

Similarly to the 1-day CaR, the general expression for the 30-day CaR is given by

$$CaR_{t,\alpha,30} = \hat{\mu}_{(30)|t} + q_{\alpha}^{\hat{\epsilon}(30)|t} \tag{11}$$

where, with respect to time t, $\hat{\mu}_{(30)|t}$ denotes the prediction of costs borne by the TSO over the next 30 days while $q_{\alpha}^{\epsilon(30)|t}$ is the $\alpha-$quantile of the 30-day-ahead error distribution.

To obtain the 30-day-ahead CaR we apply the following procedure:

1. for a given time t, we first get a 30-day-ahead prediction of the conditional mean for the time series of daily costs using a GAM model. Let us denote the predictions as $\hat{\mu}_{t+i}$, (i=1,2, …,30). Then we sum the daily predicted costs in order to obtain the total predicted cost over the 30 days: $\hat{\mu}_{(30)|t} = \sum_{i=1}^{30} \hat{\mu}_{t+i}$;
2. after that, we need to compute the right quantile of the 30-day-ahead error around the predicted conditional mean. To this purpose, we assume that daily residuals, ϵ_t, follow a GARCH(1,1) and:
   - using the estimated GARCH model up to time t, we simulate 1000 trajectories $\epsilon_{t+1}^{(j)}, \dots, \epsilon_{t+30}^{(j)}$, $(j = 1, \dots, 1000)$ according to the (Barone-Adesi, 2015) approach, which is described in Appendix A. In this context, each trajectory represents a sequence of costs not accounted by the conditional mean;
   - for each trajectory, we sum up the simulated values, given by $\hat{\epsilon}_{(30)|t}^{(j)} = \sum_{i=1}^{30} \hat{\epsilon}_{t+i}^{(j)}$. This provides a single realization of the 30-day-ahead prediction errors;
   - the 1000 simulated values $\hat{\epsilon}_{(30)|t}^{(j)}$ allow us to compute the $\alpha-$quantile $q_{\alpha}^{\hat{\epsilon}(30)|t}$

For the 30-day-CaR computation, we identify the best model following the same approach as for the 1-day-CaR. In this case, the model for the conditional mean which gives the best in-sample results and which was hence used out-of-sample is the same used for 1-day but without including vri_t. The rest of the procedure is the same but, clearly, with different estimated GARCH parameters.

The application of this methodology leads to the results listed in Table 5. They show a good unconditional coverage but they also point out the presence of an inadequate distribution of the hits, which highlights some problems concerning their independence. This unsatisfying output is very clearly shown in Fig. 6 and, particularly, in the right panel where we can see that, in the out-of-sample period, the hits are mainly concentrated in a window of 45 days. As for the 1-day-CaR, Fig. 6 lists the descriptive statistics for 30-day-ahead CaR using the GAM-GARCH model.

Finally, as for 1-day-Car, Table 6 contains some descriptive statistics for the 30-day-ahead CaR.

6. Conclusions

In this paper we have dealt with the issue of assessing the cost risk to which a Transmission System Operator (TSO) is exposed. To the best of our knowledge this topic has been considered only in very few works.

First, we have defined the notion of Cost-at-Risk, then we have looked for a suitable model able to evaluate CaR as a time-varying value depending on some calendar and market variables.

Five competitive models, with different features, have been estimated and their performance in CaR evaluation has been studied both in-sample and out-of-sample. Four models are based on a two-step procedure, where the conditional mean is modeled using a nonparametric GAM model and the final quantile is obtained assuming for the residuals: homoscedasticity, a GARCH dynamics and a quantile regression dynamics declined in two different ways. The fifth model, on the other hand, directly estimates, in a nonparametric way, the quantile of costs.

All these models are quite robust because they are nonparametric or work under very weak conditions. Nonparametric methods are not based on specific functional or distributional assumptions and, as they locally fit the data, can account for possible structural changes. On the other hand, also methods which are not fully nonparametric have been estimated without using strong assumptions, as well. For example, in the GAM-GARCH model, the conditional quantile of ϵ_t is based on the empirical quantile of z_t and does not assume any specific distribution. Methods based on quantile regression are also robust because they do not rely on distributional assumptions.

For 1-day-ahead CaR computation, results suggest that the GAM-GARCH model is the best one, non only because it shows the best and satisfactory out-of-sample performance, but also because it turned out to be the most stable among the considered models.

To estimate a 30-day-ahead CaR we have resorted to the historical filtered simulation method in a GAM-GARCH context. This approach, although it works well in terms of correct coverage, shows some problems with respect to the independence of the hits and, thus, it should be improved in this direction.

Thus, finding other methods for h-day-ahead computation of CaR is one of the future research targets. Further research can focus on including in the models other possible market variables and/or accounting for possible interactions between variables, for example introducing bivariate splines within GAM models. Another point for future research is to test models using predicted regressor variables, when available. In turn, this may require to find suitable methods to predict the costs drivers. Finally, it would be interesting to study the performance of these methods in different periods, for example in the last two turbulent years, and in other different markets across the world.

CRediT authorship contribution statement

Francesco Lisi: Conceptualization, Methodology, Software, Data curation, Writing – original draft, Reviewing and editing. **Luigi Grossi:** Conceptualization, Methodology, Software, Data curation, Writing – original draft, Reviewing and editing. **Federico Quaglia:** Conceptualization, Methodology, Software, Data curation, Writing – original draft, Reviewing and editing.

Acknowledgment

The authors wish to thank two anonymous reviewers and the editor for their work and their comments.

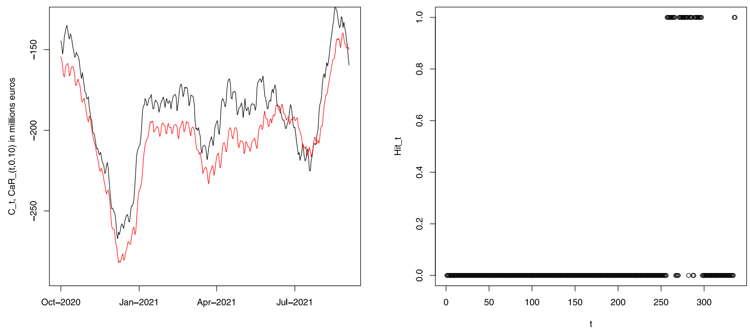

Fig. 6. Left panel: time series of 30-day daily costs (million euros), October 1, 2020–September 30, 2021 (in black) and related 30-day-CaR series (in red). Right panel: hits sequence in the out-of-sample period. $Hit_t = 1$: a CaR violation occurred at time t; $Hit_t = 0$, otherwise. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 6
Descriptive statistics for 30-day-ahead CaR using the GAM-GARCH model in the out-of-sample period (in million euros).

Average $CaR_{t,0.10}$	Median $CaR_{t,0.10}$	$Q_{0.25}(CaR_{t,0.10})$	$Q_{0.75}(CaR_{t,0.10})$	ES
−201.583	−205.004	−198.375	−212.650	−189.839

Appendix A

In this Appendix we are going to describe the procedure for the application of the filtered historical simulation used to obtain the quantile of the h step-ahead distribution, when the underlying model is a zero mean GARCH(1,1), i.e $\varepsilon_t = \sigma_t \cdot z_t$, with $\sigma_t^2 = \gamma_0 + \gamma_1 \varepsilon_{t-1}^2 + \gamma_2 \sigma_{t-1}^2$ (Barone-Adesi, 2015).

The procedure was originally proposed for the VaR computation in a financial context where the variable of interest is a financial return. It requires the following steps:

1. assume to have the time series ε_t, $t = 1, \dots, n$ and to have estimated the parameters of a GARCH(1,1) using these data;
2. at $t = n$, simulate the ε_t for times $n+1, \dots, n+h$ according to this scheme:

$$\sigma_{n+1}^2 = \hat{\gamma}_0 + \hat{\gamma}_1 \varepsilon_n^2 + \hat{\gamma}_2 \sigma_n^2 \implies \hat{\varepsilon}_{n+1} = z_1^* \sigma_{n+1}$$
$$\sigma_{n+2}^2 = \hat{\gamma}_0 + \hat{\gamma}_1 \hat{\varepsilon}_{n+1}^2 + \hat{\gamma}_2 \sigma_{n+1}^2 \implies \hat{\varepsilon}_{n+2} = z_2^* \sigma_{n+2}$$
$$\dots \quad \dots \quad \dots$$
$$\sigma_{n+h}^2 = \hat{\gamma}_0 + \hat{\gamma}_1 \hat{\varepsilon}_{n+h-1}^2 + \hat{\gamma}_2 \sigma_{n+h-1}^2 \implies \hat{\varepsilon}_{n+h} = z_h^* \sigma_{n+h}$$

3. z_i^* are sampled independently either from a parametric distribution or from the empirical distribution of $\hat{z}_t$.
4. In the original work, the value $\hat{\varepsilon}_{(30)|t} = \sum_{i=1}^{h} \hat{\varepsilon}_{n+i}$ represents the h–day-ahead return of a financial asset; in our case, it is a realization of the sum of the daily errors made in a 30-day-ahead prediction at time $t = n$
5. Iterating M times steps 1–4 leads to M simulated values $\hat{\varepsilon}_{(30)|t}^{(j)}$ which allow to find the quantile of the h-day-ahead cumulated error.

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.eneco.2023.106625.

References

Alexander, C., 2009. Value-at-Risk Models. Wiley.
Barone-Adesi, G., 2015. Simulating Security Returns: A Filtered Historical Simulation Approach. Palgrave Macmillan.
Bigerna, S., Bollino, C.A., Ciferri, D., Polinori, P., 2017. Renewables diffusion and contagion effect in Italian regional electricity markets: Assessment and policy implications. Renew. Sustain. Energy Rev. 68, 199–211.
Caprabianca, M., Falvo, M.C., Papi, L., Promutico, L., Rossetti, V., Quaglia, F., 2020. Replacement reserve for the Italian power system and electricity market. Energies 13, 2916.
Chiu, Y.-C., Chuang, I.-Y., Lai, J.-Y., 2010. The performance of composite forecast models of value-at-risk in the energy market. Energy Econ. 32, 423–431.
Christoffersen, P., 1998. Evaluating interval forecasts. Internat. Econom. Rev. 39, 841–862.
Engle, R.F., Manganelli, S., 2004. CAViaR: Conditional autoregressive value at risk by regression quantiles. J. Bus. Econom. Statist. 22, 367–381.
Falvo, M.C., Panella, S., Caprabianca, M., Quaglia, F., 2022. A review on unit commitment algorithms for the Italian electricity market. Energies 15 (1).
Fasiolo, M., Wood, S.N., Zaffran, M., Nedellec, R., Goude, Y., 2021. Fast calibrated additive quantile regression. J. Amer. Statist. Assoc. 116 (535), 1402–1412.
Fiorenzani, S., 2006. Quantitative Methods for Electricity Trading and Risk Management. Palgrave Macmillan.
Ghiani, E., Galici, M., Mureddu, M., Pilo, F., 2020. Impact on electricity consumption and market pricing of energy and ancillary services during pandemic of COVID-19 in Italy. Energies 13, 3357.
Godoy-González, D., Gil, E., Gutiérrez-Alcaraz, G., 2020. Ramping ancillary service for cost-based electricity markets with high penetration of variable renewable energy. Energy Econ. 85, 104556.
Graf, C., Quaglia, F., Wolak, F.A., 2020. Simplified Electricity Market Models with Significant Intermittent Renewable Capacity: Evidence from Italy. NBER Working Paper Series N.2726.
Graf, C., Quaglia, F., Wolak, F.A., 2021. (Machine) learning from the COVID-19 lockdown about electricity market performance with a large share of renewables. J. Environ. Econ. Manag. 105, 102398.
Hadzic, O., Bisanovic, S., 2019. Risk assessment for ancillary services. Int. J. Electr. Comput. Eng. 9, 1561–1568.
Hastie, T., Tibshirani, R., 1986. Generalized additive models. Stat. Sci. 1, 297–318.
Jorion, P., 2006. Value-at-Risk, third ed. McGrawHill.
Kaushal, A., Van Hertem, D., 2019. An overview of ancillary services and HVDC systems in European context. Energies 12, 3481.
Koenker, R., 2005. Quantile Regression. Cambridge University Press.
Koenker, R., Bassett, G., 1978. Regression quantiles. Econometrica 46, 33–50.
Koenker, R., d'Orey, J., 1994. Computing regression quantiles. Appl. Stat. 43, 410–414.
Kupiec, P., 1995. Techniques for verifying the accuracy of risk measurement models. J. Deriv. 3, 73–84.
Lamadrid, A.J., Mount, T., 2012. Ancillary services in systems with high penetrations of renewable energy sources, the case of ramping. Energy Econ. 34 (6), 1959–1971.
Laporta, A., Merlo, L., Petrella, L., 2018. Selection of value-at-risk models for energy commodities. Energy Econ. 74, 628–643.
Li, J., Ho, M.S., 2022. Indirect cost of renewabel energy: insighte from dispatching. Energy Econ. 105, 105778.
Liu, M., Wu, F.F., 2007. Risk management in a competitive electricity market. Int. J. Electr. Power Energy Syst. 29, 690–697.
Lobato Miguélez, E., Egido Cortés, I., Rouco Rodríguez, L., López Camino, G., 2008. An overview of ancillary services in Spain. Electr. Power Syst. Res. 78, 515–523.
Sadeghi, M., Shavvalpour, S., 2006. Energy risk management and value at risk modeling. Energy Policy 34, 3367–3373.
Silverman, B., 1986. Density Estimation for Statistics and Data Analysis. Chapman & Hall/CRC, London.
Wood, S.N., 2006. Generalized Additive Models: An Introduction with R. Chapman and HallCRC.