Measuring univariate VaR for the 'Prezzo Unico Nazionale' hourly losses in the Italian electricity market ('Mercato del Giorno Prima'): a backtesting exercise from July 2012 to October 2020 by means of asymmetric GARCH (GJR-GARCH) processes and Extreme Value Theory

Fabrizio Miorelli

According to the number of studies available, the spot price in the electricity market has been a very prolific topic in academic research. As pointed out by many of these studies, the prices (the returns or losses) in the electricity market do exhibit some stylized facts, which seems to be in some way similar to these showed by other time series. Some of these facts are seasonality, volatility clustering and asymmetry, extreme events, peaks and fat tails.

Following some talks that I have had with some recruiters, I decided to apply and extend some procedures and statistical methods that I've studied and applied during the preparation of my dissertation, back in 2009. The aim of this exercise should be thought as a refreshening and an extension of some skills that I have gained throughout the years. In particular, I've decided to apply the Peaks over Threshold procedure to high quantile estimation in a conditional framework, using an ARMA-GJRGARCH asymmetric volatility process to filter the hourly losses of the 'Prezzo Unico Nazionale' (PUN), the reference price of the Italian Electricity Market ('Mercato del Giorno Prima'), in the attempt to estimate some VaR figures that might be able to capture some of the stylized facts that the electricity price returns/losses seem to exhibit.

Here in this note I will briefly explain the procedure that I've adopted in my analysis and report some results that I've achieved by performing a backtesting exercise on the Value at risk estimates.

1. Moving windows and losses filtering

The first step was to calculate the hourly losses for the PUN. I chose a wide set of prices, ranging from 01-01-2010 to 08-10-2020 and then I calculated the hourly losses:

$$\left(\frac{P_t}{P_{t-1}} - 1\right) * (-1)$$ where *t* is an hour [1:24]

In order to take into account for hourly seasonality, the volatility and the VaR estimates are based on the hourly losses for each hour in the entire set of days. This means that, rather than estimating the GJR-GARCH volatility on subsequent hourly losses, I splitted the main sample extracting 24 sub-samples, each of them containing the losses on a certain hour of the day for multiple days. For example, to estimate the VaR figure for the 3rd hour (3:00 am) I extracted a subsample containing all the losses/returns that have been generated at 3:00 am for all the days in the main sample.

Each sub-sample has been splitted into 2.160 training sets using a moving window of 650 observations of hourly losses to provide 2.159 VaR estimates for the 651th losses. For each hourly sub-sample, at each iteration I extracted the *i*-th training set from the subsample and then applied the GJR-GARCH filter to obtain the volatility estimate and the residuals, which are then used to estimate a high quantile.

2. The models

To take into account for volatility clustering, asymmetry and extreme events in hourly losses, I referred to some articles and books that deals with GARCH volatility modeling and peaks over threshold method. For the scope of my analysis I assumed that the losses were driven by an ARMA(1,1) + GJR-GARCH(1,1) process:

$$
\begin{cases}
x_t = \theta_0 + \theta_1 x_{t-1} + \sigma_t z_t \\
\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \gamma_1 I_{t-1} \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \\
\epsilon_t = \sigma_t z_t
\end{cases}
$$

Where z_t is a random variable.

The GJR-GARCH models have been fitted using the 'rugarch' package (1). For fitting purposes, I assumed that z_t was a standard normal random variable.

I made use of two different models for VaR estimation (quantile estimation). The first one, assumes that z_t is standard normal, so I simply calculated the 99th quantile for a standard normal distribution. The second one, which is the main model, uses the Peak Over Threshold Method to estimate a high quantile (99%). With this method, the distribution of the excess over a predefined threshold (x-u) is modelled with a General Pareto distribution and fitted via MLE with the package 'fExtremes'(2):

$$
G_{\xi,\beta,\mu}(y) = \begin{cases}
1-\left(1+\xi\,\frac{y-\mu}{\beta}\right)^{-\frac{1}{\xi}} & \xi \neq 0 \\
1-\exp\left(-\frac{y-\mu}{\beta}\right) & \xi = 0
\end{cases}
$$

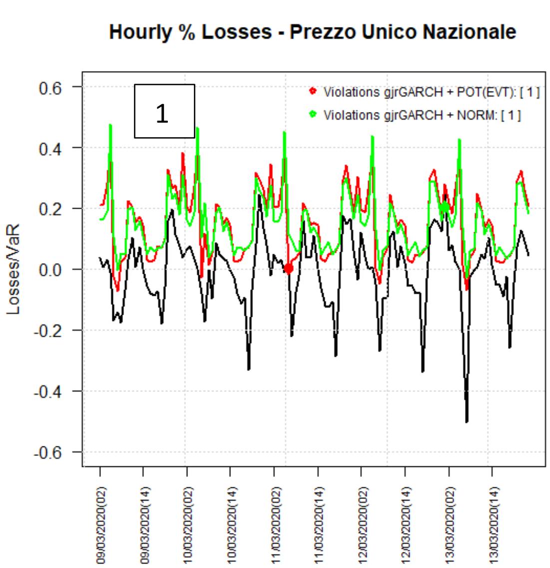
3. Results

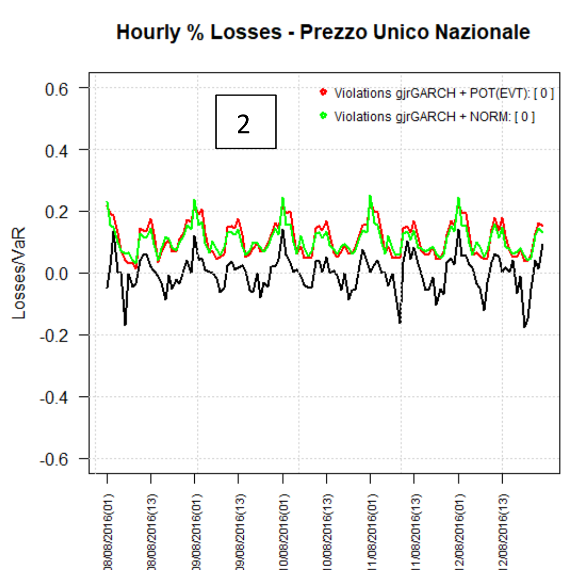
The test set ranges from 02/07/2012 to 09/10/2020 and contains 51.186 hourly losses. For each hour and each date in the test set, two VaR figures are provided (VaRevt and VaRnorm).

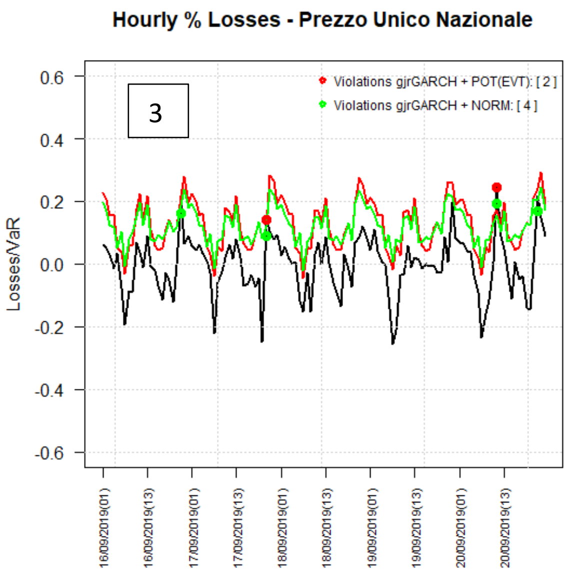
a. Counting VaR violations

The counting of VaR violations revealed that the **VaRevt** estimates have been exceeded **631 times (1,21%)** over the entire test set, while the **VaRnorm** estimates have been exceeded **867 times (1,67%).** Due to the high frequency of the test set I was not able to provide a complete chart, so I decided to take a few relevant 5 days samples and plot the VaR figures over the losses:

1) 09/03/2020 → 14/03/2020 (Italy's COVID19 lockdown start)
2) 08/08/2016 → 12/08/2016 (a hot summer week)
3) 16/09/2019 → 20/09/2019 (random chosen sample)







b. Testing VaR accuracy

To test the VaR estimates accuracy I've tried to apply some statistical coverage tests (Christoffersen, Kupiec), but due to the high frequency and numerosity of data it seems that the LR statistics does not provide a reliable value, so I decided to perform a simple binomial test. The results showed that both VaRevt and VaRnorm correct unconditional coverage are rejected.

VaR model	**Violations**	**Z (binomial test)**
VaRevt	631 (1.21%)	4.982
VaRnorm	867 (1.67%)	15.402

Conclusion

According to the binomial test which has been performed on VaR estimates, both models were rejected, although the VaRevt test statistic resulted to be lower than the VaRnorm statistic and the VaRevt model performed slightly better in violations counting as resulted more closer to the theorical proportion of violations (1,21% vs 1,67%, were the reference is 1%). Both models seem to capture the high volatility of the PUN losses, although the VaRevt model seem to perform slightly better, especially in high volatile periods with jumps and peaks.

References:

1. "rugarch: Univariate GARCH models", Alexios Ghalanos, 2013
2. "fExtremes: Rmetrics - Modelling Extreme Events in Finance", D. Wuertz, T. Setz, Y. Chalabi, 2017
3. McNeil, A. J., Frey, R., and Embrechts, P. _Quantitative Risk Management: Concepts, Techniques and Tools. Princeton University Press, (2005).
4. de Haan, L., and Ferreira, A. Extreme Value Theory. An Introduction. Springer Series in Operations Research and Financial Engineering, (2006).
5. Engle, R. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Infation . Econometrica 50 (1982), 987 1008.
6. McNeil, A. J., and Frey, R. Estimation of Tail-related Risk Measures for Heteroscedastic Financial Time Series: an Extreme Value Approach. Journal of Empirical Finance 7 (2000).

7. McNeil, A. J., and Saladin, T. The Peaks over Thresholds Method for Estimating High Quantiles of Loss Distributions. Proceedings of 28th International ASTIN Colloquium (1997), Departement Mathematik ETH Zentrum CH8092 Zürich.
8. N.Hamzaoui, B.Regaieg, The Glosten-Jagannathan-Runkle-Generalized Autoregressive Conditional heteroscedastic approach to investigating the foreign exchange forward premium volatility, International Journal of Economics and Financial Issues, 2016, 6(4), 1608-1615
9. Dataset: https://www.mercatoelettrico.org/It/Tools/Accessodati.aspx?ReturnUrl=%2fit%2fDownload%2fDownloadDati.aspx%3fval%3dMGP_Prezzi&val=MGP_Prezzi