Probabilistic forecasting of electricity spot prices

Olivier Féron[1,2]

[1] *Electricité de France R&D, Palaiseau, France*
[2] *Laboratoire Finance des marchés de l'énergie, www.fime-lab.org*

From a joint work with G. Dutot, M. Zaffran and Y. Goude[1]

1 Introduction

Electricity price forecasting (EPF) plays a major role for electricity companies as a fundamental entry for trading decisions or energy management operations. As electricity can not be stored, electricity prices are highly volatile which makes EPF a particularly difficult task.

The increase of renewable production in many countries (RTE [2022], IEA [2022a]), the development of storage devices or more generally demand response programs (e.g. electrical vehicle smart charging Nassar et al. [2022], electric water heater management Amabile et al. [2021], Marin Moreno et al. [2023]) simultaneously entails a need for good EPF and generates more complexity for price modelling. Furthermore, prices can be affected by fortuitous events such as COVID pandemic in 2020-2021 IEA [2021], the stress corrosion issue which affected French nuclear power plants in 2022 or the crisis of the gas markets triggered by Russia's invasion of Ukraine IEA [2022b]. These last events have highlighted two main characteristics of the electricity prices: high volatility and non-stasionnarity.

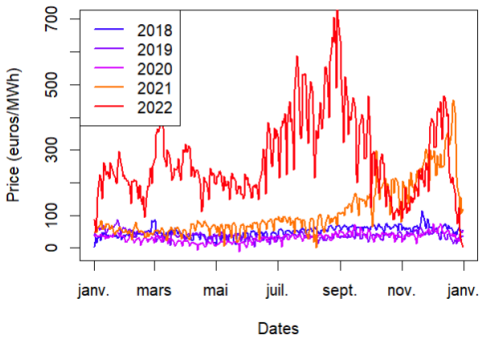

Figure 1: electricity prices in the French spot market from 2018 to 2022

The literature on EPF is growing rapidly and most papers deals with point forecasts Weron [2014], Lago et al. [2021] on day-ahead (spot) markets as the short-term power trading in Europe is the day-ahead market. Models used for forecasting electricity prices are mainly statistical methods, encompassing the recent developments in machine learning. Statistical models are dominated by auto-regressive models and their variants, in particular the state ot the art Lasso Estimated AutoRegressive (LEAR) model proposed by Uniejewski et al. [2016] and recently used as state of the art benchmark in Lago et al. [2021], Tschora et al. [2022]. It consists in a high dimensional ARX model where the fitting process is done by minimizing an elastic net regularization. The high dimension (arround 250 parameters) comes from a large number of lags of prices and forecasts of variable of interests (generation, zonal prices, consumption). As highlighted by Lago et al. [2021] pre-processing of EP such as log transformations or more generally variance stabilizing transformations Uniejewski et al. [2018] are a common practice to deal with heavy tailed distribution. Regarding non-stationarity of the prices, regime switching ARX models are proposed in Nitka et al. [2021]. Marcjasz et al. [2018] propose to average a set of point forecasts obtained from learning with different time windows to derive probabilistic forecasts.

[1] G. Dutot, M. Zaffran, O. Féron and Y. Goude, "Adaptive probabilistic forecasting of French electricity spot prices in 2021", working paper, 2023 Dutot et al. [2023]

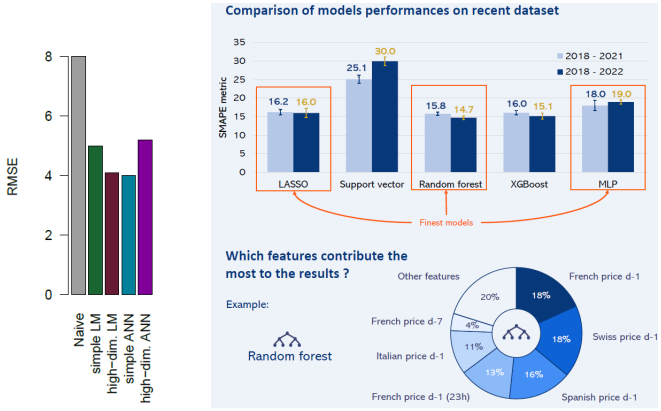

Figure 2: Illustration of results of point forecasting methods. The right figure comes from Ziel [2022]

The utilization of machine learning tools Tschora et al. [2022] for electricity price forecasting has grown over the past decade. However, it has not, at this time, shown significant improvements in EPF, as highlighted in Ziel [2022]. The smallest errors in Electricity price forecasting stil remain in the order of 10%. Let us note however that the recent develoments of hybrid methods with machine learning techniques including fundamental modeling Ziel [2022] may be promising to imporve the performance of EPF.

Probabilistic forecasting, applied to electricity prices, has grown rapidly since 2016. We refer to Nowotarski and Weron [2018] for an introduction and a review. The probabilistic description of future values is essential for trading and more generally energy risk management Bunn et al. [2016].

2 Probabilisitc forecasting methods

2.1 Framework

We are interested in the (unknown) value of a variable Y (e.g. the future price of electricity, the future value of power consumption...) and, for simplicity, we suppose in the following that $Y \in \mathbb{R}$. The objective is to build forecats of Y from the available information noted X, and we assume we have a data set $\{(y_n, x_n); n = 1, \dots, N\}$ in order to develop estimation procedures.

In the classical point forecasting framework, the objective is to estimate a value $\hat{Y}$ corresponding to the best anticipation of the value of Y given all the available information X. In the case of probabilistic forecasting, in the most general framework, the objective is to estimate a *probability law* $\hat{\mathbb{P}}$ such that, for any subset $\mathscr{S} \in \mathbb{R}$, the probability that Y belongs to $\mathscr{S}$ is $\hat{\mathbb{P}}(\mathscr{S})$. This probability $\hat{\mathbb{P}}$ then represents all the information we have on the value of Y. When this probablity is built from the available current information X, it is noted $\hat{\mathbb{P}}(.|X)$.

In a simpler and more practical framework, probabilistic forecasting can be reduced to estimate some characteristics of $\hat{\mathbb{P}}$.

- One can estimate a set of quantiles of $\hat{\mathbb{P}}(X)$: let us note the $\hat{Q}_\beta(X)$ the quantile of probability $\hat{\mathbb{P}}$ at a level β, it corresponds to

$$\hat{\mathbb{P}}(Y \leq \hat{Q}_\beta(X)|X) = 1 - \beta \tag{1}$$

  Such an estimated quantile $\hat{Q}_\beta$ is expected to be greater than approximately $1-\beta$ of the realisations of Y.

- One can built *Prediction Intervals* (PIs) for Y depending on the covariates X. Let $\alpha \in [0,1]$ be a *miscoverage rate*. A PI at the $1-\alpha$ level, noted $\hat{C}_\alpha(X)$ is such that

$$\hat{\mathbb{P}}(Y \in \hat{C}_\alpha(X)|X) = 1 - \alpha \tag{2}$$

  Such an estimated PI is expected to contain approximately $1-\alpha$ of the realisations. A natural choice for the PI is $\hat{C}_\alpha(X) = \left[\hat{Q}_{\alpha/2}(X)\ ;\ \hat{Q}_{1-\alpha/2}(X)\right]$

In both cases, probabilistic forecasting can be viewed as point forecasting tasks, where the objective is not to estimate the future value $\hat{Y}$ but instead quantiles $\hat{Q}_\beta(X)$ or the boundaries of intervals $\hat{C}_\alpha(X)$ which can also be quantiles. That is why a lot of methods dedicated to probabilistic forecasting are called "quantile regression" methods, referring to classical regression models used in point forecasting.

2.2 Review of some quantile regression methods

As said in the previous section, lots of quantile estimation methods are based on an adaptation of classical point estimation. In the following we describe the most classical methods for estimating quantiles.

2.2.1 Linear Quantile Regression Models

Quantile Regression (QR) Koenker [2005] uses the same methodology as classical regression, but replaces the quadratic loss by the *pinball loss* to forecast a conditional quantile of the distribution of Y given the covariates X:

$$\hat{Q}_\beta(x) = \arg\min_{Q \in \mathcal{Q}} \mathbb{E}\left[\rho_\beta(Y - Q(X)) | X = x\right],$$

where $\mathcal{Q}$ is the set of admissible function $Q(x)$ (the class of regressors considered), and with the pinball loss function defined as follows:

$$\rho_\beta(u) = \begin{cases} (1-\beta)u, & \text{if } u \geq 0 \\ -\beta u, & \text{if } u < 0 \end{cases}$$

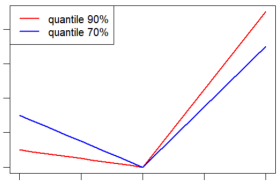

Figure 3: Illustration of the pinball loss

When the class of regressors $\mathcal{Q}$ is restricted to linear models, we speak of *Quantile Linear Regression (QLR)*. *QRL-LASSO* is used when we add specific LASSO techniques. Some other methods exist, e.g. Quantile (tree based) Gradient Boosting (QGB) and can be easily implemented through the `scikit-learn` package Pedregosa et al. [2011] in Python.

2.2.2 Quantile Generalized Additive Models

Generalized Additive Models (GAMs) Hastie and Tibshirani [1986] explain Y by a function of the set of covariates X

$$Y = \mu(X) + \varepsilon$$

with μ an unknown function to be estimated and ε a zero mean i.i.d noise from some exponential family. GAMs assume that there exists a known link function g such that

$$g(\mu(X)) = \sum_{j=1}^{J} f_j(X)$$

where the f_j are smooth functions. The function g can be considered as the identity function if we suppose all the variables X and Y normmalized. The functions f_j are cubic splines (i.e. polynomials of degree 3 joined continuously at "knots", see Wood [2006] for details). The estimation of GAMs is based on a (regularized) mean squared error (MSE) criterion.

In order to use this kind of models for quantile estimation, one could replace the MSE criterion by the pinball loss function in the estimation process as described before. However, Fasiolo et al. [2020] demonstrates that the pinball loss is statistically suboptimal in this framework and proposes a procedure based on the smooth Extended Log-F loss. The latter can be implemented through the `R` package **qgam** Fasiolo et al. [2021].

2.2.3 Quantile Random Forests (QRF)

Meinshausen [2006] adapts Random Forests (RF) to the QR task. In the mean regression setting, a set of B decision trees $\left(\hat{\mu}_b(\cdot)\right)_{b=1}^{B}$ is fitted on the training data $\left(x_n, y_n\right)_{n=1}^{N}$, predicting

$$\hat{\mu}_b(x) := \sum_{n=1}^{N} y_n \frac{\mathbb{1}\left(x_n \in L_b(x)\right)}{\#\{k : x_k \in L_b(x)\}},$$

with $L_b(x)$ the leaf where x falls in the b-th tree. Then, the RF aggregates the B tree's predictions to output their average as unique prediction:

$$\hat{\mu}(x) := \frac{1}{B} \sum_{b=1}^{B} \hat{\mu}_b(x).$$

Meinshausen [2006] uses the same B decision trees and adapts the aggregation to provide an estimator of the cumulative disribution function (cdf).:

$$\hat{F}(y|X=x) := \frac{1}{B} \sum_{b=1}^{B} \sum_{n=1}^{N} \mathbb{1}\{y_n \leq y\} \frac{\mathbb{1}\left(x_n \in L_b(x)\right)}{\#\{k : x_k \in L_b(x)\}}.$$

The inverse of this estimated cdf is then computed to obtain the estimated conditional quantiles:

$$\hat{Q}_\beta(x) = \inf\{y : \hat{F}(y|X=x) \geq \beta\}.$$

2.2.4 Extremal Random Forests (ERF)

ERF Gnecco et al. [2022] aims at providing estimates of conditional quantiles outside the range of the data by modifying the estimated quantile function $\hat{Q}_\beta(x)$ obtained with a GRF for the extreme values of β. Indeed, classical QR techniques, and QRF, usually perform poorly for extreme values of β, due to the scarcity of observed data in this tail region. By contrast, ERF seeks to approximate the cdf in the tails by a Generalized Pareto Distribution (GPD). In Extreme value theory, given a high threshold u, the GPD is used to approximate the distribution of the threshold exceedance

$$Y - u | Y > u \sim H_{\gamma,\sigma}(y) = 1 - \left(1 + \frac{\gamma y}{\sigma}\right)_{+}^{-\frac{1}{\gamma}}$$

with $\gamma \in \mathbb{R}$ and $\sigma > 0$ (shape and scale parameters respectively). In this case, under mild assumptions and when $\beta \to 1$, the conditional quantiles of $Y|X=x$ are approximated by:

$$Q_\beta(x) \approx \hat{Q}_{\beta_0}(x) + \frac{\sigma(x)}{\gamma(x)} \left[\left(\frac{1-\beta}{1-\beta_0} \right)^{-\gamma(x)} - 1 \right],$$

for $\beta > \beta_0$ and with β_0 some fixed threshold and $\hat{Q}_{\beta_0}$ a GRF trained on the data. β_0 is chosen small enough so that it can be approximated by a classical GRF technique, but large enough to ensure the validity of the previous approximation. ERF estimates the parameters $\hat{\sigma}$ and $\hat{\gamma}$ by minimising the negative log-likelihood of each exceedance $z_n = \max\left(y_n - Q_{\beta_0}(x_n)\ ;\ 0\right)$, $n = 1, \ldots, N$. The other quantiles (of levels $\beta \in [0.5, \beta_0]$) are estimated thanks to the initial GRF. Finally, the lower quantiles are estimated similarly: the vector Y is multiplied by -1 and then it proceeds as above, training two other GRFs.

In practice, ERF method provided by the authors of Gnecco et al. [2022] needs a large enough training set to provide predictions without any numerical errors.

2.3 Conformal methods

Conformal Prediction (CP) Vovk et al. [1999], Papadopoulos et al. [2002], Vovk et al. [2005] builds PI around any kind of prediction model. These intervals are valid (achieving marginal nominal coverage) in finite sample under the only assumption of exchangeability of the data.

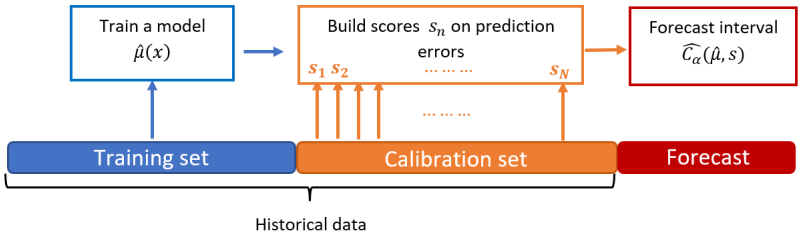

Figure 4: Illustrative scheme of conformal method principle.

Suppose that we have N random variables $\{(X_n, Y_n)\}_{n=1}^N$. For a given miscoverage rate $\alpha \in [0,1]$, we aim at building a *marginally valid* PI $\widehat{C}_\alpha$ of Y_{N+1}, i.e. $\widehat{C}_\alpha$ should satisfy:

$$\mathbb{P}\left(Y_{N+1} \in \widehat{C}_\alpha(X_{N+1})\right) \geq 1-\alpha. \tag{3}$$

To achieve this, Split Conformal Prediction (SCP) Papadopoulos et al. [2002], Lei et al. [2018] randomly splits the N data points into a training set Tr and a calibration set Cal. A regression model $\hat{\mu}$ is then fitted on Tr and used to predict on Cal to obtain a set of conformity scores $\mathscr{S}_{\text{Cal}} = \left\{S_n := s\left(X_n, Y_n; \hat{\mu}\right),\ n \in \text{Cal}\right\}$. These scores assess the conformity between the calibration's observed values and the predicted ones: the smaller the better. In the case of regression, they are usually computed using the absolute value of the residuals, i.e.

$$S_n := s\left(X_n, Y_n; \hat{\mu}\right) = |\hat{\mu}(X_n) - Y_n|.$$

A corrected[2] $(1-\tilde{\alpha})$-th empirical quantile of the conformity scores $Q_{1-\tilde{\alpha}}(\mathscr{S}_{\text{Cal}})$ is obtained, to finally build the prediction interval:

$$\widehat{C}_\alpha := \left\{y : s_{\hat{\mu}}(X_{N+1}, y) \leq Q_{1-\tilde{\alpha}}(\mathscr{S}_{\text{Cal}})\right\}.$$

In the standard regression case, it boils down to $\widehat{C}_\alpha(X_{N+1}) = \left[\hat{\mu}(N_{T+1}) \pm Q_{1-\tilde{\alpha}}(\mathscr{S}_{\text{Cal}})\right]$. This procedure is guaranteed theoretically to satisfy Equation (3) for any model $\hat{\mu}$, any sample size T, as long as the calibration and test data are exchangeable.

Proposed by Romano et al. [2019], Conformalized Quantile Regression (CQR) benefits simultaneously from the adaptiveness of classical QR methods and from the theoretical guarantees ensured by CP. Instead of training a mean regression model on the training set Tr, CQR requires to fit two conditional quantile regression models $\hat{q}_\ell(\cdot), \hat{q}_u(\cdot)$[3]. In this context, the conformity scores now quantify the error made by the fitted PI $\widehat{C}(x) := [\hat{q}_\ell(x), \hat{q}_u(x)]$. Precisely, $S_t := s\left(X_t, Y_t; \hat{q}_\ell, \hat{q}_u\right) = \max\left\{\hat{q}_\ell(X_t) - Y_t\ ;\ Y_t - \hat{q}_u(X_t)\right\}$. Accordingly, the PI becomes

$$\widehat{C}_\alpha(X_{N+1}) = \left[\hat{q}_\ell(X_{N+1}) - Q_{1-\tilde{\alpha}}(\mathscr{S}_{\text{Cal}})\ ;\ \hat{q}_u(X_{N+1}) + Q_{1-\tilde{\alpha}}(\mathscr{S}_{\text{Cal}})\right].$$

To account for the temporal aspect of time series, an online and sequential version of SCP is usually considered, in which the split leading to Tr and Cal is not random, but constrained so that any point in Tr occurs before any point in Cal Wisniewski et al. [2020], Zaffran et al. [2022].

[2] The correction $1-\tilde{\alpha} = (1-\alpha)(1+\frac{1}{\#\text{Cal}})$ is needed to ensure finite sample validity, because of the inflation of the quantiles.

[3] Usually $\ell = \alpha/2$ and $u = 1-\alpha/2$, but this is not necessary. Romano et al. [2019] suggests to choose these values by cross-validation, to improve PI's efficiency.

3 Evaluation

The main challenge of evaluating a probabilistic forecast is that the true distribution of the underlying process cannot be observed, hence the impossibility to compare the estimated probability law $\hat{\mathbb{P}}(.|X)$ to the actual law of the true value Y given X.

Yet, a lot of metrics are available (for PEPF, see Nowotarski and Weron [2018]) which can be categorized in two classes: metrics for *sharpness* and metrics for *reliability*. A common paradigm in evaluating probabilistic forecasts is "to maximize the sharpness of the predictive distributions, subject to reliability" Gneiting and Raftery [2007]. Sharpness refers to how tightly the predictive distribution covers the true one, while reliability measures the statistical consistency between the distributional forecasts and the observations (for example, whether the 90 % prediction interval covers 90 % of the observations or not).

3.1 Reliability metrics

We first describe the common empirical *coverage metric*. Given $\alpha \in [0\ ;\ 1]$ the coverage metric is defined as:

$$Cov_\alpha = \frac{1}{N}\sum_{n=1}^{N} I_n$$

with:

$$I_n = \begin{cases} 1 & \text{if } y_n \in C_\alpha(x_n) \\ 0 & \text{otherwise} \end{cases}$$

This metric is then the averaged number of observations y_n which are in the predicted interval $C_\alpha(x_n)$.

One can also compute the **Kupiec test** Kupiec [1995], which checks for a given time n and miscoverage rate α whether $\hat{\mathbb{P}}(I_n) = 1-\alpha$ under the assumptions that the violations (i.e. the points external to the predicted interval) are i.i.d according to a Bernoulli law of mean $1-\alpha$. The test rejects the null hypothesis of an accurate prediction interval if the fraction of violations is statistically greater than α.

3.2 Sharpness metrics

To evalutate the sharpness of the quantiles forecasts, one can use the classical **pinball score** ρ_β already defined in section 2.2, which can be rewritten here: consider an estimated quantile $\hat{Q}_\beta$ characterized by eq. (1) and the realization y, the corresponding pinball score is:

$$\rho_\beta(y-\hat{Q}_\beta) = \begin{cases} (1-\beta)(y-\hat{Q}_\beta) & \text{if } y \geq (y-\hat{Q}_\beta) \\ \beta((y-\hat{Q}_\beta)-y) & \text{otherwise.} \end{cases}$$

In the case the probabilistic forecast reduces to the estimation of an interval (of miscoverage α), one can measure the **interval width**: the smaller the interval, the better the interval forecat (provided that the interval forecast satisfy reliability).

One also use the **Winkler score** Gneiting and Raftery [2007]. Consider a prediction interval $\hat{C}_\alpha = [\hat{L}_\alpha\ ;\ \hat{U}_\alpha]$ with $\hat{C}_\alpha$ as defined in (2), the Winkler score is given by:

$$Winkler(y,\hat{C}_\alpha) = \begin{cases} (\hat{U}_\alpha - \hat{L}_\alpha) + \frac{2}{\alpha}(\hat{L}_\alpha - y) & \text{if } y < \hat{L}_\alpha \\ (\hat{U}_\alpha - \hat{L}_\alpha) & \text{if } y \in \hat{C}_\alpha \\ (\hat{U}_\alpha - \hat{L}_\alpha)\frac{2}{\alpha}(y - \hat{U}_\alpha) & \text{if } y > \hat{U}_\alpha \end{cases}$$

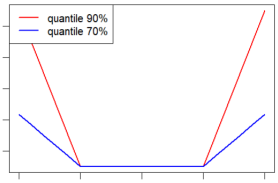

Figure 5: Illustration of the Winker score

A common drawback of these two metrics is that they need either a given quantile or a given miscoverage rate. The classical measure that is aggregated over all quantiles is the Continuous Ranked Probability Score (**CRPS**). This score is originally described in terms of the predictive distribution function. Consider again the

realization y and the estimated distribution function $\hat{F}(u) = \hat{\mathbb{P}}(Y < u)$:

$$CRPS(\hat{F}, y) = \int_{-\infty}^{\infty} \left(\hat{F}(u) - \mathbb{1}_{\{y \leq u\}}\right)^2 du$$

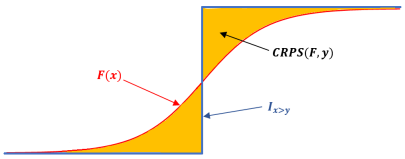

Figure 6: Illustration of the CRPS function

Interestingly, the CRPS can be reformulated (to a multiplicative constant) as :

$$CRPS(\hat{F}, y) = \int_0^1 \rho_\beta \left(y, \hat{F}^{-1}(\beta)\right) d\beta$$

where $\hat{F}^{-1}(\beta)$ actually corresponds to the predicted value at quantile β. By approximating this integral as an Riemann sum, we are able to transform pinball scores over multiple quantiles into one single metric.

4 Illustration

In this section we propose an illustration of probabilisitic forecasting applied to the French electricity spot prices. This illustration comes from the working paper Dutot et al. [2023] where a more exhaustive comparison is done. For this illustration we will study the comparison of only two quantile regression methods and the possible additional improvements of conformatlization approach. More precisely we consider the following methods:

- Linear Quantile regressors (**Linear**)
- Quantile Random Forest (**QRF**)
- Conformal quantile regressor (**CQR**)
- Conformal Quantile regresor with increasing calibration set (**OSSCQR**) and online updating parameters (**OSSCQR-horizon**)

The last method is an extension of the classical CQR methods in order to make the model more adaptive. Also, all these methods were calibratated on different size of historical data from 90 days to 4 years preceeding the forecast date.

We consider a dataset from 2016 to 2021. In order to forecast the electricity spot price at date D, we consider the following explanatory variables:

- French day-ahead prices at day D-1 and D-7
- Daily Gas prices at day D-1 and front-month prices for Oil and Coal
- Residual load[4] forecast at day D
- Available gereration capacity of french nuclear electricity at day D
- Power generation from all sources at D-2 and D-7
- EUR vs GBP and EUR vs USD exchange rate (last observed before closing)
- Total electricity exchange between France and its neighbors, and the exchange beween france and Germany (observed at D-2) ;
- Dummy variables for considering French holiday, holiday bridges, Week-ends (or week day for Gams and ERF), the time of year and a clock variable

[4]The residulal load is the power demand minus all the fatal generation (renewable production, fatal hydro...

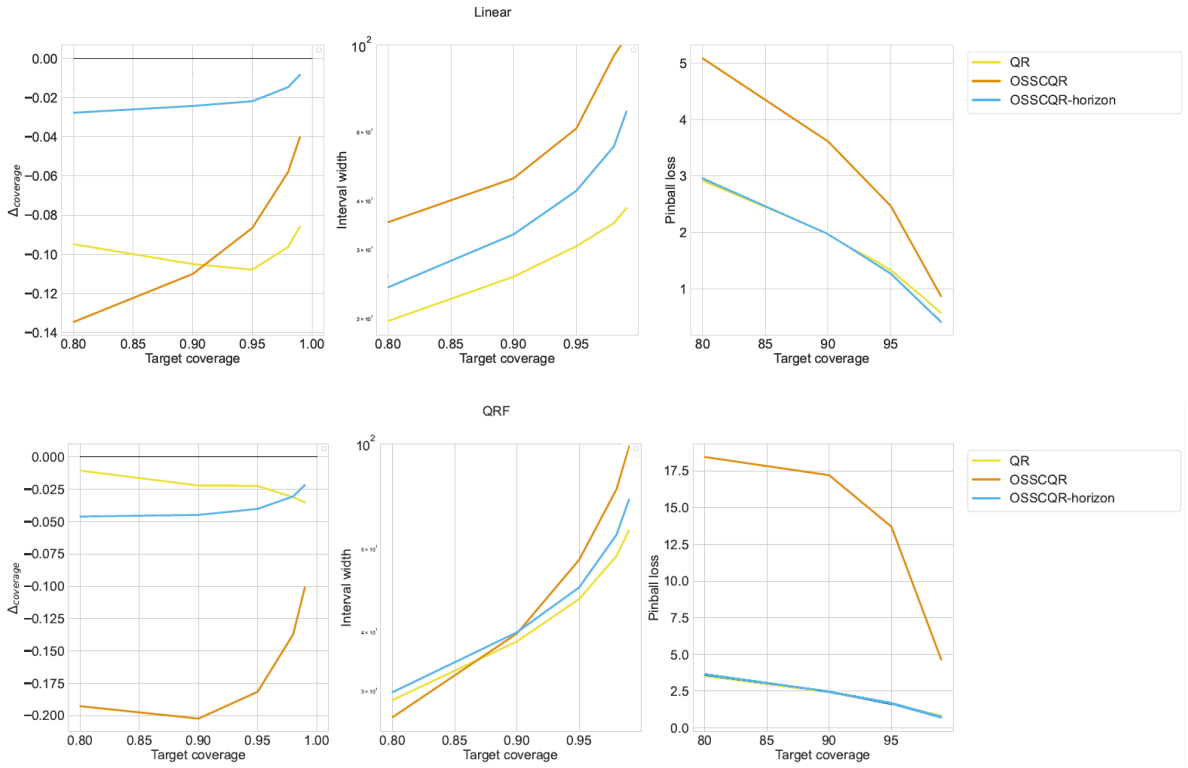

Figure 7: Exemple of results for linear quantile regression (up) and Quantile Random Forest (down): QR is the classical methods, OSSCQR and OSSCQR-Horizon are extended conformal approaches.

Figure 7 gives an illustration of the results of linear quantile regression and Quantile random forests in terms coverage, interval width and pinball loss for quantiles from 0.8 to 0.99. Concerning coverage, the zero horizontal line corresponds to the targeted coverage for each quantile: getting negative $\Delta_{coverage}$ implies that the method is theoretically not valid in termes of reliability. Therefore, we can can oberve that none of the methods reach the targetted coverage. However, For linear quantile regression the method OSSCQR-Horizon seems to get closer to the reliability requirement, without increasing the interval width. For the quantile random forest, the conformalization seems to not significantly improve the results.

5 Extension: towards aggregations techniques

As illustrated in Section 1, electricity price are highly volatile and non stationnary. which makes the probabilistic forecast a very difficult task. There are several ways of research to try to improve the peformance of probabilistic forecast of electricity prices.

Among them, the aggregation techniques are very promizing and have already proved their relevance for forecasting methods, because they allow to increase the adaptability and the efficiency of forecasts, by averaging the forecast results of several models

$$\hat{q}_{\alpha}^{Ag} = \frac{1}{M} \sum_{m=1}^{M} w_m \hat{q}_{\alpha}^{m}$$

where $\hat{q}_{\alpha}^{m}$ is the estimated α-quantile frome Model m and w_m a weight associated to model m. The adaptability of aggregation models comes from the effciency to update in real time the weights w_m which are

estimated in order to minimize a specific metrics. For Probabilistic electricity price forecasting, one can use the metrics previously described in section 3.

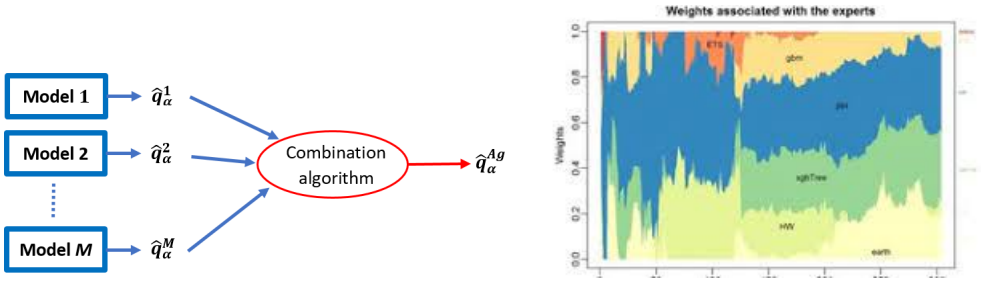

Figure 8: Illustrative scheme of aggregation approach (left) and illustrative result of estimated weights' evolution (right)

There is a need of huge research work to develop probabilistic forecasting methods adapted to (non stationnary) time series, especially applied on electricity prices. Recent work on adaptive conformal methods Zaffran et al. [2022] seem promising but miss, at this time, theoretical ensurements of convergence.

References

Loris Amabile, Delphine Bresch-Pietri, Gilbert El Hajje, Sébastien Labbé, and Nicolas Petit. Optimizing the self-consumption of residential photovoltaic energy and quantification of the impact of production forecast uncertainties. Advances in Applied Energy, 2:100020, 2021. ISSN 2666-7924. doi: https://doi.org/10.1016/j.adapen.2021.100020. URL https://www.sciencedirect.com/science/article/pii/S2666792421000135.

Derek Bunn, Arne Andresen, Dipeng Chen, and Sjur Westgaard. Analysis and forecasting of electricty price risks with quantile factor models. The Energy Journal, 37(1), 2016.

G. Dutot, M. Zaffran, O. Féron, and Y. Goude'. Adaptive probabilistic forecasting of french electricity spot prices in 2021. Working paper, 2023.

Matteo Fasiolo, Simon N. Wood, Margaux Zaffran, Raphaël Nedellec, and Yannig Goude. Fast calibrated additive quantile regression. Journal of the American Statistical Association, 116(535):1402–1412, mar 2020. doi: 10.1080/01621459.2020.1725521.

Matteo Fasiolo, Simon N. Wood, Margaux Zaffran, Raphaël Nedellec, and Yannig Goude. qgam: Bayesian nonparametric quantile regression modeling in R. Journal of Statistical Software, 100(9):1–31, 2021. doi: 10.18637/jss.v100.i09.

Nicola Gnecco, Edossa Merga Terefe, and Sebastian Engelke. Extremal random forests, 2022. URL https://arxiv.org/abs/2201.12865.

Tilmann Gneiting and Adrian E Raftery. Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association, 102(477):359–378, 2007. doi: 10.1198/016214506000001437. URL https://doi.org/10.1198/016214506000001437.

Trevor Hastie and Robert Tibshirani. Generalized additive models. Statistical Science, 1(3):297–310, 1986. ISSN 08834237. URL http://www.jstor.org/stable/2245459.

Paris IEA. Covid-19 impact on electricity. Technical report, Technical report, 2021.

Paris IEA. Renewable electricity. Technical report, Technical report, 2022a.

Paris IEA. World energy outlook 2022. Technical report, Technical report, 2022b.

Roger Koenker. Quantile Regression. Econometric Society Monographs. Cambridge University Press, 2005. ISBN 978-0-521-84573-1. doi: 10.1017/CBO9780511754098. URL https://www.cambridge.org/core/books/quantile-regression/C18AE7BCF3EC43C16937390D44A328B1.

Paul H. Kupiec. Techniques for verifying the accuracy of risk measurement models. The Journal of Derivatives, 3(2):73–84, 1995. ISSN 1074-1240. doi: 10.3905/jod.1995.407942. URL https://jod.pm-research.com/content/3/2/73.

Jesus Lago, Grzegorz Marcjasz, Bart De Schutter, and Rafał Weron. Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark. Applied Energy, 293: 116983, 2021.

Jing Lei, Max G'Sell, Alessandro Rinaldo, Ryan J. Tibshirani, and Larry Wasserman. Distribution-Free Predictive Inference for Regression. Journal of the American Statistical Association, 113(523):1094–1111, July 2018. ISSN 0162-1459. doi: 10.1080/01621459.2017.1307116.

Grzegorz Marcjasz, Tomasz Serafin, and Rafał Weron. Selection of calibration windows for day-ahead electricity price forecasting. Energies, 11(9), 2018. ISSN 1996-1073. URL https://www.mdpi.com/1996-1073/11/9/2364.

Bianca Marin Moreno, Margaux Brégère, Pierre Gaillard, and Nadia Oudjane. A MIRROR DESCENT APPROACH FOR MEAN FIELD CONTROL APPLIED TO DEMANDE-SIDE MANAGEMENT. working paper or preprint, January 2023. URL https://hal.science/hal-03972660.

Nicolai Meinshausen. Quantile regression forests. Journal of Machine Learning Research, 7(35):983–999, 2006. URL http://jmlr.org/papers/v7/meinshausen06a.html.

N. Nassar, D. Silva, and H. Morais. Hierarchical energy management solution for smart charging. In CIRED Porto Workshop 2022: E-mobility and power distribution systems, volume 2022, pages 721–725, 2022. doi: 10.1049/icp.2022.0805.

Weronika Nitka, Tomasz Serafin, and Dimitrios Sotiros. Forecasting electricity prices: Autoregressive hybrid nearest neighbors (arhnn) method. In Maciej Paszynski, Dieter Kranzlmüller, Valeria V. Krzhizhanovskaya, Jack J. Dongarra, and Peter M.A. Sloot, editors, Computational Science – ICCS 2021, pages 312–325, Cham, 2021. Springer International Publishing.

Jakub Nowotarski and Rafał Weron. Recent advances in electricity price forecasting: A review of probabilistic forecasting. Renewable and Sustainable Energy Reviews, 81:1548–1568, 2018. ISSN 1364-0321. doi: https://doi.org/10.1016/j.rser.2017.05.234. URL https://www.sciencedirect.com/science/article/pii/S1364032117308808.

Harris Papadopoulos, Kostas Proedrou, Volodya Vovk, and Alex Gammerman. Inductive Confidence Machines for Regression. In Tapio Elomaa, Heikki Mannila, and Hannu Toivonen, editors, Machine Learning: ECML 2002, pages 345–356, Berlin, Heidelberg, 2002. Springer Berlin Heidelberg. ISBN 978-3-540-36755-0.

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12:2825–2830, 2011.

Yaniv Romano, Evan Patterson, and Emmanuel Candes. Conformalized Quantile Regression. Advances in Neural Information Processing Systems, 32, 2019. URL https://papers.nips.cc/paper/2019/hash/5103c3584b063c431bd1268e9b5e76fb-Abstract.html.

J RTE. Bilan électrique 2022, 2022.

Léonard Tschora, Erwan Pierre, Marc Plantevit, and Céline Robardet. Electricity price forecasting on the day-ahead market using machine learning. Applied Energy, 313:118752, 2022. ISSN 0306-2619. doi: https://doi.org/10.1016/j.apenergy.2022.118752. URL https://www.sciencedirect.com/science/article/pii/S0306261922002057.

Bartosz Uniejewski, Jakub Nowotarski, and Rafał Weron. Automated variable selection and shrinkage for day-ahead electricity price forecasting. Energies, 9(8), 2016. ISSN 1996-1073. URL https://www.mdpi.com/1996-1073/9/8/621.

Bartosz Uniejewski, Rafał Weron, and Florian Ziel. Variance stabilizing transformations for electricity spot price forecasting. IEEE Transactions on Power Systems, 33(2):2219–2229, 2018. doi: 10.1109/TPWRS.2017.2734563.

Vladimir Vovk, Alexander Gammerman, and Glenn Shafer. Algorithmic Learning in a Random World. Springer US, 2005. ISBN 978-0-387-00152-4. doi: 10.1007/b106715.

Volodya Vovk, Alexander Gammerman, and Craig Saunders. Machine-Learning Applications of Algorithmic Randomness. In Proceedings of the Sixteenth International Conference on Machine Learning, ICML '99, pages 444–453, San Francisco, CA, USA, June 1999. Morgan Kaufmann Publishers Inc. ISBN 978-1-55860-612-8.

Rafał Weron. Electricity price forecasting: A review of the state-of-the-art with a look into the future. International journal of forecasting, 30(4):1030–1081, 2014.

Wojciech Wisniewski, David Lindsay, and Sian Lindsay. Application of conformal prediction interval estimations to market makers' net positions. In Alexander Gammerman, Vladimir Vovk, Zhiyuan Luo, Evgueni Smirnov, and Giovanni Cherubin, editors, Proceedings of the Ninth Symposium on Conformal and Probabilistic Prediction and Applications, volume 128 of Proceedings of Machine Learning Research, pages 285–301. PMLR, 09–11 Sep 2020. URL https://proceedings.mlr.press/v128/wisniewski20a.html.

Simon Wood. Generalized Additive Models: An Introduction With R, volume 66. 01 2006. ISBN 9781315370279. doi: 10.1201/9781315370279.

Margaux Zaffran, Olivier Feron, Yannig Goude, Julie Josse, and Aymeric Dieuleveut. Adaptive conformal predictions for time series. In Proceedings of the 39th International Conference on Machine Learning, volume 162 of Proceedings of Machine Learning Research, pages 25834–25866. PMLR, 17–23 Jul 2022.

F. Ziel. Electricity price forecasting: Data science meets fundamental models. In 9^th Electricity Price Modelling and Forecasting Forum, Berlin, Germany, 2022.