# Quantifying the Uncertainty of Electric Vehicle Charging with Probabilistic Load Forecasting

Yvenn Amara-Ouali [1] & Bachir Hamrouche [2] Guillaume Principato [3] & Yannig Goude [4]

[1] *University Paris-Saclay, LMO, France, yvenn.amara-ouali@universite-paris-saclay.fr*
[2] *EDF R&D, OSIRIS, Palaiseau, France, bachir.hamrouche@edf.fr*
[3] *University Paris-Saclay, LMO, France, guillaume.principato@universite-paris-saclay.fr*
[4] *EDF R&D, OSIRIS, Palaiseau, France, France et yannig.goude@edf.fr*

**Résumé.** Cet article explore les moyens de quantifier l'incertitude associée à l'utilisation croissante des véhicules électriques (VE) dans la gestion des réseaux électriques. En mettant l'accent sur les solutions de prévision de la charge, l'étude étend un premier benchmark effectué sur la prévision de la charge à l'horizon J+1 pour inclure des algorithmes de prévision probabilistes. Deux approches sont envisagées : une approche directe qui fournit des prévisions de quantiles à l'aide de modèles GAMlss, et une approche *bottom-up* qui prédit les sessions de charge individuelles avant de reconstruire la courbe de charge globale et de calculer les quantiles empiriques. Les méthodes proposées sont évaluées à l'aide de mesures telles que Pinball Loss et l'erreur quadratique, démontrant des performances comparables entre l'approche directe et l'approche ascendante. Les résultats, basés sur des données réelles de sessions de charge à Palo Alto, suggèrent les avantages potentiels de l'approche ascendante pour les quantiles élevés.

**Mots-clés.** Prévisions Quantiles, Monte-Carlo, Approche *Bottom-Up*, Modèles Additifs Généralisés, Modèles de Mélange, Prédiction Conforme, *Smart Grids*

**Abstract.** This paper explores ways of quantifying the uncertainty associated with the increasing use of electric vehicles (EVs) in the management of electricity networks. With a focus on load forecasting solutions, the study extends a benchmark on day-ahead load forecasting to include probabilistic forecasting algorithms. Two approaches are considered: a direct approach that provides quantile forecasts using GAMlss models, and a bottom-up approach that predicts individual charging session characteristics before reconstructing the load curve and calculating empirical quantiles. The proposed methods are evaluated using metrics such as Pinball Loss and RMSE, demonstrating comparable performance between the direct and bottom-up approaches. The results, based on real data from Palo Alto charging sessions, suggest potential advantages of the bottom-up approach for high quantiles.

**Keywords.** Quantile forecasts, Monte-Carlo Simulation, Bottom-up approach, Generalised Additive Models, Gaussian Mixtures, Conformal Prediction, Smart Grids

# 1 Industrial Context

A key lever for reducing greenhouse gas emissions in the transport sector is the large-scale deployment of electrical vehicles (EVs). This has led to many governments implementing strong pro-EV policies, resulting in an increase in the number of electrical vehicles in global markets. The arrival of these vehicles creates challenges in the management of the electrical network while also bringing opportunities in terms of grid flexibility. Indeed, these vehicles will become important assets for managing electricity demand whose charging can be automatically postponed when constraints are high or even used as storing batteries to reinject power when demand is high. All these operations of load optimisation are referred to as smart charging. One of the key elements of an efficient smart charging solution is a strong understanding of charging behaviours which requires the development of efficient forecasting algorithms in order to predict them.

# 2 Related Work

This work builds on a previous set of papers [1] [2] which focused on benchmarking day-ahead load and occupancy forecasting solutions, mainly using algorithms that return point estimates corresponding to the mean of the distribution of interest. The previous benchmark examined two sets of methods: direct approaches that would predict the aggregate load curve at the station level, and bottom-up approaches that would model the set of individual behaviours before aggregating them to obtain the predicted curve at the station level. The bottom-up approaches, although more complex to estimate, offer more flexibility for the deployment of smart charging solutions. In this work, we propose to extend this benchmark to probabilistic forecasting algorithms by exploring probabilistic variations. The use of probabilistic forecasts is becoming increasingly important for the efficient operation of electricity systems, as highlighted in the last two Global Energy Forecasting Competitions [3][4]. Recently, several approaches for estimating probabilistic forecasts related to energy demand have been proposed in [5] and [6]. The need for probabilistic forecasts is particularly important in the management of electric vehicles, as the optimisation of charging loads often requires a good quantification of the certainty around the forecasts in order to manage the best and worst case scenarios. [7] proposes an approach to quantify the uncertainty of parking duration forecasts in EV management.

# 3 Methods

Two approaches have been used to address the probabilistic forecasting task. A direct approach which provides quantile forecasts with a GAMlss (see Section 3.1) model trained directly on the load curve. A bottom-up approach (see Section 3.2) which predicts individual charging session characteristics to then reconstruct the load curve. With both approaches, 9 quantile forecasts are provided from 0.1 to 0.9 with 0.1 increments.

## 3.1 Direct Approach

Generalised Additive Models for location scale and shape (GAMlss) are an extension of GAM [10] which enables the fine modelling multiple parameters of a single distribution. In this study, GAMlss are used to model both the mean and variance of the load at charging points over time.

# 4 Bottom-Up Approach

Bottom-up approaches predict the characteristics of individual charging sessions occurring over time. In particular, three variables are required to reconstruct the load curve of an ensemble of charging stations in an uncontrolled charging environment: $(a_i, d_i, e_i)$ the arrival time, charging duration and energy demand of a charging sessions $i$. These three variables can be modelled using various statistical techniques. It was shown in [2] and [3] that mixture models are an adequate choice of method to represent individual charging sessions. Assuming we can predict the number of charging sessions $N$ occurring each day with a time series model, we can sample from the mixture model distribution N times to obtain a prediction for a particular day. In this work, a SARIMA model is used as the predictor for the number of daily charging sessions. A Monte-Carlo simulation is executed to obtain empirical quantiles of the SARIMA model. For each of these 9 forecasted quantiles, another Monte-Carlo simulation is led on the mixture models to reconstruct a total of 10000 load curves from which empirical quantiles are recovered for each instant.

## 4.1 Metrics

Three types of metrics have been used to evaluate quantile forecast performances. First the pinball loss defined as follows:

$$L_\tau^{\text{pinball}}(y_{\text{obs}}, (\hat{y}_\tau)) = \frac{1}{N} \sum_{i=1}^{N} \rho_\tau(y_{\text{obs}}^i - (\hat{y}_\tau^i)) \tag{1}$$

$$\rho_\tau(u) = \begin{cases} \tau \cdot u, & \text{if } u < 0 \\ (1-\tau) \cdot u, & \text{if } u \geq 0 \end{cases} \tag{2}$$

It penalises the model for deviations between the true target values $y_{obs}$ and the predicted quantiles $(y_\tau)$. The degree of penalty depends on the chosen quantile level $\tau$. Another metric which can be used for assessing the accuracy of quantile forecasts can be defined as follows:

Equation (3) provides an estimate of the probability of the observed value falling below the predicted quantile. Essentially, if $L_\tau^{emp}(y_{obs}, (y_\tau)) = \tau$ the quantile forecast is optimal on the testing sample.

$$L_\tau^{\mathrm{emp}}(y_{\mathrm{obs}}, (\hat{y}_\tau)) = \frac{1}{N}\sum_{i=1}^{N} I(y_{\mathrm{obs}}^i \leq (\hat{y}_\tau^i)) \tag{3}$$

A slightly modified version of the metric defined in equation (3) can be written as follows using block-bootstrap sampling:

$$L_\tau^{\mathrm{rmse}}(y_{\mathrm{obs}}, (\hat{y}_\tau)) = \sqrt{\frac{1}{B}\sum_{i=1}^{B}\left(\tau - \frac{1}{N}\sum_{i=1}^{N} I(y_{\mathrm{obs}}^i \leq (\hat{y}_\tau^i))\right)^2} \tag{4}$$

With B=1000 the number of bootstrap samples of size N.

# 5 Results and Discussion

Experiments were led on the city of Palo Alto (California, USA) dataset that gathers real data from charging sessions occurring. This data has been explored in [1]. Figure 1 shows the average daily quantile forecasts for both approaches. It seems that both approaches capture the general shape of the observed curve in black with a peak demand in the middle of the day and another one in the evening. Both models seem to slightly underestimate high quantiles. This is confirmed by Figure 2 where it can be observed that $L_\tau^{emp} \leq \tau$ for all quantiles except 0.1. Figure 2 also shows that both the direct approach and the bottom-up approaches yield similar performances also confirmed by Table 1. It also indicated that the bottom-up approach could be more performant for high quantiles and particularly quantile 0.9 which is more performant for the bottom-up approach across all metrics considered.

Table 1: Pinball Losses ($L_\tau^{\mathrm{Pinball}}$) and RMSE ($L_\tau^{\mathrm{rmse}}$) metrics for each quantile level

| Quantile | Pinball Loss | | RMSE | |
|---|---|---|---|---|
| | GAMlss | Bottom-Up | GAMlss | Bottom-Up |
| 0.1 | **2.59** | 2.65 | **0.0227** | 0.0356 |
| 0.2 | **4.25** | 4.36 | **0.0165** | 0.0198 |
| 0.3 | **5.36** | 5.53 | **0.0401** | 0.0503 |
| 0.4 | **6.02** | 6.27 | **0.0628** | 0.0764 |
| 0.5 | **6.31** | 6.60 | **0.0758** | 0.101 |
| 0.6 | **6.22** | 6.48 | **0.0892** | 0.115 |
| 0.7 | **5.72** | 5.92 | **0.110** | 0.122 |
| 0.8 | **4.74** | 4.82 | 0.113 | **0.107** |
| 0.9 | 3.13 | **3.05** | 0.106 | **0.0766** |

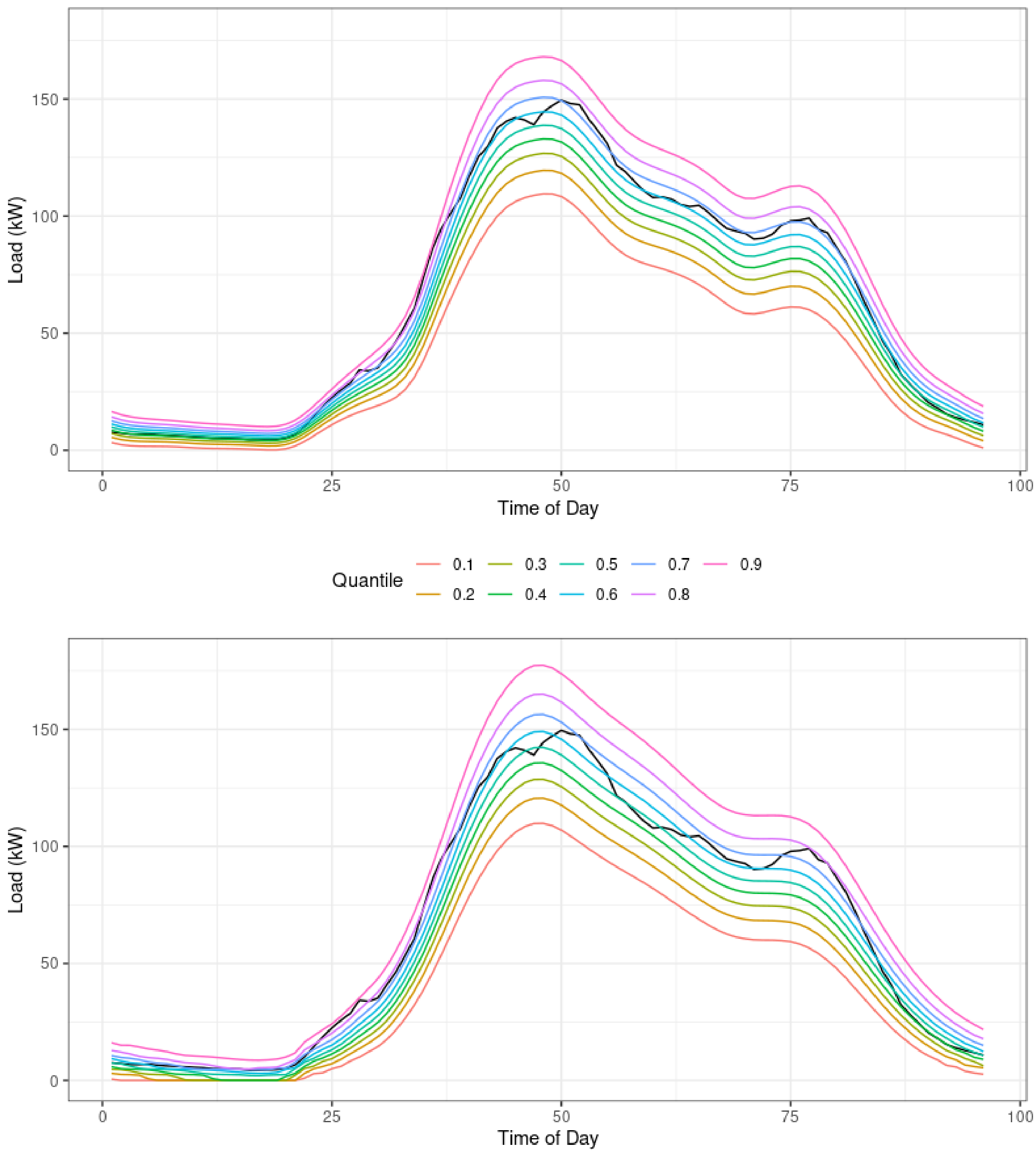


Figure 1: At the top are the GAMlss averaged quantile forecasts for each quantile levels and at the bottom the same chart for the bottom-up approach. In black is the average daily load curve over the test period.

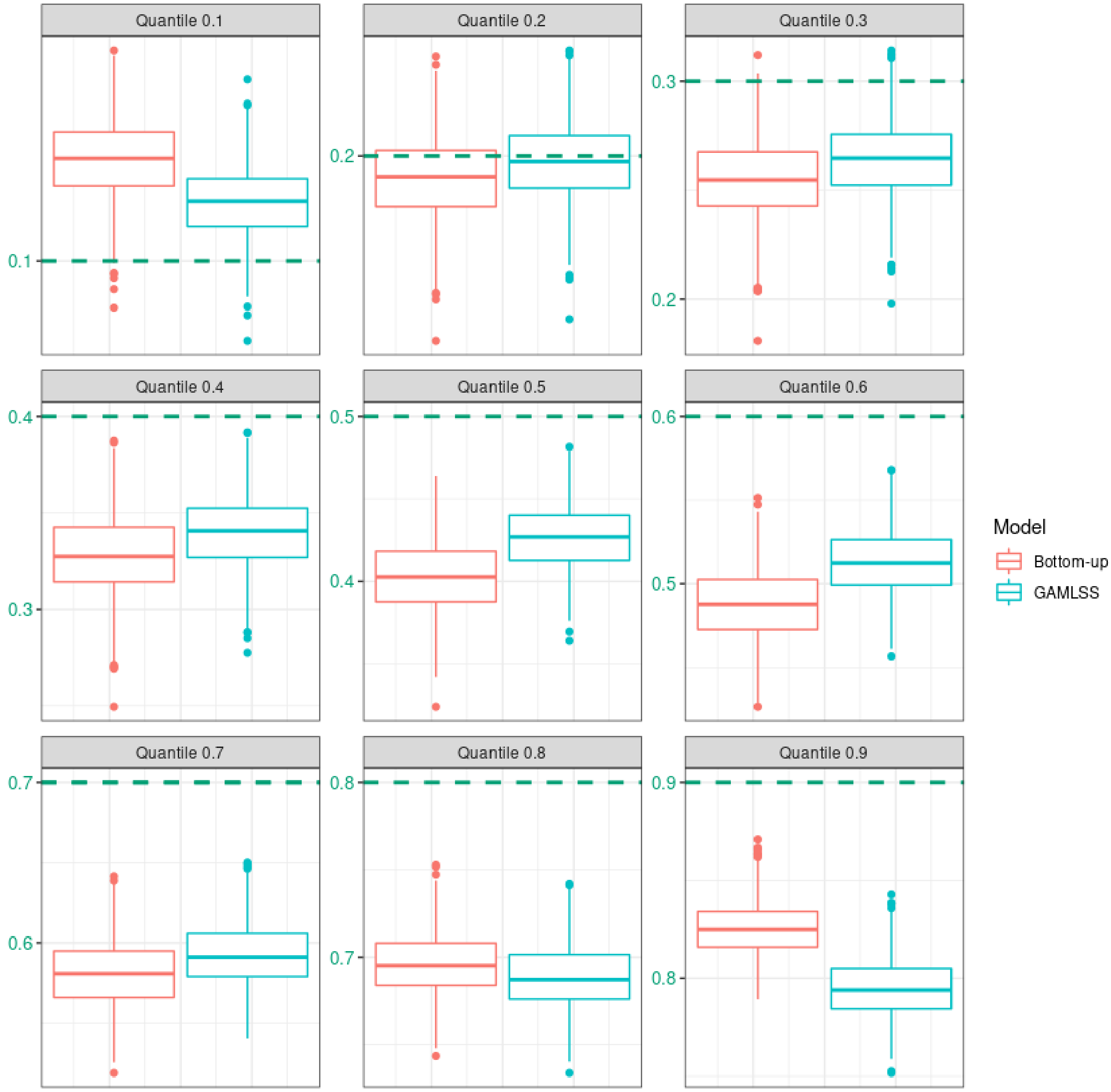


Figure 2: Boxplots of the block-bootstrap performances calculated with $L_{\tau}^{emp}$ defined in equation (3)

# Acknowledgements

The results presented in this abstract will be further detailed and complemented with conformal predictions after review and upon acceptance.

# References

[1] Amara-Ouali, Y., Goude, Y., Massart, P., Poggi, J. M., & Yan, H. (2021). A review of electric vehicle load open data and models. *Energies*, 14(8), 2233.

[2] Amara-Ouali, Y., Goude, Y., Hamrouche, B., & Bishara, M. (2022, June). A benchmark of electric vehicle load and occupancy models for day-ahead forecasting on open charging session data. In *Proceedings of the Thirteenth ACM International Conference on Future Energy Systems* (pp. 193-207).

[3] Hong, T., et al. (2016). Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond. *International Journal of forecasting*, 32(3), 896-913.

[4] Hong, T., Xie, J., & Black, J. (2019). Global energy forecasting competition 2017: Hierarchical probabilistic load forecasting. *International Journal of Forecasting*, 35(4), 1389-1399.

[5] de Vilmarest, J., et al. (2023). Adaptive Probabilistic Forecasting of Electricity (Net-) Load. arXiv preprint arXiv:2301.10090.

[6] Capezza, C., et al. (2021). Additive stacking for disaggregate electricity demand forecasting. *The Annals of Applied Statistics*, 15(2), 727-746.

[7] Phipps, K., et al. (2023). Customized Uncertainty Quantification of Parking Duration Predictions for EV Smart Charging.

[8] Flammini, M. G., et al. (2019). Statistical characterisation of the real transaction data gathered from electric vehicle charging stations. *Electric Power Systems Research*, 166, 136-150.

[9] Barman, P., et al. (2023). Renewable energy integration with electric vehicle technology: A review of the existing smart charging approaches. *Renewable and Sustainable Energy Reviews*, 183, 113518.

[10] Hastie, T., & Tibshirani, R. (1986). Generalized Additive Models. *Statistical Science*, 1(3), 297.