

Demand and residual demand modelling using quantile regression

Linh Phuong Catherine Do¹, Lars Ivar Hagfors¹, Kuan-Heng Lin², and Peter Molnár^{1,*}

¹ Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology, 7491 Trondheim, Norway

² Charles University, CERGE-EI (Center for Economic Research and Graduate Education – Economics Institute), Prague, Czech Republic

Abstract. Residual demand, the difference between demand and solar and wind production, is an important variable in predicting the future price and storage requirements. However, little is known about predicting the residual demand itself as well as its quantiles. Therefore, we model both demand and residual demand using both ordinary and quantile regression and compare the results for the hourly electricity consumption in Germany. We find that the residual demand is less predictable than demand. The effect is visible for all hours, and is higher for the lower than the upper quantiles.

1 Introduction

Over the last years, an increasing amount of renewables has been installed in Germany. This has caused a challenge for different market participants. The power producers need to consider the fluctuations from both load and renewable energy infeed when submitting daily price bids. A market with high infeed of renewable, like Germany, requires a more integrated demand model. As for the grid operators, they need to balance the demand and the supply. Since the production of renewable energy is price inelastic, it makes sense to look at the balancing problem as balancing the demand minus renewables with the supply of conventional electricity sources in order to get grid stability.

The word residual demand has not reached a common definition. We use the term residual demand as a demand minus wind and solar electricity production. This distinction is meaningful because wind and solar electricity producers supply electricity independently on the price. Therefore, we can consider wind and solar electricity as a negative demand.

Earlier studies use the residual demand in strategic price bidding in day ahead market [1,2] and forward market [3]. Reference [4] provides a residual demand model to forecast electricity prices. Reference [5] uses residual demand to understand the energy storage requirement in the future.

Our study applies ordinary and quantile regression method. The former method is useful to find the tendencies and the average relation between the demand and the explanatory variable. The quantile regression approach, introduced by [6], evaluates the tail dependencies and the risk on electricity consumption. The quantile regression

application has been widely applied in financial risk management and been recently used in energy market studies: on electricity price [7], CO₂ emission allowance price [8], household energy consumption [9] and oil prices [10]. This paper aims to contribute to the quantile regression literature by applying this method on both the aggregated electricity demand and residual demand. This analysis is relevant because it provides a more comprehensive picture of the effects from the variables.

This paper is organized as follows. Section 2 describes the data used in this study. The results from ordinary and quantile regression are presented and compared in Sections 3 and 4, respectively. Finally, Section 5 concludes the paper.

2 Data

The hourly electricity consumption in Germany can be described by following fundamental variables: trend, holidays and weather [11]. This paper uses industrial production as the trend, and incorporates holidays and weekend dummies, as well weather factor like hours of daylight and heating degree days.

The German hourly electricity load data is retrieved from the European Network of Transmission System Operators for Electricity (ENTSO-E). Our dataset contains data from July 1, 2010 to July 1, 2013. The load data is the hourly average active power consumed by all installation connected to the central and the distribution network.

The biggest share of renewables production in Germany consists of wind and solar. Moreover, production of these two renewables is completely price inelastic. We therefore focus only on these two renewables in our paper and use the term renewables as interchangeably with wind and solar.

* e-mail: peter.molnar@iot.ntnu.no

Table 1. Descriptive statistics demand, wind and PV for hour 8 am, 12 pm, and 24 pm.

Hour	Mean (MW)	Median (MW)	Max (MW)	Min (MW)	St. dev. (MW)	Skewness	Kurtosis
Demand 8 am	56,011.4	59,193.0	73,955.0	29,644.0	10,213.2	-0.6	-0.9
Demand 12 pm	63,366.4	66,004.0	76,812.0	38,327.0	7819.9	-0.7	-0.6
Demand 24 pm	49,010.4	48,767.0	67,054.0	37,101.0	5106.9	0.2	-0.3
Wind 8 am	4930.7	3492.6	24,786.1	149.5	4443.2	1.6	2.4
Wind 12 pm	5109.7	3339.7	24,724.8	38.6	4922.7	1.4	1.7
Wind 24 pm	5328.1	3978.2	23,632.8	166.6	4414.7	1.5	2.3
PV 8 am	1023.9	381.8	5524.9	0.0	1267.4	1.3	0.7
PV 12 pm	8497.6	7622.8	22,417.1	133.4	5433.0	0.5	-0.6
PV 24 pm	0.0	0.0	0.1	0.0	0.0	7.5	55.9

The actual production data is available on the website to the four Transmission System Operators in Germany: Tennet, 50 Hertz, TransnetBW and Amprion. The wind and PV data have been converted from 15 min data to hourly data.

Table 1 shows the descriptive statistics of the load, wind and solar time-series. The solar production is highest around the noon and zero during the night. In contrast to solar production, the wind production is higher during the night than during the day. In general, the wind and solar production is highly volatile.

3 Demand and residual demand modelling

We use a linear regression models for demand and residual demand specified in equations (1) and (2), respectively. We estimate 24 separate linear regression model for each hour during the day.

$$\begin{aligned}
Y_{i,t} = & a_{i,1} + a_{i,2}HDD_t + a_{i,3}IP_{t-1} + \sum_{\substack{n=1 \\ n \neq 3}}^7 a_{i,4n}W_{n,t} \\
& + a_{i,5}H_t + a_{i,6}H_{t-1} + a_{i,7}MH_t + \sum_{j=2}^{12} a_{i,8j}M_{j,t} \\
& + a_{i,9}DL_t + a_{i,10}Y_{t-1} + \varepsilon_t,
\end{aligned} \tag{1}$$

$$\begin{aligned}
Z_{i,t} = & a_{i,1} + a_{i,2}HDD_t + a_{i,3}IP_{t-1} + \sum_{\substack{n=1 \\ n \neq 3}}^7 a_{i,4n}W_{n,t} \\
& + a_{i,5}H_t + a_{i,6}H_{t-1} + a_{i,7}MH_t + \sum_{j=2}^{12} a_{i,8j}M_{j,t} \\
& + a_{i,9}DL_t + a_{i,10}Z_{t-1} + a_{i,11}AZ_t + \varepsilon_t,
\end{aligned} \tag{2}$$

where Y is demand, Z is residual demand, AZ is the average residual demand over one to seven lags, HDD is heating degree days, W_n are dummy variables for days of the week, H is a holiday variable, MH is a minor holiday variable, M_j are dummy variables for months of the year and DL is hours of daylight.

The coefficients from these regressions are presented in Table 2 for three hours of the day (other hours are omitted due to space limitations). These results illustrate significant differences in modelling demand and residual demand.

First of all, R^2 shows that the models are able to explain much more of the variation of demand than residual demand. This is due to stochastic nature of wind and solar production. This is confirmed also by looking at the coefficient of the lagged demand/residual demand. Coefficient is larger for demand. This means that past demand tells us more about future demand than past residual demand tells us about future residual demand.

Another notable difference is the impact of HDD variable (transformation of temperature). This variable has much larger impact on residual demand than on demand. This is because on average cold days are also days with low solar production.

Since the results of ordinary regression show differences in demand and residual demand modelling, we investigate this further using quantile regression.

4 Quantile regression

The quantile regression method is an extension of ordinary regression, which allows estimating the effect of explanatory variables not only on the mean, but also the effect on conditional quantiles. We use the same equation specification as equations (1) and (2), but we estimate these equations for different quantiles. The model is estimated for 5th, 25th, 50th, 75th, 95th quantile for each hour of the day. Based on previous results from ordinary regression we focus on the impact of lagged demand and lagged residual demand.

The coefficients from quantile regressions for the 5%, 25%, 50%, 75% and 95% quantiles for demand for each hour are presented in Figure 1. As we can see, median, 25th and 75th quantile are higher than 5th or 95th quantile. This means that the current demand provides more information about the future median demand than about the possible future extremely high or low demand. Moreover, this difference is more pronounced during the day than during the night, because the load variation is also higher at the day period.

Table 2. OLS estimates for demand and residual demand for hours 8 am, 12 pm, and 24 pm. ***, ** and * indicates that the coefficient is significant at 10%, 5% and 1% level, respectively.

	Demand			Residual demand		
	8 am	12 pm	24 pm	8 am	12 pm	24 pm
Demand lag	0.6*	0.6*	0.8*			
Resid. demand lag				0.5*	0.5*	0.5*
Avg. resid. demand				0.1**	0.1*	0.0
HDD	58.9*	46.7*	87.5*	133.0*	107.0*	234.4*
IP lag	-52.4***	37.6	-58.3*	-194.2*	-317.8*	-219.7*
Sunday	-14,566.8*	-10,916.3*	-1230.7*	-16,538.0*	-12,322.6*	-2927.6*
Monday	11,067.5*	9186.1*	4018.8*	8789.1*	7449.1*	1971.2*
Tuesday	229.7	141.4	304.6***	-513.3	-797.7	304.2
Thursday	-997.5*	-790.8*	-359.6**	-1165.9**	-1123.1***	-593.2
Friday	-839.0*	-749.2*	-1193.5*	-1195.3**	-1066.3***	-921.5**
Saturday	-15,432.1*	-11,400.6*	-5617.0*	-15,741.8*	-12,266.7*	-6191.6*
Holiday	-16,585.5*	-13,145.6*	-3701.4*	-17,641.8*	-13,696.3*	-5468.8*
Holiday lag	6319.4*	5719.4*	2832.2*	3499.6*	3465.5*	1325.2***
Minor holiday	-3333.5*	-2345.2*	-1114.2*	-3637.3*	-2336.4*	-2075.6*
DL	-249.2*	-122.9*	-44.1***	-55.8	-257.4*	201.7*
Constant	34,696.2*	25,329.0*	16,949.6*	46,987.8*	58,344.4*	39,755.5*
R^2	0.94	0.92	0.92	0.84	0.75	0.60

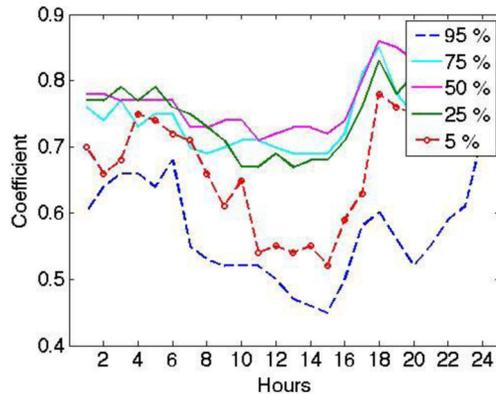
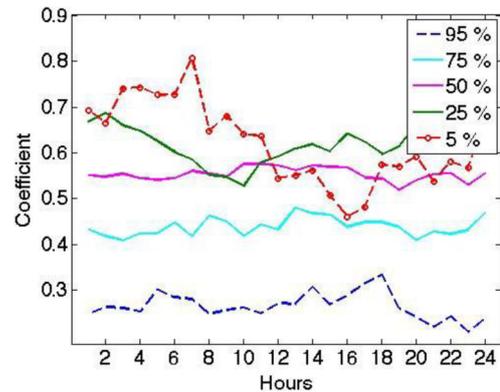
**Fig. 1.** Coefficients from quantile regression for demand lag estimated for 5%, 25%, 50%, 75% and 95% quantiles for each hour of the day.**Fig. 2.** Coefficients from quantile regression for residual demand lag estimated for 5%, 25%, 50%, 75% and 95% quantiles for each hour of the day.

Figure 2 depicts the estimated coefficients from quantile regressions estimated for residual demand. In accordance with results from ordinary regression we find that the coefficients in Figure 2 is on average lower than the coefficients in Figure 1. This confirms that the past residual demand has less predicting power for the current value than the past demand. Moreover, the quantile of residual demand are much more dispersed than quantiles of demand. This means that residual demand is much less predictable than demand. In comparison to Figure 1, the spread between the quantiles in Figure 2 increases during

the night, due to high wind production during this period. The most notable difference between Figure 1 and Figure 2 is the difference between 5th and 95th quantile. Current demand tells us relatively little about either of these quantiles for future demand. However, current residual demand tells us relatively a lot about 5th quantile of the future demand, but little about the 95th quantile of the future demand.

In order to make it easier to analyze and discuss the result, we again focus on the time-series from 8 am, 12 pm and 24 pm representing morning, day and night.

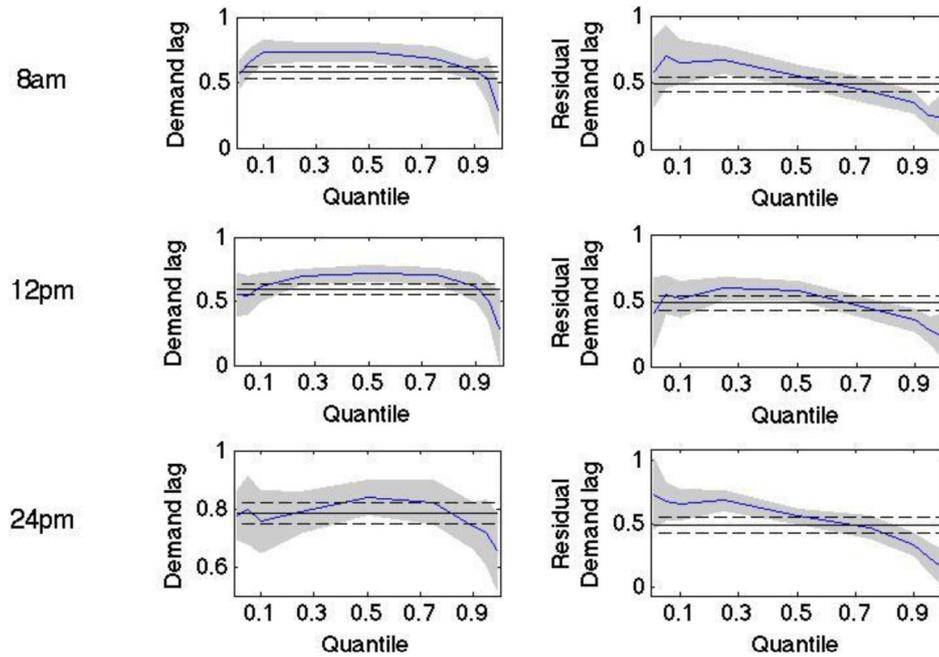


Fig. 3. Ordinary and quantile regression estimates for demand and residual demand. The shaded area represents 95% confident band for the blue quantile regression line. The dotted lines are the confident band for the ordinary regression estimates.

Figure 3 presents the conditional quantile regression estimates of the lagged demand and lagged residual demand for 8 am, 12 pm and 24 pm. For most quantiles the estimates are outside the ordinary least square confident band, suggesting that the ordinary regression estimates is not sufficient. The general pattern for the demand is an inverse U-shape, meaning that extreme quantiles (5th and 95th) have lower values than median, 25th and 75th quantiles. However, this is not the case for residual demand. In this case upper quantiles have smaller values than lower quantiles. This is in accordance with our previous discussion.

5 Conclusions

We studied the differences in modelling demand and residual demand (the difference between demand and price inelastic solar and wind production) using both ordinary regressions and quantile regressions. We find that residual demand is less predictable than demand. Similarly, the relationship between current and future value is weaker for residual demand than for demand. This result is confirmed in the quantile regression model. The quantile regression model shows that the conditional quantiles of the residual demand are more widely spread than conditional quantiles of demand. The effect was visible for all hours, and was higher for the lower than for the upper quantiles. Our results do not only confirm that the renewables lead to more challenge in predicting the load, but also illustrate how this challenge can be addressed. Our findings have implications for future research on the demand modelling, particularly in countries with increasing renewables infeed.

We recognize the Norwegian Research Centre CenSES, Centre for Sustainable Energy Studies (RCN grant 209697), and acknowledge financial support from the Research Council of Norway through project 228811.

References

1. A. Baillo, M. Ventosa, M. Rivier, A. Ramos, Optimal offering strategies for generation companies operating in electricity spot markets, *IEEE Trans. Power Syst.* **19**, 745 (2004)
2. S. Vázquez, P. Rodilla, C. Batlle, Residual demand models for strategic bidding in European power exchanges: revisiting the methodology in the presence of a large penetration of renewables, *Electr. Power Syst. Res.* **108**, 178 (2014)
3. A. Wagner, Residual demand modeling and application to electricity pricing, Available at SSRN 2018908, 2012
4. A. Motamedi, C. Geidel, H. Zareipour, W.D. Rosehart, Electricity price forecasting considering residual demand, in *2012 3rd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies (ISGT Europe)*, October (IEEE, 2012), p. 1
5. W.P. Schill, Residual load, renewable surplus generation and storage requirements in Germany, *Energy Policy* **73**, 65 (2014)
6. R. Koenker, G. Basset, Regression quantiles, *Econometrica* **46**, 33 (1978)
7. L.I. Hagfors, E. Kristoffersen, T. Staver, D. Bunn, S. Westgaard, Modelling the UK electricity price distributions using quantile regression, Working Paper, 2015

8. S. Hammoudeh, D.K. Nguyen, R.M. Sousa, Energy prices and CO₂ emission allowance prices: a quantile regression approach, *Energy Policy* **70**, 201 (2014)
9. N. Kaza, Understanding the spectrum of residential energy consumption: a quantile regression approach, *Energy Policy* **38**, 6574 (2010)
10. C.C. Lee, J.H. Zeng, The impact of oil price shocks on stock market activities: asymmetric effect with quantile regression, *Math. Comput. Simul.* **81**, 1910 (2011)
11. E.A. Feinberg, D. Genethliou, Load forecasting, in *Applied Mathematics for Restructured Electric Power Systems* (Springer, USA, 2005), p. 269

Cite this article as: Linh Phuong Catherine Do, Lars Ivar Hagfors, Kuan-Heng Lin, Peter Molnár, Demand and residual demand modelling using quantile regression, *Renew. Energy Environ. Sustain.* **1**, 41 (2016)