Stochastic Optimization of Power Market Forecast Using Non-Parametric Regression Models

Saahil Shenoy[1] and Dimitry Gorinevsky[2]

***Abstract*— The paper considers stochastic optimization of the electricity procurement in the day-ahead power market. The novelty is in addressing the random errors of time series forecasting of electrical power loads and prices in the procurement. This problem is currently important because of the increased random variability in the power grid that is caused by growing integration of renewable generation. This paper presents a methodology for stochastic optimization using data-driven models. We consider non-parametric models of multivariate distributions based on multiple quantile regressions, built from historical data sets. The statistics, such as cost expectation, required for the stochastic optimization are computed numerically using these models. Applying the methodology to utility data shows that 2% improvement of the costs is feasible.**

I. INTRODUCTION

The growing integration of renewable generation increases random variability of electrical power loads and prices in the power grid. The problem requires new analytical approaches. This paper discusses stochastic optimization in the electrical power market using time series forecasting models and stochastic models of the forecast errors. The non-parametric stochastic models are built from historical operation data.

Stochastic optimization is used in power market applications, such as wind power generation and storage in the market, see [1], [2]. It is also used for power procurement in the day ahead electricity markets, see [3], [4]. Stochastic optimization using Gaussian models of the prices has been considered in [5], [6], [7]. These papers do not address other distribution shapes. The demand and spot prices are non-Gaussian; their peaking is described by long tails. This paper presents data-driven stochastic models for the optimization.

A flexible non-parametric multivariate model is given by quantile regression, see [8]. Quantile regression is used in many applications and is included with major statistical packages as a software function. However, so far, it found limited use for building non-parametric multivariate distribution models that can be used in stochastic calculus.

This paper applies an optimization-based approach described in [9] to simultaneous estimation of multiple quantiles in the non-parametric model. The approach is scalable to extremely large training data sets because it uses the alternating direction method of multipliers (ADMM), see [10], [11]. For high (or low) quantiles, where the data is scarce, the approach uses a parametric model of the tail with fixed quantile regression slope, similar to, e.g., [12].

The stochastic optimization based on a non-parametric multivariate quantile regression model of the entire distribution appears to be new. Existing quantile regression uses are mostly for off-line analysis of the data such as risk estimation in finance [13]. The model for wind power generation in [14] could be used for forecasting; the additive quantile regression model separates each regressor variable and is not truly multi-variable.

The contributions of this paper are as follows. First, it formulates a stochastic optimization approach to power procurement using non-parametric models of the distributions with long tails. Second, it presents a computational method for building the model from historical data by formulating and solving a convex optimization problem. Third, it demonstrates the approach for electrical utility data.

As the example, we use time series of hourly electrical power loads and spot market electricity prices collected over several years. The non-parametric models of the proposed form are trained on older historical data, then backtested for stochastic optimization of power purchase decisions in the day-ahead power market. It is shown that using proposed stochastic optimization approach could lead to 2% savings.

II. SINGLE QUANTILE REGRESSION

We will consider a dataset

$$D = \{Z_i, y_i\}_{i=1}^{N}, \tag{1}$$

where scalars y_i are response variables and vectors $Z_i \in \Re^n$ are explanatory variables (regressors). Index i describes the sample and N is the number of samples available, which can be very large. In what follows, we assume that data (1) are i.i.d., and follow unknown (but fixed) conditional multivariate distribution $p(y_i|Z_i)$. In forecasting applications, i is the time sample and the i.i.d. assumption means time-invariance of the underlying process.

A. Quantile Regression Problem

We assume that the unknown generating distribution $p(y_i|Z_i)$ for (1) can be characterized through the probabilities

$$\mathbf{P}\left(y_i \le y(q)|Z_i\right) = q, \quad y(q) = Z_i\beta(q) + \alpha(q), \tag{2}$$

where $q \in (0,1)$ is the quantile level, $\beta \in \Re^n$ and $\alpha \in \Re$ define a quantile regression hyperplane in the data space.

For a given quantile level q, model (2) can be found by solving a linear programming (LP) problem, see [8]. We build a smoothed optimization formulation for joint

[1]Saahil Shenoy is a PhD student in the Department of Physics, Stanford University, Stanford, CA 94305, USA, saahils@stanford.edu

[2]Dimitry Gorinevsky is with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA, gorin@stanford.edu and Mitek Analytics LLC, Palo Alto, CA 94306, dimitry@mitekan.com

estimation of multiple quantile levels that deals with the quantile crossing problem. The formulation yields a QP problem and is described in [9]. The paper [9] discusses a scalable implementation of ADMM (see [10], [11]) to solve the QP problem for a large number of the training data points. Figure 1 illustrates the model training logic. The model is computed from Input, the dataset (1). Quantile Grid refers to set of quantile levels $\{q_j\}_{j=1}^{n_q}$, where the quantile regression is computed; n_q is the number of the levels in the grid. The ADMM Solver takes input data (1) and the quantile grid and outputs Quantile Regression Model described by the parameter set $\{\beta(q_j), \alpha(q_j)\}_{j=1}^{n_q}$.

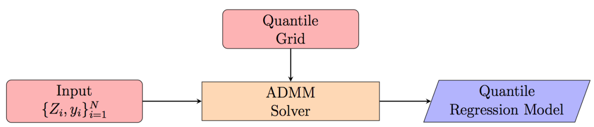

Fig. 1. Quantile regression solution logic.

B. Large Response Variable

The described smoothed model interpolates the data, where it is available. The distribution tails, i.e., the large absolute values of response variables y have to be modeled separately. The stochastic models for very low or very high quantiles can be extrapolated beyond the data range if their parametric form is known. Extreme Value Theory (EVT), predicts that in many cases the distribution tails, which describe the extreme events, follow a Pareto (power law) distribution. The tails can be estimated using peaks over threshold (POT) method.

The application examples of Sections III and V, use log coordinates. Thus, the Pareto distribution becomes an exponential distribution. Consider the first $q_L = q_1$ and the last $q_R = q_{n_q}$ quantile levels on the quantile modeling grid as the tail thresholds. The POT method exceedances are

$$e_{L,j} = y_j - Z_j\beta_1 - \alpha_1, \quad j \in J_L, \tag{3}$$

$$e_{R,k} = y_k - Z_k\beta_{n_q} - \alpha_{n_q}, \quad k \in J_R, \tag{4}$$

where $J_L \equiv \{j : y_j < Z_j\beta_1 + \alpha_1\}$ and $J_R \equiv \{k : y_k > Z_k\beta_{n_q} + \alpha_{n_q}\}$.

We model the probability distributions of $e_{L,j}$ and $e_{R,k}$ as

$$-e_{L,j} \sim q_L \cdot \mathrm{Exp}(\theta_L), \quad e_{R,k} \sim q_R \cdot \mathrm{Exp}(\theta_R). \tag{5}$$

The parameters θ_L and θ_R are estimated as a part of the smoothed optimization formulation for the quantile model; see [9] for more detail.

III. Power Load Model Example

The motivating example for development of the proposed non-parametric approach is modeling of electrical power demand for a utility. The hourly load and price data from an anonymous US utility are described in [15]. The modeling methodology was applied to the total system load. The range of the loads is 11.544 to 33.222 GW, with the average value being 18.0166 GW. The data covers a time range from January 2011 to June 2013 with sampling interval of one hour, $N = 21,696$ samples at all.

Let P_t be the load demand. The data is sampled every hour and t is the number of hours elapsed since the start of the data collection. We use logarithmic load, normalized by $P_0 = 1$ GW, as response variable $y_{P,t}$

$$y_{P,t} = \log\left(P_t/P_0\right). \tag{6}$$

We used 45 non-linear regressors Z_t that depend on time, previous load values, and previous prices. The regressors that depend on time indicate which day of the week, month, and hour of the day it is, and whether it is a holiday. We use the log load and log price values from 24 hours ago for the effect of day ahead forecasting. More detail of the time related regressors can be found in [16].

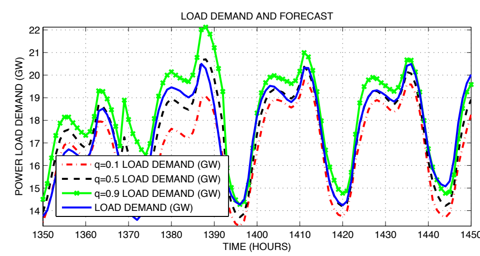

Fig. 2. Plot of quantile regression load forecast.

Median regression, described by (2) with $q = 1/2$, is used to illustrate our model for forecasting of the power load data in Figure 2, over a period of 101 hours. This model can be expressed as $y_{P,t} = Z_t\beta(1/2) + \alpha(1/2)$. We also plot other quantile forecasts of the power load distribution for $q = 0.1$ and $q = 0.9$. One can see that the median regression forecast matches the data reasonably well.

We choose a uniform quantile grid spacing of 0.01 from $q_1 = 0.01$ to $q_{n_q} = 0.99$. Figure 3 shows the PP (probablity-probablity) plot illustrating the accuracy of the data fit for the developed model. The abscissa is quantile level q in the fitted model (2), $\mathbf{P}(y_t < Z_t\beta_P(q) + \alpha_P(q)) = q$. The ordinate is the empirical quantile level estimated as the fraction of the data points in the set where the inequality $y_t < Z_t\beta_P(q) + \alpha_P(q)$ holds. The full description of the load model parameters is described in [9].

IV. Price Model Example

Dataset [15] described in Section III includes the spot price of the electricity. Let π_t be the price at time t. The price ranges from \$12.52 to \$363.80, with average value being \$48.51. It is sampled at an hourly rate, the same as the load.

We model the logarithmic price $y_{\pi,t} = \log\pi_t$. In what follows, we consider the following two conditional quantile models $\pi_t(q|Z_t)$ and $\pi_t(q|y_{P,t}, Z_t)$. The first model $\pi_t(q|, Z_t)$ has the same form as the load forecasting model in Section III. Similar to Figure 2, Figure 4 shows the quantile regression forecast of the electricity power prices for $q = 0.1, 0.5, 0.9$ along with the actual prices for a 100 hour interval. One can see that the price forecast matches the data reasonably well.

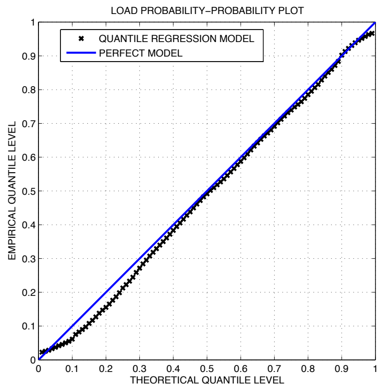

Fig. 3. PP plot for non-parametric distribution model fit for the load data.

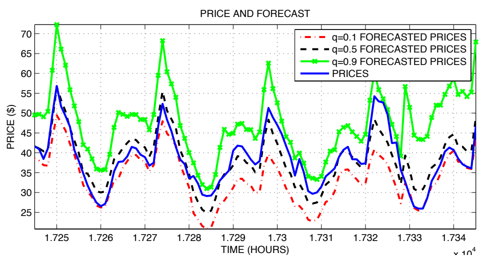

Fig. 4. Plot of quantile regression price forecast.

The second price model $\pi_t(q|y_{P,t}, Z_t)$ has the same form as the first price model $\pi_t(q|Z_t)$ with one extra regressor, y_t, added. This model assumes that the actual load is known at forecast time. We need this conditional distribution for the stochastic optimization described in Section V. We will show the extra regressor explicitly, assuming the model of the form (2) with the quantile regression hyperplane given by $Z_t\beta_\pi + y_{P,t}\gamma_\pi + \alpha_\pi$. The model is estimated similar to (2) by optimizing over decision variables $\left\{\alpha_\pi(q_i), \left[\beta_\pi(q_i)^T\ \gamma_\pi(q_i)\right]^T\right\}_{i=1}^{n_q}$. Figure 5 shows the actual prices π_t and the two median quantile models. As expected the model with the extra regressor fits the actual prices better.

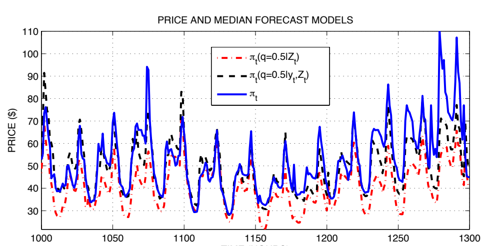

Fig. 5. Spot prices π_t vs forecasts $\pi_t(q|y_{P,t}, Z_t)$ and $\pi_t(q|Z_t)$ for $q = \frac{1}{2}$.

We used the same quantile level range and n_q quantiles as the power load models in Section III. The PP plot in Figure 6 illustrates the accuracy of the price data fit for the estimated non-parametric distribution model. The plot is similar to Figure 3. The full description of the price model parameters is described in [9].

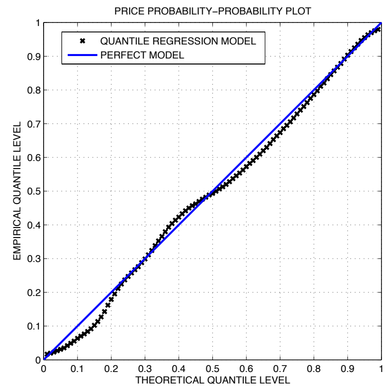

Fig. 6. PP plot for non-parametric distribution model fit for the price data.

V. Stochastic Optimization of Day-ahead Cost

The utilities order power in the electricity market a day in advance. If the actual power load is higher, the utility has to buy additional electricity at much higher spot price. The goal of the stochastic optimization approach in this section is to minimize the total expected cost. The stochastic optimization relies on the models described in Sections III and IV.

We consider stochastic optimization at given time t, when regressor Z_t is known. The future (day-ahead) log-load $y_{P,t}$ is defined by the conditional quantile model of the form (2) discussed in Section III. The future (day-ahead) log-prices $y_{\pi,t}$ are defined by the conditional quantile model of the form (2) discussed in Section IV.

The stochastic optimization requires to estimate the advance cost and the expected spot cost components of the total expected cost. These costs depend on the advance order

$$P_a(t) = P_0 e^{y_{P,t}}, \tag{7}$$

where $y_{P,t}$ is the logarithmic load that can be related to the quantile model of Section III. One can always find a quantile s such that in (7) $y_{P,t} = y_P(s)$ for Section III model

$$y_P(s) = \begin{cases} y_P(q_L) + \theta_{P,L}^{-1} \log \frac{s}{q_L}, & s < q_L \\ Z_t\beta(s) + \alpha(s), & q_L \le s \le q_R \\ y_P(q_R) - \theta_{P,R}^{-1} \log \frac{1-s}{q_R}, & s > q_R \end{cases} \tag{8}$$

The extra log terms for $s < q_L$ and $s > q_R$ come from the analytical quantile function (inverse CDF) of the exponential distribution assumed in Section II-B.

The advance cost is the deterministic value $\pi_{adv,t}P_a(t)$, where $\pi_{adv,t}$ is the advance price at time t.

$$A_t(s) = \pi_{adv,t} P_0 e^{y_{P,t}(s)}, \tag{9}$$

The spot cost is the random variable defined by the future load and spot price, in a day from the order. To compute the expected spot cost, we need the joint probability distribution over the log spot prices y_π and log load y_P conditional on the regressors Z_t, which we denote as $p(y_\pi, y_P|Z_t)$. According to the conditional probability rule,

$$p(y_\pi, y_P|Z_t) = p(y_\pi|y_P, Z_t) \cdot p(y_P|Z_t). \tag{10}$$

Similar to quantile model (8) for $y_P = y_P(q)$, we can write a quantile model for y_π. In Subsection IV, the log price y_π is modeled through the conditional distribution $p(y_\pi|y_P, Z_t)$. The conditional quantile model for y_π is then

$$y_\pi(r,q) = \begin{cases} Z_t\beta_\pi(r) + \gamma_\pi(r)y_P(q) + \alpha_\pi(r), \; r_L \le r \le r_R \\ y_\pi(r_L,q) + \log(r/r_L)/\theta_{\pi L}, \; r < r_L \\ y_\pi(r_R,q) - \log((1-r)/r_R)/\theta_{\pi R}, \; r > r_R \end{cases} \tag{11}$$

where $\theta_{\pi L}$ and $\theta_{\pi R}$ are the estimated left and right tail rates of the price distributions. The quantile level gridding for r is the same as for q in Section III.

The spot cost $C(s)$ depends on the advance order $P_a(t) = P_0 e^{y_{P,t}(s)}$ because the utility has to pay the spot price only when this order is exceeded. The spot cost expectation is

$$\mathbf{E}_{y_\pi, y_P}[C(s)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{y_\pi(r,q)} P_0 \\ \cdot \left(e^{y_P(q)} - e^{y_P(s)}\right)_+ p(y_\pi(r,q), y_P(q)) dy_\pi(r,q) dy_P(q), \tag{12}$$

where the integrand is the spot cost times the excess demand. The joint pdf in (12) can be expressed in terms of quantile levels q and r by using (10), where in accordance with (2)

$$p(y_P|Z_t) = dq/dy_P(q), \tag{13}$$
$$p(y_\pi|y_P, Z_t) = dr/dy_\pi(r,q). \tag{14}$$

Changing the integration variables in (12) to r, q and using (13), (14) yields

$$\mathbf{E}_{y_\pi, y_P}[C(s)] = P_0 \int_s^1 \int_0^1 e^{y_\pi(r,q)} \left(e^{y_P(q)} - e^{y_P(s)}\right) dr dq. \tag{15}$$

To compute integral (15), we substitute the expression (8) for $y_P(q)$ and (11) for $y_\pi(r,q)$. Each of the expressions (15), (8) breaks into three parts: the numerical model for the middle part of the distribution and two analytical models for the inverse CDFs of the tails. Their product in (15) yields the following 3×3 matrix of the partial integrals to be evaluated

$$B(s) = P_0 \begin{bmatrix} B_{L,L}(s) & B_{L,M}(s) & B_{L,R}(s) \\ B_{M,L}(s) & B_{M,M}(s) & B_{M,R}(s) \\ B_{R,L}(s) & B_{R,M}(s) & B_{R,R}(s) \end{bmatrix}, \tag{16}$$

where the subscripts L, M, and R indicate the left tail, the middle part, and the right tail respectively. The subscript combination describes an integration rectangle in the area of the integration in (15). The expected spot cost $\mathbf{E}_{y_\pi, y_L}[C(s)]$ (15) is the sum of the nine matrix entries in (16).

The logic of evaluating (15) using (16) is outlined in Figure 7. The input is the desired quantile level s at which the expected spot cost is evaluated. The two models $y_P(q)$ (8), $y_\pi(r,q)$ (11) are used to compute each matrix element of $B(s)$. The expected spot cost follows after this computation.

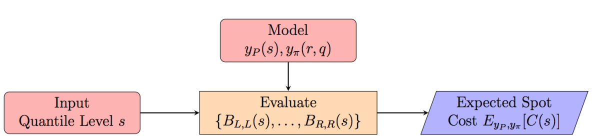

Fig. 7. Logic of evaluating expected spot cost (15) using (16).

The partial integrals in (16) can be computed as follows

$$B_{L,L}(s) = C_L G_L(s, f_L(s), \gamma_\pi(r_L)), \tag{17}$$

$$B_{L/R,M}(s) = \int_{r_L}^{r_R} e^{Z_t\beta_\pi(r)+\alpha_\pi(r)} \\ \times G_{L/R}(s, f_{L/R}(s), \gamma_\pi(r)) dr, \tag{18}$$

$$B_{L,R}(s) = C_R G_L(s, f_L(s), \gamma_\pi(r_R)), \tag{19}$$

$$B_{M,L/R}(s,s') = C_{L/R} \int_{f_M(s)}^{q_R} \Big(e^{(1+\gamma_\pi(r_{L/R}))y_P(q)} \\ - e^{y_P(s)+\gamma_\pi(r_{L/R})y_P(q)}\Big) dq, \tag{20}$$

$$B_{M,M}(s) = \int_{f_M(s)}^{q_R}\int_{r_L}^{r_R} e^{y_\pi(r|q)} \left(e^{y_P(q)} - e^{y_P(s)}\right) dr dq, \tag{21}$$

$$B_{R,L}(s) = C_L G_R(s, f_R(s), \gamma_\pi(r_L)), \tag{22}$$

$$B_{R,R}(s) = C_R G_R(s, f_R(s), \gamma_\pi(r_R)), \tag{23}$$

where the index L/R means that the expression is valid for either index L or R. The integrals in (20)–(21) are evaluated numerically for given s using the smoothed non-parametric models described above. The integrals involving the tails have been evaluated analytically. For the right tail integral to converge, the tail parameters must satisfy $\theta_{\pi,R} > 1$ and $\theta_{P,R} > \max_r\{\gamma_\pi(r)\}$. For the example dataset, these tail parameters were $\theta_{P,L} = 39.2248$, $\theta_{P,R} = 31.7821$, $\theta_{\pi,L} = 6.9101$, and $\theta_{\pi,R} = 7.2108$. The left tail integrals exist for any $s > 0$.

Formulas (17)–(23) include several interim expressions described below. The function $G_{L/R}(\cdot,\cdot,\cdot)$ is computed as

$$G_{L/R}(s, s', \gamma) = F_{L/R}(s', 1+\gamma) - e^{y_L(s)} F_{L/R}(s', \gamma), \tag{24}$$

$$F_L(s,\gamma) = \int_s^{q_L} e^{\gamma y_L(q)} dq \\ = \frac{q_L^{-\gamma/\theta_{P,L}} e^{\gamma y_P(q_L)} \theta_{P,L}}{\gamma + \theta_{P,L}} \left[q_L^{(\gamma+\theta_{P,L})/\theta_{P,L}} - s^{(\gamma+\theta_{P,L})/\theta_{P,L}}\right],$$

$$F_R(s,\gamma) = \int_s^1 e^{\gamma y_P(q)} dq \\ = q_R^{\gamma/\theta_{P,R}} e^{\gamma y_P(q_R)} \frac{\theta_{P,R}}{\theta_{P,R} - \gamma} (1-s)^{(\theta_{P,R}-\gamma)/\theta_{P,R}}.$$

The constants C_L and C_R in (17)–(23) are

$$C_R = e^{Z_t\beta_\pi(r_R)+\alpha_\pi(r_R)} \frac{\theta_{\pi,R} r_R^{1/\theta_{\pi,R}}}{\theta_{\pi,R} - 1} (1 - r_R)^{\frac{\theta_{\pi,R}-1}{\theta_{\pi,R}}},$$
$$C_L = e^{Z_t\beta_\pi(r_L)+\alpha_\pi(r_L)} \theta_{\pi,L} r_L / (\theta_{\pi,L} + 1).$$

The saturation functions in (17)–(23) are

$$f_L(s) = \min(s, q_L), \tag{25}$$
$$f_R(s) = \max(s, q_R), \tag{26}$$
$$f_M(s) = \min(\max(s, q_L)), q_R). \tag{27}$$

The total cost $T(s)$ can be computed from (9) and (15) as

$$T(s) = A(s) + \mathbf{E}[C(s)] \tag{28}$$

Based on (9), advance cost $A(s)$ is a non-decreasing function of s. Based on (15), the expected spot cost $\mathbf{E}[C(s)]$ is a non-increasing positive function of s. The optimal trade-off between the advance cost and the spot cost that minimizes the total cost for some s can be found numerically by computing $T(s)$ (28) on a grid of $s \in (0, 1)$.

VI. Cost Optimization Example

The non-parametric models for the load and price were trained on the utility data set as described in Sections III and IV. The stochastic optimization algorithm was then backtested on the last 6 months of the data, which were excluded from the model training. The constant advance price $\pi_{adv,t}$ = \$50/MWh was assumed. At each step, the optimal quantile level s_* was found by computing total cost $T(s)$ (28) on a grid of s using formulas (17)–(23). A time series segment for the computed optimal quantile level $s_*(t)$ is shown in Figure 8, where the lowest value is $s_* = 0.001$.

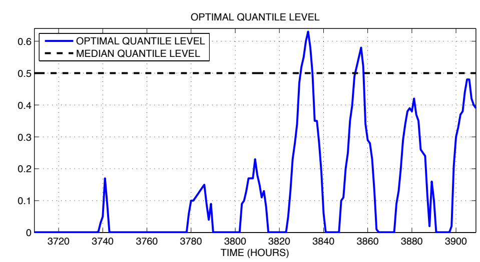

Fig. 8. Optimal quantile level sequence sample.

Figure 9 shows the actual realization of the total cost in backtesting of the advance procurement strategy that uses the optimized $s = s_*$ in (7), (8). The results from May 25, 2013 to June 2, 2013 are shown. The described optimized procurement based on the smoothed quantile regression (QR) model was compared to the baseline strategy based on the ordinary least squares (OLS) regression. The baseline model used the same regressors. The OLS regression parameters were estimated from the same data as the QR model. The baseline procurement strategy used the linear regression forecast of the log-load in (7). Figure 9 also includes the median regression strategy that used $s = 0.5$ in (7), (8). In Figure 9, when the smoothed QR strategy has a lower total cost than the baseline, it is by a significant amount; when cost is higher than OLS, it is by a much smaller amount.

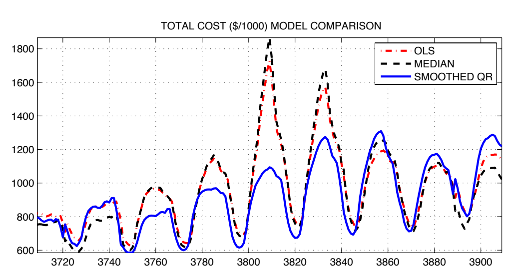

Fig. 9. The hourly total cost over a sample 200 hour period

The total cost for the 6 month period for each strategy is summarized in Table I. The total savings of the optimized strategy based on the smoothed QR model are \$85.666 million compared to the OLS baseline. The median regression strategy saves \$11.203 million, it is much closer to the OLS.

TABLE I

Cost Results (\$/Billion)

Strategy/Model	Smoothed QR	Median	OLS
Total 6-Month Cost	3.83	3.91	3.92
Percentage Savings	2.19%	0.29%	0%

References

[1] C. Abbey and G. Joós, "A stochastic optimization approach to rating of energy storage systems in wind-diesel isolated grids," *IEEE Transactions on Power Systems*, vol. 24, no. 1, pp. 418–426, 2009.

[2] J. Garcia-Gonzalez, R. R. de la Muela, L. M. Santos, and A. M. González, "Stochastic joint optimization of wind generation and pumped-storage units in an electricity market," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 460–468, 2008.

[3] S. Wong and J. D. Fuller, "Pricing energy and reserves using stochastic optimization in an alternative electricity market," *IEEE Transactions on Power Systems*, vol. 22, no. 2, pp. 631–638, 2007.

[4] S. Sen, L. Yu, and T. Genc, "A stochastic programming approach to power portfolio optimization," *Operations Research*, vol. 54, no. 1, pp. 55–72, 2006.

[5] R. Bjorgan, C.-C. Liu, and J. Lawarree, "Financial risk management in a competitive electricity market," *IEEE Transactions on Power Systems*, vol. 14, no. 4, pp. 1285–1291, 1999.

[6] A. J. Conejo, F. J. Nogales, J. M. Arroyo, and R. García-Bertrand, "Risk-constrained self-scheduling of a thermal power producer," *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1569–1574, 2004.

[7] E. Tanlapco, J. Lawarrée, and C.-C. Liu, "Hedging with futures contracts in a deregulated electricity industry," *IEEE Transactions on Power Systems*, vol. 17, no. 3, pp. 577–582, 2002.

[8] R. Koenker, *Quantile regression*. No. 38, Cambridge university press, 2005.

[9] S. Shenoy, D. Gorinevsky, and S. Boyd, "Non-parametric regression modeling for stochastic optimization of power market forecast," in *American Control Conference*, (Chicago, IL), July 2015.

[10] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[11] N. Parikh and S. Boyd, "Block splitting for distributed optimization," *Mathematical Programming Computation*, vol. 6, no. 1, pp. 77–102, 2014.

[12] H. J. Wang, D. Li, and X. He, "Estimation of high conditional quantiles for heavy-tailed distributions," *Journal of the American Statistical Association*, vol. 107, no. 500, pp. 1453–1464, 2012.

[13] L. Qian, L. Yongli, and W. Chong, "The risk linkage effects of stock indexes based on quantile regression and granger causality test," in *25th Chinese Control and Decision Conference (CCDC)*, pp. 4252–4257, May 2013.

[14] H. A. Nielsen, H. Madsen, and T. S. Nielsen, "Using quantile regression to extend an existing wind power forecasting system with probabilistic forecasts," *Wind Energy*, vol. 9, no. 1-2, pp. 95–108, 2006.

[15] Crowdanalytix.com, "Global energy forecasting competition 2014 probabilistic electricity price forecasting." Available: `https://crowdanalytix.com/contests/global-energy-forecasting-competition-2014-probabilistic-electricity-price-forecasting#`.

[16] S. Shenoy and D. Gorinevsky, "Risk adjusted forecasting of electric power load," in *American Control Conference (ACC)*, pp. 914–919, IEEE, 2014.