1 Simulation Study

We compare the performance of our quantile trend filtering model with the models of 3 others using designs proposed by Racine and Li (2017). The models compared are

- `npqw`: Racine and Li (2017) constrain the response to follow a smooth location scale model of the form $Y_i = a(X_i) + b(X_i)\epsilon_i$. They estimate the τ_{th} conditional quantile given $X_i = x$ using a kernel estimator

$$q_\tau(x) = \frac{\sum_{i=1}^n \Phi^{-1}_{(Y_i, b(X_i))}(\delta_0) K_h(X_i, x)}{\sum_{i=1}^n K_h(X_i, x)} \tag{1}$$

  defining $\Phi^{-1}_{(Y_i, b(X_i))}(\delta_0)$ as the quantile function of the Normal distribution with mean Y_i and standard deviation $b(X_i)$ evaluated at τ, δ_0 is a function of τ and chosen empirically, h is a tuning parameter and K is a kernel function. Code was obtained from the author for the `quantile-ll` method.

- `qsreg`: Oh et al. (2011) proposed a pseudo-data algorithm for a quantile spline estimator of the form

$$\sum_i \rho_\tau(y_i - g(x_i)) + \lambda \int (g''(x))^2 dx. \tag{2}$$

  If $\rho_\tau(\cdot)$ were differentiable, the solution to this equation would take a form similar to that of the squared loss smoothing spline with weights equal to $\frac{\rho'_\tau(y_i - g(x_i))}{2(y_i - g(x_i))}$. Relying on this idea, Nychka proposed to solve the problem by iteratively solving the weighted smoothing spline. To address the non-differentiability they propose an approximation

$$\rho_{\tau,\delta}(u) = [\tau I(u > 0) + (1 - \alpha) I(u < 0)] u^2/\delta \tag{3}$$

  The function `qsreg` in the textttfields R package was used.

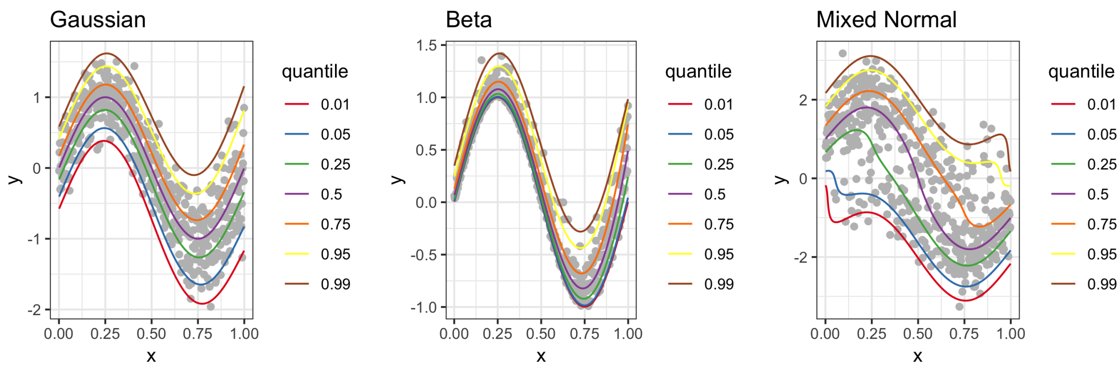
- `rqss`: Koenker et al. (1994) Koenker proposed smoothing splines using trend filtering with the second order differencing matrix which results in linear splines. The function `rqss` in the `quantreg` package implements this method. The smoothing parameter λ is chosen using a grid search and minimizing

$$SIC(p_\lambda) = log[n^{-1} \sum \rho_\tau(y_i - \widehat{g}(x_i))] + \frac{1}{2n} p_\lambda logn \tag{4}$$

  where $p_\lambda = \sum I(y_i = \widehat{g}_i(x_i))$, which can be thought of as active knots.

- `detrendr_SIC`: Our method where we minimize $\sum_i \rho_\tau(y_i - \theta_i) + \lambda ||D\theta||_1$ and λ is chosen using SIC from above.

Figure 1: Simulated data with true quantiles $\tau \in \{0.01, 0.05, 0.25, 0.5, .75, 0.95, 0.99\}$



- `detrendr_valid`: Our method where lambda is chosen by leaving out every 5th observation as a validation data set and evaluating the check loss function on the validation data.

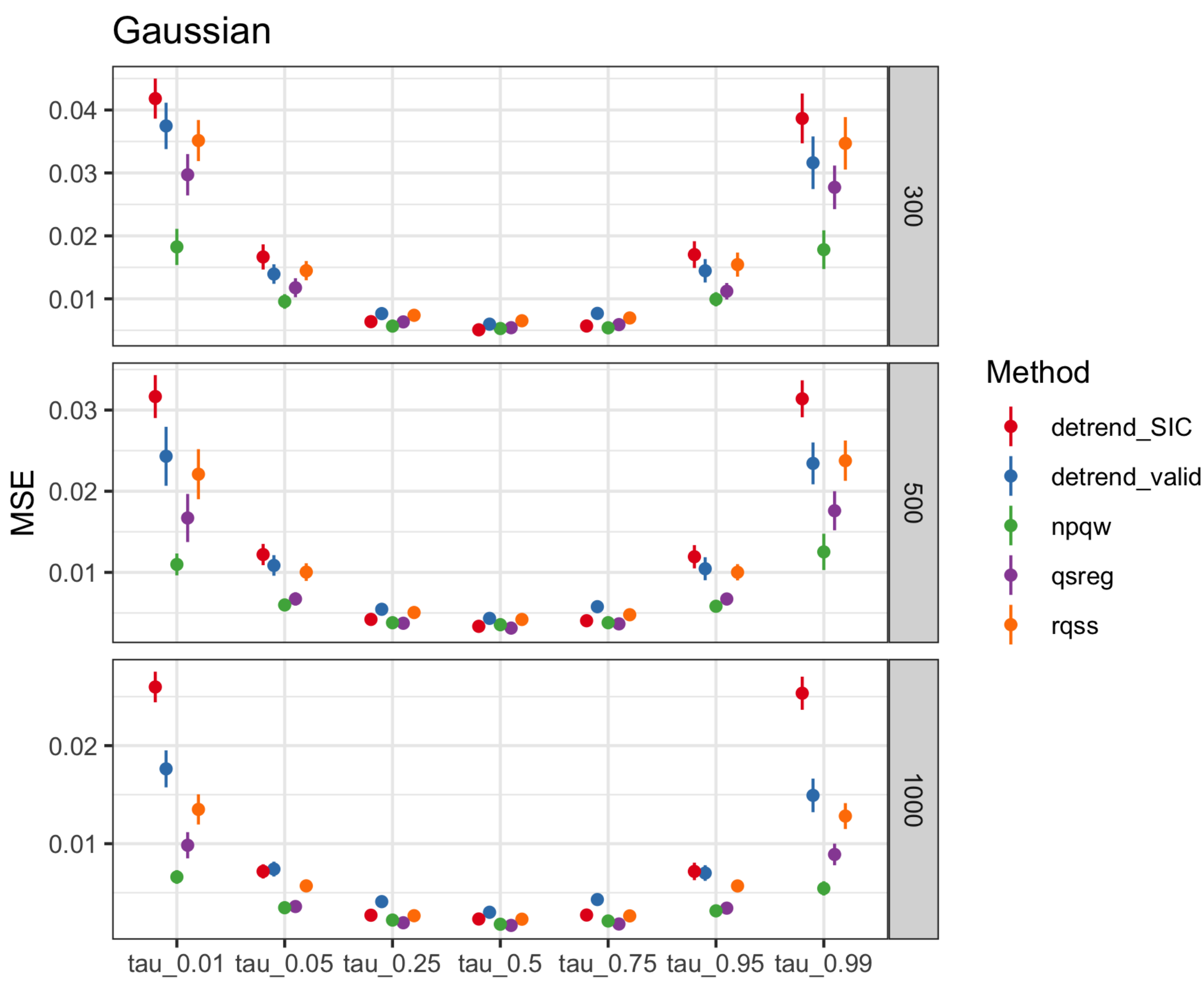
Three simulation designs from Racine and Li (2017) were considered. For all designs X_i was generated as a uniformly space sequence in $[0, 1]$ and the mean function was given as

$$Y_i = sin(2\pi X_i) + \epsilon_i(X_i)$$

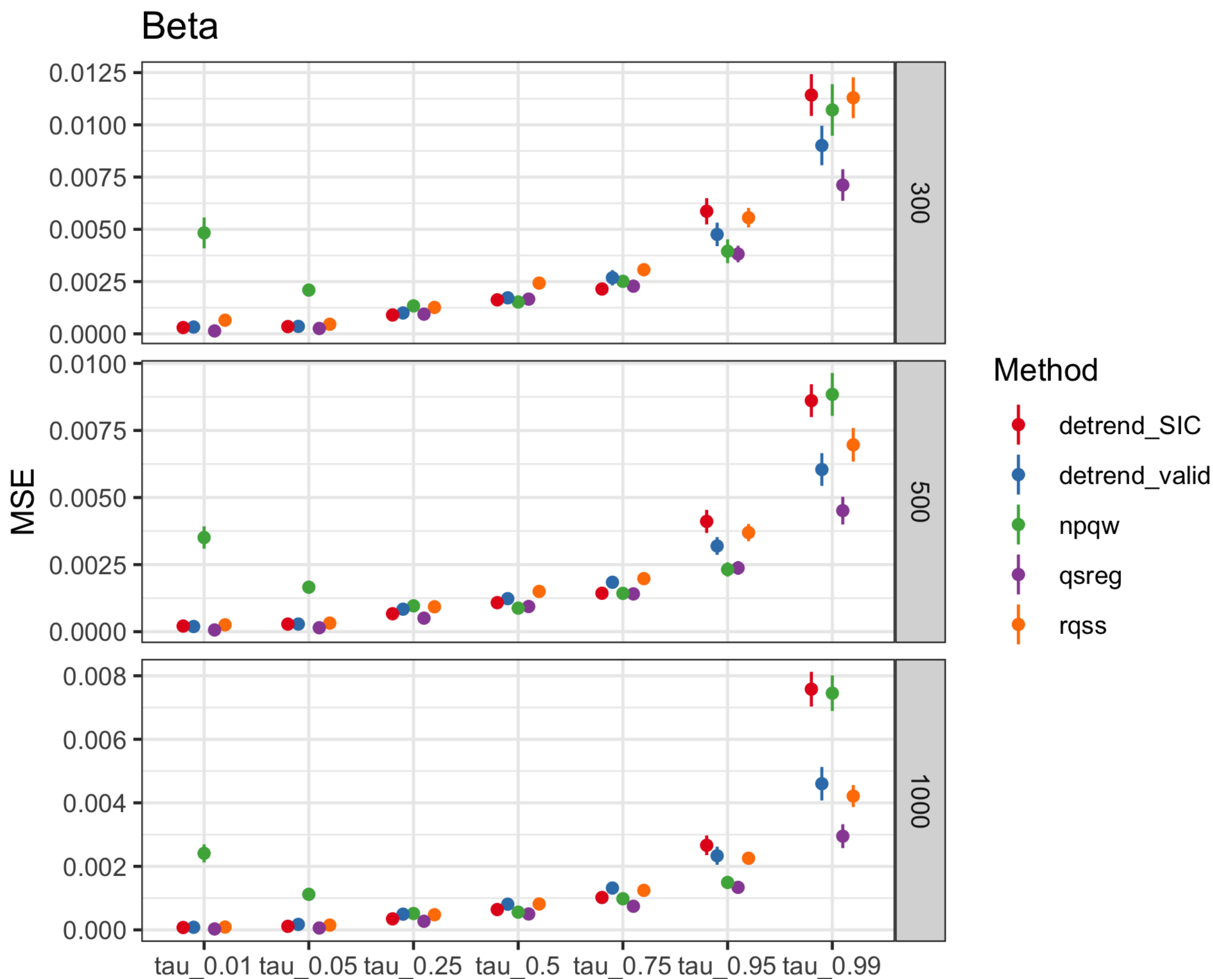
The three error distributions considered were

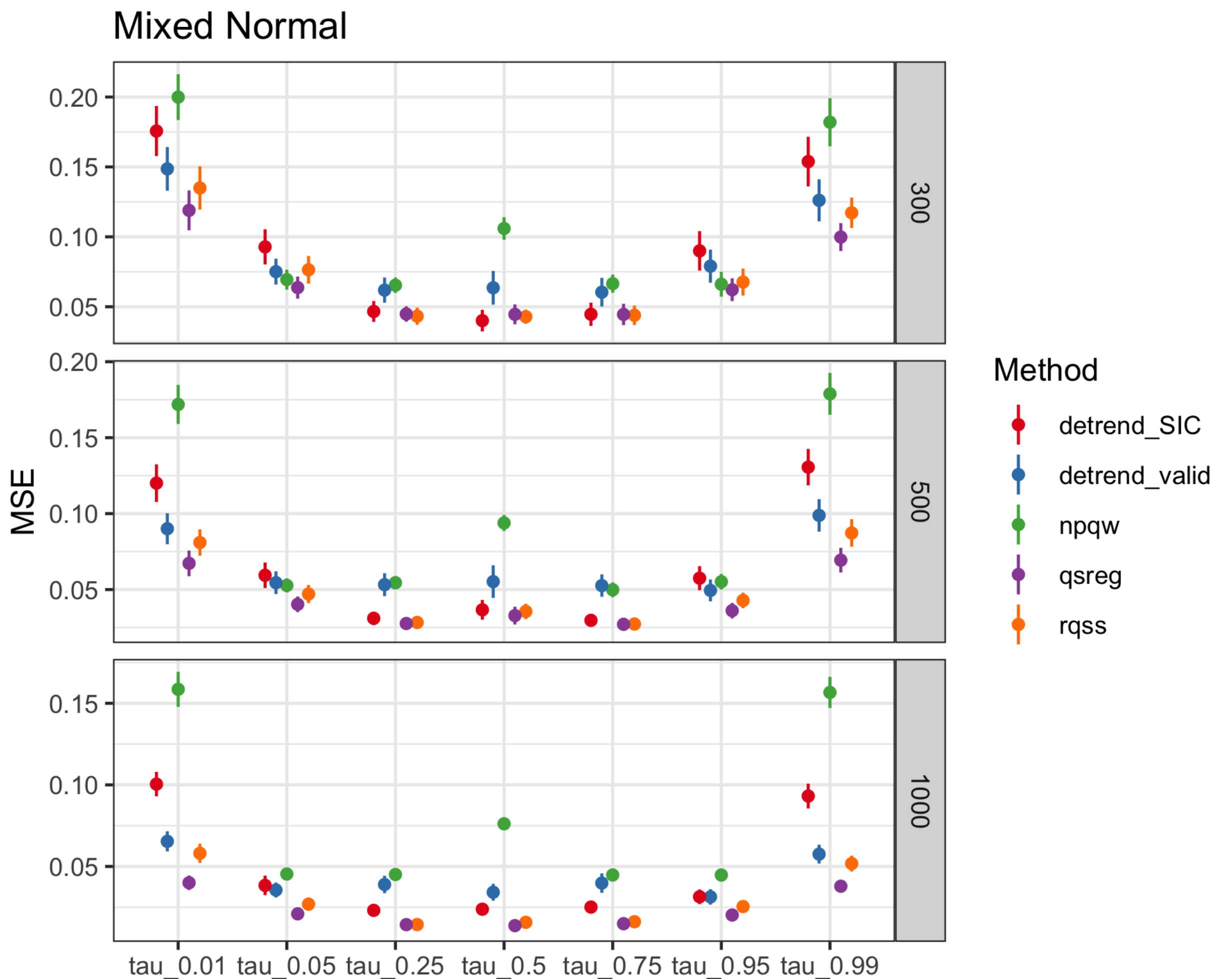
- Gaussian: $\epsilon_i \sim N\left(0, \left(\frac{1+x_i^2}{4}\right)^2\right)$
- Beta: $\epsilon_i \sim Beta(1, 11 - 10x_i)$
- Mixed normal: ϵ_i is simulated from a mixture of $N(-1, 1)$ and $N(1, 1)$ with mixing probability x_i.

100 datasets were generated of sizes 300, 500 and 1000. The MSE was calculated as $\frac{1}{n}\sum_i(\hat{q}_\tau(x_i) - q_\tau(x_i))^2$.



Gaussian
MSE
0.04
0.03
0.02
0.01
300
500
1000
tau_0.01 tau_0.05 tau_0.25 tau_0.5 tau_0.75 tau_0.95 tau_0.99
Method
detrend_SIC
detrend_valid
npqw
qsreg
rqss


Beta
MSE
0.0125
0.0100
0.0075
0.0050
0.0025
0.0000
0.0100
0.0075
0.0050
0.0025
0.0000
0.008
0.006
0.004
0.002
0.000
300
500
1000
tau_0.01 tau_0.05 tau_0.25 tau_0.5 tau_0.75 tau_0.95 tau_0.99
Method
detrend_SIC
detrend_valid
npqw
qsreg
rqss



References

Koenker, R., Ng, P., and Portnoy, S. (1994), "Quantile smoothing splines," *Biometrika*, 81, 673–680.

Oh, H.-S., Lee, T. C. M., and Nychka, D. W. (2011), "Fast Nonparametric Quantile Regression With Arbitrary Smoothing Methods," *Journal of Computational and Graphical Statistics*, 20, 510–526.

Racine, J. S. and Li, K. (2017), "Nonparametric conditional quantile estimation: A locally weighted quantile kernel approach," *Journal of Econometrics*, 201, 72–94.