A Bayesian Method for Non-Gaussian Autoregressive Quantile Function Time Series Models

Yuzhi Cai

University of Plymouth
United Kingdom

18 June 2008

Outline of The Talk

- Quantile Function
- AR Quantile Function Time Series Models
- Bayesian Approach
- Simulation Studies
- Applications
- Conclusions and Future Work

Quantile Function

- **Definition**
  - X is a continuous random variable
  - $F(x)$ is the distribution function of X

Then $x = Q(\tau) = F^{-1}(x)$ is the quantile function of X, where $0 \leq \tau \leq 1$.

Quantile Function

- **Properties**
  - Sum of two quantile functions is a quantile function
  - Product of two positive quantile functions is a quantile function
  - If $Q_1(\tau)$ and $Q_2(\tau)$ are two quantile functions, then

    $$Q(\tau) = wQ_1(\tau) + (1-w)Q_2(\tau), \quad 0 < w < 1$$

    lies between the two quantule functions.
  - If $Q(\tau)$ is a quantile function, then $\lambda + \eta Q(\tau)$ is also a quantile function.

Quantile Function

- **Examples**
  - Exponential distribution:

$$F(x) = 1 - e^{-\lambda x}, \quad x \geq 0, \; \lambda > 0$$

$$Q(\tau) = -\frac{1}{\lambda}\ln(1-\tau), \quad 0 \leq \tau \leq 1, \; \lambda > 0.$$

  - Power distribution:

$$F(x) = x^{1/\gamma_1}, \quad 0 \leq x \leq 1, \; \gamma_1 > 0.$$

$$Q_1(\tau) = \tau^{\gamma_1}, \quad \gamma_1 > 0, \; 0 \leq \tau \leq 1$$

Quantile Function

- **Examples**
  - Pareto distribution:

$$F(x) = 1 - x^{-1/\gamma_2}, \quad x \geq 1, \ \gamma_2 > 0.$$

$$Q_2(\tau) = (1-\tau)^{-\gamma_2}, \quad 0 \leq \tau < 1, \ \gamma_2 > 0.$$

  - Power-Pareto distribution:

$$Q(\tau) = Q_1(\tau)Q_2(\tau) = \tau^{\gamma_1}(1-\tau)^{-\gamma_2}, \quad \gamma_1 > 0, \ \gamma_2 > 0$$

The distribution function (inverse of $Q(\tau)$) does not have an explicit mathematical form!

AR Quantile Function Time Series Models

Let $(y_1, y_2, \ldots, y_n)$ be an observed time series.

- **Usual AR time series model:**

$$y_t = a_0 + a_1 y_{t-1} + \cdots + a_k y_{t-k} + \epsilon_t, \tag{1}$$

  where

  - ϵ_t are iid random variables with zero mean, constant variance, and given probability density/distribution function .
  - Parameter estimation: maximum likelihood, MCMC ...
  - Model (1) provides an estimate of the mean of y_t conditional on $\mathbf{y}_{t-1} = (y_1, \ldots, y_{t-1})$.

AR Quantile Function Time Series Models

- **Quantile AR model (Koenker, 2005):**

$$q^{\tau}_{y_t|\mathbf{y}_{t-1}} = a^{\tau}_0 + a^{\tau}_1 y_{t-1} + \cdots + a^{\tau}_k y_{t-k}, \qquad (2)$$

  where
  - a^{τ}_j, $j = 0, \ldots, k$, are the model parameters which depend on τ, $0 \leq \tau \leq 1$
  - Model (2) provides an estimate of the τ^{th} quantile of y_t conditional on $\mathbf{y}_{t-1}$.
  - Model (2) does not contain an error term ϵ_i, hence it is semi-parametric.

AR Quantile Function Time Series Models

- **Quantile AR model (Koenker, 2005):**
  - Parameters can be estimated by solving the minimization problem

$$\min_{\eta^\tau} \sum_{t=k+1}^{n} \rho_\tau(u_t),$$

where $\rho_\tau(u_t) = u_t(\tau - I_{[u_t<0]})$, $\eta^\tau = (a_0, \ldots, a_k)$, and

$$u_t = y_t - a_0^\tau - a_1^\tau y_{t-1} - \cdots - a_k^\tau y_{t-k}.$$

AR Quantile Function Time Series Models

- **Quantile AR model (Koenker, 2005):**
  - Main problems :
    - The number of parameters can be very large in order to estimate different quantiles of y_t
    - The estimated quantiles may violate the basic monotonic property of a quantile function
    - It is impossible to use model (2) to describe the tails of a distribution accurately. Therefore, such models are not appropriate for extreme event prediction.

AR Quantile Function Time Series Models

- **AR quantile function time series models (Gilchrist, 2000):**

$$Q_{y_t}(\tau \mid \mathbf{y}_{t-1}) = a_0 + a_1 y_{t-1} + \cdots + a_k y_{t-k} + Q(\tau, \gamma), \quad (3)$$

  - $Q(\tau, \gamma)$ is the quantile function of the error term ϵ_t
  - γ is the parameter vector involved in Q
  - Model parameters can be estimated by
    - distributional weighted least absolutes method
    - distributional weighted least squares method
    - maximum likelihood estimation method etc.

AR Quantile Function Time Series Models

- **Special AR quantile function time series models:**
  - Exponential quantile function model:

$$Q_{y_t}(\tau \mid \mathbf{y}_{t-1}) = a_0 + a_1 y_{t-1} + \cdots + a_k y_{t-k} - \frac{1}{\gamma} \ln(1 - \tau)$$

  - Pareto quantile function model:

$$Q_{y_t}(\tau \mid \mathbf{y}_{t-1}) = a_0 + a_1 y_{t-1} + \cdots + a_k y_{t-k} + (1 - \tau)^{-\gamma_2}$$

  - Power-Pareto quantile function model:

$$Q_{y_t}(\tau \mid \mathbf{y}_{t-1}) = a_0 + a_1 y_{t-1} + \cdots + a_k y_{t-k} + \tau^{\gamma_1}(1 - \tau)^{-\gamma_2}$$

Bayesian Approach

- **The model:**

  Consider exponential quantile function model:

  $$Q_{y_t}(\tau \mid \mathbf{y}_{t-1}) = a_0 + a_1 y_{t-1} + \cdots + a_k y_{t-k} - \frac{1}{\gamma} \ln(1-\tau)$$

  Let $\beta = (a_0, \ldots, a_k, \gamma)$ be the model parameter vector.

- **Lemma:** Let $Q(\tau)$ be the quantile function of X, $f(x)$ be the density function of X. Then

  $$f(x) = \frac{1}{dQ(\tau)/d\tau}, \quad 0 \leq \tau \leq 1.$$

- **Likelihood function:**

  The conditional likelihood of y_t, $t = k+1, \ldots, n$, given $\mathbf{y}_k$ is given by

$$L(y_n, \ldots, y_{k+1} \mid \mathbf{y}_k, \beta) = \prod_{t=k+1}^{n} f(y_t \mid \mathbf{y}_t, \beta) = \prod_{t=k+1}^{n} \gamma(1 - \tau_t),$$

  where τ_t is the solution of

$$y_t = a_0 + a_1 y_{t-1} + \cdots + a_k y_{t-k} - \frac{\ln(1-\tau)}{\gamma}. \qquad (4)$$

- **Posterior density function:**

$$\pi(\boldsymbol{\beta} \mid \mathbf{y}_n) \propto L(y_n, \ldots, y_{k+1} \mid \mathbf{y}_k, \boldsymbol{\beta})\pi(\boldsymbol{\beta}), \quad \boldsymbol{\beta} \in \Omega,$$

where

  - $\pi(\boldsymbol{\beta})$ is the prior density function of the parameters
  - Ω is the parameter space
  - In this paper, the prior density function is given by

$$\pi(\boldsymbol{\beta}) = \prod_{i=0}^{k} \pi(a_i)\pi(\gamma),$$

where $\pi(a_i)$ is a normal density function with zero mean and variance $\tilde{\sigma}^2_{a_i}$, and $\pi(\gamma) = \alpha e^{\alpha\gamma}$.

- **Theorem**

  The posterior distribution is well defined on $\Omega = \Omega_1 \times \Omega_2$, where $\Omega_1 = \{(a_0, \ldots, a_k) \mid a_0 + a_1 y_{t-1} + \cdots + a_k y_{t-k} \leq y_t,\ t = k+1, \ldots, n\}$, and $\Omega_2 = (0, \infty)$.

- **Notes:**
  - $\pi(a_i)$ can be taken as any proper density functions.
  - $\pi(\gamma)$ needs to be choose such that the posterior distribution is well defined.

- **Random walk sampler**
  - Obtain a proposed value β' from, $g(\beta, \beta')$, such that $\beta' \in \Omega$
  - Solve equation (4) for τ_t'
  - Accept the proposal with probability $\min\{AB, 1\}$, where

$$A = \frac{\pi(\beta' \mid \mathbf{y}_t)}{\pi(\beta \mid \mathbf{y}_t)} = \frac{\prod_{t=k}^{n} \gamma'(1-\tau_t')\pi(\beta')}{\prod_{t=k}^{n} \gamma(1-\tau_t)\pi(\beta)},$$

$$B = \frac{g(\beta', \beta) / \int_\Omega g(\beta', \beta) d\beta}{g(\beta, \beta') / \int_\Omega g(\beta, \beta') d\beta'}.$$

  - $$g(\beta, \beta') = \prod_{i=0}^{k} g_i(a_i, a_i') g_{k+1}(\gamma, \gamma'),$$

where $g_i(a_i, a_i') \sim N(a_i, \sigma_{a_i}^2)$ and $g_{k+1}(\gamma, \gamma') \sim N(\gamma, \sigma_\gamma^2)$.

Simulation Study

- **True model**

$$Q_{y_t}(\tau \mid \mathbf{y}_{t-1}) = a_0 + a_1 y_{t-1} + a_2 y_{t-2} - \frac{\ln(1-\tau)}{\gamma}. \tag{5}$$

where $a_0 = -0.6$, $a_1 = 0.3$, $a_2 = 0.6$ and $\gamma = 1.6$.

- **Simulated time series**

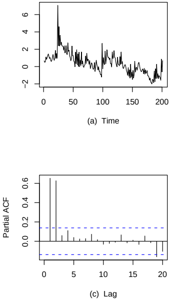


Figure: (a) Time series plot of the simulated series. (b) acf plot of the simulated series. (c) Partial acf plot of the simulated series.

- **Parameters estimation**
  - **Prior information:** $\tilde{\sigma}_{a_i} = 10$ and $\alpha = 0.5$
  - **Initial values:**

$$(a_0, a_1, a_2, \gamma) = \left(\min_{1 \le t \le n} y_t, 0, 0, \gamma_0\right) = (-2.011,\ 0,\ 0,\ 0.110),$$

    where $\gamma_0 \sim \pi(\gamma)$.
  - **Sample collection:**
    - The length of the Markov chain: 20,000 steps
    - The burn-in period: the first 10,000 steps
    - samples saved every 50 steps

- **Histograms**

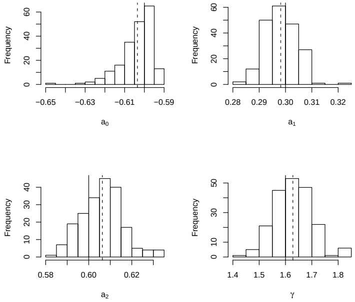


Figure: Histograms of the collected samples for each parameter in the Simulation Study. The solid vertical lines indicate the positions of the true parameter values, and the dashed vertical lines the Bayesian estimates of the parameters.

- **The fitted model**

$$Q_{y_t}(\tau \mid \mathbf{y}_{t-1}) = -0.6035 + 0.2982 y_{t-1} + 0.6063 y_{t-2} - \frac{\ln(1-\tau)}{1.6280}, \quad (6)$$

- **The fitted conditional quantiles**

  The fitted conditional quantiles from both the true model (5) and the fitted model (6) are almost the same, which can be see from the following plot.

- **The fitted conditional quantiles**

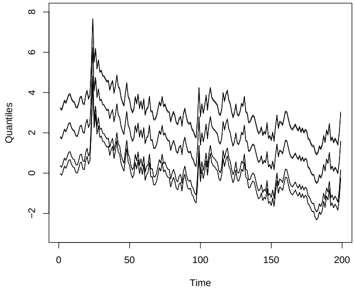


Figure: Conditional quantiles from the true model (5) and the fitted model (6) corresponding to (from the bottom to the top) $\tau = 0.05, 0.5, 0.95$ and 0.995 respectively.

- **The semi-parametric approach**
  - **The model**

$$q^{\tau}_{y_t|\mathbf{y}_{t-1}} = a^{\tau}_0 + a^{\tau}_1 y_{t-1} + a^{\tau}_2 x_{t-2}, \tag{7}$$

  - **Parameters estimation:** Use R

- **Comparisons**

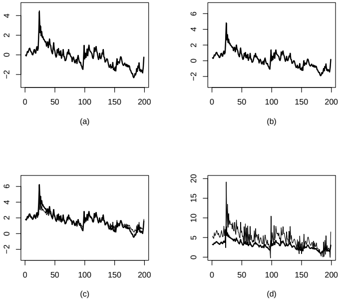


Figure: Conditional quantiles from the true model (5) (darker curves) and the fitted model (7) (lighter curves) corresponding to (a) $\tau = 0.05$, (b) $\tau = 0.5$, (c) $\tau = 0.95$ and (d) $\tau = 0.995$.

- **Models with different orders**

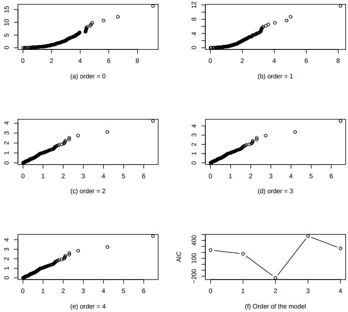


Figure: QQ-plots (a)-(e) and the AIC values (f) of the fitted models.

Applications

- **Levels of Lake Huron time series 1875-1972**

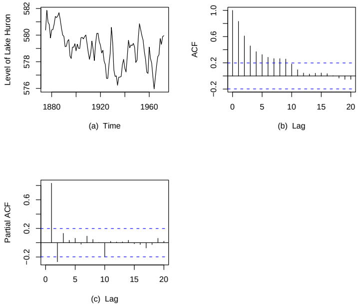


Figure: (a) Time series plot of the Lake Huron series. (b) acf plot of the Lake Huron time series. (c) Partial acf plot of the Lake Huron time series.

- **Parameters estimation**
  - **Prior information:** $\tilde{\sigma}_{a_i} = 5$ and $\alpha = 0.5$
  - **Sample collection:**
    - The length of the Markov chain: 550,000 steps
    - The burn-in period: the first 150,000 steps
    - samples saved every 10 steps
  - **The fitted model:**

$$y_t = 1.238 + 1.187 y_{t-1} - 0.537 y_{t-2} + 0.345 y_{t-3} - \frac{\log(1-\tau)}{0.767}. \tag{8}$$

- **Histograms**

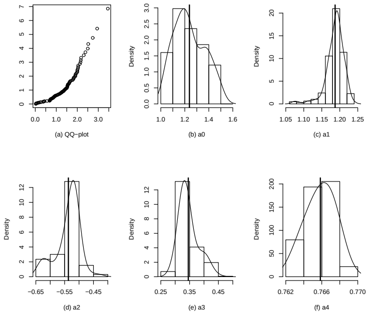


Figure: (a) The QQ-plot of the residuals of model (8). (b) $\sim$ (f): Histograms, density plots (continuous curves) and the estimated parameter values (darker vertical lines) of the collected samples.

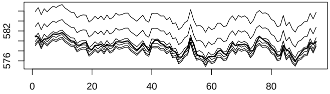
- **Semi-parametric model**

$$y_t = a_0^\tau + a_1^\tau y_{t-1} + a_2^\tau y_{t-2} + a_3^\tau y_{t-3} \qquad (9)$$

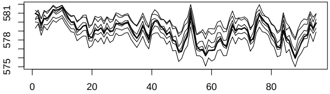
for $\tau = 0.05, 0.25, 0.5, 0.75, 0.95$ and 0.995.

- **Comparisons:** We compare the fitted τ^{th}-quantiles, where $\tau = 0.05, 0.25, 0.5, 0.75, 0.95$ and 0.995, from model (9) and model (8). See following figure.

- **Comparisons**



(a)



(b)

Figure: The fitted conditional quantiles for (from the bottom to the top) $\tau =$ 0.05, 0.25, 0.5, 0.75, 0.95 and 0.995 from (a) model (8) and (b) model (9), where the darker curves are the Lake Huron time series.

- **Comparisons**

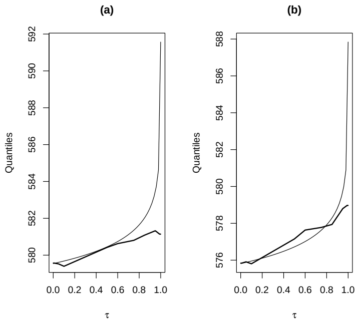


Figure: The estimated conditional quantile functions of (a) y_7 and (b) y_{33}, where the lighter curves were obtained from model (8), while the darker curves were obtained from model (9).

- **Germany-DAX Index**
  - The DAX Index is the most commonly cited benchmark for measuring the returns posted by stocks on the Frankfurt Stock Exchange.
  - The index is comprised of the 30 largest and most liquid issues traded on the exchange.
  - It is a performance-based index, which means that any dividends and other events are rolled into the index's final calculation.
  - The time series considered here is of length 400, during 1995-1996.

- **Germany-DAX Index**

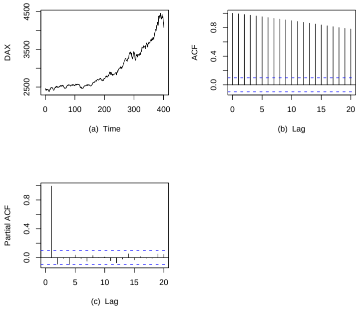


Figure: (a) Time series plot of the Germany DAX series. (b) acf plot of the Germany DAX series. (c) Partial acf plot of the Germany DAX series.

- **Fitted model**

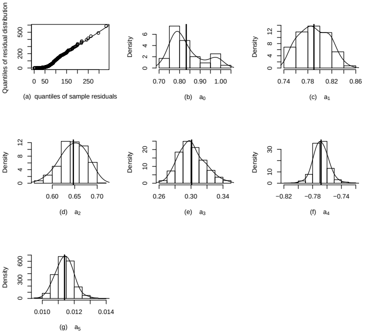


Figure: (a) The QQ-plot of the residuals of model (10). (b) $\sim$ (g): Histograms, density plots (continuous curves) and the estimated parameter values (darker vertical lines) of the collected samples.

- **Fitted model**

$$
\begin{aligned}
&Q_{y_t}(\tau \mid \mathbf{y}_{t-1}) \\
&= 0.756 + 0.944y_{t-1} + 0.343y_{t-2} + 0.250y_{t-3} \qquad (10) \\
&-0.571y_{t-4} - \frac{\log(1-\tau)}{0.011}.
\end{aligned}
$$

- **Forecasting**
  - Conditional on the history up to time 400: $y_1, \ldots, y_{400}$
  - Estimate the predictive density function of y_{401}
  - Estimate the predictive median of y_{401}
  - Estimate the predictive mean of y_{401}
  - Compare the predicted values with the actual observed value.

- **Forecasting**

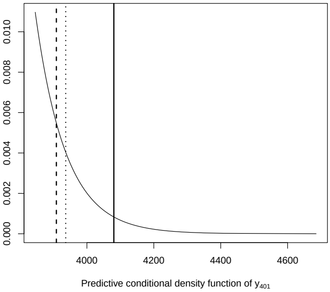


Figure: Predictive conditional density function of y_{401}. The continuous vertical line gives the position of the true observed value, the dashed and the dotted lines correspond to the predicted median and the predicted mean from the fitted model (10) respectively.

Conclusions and Future Work

- The methodology developed here can be applied to many other quantile function time series models
- The methodology developed here can be generalized
  - to deal with nonlinear time series
  - to handle the case where the scale parameter of the model also depends on the values of the time series
- Thorough comparisons between the semi-parametric and parametric approaches need to be done in the future

Conclusions and Future Work

- Comparisons between the quantile function approach and distribution function approach also need to be carried out in the future
- New methodology needs to be developed for forecasting from such quantile function models.
- The methodology also need to be developed to deal with multivariate time series.
- Research results will be reported elsewhere.

Acknowledgement

I would like to thank for the invitation and the financial support from Isaac Newton Institute for Mathematical Sciences which made it possible for me to present the paper here at the workshop.

Thank you!