

RESEARCH

Open Access



# High quantile regression for extreme events

Mei Ling Huang\* and Christine Nguyen

\*Correspondence:  
mhuang@brocku.ca  
Department of Mathematics &  
Statistics, Brock University, St.  
Catharines, Ontario, Canada

## Abstract

For extreme events, estimation of high conditional quantiles for heavy tailed distributions is an important problem. Quantile regression is a useful method in this field with many applications. Quantile regression uses an  $L_1$ -loss function, and an optimal solution by means of linear programming. In this paper, we propose a weighted quantile regression method. Monte Carlo simulations are performed to compare the proposed method with existing methods for estimating high conditional quantiles. We also investigate two real-world examples by using the proposed weighted method. The Monte Carlo simulation and two real-world examples show the proposed method is an improvement of the existing method.

**AMS 2010 Subject Classifications:** primary: 62G32; secondary: 62J05

**Keywords:** Bivariate Pareto distribution, Conditional quantile, Extreme value distribution, Generalized Pareto distribution, Linear programming, Weighted loss function

## 1 Introduction

Extreme value events are highly unusual events that can cause severe harm to people and costly damage to the environment. Examples of such harmful events are stock market crashes, equity risks, pipeline failures, large flooding, wildfires, pollution and hurricanes. The response variable,  $y$ , of an extreme event is usually distributed according to a heavy-tailed distribution. It is important to estimate high conditional quantiles of a random variable  $y$  given a variable vector  $\mathbf{x} = (1, x_1, x_2, \dots, x_k)^T \in R^p$  and  $p = k + 1$ .

The traditional mean linear regression is concerned with the estimation of the conditional expectation  $E(y|\mathbf{x})$  (Yu et al. 2003). The mean linear regression model assumes

$$\mu_{y|\mathbf{x}} = E(y|x_1, x_2, \dots, x_k) = \mathbf{x}^T \boldsymbol{\beta} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k.$$

We estimate  $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_k)^T \in R^p$  from a random sample  $\{(y_i, \mathbf{x}_i), i = 1, \dots, n\}$ , where  $\mathbf{x}_i$  is the  $p$ -dimensional design vector and  $y_i$  is the univariate response variable from a continuous distribution. The least squares (LS) estimator  $\hat{\boldsymbol{\beta}}_{LS}$  is a solution to the following equation

$$\hat{\boldsymbol{\beta}}_{LS} = \operatorname{argmin}_{\boldsymbol{\beta} \in R^p} \sum_{i=1}^n (y_i - \mathbf{x}_i^T \boldsymbol{\beta})^2, \quad (1)$$

that is,  $\hat{\boldsymbol{\beta}}_{LS}$  is obtained by minimizing the  $L_2$ -distance.

The mean linear regression provides the mean relationship between a response variable and explanatory variables (Yu et al. 2003). However, there are limitations present in the conditional mean models. Outliers significantly affect the conditional mean models and as a result, it affects the measurement of the central location, which may be misleading. Also, when analyzing extreme value events, where the response variable  $y$  has a heavy-tailed distribution, the mean linear regression cannot be extended to non-central locations (Hao and Naiman 2007). Therefore, it cannot provide insightful information for extreme events. Quantile regression offers a more complete statistical model by specifying the changes in the high conditional quantiles and it will be used to estimate values of extreme events (Yu et al. 2003; Hao and Naiman 2007). We will study two real world examples in the following sections.

### 1.1 Snowfall in Buffalo (1994-2015)

Large snowstorms can be very hazardous to people's safety, communities and their properties. They can significantly reduce visibility in an area, which makes it very dangerous for densely populated areas where major car accidents can happen on the road or accidents while flying can occur. A significant amount of snow, such as 12 inches (30 cm) or more, can cave in roofs of homes and buildings, standing trees can fall down on homes and cause the loss of electricity. There have been cases of deaths due to hypothermia, infections brought on by frostbite, car accidents caused due to slippery roads, heart attacks by overexertion while shoveling heavy snow and carbon monoxide poisoning from a power outage.

In 2006, Lake Storm "Aphid" was a lake-effect snowstorm that hit Buffalo, New York with a maximum snowfall of 24 inches (61 cm) and caused 19 fatalities. The snowstorm cost an estimated \$530 million in damages. Recently, in November 2014, severe lake-effect snowstorm heavily impacted areas in and around Buffalo with snowfall ranging from 5-7 feet (1.5-2.1 m) and the maximum snowfall of the storms was 88 inches (220 cm). There were records of 14 heart attacks due to overexertion and roofs collapsing due to the sheer weight of the snow. The data set was obtained from National Weather Service Forecast Office (2017) (Full data is available at <http://www.weather.gov/buf>) and the daily snowfall was recorded in inches for 4478 days from January 1994 – January 2015. The snowfall was converted into centimeters and a threshold of 5 cm was considered since snowfall under 5 cm is very unlikely to cause extreme damage. As a result, there are  $n = 316$  recorded data remaining. During January 1994 - January 2015, the top 10 largest daily snowfalls and maximum temperature in Buffalo, New York are shown in Table 1.

In Fig. 1(a), the y-axis represents the total daily snowfall (cm) and the x-axis represents the snowfall in the order of occurrence. The maximum daily snowfall occurred on December 10, 1995 with over 86.1 cm while the average daily snowfall was 11.96 cm. Figure 1(b) shows the scatter diagram of daily snowfalls greater or equal to 5 cm. It appears that it can be modeled with a quadratic polynomial relating snowfall to maximum temperature. The least squares method was performed on a polynomial mean regression model

$$E(y|x, x^2) = \beta_0 + \beta_1x + \beta_2x^2, \quad (2)$$

where  $y$  represents the total snowfall (cm) and  $x$  represents the maximum temperature ( $^{\circ}C$ ). The red curve represents the least squares (LS) curve  $\mu_{LS} = \hat{\mu}_{y|x} = 8.9879 - 0.2144x + 0.0040x^2$  which was obtained by using (1) and the model (2) to estimate the

**Table 1** The top 10 largest daily snowfalls with the maximum temperature in Buffalo, New York, from January 1994 to January 2015

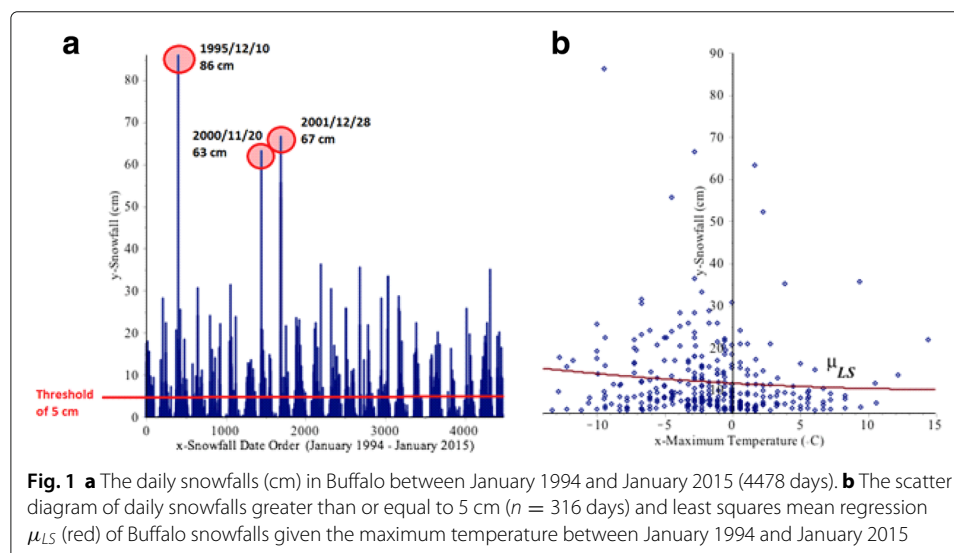
Date	Snowfall (cm)	Maximum temperature (°C)
December 10, 1995	86.11	-9.44
December 28, 2001	66.55	-2.78
November 20, 2000	63.25	1.67
December 27, 2001	55.63	-4.44
December 24, 2001	52.07	2.22
March 16, 2004	36.30	-2.78
October 13, 2006	35.61	9.39
March 12, 2014	35.10	3.89
March 8, 2008	33.30	-2.22
January 4, 1999	31.50	-6.67

mean of daily snowfall  $y$  for a given maximum temperature  $x$ . But, the least squares curve does not provide information about extreme heavy snowfalls that may cause damage. The quantile regression method will be able to estimate the high conditional quantiles. We will discuss this example further in Section 5.

### 1.2 CO<sub>2</sub> Emission

Climate change is considered to be one of the most important environmental issues as it is transforming life on Earth. It affects all aspects of our natural environment including the air and water quality, health and conservation of species at risk. It has been observed that temperatures and sea levels are rising, there are stronger storms and increased damage, and increased risk of drought, fire and floods. Climate change will rapidly alter the lands and waters that we depend upon for survival and we will no longer be able to preserve our environment for our social and economic well-being.

Natural processes and human activities can cause climate change. The recent global warming can be largely attributed to the carbon dioxide (CO<sub>2</sub>) and other greenhouse gas emissions. It was found that in 2009, CO<sub>2</sub> accounted for 82% of all European greenhouse gas emissions and about 94% of the CO<sub>2</sub> released to the atmosphere were from combusting fossil fuels (European Environment Agency (2017) at <http://www.eea.europa.eu>).



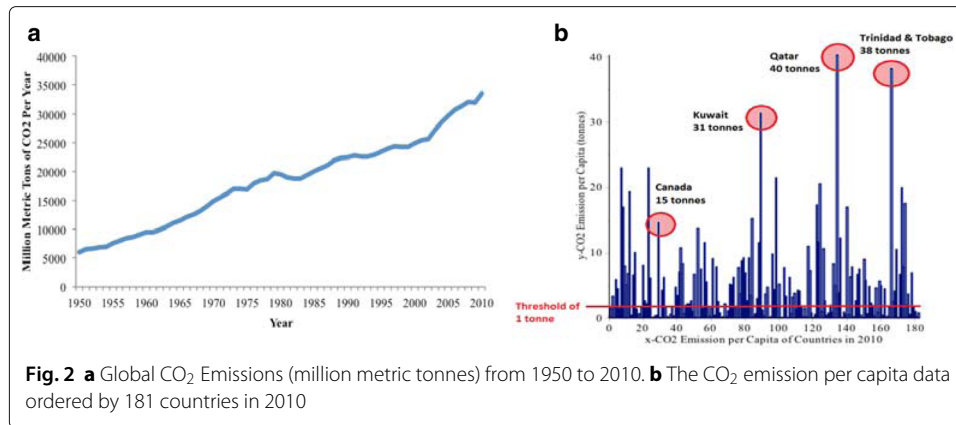


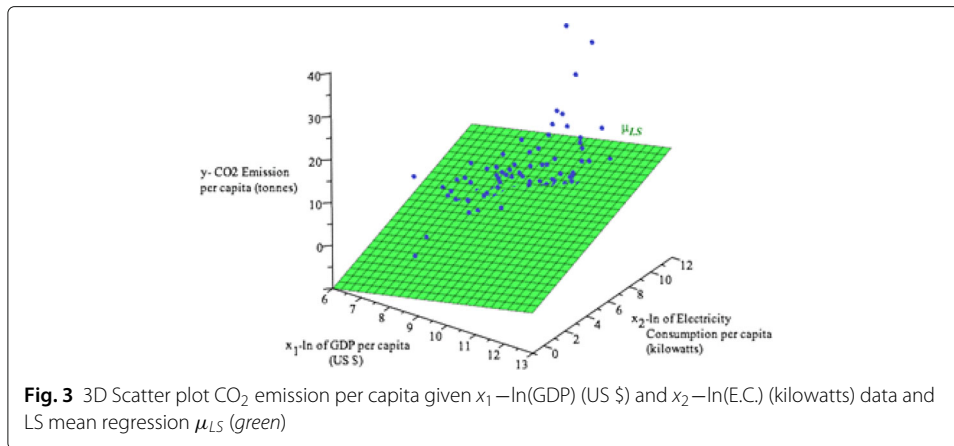
Figure 2(a) shows CO<sub>2</sub> emissions increases between 1950 and 2010; these increases are related with the increased energy use by an expanding economy, population and overall growth in emissions from electricity generation. It is important to estimate high conditional quantiles of the distribution of CO<sub>2</sub> emission in order to prevent acceleration of climate change.

In this paper, we use the 2010 data from the Carbon Dioxide Information Analysis Center (2017) at <http://cdiac.ornl.gov> for 181 countries. The CO<sub>2</sub> emission per capita was recorded in metric tonnes. There are  $n = 123$  countries remaining after the threshold of 1 tonne was applied. The threshold of 1 tonne of CO<sub>2</sub> emission was considered since values higher than 1 tonne of CO<sub>2</sub> emission would exceed the maximum allowance to emit without harming the climate. Table 2 lists the top 10 countries with the largest CO<sub>2</sub> emissions and their gross domestic product (GDP) and electricity consumption (E.C.) per capita. In Fig. 2(b), the y-axis represents the CO<sub>2</sub> emission per capita (tonnes) and x-axis represents the CO<sub>2</sub> emission ordered by country. It can be observed in Fig. 2(b) that Qatar produced the highest CO<sub>2</sub> emission per capita of 40 tonnes and Trinidad and Tobago and Kuwait produced the second and third highest CO<sub>2</sub> emissions of 38 tonnes and 31 tonnes respectively. As well, several countries emitted less than 10 tonnes of CO<sub>2</sub> in 2010.

We set the CO<sub>2</sub> emission per capita ( $y$ ) (tonnes),  $\ln(\text{GDP})$  per capita ( $x_1$ ) (US \$) and  $\ln(\text{E.C.})$  per capita ( $x_2$ ) (kilowatts) by using log-transformation. The 3D scatter diagram

**Table 2** The top 10 countries with the largest CO<sub>2</sub> emissions per capita with their GDP and Electricity Consumption (E.C.) in 2010

Country	CO <sub>2</sub> Emission per Capita (tonnes)	GDP per Capita (US \$)	Electricity Consumption (E.C.) per Capita (kilowatts)
Qatar	40.31	71,510.16	86.01
Trinidad and Tobago	38.16	15,630.05	1657.02
Kuwait	31.32	38,584.48	913.04
Brunei	22.87	30,880.34	239.40
Aruba	22.85	24,289.14	3,262.30
Luxembourg	21.36	102,856.97	2751.26
Oman	20.41	20,922.66	1562.59
United Arab Emirates	19.85	33,885.93	9007.35
Bahrain	19.34	20,545.97	10,142.73
United States	17.56	48,377.39	7588.42

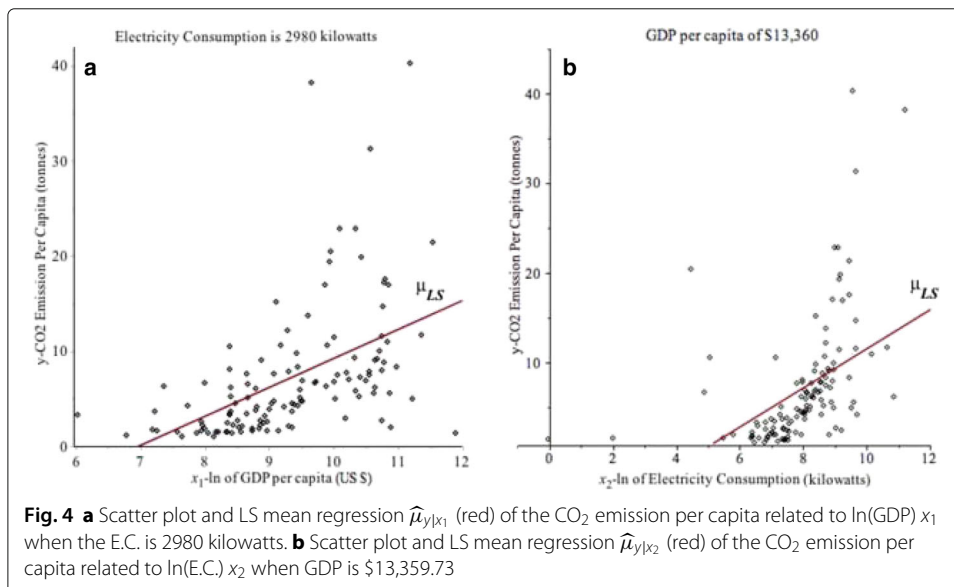


in Fig. 3 appears that it can be modeled using mean linear regression model (3) for the CO<sub>2</sub> emission per capita related with the ln(GDP) and ln(E.C.). For simplification, we do not put “per capita” for these three variables in the following text.

$$E(y|x_1, x_2) = \beta_0 + \beta_1x_1 + \beta_2x_2, \tag{3}$$

Figure 4(a) and (b) are 2D scatter plots. Figure 4(a) shows the relationship between ln(GDP) and CO<sub>2</sub> emission per capita  $\hat{\mu}_{y|x_1} = -21.0984 + 3.0255x_1$  when the E.C. per capita is 2980.96 kilowatts. Figure 4(b) demonstrates the relationship between ln(E.C.) and CO<sub>2</sub> emission per capita  $\hat{\mu}_{y|x_2} = -10.2255 + 2.1830x_2$  when the GDP per capita is \$13,359.73. The least squares mean regression curves in Fig. 4 were obtained by using (1) and the model (3).

Since the mean regression provides only the mean relationship between CO<sub>2</sub> emission per capita and GDP or E.C., it cannot provide estimation for high conditional quantiles of CO<sub>2</sub> emission. But the quantile regression method can estimate high CO<sub>2</sub> emission quantile curves. We will discuss this example further in Section 5.



### 1.3 Main methods and results

Quantile regression is an important model with applications in many fields. At first, quantile regression provides the estimates of the conditional quantiles, which are difficult to capture by a mean regression. Second, it is also more robust against outliers in the response measurements. The objective of this paper is to study and explore a new weighted quantile regression in order to improve the existing methods. In this paper, we will do three studies:

1. The theoretical approach will be investigated.
2. Monte Carlo simulations will be performed to show the efficiency of the new weighted method relative to the existing methods.
3. The new proposed method will be applied to real-world examples on extreme events and compared to mean regression and classical quantile regression.

In Section 2, we review some notation. In Section 3, we propose an optimal weighted quantile regression method, and give its good asymptotic properties for any uniformly bounded positive weight independent of response variable  $y$ , with conditional density as the weight. In Section 4, the results of Monte Carlo simulations generated from the bivariate Pareto distribution show that the proposed weighted method produces high efficiencies relative to existing methods. In Section 5, the three regression methods: mean regression, classic quantile regression and proposed weighted quantile regression, are applied to the real-life examples: the Buffalo snowfall (Subsection 1.1) and CO<sub>2</sub> emission (Subsection 1.2). Three goodness-of-fit tests are used to assess the distributions of the data. Studies of the examples illustrate that the proposed weighted quantile regression model fits better to the datasets than the existing quantile regression method.

## 2 Notation

Pickands (1975) first introduced the Generalized Pareto Distribution (GPD). (Also see de Haan and Ferreira 2006).

**Definition 2.1** *The cumulative distribution function (c.d.f.)  $F(x)$  and its corresponding probability density function (p.d.f.)  $f(x)$  of the two-parameter GPD( $\gamma, \sigma$ ) with shape parameter  $\gamma > 0$  and scale parameter  $\sigma$  of a random variable  $X$  are given by*

$$F(x) = 1 - \left(1 + \gamma \frac{x}{\sigma}\right)^{1/\gamma}, \quad \gamma, \sigma > 0, \quad x > 0; \tag{4}$$

$$f(x) = \sigma^{-1} \left(1 + \gamma \frac{x}{\sigma}\right)^{\frac{1}{\gamma}-1}, \quad \gamma, \sigma > 0, \quad x > 0.$$

**Definition 2.2** *The  $\tau$ th quantile of a continuous random variable  $Y$  with c.d.f.  $F(y)$  is defined as*

$$Q(\tau) = F^{-1}(\tau) = \inf\{y : F(y) \geq \tau\}, \quad 0 < \tau < 1.$$

**Definition 2.3** *The  $\tau$ th conditional linear quantile regression of  $y$  for given  $\mathbf{x} = (1, x_1, x_2, \dots, x_k)^T$  is defined as*

$$\begin{aligned} Q_y(\tau|\mathbf{x}) &= Q_\tau(y|x_1, x_2, \dots, x_k) = F^{-1}(\tau|\mathbf{x}) \\ &= \mathbf{x}^T \boldsymbol{\beta}(\tau) = \beta_0(\tau) + \beta_1(\tau)x_1 + \dots + \beta_k(\tau)x_k, \quad 0 < \tau < 1, \end{aligned}$$

where  $\boldsymbol{\beta}(\tau) = (\beta_0(\tau), \beta_1(\tau), \beta_2(\tau), \dots, \beta_k(\tau))^T$ .

Koenker and Bassett (1978) proposed a  $L_1$ -loss function to obtain estimator  $\widehat{\boldsymbol{\beta}}(\tau)$  by solving

$$\widehat{\boldsymbol{\beta}}(\tau) = \arg \min_{\boldsymbol{\beta}(\tau) \in R^p} \sum_{i=1}^n \rho_{\tau}(y_i - \mathbf{x}_i^T \boldsymbol{\beta}(\tau)), \quad 0 < \tau < 1, \tag{5}$$

where  $\rho_{\tau}$  is a loss function

$$\rho_{\tau}(u) = u(\tau - I(u < 0)) = \begin{cases} u(\tau - 1), & u < 0; \\ u\tau, & u \geq 0. \end{cases}$$

Quantile regression problem can be formulated as a linear program

$$\min_{(\boldsymbol{\beta}(\tau), \mathbf{u}, \mathbf{v}) \in R^p \times R_+^{2n}} \left\{ \tau \mathbf{1}_n^T \mathbf{u} + (1 - \tau) \mathbf{1}_n^T \mathbf{v} \mid \mathbf{X} \boldsymbol{\beta}(\tau) + \mathbf{u} - \mathbf{v} = \mathbf{y} \right\},$$

where  $\mathbf{1}_n^T$  is an  $n$ -vector of 1s,  $\mathbf{X}$  denotes the  $n \times p$  design matrix, and  $\mathbf{u}, \mathbf{v}$  are  $n \times 1$  vectors with elements of  $u_i, v_i$  respectively (Koenker 2005).

### 3 Proposed weighted quantile regression

#### 3.1 Proposed weighted quantile regression

Huang et al. (2015) proposed a weighted quantile regression method

$$\widehat{\boldsymbol{\beta}}_w(\tau) = \arg \min_{\boldsymbol{\beta}(\tau) \in R^p} \sum_{i=1}^n w_i(\mathbf{x}_i, \tau) \rho_{\tau}(y_i - \mathbf{x}_i^T \boldsymbol{\beta}(\tau)), \quad 0 < \tau < 1, \tag{6}$$

where  $w_i(\mathbf{x}_i, \tau)$  is any uniformly bounded positive weight function independent of  $y_i$ ,  $i = 1, \dots, n$ , for  $\mathbf{x}_i = (1, x_{i1}, x_{i2}, \dots, x_{ik})^T$ .

In this paper, we propose a weight as the local conditional density  $f(y|\mathbf{x})$  of  $y$  for given  $\mathbf{x}$  at the  $\tau$ th quantile point  $\xi_i(\tau, \mathbf{x}_i)$ , which is

$$w_i(\mathbf{x}_i, \tau) = f_i(\xi_i(\tau, \mathbf{x}_i)), \quad i = 1, 2, \dots, n, \quad 0 < \tau < 1, \tag{7}$$

where  $f_i(\xi_i(\tau, \mathbf{x}_i))$  is uniformly bounded at the quantile points  $\xi_i(\tau, \mathbf{x}_i)$ .

The following are the four reasons for the proposed weight in (7):

- (1) Koenker (2005, Chapter 5, Subsection 5.3) suggested that when the conditional densities of the response are heterogeneous, it is natural to consider whether weighted quantile regression might lead to efficiency improvements.
- (2) The error function  $\rho_{\tau}(y_i - \mathbf{x}_i^T \boldsymbol{\beta}(\tau))$  in (6) is an absolute error measure between  $y_i$  and the  $\tau$ th conditional quantile  $\xi_i(\tau, \mathbf{x}_i)$  at the  $i$ th sample point  $(y_i, \mathbf{x}_i)$ ,  $i = 1, 2, \dots, n$ .  $f_i(\xi_i(\tau, \mathbf{x}_i))$  can be interpreted as providing the local relative likelihood that the response variable  $y$  takes values in a neighborhood of  $y = \xi_i(\tau, \mathbf{x}_i) : (\xi_i(\tau, \mathbf{x}_i) - \varepsilon, \xi_i(\tau, \mathbf{x}_i) + \varepsilon)$ , for small  $\varepsilon > 0$ . Giving the weight  $f_i(\xi_i(\tau, \mathbf{x}_i))$  on  $\rho_{\tau}(y_i - \mathbf{x}_i^T \boldsymbol{\beta}(\tau))$  will make the total error  $\sum_{i=1}^n w_i(\mathbf{x}_i, \tau) \rho_{\tau}(y_i - \mathbf{x}_i^T \boldsymbol{\beta}(\tau))$  more reasonable.
- (3) The weighted estimator  $\widehat{\boldsymbol{\beta}}_w(\tau)$  in (6) using weight (7) has good properties, which we will discuss in Subsection 3.2 below.
- (4) Is weight (7) an optimal weight? It is a difficult problem in the field as described in Chapter 5, Subsection 5.3 in Koenker (2005). Also,  $f_i(\xi_i(\tau, \mathbf{x}_i))$  is difficult to

estimate. In this paper, we explore these two difficulties, we estimate the  $f_i(\xi_i(\tau, \mathbf{x}_i))$ , then obtain positive results by using weight (7).

In Section 4, in simulations, we compare using weight (7) with the weight given in Huang et al. (2014) as

$$w_i(\mathbf{x}_i, \tau) = \frac{\|\mathbf{x}_i\|^{-1}}{\sum_{i=1}^n \|\mathbf{x}_i\|^{-1}}, \quad 0 < \tau < 1, \tag{8}$$

where  $w_i(\mathbf{x}_i, \tau) \in [0, 1]$  and  $\sum_{i=1}^n w_i(\mathbf{x}_i, \tau) = 1$ ,  $i = 1, \dots, n$ ,  $\|\mathbf{x}_i\| = \sqrt{x_{i1}^2 + x_{i2}^2 + \dots + x_{ik}^2}$ ,  $k$  is the number of regressors.

In this paper, we are looking for improvement of efficiency by using weights (7) in Section 4 simulations, and applications of the Buffalo snowfall and CO<sub>2</sub> emission examples in Section 5.

### 3.2 Properties of weighted quantile regression

The following regularity conditions are necessary in deriving the asymptotic distribution of  $\hat{\beta}_{n(w)}(\tau)$  in (6) with weight  $w_i(\mathbf{x}_i, \tau) = f_i(\xi_i(\tau, \mathbf{x}_i))$ ,  $i = 1, 2, \dots, n$ ,  $0 < \tau < 1$ , in (7). Let  $Y_1, Y_2, \dots$  be independent random variables with distribution function  $F_1, F_2, \dots$

**Condition 1 (C1).** *The  $F_i$ 's are absolutely continuous, with continuous densities  $f_i(\xi)$  uniformly bounded away from 0 and  $\infty$  at the quantile points  $\xi_i(\tau, \mathbf{x}_i)$ ,  $i = 1, 2, \dots$*

**Condition 2 (C2).** *There exist positive definite matrices  $\mathbf{D}_0(\tau)$  such that the following three subconditions are satisfied*

- (i)  $\lim_{n \rightarrow \infty} n^{-1} \sum f_i^2(\xi_i(\tau, \mathbf{x}_i)) \mathbf{x}_i \mathbf{x}_i^T = \mathbf{D}_0(\tau)$ , and
- (ii)  $\lim_{i=1, \dots, n} \|f_i(\xi_i(\tau, \mathbf{x}_i))\| / \sqrt{n} \rightarrow 0$ .

We have the main asymptotic results for  $\hat{\beta}_{n(w)}(\tau)$  in (6) using weight (7). In this case, we let  $\hat{\beta}_{n(w)}(\tau) = \hat{\beta}_{n(f)}(\tau)$  in the following theorem.

**Theorem 3.1** *Under Conditions C1 and C2, we have*

$$\sqrt{n} \left( \hat{\beta}_{n(f)}(\tau) - \beta(\tau) \right) \xrightarrow{D} N(0, \tau(1 - \tau) \mathbf{D}_0^{-1}(\tau)), \text{ as } n \rightarrow \infty.$$

The proof of Theorem 3.1 is similar as the proof has been provided in Huang et al. (2015).

### 3.3 Comparison of quantile regression models

In order to compare the regular and weighted quantile regression models in (5) and (6), we extend the idea of measure of goodness of fit by Koenker and Machado (1999), and suggest the use of a Relative  $R(\tau)$ , which is defined as

$$\text{Relative } R(\tau) = 1 - \frac{V_{\text{weighted}}(\tau)}{V_{\text{regular}}(\tau)}, \quad -1 \leq R(\tau) \leq 1, \tag{9}$$

where

$$V_{\text{regular}}(\tau) = \sum_{y_i \geq \mathbf{x}_i^T \hat{\beta}(\tau)} \frac{\tau}{n} \left| y_i - \mathbf{x}_i^T \hat{\beta}(\tau) \right| + \sum_{y_i < \mathbf{x}_i^T \hat{\beta}(\tau)} \frac{(1 - \tau)}{n} \left| y_i - \mathbf{x}_i^T \hat{\beta}(\tau) \right|,$$



where  $\widehat{\beta}(\tau)$  is obtained by (5).

$$V_{weighted}(\tau) = \sum_{y_i \geq \mathbf{x}_i^T \widehat{\beta}_w(\tau)} w_i \tau \left| y_i - \mathbf{x}_i^T \widehat{\beta}_w(\tau) \right| + \sum_{y_i < \mathbf{x}_i^T \widehat{\beta}_w(\tau)} w_i (1 - \tau) \left| y_i - \mathbf{x}_i^T \widehat{\beta}_w(\tau) \right|,$$

where  $w_i$  and  $\widehat{\beta}_w(\tau)$  are given by (6).

#### 4 Simulations

In this Section, Monte Carlo simulations are performed. We generate  $m$  random samples with size  $n$  each from the bivariate Pareto distribution (Mardia 1962) with c.d.f.

$$F(x, y) = 1 - x^{-\alpha} - y^{-\alpha} - (x + y - 1)^{-\alpha}, \quad x > 1, y > 1, \alpha > 0, \tag{10}$$

and the conditional quantile function of (10) is

$$\xi(\tau|x) = Q_y(\tau|x) = 1 - x \left( 1 - \frac{1}{(1 - \tau)^{-1/(\alpha+1)}} \right), \quad x > 1, \alpha > 0, \quad 0 < \tau < 1. \tag{11}$$

The conditional density of  $y$  for given  $x$  is

$$f(y|x) = \frac{(\alpha + 1)x^{(\alpha+1)}}{(x + y - 1)^{(\alpha+2)}}, \quad x > 1, y > 1, \alpha > 0,$$

and the  $\tau$ th conditional density of  $y$  for given  $x$  at the  $\tau$ th quantile is

$$f(\xi(\tau|x)) = \frac{(\alpha + 1)(1 - \tau)^{(\alpha+2)/(\alpha+1)}}{x}, \quad x > 1, \alpha > 0, \quad 0 < \tau < 1. \tag{12}$$

Assume that the true conditional quantile is  $Q_y(\tau|x) = \beta_0(\tau) + \beta_1(\tau)x$ . We use two quantile regression methods:

1. The regular quantile regression  $Q_R(\tau|x)$  estimation based on (5),

$$Q_R(\tau|x) = \widehat{\beta}_0(\tau) + \widehat{\beta}_1(\tau)x \tag{13}$$

2. The weighted quantile regression  $Q_W(\tau|x)$  estimation based on (6)

$$Q_W(\tau|x) = \widehat{\beta}_{w0}(\tau) + \widehat{\beta}_{w1}(\tau)x. \tag{14}$$

For each method, we generate size  $n = 300, m = 1,000$  samples.  $Q_{R,i}(\tau|x)$  or  $Q_{W,i}(\tau|x), i = 1, \dots, m$ , are estimated in the  $i$ th sample. Let  $\alpha = 3$  in (12), then the weights in (7) are

$$w_i(\mathbf{x}_i, \tau) = \frac{4(1 - \tau)^{(5/4)}}{x_i}, \quad x_i > 1, i = 1, 2, \dots, n. \tag{15}$$

The simulation mean squared errors (SMSE) of the estimators (13) and (14) are:

$$SMSE(Q_R(\tau)) = \frac{1}{m} \sum_{i=1}^m \int_1^N (Q_{R,i}(\tau|x) - Q_y(\tau|x))^2 dx; \tag{16}$$

$$SMSE(Q_W(\tau)) = \frac{1}{m} \sum_{i=1}^m \int_1^N (Q_{W,i}(\tau|x) - Q_y(\tau|x))^2 dx, \tag{17}$$

where the true  $\tau$ th conditional quantile  $Q_y(\tau|x)$  is defined in (11).  $N$  is a finite  $x$  value such that the c.d.f. in (10)  $F(N, N) \approx 1$ . We take  $N = 1000$  and the simulation efficiencies (SEFF) are given by

$$SEFF(Q_W(\tau)) = \frac{SMSE(Q_R(\tau))}{SMSE(Q_W(\tau))},$$

where  $SMSE(Q_R(\tau))$  and  $SMSE(Q_W(\tau))$  are defined in (16) and (17) respectively.

**Table 3** Simulation mean square errors (SMSE) and efficiencies (SEFF) of estimating  $Q_Y(\tau|x), m = 1000, n = 300, N = 1000$

$\tau$	0.95	0.96	0.97	0.98	0.99
$SMSE(Q_R(\tau))$	$2.0533 \times 10^8$	$2.0103 \times 10^8$	$2.7731 \times 10^8$	$5.1199 \times 10^8$	$1.7458 \times 10^9$
$SMSE(Q_{W(f)}(\tau))$	$1.4939 \times 10^8$	$1.4146 \times 10^8$	$2.2462 \times 10^8$	$4.5520 \times 10^8$	$1.7161 \times 10^9$
$SEFF(Q_{W(f)}(\tau))$	1.6770	1.4211	1.2346	1.1248	1.0174

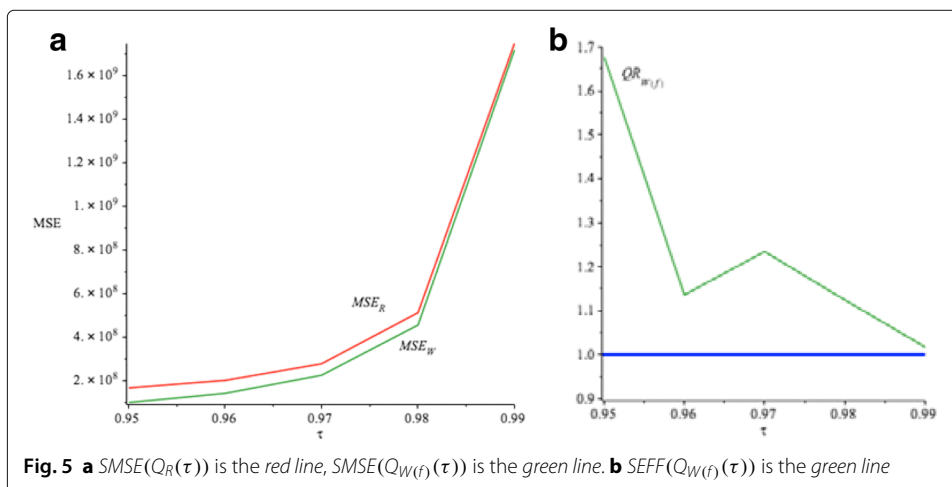
Table 3 displays the  $SEFF(Q_{W(f)}(\tau))$  for varying  $\tau$  values by using the weight in (15). It shows that the  $SEFF(Q_{W(f)}(\tau))$  are larger than 1 when  $\tau = 0.95, \dots, 0.99$ .

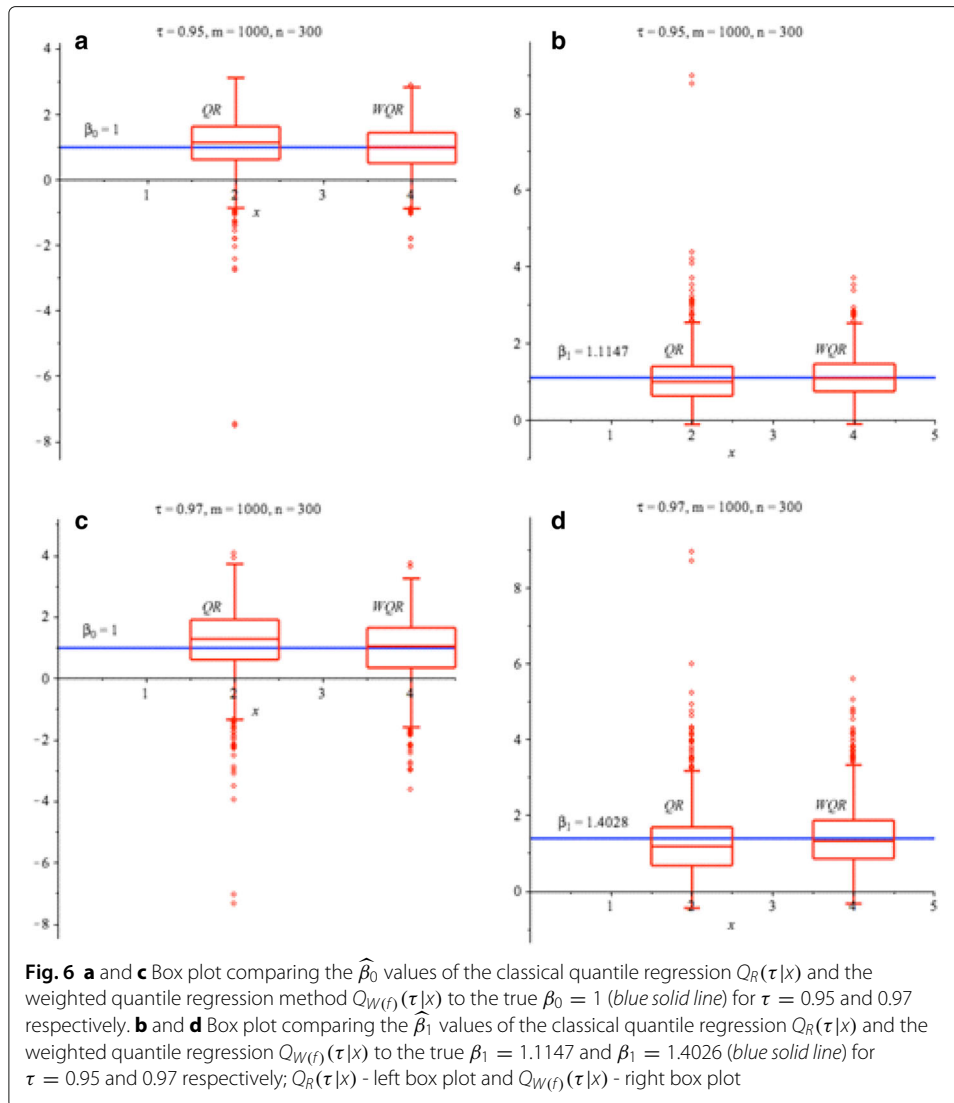
Figure 5 compares the  $SMSE(Q_R(\tau))$  with the  $SMSE(Q_{W(f)}(\tau))$  for  $\tau = 0.95, \dots, 0.99$ . It demonstrates that all  $SMSE(Q_{W(f)}(\tau))$  for our proposed weight in (15) have smaller values than  $SMSE(Q_R(\tau))$ . Furthermore, Fig. 6 shows the box plots for estimating the true  $\beta_0$  and  $\beta_1$  when  $\alpha = 3$  by using  $Q_R(\tau|x)$  and  $Q_{W(f)}(\tau|x)$  for  $\tau = 0.95$  and  $0.97$  respectively. It reveals that the proposed  $Q_{W(f)}(\tau|x)$  is unbiased and produces more accurate  $\hat{\beta}_{W(f)0}(\tau)$  and  $\hat{\beta}_{W(f)1}(\tau)$  estimators to the true  $\beta_0$  and  $\beta_1$  for  $\tau = 0.95$  and  $0.97$ . As well, the variances of  $Q_{W(f)}(\tau|x)$  are smaller relative to  $Q_R(\tau|x)$  for  $\tau = 0.95$  and  $0.97$ .

Next, we want to compare our simulation results with the following proposed weights presented in Huang et al. (2014) in (8). Table 4 compares the simulation  $SEFF(Q_{W(1)}(\tau))$  by using weight  $w_i(\mathbf{x}_i, \tau) = \|\mathbf{x}_i\|^{-1} / \sum_{i=1}^n \|\mathbf{x}_i\|^{-1}$  in (8) and  $SEFF(Q_{W(f)}(\tau))$  by using weight in (15) for different  $\tau$  values. Also, Fig. 7 compares the  $SEFF(Q_{W(1)}(\tau))$  with  $SEFF(Q_{W(f)}(\tau))$  with proposed weight in (15). It reveals that the  $SEFF(Q_{W(f)}(\tau))$  are larger than 1 and larger than  $SEFF(Q_{W(1)}(\tau))$ . Thus,  $Q_{W(f)}(\tau|x)$  is more efficient than  $Q_{W(1)}(\tau|x)$  when  $\tau = 0.95$  and up to  $0.99$ .

From the overall results of simulation, we can conclude that:

1. Table 3, Figs. 5 and 6 show that for  $\tau = 0.95, \dots, 0.99$ , the proposed weighted regression  $Q_{W(f)}(\tau|x)$  with the weight (15) is more efficient relative to the regular regression  $Q_R(\tau|x)$ .
2. Table 4 and Fig. 7 show that for  $\tau = 0.95, \dots, 0.99$ ,  $Q_{W(f)}(\tau|x)$  with the proposed weight (15) is more efficient relative to  $Q_{W(1)}(\tau|x)$  with  $w_i(\mathbf{x}_i, \tau) = \|\mathbf{x}_i\|^{-1} / \sum_{i=1}^n \|\mathbf{x}_i\|^{-1}$  in (8).





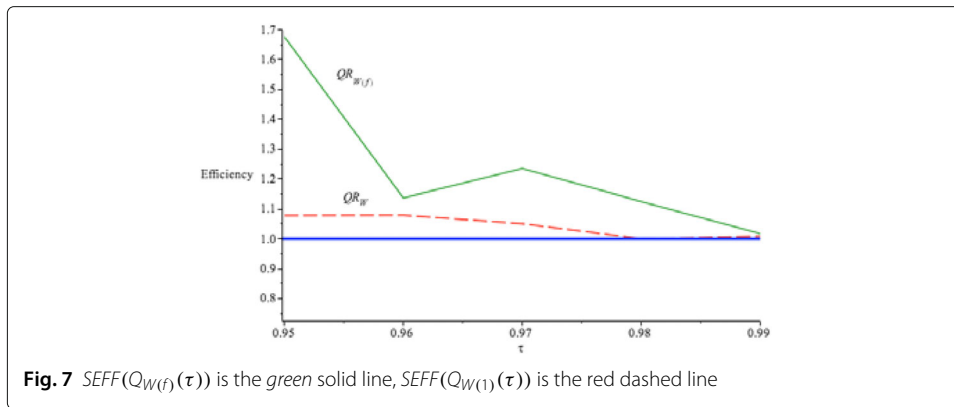
### 5 Real examples of applications

In this section, we applied the following three regression models to the Buffalo snowfall and CO<sub>2</sub> emission examples in Section 1:

1. The traditional mean linear regression (LS) estimator  $\hat{\beta}_{LS}$  in (1);
2. The regular quantile regression  $Q_R$  estimator  $\hat{\beta}(\tau)$  in (5);
3. The proposed weight quantile regression  $Q_W$  estimator  $\hat{\beta}_W(\tau)$  in (6) with weight  $w_i(\mathbf{x}_i, \tau) = f_i(\xi_i(\tau, \mathbf{x}_i))$  in (7).

**Table 4** Simulation efficiency (SEFF) of estimating  $Q_y(\tau|x)$ ,  $m = 1000$ ,  $n = 300$ ,  $N = 1000$

$\tau$	0.95	0.96	0.97	0.98	0.99
$SEFF(Q_{W(1)}(\tau))$	1.0783	1.0729	1.0504	1.0008	1.0086
$SEFF(Q_{W(f)}(\tau))$	1.6770	1.4211	1.2346	1.1248	1.0174



To estimate the proposed local conditional density weight  $w_i(\mathbf{x}_i, \tau) = f_i(\xi_i(\tau, \mathbf{x}_i))$  in (7), we use kernel density estimation (Silverman 1986; Scott 1992).

$$\widehat{w}_i(\mathbf{x}_i, \tau) = \widehat{f}_i(\widehat{\xi}_i(\tau, \mathbf{x}_i)), \text{ where } \widehat{f}(y|\mathbf{x}) = \frac{\widehat{f}(y, \mathbf{x})}{\widehat{\mu}(\mathbf{x})}, \tag{18}$$

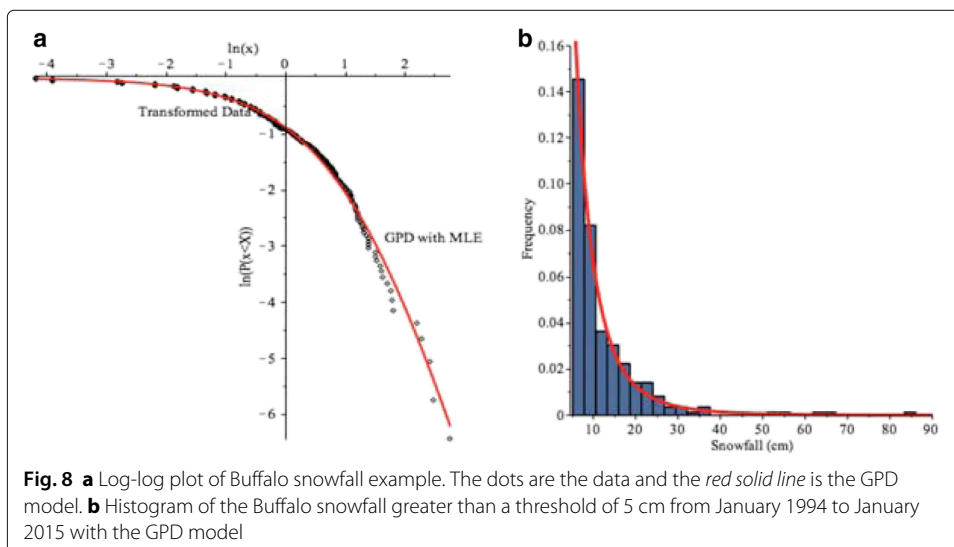
where  $\widehat{f}(y, \mathbf{x})$  is an estimator of the joint density of  $y$  and  $\mathbf{x}$ , and  $\widehat{\mu}(\mathbf{x})$  is an estimator of marginal density of  $\mathbf{x}$ . We estimate the conditional quantile function  $\xi(\tau|\mathbf{x})$  by inverting an estimated conditional c.d.f.  $\widehat{F}(y|\mathbf{x})$  (Li and Racine 2007)

$$\widehat{\xi}(\tau|\mathbf{x}) = \widehat{Q}_y(\tau|\mathbf{x}) = \inf\{y : \widehat{F}(y|\mathbf{x}) \geq \tau\} = \widehat{F}^{-1}(\tau|\mathbf{x}),$$

where  $\widehat{F}(y|\mathbf{x})$  is the estimated conditional c.d.f.  $F(y|\mathbf{x})$ .

Note that for a *one*-dimensional random sample  $X_1, X_2, \dots, X_n$  from the distribution  $\mu(x)$ , a kernel density estimator for  $\mu(x)$  is given by

$$\widehat{\mu}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right), \quad -\infty < x < \infty,$$



**Table 5** The goodness of fit tests for Buffalo snowfall example

$K - S$		$A - D$		$C - v - M$	
Test statistic	$p$ -value	Test statistic	$p$ -value	Test statistic	$p$ -value
0.0390	0.6055	0.6631	0.5905	0.0645	0.7856

where  $h$  is the window width, and  $K(x)$  is the kernel function, which is a symmetric probability density function with the conditions

$$\int_{-\infty}^{\infty} K(x)dx = 1, \int tK(t)dt = 0, \int t^2K(t)dt = k_2 \neq 0.$$

The optimal window width can be found by

$$h_{opt} = k_2 \left\{ \int K(t)^2 \right\}^{1/5} \left\{ \int \mu''(x)^2 dx \right\}^{-1/5} n^{-1/5}.$$

The  $d$ -dimensional multivariate kernel density estimator is defined by

$$\hat{\mu}(\mathbf{x}) = \frac{1}{nh^d} \sum_{i=1}^n K \left\{ \frac{\mathbf{x} - \mathbf{X}_i}{h} \right\},$$

where  $h$  is the window width and the kernel function  $K(\mathbf{x})$  is a function, defined for  $d$ -dimensional  $\mathbf{x}$ , satisfying  $\int_{R^d} K(\mathbf{x})d\mathbf{x} = 1$ . Fukunaga (1972) suggested

$$\hat{f}(\mathbf{x}) = \frac{(\det \mathbf{S})^{-1/2}}{nh^d} \sum_{i=1}^n k \left\{ \frac{(\mathbf{x} - \mathbf{X}_i)^T \mathbf{S}^{-1} (\mathbf{x} - \mathbf{X}_i)}{h^2} \right\},$$

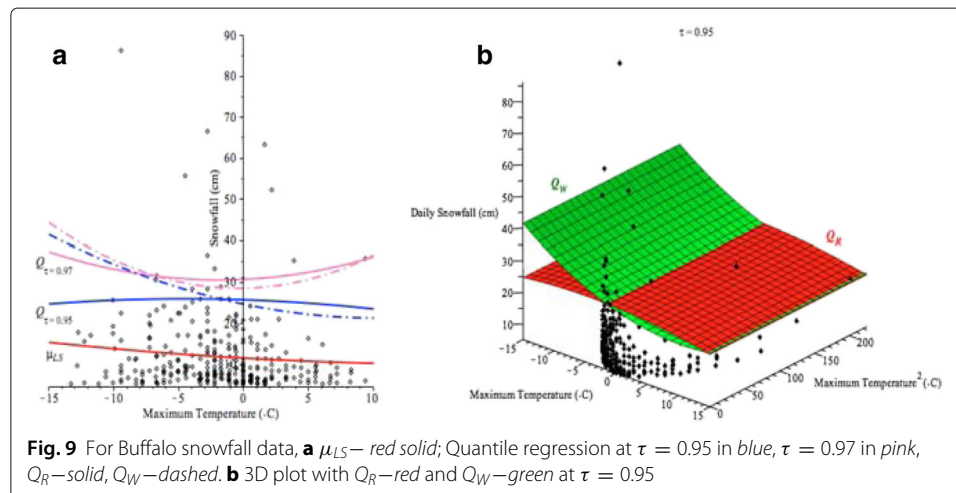
where  $\mathbf{S}$  is the sample covariance matrix of the data,  $K$  is the normal kernel and the function  $k$  is given by

$$k(u) = \left( \frac{1}{2\pi} \right)^{d/2} \exp \left( -\frac{u}{2} \right), \quad k(\mathbf{x}^T \mathbf{x}) = K(\mathbf{x}) = (2\pi)^{-d/2} \exp \left( -\frac{1}{2} \mathbf{x}^T \mathbf{x} \right).$$

An estimator for the optimal window width  $h$  will be given by

$$\hat{h}_{opt} = A(K)n^{-1/(d+4)},$$

where  $A(K) = \{4/(d + 2)\}^{1/(d+4)}$  is the constant for a multivariate normal kernel.



**Table 6** Buffalo daily snowfalls (cm) at high quantile using  $Q_R$  and  $Q_W$

Temperature (°C)	$\tau = 0.95$		$\tau = 0.97$	
	$Q_R$	$Q_W$	$Q_R$	$Q_W$
-15	24.72	41.57	37.38	44.38
-10	25.65	34.37	33.19	35.42
-5	26.00	28.74	30.98	30.15
0	25.76	24.69	30.73	28.58
5	24.93	22.21	32.47	30.70
10	23.52	21.30	36.17	36.52

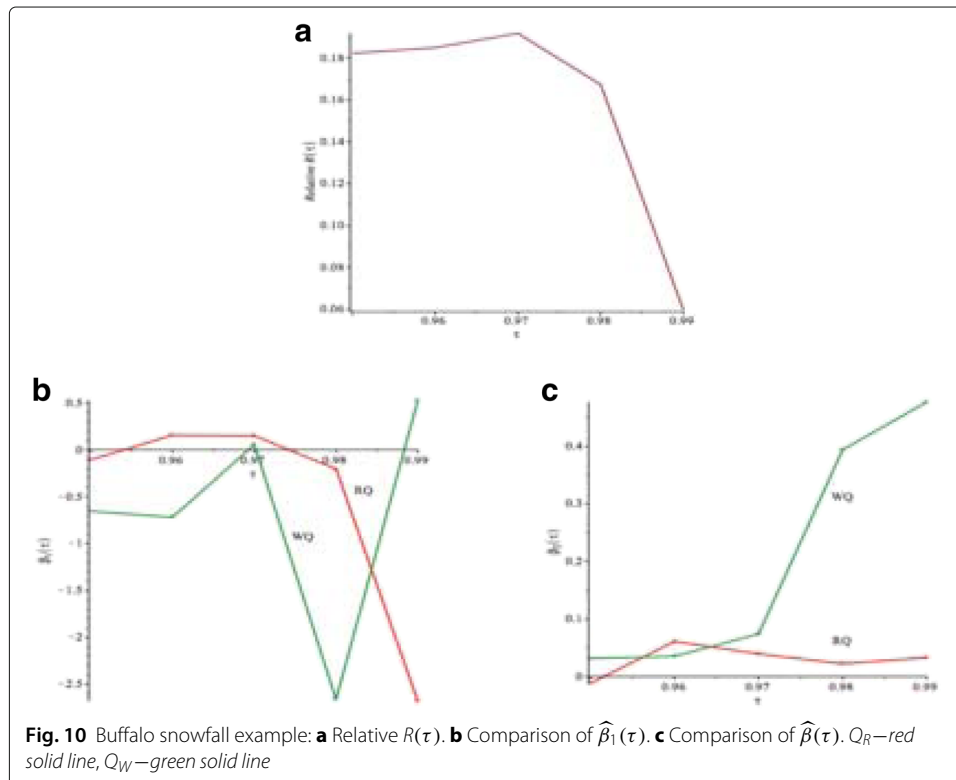
**5.1 Buffalo snowfall example**

Now, recall the Buffalo snowfall example in Subsection 1.1. We use a polynomial mean regression model (2)

$$E(y|x, x^2) = \beta_0 + \beta_1x + \beta_2x^2,$$

where  $y$  is the daily snowfall (cm) and  $x$  is the maximum temperature (°C). But the least squares curve only estimates average daily snowfall for a given maximum temperature; it cannot estimate extreme heavy snowfalls. The quantile regression method can estimate high conditional quantile curves and will be shown this Section.

In order to fit the Buffalo snowfall data to the GPD model (4), the data was transformed to  $y = \frac{x-\mu}{\sigma}$ , where  $\mu = 5$  cm, and then, we used the maximum likelihood estimates (MLEs) of the parameters,  $\hat{\sigma}_{MLE} = 5.1552, \hat{\gamma}_{MLE} = 0.2636$ , for the 2-parameter GPD model from the Buffalo snowfall data.



**Table 7** Relative  $R(\tau)$  values for the Buffalo snowfall example

	$\tau = 0.95$	$\tau = 0.96$	$\tau = 0.97$	$\tau = 0.98$	$\tau = 0.99$
Relative $R(\tau)$	0.1824	0.1851	0.1919	0.1673	0.0595

Furthermore, Fig. 8(a) and (b) shows the log-log plot and histogram of Buffalo snowfall with GPD model with the MLEs of the parameters. It illustrates that most daily snowfalls in Buffalo are between 0 and 10 centimeters and there are some occurrences of heavy snowfall, such as 50-90 centimeters. The GPD curve follows the shape of the Buffalo snowfall data very well.

Three goodness-of-fit tests are performed: the Kolmogorov-Smirnov test ( $K - S$ ) (Kolmogorov 1933), Anderson-Darling test ( $A - D$ ) and Cramé r-von-Mises test ( $C - \nu - M$ ) (Anderson and Darling 1952) respectively.

$$H_0 : F(y) = F^*(y), \text{ for all values of } y;$$

$$H_1 : F(y) \neq F^*(y), \text{ for at least one value of } y,$$

where  $F(y)$  is the true but unknown distribution function of the sample and  $F^*(y)$  is the theoretical distribution function, GPD in (4).

In Table 5, the  $K - S$ ,  $A - D$  and  $C - \nu - M$  tests show that the GPD model fits the data with a probability of 60.55%, 59.05% and 78.56% respectively.

Instead of using model (2), we use the following quantile regression model:

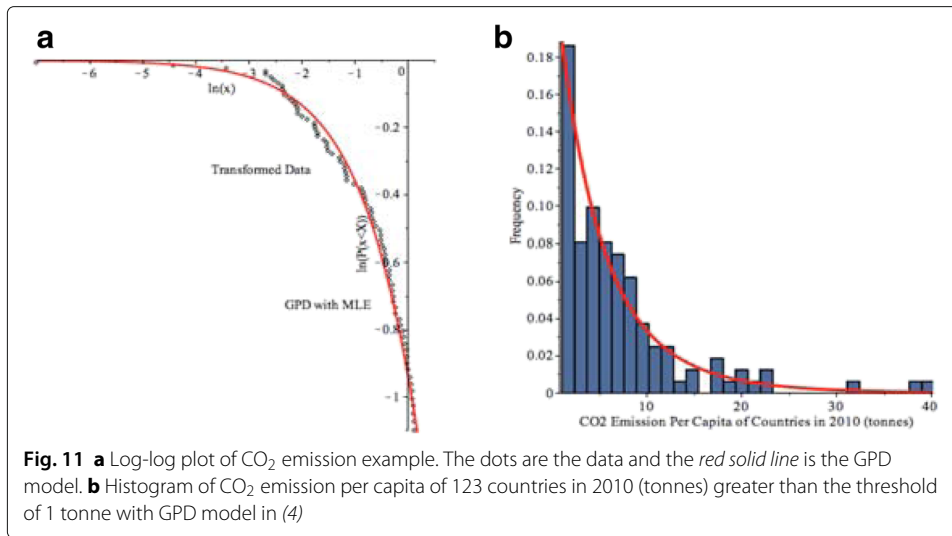
$$Q_y(\tau|x) = \beta_0(\tau) + \beta_1x(\tau) + \beta_2(\tau)x^2,$$

where we use the estimated weight  $\widehat{w}_i(\mathbf{x}_i, \tau) = \widehat{f}_i(\widehat{\xi}(\tau))$  in (18). Figure 9 shows the scatter plot of the daily snowfall with the fitted  $\mu_{LS}$ ,  $Q_R$  and  $Q_W$  curves at two high 0.95th and 0.97th quantiles. It is interesting to note that at the 0.95th and 0.97th quantiles, the  $Q_R$  and  $Q_W$  curves appear to fit the data. Table 6 lists the estimated Buffalo snowfall quantile values at a given maximum temperature for  $\tau = 0.95$  and 0.97. Both Fig. 9 and Table 6 demonstrate that when quantiles are high,  $Q_W$  have heavier snowfall than  $Q_R$ .

Figure 10(a) and Table 7 show the values of the Relative  $R(\tau)$  in (9) for given  $\tau = 0.95, \dots, 0.99$ . We note that  $R(\tau) > 0$ , which means that  $V_{weighted}(\tau) < V_{regular}(\tau)$ , and the  $Q_W$  is a better fit to the data than the  $Q_R$ . Figure 10(b), (c) and Table 8 show the values of  $\widehat{\beta}_1(\tau)$  and  $\widehat{\beta}_2(\tau)$ . The values of  $\widehat{\beta}_1(\tau)$  and  $\widehat{\beta}_2(\tau)$  are consistent with Fig. 9(a), (b) and Table 6.

**Table 8** Coefficients of the  $Q_R, Q_W$  and LS  $\mu_{LS}$  regression for buffalo snowfall example

$\tau$	Weight	$\widehat{\beta}_0(\tau)$	$\widehat{\beta}_1(\tau)$	$\widehat{\beta}_2(\tau)$
LS	—	11.5280	-0.1777	0.0053
0.95	$Q_R$	25.7589	-10.1068	-0.0117
	$Q_W$	24.6887	-0.6538	0.0315
0.96	$Q_R$	28.7869	0.1543	0.0610
	$Q_W$	24.8614	-0.7200	0.0352
0.97	$Q_R$	30.7341	0.1488	0.0395
	$Q_W$	28.5776	0.0551	0.0739
0.98	$Q_R$	35.5582	-0.2039	0.0223
	$Q_W$	25.8718	-2.6593	0.3937
0.99	$Q_R$	57.8614	-2.6793	0.0330
	$Q_R$	48.5464	0.5261	0.4768



The proposed weighted quantile regression  $Q_W$  predicts that for moderate temperatures, such as 5 °C to 10 °C, it is likely to have small snowfalls in Buffalo, and for every low temperatures, such as -15 °C to 0 °C, it is more likely to have heavy snowfalls that may cause damage. Predicting heavy snowfalls is related to cold weather forecasts. Quantile regression is useful for predicting extreme heavy snowfalls.

### 5.2 CO<sub>2</sub> emission example

In Subsection 1.2, there is a relationship between  $\ln(\text{GDP})$   $x_1$  and  $\ln(\text{E.C.})$   $x_2$  and CO<sub>2</sub> emissions per capita  $y$ . The least squares estimate is:

$$\mu_{LS} = -22.5009 + 2.0708x_1 + 1.2998x_2.$$

The quantile regression method can estimate high conditional quantile curves and will be shown in detail in this Section.

Similar to the Buffalo snowfall example, we fit the GPD model in (4) with MLEs of the parameters,  $\hat{\sigma}_{MLE} = 5.3011$ ,  $\hat{\gamma}_{MLE} = 0.1234$ , to the CO<sub>2</sub> emission data, which is demonstrated in Fig. 11(a), (b) by the log-log plot and histogram. The GPD model follows the shape of the CO<sub>2</sub> emission data very well. Table 9 shows the results of the three goodness-of-fit tests.

We use the proposed weight  $\hat{w}_i(x_i, \tau) = \hat{f}_i(\hat{\xi}(\tau))$  in (18) on the quantile regression model:

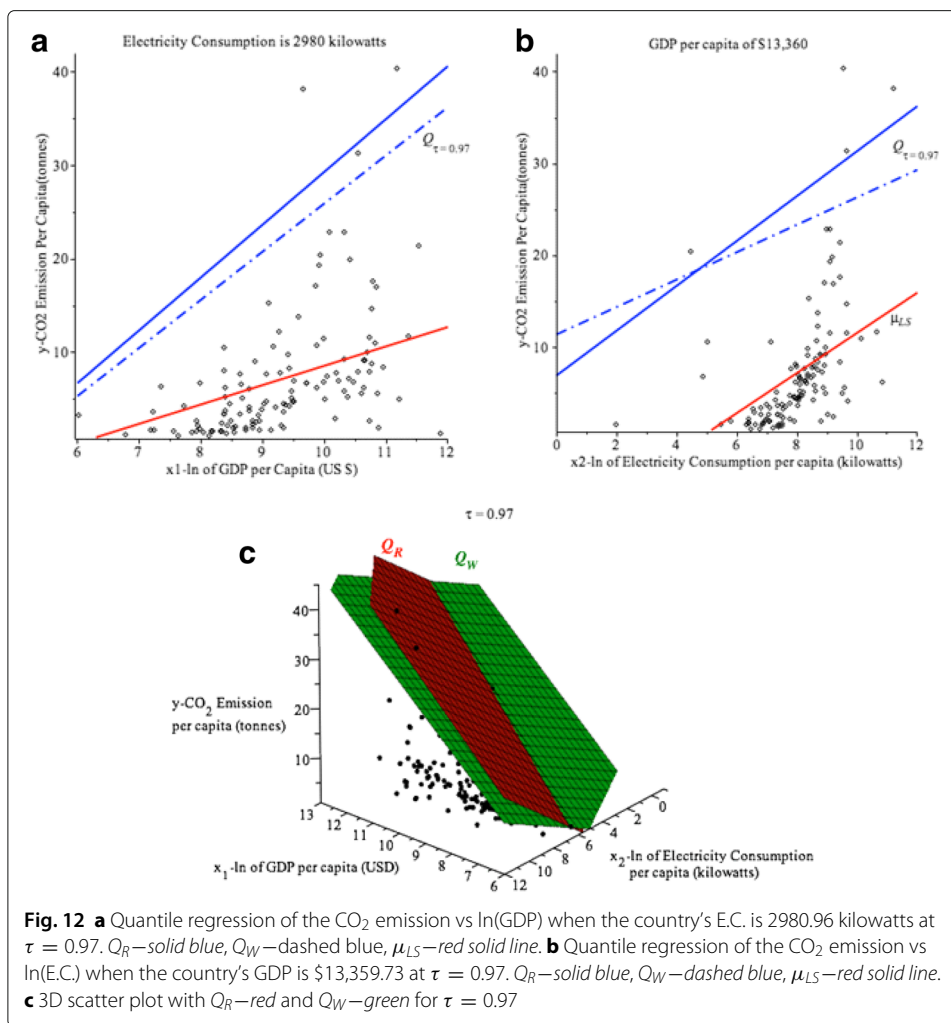
$$Q_y(\tau|x_1, x_2) = \beta_0(\tau) + \beta_1(\tau)x_1 + \beta_2(\tau)x_2.$$

Figure 12(a) shows the scatter plot of CO<sub>2</sub> emission vs  $\ln(\text{GDP})$  when the country's E.C. is 2980.96 kilowatts with the fitted  $\mu_{LS}$ ,  $Q_R$  and  $Q_W$  curves at the 0.97th quantile. Figure 12(b) shows the scatter plot of CO<sub>2</sub> emission vs  $\ln(\text{E.C.})$  when the country's GDP

**Table 9** The goodness of fit tests for CO<sub>2</sub> emission example

$K - S$		$A - D$		$C - v - M$	
Test statistic	$p$ -value	Test statistic	$p$ -value	Test statistic	$p$ -value
0.0443	0.8397	0.3619	0.8855	0.0517	0.8662





**Fig. 12** **a** Quantile regression of the CO<sub>2</sub> emission vs ln(GDP) when the country's E.C. is 2980.96 kilowatts at  $\tau = 0.97$ .  $Q_R$ —solid blue,  $Q_W$ —dashed blue,  $\mu_{LS}$ —red solid line. **b** Quantile regression of the CO<sub>2</sub> emission vs ln(E.C.) when the country's GDP is \$13,359.73 at  $\tau = 0.97$ .  $Q_R$ —solid blue,  $Q_W$ —dashed blue,  $\mu_{LS}$ —red solid line. **c** 3D scatter plot with  $Q_R$ —red and  $Q_W$ —green for  $\tau = 0.97$

is \$13,359.73 with the fitted  $\mu_{LS}$ ,  $Q_R$  and  $Q_W$  curves at the 0.97th quantile. Figure 12(c) shows the 3D scatter plot with  $Q_R$  (red) and  $Q_W$  (green) of CO<sub>2</sub> emission given the ln(GDP) and ln(E.C.) at  $\tau = 0.97$ . It is important to note that the  $\mu_{LS}$  is the red solid line and the  $Q_R$  and  $Q_W$  quantile regression lines appear to fit the data. In general, the  $Q_W$  line produces a different estimated CO<sub>2</sub> emissions than  $Q_R$  curve at high quantiles. Tables 10

**Table 10** CO<sub>2</sub> emission per capita high quantile given ln(GDP) estimators  $Q_R$  and  $Q_W$  at 2980.96 Kilowatts of electricity consumed per capita

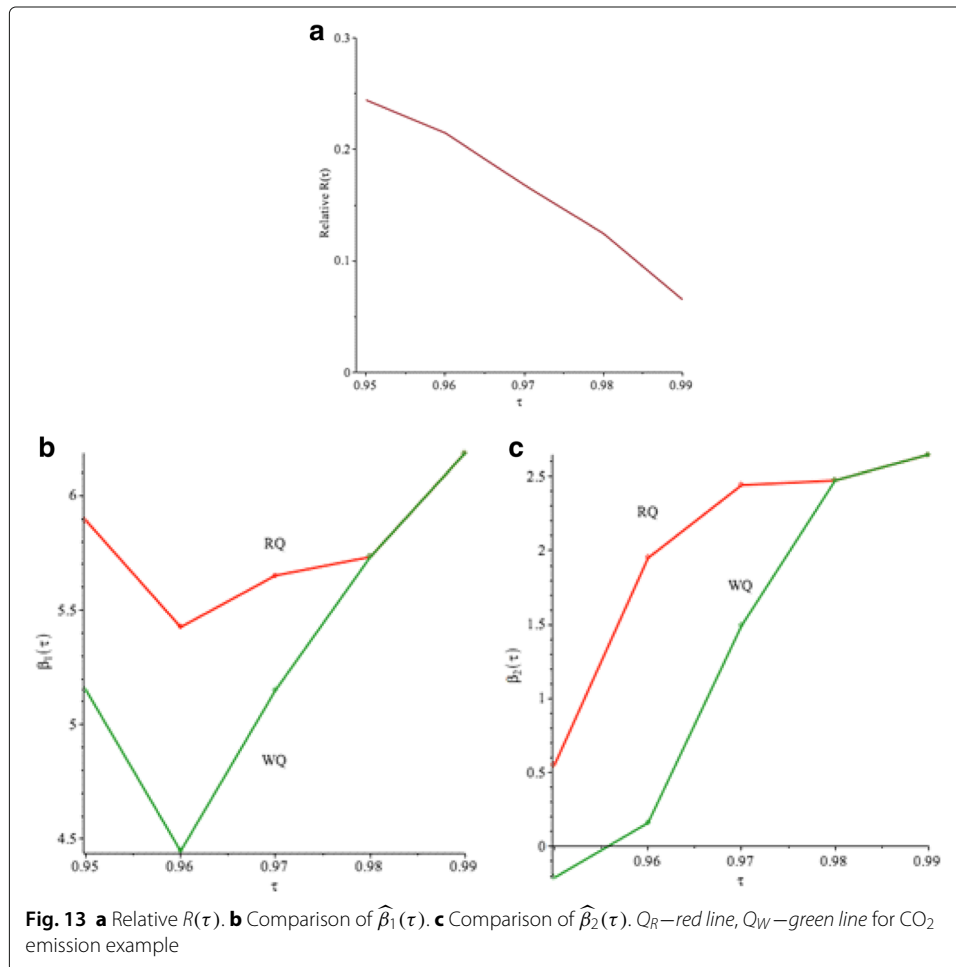
ln of GDP per capita (\$)	$\tau = 0.97$	
	$Q_R$	$Q_W$
7.5	15.2181	13.0840
8	18.0437	15.6591
8.5	20.8693	18.2342
9	23.6950	20.9093
9.5	26.5206	23.3844
10	29.3462	25.9595
10.5	32.1718	28.5346
11	34.9975	31.1097
11.5	37.8231	33.5848
12	40.6487	36.2599

**Table 11** CO<sub>2</sub> Emission per capita high quantile given ln(E.C.) estimators  $Q_R$  and  $Q_W$  at GDP per capita of \$13,359.73

ln of Electricity Consumption per capita (kilowatts)	$\tau = 0.97$	
	$Q_R$	$Q_W$
0	6.9775	11.4376
2	11.8632	14.4243
4	16.7490	17.4110
6	21.6348	20.3977
8	26.5206	23.3844
10	31.4064	26.3711
12	36.2921	29.3578

and 11 provide details about countries' CO<sub>2</sub> emission at high quantile ( $\tau = 0.97$ ) when the countries consume 2980.96 kilowatts of electricity and have a GDP of \$13,359.73 respectively.

Figure 13(a) shows the Relative  $R(\tau)$ , which is defined in (9) and Table 12 shows the values for Relative  $R(\tau)$  for  $\tau \geq 0.95$ . All values of Relative  $R(\tau)$  are larger than 0, which signifies that  $V_{weighted}(\tau) < V_{regular}(\tau)$  and as well, it suggests that the weighted quantile regression model  $Q_W$  is a better fit to the CO<sub>2</sub> emission data than the regular quantile regression model  $Q_R$ . Figure 13(b) and (c) shows the values of  $\hat{\beta}_1(\tau)$  and  $\hat{\beta}_2(\tau)$  for a given  $\tau$



**Table 12** Relative  $R(\tau)$  values for CO<sub>2</sub> emission example

	$\tau = 0.95$	$\tau = 0.96$	$\tau = 0.97$	$\tau = 0.98$	$\tau = 0.99$
Relative $R(\tau)$	0.24425	0.21478	0.16810	0.12459	0.06521

respectively, which is also shown in Table 13. The values of  $\hat{\beta}_1(\tau)$  and  $\hat{\beta}_2(\tau)$  are consistent with Tables 10 and 11.

It can be concluded that countries with higher gross domestic product and higher amounts of electricity produce higher CO<sub>2</sub> emissions. Since CO<sub>2</sub> is not destroyed over time, it can remain in the atmosphere for thousands of years due to the very slow process by which carbon is transferred to ocean sediments. As a result, countries should monitor their CO<sub>2</sub> emissions in order to prevent further damages to the environment. Countries can consider producing more energy from renewable sources, such as wind, solar, hydro and geothermal heat and using fuels with lower carbon content to reduce carbon emissions.

### 6 Conclusions

After the studies above, we can conclude:

1. Traditional mean regression are concerned with estimating the conditional mean by using the L<sub>2</sub>-loss function. Quantile regression with a L<sub>1</sub>- loss function overcomes the limitations of traditional mean regression. It gives estimates of  $\tau$ th conditional quantiles besides the measures of central tendency. Estimation of high conditional quantiles is very useful for the analysis of extreme events.
2. The proposed weighted quantile regression method with the local conditional density as the weight has good mathematical asymptotic properties.
3. The Monte Carlo computational simulation results show that the proposed weighted quantile regression with the local conditional density as the weight is more efficient relative to the classical quantile regression and some existing weighted quantile regression estimators.
4. The proposed weighted quantile regression can be used to predict extreme values of snowfall and CO<sub>2</sub> emission real world examples. In the Buffalo snowfall example, communities can use the information that quantile regression provides to prevent car accidents on roads, overexertion, and collapsing of homes. In the CO<sub>2</sub> emission

**Table 13** Coefficients of the  $Q_R, Q_W$  and LS  $\mu_{LS}$  regression for CO<sub>2</sub> emission example

$\tau$	Weight	$\hat{\beta}_0(\tau)$	$\hat{\beta}_1(\tau)$	$\hat{\beta}_2(\tau)$
LS	—	-22.5009	2.0708	1.2998
0.95	$Q_R$	-41.6856	5.8924	0.5527
	$Q_W$	-29.9131	5.1521	-0.2094
0.96	$Q_R$	-44.8147	5.4258	1.9505
	$Q_W$	-24.5504	4.4470	0.1609
0.97	$Q_R$	-46.7095	5.6513	2.4429
	$Q_W$	-37.4893	5.1502	1.4934
0.98	$Q_R$	-47.4004	5.7323	2.4739
	$Q_W$	-47.4004	5.7323	2.4739
0.99	$Q_R$	-51.2657	6.1856	2.6475
	$Q_W$	-51.2657	6.1856	2.6475

example, the countries' increase in gross domestic product and electricity consumption will likely cause an increase in the CO<sub>2</sub> emissions. CO<sub>2</sub> emission levels should be monitored to reduce the amount of carbon dioxide in the atmosphere and its long term effects.

5. It is difficult to estimate the proposed conditional density weight. The nonparametric kernel density estimation method is successful in this paper. Further studies on optimal weights are suggested.

#### Acknowledgements

We appreciate for the reviewers' and Editors' very helpful comments, which helped to improve the paper. This research is supported by the Natural Science and Engineering Research Council of Canada (NSERC) grant MLH, RGPIN-2014-04621.

#### Authors' contributions

The authors MLH and CN carried out this work and drafted the manuscript together. Both authors read and approved the final manuscript.

#### Competing interests

The authors declare that they have no competing interests.

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 24 December 2016 Accepted: 10 April 2017

Published online: 03 May 2017

#### References

- Anderson, TW, Darling, DA: Asymptotic theory of certain "Goodness of fit" criteria based on stochastic processes. *Ann. Stat.* **23**(2), 193–212 (1952)
- Carbon Dioxide Information Analysis Center (2017). <http://cdiac.ornl.gov>. Accessed 20 Oct 2014
- de Haan, L, Ferreira, A: *Extreme value theory: An introduction*. Springer, New York (2006)
- European Environment Agency (2017). <http://www.eea.europa.eu>. Accessed 20 Oct 2014
- Fukunaga, K: *Introduction to Statistical Pattern Recognition*. Academic Press, New York (1972)
- Hao, L, Naiman, DQ: *Quantile regression, Quantitative Applications in the Social Sciences Series, Vol. 149*. Sage Publications, Inc, USA (2007)
- Huang, ML, Xu, Y, Yuen, WK: On quantile regression for extremes. In: *JSM Proceedings, Statistical Computation Section*, pp. 561–601. American Statistical Association, Alexandria, VA, (2014)
- Huang, ML, Xu, X, Tashnev, D: A weighted linear quantile regression. *J. Stat. Comput. Simul.* **85**(13), 2596–2618 (2015)
- Koenker, R: *Quantile regression*. Cambridge University Press, New York (2005)
- Koenker, R, Bassett, GW: Regression Quantile. *Econometrica.* **46**, 33–50 (1978)
- Koenker, R, Machado, JAF: Goodness of fit and related inference processes for quantile regression. *J. Am. Stat. Assoc.* **96**(454), 1296–1311 (1999)
- Kolmogorov, AN: Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Istit. degli att.* **4**, 83–91 (1933)
- Li, Q, Racine, JS: *Nonparametric econometrics, theory and practice*. Princeton University Press, Princeton and Oxford (2007)
- Mardia, KV: Multivariate Pareto Distribution. *Ann. Math. Stat.* **33**, 1008–1015 (1962)
- National Weather Service Forecast Office (2017). <http://www.weather.gov/buf>. Accessed 22 Sept 2014
- Pickands, J: Statistical inference using extreme order statistics. *Ann. Stat.* **3**, 119–131 (1975)
- Scott, DW: *Multivariate Density Estimation, Theory, Practice and Visualization*. John Wiley & Sons, New York (1992)
- Silverman, BW: *Density estimation for statistics and data analysis*. Chapman & Hall, London, UK (1986)
- Yu, K, Lu, Z, Stander, J: Quantile regression: applications and current research areas. *Statistician.* **52**(3), 331–350 (2003)