Statistical Modelling xxxx; **xx(x)**: 1–23

Parametric estimation of non-crossing quantile functions

Gianluca Sottile[1] and Paolo Frumento[2]
[1]Department of Economics, Business and Statistics, University of Palermo, Italy
[2]Department of Political Science, University of Pisa, Italy

Abstract: Quantile regression (QR) has gained popularity during the last decades, and is now considered a standard method by applied statisticians and practitioners in various fields. In this work, we applied QR to investigate climate change by analysing historical temperatures in the Arctic Circle. This approach proved very flexible and allowed to investigate the tails of the distribution, that correspond to extreme events. The presence of quantile crossing, however, prevented using the fitted model for prediction and extrapolation. In search of a possible solution, we first considered a different version of QR, in which the QR coefficients were described by *parametric* functions. This alleviated the crossing problem, but did not eliminate it completely. Finally, we exploited the imposed parametric structure to formulate a constrained optimization algorithm that enforced monotonicity. The proposed example showed how the relatively unexplored field of parametric quantile functions could offer new solutions to the long-standing problem of quantile crossing. Our approach is particularly convenient in situations, like the analysis of time series, in which the fitted model may be used to predict extreme quantiles or to perform extrapolation. The described estimator has been implemented in the R package `qrcm`.

Key words: parametric quantile functions, quantile regression coefficients modelling (QRCM), quantile crossing; constrained optimization, R package`qrcm`

Received 15 May 2020; **Revised** 15 June 2021; **Accepted** 15 July 2021

1 Introduction

Often, researchers conduct their statistical analyses by considering global descriptors, like the mean and the variance, and simple measures of association such as mean differences or hazard ratios. In many situations, however, scientists want to describe various aspects of a distribution, and not just a single feature like the mean, aiming to identify subgroups of the population, for example the extremes, that would otherwise be diluted in a global summary measure. Estimating quantiles appear as the most natural alternative to various commonly used parametric approaches (e.g., Buchinsky 1998; Cook and Manning 2013). Quantile regression (QR; Koenker and Bassett 1978, Koenker 2005) allows to describe how the quantiles of a response variable depend on

Address for correspondence: Gianluca Sottile, Department of Economics, Business and Statistics University of Palermo, Via E. Basile, building 13, 90128 Palermo, Italy.
E-mail: gianluca.sottile@unipa.it

 10.1177/1471082X211036517

predictors. Unlike other regression methods, QR does not require strong parametric assumptions and is considered a 'distribution-free' approach.

In this work, we wanted to analyse meteorological data from European Climate Assessment and Dataset (ECAD). The data are described in Klein Tank et al. (2002), and contain daily information on minimum, maximum, and average temperature recorded at ∼4 000 meteorological stations in Europe, during a period of time ranging between few years and over a century. The objective of the analysis was to look for evidence of climatic change, and in particular to investigate whether the extreme quantiles of temperature in the Arctic Circle have risen over the past century. We estimated separated QR models for each of the available stations, using various spline functions to disentangle the seasonal variations from the long-term trend. When we tried to make predictions, however, we found that most regression models suffered to some extent from quantile crossing. For example, in some stations, the estimate of the 98th percentile was larger than that of the 99th percentile. The problem was particularly severe when we attempted to extrapolate future trends.

The literature is rich of solutions. Dette and Volgushev (2008) and Chernozhukov et al. (2009, 2010) suggested enforcing monotonicity by rearranging the estimated quantile curves. Neocleous and Portnoy (2008) showed how to construct a strictly monotone linear interpolator of the estimated quantile function. Approaches based on constrained optimization have been proposed in He (1997), Wu and Liu (2009), Bondell et al. (2010), Liu and Wu (2011), Muggeo et al. (2013), Andriyana et al. (2016) and Hutson (2018). These works differ in their identifying assumptions and utilize a variety of methods, including sequential algorithms and penalized approaches.

As pointed out by He (1997) and Bondell et al. (2010), crossing occurs because quantiles are estimated individually, and no parametric structure is assigned to the quantile function. Suppose, however, that the conditional quantile function could be described by a parametric model, say $Q(\tau \mid \boldsymbol{x}, \boldsymbol{\theta})$, in which $\tau \in (0, 1)$ is the order of the quantile, $\boldsymbol{x}$ is a vector of observed covariates, and $\boldsymbol{\theta}$ is an unknown parameter vector. In these new settings, monotonicity can be enforced by constraining $\boldsymbol{\theta}$ to lie in a feasible region, that is, a subset of the parameters' space in which $Q(\tau \mid \boldsymbol{x}, \boldsymbol{\theta})$ is a monotonically increasing function of τ. Various forms of parametric quantile functions were considered by Parzen (1980), Gilchrist (2000), Cai (2010), Cai and Reeve (2013) and Dong et al. (2015), among others. Single-index models have been described in the literature on composite QR(Zou and Yuan 2008, Kai et al. 2010, 2011, Jiang et al. 2016), while more flexible models have been used to perform Bayesian simultaneous QR(Tokdar and Kadane 2012, Reich 2012, Reich and Smith 2013, Yang and Tokdar 2017, Das and Ghosal 2017, 2018, Rodrigues et al. 2019, El Adlouni and Baldé 2019).

In our work, we follow the *quantile regression coefficients modelling* framework described in Frumento and Bottai (2016, 2017), Bottai and Cilluffo (2020) and Sottile et al. (2020), in which the coefficients of a linear QR, $Q(\tau \mid \boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta}(\tau)$, are expressed by parametric functions, $\boldsymbol{\beta}(\tau) = \boldsymbol{\beta}(\tau \mid \boldsymbol{\theta})$, and admit a linear representation of the form $\boldsymbol{\beta}(\tau \mid \boldsymbol{\theta}) = \boldsymbol{\theta}\boldsymbol{b}(\tau)$, where $\boldsymbol{b}(\tau)$ is a known set of 'basis' functions. In this model, $\boldsymbol{\theta}$ can be estimated by minimizing a simultaneous version of the objective function of ordinary QR. This method is currently implemented in two R packages, `qrcm` (Frumento 2021) and `qrcmNP` (Sottile 2021).

Our contribution to the methodology is twofold. First, we show that imposing a parametric structure can alleviate the crossing problem in comparison with the nonparametric, unstructured estimator based on standard QR. Second, we describe an algorithm for constrained optimization that can be used to enforce the non-crossing property in parametric quantile functions.

The article is structured as follows. In Section 2, we describe the data that motivated this project and provide arguments in favour of a quantile-based approach. In Section 3, we review the existing literature on QR, and describe how the coefficients could be modelled parametrically. In Section 4, we revisit the quantile crossing problem, and in Section 5, we present a method to estimate non-crossing quantile functions in a parametric world. In Section 6, we illustrate a simulation study. In Section 7, we present a sample of our results on the ECAD data. Section 8 concludes the article. The article is accompanied by a new version of the R package `qrcm`, in which crossing can be diagnosed and eliminated using the method described in the article.

2 Investigating climate change using the ECAD data

In the last decades, global warming has been the subject of an increasingly large number of articles (e.g., Cox et al. 2000, Root et al. 2003) and scientific reports (e.g., Shaftel 2016), and is now considered one of the most important challenges of mankind in the coming century. The global rise of temperature has been associated with an increased frequency of extreme meteorological events (e.g., Kuswanto et al. 2015, Vavrus et al. 2015), including heat waves, tornadoes, and extreme winds and precipitations. Long-term trends in global temperature have been investigated by Hanssen-Bauer and Førland (1998), Meehl et al. (2000), Manton et al. (2001), Rahmstorf and Coumou (2011), Russo et al. (2014), among many others, while the consequences of extreme temperatures on global health are described in Medina-Ramon et al. (2006).

In this work, we used meteorological data from European Climate Assessment and Dataset (ECAD). The data are described in Klein Tank et al. (2002) and can be downloaded from the ECAD website (https://www.ecad.eu/). The datasets contain daily information on temperature, precipitations, humidity, wind, sea level pressure and cloud cover from over 4 000 meteorological stations across Europe. Most time series start in the early 1900s and can be utilized to investigate long-term trends. We considered two response variables, the minimum and the maximum daily temperature. We were particularly interested in extreme events, that correspond to very low or high quantiles of the distribution.

The analysis of climatological data is often based on relatively simple indicators, such as the frequency of days in which the minimum (maximum) temperature is below (above) a certain threshold of interest. While these quantities can be summarized and represented graphically using moving averages, fitting a statistical model makes it possible to perform a more sophisticated and in-depth analysis. Regression methods provide natural tools to separate the seasonal variations from the long-term trend. Additionally, statistical models allow to describe the entire distribution, and not just a single indicator like the mean or the median. Finally, describing time trends by smooth functions, such as polynomials and splines, eliminates the noise and allows for extrapolation of future events.

To avoid unrealistic assumptions, such as symmetry, normality, and homoskedasticity, one could use some form of flexible distribution. An alternative option is to not specify the distribution at all, and to apply QR. QR estimators are essentially nonparametric, and can be applied like a sort of 'black box' method to all of the available meteorological stations. In our case, the only parametric assumption was the form of the linear predictor that, however, was entirely composed of polynomials splines and their interactions. By computing the percentiles, corresponding to quantiles of order $\tau = \{0.01, 0.02, \dots, 0.99\}$, we could estimate the full conditional distribution of the minimum and maximum daily temperature. Note that this approach generates a direct estimate of the extreme quantiles, that correspond to very large/small values of τ. Similar ideas were used in a recent article by Do Nascimento and Bourguignon (2020).

While using QR seems very appealing, it also comes with some complications. In particular, QR estimators suffered from severe crossing, as detailed in Section 7. Crossing occurred more frequently in the tails, and was more pronounced when the model was used for extrapolation. In search of a solution, we first considered replacing the coefficient functions, $\boldsymbol{\beta}(\tau)$, with a flexible parametric model, $\boldsymbol{\beta}(\tau \mid \boldsymbol{\theta})$, described by orthogonal polynomials. With this approach, crossing dropped drastically but was not completely eliminated. We then applied a constrained optimization algorithm to ensure that the model parameters, $\boldsymbol{\theta}$, lie in a feasible region where no crossing occurs. All technical details are provided in the reminder of this article.

3 Parametric and nonparametric estimation of conditional quantile functions

We review the existing methodology for QR, which can be seen as a nonparametric estimator of a quantile function, and describe its parametric counterpart, referred to as *quantile regression coefficients modelling* (QRCM).

3.1 Quantile regression (QR)

We denote by $Q(\tau \mid \boldsymbol{x})$ the quantile function of a response variable Y, conditional on a q-dimensional vector $\boldsymbol{x}$ of covariates. A standard representation of $Q(\tau \mid \boldsymbol{x})$ is the linear quantile function

$$Q(\tau \mid \boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{\beta}(\tau), \tag{3.1}$$

in which $\boldsymbol{\beta}(\tau)$ is a vector of unknown regression coefficients. Although assuming a linear effect of covariates is not crucial, it simplifies computation, inference and interpretation of the parameters. To achieve a sufficient flexibility, the vector $\boldsymbol{x}$ may include transformed predictors, such as polynomials, splines and interaction terms, and the model can be applied to a monotone transformation of the response variable. Assume model (3.1) holds, and is denoted by $\{y_i, \boldsymbol{x}_i\}_{i=1,\ldots,n}$ a random sample from the joint distribution of $(Y, \boldsymbol{x})$. In standard QR, an estimate $\hat{\boldsymbol{\beta}}(\tau)$ of $\boldsymbol{\beta}(\tau)$ is computed as the minimizer of

$$l(\boldsymbol{\beta}(\tau)) = \sum_{i=1}^{n} \rho_\tau(y_i - \boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}(\tau)), \tag{3.2}$$

where $\rho_\tau(u) = (\tau - I(u \leq 0))u$. Equivalently, $\hat{\boldsymbol{\beta}}(\tau)$ is the q-dimensional root of

$$\boldsymbol{s}(\boldsymbol{\beta}(\tau)) = \nabla_{\boldsymbol{\beta}(\tau)} l(\boldsymbol{\beta}(\tau)) = \sum_{i=1}^{n} \phi_\tau(y_i - \boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}(\tau)), \tag{3.3}$$

where $\phi_\tau(u) = (I(u \leq 0) - \tau)$. Because quantiles of different order are estimated individually, standard QR is thought of as a nonparametric method, the only parametric assumption being the functional form of the linear predictor.

3.2 Quantile regression coefficients modelling (QRCM)

Frumento and Bottai (2016) suggested replacing $Q(\tau \mid \boldsymbol{x})$ by a parametric quantile function $Q(\tau \mid \boldsymbol{x}, \boldsymbol{\theta})$, and reformulated model (3.1) as follows:

$$Q(\tau \mid \boldsymbol{x}, \boldsymbol{\theta}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta}(\tau \mid \boldsymbol{\theta}). \tag{3.4}$$

Their approach is referred to as *quantile regression coefficients modelling* (QRCM), and is currently implemented in two R packages, `qrcm` and `qrcmNP`. A convenient parametrization of model (3.4) is obtained when $\boldsymbol{\beta}(\tau \mid \boldsymbol{\theta})$ is a linear combination of known 'basis' functions, $\boldsymbol{b}(\tau) = [b_1(\tau), \ldots, b_k(\tau)]^{\mathrm{T}}$, such that $\boldsymbol{\beta}(\tau \mid \boldsymbol{\theta}) = \boldsymbol{\theta}\boldsymbol{b}(\tau)$. In this model, $\boldsymbol{\theta}$ is a $q \times k$ matrix, and the quantile function can be rewritten as

$$Q(\tau \mid \boldsymbol{x}, \boldsymbol{\theta}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\theta}\boldsymbol{b}(\tau). \tag{3.5}$$

For instance, if $\boldsymbol{b}(\tau) = [1, \tau, \tau^{1/2}]^{\mathrm{T}}$, the regression coefficients are given by $\beta_j(\tau \mid \boldsymbol{\theta}) = \theta_{0j} + \theta_{1j}\tau + \theta_{2j}\tau^{1/2}$, $j = 1, \ldots, q$. This type of parametrization simplifies computation and inference, without constituting a real limitation in terms of model flexibility. An estimate $\hat{\boldsymbol{\theta}}$ of the unknown parameter vector $\boldsymbol{\theta}$ is computed as the minimizer of

$$L(\boldsymbol{\theta}) = \sum_{i=1}^{n} \int_0^1 \rho_\tau(y_i - \boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}(\tau \mid \boldsymbol{\theta}))\mathrm{d}\tau, \tag{3.6}$$

which is the integral, with respect to τ, of the loss of ordinary QR defined in (3.2). Equivalently, $\hat{\boldsymbol{\theta}}$ is the root of

$$S(\boldsymbol{\theta}) = \nabla_{\boldsymbol{\theta}} L(\boldsymbol{\theta}) = \int_0^1 \sum_{i=1}^{n} \phi_\tau(y_i - \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{\beta}(\tau \mid \boldsymbol{\theta}))\boldsymbol{b}(\tau)\mathrm{d}\tau. \tag{3.7}$$

A closed-form expression for $L(\boldsymbol{\theta})$ and $\boldsymbol{S}(\boldsymbol{\theta})$ under model (3.5) is derived in Frumento and Bottai (2016). Discrete versions of $L(\boldsymbol{\theta})$, in which the integral was replaced by a (weighted) sum, were used by Koenker (2004) and Zou and Yuan (2008) to improve estimation of quantile-invariant parameters. An extension to censored and truncated data is described in Frumento and Bottai (2017).

The advantages of the described parametric approach are numerous. While standard QR generates a non-smooth, grid-based estimate of the quantile function, $Q(\tau \mid \boldsymbol{x}, \boldsymbol{\theta})$ is usually defined to be a continuous and differentiable function of τ. For the same reason, $\boldsymbol{S}(\boldsymbol{\theta})$ is a smooth function of $\boldsymbol{\theta}$, while $\boldsymbol{s}(\boldsymbol{\beta}(\tau))$ is not. This permits solving $\boldsymbol{S}(\boldsymbol{\theta})$ by standard Newton-type algorithms, that are usually faster (and simpler) than the modified simplex method used in ordinary QR (Barrodale and Roberts 1973). Moreover, the differentiability of the objective function allows applying a simpler asymptotic theory (e.g., Newey and McFadden 1994), and avoids the problem of estimating the sparsity function as in ordinary QR (Koenker and Machado 1999). The closed-form analytical expression of $Q(\tau \mid \boldsymbol{x}, \boldsymbol{\theta})$ makes it simple to predict or extrapolate from the fitted model, to compute extreme quantiles and, in the context of this article, to detect quantile crossing.

3.3 An example

Consider the simulated data represented in Figure 1, and assume to fit a QR model $Q(\tau \mid x) = \beta_0(\tau) + \beta_1(\tau)x$. In the bottom-left panels of the figure, we report ordinary QR coefficients estimated at a grid of percentiles, $\tau = \{0.01, 0.02, \dots, 0.99\}$. In the bottom-right panels, we show how coefficients could be modelled parametrically. In the example, the intercept is described by the quantile function of an Exponential distribution, $\beta_0(\tau \mid \boldsymbol{\theta}) = -\theta_{01} \log(1 - \tau)$, and the slope is parametrized by a cubic function, $\beta_1(\tau \mid \boldsymbol{\theta}) = \theta_{11}\tau + \theta_{12}\tau^2 + \theta_{13}\tau^3$.

The 'nonparametric' estimators, based on ordinary QR, show a relatively simple trend contaminated by a significant background noise, which could reasonably be eliminated with no real loss of information. Due to data sparsity, standard errors become very large as τ approaches one. On the other hand, using a model-based approach generates a significant gain in efficiency, especially in the tails, and permits characterizing the entire quantile function with only few parameters. The closed-form mathematical expression of the regression coefficients makes it very easy to extrapolate extreme quantiles [e.g., to compute $Q(0.999 \mid \boldsymbol{x}, \hat{\boldsymbol{\theta}})$], and can be used to verify the presence of quantile crossing. In the example, the estimated intercept is given by $\hat{\beta}_0(\tau) = \beta_0(\tau \mid \hat{\boldsymbol{\theta}}) = -0.79 \log(1 - \tau)$, while the slope is $\hat{\beta}_1(\tau) =$

$\beta_1(\tau \mid \hat{\boldsymbol{\theta}}) = 0.12\tau + 0.43\tau^2 + 0.68\tau^3$. Since $\hat{\beta}'_0(\tau)$ and $\hat{\beta}'_1(\tau)$ are always positive, no crossing is possible for $x \geq 0$. The cost of the described QRCM framework is that it requires to formulate a parametric model for $\boldsymbol{\beta}(\tau \mid \boldsymbol{\theta})$. Meaningful options include polynomials, trigonometric or logarithmic functions, quantile functions of known distributions, and combinations of the above. More flexible models can be specified by using splines, high-degree polynomials, or piecewise linear functions. Particularly simple parametrizations include the linear coefficient, $\beta(\tau \mid \boldsymbol{\theta}) = \theta_0 + \theta_1\tau$, and the constant-slope model, $\beta(\tau \mid \boldsymbol{\theta}) = \theta_0$. On the opposite extreme, standard QRcan be seen as a special case of QRCM in which all coefficients are modelled by 'very flexible' parametric functions.

4 Crossing in parametric quantile functions

Assume to estimate a quantile function $Q(\tau \mid \boldsymbol{x}, \boldsymbol{\theta})$, and denote by $Q'(\tau \mid \boldsymbol{x}, \boldsymbol{\theta})$ its first derivative. In the standard QRmodel, $Q(\tau \mid \boldsymbol{x}, \boldsymbol{\theta}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta}(\tau \mid \boldsymbol{\theta})$ and $Q'(\tau \mid \boldsymbol{x}, \boldsymbol{\theta}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta}'(\tau \mid \boldsymbol{\theta})$, which simplify to $Q(\tau \mid \boldsymbol{x}, \boldsymbol{\theta}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\theta}\boldsymbol{b}(\tau)$ and $Q'(\tau \mid \boldsymbol{x}, \boldsymbol{\theta}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\theta}\boldsymbol{b}'(\tau)$ under model (3.5). Denote by $\hat{\boldsymbol{\theta}}$ the minimizer of the loss function defined in equation (3.6). Quantile crossing occurs if the set $\{\tau : Q'(\tau \mid \boldsymbol{x}, \hat{\boldsymbol{\theta}}) < 0\}$ is non-empty. It should be clarified that all linear quantile functions will cross at some $\boldsymbol{x}$, unless all regression lines are forced to be parallel. To be considered relevant, crossing must occur at meaningful covariates' values. Crossing may be due to a variety of reasons, some of which are illustrated by the following examples, with the graphical support of Figure 2.

Example (a). Misspecified model with empty feasible region. Consider the quantile function defined by $Q(\tau \mid x, \theta) = \theta\tau x$. As shown in Figure 2**a**, all regression lines cross at $x = 0$, where the model is assumed to be degenerated. If x only takes either positive or negative values, this model is guaranteed to be non-crossing. Otherwise, crossing occurs at all values of θ and cannot be avoided. Other examples of ill-defined quantile functions are $Q(\tau \mid \theta) = \theta\cos(2\pi\tau)$ and $Q(\tau \mid \theta) = \theta(\tau - 0.5)^2$.

Example (b). Crossing caused by outliers. Define $Q(\tau \mid x, \boldsymbol{\theta}) = \beta_0(\tau \mid \boldsymbol{\theta}) + \beta_1(\tau \mid \boldsymbol{\theta})x$ with $\beta_0(\tau \mid \boldsymbol{\theta}) = \theta_{00} + \theta_{01}\tau$ and $\beta_1(\tau \mid \boldsymbol{\theta}) = \theta_{10} + \theta_{11}\tau$, and $x \in (0, 1)$. A simple condition for monotonicity is that $(\hat{\theta}_{01}, \hat{\theta}_{11}) \geq 0$. However, as shown in Figure 2**b**, the estimated regression lines may cross in correspondence of outlying values.

Example (c). Non-monotone coefficient. Define $Q(\tau \mid x, \boldsymbol{\theta}) = \theta_0\tau - \theta_1(\tau - 0.75)^2 x$, and assume $\boldsymbol{\theta} > \mathbf{0}$ and $x \geq 0$. The regression model is illustrated in Figure 2**c**. The coefficient associated with x is assumed to be a non-monotone function, $\beta_1(\tau \mid \boldsymbol{\theta}) = -\theta_1(\tau - 0.75)^2$, which may obviously induce crossing. A non-crossing quantile function can be obtained by imposing $2\hat{\theta}_0 \geq \hat{\theta}_1 \max_i(x_i)$.

Example (d). Crossing in a flexible model. In most situations, the true model is not known and the coefficients are described by splines, polynomials, or other types of smooth flexible functions. Consider, for example, the data illustrated

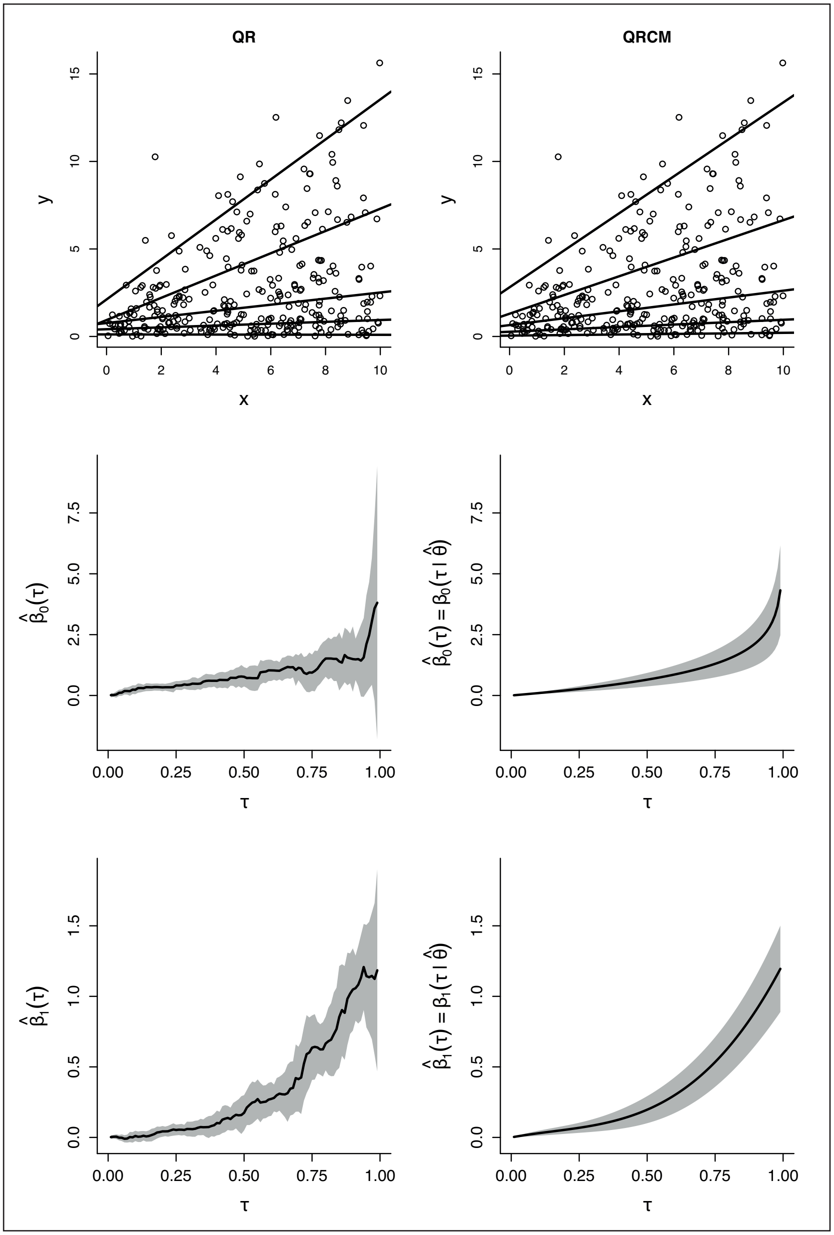

Figure 1 The same data are analysed using QR (left) and QRCM (right), with parametrization $\beta_0(\tau \mid \boldsymbol{\theta}) = -\theta_{01}\log(1-\tau)$, $\beta_1(\tau \mid \boldsymbol{\theta}) = \theta_{11}\tau + \theta_{12}\tau^2 + \theta_{13}\tau^3$. Top: scatterplot of y versus x (simulated data) with regression lines corresponding to quantiles 0.05, 0.25, 0.5, 0.75, 0.95. Bottom: estimated regression coefficients with 95% confidence intervals

in Figure 2**d**, and suppose to fit a quantile function $Q(\tau \mid x, \boldsymbol{\theta}) = \beta_0(\tau \mid \boldsymbol{\theta}) + \beta_1(\tau \mid \boldsymbol{\theta})x + \beta_2(\tau \mid \boldsymbol{\theta})x^2 + \beta_3(\tau \mid \boldsymbol{\theta})x^3$, in which $\beta_j(\tau \mid \boldsymbol{\theta}) = \theta_{j0} + \theta_{j1}\tau + \theta_{j2}\tau^2 + \theta_{j3}\tau^3$, $j = 0, \ldots, 3$. In this example, quantiles are cubic functions of x, and regression coefficients are cubic functions of τ. Crossing at extreme quantiles arises from the combination between a very flexible parametric structure and a relatively small sample size. Although analytical identification of the feasible region [as in Examples (b) and (c)] is not possible here, crossing can be easily diagnosed using the fact that $Q'(\tau \mid \boldsymbol{x}, \boldsymbol{\theta})$ has a simple closed-form expression.

Situations like those described in Example (a), in which the feasible region is empty, must be considered pathological and may only occur if the model is severely misspecified and is subject to restrictive assumptions. Example (c) suggests that crossing can be attenuated or completely eliminated by a suitable parametrization of the quantile function. For example, a general form of non-crossing quantile function is obtained when $\boldsymbol{x} \geq \mathbf{0}$ and $\boldsymbol{\beta}'(\tau \mid \boldsymbol{\theta}) > 0$, that is, all coefficients are monotonically increasing functions of τ. This model can be used to build growth charts (e.g., Muggeo et al. 2013). As suggested by Example (d), the risk of crossing tends to increase as the fitted model approaches the 'nonparametric' estimator obtained with standard quantile regression. In particular, crossing is often caused by an erratic tail behaviour, which can be mitigated by imposing a parametric structure.

5 Parametric estimation of non-crossing quantile functions

We describe a numerical method to estimate a parametric quantile function with the non-crossing property. The problem to be solved is the following:

$$\begin{aligned} &\min L(\boldsymbol{\theta}) \\ &\text{s.t. } Q'(\tau \mid \boldsymbol{x}_i, \boldsymbol{\theta}) \geq 0, \quad i = 1, \ldots, n, \end{aligned} \tag{5.1}$$

where $L(\boldsymbol{\theta})$ is the loss function defined in equation (3.6). Because non-negativity of $Q'(\tau \mid \boldsymbol{x}_i, \boldsymbol{\theta})$ must hold at all values of τ in $(0, 1)$, problem (5.1) has an infinite number of constraints. An equivalent problem with only n constraints is given by

$$\begin{aligned} &\min L(\boldsymbol{\theta}) \\ &\text{s.t. } \int_0^1 |Q'_-(\tau \mid \boldsymbol{x}_i, \boldsymbol{\theta})| d\tau = 0, \quad i = 1, \ldots, n, \end{aligned} \tag{5.2}$$

where $Q'_-(\tau \mid \boldsymbol{x}_i, \boldsymbol{\theta}) = \min\{0, Q'(\tau \mid \boldsymbol{x}_i, \boldsymbol{\theta})\}$ denotes the negative part of $Q'(\tau \mid \boldsymbol{x}_i, \boldsymbol{\theta})$. Each constraint is a function of $\boldsymbol{\theta}$, and is equal to zero if no crossing occurs, and a positive quantity otherwise. To solve problem (5.2), we suggest using a penalized approach as in Richardson et al. (1989), Siedlecki and Sklansky (1989) and Baeck

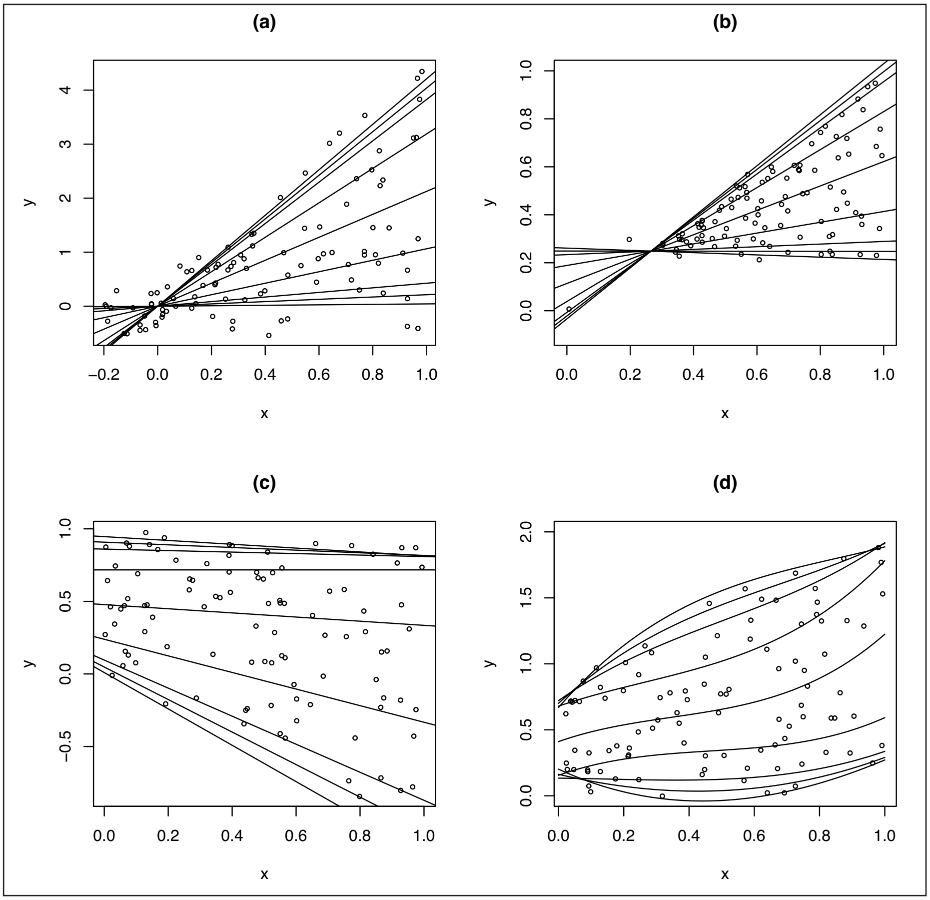

Figure 2 Simulated data with superimposed parametric models suffering from quantile crossing. Lines represent estimated quantiles of order $\tau = \{0.01, 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95, 0.99\}$. Figures a–d correspond to examples a–d in the text. Figure a: misspecified model with empty feasible region. Figure b: crossing caused by outliers. Figure c: crossing caused by a non-monotone coefficient. Figure d: flexible model with crossing at extreme quantiles

et al. (1997), and estimating $\boldsymbol{\theta}$ as the minimizer of

$$\begin{aligned} L_\lambda(\boldsymbol{\theta}) &= L(\boldsymbol{\theta}) + \lambda P(\boldsymbol{\theta}) \\ &= \sum_{i=1}^{n} \int_0^1 \rho_\tau(y_i - Q(\tau \mid \boldsymbol{x}_i, \boldsymbol{\theta}))\mathrm{d}\tau + \lambda \sum_{i=1}^{n} \int_0^1 |Q'_-(\tau \mid \boldsymbol{x}_i, \boldsymbol{\theta})|\mathrm{d}\tau, \end{aligned} \tag{5.3}$$

where λ is a tuning constant, and $P(\boldsymbol{\theta})$ is a penalty term computed as the sum of all constraints. The value of $P(\boldsymbol{\theta})$ reflects both the sign and the absolute size of $Q'(\tau \mid \boldsymbol{x}_i, \boldsymbol{\theta})$, and is zero at values of $\boldsymbol{\theta}$ that generate a non-crossing quantile function.

The tuning parameter λ defines the balance between the two ingredients of $L_\lambda(\boldsymbol{\theta})$, namely the unpenalized loss function, $L(\boldsymbol{\theta})$, and the penalty term, $P(\boldsymbol{\theta})$. If λ is too small, the penalization has no effect and the constraints are ignored. As λ increases, the relative weight of $P(\boldsymbol{\theta})$ becomes larger, and $\hat{\boldsymbol{\theta}}$ is pushed towards the feasible region. However, if λ is unnecessarily large, convexity of $L_\lambda(\boldsymbol{\theta})$ is no longer guaranteed, because $P(\boldsymbol{\theta})$, unlike $L(\boldsymbol{\theta})$, is not generally a convex function. The minimization algorithm is initialized with a relatively small value of λ, which is progressive increased until one of the following conditions has been reached: (a) a well-defined, non-crossing solution has been identified; (b) any further increase of λ generates an ill-defined problem with a non-convex objective function; (c) the process is stopped by the user. Case (a) is by far the most common and desirable outcome. Case (b) may interrupt the algorithm if the feasible region is either empty, or very difficult to reach due to severe model misspecification. Although crossing has not been eliminated completely in this scenario, the current solution is usually better than the unconstrained model. Finally, Case (c) may be the consequence of a cost-benefit analysis. For instance, if the minimised loss function of the constrained model is much larger than that of the unpenalized fit, one may consider choosing an intermediate solution in which the estimated quantile function is allowed to cross at some values in the support.

5.1 Asymptotic properties

We denote by $\boldsymbol{\theta}_0$ the true population parameter, and by $\hat{\boldsymbol{\theta}}$ the minimizer of the penalized loss function, $L_\lambda(\boldsymbol{\theta}) \triangleq L(\boldsymbol{\theta}) + \lambda P(\boldsymbol{\theta})$ (equation 5.3) or the solution to the estimating equation represented by $\boldsymbol{S}_\lambda(\boldsymbol{\theta}) \triangleq \boldsymbol{S}(\boldsymbol{\theta}) + \lambda \nabla_{\boldsymbol{\theta}} P(\boldsymbol{\theta})$. Note that, by definition, $P(\boldsymbol{\theta}_0) = 0$. This shows that, assuming that the model is correct, quantile crossing is a finite-sample problem. As n tends to ∞, solutions that generate quantile crossing will not be supported by the data, and the properties of $\hat{\boldsymbol{\theta}}$ will be identical to those of the unconstrained estimator. Under standard regularity conditions (e.g., Newey and McFadden 1994),

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} N\left(0, (\boldsymbol{H}^{\mathrm{T}} \boldsymbol{\Omega}^{-1} \boldsymbol{H})^{-1}\right)$$

where $\boldsymbol{H} = E[\nabla_{\boldsymbol{\theta}} \boldsymbol{S}(\boldsymbol{\theta}_0)]$, and $\boldsymbol{\Omega} = E[\boldsymbol{S}(\boldsymbol{\theta}_0)\boldsymbol{S}(\boldsymbol{\theta}_0)^{\mathrm{T}}]$. As shown by Frumento and Bottai (2016, 2017), to whom we refer for the technical details of the proof, a non-trivial condition is that $\int Q(\tau \mid \boldsymbol{x}, \boldsymbol{\theta}) d\tau$ must be finite. This rules out models in which coefficients are described by non-integrable functions, such as $\tan(\pi(\tau - 0.5))$, $1/\tau$, and $1/(1-\tau)$.

6 Simulation results

We conducted a simulation study to address two equally important research questions. First, we wanted to compare unconstrained QRCM estimators with ordinary

QR, to verify if the sheer fact of imposing a parametric structure can stabilize the tail behaviour and alleviate the crossing problem. Second, we wanted to assess the performance of the constrained QRCM estimator introduced in Section 5. All the necessary software is provided and documented in the R package `qrcm`, version 3.0.

We considered the quantile function defined by

$$Q(\tau \mid \boldsymbol{x}, \boldsymbol{\theta}) = \beta_0(\tau \mid \boldsymbol{\theta}) + \beta_1(\tau \mid \boldsymbol{\theta})x_1 + \cdots + \beta_5(\tau \mid \boldsymbol{\theta})x_5,$$

in which $(x_1, x_2, \ldots, x_5) \sim \mathcal{N}^+(\mathbf{0}, \mathcal{I})$. Here, $\mathcal{N}^+$ identifies a truncated version of the multivariate normal distribution in which the support only includes the positive real axis. We defined

$$\beta_0(\tau) = 10(1 - (1 - \tau)^{1/10}), \quad \beta_1(\tau) = 2\tau, \quad \beta_2(\tau) = 2.5\tau + 0.75\sin(\pi\tau),$$

$$\beta_3(\tau) = 0.3\tau + \tau/(1.3 - \tau), \quad \beta_4(\tau) = 3Q_{\text{Beta}}(\tau, 2, 3), \quad \beta_5(\tau) = 0,$$

where $Q_{\text{Beta}}(\tau, a, b)$ denotes the quantile function of a Beta distribution. We used two different sample sizes, namely $n = 300$ and $n = 600$. For each sample size, we generated $B = 1\,000$ simulated datasets. We then estimated the regression model using three different methods: (a) standard QR; (b) the unconstrained QRCM estimator, obtained as the minimizer of $L(\boldsymbol{\theta})$; and (c) the constrained QRCM estimator, obtained as the minimizer of the penalized loss function $L_\lambda(\boldsymbol{\theta})$ described in Section 5 and denoted by QRCM$_c$.

QR estimators were evaluated at a fine grid of quantiles, namely $\tau = \{0.001, 0.002, \ldots\}$. To apply QRCM, we assumed to be ignorant about the data-generating process and described each regression coefficient by a 5th-degree shifted Legendre polynomial, an orthogonal polynomials in $(0, 1)$ (Abramovitz 1965). With this definition, the model had 36 free parameters to be estimated. The three estimators mainly differed in their tail behaviour. In Table 1, we report the average value and empirical standard errors of the estimated QR coefficients at $\tau = \{0.90, 0.95, 0.975\}$. As expected, parametric models are typically more efficient than fully nonparametric estimators. Importantly, the constrained QRCM estimator did not appear to introduce bias, and was often superior to the unconstrained model in terms of standard errors.

Assessing monotonicity of the estimated quantile function is not trivial, because crossing may or may not be detected depending on the resolution at which $\hat{Q}(\tau \mid \boldsymbol{x})$ is measured. Suppose, for example, that $\hat{Q}(0.90 \mid \boldsymbol{x})$ is smaller than $\hat{Q}(0.91 \mid \boldsymbol{x})$, but larger than $\hat{Q}(0.901 \mid \boldsymbol{x})$. In this situation, the percentiles appear to form a monotonically increasing function, and the presence of crossing is only revealed by a sufficiently dense grid of quantiles, for example, $\tau = \{0.001, 0.002, \ldots, 0.999\}$. One may argue that the number of quantiles being evaluated should not exceed the sample size. Intuitively, if crossing is only detected using a grid step of 0.001, it should not be considered relevant unless $n > 1\,000$. In our analysis, we considered three different grids of quantiles, namely $\tau_1 = \{0.05, 0.10, \ldots, 0.95\}$, $\tau_2 = \{0.01, 0.02, \ldots, 0.99\}$, and $\tau_3 = \{0.001, 0.002, \ldots, 0.999\}$.

Table 1 Average estimates and empirical standard errors

	n		β_0	β_1	β_2	β_3	β_4	β_5
$\tau = 0.90$		true	2.06	1.80	2.48	2.52	2.04	0.00
	300	QR	2.14 (0.93)	1.78 (0.56)	2.44 (0.53)	2.46 (0.65)	2.04 (0.59)	−0.02 (0.51)
		QRCM	2.21 (0.82)	1.77 (0.50)	2.44 (0.48)	2.45 (0.58)	2.02 (0.52)	−0.01 (0.45)
		QRCM$_c$	2.21 (0.80)	1.76 (0.49)	2.44 (0.47)	2.45 (0.57)	2.02 (0.51)	−0.02 (0.43)
	600	QR	2.10 (0.64)	1.79 (0.39)	2.49 (0.37)	2.49 (0.46)	2.01 (0.41)	−0.02 (0.36)
		QRCM	2.15 (0.56)	1.78 (0.34)	2.49 (0.33)	2.51 (0.40)	2.03 (0.36)	−0.02 (0.32)
		QRCM$_c$	2.14 (0.55)	1.78 (0.34)	2.49 (0.32)	2.51 (0.39)	2.03 (0.35)	−0.03 (0.31)
$\tau = 0.95$		true	2.59	1.90	2.49	3.00	2.25	0.00
	300	QR	2.68 (1.04)	1.87 (0.62)	2.47 (0.60)	2.88 (0.73)	2.19 (0.65)	0.01 (0.59)
		QRCM	2.79 (0.97)	1.87 (0.56)	2.48 (0.55)	2.86 (0.66)	2.21 (0.59)	−0.01 (0.51)
		QRCM$_c$	2.78 (0.95)	1.87 (0.55)	2.48 (0.54)	2.85 (0.65)	2.21 (0.58)	−0.01 (0.50)
	600	QR	2.64 (0.74)	1.89 (0.42)	2.49 (0.42)	2.94 (0.49)	2.23 (0.45)	−0.03 (0.41)
		QRCM	2.70 (0.67)	1.89 (0.38)	2.51 (0.38)	2.95 (0.44)	2.24 (0.41)	−0.04 (0.38)
		QRCM$_c$	2.70 (0.67)	1.89 (0.38)	2.52 (0.37)	2.95 (0.44)	2.24 (0.41)	−0.04 (0.37)
$\tau = 0.975$		true	3.08	1.95	2.50	3.29	2.42	0.00
	300	QR	3.21 (1.20)	1.90 (0.71)	2.46 (0.69)	3.14 (0.81)	2.33 (0.74)	0.01 (0.63)
		QRCM	3.16 (1.17)	1.91 (0.66)	2.50 (0.65)	3.10 (0.76)	2.33 (0.70)	−0.02 (0.61)
		QRCM$_c$	3.16 (1.11)	1.93 (0.62)	2.51 (0.62)	3.08 (0.74)	2.32 (0.67)	0.00 (0.58)
	600	QR	3.12 (0.86)	1.94 (0.48)	2.52 (0.48)	3.22 (0.55)	2.38 (0.51)	−0.04 (0.48)
		QRCM	3.06 (0.83)	1.95 (0.45)	2.52 (0.45)	3.21 (0.52)	2.37 (0.48)	−0.05 (0.45)
		QRCM$_c$	3.06 (0.80)	1.95 (0.44)	2.53 (0.44)	3.20 (0.51)	2.37 (0.47)	−0.04 (0.43)

Notes: True coefficients, average estimates across the simulation, and standard errors (in brackets), for $\tau = \{0.90, 0.95, 0.975\}$. In the table, QR refers to standard quantile regression, while QRCM and QRCM$_c$ indicate unconstrained and constrained parametric estimators, respectively.

In Table 2, we report the proportion of datasets in which crossing was detected, and summary statistics (mean and standard deviations across the simulation) of the following measures of crossing:

$$P_{\text{cross}} = n^{-1} \sum_{i=1}^{n} I\Big(\int_0^1 |Q'_{-}(\tau \mid \boldsymbol{x}_i, \hat{\boldsymbol{\theta}})| \mathrm{d}\tau > 0 \Big), \tag{6.1}$$

$$L_{\text{cross}} = n^{-1} \sum_{i=1}^{n} \int_0^1 I\Big(Q'(\tau \mid \boldsymbol{x}_i, \hat{\boldsymbol{\theta}}) < 0 \Big) \mathrm{d}\tau, \tag{6.2}$$

where $Q'_{-}(\tau \mid \boldsymbol{x}_i, \boldsymbol{\theta}) = \min\{0, Q'(\tau \mid \boldsymbol{x}_i, \hat{\boldsymbol{\theta}})\}$ is the negative part of $Q'(\tau \mid \boldsymbol{x}_i, \hat{\boldsymbol{\theta}})$ as in Section 5. P_{cross} can be interpreted as the proportion of observations for which the estimated quantile function is non-monotone, and was measured using each of the three grids of quantiles defined earlier. L_{cross} is the average length of the crossing region on the τ scale, and was only measured using the dense grid τ_3. Results showed that crossing occurs very frequently in standard QR, and can be significantly alleviated

Table 2 Bias and variance indicators in simulation

	n	τ	QR	QRCM	QRCM$_c$
Datasets with crossing (%)	300	$\tau 1$	100.00	89.10	0.00
		$\tau 2$	100.00	99.50	0.00
		$\tau 3$	100.00	99.70	0.00
	600	$\tau 1$	100.00	56.80	0.00
		$\tau 2$	100.00	89.10	0.00
		$\tau 3$	100.00	92.30	0.00
$100 \times P_{\text{cross}}$	300	$\tau 1$	37.59 (6.27)	2.24 (2.42)	0.00 (0.00)
		$\tau 2$	99.47 (0.49)	11.00 (7.73)	0.00 (0.00)
		$\tau 3$	100.00 (0.00)	13.79 (9.07)	0.00 (0.00)
	600	$\tau 1$	13.42 (3.88)	0.37 (0.59)	0.00 (0.00)
		$\tau 2$	95.53 (1.27)	3.39 (3.78)	0.00 (0.00)
		$\tau 3$	100.00 (0.00)	4.90 (5.07)	0.00 (0.00)
$100 \times L_{\text{cross}}$	300		15.83 (0.63)	0.56 (0.40)	0.00 (0.00)
	600		13.00 (0.51)	0.15 (0.17)	0.00 (0.00)

Notes: Proportion of simulated datasets in which quantile crossing was detected, and average values of P_{cross} and L_{cross} across the simulation (standard deviations in brackets). The presence of crossing and the value of P_{cross} were measured using three different grids of quantiles, namely $\tau_1 = \{0.05, 0.10, \dots, 0.95\}$, $\tau_2 = \{0.01, 0.02, \dots, 0.99\}$ and $\tau_3 = \{0.001, 0.002, \dots, 0.999\}$. The value of L_{cross} was only computed using the dense grid τ_3.

by fitting a parametric model. Using QR, the average values of P_{cross} were already unacceptably high on the low-density grid $\tau_1 = \{0.05, 0.10, \dots, 0.95\}$, and achieved the maximum possible value of 100% when the dense grid $\tau_3 = \{0.001, 0.002, \dots, 0.999\}$ was used. In comparison, the average P_{cross} never exceeded 14% using QRCM. Similarly, the values of L_{cross} were much lower with QRCM than with QR. Finally, the constrained parametric estimator eliminated crossing entirely in all simulated datasets.

For comparison, we also applied to the simulated data a Normal heteroskedastic model, $\mathcal{N}(\mu = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta}, \log\sigma = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\phi})$, where $\boldsymbol{x} = [1, x_1, x_2, \dots, x_5]^{\mathrm{T}}$. This model is relatively simple and cannot suffer from quantile crossing. However, it assumes a symmetric, bell-shaped conditional distribution, that may not be sufficiently flexible in general. In Table 3, we report the predicted quantiles $\{0.90, 0.95, 0.975\}$ at $\boldsymbol{x} = [1, 1, 1, 1, 1, 1]^{\mathrm{T}}$. Results suggest that the Normal heteroskedastic model is likely to suffer from a significant bias, which can often be explained by a misspecified tail behaviour.

7 Analysis of temperatures in the Arctic region

We present our analysis of the ECAD data, described in Section 2. We considered two response variables, namely the minimum and the maximum daily temperature. Other possible outcomes of interest include mean daily temperature, wind speed, and precipitation. We were particularly interested in the extreme quantiles, corresponding

Table 3 Predicted quantiles: comparison with a Normal heteroskedastic model.

	Method	Measure	τ 0.90	0.95	0.975
$n = 300$	QR	Bias	−0.059	−0.125	−0.195
		Variance	0.235	0.290	0.365
	Norm heter	Bias	−0.618	−0.711	−0.639
		Variance	0.207	0.257	0.307
	QRCM	Bias	−0.017	−0.068	−0.302
		Variance	0.186	0.229	0.319
	QRCM$_c$	Bias	−0.030	−0.042	−0.239
		Variance	0.183	0.229	0.308
$n = 600$	QR	Bias	−0.030	−0.060	−0.099
		Variance	0.129	0.145	0.170
	Norm heter	Bias	−0.566	−0.647	−0.564
		Variance	0.072	0.089	0.106
	QRCM	Bias	0.042	0.021	−0.189
		Variance	0.093	0.113	0.155
	QRCM$_c$	Bias	0.038	0.031	−0.168
		Variance	0.092	0.113	0.155

Notes: Bias and variance of predicted quantiles of order $\tau = \{0.90, 0.95, 0.975\}$ across the simulation. Predictions are obtained at $x_1 = x_2 = \ldots = x_5 = 1$. We compare the unstructured quantile regression (QR) model with our parametric approach (QRCM and QRCM$_c$) and a Normal heteroskedastic model in which both the mean and the log-scale are allowed to depend on covariates. The Normal model does not describe well the tails of the distribution and suffers from a significant bias.

to cold/heat waves. Climatological data are often seen as a composition of a seasonal trend, that repeats with every solar year, and a long-term trajectory. We formulated the following regression model, to be applied to both responses separately:

$$Q(\tau \mid t, s) = \beta_0(\tau) + \mathbf{g_t}(\mathbf{t})^{\mathrm{T}}\boldsymbol{\beta}_{\mathbf{t}}(\tau) + \mathbf{g_s}(\mathbf{s})^{\mathrm{T}}\boldsymbol{\beta}_{\mathbf{s}}(\tau) + (\mathbf{g_t}(\mathbf{t}) \otimes \mathbf{g_s}(\mathbf{s}))^{\mathrm{T}}\boldsymbol{\beta}_{\mathbf{t:s}}(\tau). \tag{7.1}$$

In the above regression equation, t is a progressive count $(1, 2, \ldots)$, expressed in days, while $s = (t \mod 365.2422)$ counts the days within a solar year. The long-term trend corresponds to the 'effect' of t, and is modelled by the basis of a restricted natural cubic spline, $\mathbf{g_t}(\mathbf{t})$, with one internal knot every 20 years. Seasonal variations are described by a periodic spline, $\mathbf{g_s}(\mathbf{s})$, with a period of one solar year and one internal knot every 2 solar months. Additionally, the regression equation includes the tensor product $\mathbf{g_t}(\mathbf{t}) \otimes \mathbf{g_s}(\mathbf{s})$ that defines an interaction term. For example, this allows the long-term trend of January to differ from that of July by more than just an additive constant.

This flexible modelling approach permits describing the conditional distribution of the data without making strong parametric assumptions, and is much more efficient than estimating sample quantiles separately in each week or month. In particular, regression modelling makes a better use of the available information, especially in presence of missing data, and allows for prediction and extrapolation. Moreover, it makes it possible to adjust for additional predictors of the response. Finally, regression splines are sufficiently ‘parametric’ to ignore local features of limited scientific interest, that can often be explained by random chance or measurement errors.

To fit model (7.1), we first applied standard QRto all stations above the Arctic Circle with at least 80 years of data. Crossing represented a very common problem, especially in the tails of the distribution, and often prevented using the fitted model to predict future trends. Consider, for example, station n. 78, that is located at 26 metres above sea level in Kandalaksha, Russia, with coordinates N 67° 09’ 00”, E 32° 21’ 00”. Data for this station were available from 1912 to 2018. Crossing was quantified using the summary measures defined in Section 6, and appeared unacceptably high, as shown in Table 4. We formulated a parametric version of (7.1) in which each of the regression coefficients, $\boldsymbol{\beta}(\tau) = \{\beta_0(\tau), \boldsymbol{\beta}_t(\tau), \boldsymbol{\beta}_s(\tau), \boldsymbol{\beta}_{s:t}(\tau)\}$, was described by a 5th-degree shifted Legendre polynomial (Abramovitz 1965), inclusive of an intercept. The resulting quantile function had 144 model parameters. As shown in Table 4, crossing was much less severe than with standard QR, but was not eliminated completely. Finally, we applied the constrained estimator described in Section 5. Results are illustrated in Figures 3 and 4, where we represent model-based quantiles of order $\tau = \{0.01, 0.10, 0.25, 0.50, 0.75, 0.90, 0.99\}$, using a variety of predictions to describe the long-term and seasonal trends of minimum and maximum temperatures. In Table 5, we report predicted quantiles of order $\tau = \{0.01, 0.50, 0.99\}$ in two different days, 31 January (winter) and 31 July (summer) for years 1935, 1955, 1975, and 1995 (the ‘past’), 2015 (the ‘present’) and 2035 (the ‘future’).

Table 4 Indicators of crossing in the ECAD data

		QR	QRCM	QRCM$_c$
Minimum temperatures	$100 \times P_{\text{cross}}$	45.03	2.94	0.00
	$100 \times L_{\text{cross}}$	2.69	0.23	0.00
Maximum temperatures	$100 \times P_{\text{cross}}$	34.56	0.32	0.00
	$100 \times L_{\text{cross}}$	2.56	0.03	0.00

Notes: Values of P_{cross} and L_{cross} (equations 6.1 and 6.2) are obtained when model (7.1) is applied to the minimum and maximum daily temperature in Kandalaksha station, Russian Federation. In the table, QR refers to standard quantile regression, while QRCM and QRCM$_c$ indicate unconstrained and constrained parametric estimators, respectively.

While the seasonal trends appear unchanged, the long-term trends show a cooling between the 1960s and the early 1990s (Peterson et al. 2008), followed by a warming effect. Between 1935 and 2015, the minimum temperature in January increased by 0.26 degrees at the 1st percentile, 1.65 degrees at the median, and 0.93 degrees at the 99th percentile. Similarly, the maximum temperature in July increased by 0.10, 1.68, and 1.12 degrees at the 1st, 50th, and 99th percentile, respectively. Would the current trends continue in the future, in 20 years we may expect an additional warming effect well above 1°C in both winter and summer.

For comparison, we also estimated a Normal heteroskedastic model in which both the mean and the log-variance were described by the spline-based predictor presented in equation (7.1). This model is much simpler than the proposed parametric QR, and

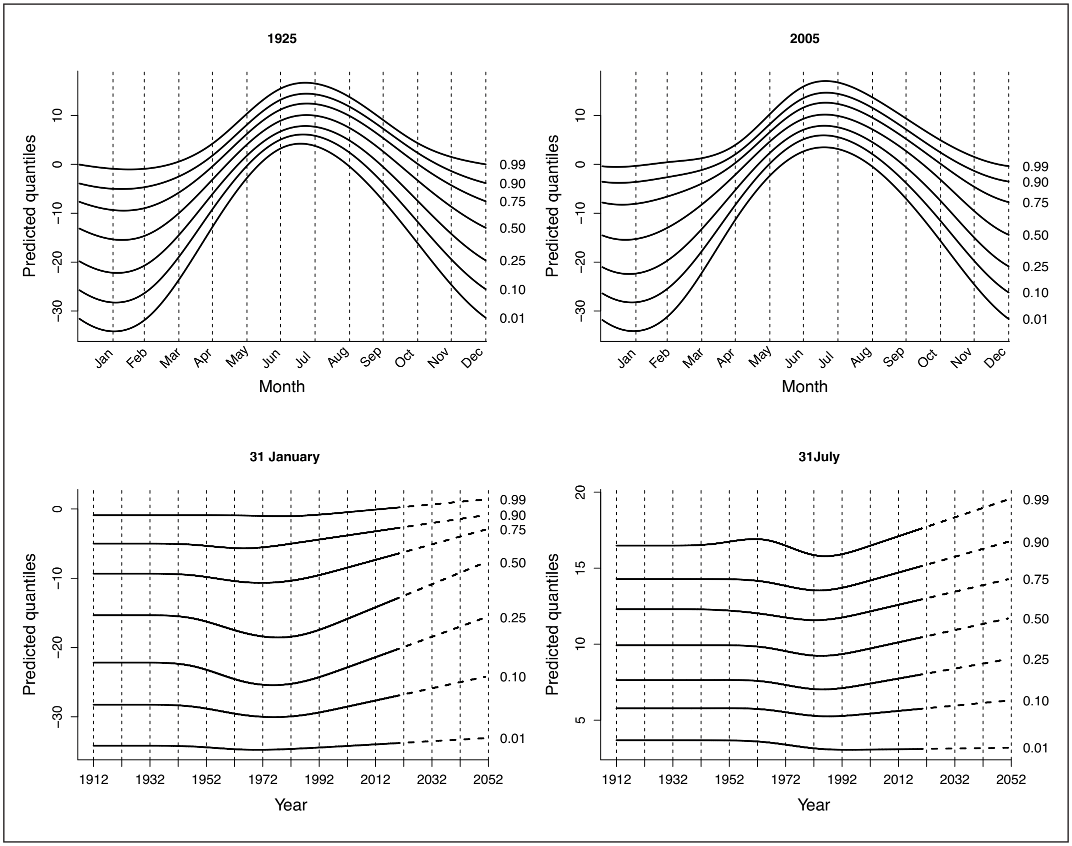

Figure 3 Estimated quantiles of order $\tau = \{0.01, 0.10, 0.25, 0.50, 0.75, 0.90, 0.99\}$ of the minimum daily temperature in Kandalaksha station, Russia. The top panels compare the seasonal trend in 1925 and that in 2005, while the bottom panels illustrate the long-term trends on two selected days, 31 January and 31 July. Dashed lines indicate extrapolation.

has only 24 parameters. However, it is less flexible and assumes that the conditional distribution is symmetric. Some important differences could be observed at the most extreme quantiles, for example, $\tau = 0.01$ and $\tau = 0.99$, where the Normal distribution did not appear to describe well the tail behaviour.

8 Concluding remarks

The described QRCM framework represents an appealing alternative to likelihood-based modelling of conditional distributions. At the same time, it allows overcoming a number of limitations of the standard, nonparametric QR estimators. Using a parametric model can significantly alleviate the crossing problem, and makes it particularly simple to enforce the non-crossing property by using a penalized

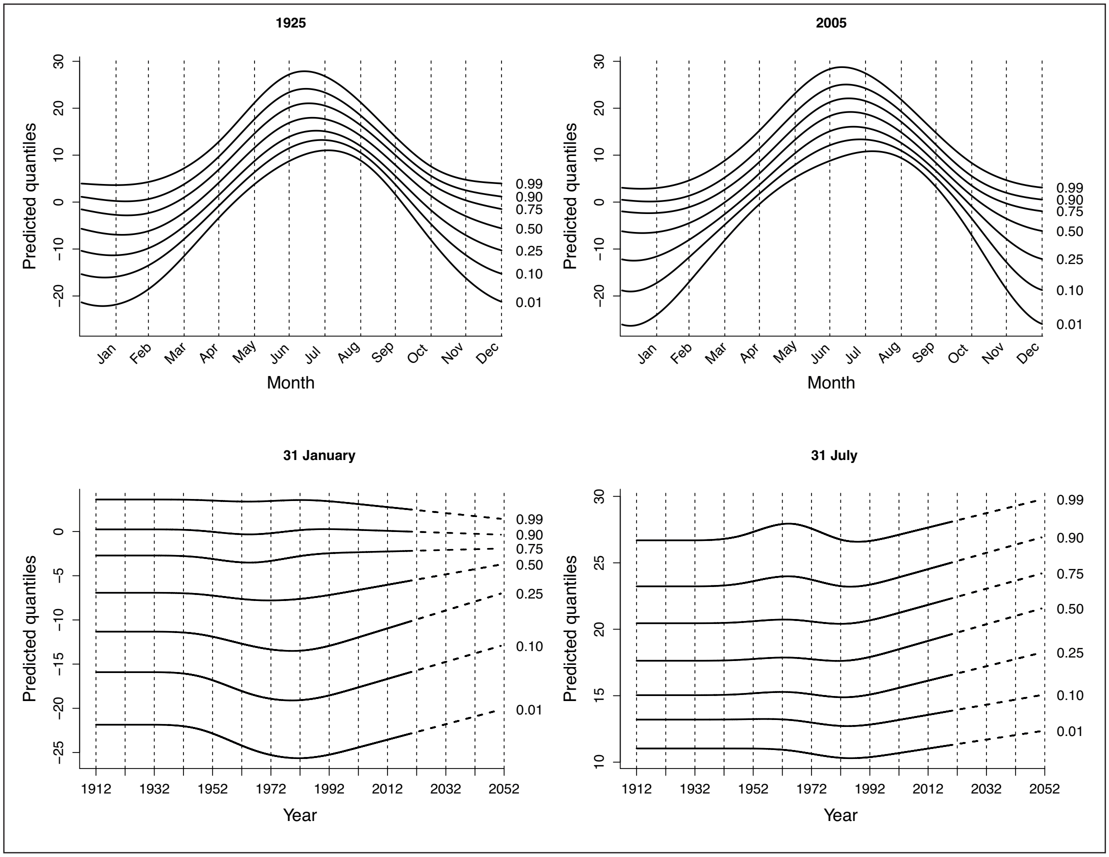

Figure 4 Estimated quantiles of the maximum temperature in Kandalaksha station, Russia

algorithm. The method described in this article can be generalized to censored and truncated data by replacing equation (3.7) with the estimating equation described in Frumento and Bottai (2017).

Considering that crossing often occurs in the tails of the distribution, the proposed method is particularly useful when the fitted model is used for prediction of extreme quantiles. The fact that standard QR estimators may not perform well in the tails is well studied in the literature (e.g., Portnoy and Jurečková 1999, Beirlant et al. 2004, Chernozhukov 2005, Wang et al. 2012, Wang and Li 2013; De Valk 2016). The presence of a parametric structure makes it very simple to extrapolate the extreme quantiles and to estimate their standard errors. The sheer fact of imposing some degree of smoothness generally improves efficiency, possibly at the cost of introducing some bias.

The R package `qrcm`, version 3.0, provides all the necessary procedures for estimation, inference, plotting, and prediction, and can be used with censored or

Table 5 Predicted quantiles

	Minimum temperature			Maximum temperature		
	$\tau = 0.01$	$\tau = 0.50$	$\tau = 0.99$	$\tau = 0.01$	$\tau = 0.50$	$\tau = 0.99$
31 Jan 1935	−34.17 (0.31)	−15.35 (0.16)	−0.91 (0.18)	−21.87 (0.28)	−6.93 (0.13)	3.62 (0.17)
1955	−34.49 (0.29)	−16.62 (0.19)	−0.91 (0.19)	−23.24 (0.33)	−7.36 (0.13)	3.46 (0.22)
1975	−34.74 (0.31)	−18.52 (0.22)	−1.03 (0.22)	−-25.47 (0.34)	−7.77 (0.16)	3.50 (0.31)
1995	−34.36 (0.32)	−17.00 (0.22)	−0.72 (0.20)	−25.02 (0.39)	−7.02 (0.16)	3.34 (0.22)
2015	−33.91 (0.56)	−13.70 (0.41)	0.02 (0.32)	−23.29 (0.60)	−5.85 (0.29)	2.66 (1.14)
2035	−33.44 (1.13)	−10.36 (0.85)	0.77 (0.66)	−21.54 (1.21)	−4.67 (0.63)	1.98 (2.42)
31 Jul 1935	3.68 (0.15)	9.93 (0.08)	16.49 (0.12)	11.03 (0.15)	17.63 (0.08)	26.70 (0.17)
1955	3.66 (0.19)	9.92 (0.08)	16.82 (0.13)	11.01 (0.19)	17.81 (0.08)	27.54 (0.20)
1975	3.31 (0.26)	9.43 (0.09)	16.26 (0.14)	10.52 (0.25)	17.69 (0.09)	27.31 (0.23)
1995	3.05 (0.20)	9.44 (0.08)	16.06 (0.14)	10.46 (0.20)	18.07 (0.09)	26.75 (0.22)
2015	3.10 (0.94)	10.24 (0.14)	17.28 (0.21)	11.13 (0.91)	19.31 (0.16)	27.82 (0.37)
2035	3.15 (2.00)	11.05 (0.30)	18.53 (0.43)	11.80 (1.97)	20.57 (0.34)	28.90 (0.76)

Notes: Estimated quantiles of the minimum and maximum daily temperature in Kandalaksha station, Russia. Predictions refer to 31 January (top table) and 31 July (bottom), and are computed for various years in the past, present, and future. Estimated standard errors in brackets.

truncated data. The package includes special functions that allow to diagnose and eliminate crossing.

Supplementary materials

Supplementary materials for this article, that is, R codes to reproduce the simulations and the analysis on climate change, the R package `qrcm` (version 3) and the ECAD data from station n. 78 (that can also be freely downloaded from https://www.ecad.eu/) are available from http://www.statmod.org/smij/archive.html

Acknowledgements

The authors are very grateful to Professor Erin Gabriel for her valuable contribution to the final version of this manuscript.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The authors received no financial support for the research, authorship and/or publication of this article.

References

Abramowitz M and Stegun IA (1964) Handbook of mathematical functions: With formulas, graphs, and mathematical tables. *Courier Corporation*, 55. URL http://www.pdmi.ras.ru/ lowdimma/BSD/abramowitz_and_stegun.pdf (last accessed 22 July 2021).

Andriyana Y, Gijbels I and Verhasselt A (2016) Quantile regression in varying-coefficient models: Non-crossing quantile curves and heteroscedasticity. *Statistical Papers*, 1–33. doi:10.1007/s00362-016-0847-7.

Baeck T, Fogel DB and Michalewicz Z (1997) *Handbook of Evolutionary Computation, 1st edition*. Boca Raton, FL: CRC Press, Section C 5.2.

Barrodale I and Roberts FDK (1973) An improved algorithm for discrete l1 linear approximation. *S.I.A.M. Journal on Numerical Analysis*, **10**, 839–48.

Beirlant J, Wet TD and Goegebeur Y (2004) Nonparametric estimation of extreme conditional quantiles. *Journal of Statistical Computation and Simulation*, **74**, 567–80.

Bondell HD, Reich BJ and Wang H (2010) Non-crossing quantile regression curve estimation. *Biometrika*, **97**, 825–38.

Bottai M and Cilluffo G (2020) Nonlinear parametric quantile models. *Statistical Methods in Medical Research*, **29**, 3757–69.

Buchinsky M (1998) Recent advances in quantile regression models: A practical guideline for empirical research. *Journal of Human Resources*, **33**, 88–126.

Cai Y (2010) Polynomial power-Pareto quantile function models. *Extremes*, **13**, 291–314.

Cai Y and Reeve DE (2013) Extreme value prediction via a quantile function model. *Coastal Engineering*, **77**, 91–98.

Chernozhukov V (2005) Extremal quantile regression. *The Annals of Statistics*, **33**, 806–39.

Chernozhukov V, Fernandez-Val I and Galichon A (2009) Improving point and interval estimators of monotone functions by rearrangement. *Biometrika*, **96**, 559–75.

——— (2010) Quantile and probability curves without crossing. *Econometrica*, **78**, 1093–1125.

Cook BL and Manning WG (2013) Thinking beyond the mean: A practical guide for using quantile regression methods for health services research. *Shanghai Archives of Psychiatry*, **25**, 55.

Cox PM, Betts RA, Jones CD, Spall SA and Totterdell IJ (2000) Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, **408**, 184.

Das P and Ghosal S (2017) Bayesian quantile regression using random B-spline series prior. *Computational Statistics and Data Analysis*, **109**, 121–43.

——— (2018) Bayesian non-parametric simultaneous quantile regression for complete and grid data. *Computational Statistics and Data Analysis*, **127**, 172–86.

De Valk C (2016) Approximation of high quantiles from intermediate quantiles. *Extremes*, **19**, 661–86.

Dette H and Volgushev S (2008) Non-crossing non-parametric estimates of quantile curves. *Journal of the Royal Statistical Society, Series B*, **70**, 609–27.

Do Nascimento F and Bourguignon M (2020) Bayesian time-varying quantile regression to extremes. *Environmetrics*, **31**, e2596.

Dong AXD, Chan JSK and Peters GW (2014) Risk margin quantile function via parametric and non-parametric Bayesian quantile regression. *ASTIN Bulletin*, **45**, 503–50.

El Adlouni S and Balde I (2019) Bayesian non-crossing quantile regression for regularly varying distributions. *Journal of Statistical Computation and Simulation*, **89**, 884–98.

Frumento P (2021) *qrcm: Quantile Regression Coefficients Modeling*. R package version 3.0. URL http://CRAN.R-project.org/package=qrcm (last accessed 22 July 2021).

Frumento P and Bottai M (2016) Parametric modeling of quantile regression coefficient functions. *Biometrics*, **72**, 74–84.

——— (2017) Parametric modeling of quantile regression coefficient functions with censored and truncated data. *Biometrics*, **73**, 1179–85.

Gilchrist W (2000) *Statistical Modeling with Quantile Functions*. London: Chapman & Hall.

Hanssen-Bauer I and Førland EJ (1998) Long-term trends in precipitation and temperature in the Norwegian Arctic: Can they be explained by changes in atmospheric circulation patterns? *Climate Research*, **10**, 143–53.

He X (1997) Quantile curves without crossing. *The American Statistician*, **51**, 186–92.

Hutson AD (2018) Quantile planes without crossing via nonlinear programming. *Computer Methods and Programs in Biomedicine*, **153**, 185–90.

Jiang R, Qian WM and Zhou ZG (2016) Weighted composite quantile regression for single-index models. *Journal of Multivariate Analysis*, **148**, 34–48.

Kai B, Li R and Zou H (2010) Local composite quantile regression smoothing: An efficient and safe alternative to local polynomial regression. *Journal of the Royal Statistical Society, Series B*, **72**, 49–69.

——— (2011) New efficient estimation and variable selection-methods for semiparametric varying-coefficient partially linear models. *Annals of Statistics*, **39**, 305–32.

Klein Tank AMG, Wijngaard JB, Können GP, Böhm R, Demarée G, Gocheva A, Mileta M, Pashiardis S, Hejkrlik L, Kern-Hansen C, Heino R, Bessemoulin P, MÃ¼ller-Westermeier G, Tzanakou M, Szalai S, PÃ¡lsdÃ³ttir T, Fitzgerald D, Rubin S, Capaldo M, Maugeri M, et al. (2002) Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *International Journal of Climatology*, **22**, 1441–53.

Koenker R (2004) Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, **91**, 74–89.

——— (2005) *Quantile Regression, Econometric Society Monograph Series*. Cambridge: Cambridge University Press.

Koenker R and Bassett G Jr (1978) Regression quantiles. *Econometrica*, 46, 33–50.

Koenker R and Machado JAF (1999) Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association*, **94**, 1296–1310.

Kuswanto H, Andari S and Permatasari EO (2015) Identification of extreme events in climate data from multiple sites. *Procedia Engineering*, **125**, 404–310.

Liu Y and Wu Y (2011) Simultaneous multiple non-crossing quantile regression estimation using kernel constraints. *Journal of Nonparametric Statistics*, **23**, 415–37.

Manton MJ, Della-Marta PM, Haylock MR, Manton, MJ, Della-Marta PM, Haylock, MR, Hennessy KJ, Nicholls N, Chambers LE, Collins DA, Daw G, Finet A, Gunawan D, Inape K, Isobe H, Kestin TS, Lefale P, Leyu CH, Lwin T, Maitrepierre L, Ouprasitwong N, Page CM, Pahalad J, et al. (2001) Trends in extreme daily rainfall and temperature in Southeast Asia and the South Pacific: 1961–1998. *International Journal of Climatology*, **21**, 269–84.

Medina-Ramon M, Zanobetti A, Cavanagh DP and Schwartz J (2006) Extreme temperatures and mortality: Assessing effect modification by personal characteristics and specific cause of death in a multi-city case-only analysis. *Environmental Health Perspectives*, **114**, 1331.

Meehl GA, Zwiers F, Evans J, Knutson T, Mearns L and Whetton P (2000) Trends in extreme weather and climate events: Issues related to modeling extremes in projections of future climate change. *Bulletin of the American Meteorological Society*, **81**, 427–36.

Muggeo VMR, Sciandra M, Tomasello A and Calvo S (2013) Estimating growth charts via nonparametric quantile regression: A practical framework with application in ecology. *Environmental and Ecological Statistics*, **20**, 519–31.

Neocleous T and Portnoy S (2008) On monotonicity of regression quantile functions. *Statistics and Probability Letters*, **78**, 1226–29.

Newey WK and McFadden D (1994) Large Sample Estimation and Hypothesis Testing. In *Handbook of Econometrics*, edited by RF Engle and DL McFadden. Vol. 4, Ch. 36, pages 2111–2245. Amsterdam: Elsevier.

Parzen E (1980) *Data modeling using quantile and density-quantile functions* (Technical Report No. B-2, , Project No. 4226). College Station, TX: Texas A&M Research Foundation.

Peterson TC, Connolley WM and Fleck J (2008). The mtyth of the 1970s global cooling scientific consensus. *American Meteorological Society*, **89**, 1325–37.

Portnoy S and Jureckova J (1999) On extreme regression quantiles. *Extremes*, **2**, 227–43.

Rahmstorf S and Coumou D (2002) Increase of extreme events in a warming world. *Proceedings of the National Academy of Sciences*, **108**, 17905.

Reich BJ (2012). Spatiotemporal quantile regression for detecting distributional changes in environmental processes. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **61**, 535–53.

Reich BJ and Smith LB (2013) Bayesian quantile regression for censored data. *Biometrics*, **69**, 651–60.

Richardson JT, Palmer MR, Liepins G and Hilliard M (1989) Some guidelines for genetic algorithms with penalty functions. *Proceedings of the Third International Conference on Genetic Algorithms*, 191–97. URL https://dl.acm.org/doi/abs/10.5555/645512.657233 (22 July 2021).

Rodrigues T, Dortet-Bernadet J.L., and Fan, Y. (2019). Simultaneous fitting of Bayesian penalised quantile splines. *Computational Statistics and Data Analysis*, **134**, 93–109.

Root TL, Price JT, Hall KR, Schneider SH, Rosenzweig C and Pounds JA (2003) Fingerprints of global warming on wild animals and plants. *Nature*, **421**, 57.

Russo S, Dosio A, Graversen RG, Sillmann J, Carrao H, Dunbar MB, Singleton A, Montagna P, Barbola P and Vogt JV (2014) Magnitude of extreme heat waves in present climate and their projection in a warming world. *Journal of Geophysical Research: Atmospheres*, **119**, 500–12.

Shaftel H (2016) *What's in a name? Weather, global warming and climate change.* NASA Climate Change: Vital Signs of the Planet.

Siedlecki W and Sklansky J (1989) Constrained genetic optimization via dynamic reward-penalty balancing and its use in pattern recognition. *Proceedings of the Third International Conference on Genetic Algorithms*, 141–150. URL https://doi.org/10.1142/9789814343138_0006 (22 July 2021).

Sottile G (2021) *qrcmNP: Nonlinear and Penalized Quantile Regression Coefficients Modeling.* R package version 0.2.0. URL https://CRAN.R-project.org/package=qrcmNP (last accessed 22 July 2021).

Sottile G, Frumento P, Chiodi M and Bottai M (2020) A penalized approach for parametric modeling of quantile regression coefficient functions. *Statistical Modelling*, **20**, 369–85.

Tokdar S and Kadane JB (2012) Simultaneous linear quantile regression: A semiparametric Bayesian approach. *Bayesian Analysis*, **7**, 51–72.

Vavrus SJ, Notaro M and Lorenz DJ (2015) Interpreting climate model projections of extreme weather events. *Weather and Climate Extremes*, **10**, 10–28.

Wang H, Li D and He X (2012) Estimation of high conditional quantiles for heavy-tailed distributions. *Journal of the American Statistical Association*, **107**, 1453–64.

Wang HJ and Li D (2013) Estimation of extreme conditional quantiles through power transformation. *Journal of the American Statistical Association*, **108**, 1062–74.

Wu Y and Liu Y (2009) Stepwise multiple quantile regression estimation using non-crossing constraints. *Statistics and its Interface*, **2**, 299–310.

Yang Y and Tokdar S (2017) Joint estimation of quantile planes over arbitrary predictor spaces. *Journal of the American Statistical Association*, **112**, 1107–20.

Zou H and Yuan M (2008) Composite quantile regression and the oracle model selection theory. *The Annals of Statistics*, **36**, 1108–26.