文章编号:1671-9352(2025)03-0100-07 DOI:10.6040/j.issn.1671-9352.0.2024.051

不完全观测的部分函数型线性分位数回归模型及应用

杨玉杰,凌能祥*
(合肥工业大学数学学院, 安徽 合肥 230009)

摘要:首先基于不完全观测的函数型数据,介绍部分函数型线性分位数回归模型及模型的估计方法和实施预测的步骤。其次,由于不完全观测的函数型变量广泛存在,利用尼泊尔 2019 年 4 月 8 日至 2020 年 8 月 31 日期间每 10 分钟一次的测风塔记录数据进行实证分析。针对风速数据的不完全函数型特征,构建以此为协变量、以日均气压为响应变量的不完全部分函数型线性分位数回归模型,获得模型未知斜率函数和未知参数的估计量,并且对日均气压进行预测分析,进一步说明模型及方法的有效性。

关键词:函数型数据分析;不完全观测;部分函数线性分位数回归;主成分分析
中图分类号:O212.7 **文献标志码**:A
引用格式:杨玉杰,凌能祥. 不完全观测的部分函数型线性分位数回归模型及应用[J]. 山东大学学报(理学版),2025,60(3):100-106.

Partially functional linear quantile regression model and its application for incomplete observations

YANG Yujie, LING Nengxiang*
(School of Mathematics, Hefei University of Technology, Hefei 230009, Anhui, China)

Abstract: Firstly, based on incomplete observed functional data, partially functional linear quantile regression model, the estimation method and prediction step are introduced. Secondly, due to the widespread existence of incomplete observed functional variables, the mast records every 10 minutes from April 8, 2019 to August 31, 2020 in Nepal for empirical analysis are used. Aiming at the incomplete functional characteristics of wind speed data, an incomplete partial functional linear quantile regression model with daily mean air pressure as the response variable is constructed, and the estimators of the unknown slope function and unknown parameters of the model are obtained, and the daily mean air pressure is predicted and analyzed, which further illustrates the effectiveness of the model and the method.

Key words: functional data analysis; incomplete observation; partial linear quantile regression; principal component analysis

0 引言

随着现代科学技术的发展以及测量设备仪器的进步,收集到的数据更多呈现出函数的形式,这类数据被称为函数型数据。函数型数据分析(functional data analysis, FDA)在过去几十年也得到了快速发展,广泛应用于经济、金融、环境、医学、化学计量学等领域。更多有关函数型数据分析的内容可以参考文献[1-2]。作为函数型数据分析的重要组成部分,函数型线性回归模型已经受到很多学者的关注,并被广泛研究和应用,主要集中于构造模型的估计量和研究估计量的大样本性质[3-6]。

事实上,上述函数型线性模型的估计方法都集中于均值回归,容易受异常点的影响。Koenker 等[7]在 1978 年首次提出了分位数回归,它提供了比均值更全面的响应变量的分布描述,并且具有较好的稳健性。

收稿日期:2024-02-24;**网络出版时间**:2024-08-05 10:49:34
基金项目:国家自然科学基金资助项目(72071068)
第一作者:杨玉杰(1998—),女,硕士研究生,研究方向为概率论与数理统计. E-mail:yangyujie225@163.com
*通信作者:凌能祥(1964—),男,教授,博士生导师,研究方向为非参数统计、函数型数据分析. E-mail:hfut.lnx@163.com

Cardot 等[8]利用 B-样条的线性组合和比例惩罚项构造了函数线性分位数回归模型中的估计量,并研究了估计量的渐近性质。Kato[9]研究了函数线性分位数回归的估计,并建立了估计量的收敛速率。杜艳芳[10]通过分位数回归方法对甘肃省三大城市的空气质量指数进行评价和预测,为相关部门政策的出台提供意见和建议。孙鸣茜[11]针对自变量和响应变量均为函数型数据的分位数回归模型进行了研究,将此模型应用于基于温度曲线的降水量预测问题中,验证了模型优越性。Tang 等[12]研究了部分函数线性分位数回归,得到了斜率函数的收敛速度和常数斜率函数的渐近正态性。

一方面,注意到在数据收集过程中,由于一些设备损坏或不可控原因,因此会出现收集到不完全观测的函数数据,如心率曲线收集时患者因身体不适摘下设备,收集环境数据时设备发生故障导致数据收集不完整等。对于这类型数据的研究也引起了越来越多学者的关注:Kraus[13]提出了不完全函数型数据的修复方法,通过函数主成分分析(functional principal component analysis, FPCA),对函数曲线缺失部分进行最佳线性预测;王楚[14]研究了数值解释变量缺失、响应变量具有函数特征的部分函数型线性回归模型,并将模型应用于家电耗能、车流量等数据分析,说明了模型的有效性;Kneip 等[15]提出了一种新的基于重构算子恢复不完全观测函数型解释变量的缺失部分;Xiao 等[16]将文献[13]的方法应用在解释变量是部分观测的函数线性分位数回归模型中;Crambes 等[17]考虑了一个函数协变量部分观察且响应变量缺失的函数线性回归模型。然而,在许多实际问题的建模中,解释变量常常包括数值向量和函数型解释变量,这类数据被称为混合数据场景。另一方面,气压是大气科学研究中的重要指标,可以预测天气的变化趋势,并提供准确的预报服务;同时,气压在农业生产中也是一项不可缺少的监测指标,科研人员通过气压数值及时预测气象灾害,对农田及相关设施提前做好防护,降低损失。本文针对尼泊尔测风塔的不完全观测函数型数据,构建不完全观测的部分函数型线性分位数回归模型,研究风速和相对湿度对气压的影响。该模型同时考虑函数型变量和数值标量对响应变量的效应,并利用分位数回归增加了模型的稳健性,不同分位点的回归也使对响应变量的描述更加全面。实证分析结果也进一步验证了模型及方法的有效性,具有较好的实际意义。

1　模型及预测

1.1　模型介绍

对分位水平 $\tau\in(0,1)$,考虑如下部分函数型线性分位数回归模型:

$$Y=\boldsymbol{b}_\tau^{\mathrm{T}}\boldsymbol{Z}+\int_a^b a_\tau(t)X(t)\,\mathrm{d}t+\varepsilon_\tau,\tag{1}$$

其中 Y 是一个实值的随机变量,$\boldsymbol{Z}=(Z_1,Z_2,\cdots,Z_d)^{\mathrm{T}}$ 是一个 d 维的协变量,$\{X(t),t\in[a,b]\}$ 是一个均值为 0 的随机函数。不失一般性,假设 $[a,b]=[0,1]$,即 $X(t)\in L^2([0,1])$,$L^2([0,1])$是$[0,1]$上平方可积的随机函数构成的希尔伯特空间,具有内积$\langle x(t),y(t)\rangle=\int_0^1 x(t)y(t)\,\mathrm{d}t$,其范数为 $\|x(t)\|=\langle x(t),x(t)\rangle^{\frac{1}{2}}$。$a_\tau(t)$为$[0,1]$上平方可积的斜率函数,$\boldsymbol{b}_\tau=(b_{\tau1},b_{\tau2},\cdots,b_{\tau d})^{\mathrm{T}}$ 为 d 维的未知参数,ε_τ 独立于$\{X(\cdot),\boldsymbol{Z}\}$并且 $P(\varepsilon_\tau\leqslant0)=\tau$。

1.2　模型估计

假设$\{Y_i,Z_i,X_i(t),i=1,2,\cdots,n\}$是由模型(1)生成的一组独立样本,定义 $X(\cdot)$的协方差函数为 $k_X(s,t)=\mathrm{Cov}(X(s),X(t))$,协方差算子为 $K_X f(s)=\int_0^1 k_X(s,t)f(t)\,\mathrm{d}t$,其中 $\forall f(\cdot)\in L^2([0,1])$。根据 Mercers定理[18],对协方差函数谱分解

$$k_X(s,t)=\sum_{j=1}^{\infty}\lambda_j v_j(t)v_j(s),$$

其中,$\lambda_1>\lambda_2>\cdots>0$ 为协方差算子 K_X 的特征值,$\{v_j(\cdot)\}_{j=1}^{\infty}$为相应的标准正交特征函数。

首先,本文基于函数曲线是部分观测的,即 $X_i(\cdot),i=1,2,\cdots,n$ 在 $O_i\subset[0,1]$上观测到,在区间 $M_i=[0,1]\backslash O_i$ 观测不到。类似于文献[13,16],将曲线的观测部分记为 $X_{iO_i}(\cdot)$,曲线的缺失部分记为 $X_{iM_i}(\cdot)$,并且数据是完全随机缺失的,即观测区间独立于函数曲线 $X_1(\cdot),X_2(\cdot),\cdots,X_n(\cdot)$。但在实际中,每条函数曲线 $X_i(\cdot)$在离散网格点 $0\leqslant t_{i1}\leqslant t_{i2}\leqslant\cdots\leqslant t_{i,L_i+1}\leqslant1$ 观测到。缺失区间 M_i 由以下几个部分组成:第一次观测时间

之前和最后一次观测时间之后,以及 2 个连续的观测时间之间的所有大于一定阈值 Δ_n 的间隙,其中 $\Delta_n=\max\limits_{1<i<n}\max\limits_{2<l<L_i}\min(t_{il}-t_{i,l-1},t_{i,l+1}-t_{il})$,且 $\Delta_n\to 0$, $n\to\infty$,因此,曲线 $X_{iO_i}(\cdot)$ 的估计可以用插值函数表示为

$$\hat{X}_{iO_i}(t)=\sum_{j=1}^{L_i}X_i(t_{ij})I(t\in O_i\cap[t_{ij},t_{i,j+1}))\,,\quad i=1,2,\cdots,n。$$

基于文献[13],协方差函数的估计为

$$\hat{k}_X(s,t)=\frac{I(s,t)}{\sum\limits_{i=1}^{n}U_i(s,t)}\sum_{i=1}^{n}U_i(s,t)\hat{X}_{iO_i}(s)\hat{X}_{iO_i}(t)=\sum_{j=1}^{\infty}\hat{\lambda}_j\hat{v}_j(t)\hat{v}_j(s)\,,$$

其中 $I(s,t)=I(\sum\limits_{i=1}^{n}U_i(s,t)>0)$, $U_i(s,t)=O_i(s)O_i(t)$, $O_i(s)$ 为 $I(s\in O_i)$。协方差算子 K_X 的估计值用 $\hat{K}_X$ 表示,以及估计的标准正交特征函数和特征值为 $\{\hat{v}_j(\cdot)\}_{j=1}^{\infty}$ 和 $\hat{\lambda}_1\geqslant\hat{\lambda}_2\geqslant\cdots\geqslant 0$。

更进一步地,由 Karhunen-Loève 表示定理[18]有

$$X_i(t)=\sum_{j=1}^{\infty}\xi_{ij}v_j(t)\,,\quad a_\tau(t)=\sum_{j=1}^{\infty}a_{\tau j}v_j(t)\,,\tag{2}$$

其中 $\xi_{ij}=\langle X_i(\cdot),v_j(\cdot)\rangle$ 是均值为 0 的不相关变量并且 $E\xi_{ij}^2=\lambda_j$, $a_{\tau j}=\langle a_\tau(t),v_j(t)\rangle$。在主成分分析中,主成分得分的估计至关重要。当函数数据完全观测时,ξ_{ij} 可以用 $\langle X_i(\cdot),\hat{v}_j(\cdot)\rangle$ 估计。当函数不完全观测时,利用 Kraus[13] 提出的估计方法,ξ_{ij} 的估计表示为 $\hat{\xi}_{ij}^{(\alpha)}=\hat{\xi}_{ijO_i}+\hat{\xi}_{ijM_i}^{(\alpha)}$, $\hat{\xi}_{ijO_i}=\langle\hat{X}_{iO_i}(\cdot),\hat{v}_{jO_i}(\cdot)\rangle$, $\hat{\xi}_{ijM_i}^{(\alpha)}=\langle\hat{e}_{ij}^{(\alpha)},\hat{X}_{iO_i}\rangle$,其中 $\alpha>0$ 为正则化参数,更多详细内容见文献[13],其中的正则化参数 α 通过广义交叉验证方法进行选择。具体来说,通过最小化交叉验证分数 $\mathrm{gcv}(\alpha)$ 来选择第 i 个函数的第 j 个分数的 α,

$$\mathrm{gcv}(\alpha)=\frac{\sum\limits_{k\in N}(\langle\hat{X}_{kM_i},\hat{v}_{jM_i}\rangle-\langle\hat{e}_{ij}^{(\alpha)},\hat{X}_{kO_i}\rangle)^2}{\{1-(1/|N|)\mathrm{d}f_i(\alpha)\}^2}\,,$$

其中 N 为完全观测到的函数曲线的数目,并且 $\mathrm{d}f_i(\alpha)=\sum\limits_{k=1}^{\infty}\dfrac{\hat{\lambda}_{O_iO_ik}}{\hat{\lambda}_{O_iO_ik}+\alpha}$。

最后,将式(2)代入式(1),模型为

$$Y_i=\boldsymbol{b}_\tau^{\mathrm{T}}\boldsymbol{Z}_i+\int_0^1 a_\tau(t)X_i(t)\,\mathrm{d}t+\varepsilon_{\tau i}\,,\quad i=1,2,\cdots,n。\tag{3}$$

分位数估计 $\hat{\boldsymbol{b}}_\tau=(\hat{b}_{\tau 1},\hat{b}_{\tau 2},\cdots,\hat{b}_{\tau d})^T$ 和 $\hat{a}_\tau(t)=\sum\limits_{j=1}^{m}\hat{a}_{\tau j}\hat{v}_j(t)$ 可以通过解以下最小化问题获得

$$\min_{b_1,b_2,\cdots,b_d;a_1,a_2,\cdots,a_m}\sum_{i=1}^{n}\rho_\tau\Big(Y_i-\boldsymbol{b}^{\mathrm{T}}\boldsymbol{Z}_i-\sum_{j=1}^{m}a_j\hat{\xi}_{ij}^{(\alpha)}\Big)\,,$$

其中:$\rho_\tau(u)=u(\tau-I_{(u<0)})$ 为分位数损失函数;$m=m_n$ 为截断参数。采用估计特征值的累积百分比(cumulative percentage of variance, CPV)进行选取

$$\mathrm{CPV}(m)=\sum_{j=1}^{m}\hat{\lambda}_j\Big/\sum_{j=1}^{\infty}\hat{\lambda}_j。$$

1.3 预测的主要步骤

步骤 1 通过插值函数将离散观测到的函数拟合为 $\hat{X}_{iO_i}(\cdot)$,对估计的协方差核进行谱分解 $\hat{k}_X(s,t)=\sum\limits_{j=1}^{\infty}\hat{\lambda}_j\hat{v}_j(t)\hat{v}_j(s)$;

步骤 2 选取正则化参数 α 和截断参数 m;

步骤 3 求主分数估计量 $\hat{\xi}_{ij}^{(\alpha)}=\hat{\xi}_{ijO_i}+\hat{\xi}_{ijM_i}^{(\alpha)}$,其中 $\hat{\xi}_{ijO_i}=\langle\hat{X}_{iO_i}(\cdot),\hat{v}_{jO_i}(\cdot)\rangle$, $\hat{\xi}_{ijM_i}^{(\alpha)}=\langle\hat{e}_{ij}^{(\alpha)},\hat{X}_{iO_i}\rangle$;

步骤 4 求最小值问题 $\min\limits_{b_1,b_2,\cdots,b_d;a_1,a_2,\cdots,a_m}\sum\limits_{i=1}^{n}\rho_\tau\Big(Y_i-\boldsymbol{b}^{\mathrm{T}}\boldsymbol{Z}_i-\sum\limits_{j=1}^{m}a_j\hat{\xi}_{ij}^{(\alpha)}\Big)$,得到估计量 $\hat{a}_\tau(t)$ 和 $\hat{\boldsymbol{b}}_\tau$;

步骤 5 给定随机变量 $\boldsymbol{Z}$ 和随机函数 $X(t)$ 时,响应变量 $\boldsymbol{Y}$ 在分位数 τ 下的预测值为

$$\hat{\boldsymbol{Y}}=\hat{\boldsymbol{b}}_\tau^{\mathrm{T}}\boldsymbol{Z}+\int_0^1\hat{a}_\tau(t)X(t)\,\mathrm{d}t。$$

2　实证分析

在本文中,根据收集到的来自尼泊尔测风塔的风力测量数据集进行实证分析研究,数据来源为 https://energydata.info/dataset/nepal-wind-measurement-data。此数据集中包含尼泊尔 10 个测风塔的风力测量数据,每个数据集都包含风速、风向、气压、相对湿度和温度的每日报告,并且数据以每 10 min 一次的频率记录。本文感兴趣的是风速和相对湿度对气压的影响,其中风速是一个连续变化的过程,可看作函数型数据。进一步地,10 个测风塔之一的 Tangbe 测风塔(2019 年 4 月 8 日—2020 年 8 月 31 日)在高度为 80 m 的风速计处测得的风速数据存在大量不完全观测,因此以此站点的数据集为样本建立模型

$$Y_i=\boldsymbol{b}_\tau\boldsymbol{Z}_i+\int_0^{144}a_\tau(t)X_i(t)\,\mathrm{d}t+\varepsilon_{\tau i},\quad i=1,2,\cdots,285, \tag{4}$$

其中 Y_i 为每日气压的平均值,$\boldsymbol{Z}_i$ 为每日相对湿度的平均值,$X_i(t)$ 为每 10 min 记录一次的风速。风速曲线分为以下几类:完全观测、部分观测、完全观测不到。为了对模型进行估计,删除风速完全观测不到的样本(剩余 335 个样本),并在风速完全观测到的样本中随机选取 50 个作为测试集,剩余的 285 个样本作为训练集建立模型(4)。为了减少测试集和训练集划分的影响,将上述随机选择进行 $L=100$ 次。图 1 展示了气压和相对湿度的数据,随机挑选一些不完全观测的风速曲线展示在图 2。

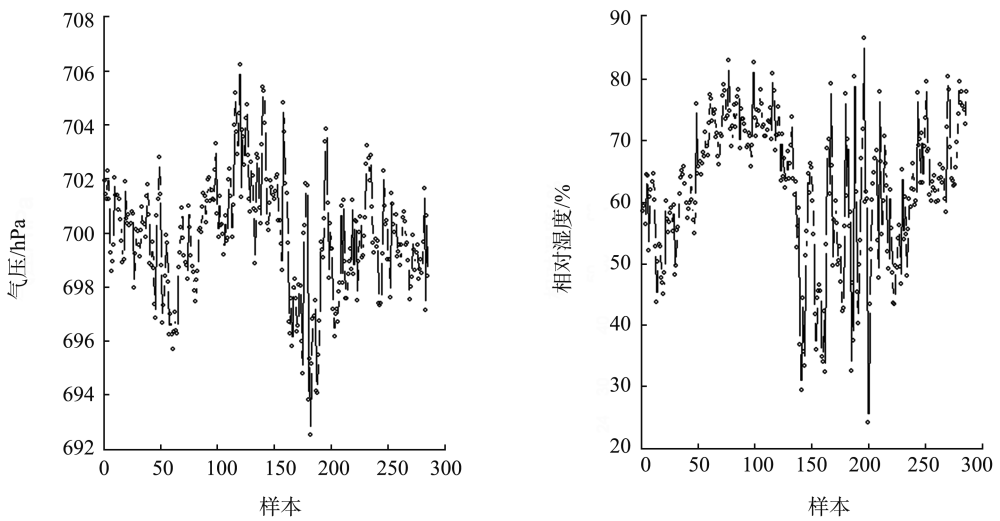


图 1　气压和相对湿度数据散点图

Fig.1　Scatter plots of barometric pressure data and relative humidity data

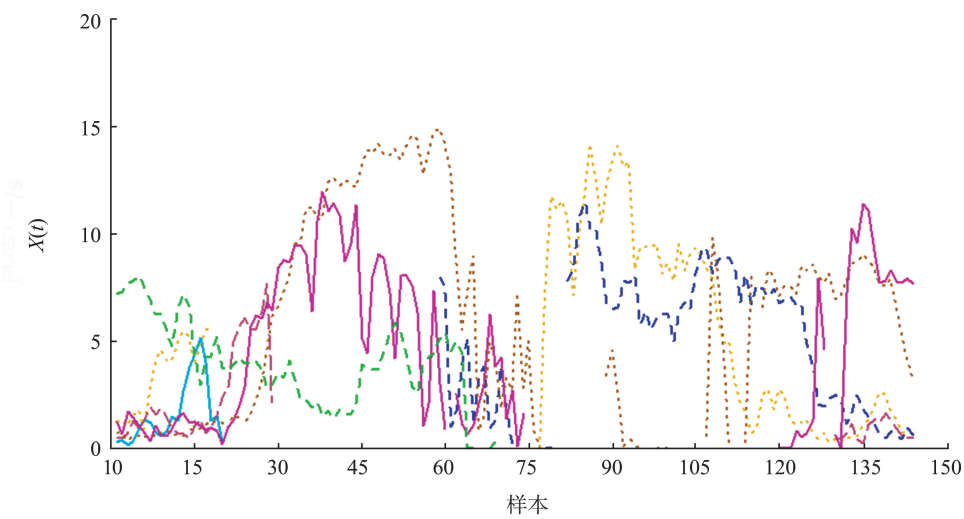


图 2　不完全观测的风速曲线图

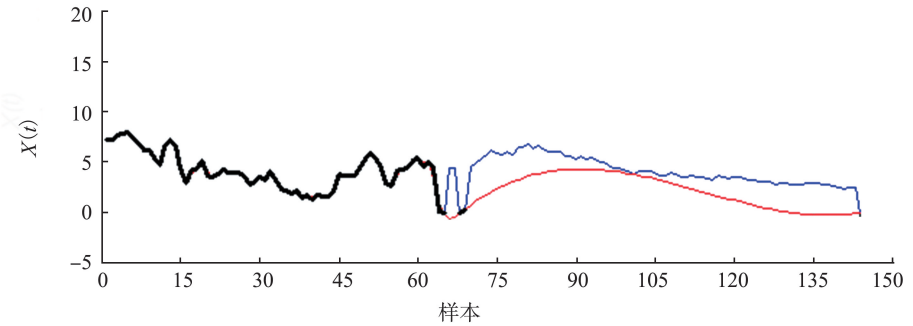
Fig.2　Wind speed graph with incomplete observations

首先,对模型(4)中心化处理得到

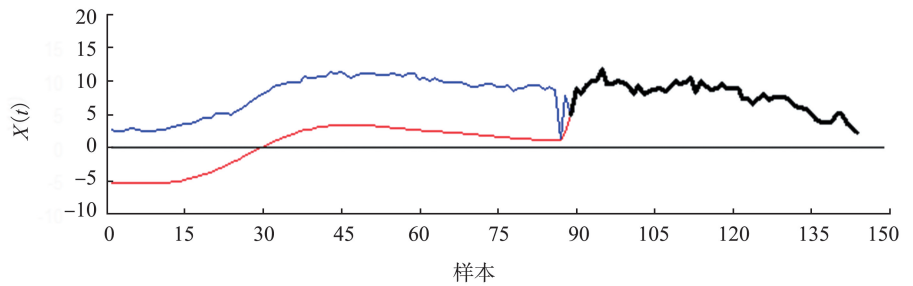
$$Y_i=a+\boldsymbol{b}_\tau\boldsymbol{Z}_i+\int_0^{144}a_\tau(t)\tilde{X}_i(t)\,\mathrm{d}t+\varepsilon_{\tau i},\quad i=1,2,\cdots,285。\tag{5}$$

其中 $\tilde{X}_i(t)=X_i(t)-E[X_i(t)]$, $a=\int_0^{144}a_\tau(t)E[X_i(t)]\,\mathrm{d}t$。

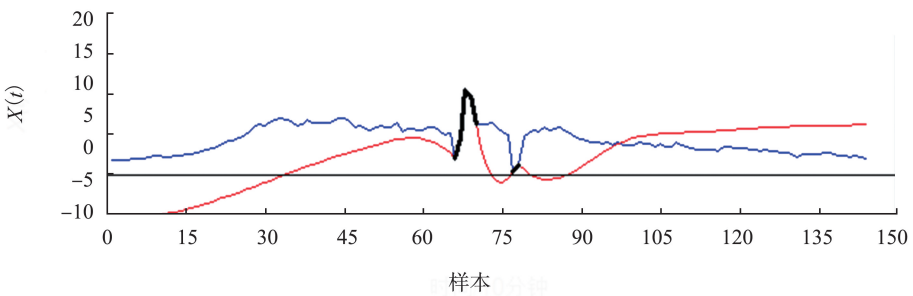
其次,对部分观测的风速曲线进行恢复。如 1.2 节所述,使用 Kraus[13] 提出的正则化回归方法来重构函数曲线。根据重构的函数曲线,采用 1.2 节中给出的方法对未知参数 b_τ 和斜率函数 $a_\tau(t)$ 进行分位数估计(此过程记为 QRP),同时并为了比较同时采用 Kneip 等[15] 提出的方法对曲线进行重构并进行分位数回归(此过程记为 QRO)。图 3 展示了几条观测率不同的曲线的重构情况,其中黑色为原始观测到的曲线,红色为文献[15]方法的重构情况,蓝色为文献[13]方法的重构情况。结合图 2、3 可以看出,对于风速的非负曲线,文献[15]的方法重构的曲线会出现负值的异常情况。从总体来看,本文使用文献[13]的方法重构的曲线与观测到的曲线一样更为尖锐,而文献[15]的方法重构的曲线更为光滑。



(a) 曲线观测率为47.2%



(b) 曲线观测率为39.6%



(c) 曲线观测率为4.9%

注:黑色为原始观测到的曲线,红色为文献[15]方法的重构情况,蓝色为文献[13]方法的重构情况。

图 3 风速曲线的重构图

Fig.3 Reconstruction of the wind speed curves

最后,定义相对预测误差(relative prediction error, RPE)来检验方法的优越性

$$\mathrm{RPE}=\frac{1}{N}\sum_{i=1}^{N}\frac{(Y_i^*-\hat{Y}_i^*)^2}{\mathrm{Var}\,X(t)_{Y^*|Z^*,X^*}},$$

其中 $\{Y_i^*,\boldsymbol{Z}_i^*,X_i^*,\ i=1,2,\cdots,N\}$ 为测试集,$N=50$, $\hat{Y}_i^*=\hat{\boldsymbol{b}}_\tau^{\mathrm{T}}\boldsymbol{Z}_i^*+\int_0^{144}\hat{a}_\tau(t)X_i^*(t)\,\mathrm{d}t$, $\mathrm{Var}_{Y^*|Z^*,X^*}$ 是 Y^* 的条件方

差。RPE 的 $L=100$ 次结果的平均值展示在表 1 中。另外,针对重复 100 次的结果,图 4 绘制了 2 种方法相对预测误差的箱线图。结合表 1 及图 4 可以看出,本文提出的 QRP 方法在极端分位点处更加稳健,具有更好的表现。

表 1　不同分位点的 RPE

Table 1　RPE at different quantiles

方法	$\tau=0.15$	$\tau=0.25$	$\tau=0.50$	$\tau=0.75$	$\tau=0.85$
QRP	2.310	2.056	1.076	1.345	1.366
QRO	2.678	1.458	0.996	1.652	1.787

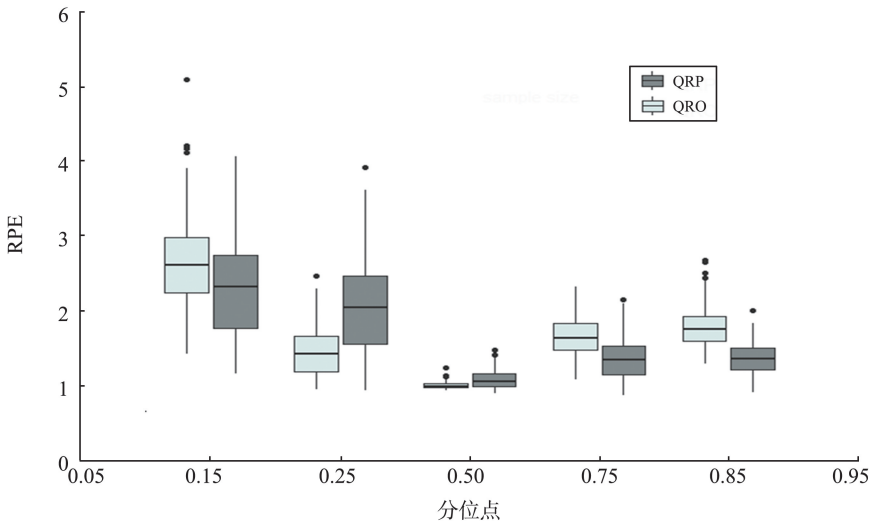


图 4　2 种方法 RPE 结果的箱线图

Fig.4　Boxplot of the RPE results of the two methods

同时,图 5 展示了斜率函数 $a_\tau(t)$ 在不同分位数水平 $\tau=0.50,0.75,0.85$ 的估计曲线。从中可以发现,不同分位数对斜率函数之间存在显著差异,表明分位数回归对响应变量进行了更加全面地描述。

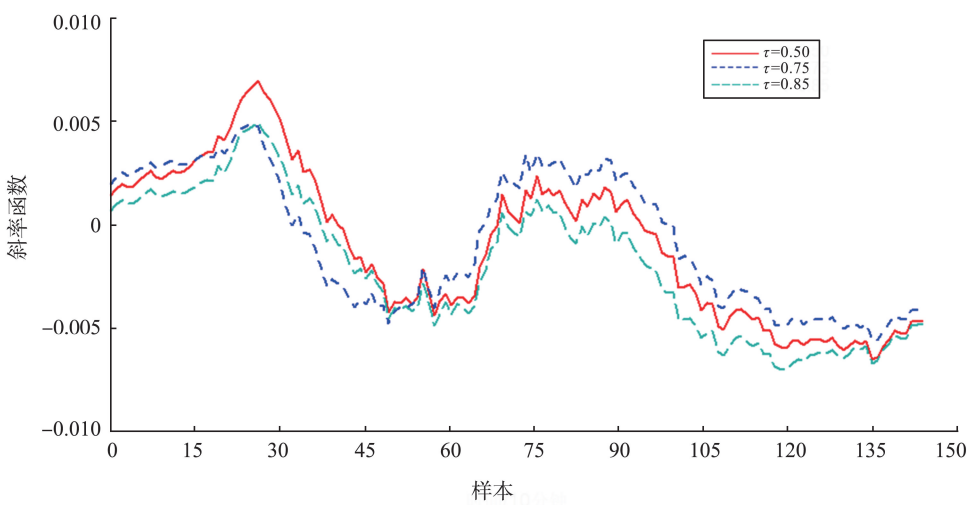


图 5　在分位数 $\tau=0.50,0.75,0.85$ 下估计的斜率曲线图

Fig.5　The estimated curves of the slope function at $\tau=0.50,0.75,0.85$

3　结语

本文利用不完全观测的部分函数型线性分位数回归模型创新地研究风速和相对湿度对气压的影响。与文献[16]相比,本文模型可以建立响应变量与数值协变量、函数型协变量之间的联系,测风塔数据的实证分析进一步说明了模型及方法的有效性。

参考文献:

[1] RAMSAY J, SILVERMAN B. Functional data analysis[M]. New York: Springer, 2005.

[2] FERRATY F, VIEU P. Nonparametric functional data analysis:theory and practice[M]. New York: Springer, 2006.

[3] HALL P, HOROWITZ J L. Methodology and convergence rates for functional linear regression[J]. The Annals of Statistics, 2007, 35:70-91.

[4] CAI T T, HALL P. Prediction in functional linear regression[J]. The Annals of Statistics, 2006, 34:2159-2179.

[5] ZHANG D W, LIN X H, SOWERS M. Two-stage functional mixed models for evaluating the effect of longitudinal covariate profiles on a scalar outcome[J]. Biometrics, 2007, 63(2):351-362.

[6] SHIN H. Partial functional linear regression[J]. Journal of Statistical Planning and Inference, 2009, 139:3405-3418.

[7] KOENKER R, BASSETT G. Regression quantiles[J]. Econometrica, 1978, 46:33-51.

[8] CARDOT H, CRAMBES C, SARDA P. Quantile regression when the covariates are functions[J]. Journal of Nonparametric Statistics, 2005, 17(7):841-856.

[9] KATO K. Estimation in functional linear quantile regression[J]. The Annals of Statistics, 2012, 40:3108-3136.

[10] 杜艳芳. 基于分位数回归的空气质量指数分析[D]. 兰州:兰州大学,2016.
DU Yanfang. The analysis of air quality index based on quantile regression[D]. Lanzhou: Lanzhou University, 2016.

[11] 孙鸣茜. 函数型数据分位数回归模型及其应用[D]. 武汉:华中科技大学,2019.
SUN Mingxi. Quantile regression of functional data and its application[D]. Wuhan: Huazhong University of Science and Technology, 2019.

[12] TANG Qingguo, CHENG Longsheng. Partial functional linear quantile regression[J]. Science China Mathematics, 2014, 57(12):2589-2608.

[13] KRAUS D. Components and completion of partially observed functional data[J]. Journal of the Royal Statistical Society, Series B: Statistical Methodology, 2015, 77(4):777-801.

[14] 王楚. 带有缺失数据的函数型响应部分函数线性回归模型的估计与应用[D]. 南京:南京理工大学,2021.
WANG Chu. Estimation andapplication of partial functional linear regression model for function response with missing data [D]. Nanjing: Nanjing University of Science and Technology, 2021.

[15] KNEIP A, LIEBL D. On the optimal reconstruction of partially observed functional data[J]. The Annals of Statistics, 2020, 4:1692-1717.

[16] XIAO Juxia, XIE Tianfa, ZHANG Zhongzhan. Estimation in partially observed functional linear quantile regression[J]. Journal of Systems Science and Complexity, 2020, 35:313-341.

[17] CRAMBES C, DAAYEB C, GANNOUN A, et al. Functional linear model with partially observed covariate and missing values in the response[J]. Journal of Nonparametric Statistics, 2023, 35(1):172-197.

[18] BOSQ D. Linear processes in function spaces: theory and applications[M]. New York: Springer Science and Business Media, 2000.

(编辑:李艺)