

doi:10.6041/j.issn.1000-1298.2017.07.021

# 基于 VTCI 和分位数回归模型的冬小麦单产估测方法

王蕾<sup>1,2</sup> 王鹏新<sup>1,2</sup> 李俐<sup>1,2</sup> 张树誉<sup>3</sup>

(1. 中国农业大学信息与电气工程学院, 北京 100083; 2. 农业部农业灾害遥感重点实验室, 北京 100083;

3. 陕西省气象局, 西安 710014)

**摘要:** 条件植被温度指数(VTCI)是一种综合了归一化植被指数(NDVI)与地表温度(LST)的遥感干旱监测方法,在关中平原的近实时干旱监测中具有其适用性。分位数回归能全面反映因变量的条件分布在不同分位数处的特征,回归结果稳健可靠。为了进一步研究 VTCI 干旱监测结果与小麦单产之间的关系及提高冬小麦单产估测精度,构建了不同分位数  $\tau(0.1, 0.3, 0.5, 0.7, 0.9)$  下关中平原各市 2008—2014 年的冬小麦主要生育期 VTCI 与单产之间的线性回归模型,并基于中位数( $\tau=0.5$ )回归模型对研究区域的冬小麦单产进行了估测。结果表明,分位数回归模型比较全面地反映了不同分位数下冬小麦单产分布与 VTCI 之间的相关程度,弥补了最小二乘估产模型回归结果单一、易受异常值影响等的不足。中位数回归模型的单产估测结果与实际单产之间的相对误差和均方根误差的最小值及平均值均低于最小二乘回归模型,估测精度较高。此外,中位数单产估测模型获取的冬小麦估产结果在年际变化规律与空间分布特征上与实际产量均较相符,说明分位数回归在研究 VTCI 与产量之间的关系及冬小麦单产估测中具有其适用性与可靠性。

**关键词:** 冬小麦; 分位数回归; 条件植被温度指数; 遥感; 估产

**中图分类号:** S127; TP79      **文献标识码:** A      **文章编号:** 1000-1298(2017)07-0167-07

## Winter Wheat Yield Estimation Method Based on Quantile Regression Model and Remotely Sensed Vegetation Temperature Condition Index

WANG Lei<sup>1,2</sup> WANG Pengxin<sup>1,2</sup> LI Li<sup>1,2</sup> ZHANG Shuyu<sup>3</sup>

(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

2. Key Laboratory of Remote Sensing for Agri-Hazards, Ministry of Agriculture, Beijing 100083, China

3. Shaanxi Provincial Meteorological Bureau, Xi'an 710014, China)

**Abstract:** Vegetation temperature condition index (VTCI) combines normalized difference vegetation index (NDVI) and land surface temperature (LST), and is applicable to a more accurate monitoring of droughts in Guanzhong Plain, Shaanxi Province, China. Quantile regression is a tool for comprehensively reflecting the conditional distribution characters under different quantiles, and its regression results are steady and reliable. In order to achieve a better correlation between winter wheat yield and the weighted VTCI as well as a higher yield estimation accuracy, linear regression models between the weighted VTCI and yields in the cities of Guanzhong Plain in the years from 2008 to 2014 were analyzed by using the quantile regression whose quantiles were set to be 0.1, 0.3, 0.5, 0.7 and 0.9, respectively. These quantile regression results roundly reflected the distribution of the yields under different drought conditions and were beneficial supplement of the linear regression from which the single fitted line and impressionable results from outliers were obtained. The wheat yield estimation model based on the median regression (quantile equalled to 0.5) was used to monitor the wheat yields in the cities of Guanzhong Plain from 2008 to 2014, the average and minimum values of the relative errors and the root mean square errors (RMSE) between the estimated yields and the actual yields were all lower than those derived from the ordinary least square method. Additionally, the characteristics of inter-annual evolution and spatial distribution of the estimated yields using the median regression model were in good agreement with the

收稿日期: 2016-11-21 修回日期: 2016-12-14

基金项目: 国家自然科学基金项目(41371390)

作者简介: 王蕾(1988—),女,博士生,主要从事定量遥感及其在干旱预测中的应用研究,E-mail: 409118258@qq.com

通信作者: 王鹏新(1965—),男,教授,博士生导师,主要从事定量遥感及其在农业中的应用研究,E-mail: wangpx@cau.edu.cn

actual situation, which indicated that the quantile regression was feasible and reliable in the research of winter wheat yield estimation and the relationship between yield and drought.

**Key words:** winter wheat; quantile regression; vegetation temperature condition index; remote sensing; yield estimation

## 引言

作物估产信息是国民经济宏观调控的重要信息,及时、准确、大范围地获取作物估产信息对农业经济发展与粮食政策的制定具有重要的现实意义<sup>[1-3]</sup>。遥感技术可充分利用地物表面的光谱、时间、空间和方向信息,具有准确客观、宏观快速的特点,成为大面积农业资源调查、作物估产、精准农业中应用最为广泛的技术手段之一<sup>[4]</sup>。常用的基于遥感信息的作物产量估测方法主要有机理模型法、经验模型法及半机理半经验模型法等<sup>[2]</sup>。其中,经验模型法通过分析遥感数据与作物长势指标之间的线性关系及作物长势指标与产量之间的相关性实现作物产量的间接估测,是一种简单易行的估产方法<sup>[5]</sup>;机理模型法参数与过程复杂,目前多利用遥感数据与作物模型的耦合进行作物产量的估测与预测<sup>[6]</sup>;半经验半机理模型中以光能利用效率模型应用最广泛,但存在一些关键参数难以定量模拟的问题<sup>[7]</sup>。

在全球变暖的背景下,农业干旱已经成为影响作物生长发育与产量形成的主要因素<sup>[8-9]</sup>。目前,通过构建遥感干旱指数与作物单产之间的关系可初步实现农业干旱对粮食安全影响的定量分析<sup>[10-11]</sup>。常见的遥感干旱监测指数主要有基于能量平衡、微波、植被状态和热红外的遥感监测方法等<sup>[12]</sup>。其中,条件植被温度指数(Vegetation temperature condition index, VTCI)是基于归一化植被指数(NDVI)和地表温度(LST)的散点图呈三角形区域分布的空间特征提出的一种干旱监测方法<sup>[13-14]</sup>,适用于监测一特定年内某一时期区域级的干旱程度,可通过构建VTCI与作物长势指标之间的定量关系进行干旱影响评估、作物产量估测及预测等研究<sup>[15-17]</sup>。

在以往的作物估产研究中,大多应用最小二乘法(Ordinary least squares, OLS)的经典线性回归进行模型构建与数据分析,以及遥感干旱监测指数与作物产量之间的关系研究<sup>[18-20]</sup>。然而,基于经典最小二乘法的线性回归模型只能基于因变量与自变量在均值水平上的相关关系得到一条回归直线,挖掘的信息量有限。此外,利用最小二乘法进行回归分析时对随机误差的分布特征要求严格,对于一些实

际问题,最小二乘法的线性假设可能过于苛刻,难以得到无偏、有效的参数估计值<sup>[21]</sup>。为了弥补经典最小二乘法在回归分析中的缺陷, KOENKER 和 BASSETT 于 1978 年提出了分位数回归(Quantile regression, QR)的理论<sup>[22]</sup>。分位数回归利用因变量的条件分位数建模,较之经典的最小二乘法应用条件更加宽松,挖掘信息更加丰富,可以更精确地描述自变量对于因变量的变化范围以及条件分布形状的影响,在一定程度上比最小二乘的回归结果更加稳健<sup>[21,23]</sup>。这一方法最初主要应用于经济领域,随着计算机的普及被迅速应用于医学、教育、社会、环境科学等多个领域<sup>[24-26]</sup>,但在遥感干旱影响评估与作物单产估测方面的应用还未见报道。本文基于分位数回归理论构建关中平原冬小麦主要生育期的条件植被温度指数与冬小麦单产之间的分位数回归模型,研究不同分位数下的冬小麦单产受干旱的影响程度,并通过对比最小二乘法估产模型与分位数估产模型的精度,验证分位数回归模型在干旱影响评估与冬小麦单产估测中的适用性。

## 1 材料与方 法

### 1.1 研究区域

研究区域为位于陕西省中部的关中平原,总面积  $5.55 \times 10^4 \text{ km}^2$ ,东西长约 300 km,西窄东宽,地势西高东低,土壤肥沃,垦植指数达 70% 以上,主要包括渭南市、西安市、咸阳市、宝鸡市、铜川市 5 个地级市及杨凌农业高新技术产业示范区。该区域地处季风区的边缘,气候温暖,年平均气温  $6 \sim 13^\circ\text{C}$ ,年均降水量  $550 \sim 700 \text{ mm}$ ,降水量偏少,且降水分布西部优于东部,在空间分布上不均匀,年际变化与年内变化均较大<sup>[27]</sup>。尤其自 20 世纪 90 年代以来,关中平原整体呈现暖干化趋势,春旱、伏旱显著,关中平原的农作物在全生育期呈现不同程度的水分不足现象,进而造成该区域的粮食减产,干旱已成为该地区农业增产增收的主要制约因素<sup>[17]</sup>。

### 1.2 数据处理

VTCI 已被证实是一种近实时的干旱监测方法,其定义为<sup>[13-14]</sup>

$$V = \frac{L_{N_i, \max} - L_{N_i}}{L_{N_i, \max} - L_{N_i, \min}} \quad (1)$$

其中

$$L_{N_i, \max} = a + bN_i \quad (2)$$

$$L_{N_i, \min} = a' + b'N_i \quad (3)$$

式中  $N_i$ ——某一像素的 NDVI

$L_{N_i}$ ——某一像素的 NDVI 值为  $N_i$  时的地表温度

$L_{N_i, \max}$ ——NDVI 为  $N_i$  时研究区域内所有像素地表温度的最大值

$L_{N_i, \min}$ ——NDVI 为  $N_i$  时研究区域内所有像素地表温度的最小值

$a, b, a', b'$ ——待定系数, 由研究区域的 NDVI 和 LST 散点图近似获得

采用关中平原 2008—2014 年 3—5 月份的 Aqua—MODIS 的日地表温度产品 (MYD11A1) 和日地表反射率产品 (MYD09GA) 反演得到日 LST 和日 NDVI 产品, 应用最大值合成技术, 分别生成旬 NDVI 和 LST 最大值合成产品, 通过对多年某一旬的 LST 最大值合成产品逐像素取最小值, 生成多年的旬 LST 最大-最小值合成产品。经上述方法确定冷、热边界后, 根据 VTCI 的计算公式 (式(1)), 生成以旬为单位的 VTCI 时间序列数据。依据关中平原冬小麦的生长情况, 将冬小麦越冬后的主要生育期划分为返青期 (3 月上旬—3 月中旬)、拔节期 (3 月下旬—4 月中旬)、抽穗-灌浆期 (4 月下旬—5 月上旬) 和乳熟期 (5 月中旬—5 月下旬) 4 个生育时期<sup>[19]</sup>, 取每个生育时期内多旬 VTCI 的均值作为该生育时期的 VTCI 值, 计算出关中平原各市 2008—2014 年各生育时期 VTCI。然后运用基于熵值法与因子权重排序法的归一组合赋权法<sup>[28]</sup> 确定关中平原五市 (除铜川市) 的冬小麦各个生育时期的权重, 进而得到关中平原冬小麦每年的加权 VTCI, 即冬小麦的主要生育期 VTCI。

关中平原各市 2008—2014 年的冬小麦单产数据主要来源于研究年份的陕西省统计年鉴。

### 1.3 分位数回归

#### 1.3.1 分位数回归理论

分位数回归方法是将回归方法与条件分位数进行结合, 通过最小化离差绝对值的加权和, 依据因变量的条件分位数对自变量进行回归, 进而得到所有分位数下的回归模型<sup>[23]</sup>。其定义为: 设随机变量  $Y$  的一组随机变量为  $\{y_1, y_2, \dots, y_n\}$ ,  $Y$  的分布函数为  $F(y) = P(Y \leq y)$ , 则对于任意的  $\tau (0 < \tau < 1)$ ,  $Y$  的第  $\tau$  分位数函数为

$$Q(\tau) = \inf \{y: F(y) \geq \tau\} \quad (4)$$

对于任意的  $\tau (0 < \tau < 1)$ , 定义损失函数  $\rho_\tau(z)$  为

$$\rho_\tau(z) = z(\tau - I(z < 0)) = \begin{cases} \tau z & (z \geq 0) \\ (\tau - 1)z & (z < 0) \end{cases} \quad (5)$$

式中  $I(\cdot)$ ——示性函数

由式(5)可知, 损失函数为分段函数 (图 1), 且  $\rho_\tau(z) \geq 0$ 。

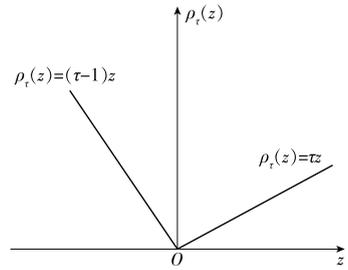


图 1 分位数回归的损失函数  $\rho_\tau(z)$

Fig. 1 Loss function of quantile regression

为方便积分, 将  $\rho_\tau(z)$  表示为

$$\rho_\tau(z) = \tau z I(z \geq 0) - (1 - \tau) z I(z < 0) \quad (6)$$

当  $z$  取  $y - \hat{y}$  时, 则有

$$\rho_\tau(y - \hat{y}) = \tau(y - \hat{y}) I(y - \hat{y} \geq 0) + (\tau - 1)(y - \hat{y}) I(y - \hat{y} < 0) \quad (7)$$

则  $Y$  的分位数回归就是通过最小化期望损失求  $\hat{y}$ , 即找到满足

$$\min_{\hat{y} \in \mathbf{R}} E[\tau(y - \hat{y}) I(y - \hat{y} \geq 0) + (\tau - 1)(y - \hat{y}) I(y - \hat{y} < 0)] \quad (8)$$

的  $\hat{y}$  即为  $Y$  的  $\tau$  分位数回归。

#### 1.3.2 线性分位数回归

线性分位数回归模型的方程为

$$Y = \mathbf{X}^T \boldsymbol{\beta} + \varepsilon \quad (9)$$

式中  $Y$ ——因变量  $\boldsymbol{\beta}$ ——未知参数

$\varepsilon$ ——误差项

对于一般的线性条件分位数函数, 为了分析自变量  $\mathbf{X}$  对于因变量  $Y$  在其各分位数  $\tau$  上的影响, 需要求解满足

$$\min_{\boldsymbol{\beta} \in \mathbf{R}} \left( \sum_{i=1}^n y_i - \mathbf{x}_i^T \boldsymbol{\beta}(\tau) \right) \quad (10)$$

的  $\boldsymbol{\beta}(\tau)$ 。目前对式(10)的算法主要包括单纯形算法、内点算法与平滑算法等<sup>[23]</sup>。进而得到  $Y$  的  $\tau (0 < \tau < 1)$  分位数函数为

$$Q_y(\tau | x) = \mathbf{x}^T \boldsymbol{\beta}(\tau) \quad (11)$$

式中  $\boldsymbol{\beta}(\tau)$ ——式(10)中极小化问题的解

由此, 在不同的  $\tau (0 < \tau < 1)$  下就能得到不同的分位数函数, 即所有  $y$  在  $x$  上的条件分布的一簇直线。

## 2 结果与分析

### 2.1 分位数回归模型的构建

为了研究干旱对不同地区的小麦产量分布产生的影响, 本文对关中平原四市 (除铜川市) 2008—2014 年的冬小麦主要生育期 VTCI 与冬小麦单产的 28 组数据进行分析, 以小麦单产作为因变量 ( $Y$ ), 以主要生育期 VTCI 的干旱监测结果作为自变量 ( $X$ ),

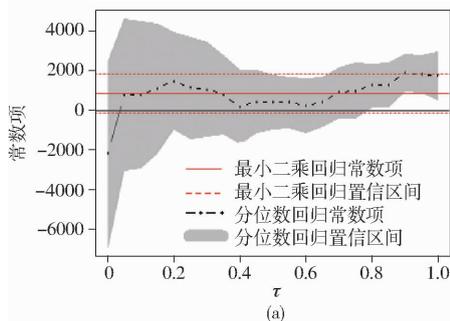
建立 VTCI 与小麦单产在不同分位数  $\tau$  (0.1, 0.3, 0.5, 0.7, 0.9) 下的线性分位数回归模型, 通过参数估计算法得到常数项与  $X$  系数 ( $\beta(\tau)$ ) (表 1)。由表 1 可知, 不同分位数下的常数项与  $X$  系数的估计值均落在了置信区间内, 且以中位数 ( $\tau = 0.5$ ) 回归与高分位数 ( $\tau = 0.7$  与  $\tau = 0.9$ ) 回归的置信区间较小, 说明中位数与高分位数参数估计的置信水平高于低分位数处, 回归模型的拟合优度较好。

表 1 分位数回归的系数

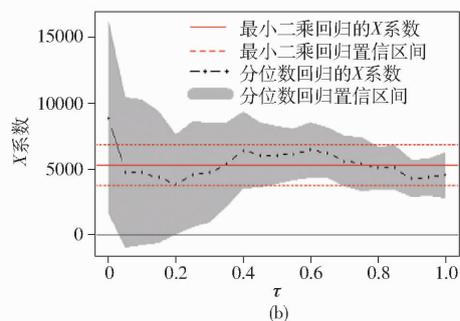
Tab. 1 Coefficients of quantile regression

$\tau$	常数项	$X$ 系数
0.1	787.9 (-6 288.6, 1 468.7)	4 779.1 (3 677.7, 14 538.7)
0.3	1 016.8 (-2 409.9, 1 931.8)	4 768.0 (3 061.1, 10 169.7)
0.5	435.9 (-609.0, 1 847.0)	6 094.5 (3 605.5, 7 703.6)
0.7	882.6 (202.0, 2 374.0)	5 623.1 (4 277.3, 6 638.4)
0.9	1 885.0 (1 187.0, 2 185.6)	4 285.7 (3 888.6, 5 319.5)

以 0.05 为步长, 估计各个分位数 (0.05, 0.10, 0.15, ..., 0.95) 下的常数项与  $X$  系数, 及其对应的置信区间, 得到了分位数回归在不同分位数下的系



数 (图 2)。由图 2 可知, 常数项的估计值 (图 2a) 主要介于 (400, 2 000) 之间, 在最小二乘回归常数项估计值的附近上下波动, 并随着分位数的增加呈现出先降低后增加的变化趋势, 置信区间在低分位数较大, 在高分位处逐渐缩小。自变量  $X$  系数 (图 2b) 在低分位数处 ( $\tau < 0.4$ ) 随着分位数的增加呈现先降低后显著增加的趋势, 并在大约  $\tau = 0.35$  处, 其取值与最小二乘线性回归相同; 在  $\tau \in (0.4, 0.6)$  时,  $X$  系数的估计值较大且发展趋势平稳; 当  $\tau > 0.6$  时, 随着分位数的增加,  $X$  系数的估计值缓慢下降, 且  $\tau = 0.75$  处的  $X$  系数估计值与最小二乘法相同; 其置信区间的变化规律与常数项类似, 低分位数处较大, 随着分位数的增加逐渐缩小。上述分析表明, 最小二乘法的线性回归使用简单的线性回归参数估计值对所有的数据进行估算, 只能得到 VTCI 干旱监测结果与冬小麦单产的条件分布在均值位置上的一条拟合直线, 这样势必会造成产量估测结果的偏高或偏低。而分位数回归在每个分位数处的系数估计值及其置信区间不尽相同, 说明在冬小麦减产、正常或丰产的年份或地区, VTCI 值与冬小麦产量之间的线性关系不同, 由此可依据研究年份间关中平原各市冬小麦的生长与干旱状况, 构建不同分位数下的小麦估产模型, 得到更加切合实际情况的冬小麦单产估测结果。

图 2 不同分位数下的常数项及  $X$  系数Fig. 2 Coefficients of intercept and  $X$  under different quantiles

不同分位数下的回归直线 (图 3) 中, 最小二乘回归、中位数回归 ( $\tau = 0.5$ ) 以及  $\tau$  为 0.9、0.7、0.3、0.1 下的分位数回归直线的变化趋势相同, 均揭示了冬小麦单产在旱情严重时偏低, 而旱情较轻或无旱时冬小麦单产较高的规律。但不同分位数下回归直线的斜率却不尽相同, 中位数回归时的斜率最大, 高于中高分位 ( $\tau = 0.7, 0.9$ ) 与中低分位 ( $\tau = 0.3, 0.1$ ) 时的直线斜率, 说明在小麦产量正常的年份或地区, 小麦产量与干旱之间的线性关系最为密切; 中位数回归与最小二乘法的回归直线的位置不同, 说明条件密度的不对称性, 且最小二乘回归受到异常值 (VTCI 值较大时产量却偏低的点) 的影响, 回归

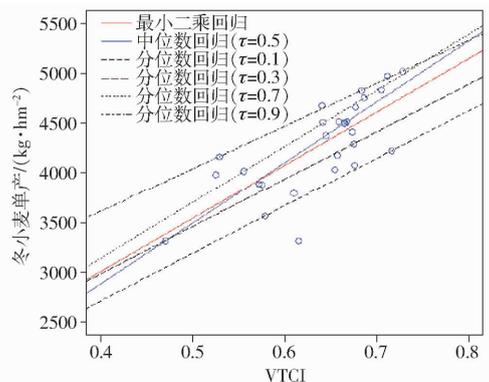


图 3 不同分位数下的拟合线

Fig. 3 Fitted lines under different quantiles

的稳健度受到削弱,基于最小二乘法的估产结果会在无旱情发生(VTCI大于0.57)时偏低,而在有旱情发生(VTCI小于0.57)时略微偏高,这样会导致有旱灾发生的年份或区域的冬小麦产量被高估,而无旱年份或区域的冬小麦产量被低估;分位数回归可弥补最小二乘回归的局限性,基于不同的小麦单产分布条件,构建了不同的单产估测模型,比经典的最小二乘回归反映更多的局部信息,从而实现对于旱与冬小麦产量之间关系的深入分析及小麦估产精度的提高。

## 2.2 分位数回归模型的精度评价

分位数回归模型中,0.5分位点处的中位数回归模型可以较好地解决最小二乘法中某些“离群值”影响回归显著性的问题,是一种较为稳健的线性回归模型<sup>[23]</sup>。同时,由于0.5分位点处于因变量的中间位置,在对所有的数据进行拟合时较为适宜。故本文运用分位数回归模型中的0.5分位点进行的研究区域2008—2014年的冬小麦单产估测,并结合实际单产对线性分位数估产模型的估测精度进行评价。在 $\tau=0.5$ 处构建的分位数回归模型为

$$Y = 6\,094.5X + 435.9 \quad (12)$$

中位数回归结果 $t$ 检验的 $P$ 值小于0.001,达到极显著水平。类似地,令变量 $Y_{lm}$ 表示小麦单产,变量 $X_{lm}$ 表示VTCI的干旱监测结果,用最小二乘法求得回归方程,建立的最小二乘回归模型为

$$Y_{lm} = 5\,349.0X_{lm} + 854.7 \quad (13)$$

结果表明,最小二乘回归结果的 $F$ 显著性检验的 $P$ 值小于0.001,达到极显著水平。

为了检验上述两种回归模型对小麦单产估测精度的相对优劣,分别应用最小二乘法及分位数回归模型计算得到关中平原四市2008—2014年的冬小麦单产估测值。统计得到了两模型的冬小麦估测单产与实际单产之间的相对误差和均方根误差的最大值、最小值及平均值(表2)。由表2可知,应用分位数回归模型获取的2008—2014年的冬小麦单产估测结果与实际单产之间相对误差的最小值(0)与平均值(5.6%)均低于最小二乘回归,分位数回归的相对误差最大值(26.7%)受异常值的影响略高于最小二乘的最大值(25.5%);类似地,分位数回归的均方根误差最大值为623.0 kg/hm<sup>2</sup>,高于最小二乘的最大值(594.9 kg/hm<sup>2</sup>),最小值仅为0,低于最小二乘回归的最小值(25.0 kg/hm<sup>2</sup>),分位数回归的平均值为160.3 kg/hm<sup>2</sup>,低于最小二乘的平均值(172.1 kg/hm<sup>2</sup>)。上述结果表明,分位数回归的单产估测结果的相对误差与均方根误差整体上低于最小二乘回归,基于分位数回归的估产精度优于最小

二乘法,故分位数回归模型可用于VTCI与冬小麦单产之间的相关关系及冬小麦产量估测等的研究。

表2 最小二乘与分位数回归的估测单产与实际单产之间的相对误差与均方根误差

Tab.2 Relative errors and root mean square errors between estimated and actual yields using the least squares and quantile regressions

误差模型	相对误差/%			均方根误差/(kg·hm <sup>-2</sup> )		
	最大值	最小值	平均值	最大值	最小值	平均值
最小二乘回归	25.5	0.9	5.9	594.9	25.0	172.1
分位数回归	26.7	0	5.6	623.0	0	160.3

## 2.3 基于分位数回归的冬小麦单产估测

基于 $\tau=0.5$ 处的分位数回归模型估测关中平原2008—2014年的冬小麦单产(图4)。从冬小麦单产的空间分布特征来看,冬小麦单产以中部最高,西部次之,东部最低;从年际变化规律来看,关中平原2008—2014年关中平原冬小麦的单产呈现个别年份波动,总体增长的趋势,2013年关中平原冬小麦较之前几年减产严重,这是由于2013年关中平原降水较少,旱情偏重,造成冬小麦生育期内的水分不足,进而形成一定程度的减产。由此,基于分位数回归模型的冬小麦单产估测结果与实际情况相符,模型的估测精度较高。

## 3 讨论

分位数回归较之经典的最小二乘回归具有其独特的优势,原因在于:最小二乘回归系数的估计为最佳线性无偏估计,对误差的分布要求严格,且对于存在重尾或离群值的数据,其稳健性会被削弱;此外,最小二乘回归模型的回归结果为基于自变量与因变量在均值水平上的相关关系得到的一条拟合直线,所反映的信息不够丰富。分位数回归采用加权的最小一乘回归法,在某些情况下比经典的最小二乘回归法更加稳健,且分位数回归可以得到不同分位数下因变量在自变量上的条件分布的轨迹,有效突出局部之间的相关关系,故可以作为经典最小二乘法线性回归的有益补充。本文中,分位数回归的方法能够针对不同的产量分布条件建立各个分位点的回归模型,分位数回归结果不仅包含VTCI干旱监测数据与冬小麦单产中心位置分布的相关关系,还可以度量干旱与冬小麦单产在各个不同位置(产量偏高或偏低)下的线性相关程度,因而,分位数回归模型对不同分位点上干旱对冬小麦单产作用差异的刻画更加全面,为干旱影响评估方面的研究提供了大量宝贵信息,这比最小二乘法表现出的平均水平所包含的信息量更有价值。

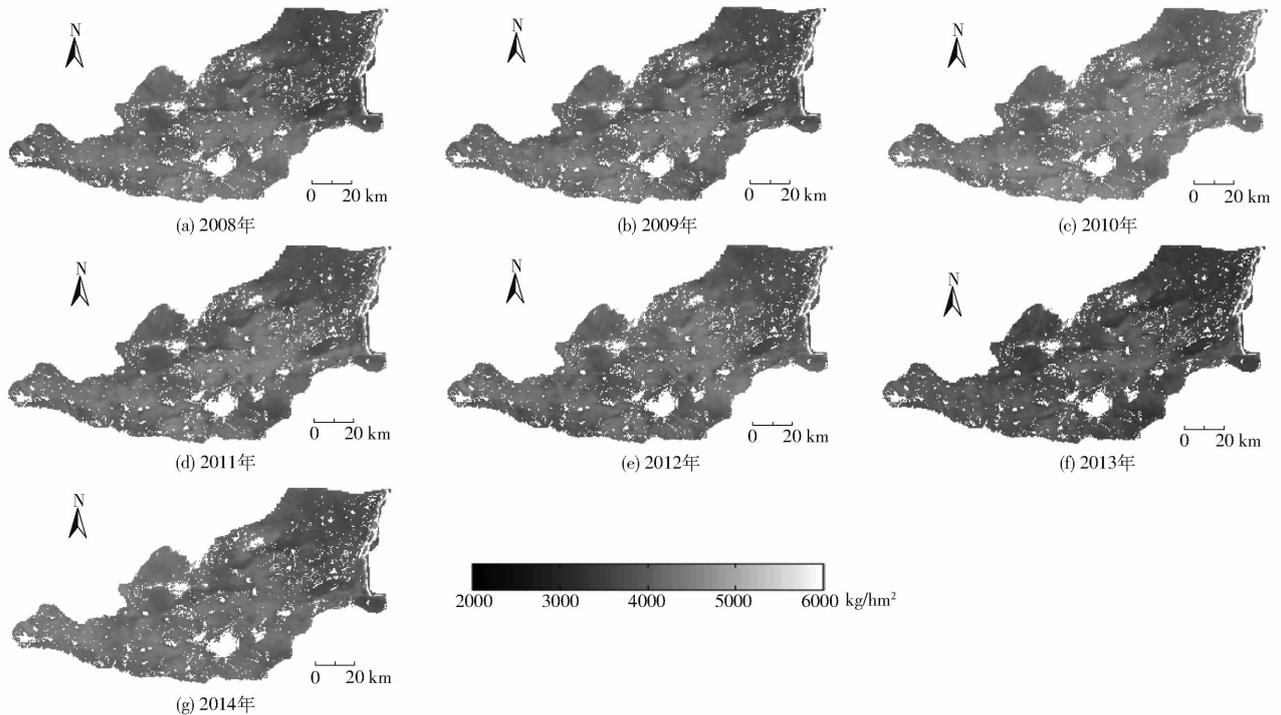


图4 关中平原2008—2014年的冬小麦单产估测结果

Fig. 4 Estimated winter wheat yields in Guanzhong Plain from 2008 to 2014

不同分位数下的小麦单产与主要生育期 VTCI 之间回归直线的整体发展趋势一致,即冬小麦单产随着旱情的加剧而降低,随着旱情的减缓而增加,这是由于研究区域冬小麦的生长和最终产量主要受到干旱因素的制约,一般而言,某区域某一年份的干旱程度与该年份的冬小麦产量的高低密切相关,故利用 VTCI 干旱监测结果与冬小麦单产之间的线性关系可实现小麦的单产估测。分位数回归模型中, $\tau = 0.5$  分位数处于所有因变量的中间位置,利用该分位数对关中四市多年的冬小麦单产进行估测最为可靠,其估测精度相对最小二乘法有所提高。然而,冬小麦产量的影响因素繁多,除干旱因素外,冬小麦产量的形成还受到多种非干旱因素(病虫害、冻害、田间管理措施等)的综合作用,只有对冬小麦产量的各种影响要素进行全面考虑,才能得到更加科学、合理的产量估测结果,这些问题是今后通过分位数回归模型进行小麦单产估测研究的重点,可尝试通过多元的分位数回归模型等来解决。

#### 4 结论

以关中平原四市 2008—2014 年的冬小麦主要生育期的 VTCI 干旱监测结果作为自变量,冬小麦

单产作为因变量,分别建立最小二乘回归模型与不同分位点(0.1, 0.3, 0.5, 0.7, 0.9)下的分位数回归模型,对比分析中位数回归、最小二乘回归与其他分位数下的回归结果及模型精度,并基于中位数回归与最小二乘回归构建的单产估测模型对研究区域的冬小麦单产进行估测。主要结论如下:

(1)分位数回归得到了冬小麦单产在 VTCI 干旱监测结果下的条件分布的一簇直线,能够比较全面的度量不同分位数下干旱与冬小麦单产之间的相关关系。而且,分位数回归采用加权的最小一乘回归法,对异常值不敏感,建立的估产模型更具可靠性。由此,分位数回归克服了普通线性回归在回归结果单一、易受异常值的影响等方面的不足,适用于关中平原冬小麦干旱影响评估的研究中。

(2)基于中位数回归构建的关中平原冬小麦的单产估测模型的精度较高,估测单产与实际单产之间的相对误差均值为 5.6%,最小值为 0;均方根误差的均值为  $160.3 \text{ kg/hm}^2$ ,最小值仅为 0,估产精度整体上优于最小二乘回归的结果。此外,应用中位数回归模型获取的冬小麦估产结果在年际变化规律与空间分布特征上与实际产量均较相符,说明中位数回归及分位数回归在研究干旱与产量之间的关系及冬小麦单产估测中具有其适用性。

#### 参 考 文 献

- 1 KOWALIK W, DABROWSKA-ZIELINSKA K, MERONI M, et al. Yield estimation using SPOT-VEGETATION products: a case study of wheat in European countries[J]. International Journal of Applied Earth Observation and Geoinformation, 2014, 32:

- 228 - 239.
- 2 赵春江. 农业遥感研究与应用进展[J/OL]. 农业机械学报, 2014, 45(12): 277 - 293. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20141241&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20141241&journal_id=jcsam). DOI:10.6041/j.issn.1000-1298.2014.12.041.  
ZHAO Chunjiang. Advances of research and application in remote sensing for agriculture[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(12): 277 - 293. (in Chinese)
  - 3 DORAISWAMY P C, SINCLAIR T R, HOLLINGER S, et al. Application of MODIS derived parameters for regional crop yield assessment[J]. Remote Sensing of Environment, 2005, 97(2): 192 - 202.
  - 4 杨鹤松, 王鹏新, 孙威. 条件植被温度指数在华北平原干旱监测中的应用[J]. 北京师范大学学报:自然科学版, 2007, 43(3):314 - 318.  
YANG Hesong, WANG Pengxin, SUN Wei. Application of the vegetation temperature condition index to drought monitoring in North China Plain[J]. Journal of Beijing Normal University:Natural Science, 2007, 43(3):314 - 318. (in Chinese)
  - 5 李军玲, 郭其乐, 彭记永. 基于MODIS数据的河南省冬小麦产量遥感估算模型[J]. 生态环境学报, 2012, 21(10): 1665 - 1669.  
LI Junling, GUO Qile, PENG Jiyong. Remote sensing estimation model of Henan Province winter wheat yield based on MODIS data[J]. Ecology and Environmental Sciences, 2012, 21(10): 1665 - 1669. (in Chinese)
  - 6 DE WIT A, DUVEILLER G, DEFOURNY P. Estimating regional winter wheat yield with WOFOST through the assimilation of green area index retrieved from MODIS observations[J]. Agricultural and Forest Meteorology, 2012, 164: 39 - 52.
  - 7 BASTIAANSEN W G M, ALI S. A new crop yield forecasting model based on satellite measurements applied across the Indus Basin, Pakistan [J]. Agriculture, Ecosystems & Environment, 2003, 94(3): 321 - 340.
  - 8 WU J J, ZHOU L, MO X Y, et al. Drought monitoring and analysis in China based on the integrated surface drought index (ISDI) [J]. International Journal of Applied Earth Observation and Geoinformation, 2015, 41:23 - 33.
  - 9 李芬, 于文金, 张建新, 等. 干旱灾害评估研究进展[J]. 地理科学进展, 2011, 28(7): 891 - 898.  
LI Fen, YU Wenjin, ZHANG Jianxin, et al. Review of drought disaster evaluation[J]. Progress in Geography, 2011, 28(7): 891 - 898. (in Chinese)
  - 10 史舟, 梁宗正, 杨媛媛, 等. 农业遥感研究现状与展望[J/OL]. 农业机械学报, 2015, 46(2): 247 - 260. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20150237&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20150237&journal_id=jcsam). DOI: 10.6041/j.issn.1000-1298.2015.02.037.  
SHI Zhou, LIANG Zongzheng, YANG Yuanyuan, et al. Status and prospect of agriculture remote sensing[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(2): 247 - 260. (in Chinese)
  - 11 李文娟, 覃志豪, 林绿. 农业旱灾对国家粮食安全影响程度的定量分析[J]. 自然灾害学报, 2010, 19(3): 111 - 118.  
LI Wenjuan, QIN Zhihao, LIN Lü. Quantitative analysis of agro-drought impact on food security in China[J]. Journal of Nature Disasters, 2010, 19(3): 111 - 118. (in Chinese)
  - 12 周磊, 武建军, 张洁. 以遥感为基础的干旱监测方法研究进展[J]. 地理科学, 2014, 28(1): 85 - 91.  
ZHOU Lei, WU Jianjun, ZHANG Jie. Remote sensing based drought monitoring approach and research progress[J]. Scientia Geographica Sinica, 2014, 28(1): 85 - 91. (in Chinese)
  - 13 王鹏新, 龚健雅, 李小文. 条件植被温度指数及其在干旱监测中的应用[J]. 武汉大学学报:信息科学版, 2001, 26(5): 412 - 418.  
WANG Pengxin, GONG Jianya, LI Xiaowen. Vegetation temperature condition index and its application for drought monitoring [J]. Geomatics and Information Science of Wuhan University, 2001, 26(5): 412 - 418. (in Chinese)
  - 14 SUN W, WANG P X, ZHANG S Y, et al. Using the vegetation temperature condition index for time series drought occurrence monitoring in the Guanzhong Plain, PR China[J]. International Journal of Remote Sensing, 2008, 29(17): 5133 - 5144.
  - 15 陈阳, 范建容, 郭芬芬, 等. 条件植被温度指数在云南干旱监测中的应用[J]. 农业工程学报, 2011, 27(5): 231 - 236.  
CHEN Yang, FAN Jianrong, GUO Fenfen, et al. Application of the vegetation temperature condition index to drought monitoring in Yunnan Province[J]. Transactions of the CSAE, 2011, 27(5): 231 - 236. (in Chinese)
  - 16 李艳, 王鹏新, 刘峻明, 等. 基于条件植被温度指数的冬小麦主要生育时期干旱监测效果评价 I: 因子权重排序法和熵值法组合赋权[J]. 干旱地区农业研究, 2013, 31(6):159 - 163.  
LI Yan, WANG Pengxin, LIU Junming, et al. Evaluation of drought monitoring effects in the main growth and development stages of winter wheat using vegetation temperature condition index. I: factor weight sorting method and entropy method[J]. Agricultural Research in the Arid Areas, 2013, 31(6):159 - 163. (in Chinese)
  - 17 田苗, 王鹏新, 张树誉, 等. 基于条件植被温度指数的冬小麦产量预测[J/OL]. 农业机械学报, 2014, 45(2): 239 - 245. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20140240&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20140240&journal_id=jcsam). DOI: 10.6041/j.issn.1000-1298.2014.02.040.  
TIAN Miao, WANG Pengxin, ZHANG Shuyu, et al. Winter wheat yield forecasting based on vegetation temperature condition index[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(2): 239 - 245. (in Chinese)
  - 18 张顺谦, 卿清涛, 侯美亭, 等. 基于温度植被干旱指数的四川伏旱遥感监测与影响评估[J]. 农业工程学报, 2007, 23(9): 141 - 146.  
ZHANG Shunqian, QING Qingtao, HOU Meiting, et al. Remote sensing and impact estimation for Sichuan hot-drought based on temperature vegetation dryness index[J]. Transactions of the CSAE, 2007, 23(9): 141 - 146. (in Chinese)

- optimization [J]. *Computer Physics Communications*, 2014, 185(10):2622–2628.
- 13 崔文华, 刘晓冰, 王伟, 等. 混合蛙跳算法研究综述[J]. *控制与决策*, 2012, 27(4):481–486.  
CUI Wenhua, LIU Xiaobing, WANG Wei, et al. Survey on shuffled frog leaping algorithm [J]. *Control and Decision*, 2012, 27(4):481–486. (in Chinese)
- 14 金芬, 孙春华, 钟鸣. 遗传算法中适应度函数的改进[J]. *机械设计与制造*, 2010(3):218–219.  
JIN Fen, SUN Chunhua, ZHONG Ming. Improvement of fitness function in genetic algorithm [J]. *Machinery Design & Manufacture*, 2010(3):218–219. (in Chinese)
- 15 刘建锋, 肖文发, 江泽平, 等. 景观破碎化对生物多样性的影响[J]. *林业科学研究*, 2005, 18(2):222–226.  
LIU Jianfeng, XIAO Wenfa, JIANG Zeping, et al. A study on the influence of landscape fragmentation on biodiversity [J]. *Forest Research*, 2005, 18(2):222–226. (in Chinese)
- 16 覃凤飞, 安树青, 卓元午, 等. 景观破碎化对植物种群的影响[J]. *生态学杂志*, 2003, 22(3):43–48.  
TAN Fengfei, AN Shuqing, ZHUO Yuanwu, et al. Effect of landscape fragmentation on plant populations [J]. *Chinese Journal of Ecology*, 2003, 22(3):43–48. (in Chinese)
- 17 谭龙, 陈冠, 王思源, 等. 逻辑回归与支持向量机模型在滑坡敏感性评价中的应用[J]. *工程地质学报*, 2014, 22(1):56–63.  
TAN Long, CHEN Guan, WANG Siyuan, et al. Landslide susceptibility mapping based on logistic regression and support vector machine [J]. *Journal of Engineering Geology*, 2014, 22(1):56–63. (in Chinese)
- 18 于强, 岳德鹏, 张启斌, 等. 磴口县荒漠绿洲景观时空演变及其格局特征分析[J]. *生态科学*, 2016, 35(6):73–83.  
YU Qiang, YUE Depeng, ZHANG Qibin, et al. Spatial-temporal evolution and pattern characteristics of desert oasis landscape in Dengkou County [J]. *Ecological Science*, 2016, 35(6):73–83. (in Chinese)
- 19 张启斌, 岳德鹏, 于强, 等. 林业生态工程建设对磴口县景观格局演变及重心迁移的影响[J]. *浙江农业学报*, 2017, 29(2):261–269.  
ZHANG Qibin, YUE Depeng, YU Qiang, et al. Effects of forestry ecological projects on landscape pattern evolution and migration of barycenter in Dengkou County [J]. *Acta Agriculturae Zhejiangensis*, 2017, 29(2):261–269. (in Chinese)
- 20 于强, 岳德鹏, DI Yang, 等. 基于 EnKF-MCRP 模型的生态用地扩张模拟研究[J/OL]. *农业机械学报*, 2016, 47(9):285–293. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20160939&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20160939&journal_id=jcsam). DOI:10.6041/j.issn.1000-1298.2016.09.039.  
YU Qiang, YUE Depeng, DI Yang, et al. Simulation on ecological land use expansion based on EnKF-MCRP model [J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2016, 47(9):285–293. (in Chinese)

~~~~~

(上接第 173 页)

- 19 黄弘, 王鹏新, 李俐. 关中平原小麦生育期 VTCI 加权估算及其与产量的相关性研究[J]. *干旱地区农业研究*, 2011, 29(6):173–178.  
HUANG Hong, WANG Pengxin, LI Li. Correlations between weighted VTCI in key growth and development stages of winter wheat and wheat yields in the Guanzhong Plain [J]. *Agricultural Research in the Arid Areas*, 2011, 29(6):173–178. (in Chinese)
- 20 黄健熙, 罗倩, 刘晓暄, 等. 基于时间序列 MODIS NDVI 的冬小麦产量预测方法[J/OL]. *农业机械学报*, 2016, 47(2):295–301. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20160239&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20160239&journal_id=jcsam). DOI:10.6041/j.issn.1000-1298.2016.02.039.  
HUANG Jianxi, LUO Qian, LIU Xiaoxuan, et al. Winter wheat yield forecasting based on time series of MODIS NDVI [J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2016, 47(2):295–301. (in Chinese)
- 21 罗玉波. 分位数回归模型及其应用[M]. 北京: 知识产权出版社, 2009.
- 22 KOENKER R, BASSETT G J. Regression quantiles [J]. *Journal of the Econometric Society*, 1987, 46(1):33–50.
- 23 陈建宝, 丁军军. 分位数回归技术综述[J]. *统计与信息论坛*, 2008, 23(3):89–96.  
CHEN Jianbao, DING Junjun. A review of technologies on quantile regression [J]. *Statistics & Information Forum*, 2008, 23(3):89–96. (in Chinese)
- 24 FOURNIER J M, KOSKE I. Public employment and earnings inequality: an analysis based on conditional and unconditional quantile regressions [J]. *Economics Letters*, 2013, 121:263–266.
- 25 ARUNRAJ N S, AHRENS D. A hybrid seasonal autoregressive integrated moving average and quantile regression for daily food sales forecasting [J]. *International Journal of Production Economics*, 2015, 170:321–335.
- 26 ROCCHINI D, CADE B S. Quantile regression applied to spectral distance decay [J]. *IEEE Geoscience and Remote Sensing Letters*, 2008, 5(4):640–643.
- 27 贺音, 张聪娥, 张黎. 基于 SPEI 的陕西近 40 年干旱时空特性分析[J]. *陕西气象*, 2014(5):26–32.  
HE Yin, ZHANG Conge, ZHANG Li. Analyzing the temporal and spatial characteristics of drought in Shaanxi Province in recent 40 years based on standardized precipitation evapotranspiration index [J]. *Journal of Shaanxi Meteorology*, 2014(5):26–32. (in Chinese)
- 28 王蕾. 基于条件植被温度指数的冬小麦主要生育期干旱影响评估研究[D]. 北京: 中国农业大学, 2015.  
WANG Lei. Estimation of drought impact on the main growth stage of winter wheat based on vegetation temperature condition index [D]. Beijing: China Agricultural University, 2015. (in Chinese)