

کاربرد رگرسیون چندک در تحلیل منحنی سنجه بار رسوب معلق

میثم سالاری جزی^{1*}

تاریخ دریافت: 1395/5/10 تاریخ پذیرش: 1395/7/28

چکیده

روش منحنی سنجه رسوب پرکاربردترین برای بار رسوب معلق رودخانه است که با استفاده از روش رگرسیون معمولی حداقل مربعات ارتباطیمنی‌نگین شرطی بار معلق رسوب و دبی جریان را نشان می‌دهد و برای برآورد بار معلق رسوب به عنوان تابعی از دبی جریان رودخانه به کار گرفته می‌شود. مدل رگرسیون معمولی حداقل مربعات به داده‌های پرت حساس است و زمانی که پیش فرض‌های مرتبط با تحلیل باقیمانده‌ها برقرار نباشند قابل پذیرش نیست. رگرسیون چندک (کوانتایل) یک روش آماری است که می‌تواند برای غلبه بر این محدودیت‌ها در تحلیل منحنی سنجه رسوب به کار گرفته شود. در این مطالعه با استفاده از داده‌های ایستگاه هیدرومتری الگدره در استان گلستان (دوره ثبت شده سال‌های 1380-1366) روش رگرسیون کوانتایل در برآورد منحنی سنجه رسوب به کار گرفته شد و نتایج با کاربرد روش متداول رگرسیون معمولی حداقل مربعات مقایسه گردید. نتایج نشان می‌دهد کاربرد رگرسیون معمولی حداقل مربعات در تحلیل منحنی سنجه رسوب منجر به برآورد اریب شده است در حالی که رگرسیون کوانتایل بدون داشتن مفروضات محدود کننده رگرسیون معمولی حداقل مربعات خطا، می‌تواند به شکل مناسبی تأثیر دبی جریان را در کوانتایل‌های مختلف توزیع بار رسوب معلق به ویژه در کوانتایل‌های ابتدایی و انتهایی نشان دهد. به علاوه مشخص گردید که بزرگی تأثیر دبی جریان بر بار رسوب معلق به ترتیب متعلق به کوانتایل‌های بالایی، پایینی و میانی بار رسوب معلق می‌باشد و با افزایش دبی جریان رودخانه داده‌های بار رسوب معلق به مقدار بیش‌تری به سمت راست چولگی نشان می‌دهند. همچنین از مفهوم رگرسیون کوانتایل می‌توان به عنوان یک ابزار کاربردی بسیار مهم برای استخراج تابع چگالی احتمال و تابع توزیع تجمعی بار رسوب معلق به ازای مقادیر مشخص دبی جریان ارایه شد.

واژه‌های کلیدی: حداقل مربعات، رگرسیون کوانتایل، رگرسیون معمولی، منحنی سنجه بار رسوب

مقدمه

توجه قرار می‌گیرد. در رویکرد هیدرولوژیکی روش منحنی سنجه رسوب به عنوان شناخته شده‌ترین روش تخمین بار معلق رسوب در نظر گرفته می‌شود. در این روش یک رابطه رگرسیونی بین دبی رودخانه و نیز بار معلق رسوب برآزش داده می‌شود و از این روش ساده و کاربردی در بسیاری از طراحی‌های مورد استفاده قرار می‌گیرد (Kao et al., 2005; Salarijazi et al., 2016; Ghorbani et al., 2016). به طور معمول نمونه برداری‌های انجام شده در رودخانه‌ها که در واقع منجر به دسترسی به داده‌های مورد نیاز برای تحلیل منحنی‌سنجه رسوب می‌شوند به شکل محدودی به داده‌های زمان سیلاب دسترسی دارند که این موضوع منجر به محدودیت برای استفاده از این روش می‌شود. با وجود آنکه محققان بسیاری استفاده از رگرسیون معمولی حداقل مربعات را برای تحلیل منحنی سنجه رسوب، ناکافی و دارای دقت نامناسب معرفی کرده‌اند اما در عمل استفاده از این روش به شکل گسترده‌ای مرسوم است (Walling., 1997; McBean and Al-Nassri., 1988; Cohn., 1992; Kisi., 2005). روش‌های متنوعی جهت بهبود دقت منحنی سنجه رسوب توسط محققین مختلف پیشنهاد شده است. در میان این

برآورد میزان بار رسوب که توسط رودخانه‌ها انتقال می‌یابد یک موضوع مهم و کاربردی در مدیریت و برنامه ریزی‌های هیدرولوژیکی می‌باشد زیرا طراحی ظرفیت مخزن و نیز تخمین زمان پر شدن مخزن، طراحی سازه‌های هیدرولیکی، ارزیابی معیارهای کیفیت آب، پایداری کانال‌ها برای کشتی‌رانی، فعالیت‌های اکولوژیکی و زیست محیطی و بسیاری از مسایل دیگر در ارتباط با فرآیند انتقال بار رسوبی رودخانه می‌باشند (Morehead et al., 2003; Yang et al., 2007; Huang and Montgomery., 2013; Warrick., 2015). برای برآورد بار رسوب رودخانه دو رویکرد هیدرولیکی و هیدرولوژیکی وجود دارد. علی‌رغم برتری نظری رویکرد هیدرولیکی در اغلب موارد، رویکرد هیدرولوژیکی به علت نیاز به دسترسی به حداقل داده‌ها و نیز سادگی و سهولت کاربرد آن در کارهای عملی بیش‌تر مورد

1- استادیار گروه مهندسی آب، دانشکده مهندسی آب و خاک، دانشگاه علوم کشاورزی و منابع طبیعی گرگان
* -نویسنده مسئول: (Email: meysam.salarijazi@gau.ac.ir)

تغییرات تمام قسمت‌های توزیع داده‌ها را در نظر بگیرد در سالیان اخیر مورد توجه محققین در اقتصاد و مباحث مالی (Meligkotsidou et al., 2009; Gaglianone et al., 2011; Alagidede and Cade and Panagiotidis., 2012) زیست شناسی و اکولوژی (Noon., 2003; Austin., 2007; Cozzoli et al., 2013 Jagger and Elsner., 2009; Hirschi et al., 2011) قرار گرفته است. هدف این پژوهش آن است که کاربرد رگرسیون کوانتایل برای برآورد احتمالاتی بار رسوب معلق روزانه با استفاده منحنی سنج رسوب مورد بررسی قرار گیرد و نتایج آن با نتایج روش متداول کاربرد رگرسیون معمولی حداقل مربعات مقایسه گردد. همچنین کاربردهای متنوع رگرسیون کوانتایل در برآورد منحنی‌های متفاوت سنجه رسوب به ازای کوانتایل‌های متفاوت ارایه می‌شود و کاربرد آن در برآورد تابع توزیع احتمال و تجمعی تجربی بار رسوب معلق به ازای دبی‌های مشخص جریان رودخانه معرفی می‌گردد. برای بررسی‌های بیان شده از داده‌های ثبت شده یک ایستگاه هیدرومتری در استان گلستان استفاده می‌شود.

مواد و روش‌ها

داده‌های مورد استفاده: داده‌های مورد استفاده در این تحقیق از سری داده‌های ثبت شده ایستگاه هیدرومتری النگ‌دره در شهرستان گرگان در استان گلستان می‌باشد. این مجموعه داده شامل 122 جفت داده ثبت شده دبی جریان رودخانه و بار رسوب معلق در بازه زمانی سال‌های 1366-1380 است. مشخصه‌های داده‌های ثبت شده در جدول 1 ارایه شده است.

روش‌های توان به فاکتور تصحیح اریب (Ferguson., 1986; Cohn et al., 1989; Philips et al., 1999; Asselman., 2000) سنجه چند متغیره (Cohn et al., 1992; Wang and Linker., 2008; Wang and Tian., 2013) و روش‌های متنوع رگرسیون (Hicks et al., 2000; Krishnaswamy et al., 2001; Tarras-Wahlberg et al., 2001) اشاره کرد. در ایران نیز تحقیقات مختلفی در زمینه برآورد بار رسوب با استفاده از منحنی سنجه رسوب انجام شده است که در این تحقیقات کارایی روش پایه و روش‌های اصلاحی مورد توجه قرار گرفته است (کالوندی و همکاران، 1389؛ کیا و عمادی، 1392). همچنین در سال‌های اخیر تحقیقات متعددی در استفاده از روش‌های متنوع جدید برای پیش‌بینی بار رسوب معلق بر اساس دبی رودخانه انجام شده است که این روش‌ها شامل شبکه عصبی مصنوعی (Jain., 2001; Nagy et al., 2002; Alp and Cigizoglu., 2007; Rai and Mathur., 2008; Afan et al., 2015)، برنامه ریزی ژنتیک (Aytek and Kisi., 2008; Guven and Kisi., 2011; Kitsikoudis et al., 2014) و شبکه عصبی-موجکی (Kisi., 2004; Lohani et al., 2007) و ماشین‌های بردار پشتیبان (Partal and Cigizoglu., 2008) است. رگرسیون کوانتایل یک روش آماری برای رگرسیون برای برآورد کوانتایل‌های مختلف متغیر وابسته است درحالی‌که رگرسیون معمولی حداقل مربعات بر میانگین استوار است. درمقایسه با روش رگرسیون معمولی حداقل مربعات که در آن برآورد توابع شرطی میانگین بر اساس حداقل‌سازی مجموع مربعات باقیمانده‌ها انجام می‌شود، روش رگرسیون کوانتایل بر اساس حداقل‌سازی نامتقارن قدرمطلق موزون باقیمانده‌ها و با هدف برآورد توابع شرطی میانه و دامنه متنوعی از سایر کوانتایل‌ها مطرح می‌گردد. روش رگرسیون کوانتایل به علت آن‌که می‌تواند

جدول 1- ویژگی‌های آماری داده‌های مورد استفاده در تحلیل منحنی سنجه رسوب

انحراف معیار	حداکثر	متوسط	حداقل
دبی (مترمکعب بر ثانیه)	۱/۱۶	۰/۱۴	۰/۰۱
بار معلق رسوب (تن بر روز)	۳۷	۰/۷۳	۰/۰۰۰۰۰۱۷۳

$$E(Y_i) = \alpha + \beta x_i \quad (2)$$

که $E(Y_i)$ میانگین شرطی (Conditional Mean) متغیر تصادفی Y است با $E(Y | X)$ نیز نمایش داده می‌شود. به عبارت دیگر، متغیر تصادفی Y در هر سطح از متغیر مستقل دارای توزیعی است که میانگین‌های این توزیع‌ها روی یک خط راست جای گرفته‌اند. نیز در ادامه مفروضات مربوط به بخش خطا مورد بررسی قرار می‌گیرند. این مفروضات عبارتند از اینکه ε_i ها دارای توزیع نرمال، مستقل از

رگرسیون معمولی حداقل مربعات: مدل رگرسیون معمولی حداقل مربعات خطی با یک متغیر مستقل به فرم زیر می‌باشد.

$$Y_i = \alpha + \beta x_i + \varepsilon_i \quad (1)$$

که در آن ε_i متغیر تصادفی و α, β پارامترهای نامعلوم مدل که باید برآورد شوند و x_i مقادیر معلوم از متغیر مستقل هستند.

در صورتیکه $E(\varepsilon_i) = 0$ باشد آنگاه مدل 1 به فرم رابطه 2 بازنویسی می‌شود:

مجموع قدرمطلق باقیمانده‌های موزون برای برآورد پارامترهای مدل استفاده می‌شود که به آن روش حداقل قدرمطلق باقیمانده‌ها یا LAD گفته می‌شود. این معیار برای مدل 3 به قرار رابطه 4 است:

$$\phi_{LAD} = \theta \sum_{i/y_i \leq \alpha_\theta + \beta_\theta x_i} (y_i - \alpha_\theta - \beta_\theta x_i) + (1-\theta) \sum_{i/y_i > \alpha_\theta + \beta_\theta x_i} (y_i - \alpha_\theta - \beta_\theta x_i) \quad (4)$$

به این ترتیب، خط برازشی به گونه‌ای است که $100 \times \theta$ درصد نقاط تقریباً زیر آن و باقی آن‌ها بالای خط قرار می‌گیرند. بر خلاف روش حداقل مربعات، روش حداقل قدرمطلق باقیمانده‌ها نسبت به داده‌های دورافتاده استوار است. این ویژگی ناشی از آن است که بر خلاف اهمیت اندازه باقیمانده‌ها در روش حداقل مربعات، در این روش فقط علامت باقیمانده‌ها مورد توجه قرار می‌گیرد. بنابراین، تعداد باقیمانده‌هایی که بیش‌تر (مثبت) یا کم‌تر (منفی) از کوانتایل مورد نظرند و نه مقدار بزرگی آن‌ها در برآورد اثر گذار است. پس، داده‌های دور افتاده که تأثیر خود را از طریق بزرگی باقیمانده‌ها نشان می‌دهند، نمی‌توانند برآوردهای LAD را متأثر سازند. همچنین شکل بسته‌ای برای برآورد پارامترهای این مدل وجود ندارد و از روش‌های عددی برای برآورد آن‌ها استفاده می‌شود. همچنین، جواب‌های نهایی مدل رگرسیون کوانتایل می‌تواند یکتا نباشد. البته یافتن جواب یکتا با انتخاب یک معیار مناسب میسر است (انصاری و همکاران، 1385؛ Koenker., 2005). در این مطالعه برای بررسی دقت برازش مدل از معیارهای R^2 (ضریب تبیین)، Adjusted R^2 (ضریب تبیین تعدیل شده)، MAPE (میانگین قدر مطلق درصد خطا) و AIC (معیار اطلاعات آکائیک) استفاده شده است که به صورت رابطه 5 تا 8 تعریف می‌گردند.

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad SST = \sum_{i=1}^n (y_i - \bar{y}_i)^2, \quad R^2 = 1 - \frac{SSE}{SST} \quad (5)$$

$$Adjusted R^2 = 1 - \left(\frac{(n-1)}{(n-k)} \right) \cdot (1 - R^2) \quad (6)$$

$$MAPE = \frac{100}{n} \sum_{i=1}^n |(y_i - \hat{y}_i) / y_i| \quad (7)$$

$$MSE = [1/n] \cdot SSE, \quad AIC = n \cdot \ln(MSE) + 2k \quad (8)$$

نتایج و بحث

در ابتدا با در نظر گرفتن تبدیل‌های لگاریتمی داده‌های دبی جریان و نیز بار رسوب معلق از یک معادله خطی متناظر با منحنی سنجه رسوب استفاده و کارایی آن مورد ارزیابی قرار گرفت. نتایج به دست آمده از تحلیل مدل رگرسیون خطی در بخش‌های مختلف جدول 2 و شکل 1 ارائه شده است. معادله فارسی اصطلاحات مهم به-کار رفته در جداول و شکل‌های این متن به صورت ذیل می‌باشد:

DF: درجه آزادی، Intercept: عرض از مبدا، Coefficient:

یکدیگر و دارای واریانس یکسان هستند. در بعضی موارد ممکن است با وجود ضریب تبیین مناسب تحلیل باقیمانده‌ها منجر به نتیجه‌گیری این موضوع شود که مدل رگرسیونی مورد بررسی مناسب نمی‌باشد. میانگین یکی از شاخص‌های مرکزی است و آگاهی از آن به تنهایی نمی‌تواند اطلاعات کاملی از شکل توزیع به همراه داشته باشد بنابراین رگرسیون معمولی نیز ممکن است نتواند اطلاعات کافی درباره شکل توزیع متغیر تصادفی تحت بررسی را در سطوح مختلف متغیر مستقل به دست دهد. در این جهت، کوانتایل‌ها که معیارهای دیگری برای تمرکز هستند (در کنار هم) و نیز می‌توانند شکل توزیع را جامع‌تر به تصویر بکشند. برای مثال، اگر فاصله دهک‌ها از یکدیگر تقریباً برابر، انتظار داریم توزیع (نسبتاً هموار) یا یکنواختی داشته باشیم. همچنین اگر دهک‌های بالایی دارای فاصله زیاد و دهک‌های پایینی فاصله کمی از یکدیگر داشته باشند، توزیع به سمت راست چوله خواهد بود. مشخص است اگر مانند رگرسیون معمولی که برای میانگین به کار می‌رود، یک شیوه رگرسیونی برای کوانتایل‌ها نیز استفاده شود موجب می‌شود شکل توزیع در سطوح مختلف متغیر مستقل برآورد گردد (انصاری و همکاران، 1385)

رگرسیون کوانتایل: مدل 2 برای میانگین شرطی برازش داده می‌شود. مدل رگرسیون کوانتایل با ایده‌ای مشابه برای کوانتایل‌های شرطی (Conditional Quantiles) به کار می‌رود. مانند رگرسیون معمولی (میانگین)، کاربردهایی نظیر بررسی رابطه متغیرهای وابسته با کوانتایل‌ها و همچنین پیش‌بینی آن‌ها برای این نوع از رگرسیون نیز میسر است. با وجود این، شاید مهم‌ترین کاربرد رگرسیون کوانتایل، شناسایی شکل توزیع متغیر وابسته مدل در سطوح گوناگون متغیر مستقل باشد؛ اینکار با برازش مدل رگرسیونی متعدد بر یک مجموعه داده به ازای کوانتایل‌های مختلف صورت می‌گیرد. برای ارایه تعریف دقیقی از مدل رگرسیون کوانتایل $\theta \in (0,1)$ ام ابتدا مدل 1 را به صورت $Y_i = \alpha_\theta + \beta_\theta X_i + \varepsilon_{\theta i}$ می‌توان بازنویسی نمود. مدل رگرسیون ساده خطی کوانتایل $\theta \in (0,1)$ ام (کونکر و باست، 1978) عبارتست از:

$$Q_{\theta_i} = \alpha_\theta + \beta_\theta X_i \quad i = 1, 2, \dots, n \quad (3)$$

که در آن Q_{θ_i} کوانتایل $\theta \in (0,1)$ ام متغیر تصادفی Y در سطح i ام متغیر مستقل نشان می‌دهد. توجه کنید که این مدل معادل مدل 1 است هنگامی که کوانتایل $\theta \in (0,1)$ ام متغیر تصادفی ε به ازای تمام سطوح متغیر مستقل برابر با صفر است. شیوه برآورد پارامترهای مدل رگرسیون معمولی بر حداقل کردن مربع باقیمانده‌های مدل مبتنی است که روش حداقل مربعات (Least Squares) نامیده می‌شود. در این روش، منحنی رگرسیونی به گونه‌ای برازش داده می‌شود که در مجموع فاصله نقاط از آن به حداقل برسد. در رگرسیون کوانتایل بر خلاف رگرسیون معمولی از حداقل نمودن

دبی جریان، Log Y: داده‌های تبدیل لگاریتمی شده رسوب معلق، Discharge (m³/s): دبی (جریان) (مترمکعب بر ثانیه)، Suspended Sediment Load (Ton/Day): بار رسوب معلق (تن بر روز)

شیب، Standard Error: خطای استاندارد، Lower Bound: کران پایین، Upper Bound: کران بالا، Value: مقدار، Conf. Interval: بازه اطمینان، Residual: باقیمانده، Standardized Residuals: باقیمانده‌های استاندارد شده، Log X: داده‌های تبدیل لگاریتمی شده

جدول 2- نتایج کاربرد رگرسیون خطی در برآورد منحنی سنجه رسوب (تحلیل واریانس، پارامترهای مدل، معیارهای دقت برازش) تحلیل واریانس

منبع	درجه آزادی	مجموع مربعات	میانگین مربعات	F	Pr > F
مدل	۱	۲۵۲/۱۸۸	۲۵۲/۱۸۸	۱۱۷۲/۲۳۵	< ۰/۰۰۰۱
خطا	۱۲۰	۲۵/۸۱۶	۰/۲۱۵		
کل	۱۲۱	۲۷۸/۰۰۵			

منبع	مقدار	خطای استاندارد	t	Pr > t	باند پایین (0/95)	باند بالا (0/95)
عرض از مبدأ	۰/۳۱۱	۰/۱۰۱	۳/۰۶۴	۰/۰۰۳	۰/۱۱	۰/۵۱۲
شیب خط	۲/۶۰۶	۰/۰۷۶	۳۴/۲۳۸	< ۰/۰۰۰۱	۲/۴۵۶	۲/۷۵۷

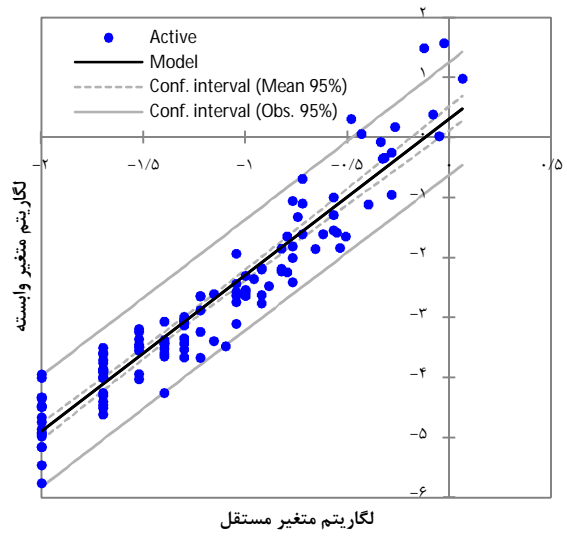
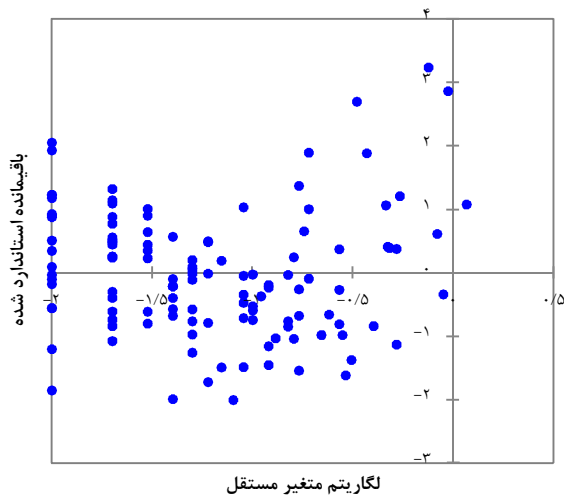
ضریب تعیین	ضریب تعیین تعدیل شده	میانگین قدر مطلق درصد خطا	معیار اطلاعات آکائیک
۰/۹۰۷	۰/۹۰۶	۴۵/۴۴۵	-۱۸۵/۴۶۸

است و نیز همه معادلات خطی برآورد شده به ازای کوانتایل‌های مختلف از نظر آماری معنی‌دار می‌باشند. بررسی ضریب تبیین کوانتایل‌های مختلف بیانگر آن است که این مشخصه آماری برای کوانتایل 0/1 دارای بیش‌ترین مقدار یعنی 0/97 می‌باشد و با افزایش مقادیر کوانتایل‌ها ضریب تبیین کاسته شده و به مقدار 0/79 برای کوانتایل 0/9 می‌رسد. الگوی تغییرات معیار اطلاعات آکائیک متفاوت از الگوی تغییرات ضریب تبیین می‌باشد. بیش‌ترین مقدار معیار اطلاعات آکائیک برای کوانتایل 0/1 است که برابر با مقدار 317/26- است و به ترتیب کوانتایل‌های 0/9، 0/2، 0/8، 0/25، 0/75، 0/3، 0/7، 0/4، 0/6 و 0/5 دارای رتبه‌های بعدی در این معیار می‌باشند. به عبارت دیگر به طور مشخص کوانتایل‌های بالا و پایین دارای دقت برازش بیش‌تری نسبت به کوانتایل‌های مرکزی بوده‌اند. الگوی تغییرات معیار میانگین قدر مطلق درصد خطا دارای الگوی تغییراتی تقریباً مشابه با الگوی تغییرات معیار اطلاعات آکائیک می‌باشد. با در نظر گرفتن این معیار دقت برازش‌ها به ترتیب به صورت 0/9، 0/1، 0/2، 0/25، 0/8، 0/3، 0/6، 0/75، 0/7، 0/5 و 0/4 رتبه‌بندی می‌شود به طوری‌که کم‌ترین خطا در کوانتایل 0/1 برابر با 5/72 و بیش‌ترین خطا در کوانتایل 0/4 برابر با 15/82 می‌باشد. مقادیر عرض از مبدأ معادلات خطی رگرسیون کوانتایل به ازای کوانتایل‌های بین

بررسی جدول 2 نشان می‌دهد که مدل رگرسیون خطی اعمال شده بر داده‌های دبی جریان - بار رسوب معلق از نظر آماری معنی‌دار است و همچنین مقدار ضریب تبیین بیانگر برازش مناسب مدل خطی بر داده‌ها می‌باشد. با توجه به نتایج بیان شده تحلیل‌های گرافیکی برازش مدل رگرسیون خطی مدنظر قرار گرفت. بررسی شکل 1 نشان می‌دهد با وجود ضریب تبیین کاملاً مناسب و نیز معنی‌داری آماری رگرسیون خطی، باقیمانده‌های مدل دارای توزیع تقریباً نرمال نیستند و نیز همچنین دارای واریانس ثابت نمی‌باشند. عبارتی پیش‌فرض‌های مدل رگرسیون خطی برقرار نمی‌باشد. در گام بعدی مدل رگرسیون کوانتایل مورد بررسی قرار گرفت. کاربرد رگرسیون کوانتایل منجر به استخراج مجموعه‌ای از معادلات خطی متناظر با کوانتایل‌های منتخب می‌گردد. در این مطالعه کوانتایل‌های 0/9، 0/8، 0/75، 0/7، 0/6، 0/5، 0/4، 0/3، 0/25، 0/2 و 0/1 می‌باشد. همان‌طور که مشاهده می‌شود در واقع کوانتایل‌های مورد بررسی شامل دهک‌ها و چارک‌های مختلف می‌باشد. یافته‌های کاربرد مدل رگرسیون کوانتایل در جدول 3 و شکل 2 ارایه شده است. همان‌طور که نتایج جدول 3 مشخص می‌کند به ازای تمام کوانتایل‌های مورد بررسی مقادیر ضریب تبیین کاملاً قابل پذیرش

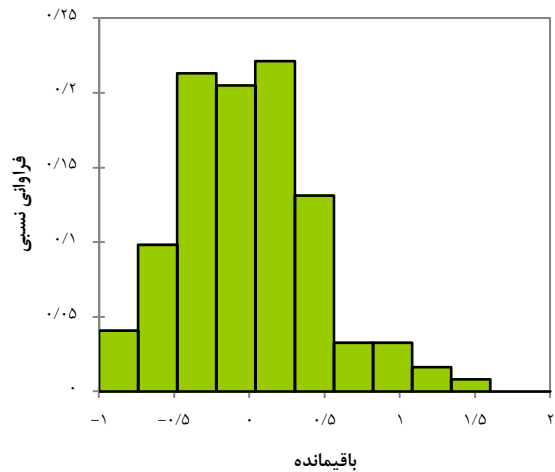
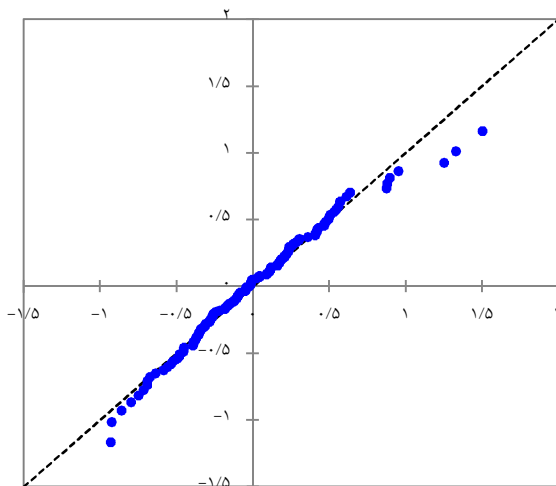
0/1 تا 0/4 دارای مقادیر منفی و برای کوانتایل‌های بین 0/5 تا 0/9

دارای مقادیر مثبت می‌باشند.

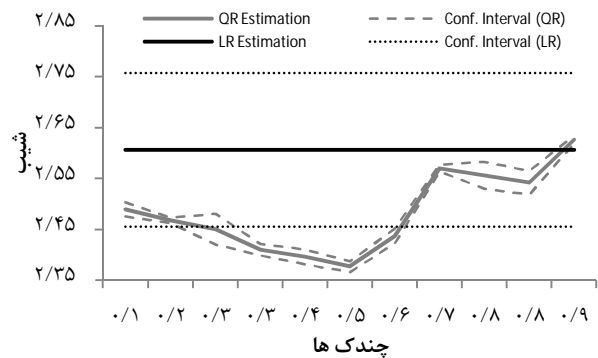
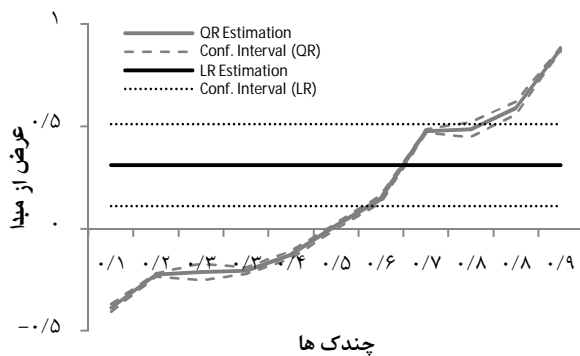


نمودار Q-Q (باقیمانده‌ها)

هیستوگرام (باقیمانده‌ها)



شکل 1- معیارهای گرافیکی بررسی دقت کاربرد رگرسیون خطی در برآورد منحنی سنجه رسوب



شکل 2- نمودار پارامترهای مدل رگرسیون کوانتایل به ازای کوانتایل‌های مختلف مورد بررسی

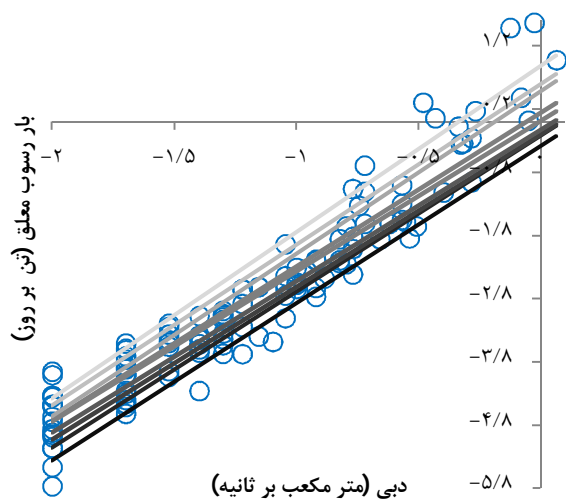
جدول 3- نتایج کاربرد رگرسیون کوانتایل در برآورد منحنی سنجه رسوب (پارامترهای مدل‌ها، دقت برازش، سطح معنی‌داری آماری)

چندک-ها	عرض از مبدأ	شیب	ضریب تیبین	ضریب تبیین شده	میانگین قدر مطلق درصد خطا	معیار اطلاعات آکائیک	p
۰/۱۰۰	-۰/۳۹	۲/۴۹	۰/۹۷	۰/۹۷	۵/۷۲	-۳۱۷/۲۶	<۰/۰۰۰۱
۰/۲۰۰	-۰/۲۲	۲/۴۷	۰/۹۵	۰/۹۵	۹/۳۲	-۲۶۱/۸۹	<۰/۰۰۰۱
۰/۲۵۰	-۰/۲۱	۲/۴۵	۰/۹۴	۰/۹۴	۱۱/۴۰	-۲۴۴/۹۴	<۰/۰۰۰۱
۰/۳۰۰	-۰/۲۱	۲/۴۱	۰/۹۲	۰/۹۳	۱۳/۴۳	-۲۳۲/۴۱	<۰/۰۰۰۱
۰/۴۰۰	-۰/۱۳	۲/۴۰	۰/۹۰	۰/۹۰	۱۵/۸۲	-۲۱۶/۱۹	<۰/۰۰۰۱
۰/۵۰۰	۰/۰۱	۲/۳۸	۰/۸۸	۰/۸۸	۱۵/۲۸	-۲۱۰/۳۱	<۰/۰۰۰۱
۰/۶۰۰	۰/۱۴	۲/۴۴	۰/۸۶	۰/۸۶	۱۳/۶۶	-۲۱۲/۵۵	<۰/۰۰۰۱
۰/۷۰۰	۰/۴۸	۲/۵۷	۰/۸۵	۰/۸۵	۱۴/۷۰	-۲۲۳/۶۰	<۰/۰۰۰۱
۰/۷۵۰	۰/۴۸	۲/۵۶	۰/۸۳	۰/۸۳	۱۴/۶۹	-۲۳۳/۱۲	<۰/۰۰۰۱
۰/۸۰۰	۰/۵۹	۲/۵۴	۰/۸۲	۰/۸۲	۱۳/۴۲	-۲۴۶/۸۰	<۰/۰۰۰۱
۰/۹۰۰	۰/۸۸	۲/۶۳	۰/۷۹	۰/۷۹	۸/۹۲	-۳۰۱/۲۸	<۰/۰۰۰۱

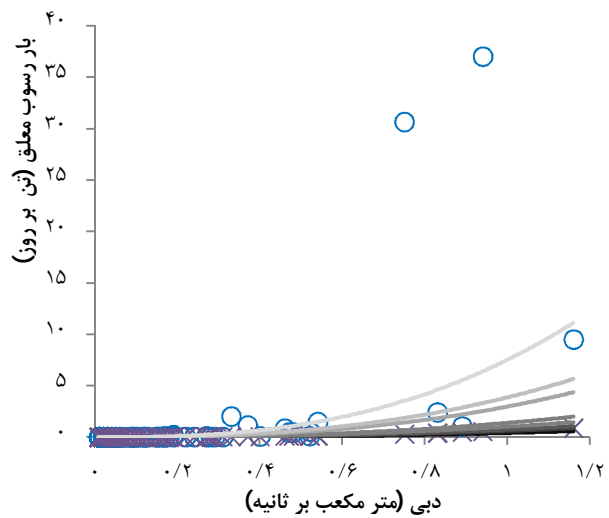
مقادیر $p < 0/05$ از نظر آماری معنی‌دار است.

معمولی حداقل مربعات خطی از مقادیر عرض از مبدأ معادلات خطی رگرسیون کوانتایل کم‌تر می‌باشد. همچنین بررسی این شکل به خوبی مشخص می‌کند که فاصله اطمینان رگرسیون معمولی حداقل مربعات خطی برای مقدار شیب خط و عرض از مبدأ به مراتب بزرگ‌تر از مقادیر متناظر برای معادلات خطی رگرسیون کوانتایل در کوانتایل‌های مختلف مورد بررسی است. همچنین در شکل 3 خطوط و منحنی‌های مختلف به ازای کوانتایل‌های 0/6، 0/7، 0/8، 0/9 و 0/5، 0/4، 0/3، 0/2 و 0/1 ارایه شده است. در این شکل‌ها کوانتایل‌های 0/25 و 0/75 ترسیم نشده است زیرا به علت فاصله کم بین منحنی‌ها و خطوط شکل نسبتاً متراکم می‌گردد که موجب می‌شود تفسیر شکل مشکل‌تر گردد.

در شکل 2 شیب و نیز عرض از مبدأ معادلات خطی رگرسیون کوانتایل و نیز فاصله اطمینان آن‌ها به ازای کوانتایل‌های مختلف ترسیم شده است. بررسی این شکل نشان می‌دهد که شیب معادلات خطی رگرسیون کوانتایل متناظر با کوانتایل‌های مختلف دارای تغییرات محدودی می‌باشد در حالی که مقادیر عرض از مبدأ یا بخش ثابت این معادلات خطی با افزایش کوانتایل‌ها همواره افزایش می‌یابد. همچنین مشخص است که شیب معادلات خطی در کوانتایل‌های مختلف به جز کوانتایل 0/9 همواره از شیب رگرسیون معمولی حداقل مربعات خطی کم‌تر می‌باشد لیکن برای مقادیر عرض از مبدأ شیب خط معادلات خطی رگرسیون کوانتایل از کوانتایل 0/1 تا کوانتایل 0/6 از عرض از مبدأ مدل رگرسیون معمولی حداقل مربعات خطی کم‌تر است و به ازای کوانتایل‌های بالاتر مقدار عرض از مبدأ رگرسیون



شکل 3- منحنی‌های سنجه رسوب به ازای کوانتایل‌های مختلف مورد بررسی



های متغیر وابسته به ازای مقادیر متغیر وابسته می‌باشند. برای نمونه اگر کوانتایل 0/2 و کوانتایل 0/70 را در نظر بگیریم با توجه به جدول 3 به معادلات خطی $Log(Q_s) = -0.22 + 2.47 Log(Q_w)$ و $Log(Q_s) = 0.48 + 2.57 Log(Q_w)$ دست خواهیم یافت. بر اساس حالت تبدیل یافته غیرخطی این معادلات برای کوانتایل 0/2 می‌توان نتیجه گرفت که به ازای دبی جریان برابر با 0/75 مترمکعب بر ثانیه انتظار می‌رود 20 درصد مقادیر بار رسوب معلق رودخانه کم‌تر از 0/29 تن بر روز باشد و 80 درصد آن بیش از مقدار بیان شده باشد. نیز برای کوانتایل 0/7 انتظار می‌رود به ازای دبی 0/75 مترمکعب بر ثانیه 70 درصد مقادیر بار رسوب معلق رودخانه کم‌تر از 1/43 تن بر روز باشد و 30 درصد آن بیش از این مقدار باشد. به طور مشخص می‌توان نتیجه گرفت که انتظار می‌رود که 50 درصد مقادیر بار رسوب معلق به ازای دبی جریان برابر با 0/75 مترمکعب بر ثانیه در بازه بین 0/29 تا 1/43 تن بر روز باشد. باید توجه داشت که چنین تحلیل‌های بار ارزشی تنها با مدل رگرسیون کوانتایل قابل دستیابی است و مدل رگرسیون معمولی حداقل مربعات خطی دارای چنین قابلیت‌هایی نمی‌باشند.

با استفاده از پارامترهای معادله‌های خط برآورد شده به ازای کوانتایل‌های مختلف می‌توان به برآورد توصیفی مناسبی از چولگی داده‌ها به ازای سطوح مختلف متغیر مستقل (دبی جریان رودخانه) دست یافت. در جدول ذیل مقادیر برآورد شده بار رسوب معلق به ازای کوانتایل‌های مختلف متغیر دبی جریان رودخانه در سطح مقدار 0/52 و 1/16 مترمکعب بر ثانیه ارائه شده است. همان‌طور که نتایج جدول 4 مشخص می‌کند داده‌های بار رسوب معلق دارای چولگی به سمت راست می‌باشد.

در نمودارهای ارائه شده منحنی‌های سنجه بار رسوب معلق به ازای کوانتایل‌های مختلف بیان شده در دو فرم منحنی (رابطه غیرخطی توانی) و خطی (استفاده از داده‌ها تحت تبدیل لگاریتمی) ترسیم شده است. نمودارها به ترتیب از پایین به بالا، که به صورت خط‌ها و منحنی‌هایی پر رنگ به کم رنگ ترسیم شده‌اند، بیانگر کوانتایل‌های 0/1 تا 0/9 می‌باشند به عبارت دیگر پایین‌ترین منحنی یا خط (پررنگ‌ترین) مربوط به کوانتایل 0/1 و بالاترین منحنی یا خط (کم‌رنگ‌ترین) مربوط به کوانتایل 0/9 می‌باشد. به طور واضح مشخص است که در فرم منحنی که حاصل کاربرد داده‌ها بدون تبدیل لگاریتمی است کوانتایل‌های پایینی دارای فواصل بسیار نزدیک هستند و به عبارتی تفاوت آن‌ها در نمودار به سختی قابل تشخیص است اما در فرم خطی که حاصل کاربرد داده‌ها تحت تبدیل لگاریتمی است تفاوت بین خطوط کوانتایل‌های مختلف به راحتی قابل تشخیص هستند. مشخص است که خطوط رسم شده در نمودار خطی بیانگر معادلات بیان شده در جدول 3 می‌باشند. بررسی نمودارهای فوق در هر دو فرم توانی و خطی بیانگر آنست که با افزایش مقدار دبی به عنوان متغیر مستقل فاصله بین کوانتایل‌های پایین‌تر یعنی کوانتایل‌های 0/1 تا 0/6 کاهش می‌یابد. این بدان معنی است که با افزایش مقدار متغیر مستقل توزیع داده‌های بار رسوب معلق در سمت چپ توزیع شرطی دارای تراکم بیش‌تری نسبت به سمت راست آن می‌باشد. بررسی نمودار خطی شکل 3 و جدول 3 نشان می‌دهد که بیش‌ترین شیب خط به ترتیب مربوط به کوانتایل‌های 0/9، 0/7، 0/75، 0/8، 0/1، 0/2، 0/25، 0/6، 0/3، 0/4 و 0/5 می‌باشد. به عبارتی متغیر دبی رودخانه بر کوانتایل‌های که در بالا و پایین مقادیر مورد بررسی قرار گرفته‌اند بیش‌تر اثر دارد.

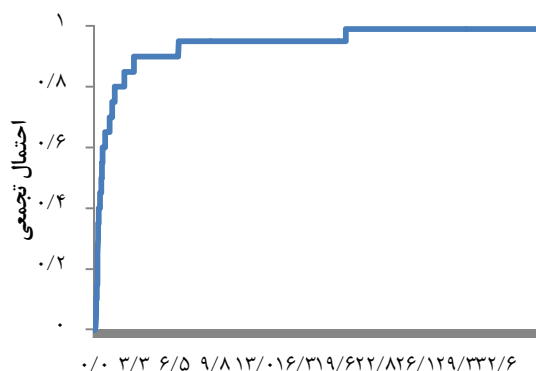
همچنین معادلات بدست آمده ابزار مناسبی برای برآورد کوانتایل‌ها

جدول 4-مقادیر کوانتایل‌های مختلف بر رسوب معلق به ازای مقادیر مشخص دبی جریان رودخانه کوانتایل‌ها

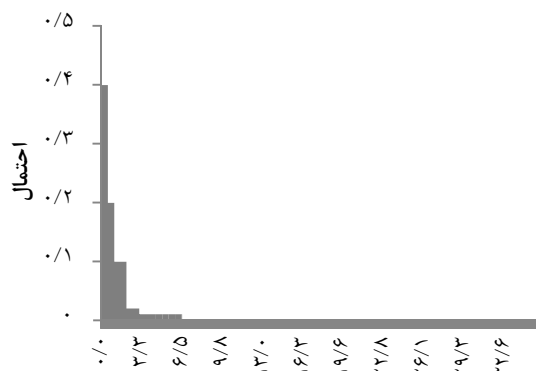
دبی (متر مکعب بر ثانیه)	0/1	0/2	0/3	0/4	0/5	0/6	0/7	0/8	0/9
0/52	0/08	0/12	0/13	0/16	0/22	0/28	0/56	0/74	1/25
1/16	0/59	0/86	0/89	1/07	1/46	2/00	4/37	5/66	11/10

توزیع احتمال و تابع توزیع تجمعی تجربی داده‌ها بار رسوب معلق در شکل 4 ارائه شده است. بدیهی است فرآیند مشابهی می‌تواند برای استخراج تابع توزیع احتمال و تابع توزیع تجمعی تجربی داده‌ها بار رسوب معلق به ازای تمام مقادیر دلخواه دبی جریان رودخانه انجام شود.

با توجه به اینکه با استفاده از رگرسیون کوانتایل می‌توان مقادیر بار رسوب معلق رودخانه را به ازای کوانتایل‌های مختلف برآورد نمود این مفهوم کاربردی در برآورد تابع توزیع احتمال و تابع توزیع تجمعی تجربی داده‌ها بار رسوب معلق به ازای سطوح مختلف مقادیر دبی جریان به کار گرفته شد. با در نظر گرفتن مقدار 0/75 مترمکعب بر ثانیه برای داده‌های ثبت شده در ایستگاه هیدرومتری النگ دره، تابع



بار رسوب معلق (تن بر روز)



بار رسوب معلق (تن بر روز)

شکل 4- تابع توزیع احتمال و توزیع تجمعی تجربی بار معلق رسوب به ازای دبی 0/75 مترمکعب بر ثانیه

نتیجه گیری

کاربرد رگرسیون معمولی حداقل مربعات خطی و کوانتایل در تحلیل منحنی سنجه رسوب بار معلق به طور مشخص منجر به نتایج ذیل شده است.

-نتایج کاربرد رگرسیون معمولی حداقل مربعات در برآورد منحنی سنجه رسوب نشان می‌دهد این روش نمی‌تواند به خوبی پراکندگی داده‌های ثبت شده را بازسازی کند و نیز برخی فرضیات لازم برای این روش مانند ثابت بودن واریانس و نرمال بودن باقی‌مانده‌ها برقرار نمی‌باشد.

-بررسی ضریب تبیین معادلات در کوانتایل‌های مختلف بیانگر آنست که این مشخصه آماری با افزایش مقادیر کوانتایل‌ها کاهش می‌یابد. نیز بررسی معیار اطلاعات آکائیک و میانگین قدر مطلق درصد خطا بیانگر دقت بالاتر برازش مدل‌های خطی در کوانتایل‌های بالا و پایین است درحالی‌که کوانتایل‌های میانی بر اساس مشخصه‌های بیان شده دارای دقت برآزش کم‌تری می‌باشند.

-تحلیل‌های معادله‌های خطوط رگرسیون به ازای کوانتایل‌های مختلف به وضوح بیان می‌کند که تأثیر دبی جریان بر بار رسوب معلق رودخانه در کوانتایل‌های مختلف متفاوت می‌باشد. در این مطالعه بار رسوب معلق رودخانه در کوانتایل‌های حدی نسبت به کوانتایل‌های میانی به میزان بیش‌تری تحت‌تأثیر دبی رودخانه می‌باشند. همچنین در بین کوانتایل‌های حدی بار رسوب معلق در کوانتایل‌های بالایی نسبت به کوانتایل‌های پایینی تأثیرپذیری بیش‌تری از دبی رودخانه دارند.

- بررسی فاصله خطوط رگرسیونی کوانتایل‌های مختلف بار رسوب معلق می‌تواند فاصله کوانتایل‌ها را از یکدیگر در سطوح مختلف متغیر دبی جریان رودخانه نشان دهد بنابراین با استفاده از رگرسیون کوانتایل می‌توان در هر سطح از مقدار متغیر جریان رودخانه

به توصیفی از چولگی متغیر وابسته دست یافت. بررسی تغییرات فاصله بین خطوط و منحنی‌های کوانتایل‌های مختلف نشان داد که با افزایش دبی جریان رودخانه داده‌های بار رسوب معلق به مقدار بیش‌تری به سمت راست چولگی نشان می‌دهند.

-معادلات استخراج شده از رگرسیون کوانتایل ابزار مناسب و کارایی برای برآورد کوانتایل‌های متفاوت و متنوع بار رسوب معلق به ازای تمام مقادیر دبی جریان رودخانه می‌باشد. این کاربرد موجب می‌شود که به ازای یک مقدار خاص از دبی جریان بتوان تابع توزیع احتمال و تابع توزیع تجمعی تجربی داده‌ها بار رسوب معلق رودخانه را برآورد نمود که ابزار بسیار ارزشمندی در برآورد ریسک و عدم قطعیت در برآورد بار رسوب معلق رودخانه می‌باشد.

- نتایج نشان می‌دهد که رگرسیون کوانتایل می‌تواند جایگزین شایسته‌ای برای رگرسیون معمولی حداقل مربعات در تحلیل منحنی سنجه بار معلق رسوب باشد و با مدل‌سازی تمام قسمت‌های توزیع بار رسوب معلق و ارایه شکل رگرسیونی کامل نسبت به حالت متداول منجر به ارایه نتایج ارزشمندتری می‌گردد.

-نتایج این تحقیق نشان می‌دهد که برای داده‌های دارای چولگی شدید، استفاده از رگرسیون کوانتایل در مقایسه با رگرسیون معمولی یافته‌های بهتر و کامل‌تری را ارایه می‌دهد که این مورد با نتایج اشاره شده در کارهای (Koenker., 2005; Yu et al., 2003; Eideand Showalter) مطابقت دارد.

رگرسیون کوانتایل علاوه بر نقاط تکیه‌بها ناشاره‌شده، دارای محدودیت‌هایی نیز می‌باشد که از آنجمله می‌توان به غیر واقعی بودن خطی فرض کردن کوانتایل‌ها اشاره نمود که تعیین متغیرهای مستقل مناسب را با مشکل مواجه می‌سازد. محدودیت دیگر این است که در برآورد کوانتایل‌های متعدد، احتمال تقاطع کوانتایل‌ها وجود دارد، به ویژه زمانی که تعداد متغیرهای مستقل زیاد و یا تعداد داده‌ها کم باشد

- Cimen, M. 2008. Estimation of daily suspended sediments using support vector machines. *Hydrological Sciences Journal*. 53.3: 656-666.
- Cohn, T.A., Caulder, D.L., Gilroy, E.J., Zynjuk, L.D., Summers, R.M. 1992. The validity of a simple statistical model for estimating fluvial constituent loads: An empirical study involving nutrient loads entering Chesapeake Bay. *Water Resources Research*. 28.9: 2353-2363.
- Cohn, T.A., DeLong, L.L., Gilroy, E.J., Hirsch, R.M., Wells, D.K. 1989. Estimating constituent loads. *Water Resources Research*. 25.5: 937-942.
- Cozzoli, F., Bouma, T.J., Ysebaert, T., Herman, P.M.J. 2013. Application of non-linear quantile regression to macrozoobenthic species distribution modelling: comparing two contrasting basins. *Marine Ecology Progress Series*. 475: 119-133.
- Eide, E., Showalter, M.H. 1998. The effect of school quality on student performance: A quantile regression approach. *Economics Letters*. 58.3: 345-350.
- Ferguson, R.I. 1986. River loads underestimated by rating curves. *Water Resources Research*. 22.1: 74-76.
- Gaglianone, W.P., Lima, L.R., Linton, O., Smith, D.R. 2012. Evaluating value-at-risk models via quantile regression. *Journal of Business and Economic Statistics*. 29.1: 150-160.
- Güven, A., Kisi, Ö. 2011. Estimation of suspended sediment yield in natural rivers using machine-coded linear genetic programming. *Water resources management*. 25.2: 691-704.
- Hicks, D.M., Gomez, B., Trustrum, N.A. 2000. Erosion thresholds and suspended sediment yields, Waipaoa River basin, New Zealand. *Water Resources Research*. 36.4: 1129-1142.
- Hirschi, M., Seneviratne, S.I., Alexandrov, V., Boberg, F., Boroneant, C., Christensen, O.B., Stepanek, P. 2011. Observational evidence for soil-moisture impact on hot extremes in southeastern Europe. *Nature Geoscience*. 4.1: 17-21.
- Huang, M.Y.F., Montgomery, D.R. 2013. Altered regional sediment transport regime after a large typhoon, southern Taiwan. *Geology*. 41.12: 1223-1226.
- Jagger, T.H., Elsner, J.B. 2009. Modeling tropical cyclone intensity with quantile regression. *International Journal of Climatology*. 29.10: 1351-1361.
- Jagger, T.H., Elsner, J.B. 2009. Modeling tropical cyclone intensity with quantile regression. *International Journal of Climatology*. 29.10: 1351-1361.
- Jain, S.K. 2001. Development of integrated sediment (Bjrnar Bremnes., 2004).
- منابع**
- انصاری، م.ت.، بامنی مقدم، م.، خوش‌گویان‌فرد، ع.، سام آرام، ع. 1385. کاربرد رگرسیون چندک در تحلیل سلامت روانی. *رفاه اجتماعی*. 20.5: 49-60.
- کالوندی، س.م.، خدائشناس، س.م.، قهرمان، ب.، طهماسبی، ر.، بوستانی، آ. 1389. آنالیز روش‌های مختلف منحنی سنجه در برآورد رسوب ورودی به سدها (مطالعه موردی: سد دوستی). *مهندسی آبیاری و آب*. 1.1: 10-20.
- کیا، ع.، عمادی، ع. 1392. مقایسه روش‌های مختلف رگرسیون آماری در برآورد بار رسوب معلق درازمدت سالانه (مطالعه موردی: نابل-رود). *مدیریت حوضه و آبخیز*. 8.4: 15-27.
- Afan, H.A., El-Shafie, A., Yaseen, Z.M., Hameed, M.M., Mohtar, W.H.M.W., Hussain, A. 2015. ANN based sediment prediction model utilizing different input scenarios. *Water Resources Management*. 29.4: 1231-1245.
- Alagidede, P., Panagiotidis, T. 2012. Stock returns and inflation: Evidence from quantile regressions. *Economics Letters*. 117.1: 283-286.
- Alp, M., Cigizoglu, H.K. 2007. Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. *Environmental Modelling and Software*. 22.1: 2-13.
- Asselman, N.E.M. 2000. Fitting and interpretation of sediment rating curves. *Journal of Hydrology*. 234.3: 228-248.
- Austin, M. 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecological modeling*. 200.1: 1-19.
- Aytek, A., Kisi, O. 2008. A genetic programming approach to suspended sediment modelling. *Journal of Hydrology*. 351.3: 288-298.
- Barbosa, S.M., Scotto, M.G., Alonso, A.M. 2011. Summarising changes in air temperature over Central Europe by quantile regression and clustering. *Natural Hazards and Earth System Sciences*. 11.12: 3227-3233.
- Bjrnar Bremnes, J. 2004. Probabilistic forecasts of precipitation in terms of quantiles using NWP model output. *Monthly Weather Review*. 132.1: 338-347.
- Cade, B.S., Noon, B.R. 2003. A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment*. 1.8: 412-420.

- Nagy,H.M., Watanabe,K.A.N.D., Hirano,M. 2002. Prediction of sediment load concentration in rivers using artificial neural network model. *Journal of Hydraulic Engineering*. 128.6: 588-595.
- Partal,T., Cigizoglu,H.K. 2008. Estimation and forecasting of daily suspended sediment data using wavelet-neural networks. *Journal of hydrology*. 358.3: 317-331.
- Phillips,J.M., Webb,B.W., Walling,D.E., Leeks,G.J.L. 1999. Estimating the suspended sediment loads of rivers in the LOIS study area using infrequent samples. *Hydrological processes*. 13.7: 1035-1050.
- Rai,R.K., Mathur,B.S. 2008. Event-based sediment yield modeling using artificial neural network. *Water Resources Management*. 22.4: 423-441.
- Salarijazi,M., Abdolhosseini,M., Ghorbani,Kh., Eslamian,S. 2016. Evaluation of Quasi-Maximum Likelihood and Smearing Estimator to Improve Sediment Rating Curve Estimation. *International Journal of Hydrology Science and Technology*. 6.4: 359-370
- Tarras-Wahlberg,N.H., Lane,S.N. 2003. Suspended sediment yield and metal contamination in a river catchment affected by El Niño events and gold mining activities: the Puyango River Basin, southern Ecuador. *Hydrological Processes*. 17.15: 3101-3123.
- Walling,D.E. 1977. Assessing the accuracy of suspended sediment rating curves for a small basin. *Water Resources Research*. 13.3: 531-538.
- Wang,P., Linker,L.C. 2008. Improvement of regression simulation in fluvial sediment loads. *Journal of Hydraulic Engineering*. 134.10: 1527-1531.
- Wang,Y.G., Tian,T. 2013. Sediment concentration prediction and statistical evaluation for annual load estimation. *Journal of Hydrology*. 482: 69-78.
- Warrick,J.A. 2015. Trend analyses with river sediment rating curves. *Hydrological Processes*. 29.6: 936-949.
- Yang,G., Chen,Z., Yu,F., Wang,Z., Zhao,Y., Wang,Z. 2007. Sediment rating parameters and their implications: Yangtze River, China. *Geomorphology*. 85.3: 166-175.
- Yu,K., Lu,Z., Stander,J. 2003. Quantile regression: applications and current research areas. *Journal of the Royal Statistical Society: Series D (The Statistician)*. 52.3: 331-350.
- rating curves using ANNs. *Journal of hydraulic engineering*. 127.1: 30-37.
- Kao,S., Lee,T., Milliman,J.D. 2005. Calculating highly fluctuated suspended sediment fluxes from mountainous rivers in Taiwan. *Terrestrial Atmospheric and Oceanic Sciences*. 16.3: 653.
- Kii,OÖ. 2004. Daily suspended sediment modelling using a fuzzy differential evolution approach/Modélisation journalière des matières en suspension par une approche d'évolution différentielle floue. *Hydrological sciences journal*. 49.1: 183-197.
- Kisi,O. 2005. Suspended sediment estimation using neuro-fuzzy and neural network approaches/Estimation des matières en suspension par des approches neuro-floues et à base de réseau de neurones. *Hydrological Sciences Journal*. 50.4: 683-696.
- Kitsikoudis,V., Sidiropoulos,E., Hrisanthou,V. 2014. Machine learning utilization for bed load transport in gravel-bed rivers. *Water resources management*. 28.11: 3727-3743.
- Koenker,R. 2005. *Quantile regression* (No. 38). Cambridge university press.
- Krishnaswamy,J., Richter,D.D., Halpin,P.N., Hofmockel,M.S. 2001. Spatial patterns of suspended sediment yields in a humid tropical watershed in Costa Rica. *Hydrological Processes*. 15.12: 2237-2257.
- Lafdani,E.K., Nia,A.M., Ahmadi,A. 2013. Daily suspended sediment load prediction using artificial neural networks and support vector machines. *Journal of Hydrology*. 478: 50-62.
- Lohani,A.K., Goel,N.K., Bhatia,K.S. 2007. Deriving stage-discharge-sediment concentration relationships using fuzzy logic. *Hydrological Sciences Journal*. 52.4: 793-807.
- McBean,E.A., Al-Nassri,S. 1988. Uncertainty in suspended sediment transport curves. *Journal of Hydraulic Engineering*. 114.1: 63-74.
- Meligkotsidou,L., Vrontos,I.D., Vrontos,S.D. 2009. Quantile regression analysis of hedge fund strategies. *Journal of Empirical Finance*. 16.2: 264-279.
- Morehead,M.D., Syvitski,J.P., Hutton,E.W., Peckham,S.D. 2003. Modeling the temporal variability in the flux of sediment from ungauged river basins. *Global and Planetary Change*. 39.1: 95-110.

Application of Quantile Regression in Analysis of Suspended Sediment Load Rating Curve

M. Salarijazi^{1*}

Received: Jul.31, 2016

Accepted: Oct.19, 2017

Abstract

The sediment rating curve is the most widely used method to estimate river suspended sediment load that shows the relation between conditional mean of suspended sediment load and river discharge using Ordinary Least Square (OLS) regression and is applied to estimate suspended sediment load as a function of the river discharge. The OLS regression model is sensitive to outliers and when its assumptions including assumptions related to the residuals analysis are not satisfied, it is not acceptable. Quantile regression is a statistical method that can be used to overcome these limitations in sediment rating curve analysis. In this study, quantile regression method was used to estimate sediment rating curve using data from Alang-Darreh hydrometry station in Golestan province (recorded period years 1987-2001) and the results were compared with the conventional OLS regression method. The results show that application of OLS regression in sediment rating curve analysis led to bias estimation while quantile regression without OLS regression's limiting assumptions can be appropriately show the effect of river discharge on different quantiles of suspended sediment load distributions especially in upper and lower tail. In addition, it was found that the magnitude of impact of river discharge belonging to upper, lower and central quantiles of suspended sediment load respectively and with increase in river discharge, the suspended sediment load shows more skewness to the right. Moreover, the quantile regression concept is presented as a very important tool to extract the probability density and cumulative distribution functions of suspended sediment load for specific values of river discharge.

Keywords: Ordinary Least Square Regression, Sediment Rating Curve, Quantile Regression

1- Assistant Professor, Department of Water Engineering, Gorgan University of Agricultural Sciences and Natural Resources.
(*- Corresponding Author Email: meysam.salarijazi@gau.ac.ir, meysam.salarijazi@gmail.com)