

دبی رسوبات معلق متناظر، عمق آب و سرعت جریان است. دبی رسوب بستر تنها متغیر خروجی بوده است. نتایج کمی بهترین درخت تصمیم نشان داد که مقدار  $MAE=2134$ ،  $RMSE=2668$ ،  $R=89\%$  می باشد که مقایسه نتایج درختان رگرسیونی با فرمول های تجربی (اینشتین-براون، میر پتر مولر، راتنر، کالینسکی، شوکلیچ و ...)، حاکی از دقت بیش تر درخت تصمیم در برآورد رسوبات بستر است.

واژه های کلیدی: بار بستر، برآورد رسوب، درختان رگرسیونی، روش های تجربی.

#### مقدمه

هرساله هزاران تن مواد جامد از سطح حوزه های آبخیز شسته شده و از محلی به محل دیگر انتقال می یابند، به طوری که در سال بیش از ۲۰ میلیارد تن رسوب توسط رودخانه های جهان انتقال و در آب های ساکن ته نشین می شود. در ایران بالغ بر ۱۰۰ میلیون متر مکعب از گنجایش سدهای مخزنی در سال به خاطر رسوب گذاری کاسته می شود [۸]. ورود رسوب به رودخانه ها باعث ایجاد بسیاری از مشکلات از جمله کاهش ظرفیت ذخیره مخازن از طریق رسوب گذاری و افزایش گل آلودگی در سیستم های توزیع آب می شود. در بسیاری از حوضه ها، وقایع بارشی کوتاه مدت و شدید، دلیل بخش بزرگی از انتقال رسوب کل حوضه است [۲]. به این دلیل که بررسی پدیده رسوب و برآورد رسوب حمل شده توسط رودخانه اهمیت خاصی داشته و گسترش شیوه های نوین تخمین رسوب که دارای سهولت کاربرد باشند در این میان نقش مهمی خواهد داشت. استفاده از روش درخت تصمیم گیری به عنوان روشی نو و ابزاری قدرتمند و قابل اعتماد می تواند یاری گر مسئولان و مهندسان امور آب برای برنامه ریزی و استفاده بهینه از منابع آبی کشور باشد. در پژوهش های مختلف و در زمینه های متعددی از روش درخت تصمیم گیری استفاده شده است که از آن جمله می توان به موارد زیر به صورت موضوعی اشاره کرد. طالبی و اکبری [۱۱] به بررسی کارایی مدل درختان تصمیم گیری در برآورد رسوبات معلق رودخانه گل گل حوضه سد ایلام پرداختند. جهت ارزیابی دقت و صحت نتایج این مدل از معیارهای آماری  $RMSE^5$ ،  $BIAS^6$  و  $R^2$

## برآورد بار رسوب بستر با استفاده از روش درخت تصمیم گیری و مقایسه با روش های تجربی

نصرت الله امانیان<sup>۱\*</sup>، علی طالبی<sup>۲</sup>، پریسا حجه بخش<sup>۳</sup> و سارا پرویزی<sup>۴</sup>  
تاریخ دریافت: ۹۹/۰۳/۰۲ تاریخ پذیرش: ۹۹/۰۸/۱۱

#### چکیده

برآورد مقدار مواد رسوبی که یک جریان مشخص قادر به حمل آن است یکی از موضوعات اصلی تحقیقات رسوب می باشد که در بسیاری از پروژه های مهندسی هم چون برنامه ریزی و طراحی منابع ذخیره آب، مورفولوژی و تغییرات بستر رودخانه برآورد رسوب سالیانه برای آبگیرهای رودخانه، طراحی و نگهداری کانال های آبیاری پایدار، حفاظت سواحل، لایروبی کانال ها و غیره حائز اهمیت می باشد. بررسی پدیده رسوب دهی و برآورد رسوب حمل شده توسط رودخانه اهمیت خاصی دارد و گسترش شیوه های نوین تخمین رسوب که دارای سهولت کاربرد باشند در این میان نقش مهمی خواهد داشت. در این تحقیق سعی بر آن است با استفاده از شیوهی درختان تصمیم گیری رگرسیونی، بار رسوب بستر برآورد شود و سپس بار رسوب بستر برآورد شده با چند روش تجربی رایج (مانند روابط میر پتر مولر، شیلدز، اینشتین براون، شوکلیچ و ...) مقایسه گردد تا میزان کارایی این روش در مقایسه با شیوه های رایج برآورد رسوب مشخص گردد. داده های ایستگاه پل یزدکان واقع بر رودخانه قطورچای در استان آذربایجان غربی که رودخانه ای کوهستانی با بستر شنی است و حجم رسوبات بالایی دارد، استفاده شد. بعد از بررسی صحت و پالایش داده ها، طول دوره آماری برابر با ۸ سال (۱۳۸۸ - ۱۳۸۰) و ۷۶ داده در نظر گرفته شد. از این میزان داده ها ۸۰ درصد آن برای آموزش و ۲۰ درصد آن برای آزمون مورد استفاده قرار گرفته است. بهترین ترکیب ورودی که بهترین انطباق نتایج را نشان داد، شامل دبی آب متناظر (هم زمان با برداشت رسوب)،

۱- استادیار گروه مهندسی عمران، دانشکده عمران، دانشگاه یزد (نویسنده مسئول)  
Email: Namanian@yazd.ac.ir

۲- استاد گروه آبخیزداری، دانشکده منابع طبیعی و کویر شناسی، دانشگاه یزد

۳- فارغ التحصیل کارشناسی ارشد عمران آب، دانشگاه یزد.

۴- دانشجوی دکتری آبخیزداری، دانشکده منابع طبیعی و کویر شناسی، دانشگاه یزد

5. Root Mean Square Error

6. Bias Error

7. Mean Absolute Error

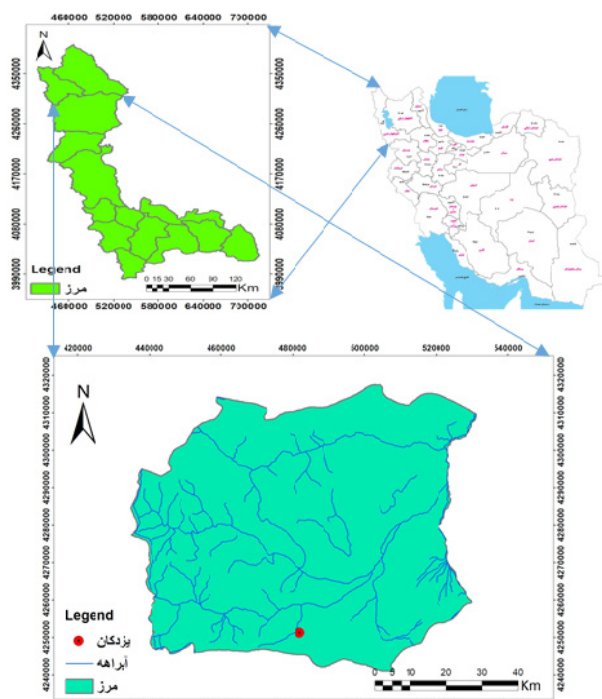
MAE مورد استفاده قرار گرفته است. نتایج نشان داده که به طور کلی مدل درخت تصمیم‌گیری در مقایسه با روش مرسوم منحنی سنج رسوب تطابق بسیار بیشتری با مقادیر اندازه‌گیری شده داشته است. حلییان و همکاران [۶] به ارزیابی کارایی مدل درختان تصمیم‌گیری در برآورد رسوبات معلق در حوضه‌ی میمه پرداختند. داده‌های مورد استفاده شامل دبی رسوب، دبی آب، بارش و دبی‌های روزانه بود. نتایج بدست آمده، مدل درخت تصمیم‌گیری نتایج قابل قبولی را در جهت شبیه‌سازی بار معلق در ایستگاه مورد مطالعه ارائه کرده است، به‌طوری که درخت تصمیم‌گیری نسبت به منحنی سنج رسوب با خطای کم‌تری رسوبات را برآورد کرده است. ثقیبان و همکاران [۱۰] با استفاده از روش درخت تصمیم‌گیری برای پیش‌بینی کیفیت آب‌های زیرزمینی براساس نمودار ایالات متحده آمریکا (USSL) در استان اردبیل پرداختند هم‌چنین برای تعیین پارامترهای مهم طبقه‌بندی کیفیت آب زیرزمینی از  $PCA^2$  استفاده نمودند. نتایج نشان داد طبقه‌بندی کیفیت آب زیرزمینی با درخت تصمیم کارآمدتر از PCA است. برزگری و همکاران [۱] به پیش‌بینی رسوبات معلق با استفاده از درخت‌های تصمیم‌گیری، شبکه عصبی مصنوعی و مدل اصلاح رسوب در ۱۰ ایستگاه هیدرومتری واقع در استان لرستان پرداختند. نتایج نشان داد که دقت شبکه عصبی مخصوصاً در دبی‌های بالا بیش‌تر از دو مدل دیگر است. دستورانی و همکاران [۴] برای شبیه‌سازی رواناب در حوزه زاینده‌رود از الگوریتم  $CART^3$ ، الگوریتم  $M5$ ،  $SVM^4$ ،  $ANN^5$  استفاده کردند. نتایج نشان داد مقدار  $RMSE$  ارائه شده توسط  $SVM$ ، الگوریتم  $CART$ ، الگوریتم  $M5$  و شبکه عصبی به ترتیب ۲/۴، ۶/۷۱، ۳/۲، ۳/۰۴ است. دقیق‌ترین نتیجه مربوط به  $SVM$  می‌باشد که کم‌ترین مقدار  $RMSE$  را دارد. طهماسبی نسب و دهقانی [۱۲] با استفاده از داده‌های اندازه‌گیری شده صحرایی از دو رودخانه کورا و آرا واقع در کشور مالزی به بررسی دقت ۱۲ معادله‌ی نرخ انتقال بار بستر پرداختند. نتایج این پژوهش نشان داد که معادله‌های میر-پیتر و مولر، جولین و ویلسون به ترتیب نتایج مناسب‌تری نسبت به سایر معادلات ارائه می‌دهند و از بین معادلات برآورد بار بستر، معادله میر-پیتر و مولر که در آن ۶۱/۴ درصد از داده‌ها در دامنه نسبت ناجوری بین ۰/۵ تا ۲ قرار می‌گیرند بهترین نتایج را برای رودخانه‌های محدوده‌ی مطالعاتی ارائه می‌دهند.

در مطالعات مشابه بیش‌تر از روش‌های تجربی استفاده شده است اما در این تحقیق سعی بر آن است تا با استفاده از شیوه‌ی درختان تصمیم‌گیری، بار رسوب بستر برآورد گردد و سپس با بار رسوب بستر برآورد شده با چند روش تجربی رایج و داده‌های اندازه‌گیری شده بار رسوب مقایسه گردد تا خطای این روش در قیاس با شیوه‌های رایج برآورد رسوب مشخص گردد.

1. United States Salinity Laboratory
2. Principal component analysis
3. Classification and Regression Tree
4. Support Vector Machines
5. Artificial Neural Network

### منطقه مطالعاتی

رودخانه قطور در حوزه آبریز ارس می‌باشد و حوزه آبخیز آن با مساحت ۳۴۷۱/۹ کیلومتر مربع از چین‌خوردگی‌های شمال غربی سلسله جبال زاگرس بوجود آمده به‌طوری که اراضی محدوده‌ای دارای شیب تند بوده و کمتر اراضی مسطح در این منطقه مشاهده شده است. مساحت حوزه رودخانه قطور تا میله مرزی ۲۴۸ (محل ورودی رودخانه قطور به ایران) در ارتفاع ۲۹۰ متری، ۸۱۰ کیلومتر مربع و طول آبراهه‌های اصلی در ترکیه ۱۱۵ کیلومتر می‌باشد. منطقه مطالعاتی در شکل (۱) نشان داده شده است.



شکل ۱: شمایی از منطقه مورد مطالعه

### روش کار

جهت برآورد رسوب بار بستر با استفاده از درخت تصمیم رگرسیونی در این پژوهش از توابع نرم‌افزار MATLAB استفاده گردیده است. در مرحله اول می‌بایست داده‌ها به دو قسمت آموزش و آزمون تقسیم شوند. سپس هرکدام از این دو قسمت به صورت جداگانه از محیط اکسل فراخوانی شده و درخت تصمیم با دستور  $T = \text{classregtree}(X, Y)$  بر روی مقادیر آموزشی ساخته می‌شود. در واقع دستور  $T = \text{classregtree}(X, Y)$  درخت تصمیمی با نام  $T$ ، برای پیش‌بینی  $Y$  به‌عنوان تابعی از ستونهای  $X$  می‌سازد.  $X$  یک ماتریس  $n \times m$  از مقادیر پیش‌بینی کننده است. اگر  $Y$  برداری  $n$  تایی از مقادیر پاسخ باشد، این تابع، رگرسیون انجام می‌دهد و اگر  $Y$  غیر از این باشد، تابع به‌جای رگرسیون، دسته‌بندی انجام می‌دهد. به هر حال  $T$  یک درخت باینری است و هر گره بر اساس مقادیر ستونهای

X منشعب می‌شود.

با استفاده از دستور View(T) می‌توان درخت ایجاد شده را به صورت گرافیکی مشاهده نمود. و با استفاده از دستور  $yfit=eval(T,X)$  مقادیر پیش‌بینی درخت برای ماتریس ورودی X در بردار  $yfit$  قرار داده خواهد شد. برای هرس کردن درخت، می‌توان از دستور  $Tprune=prune(T)$  بهره برد که درخت Tprune هرس شده را ایجاد می‌کند، البته دستورات دیگری نیز وجود دارند که سطح هرس را نیز می‌توان به عنوان ورودی تعیین نمود. بهتر است بدین طریق عمل گردد که از هر پنج داده، به ترتیب چهار مورد را برای آموزش و یکی را برای آزمون به کار رود، در این صورت توزیع داده‌های آموزشی و آزمایشی مثل هم خواهد بود. بنابراین به طور معمول ۸۰ درصد داده‌ها برای آموزش و ۲۰ درصد برای آزمون مورد استفاده قرار گرفته است طالبی و اکبری [۱۱].

### مواد بار کف بازه پل یزدکان

داده‌های مورد استفاده ما در این تحقیق، بار رسوب بستر ( $Q_b$ )، بار رسوب معلق ( $Q_{sus}$ )، عمق آب (Y)، سرعت جریان (V) و دبی جریان آب ( $Q_{flow}$ ) می‌باشد که همگی به طور هم‌زمان در روزهای مختلف سال اندازه‌گیری شده‌اند. این داده‌ها مربوط به ایستگاه پل یزدکان بر روی رودخانه قطورچای و برای دوره آماری سال‌های ۸۰

تا ۸۸ می‌باشد. تعداد داده‌ها جهت مدل‌سازی ۷۶ برداشت پس از پالایش داده‌ها می‌باشد. در این بازه ۱۷ نمونه مواد بار کف توسط کارشناسان سازمان آب منطقه‌ای تهیه شده است. مواد بار کف در محدوده اندازه ۱۰ تا ۹۰ درصد ذرات بین ۰/۷ تا ۲۰ میلی‌متر قرار داشته و متوسط اندازه ذرات ۳/۷ میلی‌متر می‌باشد. جنس مواد آبرفتی از نوع کوارتز با چگالی ۲/۶۵ می‌باشد. براساس طبقه بندی انجام شده، دانه‌بندی مواد بار کف در گروه شن متوسط قرار دارد. اندازه ذرات مواد بستر و بار کف پل یزدکان در جدول (۱) لحاظ گردیده است (نتایج آزمایشگاه).

### الف-برآورد بار بستر با روش درخت تصمیم‌گیری

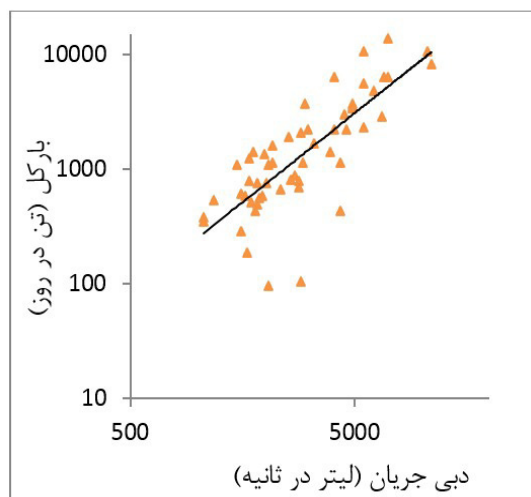
با بررسی داده‌های موجود از بار کف، معلق و کل، نتایج زیر به دست آمد: بطور طبیعی افزایش دبی جریان، افزایش بار معلق و بار کل را به همراه داشته است (شکل‌های ۲ و ۳). اما این روند افزایشی در بار کف بسیار ملایم‌تر دیده می‌شود (شکل ۴). هم‌چنین با افزایش دبی، نسبت بار کف به کل بار رسوب، کم و نسبت بار معلق به بار کل، زیاد شده است (شکل ۵). به عبارت دیگر می‌توان این‌گونه برداشت کرد که با افزایش دبی جریان، بار معلق بیش‌تر از بار کف افزایش می‌یابد.

### بحث و نتایج

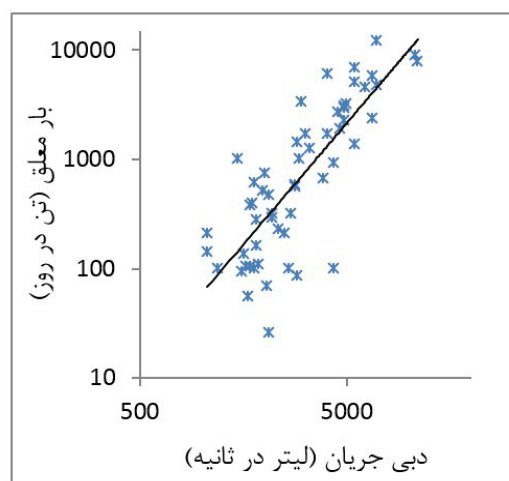
جدول ۱: اندازه ذرات مواد بستری و بار کف بازه پل

یزدکان-رودخانه قطور

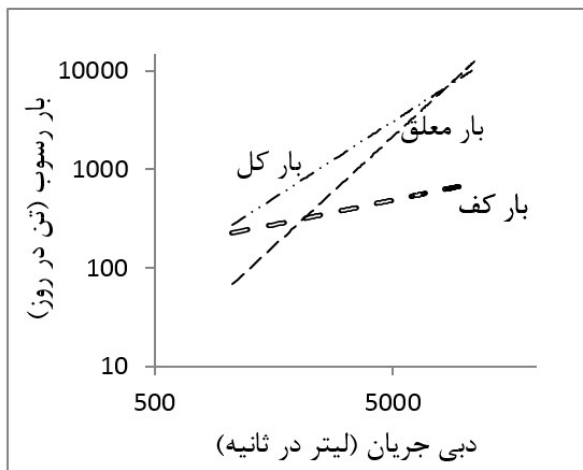
| اندازه ذرات  | D <sub>10</sub> | D <sub>16</sub> | D <sub>35</sub> | D <sub>50</sub> | D <sub>60</sub> | D <sub>65</sub> | D <sub>84</sub> | D <sub>90</sub> | D <sub>100</sub> |
|--------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|------------------|
| واحد         | (mm)            | (mm)            | (mm)            | (mm)            | (mm)            | (mm)            | (mm)            | (mm)            | (mm)             |
| مواد زیرسطحی | ۰/۶۵            | ۰/۹             | ۲               | ۳/۷             | ۶               | ۶/۸             | ۱۴/۵            | ۲۲              | ۳۹               |
| بار کف       | ۰/۷۱            | ۰/۹۵            | ۲/۲۵            | ۳/۷             | ۶               | ۶/۴             | ۱۳/۷            | ۲۰              | ۳۰               |



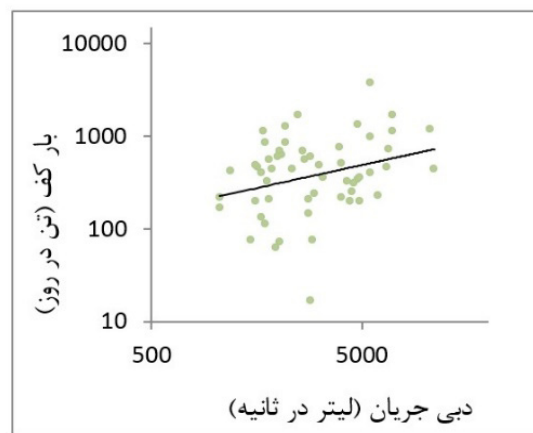
شکل ۳: افزایش بار کف با افزایش دبی جریان



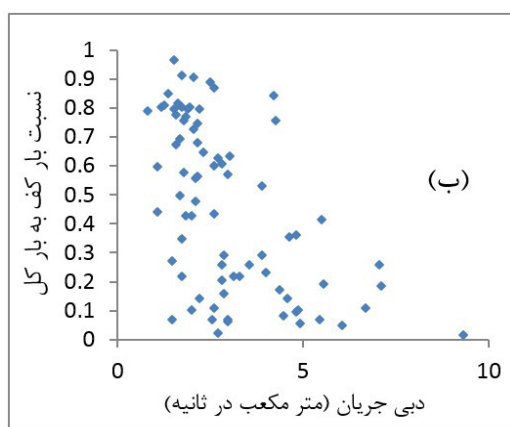
شکل ۲: افزایش بار معلق با افزایش دبی جریان



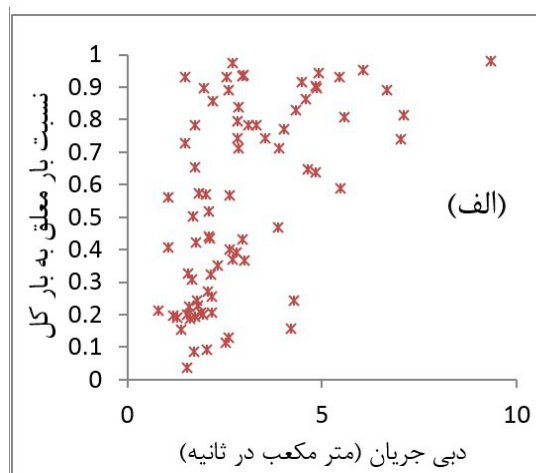
شکل ۵: خطوط برازش شده روی بار رسوب کف، معلق و کل



شکل ۴: افزایش بار کف با افزایش دبی جریان



ب- برآورد بار بستر با روش‌های تجربی



شکل ۶: (الف) نسبت بار معلق به بار کل

کالینسکی<sup>۴</sup>، شوکلیچ<sup>۵</sup>، شیلدرز<sup>۶</sup> و دوبوی<sup>۷</sup>. برای اینکه دامنه تغییرات مقادیر بار رسوب بستر به خوبی نمایش داده شود، محورهای مختصات بصورت لگاریتمی انتخاب شده‌است. نیمساز ربع اول نمودارها به عنوان خط تطابق کامل<sup>۸</sup> معرفی شده و هر چقدر نقاط در دستگاه مختصات به این خط نزدیک‌تر باشند، نشانگر درجه مطابقت خوب مقادیر اندازه‌گیری شده و مقادیر محاسباتی است و پراکندگی بیش‌تر نقاط در طرفین این خط، بیانگر اختلاف بین مقادیر محاسباتی و برآورد شده خواهد بود. نمودارهای شکل (۷) کیفیت برآورد روش‌های مختلف را در مقایسه باهم نشان می‌دهند.

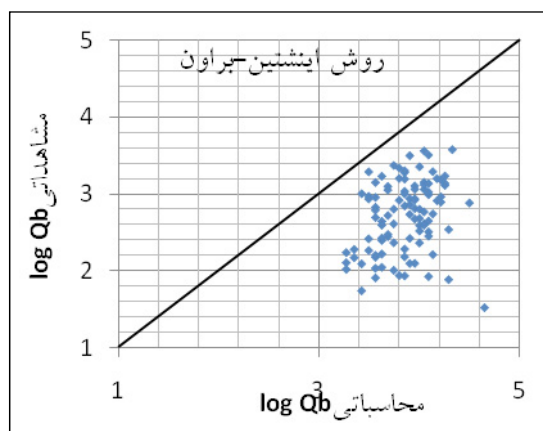
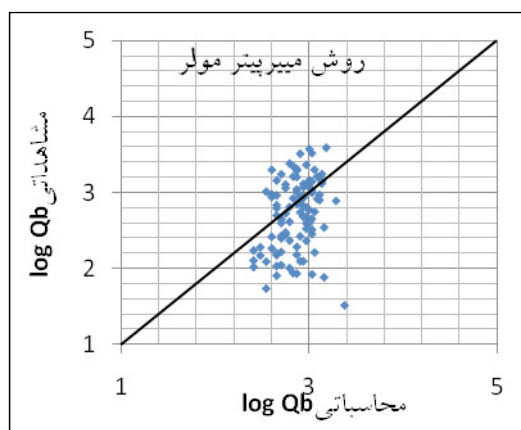
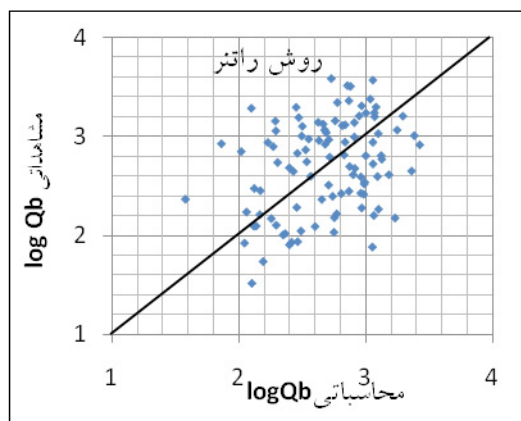
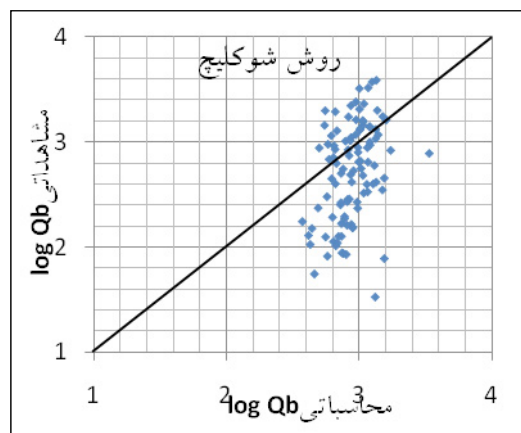
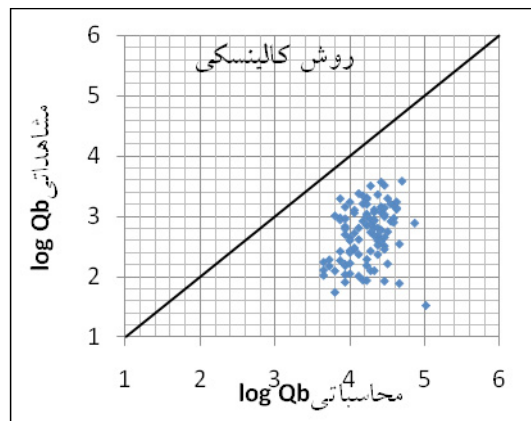
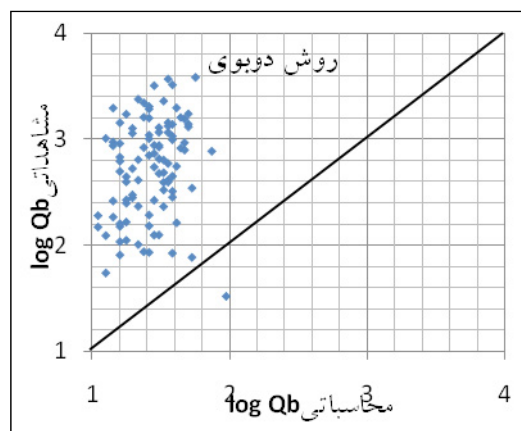
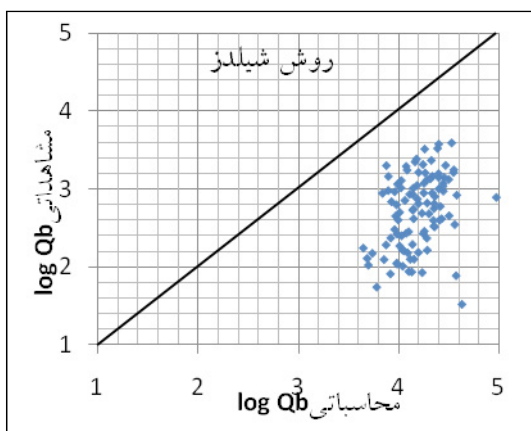
(ب) نسبت بار کف به بار کل

نسبت بار کف به بار کل به‌طور متوسط ۰/۴۵ و نسبت بار معلق به بار کل به‌طور متوسط ۰/۵۵ می‌باشد و نسبت بار کف به معلق به‌طور متوسط ۰/۸۲ است و این در حالی است که در رودخانه‌های با بستر ریزدانه بار کف به معلق بین ۰/۰۵ تا ۰/۲۵ تخمین زده می‌شود. در ۴۵ درصد از داده‌ها بار کف از بار معلق بیش‌تر بوده است که در اکثر این موارد دبی جریان از یک مقدار خاص کم‌تر بوده‌است. بنابراین همان‌طور که در شکل ۵ دیده می‌شود، می‌توان این‌طور استنباط نمود که با افزایش دبی و تجاوز از یک مقدار مشخص (یعنی محل تلاقی دو خط برازش شده‌ی بار معلق و کف) بار معلق از بار کف بیش‌تر می‌شود.

در اینجا از هفت روش رایج تجربی برآورد بار رسوب بستر استفاده شده است؛ اینشتین-براون<sup>۱</sup>، مییر پیتر مولر<sup>۲</sup>، راتنر<sup>۳</sup>،

4. Kalinske  
5. Shoklitch  
6. Shields  
7. Duboy  
8. Line of perfect agreement

1. Einstein-Brown, 1952  
2. Meyer Peter and Muller (MPM)  
3. Rottner



شکل ۷: مقایسه نتایج برآورد هفت روش تجربی

جدول ۲: نتایج روش‌های تجربی با داده‌های آزمون درخت تصمیم

| شماره ردیف | روش         | R% | RMSE   | $\overline{Dr}$ | MAE    |
|------------|-------------|----|--------|-----------------|--------|
| ۱          | دوبوی       | ۵۲ | ۸۸۴۸   | ۰/۰۴            | ۷۵۴۱   |
| ۲          | شیلدز       | ۵۹ | ۲۱۹۳۳۱ | ۲۵              | ۱۶۱۰۷۷ |
| ۳          | شوکلچ       | ۶۰ | ۶۲۲۳   | ۱/۳۳            | ۳۸۵۴   |
| ۴          | کالینسکی    | ۵۷ | ۲۶۵۹۶۹ | ۲۹              | ۱۸۳۳۶۵ |
| ۵          | راتر        | ۷۲ | ۵۶۰۲   | ۰/۹۹            | ۴۶۴۶   |
| ۶          | مبیرپترمولر | ۵۳ | ۶۲۰۸   | ۱/۱۳            | ۳۹۱۰   |
| ۷          | اینشتین     | ۵۲ | ۵۷۰۵   | ۱۱/۹۵           | ۷۲۰۳۲  |

ازای هر چهار داده برای آموزش یک داده بعد از آن برای آزمون کنار گذاشته شد، این امر باعث می‌شود که نحوه توزیع داده‌های آموزش و آزمون به هم نزدیک باشد.

پس از انتخاب ورودی‌ها، آن‌ها به صورت ستون‌های یک ماتریس برای ورود به درخت تصمیم معرفی گردیدند. خروجی‌های درخت تصمیم که همان بار رسوب بستر هستند در مقایسه با مقادیر مشاهداتی در شکل (۸) آورده شده‌اند (به ترتیب از ۱ تا ۹ شکل‌ها مربوط به ورودی‌های معرفی شده در جدول ۴ می‌باشند).

جدول ۳: ورودی‌ها و خروجی‌های در نظر گرفته شده برای الگوریتم درخت تصمیم

| شماره مدل | ورودی                               | خروجی         |
|-----------|-------------------------------------|---------------|
| ۱         | دبی جریان، عمق                      | دبی رسوب بستر |
| ۲         | دبی جریان، عمق، دبی رسوب معلق       | دبی رسوب بستر |
| ۳         | سرعت، عمق                           | دبی رسوب بستر |
| ۴         | عمق، سرعت، دبی رسوب معلق            | دبی رسوب بستر |
| ۵         | عمق، سرعت، دبی رسوب معلق، دبی جریان | دبی رسوب بستر |
| ۶         | سرعت، دبی رسوب معلق، دبی جریان      | دبی رسوب بستر |
| ۷         | سرعت، دبی جریان                     | دبی رسوب بستر |
| ۸         | دبی رسوب معلق، دبی جریان            | دبی رسوب بستر |
| ۹         | سرعت، دبی رسوب معلق                 | دبی رسوب بستر |

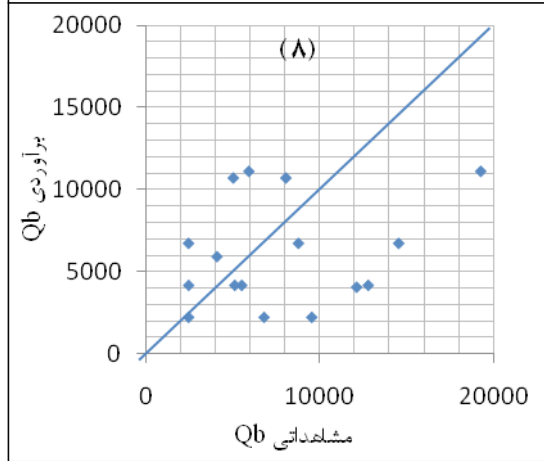
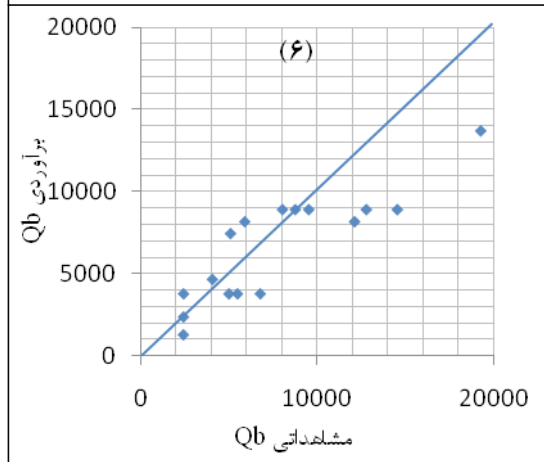
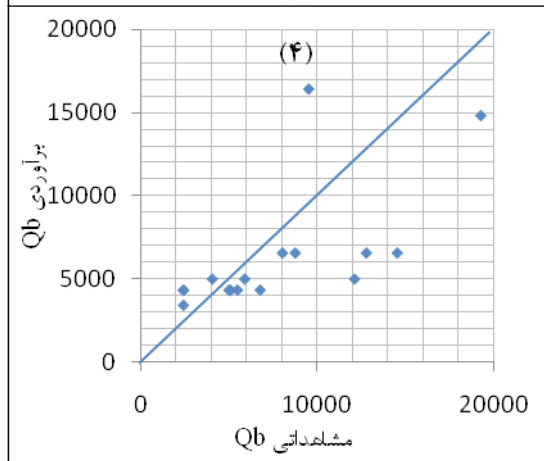
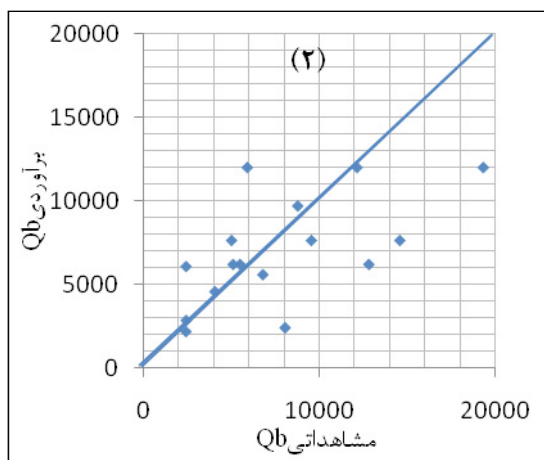
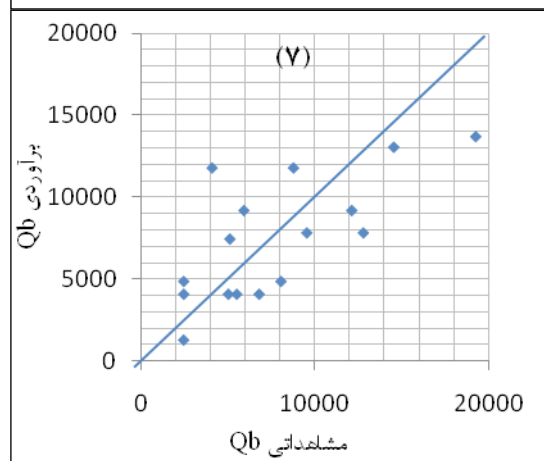
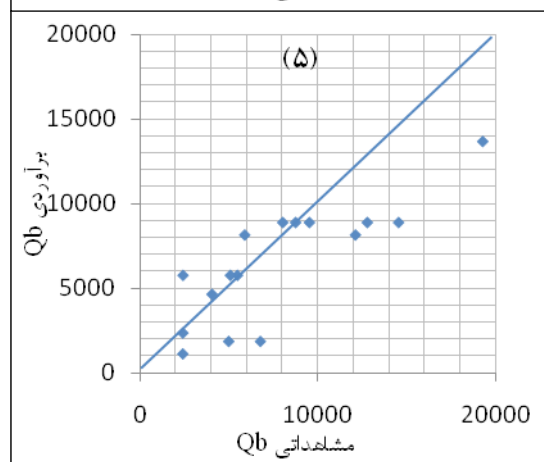
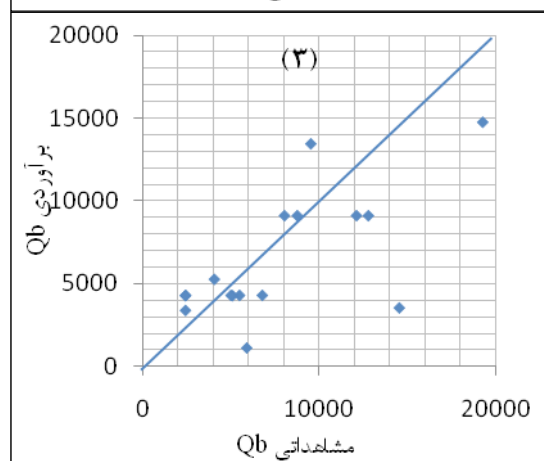
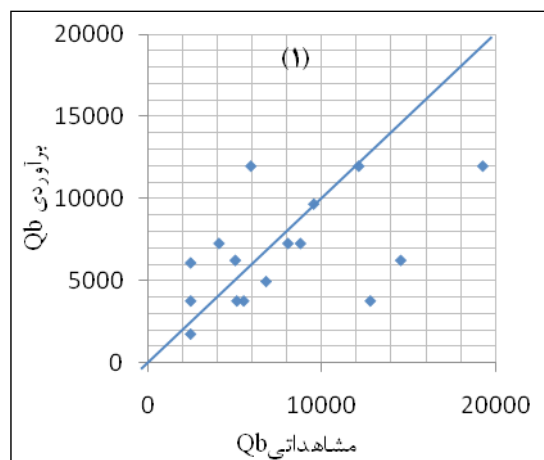
همان‌گونه که دیده می‌شود روش شیلدز، اینشتین-براون و کالینسکی مقادیر رسوب را بسیار بیش‌تر از مقادیر مشاهداتی پیش‌بینی کرده‌اند (بیش‌برازش) و روش دوبویز نیز کم‌تر از مقدار واقعی برآورد نموده‌است. جهت مقایسه بهتر نتایج روش‌های تجربی و مدل درخت تصمیم رگرسیون، پارامترهای خطای روش‌های تجربی را برای همان داده‌هایی که در آزمون درخت تصمیم استفاده شده محاسبه گردیده است. همان‌گونه که پارامترهای جدول (۲) نشان می‌دهد بالاترین ضریب همبستگی (R)، کم‌ترین مجذور میانگین مربعات خطا (RMSE)، کم‌ترین میانگین خطای مطلق (MAE) و نزدیک‌ترین ضریب توزیع خطا نتایج بهتری داشته‌است. به نظر می‌رسد سایر روش‌ها قابلیت کاربرد ندارند چرا که میزان خطای برآورد آن‌ها بسیار بالاست.

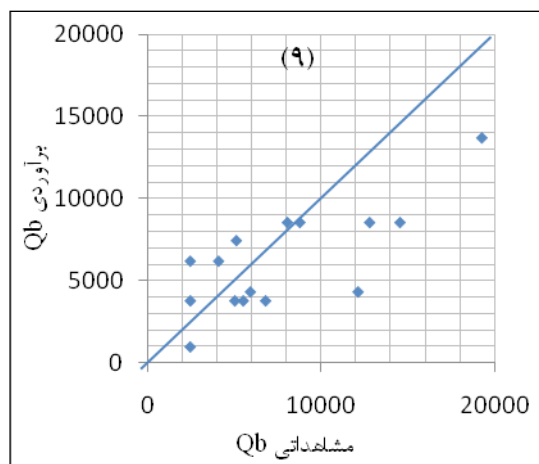
(Dr) به یک، مربوط به دو روش راتر، راتر، شوکلچ و راتر می‌باشد و به جز پارامتر میانگین خطای مطلق روش راتر در سایر پارامترهای

### نتایج مدل درخت تصمیم

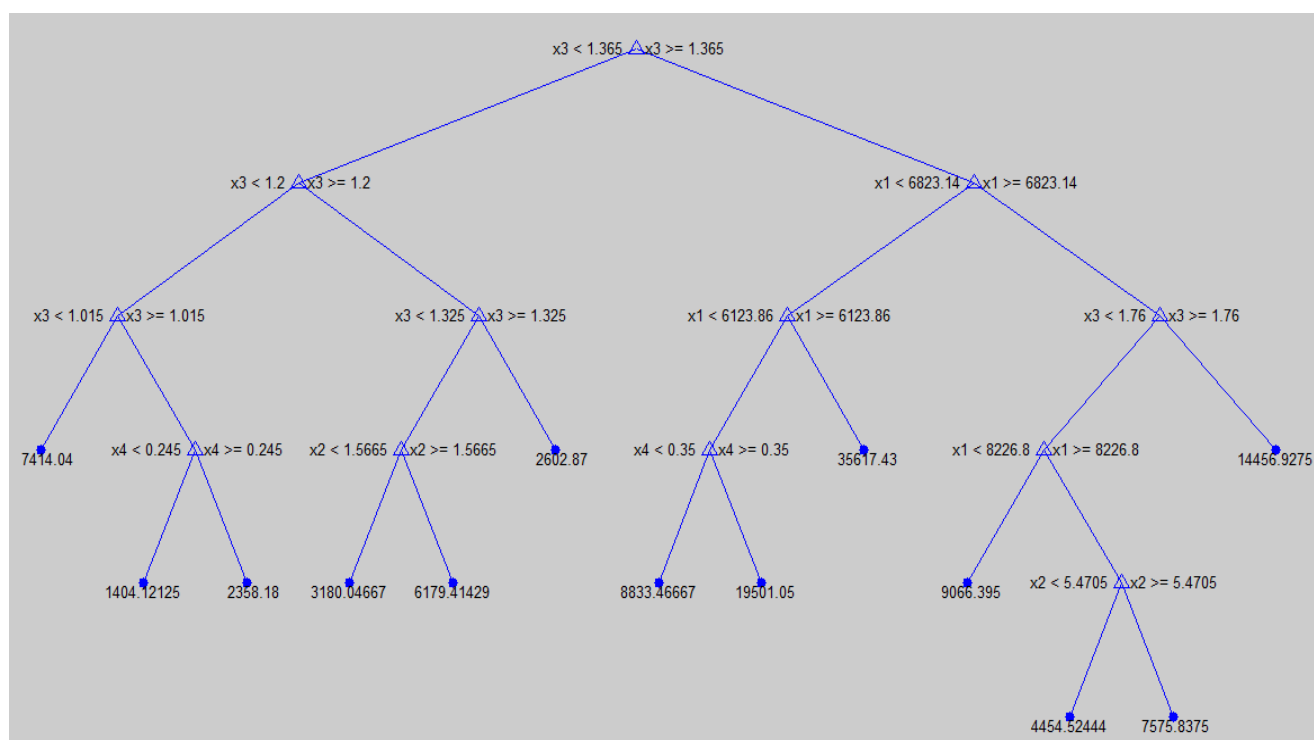
#### تعیین بهترین ترکیب ورودی

از بین پارامترهای دبی آب، دبی رسوبات معلق، سرعت و عمق آب، ۹ نوع ترکیب مختلف برای ورودی در نظر گرفته شد. ۸۰ درصد داده‌ها (۶۰ داده) برای آموزش و ۲۰ درصد بقیه برای آزمون به کار برده شده‌اند که در جدول (۳) ذکر شده است. نحوه گزینش داده‌ها به دو شکل انجام گرفت، ابتدا به صورت ساده، یعنی ۸۰ درصد ابتدایی داده‌ها برای آموزش و ۲۰ درصد انتهایی جهت آزمون به کار برده شدند و پس از آن چون داده‌ها به ترتیب زمان برداشت پشت‌سر هم قرار داشتند و جهت حذف اثر زمان برداشت در سال‌های مختلف، داده‌ها به صورت چهار تا یکی تفکیک گردیدند، یعنی به





شکل ۸: همبستگی مقادیر رسوب بستر برآورد درخت تصمیم به ازای ۹ مدل ورودی و مقادیر مشاهداتی



شکل ۹: درخت تصمیم مدل ۵،  $x_1$  دبی رسوب معلق،  $x_2$  دبی جریان،  $x_3$  سرعت جریان و  $x_4$  عمق جریان آب می‌باشد.

چرا که عمق آب در تقسیم شاخه‌های درخت تصمیم مدل ۵ که در شکل (۹) نشان داده شده است، نقش چشمگیر و مهمی نداشته است و فقط در یک تقسیم بندی حضور داشته است. از این رو از بین ترکیب ورودی ۵ که در حالت انتخاب داده‌ی ترتیبی و ترکیب ورودی ۶ در انتخاب داده ساده، بهترین نتیجه را داشته‌اند، مدل ۶ ارجح می‌باشد. بنابراین بهترین نتایج مدل مربوط به مدلی است که بیشترین تعداد ورودی‌ها را داشته‌است، اگرچه که برخی ورودی‌ها مثل عمق نقش مؤثری نداشته‌اند. در واقع می‌توان این‌گونه توضیح داد که درخت تصمیم رگرسیونی از بین کلیه پارامترهای ورودی خود به آنالیز حساسیت پرداخته و پارامترهای مؤثرتر را در اولین شکاف‌های طبقه‌بندی مورد استفاده قرار می‌دهد. به‌عنوان نمونه در

با بررسی پارامترهای ضریب همبستگی  $R$  و  $RMSE$  درمی‌یابیم که منطبق‌ترین خروجی با واقعیت مربوط به دو حالت ۵ و ۶ از جدول (۳) و (۴) می‌باشند که دارای بالاترین ضریب همبستگی و کم‌ترین جذر میانگین مربعات خطا می‌باشند. این نتیجه برای حالتی که انتخاب داده‌های آموزش و آزمون نیز به صورت ترتیبی انتخاب گردید، صادق می‌باشد (جدول ۵). این ورودی‌ها در حالت ۶ شامل دبی جریان، دبی رسوبات معلق و سرعت جریان آب و در حالت ۵ شامل همه موارد بالا به علاوه عمق آب می‌باشد. اگر از بین این دو حالت ورودی بخواهیم یکی انتخاب گردد، مدل ۶ که شامل سه ورودی ذکر شده است، انتخاب می‌گردد،

شکل (۹) که درخت مربوط به هر چهار ورودی نمایش داده شده است می‌بینیم که اولین متغیری که جهت شکاف استفاده شده است،  $X_3$  یا همان سرعت است و متغیر  $X_4$  که همان عمق است، تنها یک‌بار در شکاف داده‌ها مورد استفاده قرار گرفته است. این مطلب نشان دهنده این است که سرعت اهمیت بیش‌تری در برآورد بار رسوب بستر دارد.

#### درخت تصمیم بهینه

درخت تصمیم حداکثر می‌بایست با هرس کردن تبدیل به درخت بهینه (یعنی با حداقل شاخه و برگ ولی پیش‌بینی‌های مناسب) نمود. مدل ورودی ۶ که بهترین خروجی را بدست می‌دهد، تا اینجا یک درخت حداکثر البته با اعمال شرط حداقل تعداد داده در برگ‌هاست، این شرط را نرم افزار متلب به صورت پیش‌فرض (برابر با ۱۰)، اعمال می‌کند. این درخت که شکل (۱۲) آورده شده است، تا ۱۰ سطح می‌تواند هرس شود، یعنی تا جایی که کلیه برگ‌ها حذف شوند. در اینجا جزئیات ۵ مرحله از آن و نتایج حاصل را در جدول

جدول ۴: تعیین بهترین ترکیب ورودی درخت تصمیم با نحوه انتخاب داده ساده

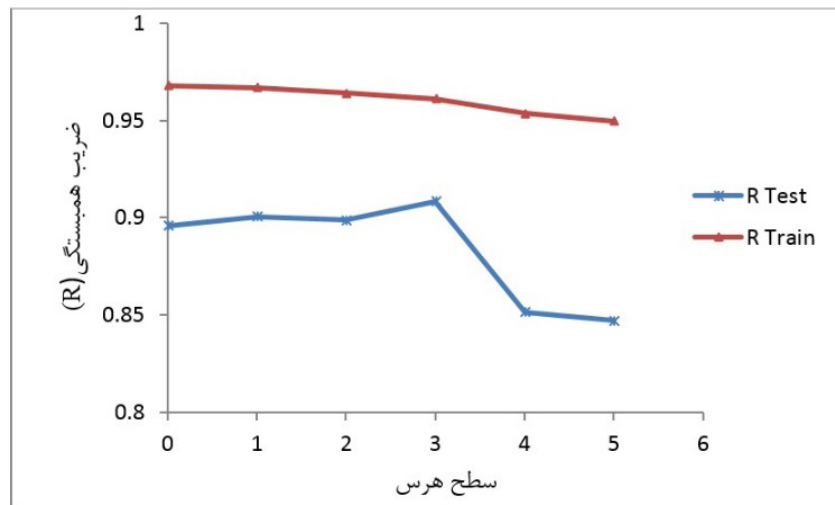
| ردیف | خروجی      | ورودی                                       | روش انتخاب | R%    |       | RMSE  |       | $\overline{Dr}$ |       | MAE   |       |
|------|------------|---------------------------------------------|------------|-------|-------|-------|-------|-----------------|-------|-------|-------|
|      |            |                                             |            | آزمون | آموزش | آزمون | آموزش | آزمون           | آموزش | آزمون | آموزش |
| ۱    | Qbed(gr/s) | Qflow(m3/s)<br>Y(m)                         | ساده       | ۷۲    | ۹۵    | ۴۱۷۷  | ۱۶۶۰  | ۱/۰۶            | ۱/۴۲  | ۳۰۱۷  | ۱۸۱۲  |
| ۲    | Qbed(gr/s) | Qflow(m3/s)<br>Y(m)<br>Qsus(gr/s)           | ساده       | ۷۷    | ۹۵    | ۳۵۶۴  | ۲۳۶۲  | ۱/۰۸            | ۱/۳۷  | ۲۸۶۹  | ۱۶۱۱  |
| ۳    | Qbed(gr/s) | Y(m)<br>V(m/s)                              | ساده       | ۸۲    | ۸۸    | ۳۷۱۱  | ۳۵۸۲  | ۰/۹۸            | ۱/۱۸  | ۲۷۰۹  | ۱۹۱۵  |
| ۴    | Qbed(gr/s) | Y(m)<br>V(m/s)<br>Qsus(gr/s)                | ساده       | ۸۰    | ۹۰    | ۳۹۲۱  | ۳۳۰۵  | ۰/۹۷            | ۱/۲۱  | ۳۰۰۵  | ۱۸۱۲  |
| ۵    | Qbed(gr/s) | Y(m)<br>V(m/s)<br>Qsus(gr/s)<br>Qflow(m3/s) | ساده       | ۹۰    | ۹۵    | ۳۰۵۳  | ۲۲۹۴  | ۰/۹۳            | ۱/۳۵  | ۲۳۳۴  | ۱۴۹۰  |
| ۶    | Qbed(gr/s) | V(m/s)<br>Qsus(gr/s)<br>Qflow(m3/s)         | ساده       | ۹۳    | ۹۵    | ۲۷۸۲  | ۲۲۰۶  | ۰/۹۲            | ۱/۱۸  | ۲۱۵۸  | ۱۴۱۱  |
| ۷    | Qbed(gr/s) | V(m/s)<br>Qflow(m3/s)                       | ساده       | ۸۳    | ۸۱    | ۳۴۰۳  | ۴۲۰۷  | ۱/۱۲            | ۱/۳   | ۲۹۱۵  | ۲۱۶۴  |
| ۸    | Qbed(gr/s) | Qflow(m3/s)<br>Qsus(gr/s)                   | ساده       | ۵۰    | ۹۴    | ۵۲۳۵  | ۲۶۹۵  | ۱/۰۵            | ۱/۴۹  | ۴۳۹۵  | ۱۹۴۶  |
| ۹    | Qbed(gr/s) | V(m/s)<br>Qsus(gr/s)                        | ساده       | ۷۱    | ۹۱    | ۴۹۵۱  | ۳۲۱۶  | ۱/۰۶            | ۱/۳۷  | ۳۵۷۷  | ۲۰۱۰  |

جدول ۵. تعیین بهترین ترکیب ورودی درخت تصمیم با نحوه انتخاب داده تریبی

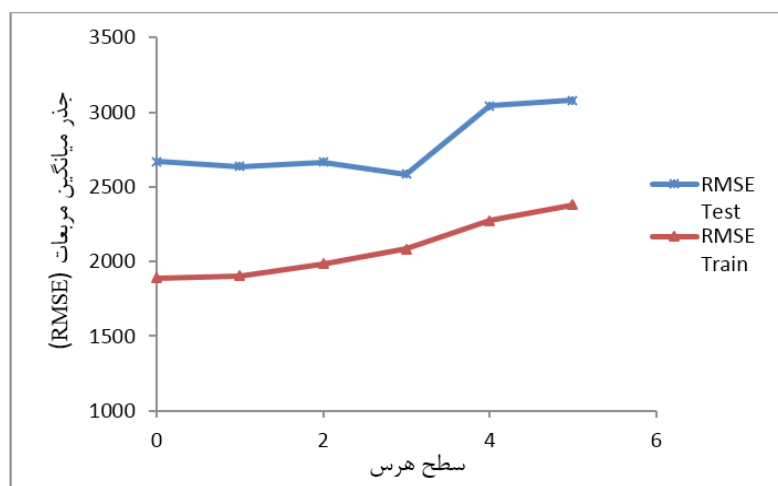
| ردیف | خروجی      | ورودی                                       | روش انتخاب | R%    |       | RMSE  |       | $\bar{Dr}$ |       | MAE   |       |
|------|------------|---------------------------------------------|------------|-------|-------|-------|-------|------------|-------|-------|-------|
|      |            |                                             |            | آزمون | آموزش | آزمون | آموزش | آزمون      | آموزش | آزمون | آموزش |
| ۱    | Qbed(gr/s) | Qflow(m3/s)<br>Y(m)                         | ۱ به ۴     | ۵۷    | ۹۵    | ۴۲۸۱  | ۳۹۸۰  | ۱/۶۹       | ۱/۵   | ۳۱۳۶  | ۱۵۹۳  |
| ۲    | Qbed(gr/s) | Qflow(m3/s)<br>Y(m)<br>Qsus(gr/s)           | ۱ به ۴     | ۷۴    | ۹۶    | ۳۴۴۵  | ۲۲۱۵  | ۱/۳۵       | ۱/۳   | ۲۷۴۶  | ۱۴۹۹  |
| ۳    | Qbed(gr/s) | Y(m)<br>V(m/s)                              | ۱ به ۴     | ۸۵    | ۸۸    | ۳۰۲۰  | ۳۶۰۲  | ۱/۵        | ۱/۳۳  | ۲۷۰۱  | ۱۹۴۸  |
| ۴    | Qbed(gr/s) | Y(m)<br>V(m/s)<br>Qsus(gr/s)                | ۱ به ۴     | ۷۷    | ۹۴    | ۶۷۳۵  | ۱۸۴۵  | ۱/۷        | ۱/۲۹  | ۴۰۰۶  | ۱۳۰۵  |
| ۵    | Qbed(gr/s) | Y(m)<br>V(m/s)<br>Qsus(gr/s)<br>Qflow(m3/s) | ۱ به ۴     | ۹۱    | ۹۷    | ۲۴۶۹  | ۱۸۹۰  | ۱/۵        | ۱/۲۹  | ۱۹۶۰  | ۱۳۶۵  |
| ۶    | Qbed(gr/s) | V(m/s)<br>Qsus(gr/s)<br>Qflow(m3/s)         | ۱ به ۴     | ۸۹    | ۹۶    | ۲۶۶۸  | ۲۰۲۸  | ۱/۴۱       | ۱/۲   | ۲۱۳۴  | ۱۴۳۲  |
| ۷    | Qbed(gr/s) | V(m/s)<br>Qflow(m3/s)                       | ۱ به ۴     | ۶۷    | ۹۰    | ۴۱۴۵  | ۳۲۵۰  | ۱/۷۱       | ۱/۳۲  | ۳۲۷۸  | ۱۸۵۹  |
| ۸    | Qbed(gr/s) | Qsus(gr/s)<br>Qflow(m3/s)                   | ۱ به ۴     | ۵۵    | ۹۱    | ۶۰۰۰  | ۳۱۷۸  | ۲/۵        | ۱/۵۶  | ۵۰۵۳  | ۲۳۶۳  |
| ۹    | Qbed(gr/s) | Qsus(gr/s)<br>V(m/s)                        | ۱ به ۴     | ۷۶    | ۹۶    | ۶۶۹۷  | ۱۹۸۷  | ۱/۶۳       | ۱/۲   | ۴۰۰۱  | ۱۳۷۳  |

جدول ۶: هرس درخت تصمیم نهایی

| سطح هرس | تعداد قوانین | R     |       | RMSE  |       |
|---------|--------------|-------|-------|-------|-------|
|         |              | آموزش | آزمون | آموزش | آزمون |
| ۰       | ۲۷           | ۹۶/۷  | ۸۹/۵  | ۲۰۲۸  | ۲۶۶۸  |
| ۱       | ۲۳           | ۹۶/۶  | ۹۰    | ۲۰۵۸  | ۲۶۳۶  |
| ۲       | ۲۱           | ۹۶/۴  | ۸۹/۸  | ۲۱۳۴  | ۲۶۶۶  |
| ۳       | ۱۹           | ۹۶    | ۹۰/۸  | ۲۲۲۵  | ۲۵۸۴  |
| ۴       | ۱۵           | ۹۵/۳  | ۸۵/۱  | ۲۴۰۵  | ۳۰۴۱  |
| ۵       | ۱۳           | ۹۴/۹  | ۸۴/۷  | ۲۵۰۵  | ۳۰۷۸  |



شکل ۱۰: تغییر ضریب همبستگی به ازای هر سطح هرس درخت تصمیم



شکل ۱۱: تغییر جذر میانگین مربعات خطا به ازای هر سطح هرس درخت تصمیم

جدول ۷: مقایسه نتایج بهترین مدل درخت تصمیم و بهترین روش تجربی

| پارامترهای خطا/ روش           | R% | RMSE | $\overline{Df}$ | MAE  |
|-------------------------------|----|------|-----------------|------|
| بهترین مدل درخت تصمیم (مدل ۶) | ۸۹ | ۲۶۶۸ | ۱/۴۱            | ۲۱۳۴ |
| بهترین روش تجربی (راتنر)      | ۷۲ | ۵۶۰۲ | ۰/۹۹            | ۴۶۴۶ |

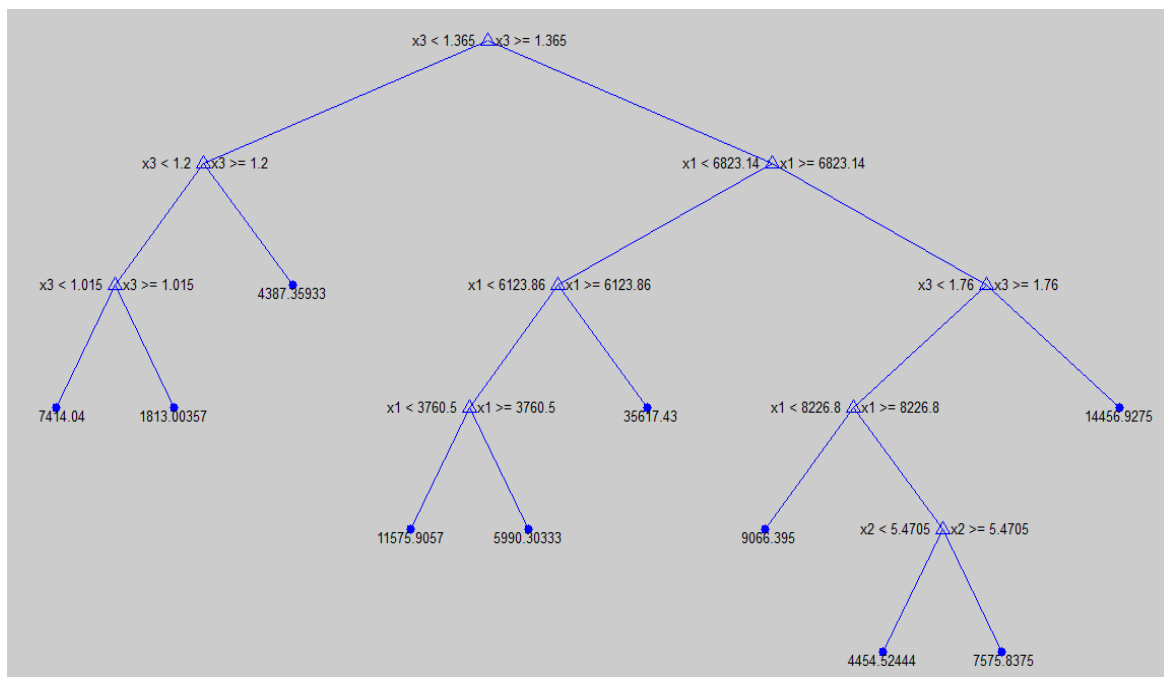
۵ آورده شده است.

### مقایسه نتایج مدل درخت تصمیم رگرسیونی و روش‌های

#### تجربی

جهت مقایسه مدل درخت تصمیم‌گیری رگرسیونی با روش‌های تجربی، بهترین روش تجربی را با بهترین ترکیب ورودی مدل درخت مقایسه کرده‌ایم. همان‌گونه که در جدول (۷) که مقادیر آن از جداول (۲) و (۵) استخراج شده است، می‌بینیم، درخت تصمیم به وضوح بهتر از روش‌های تجربی نتیجه داده است. در واقع به جز

در دو شکل (۱۰) و (۱۱) دیده می‌شود که مقدار حداکثر  $R$  و حداقل  $RMSE$  به ازای سه سطح هرس کردن درخت در قسمت آزمون به دست می‌آید، بدین معنی که برگ‌های اضافه و غیر مؤثر در افزایش راندمان درخت را حذف کرده و بدین ترتیب درخت کوچک‌تر (با قوانین کم‌تر) با کارایی بیشتر به جای درخت حداکثر استفاده می‌گردد. درخت در شکل (۱۲) آورده شده است.



شکل ۱۲: درخت تصمیم بهینه که سه سطح آن هرس شده است،  $x_1$  دبی رسوب معلق،  $x_2$  دبی جریان،  $x_3$  سرعت جریان

داده‌های آزمایشگاهی یا صحرایی محدود مربوط به یک یا چند منطقه ارائه می‌گردند که البته قریب به اتفاق همه‌ی آن‌ها مربوط به حوزه‌های خارج از کشور هستند، بنابراین انتظار نتایج عالی در مناطق دیگر با شرایط متفاوت نمی‌توان داشت.

#### پیشنهادهای

۱- دیگر روش‌های هوش مصنوعی مثل شبکه‌های عصبی نیز در همین منطقه استفاده شود و بررسی گردد کدام روش بهینه‌ترین برآورد را انجام می‌دهد. این بررسی می‌تواند محدود به بار رسوب کف، معلق و یا کل شود.

۲- سایر الگوریتم‌های درخت تصمیم مثل M5 نیز به بوته آزمایش گذاشته شود تا الگوریتم بهینه و نزدیک‌تر به نتایج برداشت‌های صحرایی بدست آید.

۳- این مطالعه اگر به‌طور جامع در مناطق مختلف کشور انجام گیرد، نتایج ارزشمندی در پی خواهد داشت و از این پس می‌توان برای برآورد رسوبات مناطق مختلف با اطمینان بیش‌تری روش‌های هوش مصنوعی را جایگزین فرمول‌های تجربی نمود.

#### منابع

1. Barzegari, F., Yosefi, M. and Talebi, A. 2015. Estimating suspended sediment by Artificial Neural Network (ANN), Decision Trees (DT) and Sediment Rating Curve (SRC) models (Case study: Lorestan Province, Iran). Civil Engineering Infrastructures Journal. 48(2): 391-398.
2. Bussi, G., Dadson, S. J. C. and Prudhomme Whitehead, P. G. 2016. "Modelling the Future Impacts of Climate and

پارامتر ضریب توزیع، بقیه پارامترهای خطای مدل درخت که شامل ضریب همبستگی، جذر میانگین مربعات خطا و میانگین قدر مطلق خطاست، بهتر بوده‌اند.

#### نتیجه‌گیری

در این تحقیق در قسمت برآورد با فرمول‌های تجربی از هر سه گروه فرمول‌های مذکور استفاده گردیده است و نتیجه نشان داد که در این میان روش شوکلچ که بر مبنای دبی است و روش راتنر که از روش‌های رگرسیونی چند متغیره است و روش میسر پیتز که براساس همبستگی با عمق جریان است بهتر از سایر روش‌ها در این منطقه نتیجه می‌دهند. البته دلیل دیگر اینکه سایر روش‌ها نتایج چندان مناسبی ارائه نکرده‌اند می‌تواند این باشد که در هر بار برداشت بار رسوب و دبی جریان، دانه‌بندی اندازه‌گیری نشده است، و تنها به تعداد محدودی دانه‌بندی در این منطقه برداشت شده است. از طرف دیگر در برآورد با درخت تصمیم نیز بهترین نتیجه مربوط به ترکیب پارامترهای دبی جریان، سرعت جریان، دبی رسوب معلق می‌باشد، که البته تاثیر سرعت بیش از سایر پارامترها بوده است. درخت تصمیم در این تحقیق در حالت بهینه شامل ۱۹ قانون می‌باشد. اگرچه که این امر نیازمند مطالعات منطقه‌ای است و ممکن است قوانین ارائه شده در یک منطقه برای منطقه دیگر قابل استفاده نباشد، اما به‌هرحال برآوردهای نزدیک‌تری به واقعیت خواهیم گرفت. در مقایسه فرمول‌های تجربی با نتایج درخت تصمیم در این پژوهش، به وضوح می‌توان دید که درخت تصمیم عملکرد بهتری داشته است که با نتایج طالبی و اکبری [۱۱] و حلبیان و همکاران [۶] هم‌خوانی دارد و این نتیجه منطقی است چرا که روابط تجربی عموماً بر مبنای

8. Kazemi, Y. 2009. Investigation of barbestor to suspension ratio in Central Alborz rivers (Jajrood and Taleghan rivers). M.Sc., University of Tehran. (In Persian)
9. Nazamara, H. 2005. Estimation of Suspended Load by Artificial Neural Network (Case Study of Ajji Chay River). Senior Thesis, Tabriz University, Faculty of Civil Engineering. (In Persian)
10. Saghebani, S. M., Sattari, M., Mirabbasi, R. and Mahesh Pal. 2014. "Ground water quality classification by decision tree method in Ardebil region". Iran, Arabian Journal of Geosciences. 7, 4767-4777.
11. Talebi, A., and Akbari, Z. 2012. Evaluation of Decision Tree Modeling in Estimation of River Suspended Sediments (Case Study of Ilam Dam Basin), Journal of Agriculture and Natural Resources Science and Technology, Soil Science, 17th Year, 63. (In Persian)
12. Tahmasebinasb, M., and Dehghani, A. 2013. Accuracy estimation of unconventional methods for estimating bed load transfer velocity using field data, Iranian Journal of Drainage Research, No. 1, 8(12): 126-116. (In Persian)
13. Yousefi, M., and Barzegari, F. 2013. Comparative Study of Suspended Sediment Estimation Using Decision Tree Model and Sedimentation Index Curves (Case Study of Lorestan Province), Watershed Research, 104. (In Persian)
- Land-Use Change on Suspended Sediment Transport in the River Thames (UK)", Hydrology. 542: 357-372.
3. Dasturani, M. T., Habibipour, A., Ekhtesasi, M., et al. R., Talebi, A. and Mahjubi, J. 2012. Investigation of the Performance of Decision Tree Model in Rainfall Prediction (Case Study of Yazd Synoptic Station), Iranian Water Resources Research, Eighth Year, 3. (In Persian)
4. Dastorani, M.T., Mahjoobi, J., Talebi, A. and Fakhar, F. 2018. "Application of Machine Learning Approaches in Rainfall-Runoff Modeling (Case Study: Zayandeh\_Rood Basin in Iran)", Civil Engineering Infrastructures Journal. 51(2): 293 – 310.
5. Dehghani, N., and Vafaakhah, M. 2013. Evaluation of Sediment Load Hydraulic Equations in Guilan Noorud River, Journal of Agricultural Science and Technology, Soil and Water Sciences, Eighteen Years, 69. (In Persian)
6. Halabyan, A. H., Jawari, M., Akbari, Z. and Akbari, G. 2017 Evaluation of Fashion Efficiency of Decision Trees in Estimation of River Suspended Sediments (Case Study of Meymeh Basin), Geography & Development, 49: 81-96. (In Persian)
7. Hamzeh Pouri, R. 2005. Evaluation of sediment load in rivers with coarse-grained materials. M.Sc. Urmia University. (In Persian)



## Abstract

**Estimation of Bedload Sediment Using Decision Trees Method and Comparison with Empirical Methods**N.L. Amanian<sup>\*1</sup>, A. Talebi<sup>2</sup>, P. Hajah Bakhsh<sup>3</sup> and S. parvizi<sup>4</sup>

Received: 2020/05/22 Accepted: 2020/11/01

Estimation of the amount of sediment material (carried by a specific stream) is one of the main topics of sediment research. Sediment research is important in many engineering projects such as planning and design of water storage resources, morphology and changes in river bed, annual sediment estimation for reservoirs, sustainable irrigation, coastal protection, channel dredging, and etc. The evaluation of the sediment occurrences, and estimation of the sediment carried by rivers has a special importance and it is necessary to develop new methods with easy application for estimating the sediment transport. This study seeks to develop a method to be able to predict the amount of bedload with classification and regression trees and compare with some well-known empirical formulas such as Meyer Pette Muller, Shields, Einstein Brown, and Shoklitch. Measured data from Yazdakan station, located in Qotour River in Azarbayjan Gharbi, northwest of Iran, has been used in this research. It is a mountainous river with gravel bed that carries a high volume of sedimentation. A set of 76 field measurements over 8 years was used in modeling after verification and refinement. The model was trained on 80 percent of data and tested on the remained 20 percent. The best input contained whole four variables; flow discharge, suspended sediment discharge, flow depth and velocity, and the output was bedload discharge. The quantitative results of the best Decision Tree include MAE = 2134, RMSE = 2668 and % R = 89. The results of Regression Trees demonstrate more accuracy in the prediction of bedload compared with empirical formulas.

**Keywords: Bedload, Empirical methods, Estimation of sediment, Regression trees**

1. Assistant Professor, Faculty of Civil Engineering, Yazd University, Yazd, Iran. Corresponding author Email University Email: Email: Namanian@yazd.ac.ir

2. Professor in Department of Watershed Management Engineering, Yazd University, Yazd, Iran.

3. M.Sc. graduate Civil water Engineering, Yazd University, Yazd, Iran.

4. Ph.D Student Watershed Management Engineering, Yazd University, Yazd, Iran.