پژوهش‌های آبخیزداری شماره ۱۰۴، پاییز ۱۳۹۳
(پژوهش و سازندگی)

بررسی مقایسه‌ای برآورد رسوبات معلق با استفاده از مدل درخت تصمیم و منحنی‌های سنجه رسوب (مطالعه موردی: استان لرستان)

• **محسن یوسفی**
دانشجوی کارشناسی ارشد آبخیزداری، دانشگاه یزد (نویسنده مسئول)
• **فاطمه برزگری**
مربی دانشگاه پیام نور دانشکده کشاورزی یزد
تاریخ دریافت: تیر ماه ۱۳۹۲ تاریخ پذیرش: آذر ماه ۱۳۹۲
تلفن تماس نویسنده مسئول: ۰۹۱۲۷۰۷۷۰۴۹
Email:Mohsenyosefi67@gmail.com

چکیده
آگاهی از میزان رسوب معلق، از عوامل مهم در برآورد فرسایش حوضه و تصمیم‌گیری در مورد احداث سازه‌های آبی بر روی رودخانه‌هاست، همچنین برآورد بار رسوبی در رودخانه‌ها با توجه به خسارات ناشی از آن، یکی از مهم‌ترین و مشکل‌ترین مطالعات انتقال رسوب و مهندسی رودخانه است. بنابراین دست یافتن به شیوه‌های نوین که بتواند در این راستا مؤثر واقع گردد دارای اهمیت زیادی است. هدف از این مطالعه برآورد رسوبات معلق با استفاده از مدل درخت تصمیم با الگوریتم کارت و مقایسه نتایج با انواع منحنی سنجه رسوب می باشد. در این پژوهش از آمار دبی و رسوب متناظر ده ایستگاه آبسنجی رودخانه‌های استان لرستان استفاده گردید. نتایج نشان داد در بین انواع روش‌های منحنی سنجه و ضرایب اصلاحی به کار گرفته شده که جمعا شامل ۲۰ روش بود، ترکیب منحنی سنجه ماهانه و ضریب اصلاحی MUVE دارای میانگین مربعات خطای کمتر و دقت بیشتری میباشد. بعد از اعمال مناسب‌ترین ضریب اصلاحی به انواع منحنی‌های سنجه رسوب، نتایج حاصل از برآورد رسوب با منحنی سنجه، با نتایج حاصل از مدل درخت تصمیم با استفاده از شاخص ضریب همبستگی، میانگین قدر مطلق خطا، ضریب دقت و میانگین مربعات خطا، مورد مقایسه قرار گرفت. نتایج، نشان دهنده مناسب بودن هر دو مدل در تخمین رسوبات معلق است که در بعضی ایستگاه‌ها منحنی سنجه و در بعضی دیگر درخت تصمیم مناسبتر به نظر میرسد. اما با توجه به مناسب‌تر بودن منحنی سنجه با اعمال ضرایب اصلاحی در بیشتر ایستگاه‌ها، می‌توان منحنی سنجه را مناسب‌تر از مدل درخت تصمیم دانست.

کلمات کلیدی: الگوریتم کارت، بار معلق، درخت تصمیم، ضرایب اصلاحی، منحنی سنجه ماهانه.

Watershed Management Research (Pajouhesh & Sazandegi) No 104 pp: 187-195

Estimation of suspeded sediment using decision tree method and copmarision the results with rating curve method (Case study: Rivers in Lorestan Province)

By: M. Yousefi: MSc Student, Yazd University (Corresponding Author; Tel:+989127077049). F. Barzegari: Scientific Board, Payame Noor University, Yazd, Iran.

Awareness of the amount of suspended sediment yield is an important factor to estimating erosion and also making decisions about the construction of hydraulic structures on rivers. Estimating suspended sediment load of the river with regard to the damage caused by it, is one of the most important and most difficult studies of sediment transport and river engineering. So finding new models that could be useful in this regard has extraordinary significance. The purpose of this study was to estimating suspended sediment using Decision Tree with CART algorithm, and compare the results with different sediment rating curves including: seasonal, monthly, Mean load within discharge classes and linear. To attain this, the sediment discharge and the corresponding water discharge data for ten hydrometric stations of Lorestan province of Iran were used. In next step, different methods of sediment rating curves along with different correction factors, a total of 20 methods were applied to data. Results showed among examined methods, monthly rating curve with MUVE correction factor has minimum mean square error and maximum accuracy index. After applying the appropriate correction factor of sediment rating curves, the results of estimating sediment load by using sediment rating curve, were compared with the results of the Decision Tree. Mean-square error, MAE, P and the correlation coefficient were applied to select more appropriate method. The results showed the suitability of the sediment rating curves in compare with Decision Tree.

Keywords: CART algorithm, suspended sediment load, Decision Tree, sediment rating curves, corrective coefficient

مقدمه

انتقال رسوب و رسوبگذاری، پیامدهایی چون ایجاد جزایر رسوبی در مسیر رودخانه و در نتیجه کاهش ظرفیت انتقال جریانهای سیلابی، کاهش عمر مفید سدها و ظرفیت ذخیره مخازن، خوردگی تأسیسات سازهای رودخانهای و وارد شدن خسارات به ابنیه آبی و مزارع، رسوبگذاری در کف کانال و بسیاری مسایل و مشکلات دیگر را در بر دارد (حجه بخش و همکاران، ۱۳۹۰). به منظور اجرای برنامههای حفاظت خاک و کاهش رسوبزایی، همچنین محاسبه و طراحی دقیق حجم سد در احداث سدهای مخزنی، ضرورت دارد که میزان تولید رسوب در یک حوزه آبخیز، ارزیابی و برآورد گردد. در کنار اهمیت اندازهگیری رسوب، انجام این امر پرهزینه بوده و امکانات خاصی را طلب میکند. روشهای تخمین میزان رسوب را میتوان در دو گروه دستهبندی کرد (اکبری و طالبی، ۱۳۸۹). رویکرد روشهای سری اول، ایجاد مدلهای ریاضی با توجه به مفاهیم فیزیکی انتقال ذرات و حل معادلات هیدرودینامیک رسوب به همراه حل میدان جریان است .با توجه به اینکه این روشها به دادههای متنوعی نیازمند بوده و حتی اگر مدلی ریاضی نیز تبیین شود، دسترسی به دادههای لازم در اکثر موارد به آسانی امکان پذیر نخواهد بود، محققان در جستجوی راههای عملی برای پیشبینی رسوب برآمدهاند و تلاش آنها باعث ایجاد روش دیگری موسوم به روش منحنی سنجه رسوب شده است. همچنین بسیاری از روابط و تکنیکهای فعلی تخمین رسوب بر اساس روابط خطی بین متغیرها استوار است که به خوبی رفتار پیچیده و واقعی رسوب را پیشبینی نمیکنند (آبراهارت و همکاران، ۲۰۰۱). فرآیند بدست آوردن یک رابطه برای تخمین میزان رسوب، بطور اساسی یک مسأله غیرخطی است و شبکههای عصبی مصنوعی و درخت تصمیم به عنوان ابزاری توانمند در حل اینگونه مسائل به شمار میآیند. بنابراین استفاده از مدلهای غیرخطی مثل درخت تصمیم یک راه حل مناسب برای مدلسازی رفتار واقعی رسوب به نظر میرسد (حجه بخش و همکاران، ۱۳۹۰).

امروزه حجم عظیمی از دادههای رسوب معلق، توسط سازمانها و ارگانهای متولی امر در ثبت دادههای هیدرومتری، در مقیاسهای زمانی مختلف ثبت میگردد. نظر به گسترش سیستم پایگاه داده و ایجاد ابزارهای متعدد برای ذخیره حجم بالایی از اطلاعات، داده کاوی به عنوان یکی از شاخههای علوم بین رشتهای توسعه روزافزونی یافته است. داده کاوی به فرایند جستجو و کشف مدلهای گوناگون، مختصرسازی و اخذ مقادیر از مجموعهای از دادههای معلوم اطلاق میگردد (کانتاردزیک، ۱۳۸۵).

اکبری و طالبی (۱۳۸۹)، به بررسی کارایی درختان تصمیم در برآورد رسوب حوزه سد ایلام پرداخت و سپس نتایج این روش با منحنی سنجه رسوب مقایسه گردید. تحقیق مزبور نشان داد که در تمام شبیهسازیها دقت مدل درخت تصمیم نسبت به روش منحنی سنجه بالاتر بوده است.

حجه بخش و همکاران (۱۳۹۰) در تحقیقی با عنوان "برآورد بار رسوب بستر با استفاده از روش درخت تصمیمگیری رگرسیونی و

مقایسه با روش‌های تجربی" به این نتیجه رسیدند که درخت تصمیم در برآورد رسوبات بستر دارای دقت بیشتر میباشد.

شعبانی و شعبانی (۲۰۱۲) در مطالعه‌ای با "عنوان تخمین رسوبات معلق با استفاده از شبکه عصبی و منحنی سنجه رسوب در حوزه خوارستان" به این نتایج دست یافتند که شبکه عصبی مصنوعی مناسبتر از منحنی سنجه می باشد.

دستورانی و همکاران (۱۳۹۱) در مطالعه‌ای "برآورد رسوبات معلق با استفاده از شبکه عصبی مصنوعی (مطالعه موردی: حوزه آبخیز جامیشان استان کرمانشاه)" به این نتایج دست یافتند که شبکه عصبی مصنوعی نتایج قابل قبولی را جهت شبیه‌سازی بار معلق در ایستگاه حیدرآباد ارائه میکند، بطوریکه در مقایسه با منحنی سنجه سوب از دقت بالاتری برخوردار است.

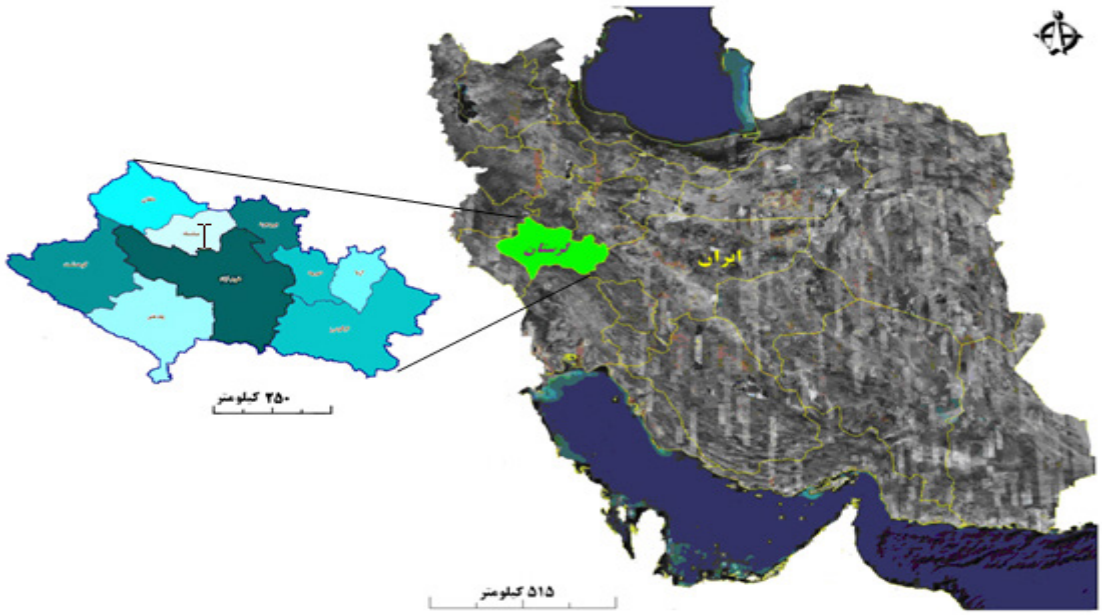
بهاتا چاریا و همکاران (۲۰۰۷)، از دو روش یادگیری ماشینی شامل شبکه عصبی مصنوعی و مدل درخت تصمیم جهت مدل‌سازی انتقال بارکل و بارکف رودخانه استفاده کردند و نتایج مدل‌های انتقال بارکف را با مدل‌های بگنولد، انیشتین، پارکر و همکاران و وان ریجن و همچنین نتایج مدلهای انتقال بارکل را با مدل‌های وایت ایکرز، بگنولد، انگلوند و هانسون و وان ریجن مقایسه کردند. طبق نتایجی که آنها گرفتند مدل‌های یاد گیری ماشینی دقت بیشتری نسبت به دیگر روش‌ها دارد و از بین مدل‌های مذکور، مدل درخت تصمیم کمترین خطا را نشان داد.

در زمینه‌های مختلف دیگر افرادی همچون محجوبی و صمغی (۱۳۸۸) و محجوبی و شهیدی (۲۰۰۸)، در ارتباط با پیش بینی ارتفاع موج شاخص، هیتوشی و همکاران (۲۰۰۹) در زمینه آنالیز حساسیت زمین لغزش در منطقه وسیع کوهستان‌های آکایشی ژاپن، روسجان و میکاس (۲۰۰۸) نقش متقابل وضعیت‌های فصلی و هیدرولوژیکی آبشویی نیترات، همچنین گیسن و همکاران (۲۰۰۷) در فرسایش آبی سطحی و زیرزمینی، زمازک و همکاران (۲۰۰۳) برای پیش‌بینی زمین لرزه و هندرسون (۲۰۰۵) در پیش‌بینی خصوصیات خاک گستره استرالیا، با استفاده از این روش تحقیقاتی را به انجام رساندهاند. نتایج این تحقیقات به طور عمده مدل درخت تصمیم را به عنوان ابزاری دقیق و قابل اعتماد معرفی کرده است.

بررسی تحقیقات پیشین، بیانگر این است که هر چند در معدود مواردی از مقایسه منحنی سنجه و درخت تصمیم استفاده شده است ولی تنها شکل متداول منحنی سنجه یعنی منحنی سنجه یک خطی مبنای مقایسه قرار گرفته که خالی از اشکال نیست، لذا در تحقیق حاضر، مناسبترین منحنی سنجه از بین منحنی سنجه رسوب فصلی، ماهانه، حد وسط دسته‌ها و روزانه انتخاب و با اعمال ضریب اصلاحی مناسب در مقایسه با مدل درخت تصمیم با الگوریتم کارت بررسی شد.

مواد و روش‌ها

موقعیت منطقه مورد مطالعه: استان لرستان در غرب ایران، در محدوده جغرافیایی '۵۱ °۴۶ تا '۳۰ °۵۰ طول شرقی و '۳۷ °۳۲ تا '۲۲ °۳۴ عرض شمالی قرار گرفته است. به منظور انجام این تحقیق از آمار دبی و رسوب متناظر ده ایستگاه هیدرومتری واقع در محدوده مورد مطالعه استفاده شد. شکل ۱ نشان دهنده موقعیت منطقه مورد مطالعه و جدول ۱ مشخصات ایستگاه‌های مورد مطالعه میباشد.



شکل ۱: موقعیت منطقه مورد مطالعه

شکل ۴ - شماتیکی از دیاگرام مدل WetSpa در مقیاس سلول (لیو، ۲۰۰۴)

ایستگاه ها	طول جغرافیایی	عرض جغرافیایی	حداکثر رسوب (تن در روز)	مساحت حوزه (Km^2)	ارتفاع حوزه (m)
تیره- درود	۴۵-۰۳-۴۹	۴۵-۲۸-۳۳	۳۲۴۲۰/۳۳	۳۴۰۰	۱۴۵۰
ماریره-دورود	۲۹-۰۴-۴۹	۲۸-۲۸-۳۳	۱۲۴۱۰۶/۹	۲۶۵۵	۱۴۵۰
آب سبزه	۳۱-۵۸-۴۸	۵۲-۳۳-۳۳	۲۳۷۸/۸۴	۳۴۵	۱۲۹۰
ماریره-دره تخت	۱۳-۲۳-۴۹	۴۰-۲۲-۳۳	۱۹۰۹۸/۹۱	۲۱۸۵	۱۸۰۰
ازنا	۱۰-۲۴-۴۹	۰۱-۲۴-۳۳	۴۹۸/۲۳	۲۰۱۰	۱۸۳۰
آبسرده	۱۶-۴۵-۴۸	۱۴-۴۴-۳۳	۱۶۲۳/۱۹	۲۲۳	۱۵۷۰
سیلاخور	۰۱-۴۸-۴۸	۰۰-۴۷-۳۳	۱۳۱۱۷/۰۷	۱۳۶۳	۱۴۹۰
گله رود	۲۴-۳۵-۴۸	۴۹-۵۴-۳۳	۳۰/۷۷	۶۰/۴	۲۰۰۰
سراب سفید	۵۷-۳۵-۴۸	۵۱-۵۴-۳۳	۳۷/۷۹	۶۴/۶	۱۹۸۰
بیاتون	۵۵-۵۸-۴۸	۲۵-۴۲-۳۳	۲۱۷/۷۴	۱۲۰	۱۶۲۰

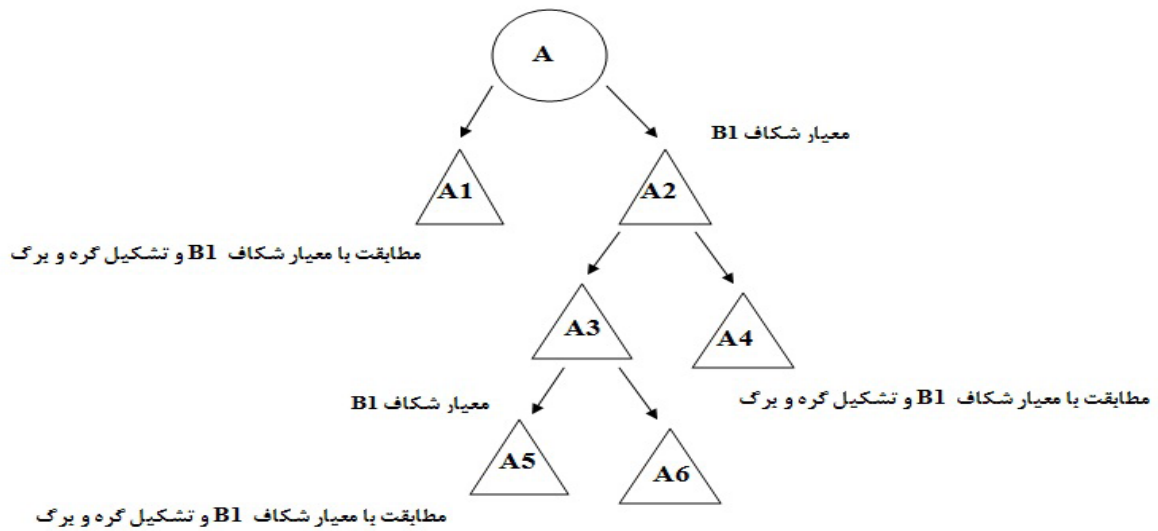
دادههای مورد بررسی

در این مطالعه از دادههای دبی به عنوان ورودی و رسوب به عنوان خروجی مدل‌ها استفاده شد. داده‌ها در بازه زمانی ۸۵ - ۱۳۷۵ به صورت روزانه، ماهانه، حد وسط دسته‌ها (دسته‌بندی شده) و فصلی از ده ایستگاه منطقه مورد مطالعه، مورد استفاده قرار گرفت. ۷۰٪ دادهها به عنوان دادههای آموزش و ۳۰٪ به منظور تست در نظر گرفته شد. به دلیل پراکنش زیاد دادههای رسوب نسبت به دادههای دبی، ابتدا به نرمال سازی دادههای رسوب پرداخته شد. به منظور نرمال سازی از رابطه ۱در بازه ۰/۱-۰/۹ استفاده گردید.

$$T= Tmin+ (Tmax-Tmin)(X-Xmin)/(Xmax-Xmin) \quad (1)$$

T: مقادیر نرمال شده داده های واقعی، X: داده‌های رسوب مشاهده شده واقعی، xmin و xmax کمینه و بیشینه داده‌های واقعی، Tmax و Tmin بیشینه و کمینه داده‌های هدف که در اینجا به ترتیب ۰/۹ و ۰/۱ است.

مدل درخت تصمیم: درخت تصمیم یکی ازروش‌های داده کاوی و از ابزارهای قوی و متداول برای دسته‌بندی و پیشبینی میباشد که برخلاف شبکه‌های عصبی به تولید قانون میپردازد. یعنی درخت تصمیم پیشبینی خود را در قالب یکسری قوانین توضیح می‌دهد شکل (۲). در حالی که در شبکه‌های عصبی تنها پیشبینی بیان میشود و چگونگی آن در خود شبکه پنهان باقی می‌ماند. علاوه بر آن در درخت تصمیم بر خلاف شبکههای عصبی، می‌توان از دادههای غیرعددی نیز استفاده نمود (برایمن و همکاران، ۱۹۸۴).



شکل ۲ : نمونهای از یک درخت تصمیم ساده

الگوریتم کارت: الگوریتم CART که یکی از معروفترین روش‌های انجام درخت تصمیم است به وسیله (برایمن و همکاران، ۱۹۸۴) ایجاد گردیـد. در اینجـا تعدادی متغیر وجود دارد که دسـته آن‌ها از قبل معلوم می‌باشـد (متغیر وابسـته در آنها معلوم می‌باشد) هدف تهیه درختی اسـت که بتوان به وسیله آن متغیر وابسته یا همان کلاس را برای یک متغیر جدید پیش‌بینی و تعیین نمود.

انواع منحنی سنجه رسوب: مدل یک خطی: در روش منحنی سنجه رسوب یک خطی، داده‌های موجود از اندازه‌گیری رسوب و دبی متناظر با آن به صفحه مختصات لگاریتمی منتقل شده و بهترین خط برازش بـر مبنای روش حداقل مربعات خطا از میان آنها عبور داده میشـود. سپس یک رابطه رگرسیونی به صورت معادله توانی استخراج میگردد که به منحنی سنجه رسوب یک خطی معروف است.

مدل ماهانه: در این مدل برآورد رسـوب معلق مشابه مدل یک خطی میباشد ولی تفکیک داده‌ها به صورت ماهانه مورد مطالعه قرار می‌گیرد.
مـدل فصلی: در این مدل برآورد رسـوب معلق مشـابه مدل خطی می‌باشـد ولی تفکیک داده‌ها بـه صورت فصلی مـورد مطالعه قرار می‌گیرد.
مدل حدوسـط داده‌ها: در این روش ابتدا دبی‌های جریان با یک نمو معین به تعدادی دسته تقسیم می‌شوند و برای دبی متوسط هر دسته، میانگین رسوب همان دسـته تعیین گردیده و منحنی سنجه رسوب با استفاده از آنها به دست می‌آید (جانسون ۱۹۹۶). در تحقیق حاضر جهت برآورد رسـوب معلق بر اساس مدل حد وسـط دسته‌ها، ابتدا دبیهای جریان بر اسـاس یک فاصله معین به ۱۲دسته تقسیم شدند و برای متوسط هر دسته، میانگین رسوب همان دسته تعیین گردید. برای اینکه رابطه سـنجه رسـوب، برآورد دقیق‌تـر و اریب کمتری نشـان دهد، تاکنون روش‌های مختلفی توسط محققین به کار گرفته شده اسـت. در این راستا ضرایب اصلاحی شامل FAO، QMLE، Smearing و MVUE در راسـتای بهبـود منحنی‌های سـنجه پیشنهاد شده است.

روش ضریـب اصلاحی فائو: در این روش ابتدا خطی به روش حداقل مربعات به مقادیر لگاریتمی غلظت و دبی متناظر آنها برازش می‌شود ولـی از معادله این خط که از نقاط میانگین log Qs و میانگین log Qw عبور می‌کند، اسـتفاده نمی‌شـود و به جای آن از معادله خطی مـوازی که از میانگین Qs و میانگین Qw عبور می‌کند، اسـتفاده می‌شـود بنابراین معادله جدید به صورت رابطـه (۲) خواهد بود. که از تقسیم میانگین دبی‌های رسـوب متناظر اندازه‌گیری شده بدست می‌آید.

$$Q_S = a' Q_w^b \qquad (۲)$$

ضریب اصلاحـی پارامتری: برای اصـلاح خطای ناشـی از تبدیل لگاریتمی داده‌ها، توماس (۱۹۸۵) به نقل از میلر ضریب اصلاحی زیر را پیشنهاد داده است که s مقدار انحراف معیار داده‌هاست.

$$CF_1 = \exp(2.65\, s^2) \qquad (۳)$$

ضریب اصلاحی غیر پارامتری: کوچ و اسـمیل (۱۹۸۶) با اذعان به این نکته که منحنی سنجه باعث برآورد کمتر مقدار رسوب می‌گردد، فاکتور اصلاحی غیر پارامتری زیر را برای تصحیح مقدار رسـوب ارائه داده‌اند.

$$C_{FZ} = \frac{1}{n}\sum 10^{\varepsilon i} \qquad (۴)$$

$$\varepsilon i = \log C_O - \log C_e \qquad (۵)$$

در ایـن روابط خطای برآورد و Ce غلظت رسـوب برآوردی (تن در روز)، Co غلظت رسوب مشاهده‌ای (تن در روز) میباشد.
ضریب اصلاحی MVUE

$$V = \frac{1}{N} + \left[\frac{(\ln(Q_x) - Q_{Bar})^2}{Q_{Var}}\right] \qquad (۶)$$

$$g_{mi} = \left(\frac{m+1}{2m}\right)\left[(1-V)S^2\right] \qquad (۷)$$

$$CF_{MVUE} = \left[\frac{\sum_{i=1}^{n} g_{mi}}{n}\right] \qquad (۸)$$

کـه در روابط فـوق، تابع فینـی که در حقیقـت نقش ضریب اصلاحی نمونه i را دارد. m درجه آزادی معادله رگرسـیونی (n-۱)، خطای اسـتاندارد ، Vتابعی از متغیرهای توزیعی، متوسط دبی‌های مشـاهده‌ای، واریانس دبی‌های جریان مقدار دبی مشاهده‌ای نمونه i می‌باشد.
پس از بـه کارگیری ضرایب اصلاحی برای انواع منحنی سـنجه رسـوب، با استفاده از ۳۰ درصد باقیمانده داده ها و روشهای مختلف محاسبه خطا شامل R2،MAE ، P و RMSE(معادلات ۹،۱۰، ۱۱ و ۱۲) بهترین روش منحنی سـنجه رسـوب در منطقه مذکور جهت برآورد بار رسوب معلق انتخاب شد.

سـنجش میزان کارآیی مدل: از تجزیه و تحلیل خطاهای باقیمانده و تفاوت‌هـای بین مقادیر اندازه‌گیری شـده و برآورد شـده می‌توان در ارزیابـی کارآیـی مدل بهره برد .شـمار زیـادی از ملاک‌ها برای ارزیابـی کارآیی مدل وجود دارند .در پایـان، نتایج حاصل از مدل‌ها (خروجی‌ها) و رخدادهای واقعی رسـوب با همدیگر مقایسـه شده و از شـاخص معیارهای همبستگی (R^2)، مجذور میانگین مربعات خطا (RMSE)، میانگین قـدر مطلق خطا (MAE) و ضریب دقت مدل

(P) بر ای نشان دادن عملکرد مدلها استفاده شده است (عرب خدری و حکیم خانی، ۱۳۸۲).

$$R^2 = \frac{\sum_{k-1}^{k} X_k Y_k}{\sqrt{\sum_{k-1}^{k} X_k^2 \sum Y_k^2}} \quad (۹)$$

$$RMSE = \sqrt{\frac{\sum_{k-1}^{k} (X_k - Y_k)^2}{K}} \quad (۱۰)$$

$$MAE = \frac{1}{n}\sum_{t=1}^{n} |O_i - P_i| \quad (۱۱)$$

$$P = \frac{\sum_{i=1}^{n} (SSC_O / SSC_e)}{n} \quad (۱۲)$$

به‌طوری‌که در این روابط Xk مقدار مشاهداتی، Yk مقدار برآورد شده، oi و pi به ترتیب مقادیر مشاهده شده و شبیه سازی شده، n و K تعداد نمونه‌ها، غلظت رسوب برآوردی و غلظت رسوب مشاهده‌ای می‌باشد.

در تمام روش‌های ذکر شده، از غلظت رسوب و دبی متناظر با آن استفاده شده است. به منظور ترسیم منحنی‌های سنجه از محیط نرم‌افزاری Excel استفاده شد و محاسبات مربوط به آن‌ها نیز در همین نرم‌افزار انجام شد و بین دبی و رسوب معادله نمایی به دست آمد. همچنین با استفاده از نرم افزار متلب، داده‌های مورد استفاده را به مدل درخت تصمیم وارد نموده و میزان ضریب همبستگی و میانگین مربعات خطا به دست آمد. ضریب همبستگی و ضریب دقت نزدیک به یک، میانگین مربعات خطا و میانگین قدرمطلق خطا نزدیک به صفر باشد، مدل از خطای کمتری برخوردار است.

نتایج و بحث

به منظور مقایسه نتایج مدل درخت تصمیم با نتایج منحنی سنجه رسوب فصلی، یک خطی، ماهانه و حد وسط دسته‌ها (جانسون)، ابتدا هر کدام از انواع منحنی سنجه رسوب در بازه ده ساله ترسیم شد که جواب‌های به دست آمده از خطای بالایی برخوردار بودند. با استفاده از ضرایب اصلاحی متداول شامل (Smearing, FAO, MUVE, QMLE) که در بخش مواد و روش‌ها مفصل آمده است، اقدام به کاهش خطا و اصلاح منحنی‌های سنجه گردید (جدول ۲). در اکثر ایستگاه‌ها ضریب اصلاحی MUVE با استفاده از شاخص آماری میانگین مربعات خطا و شاخص دقت، مناسب‌ترین ضریب اصلاحی و منحنی سنجه ماهانه به عنوان مناسب‌ترین نوع منحنی سنجه به منظور تخمین رسوب بار معلق، انتخاب شد.

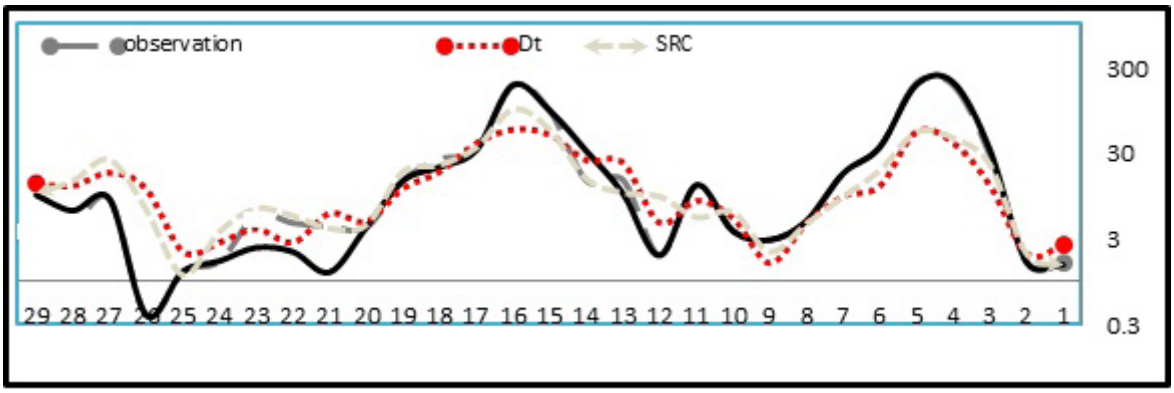
بعد از انتخاب مناسب‌ترین ضریب اصلاحی، انواع منحنی‌های سنجه (با اعمال مناسب‌ترین ضریب اصلاحی) که مورد استفاده قرار گرفته شده بود با مدل درخت تصمیم مورد مقایسه قرار گرفت (جدول ۳). به این منظور داده‌ها به همان صورت که در انواع منحنی سنجه مورد استفاده قرار گرفته شده بود به مدل درخت تصمیم اعمال شد. مثلا داده‌ها به صورت روزانه، ماهانه، فصلی و حد وسط دسته‌ها وارد مدل درخت تصمیم گردید.

جدول ۲- نتایج اعمال ضرایب اصلاحی در مدل منحنی سنجه رسوب ماهانه (مناسبترین منحنی سنجه)

	بدون ضریب		FAO		CF1		CF2		MUVE	
	RMSE	P	RMSE	P	RMSE	P	RMSE	P	RMSE	P
تیره-دورود	۳۲۳/۳۵	۳۱/۸۸	۱۷۳۲	۱۲/۰۸	۶۹۱۹	۲/۷۸	۷۰۱۶	۲/۲	۱۸۱/۶۴	۰/۳۱
ماریره-دورود	۳۴/۹۶	۲/۵۲	۳۵/۸۲	۰/۳۶	۴۲/۵۳	۱/۸۷	۴۹/۵۶	۰/۸۵	۱۲/۸	۰/۸۴
آبسبزه	۱۰/۸۶	۵/۹۴	۳۱/۶۳	۱/۷۴	۲۴۲/۴۵	۳/۹۳	۲۵۹/۰۸	۱/۴۵	۷/۵۹	۰/۴۲
ازنا	۱۹/۹۹	۷۶/۸۲	۲۵/۷۶	۱۴/۱۹	۱۱۱۴/۸۲	۲/۵۸	۱۱۲۶/۲۴	۱/۸۰	۱۹/۳۷	۰/۳۳
آبسرده	۵/۱۱	۱۷/۵۸	۲۲/۳۲	۲/۳۸	۲۳۲/۸۴	۱/۱۲	۲۳۰/۵۳	۶/۲۵	۴/۵۸	۰/۷۷
سیلاخور	۹/۹۷	۱۴/۷۲	۱۸/۰۳	۰/۵۶	۱۶۲/۰۲	۰/۲۹	۱۶۴/۰۸	۰/۲۱	۸/۱۲	۰/۳۱
گله رود	۱/۹۱	۱/۶۷	۲/۰۴	۴/۱۹	۳/۹۶	۲/۰۳	۴/۸۱	۰/۵۷	۰/۶۱	۰/۸۲
سراب سفید	۰/۶۳	۲/۸۷	۰/۷۴	۰/۴۲	۵/۵۳	۰/۷۲	۵/۶۶	۰/۶	۰/۵۴	۰/۴۱
بیاتون	۴/۹۱	۴۸/۳۸	۱۱/۵۵	۲/۰۴	۴۴/۶۳	۰/۳۷	۴۵/۱۶	۰/۱۱	۰/۸۸	۴
ماریره- دره تخت	۳/۶۴	۲۷/۹۶	۱۶۰/۶۶	۵/۵	۲۲۴۴	۲/۹	۲۲۹۲	۱/۹	۴۰/۰۳	۰/۰۱۲

جدول ۳: نتایج تخمین رسوبات معلق با استفاده از شاخصهای آماری در ایستگاههای مورد مطالعه

نام ایستگاه	مدل‌ها	درخت تصمیم				منحنی سنجه رسوب			
		ضریب همبستگی	میانگین مربعات خطا	میانگین قدرمطلق خطا	ضریب دقت	ضریب همبستگی	میانگین مربعات خطا	میانگین قدرمطلق خطا	ضریب دقت
تیره-دورود	فصلی	۰/۹۲	۰/۰۲۱	۰,۰۰۱۵	۰,۴۷	۰/۹۴	۰/۰۴	۵۰۶,۸۸	۰,۵۷
	خطی	۰/۸۰	۰/۰۰۷۲	۰,۰۰۱۷	۰,۳۹	۰/۹۲	۰/۰۳	۵۹۹,۹۰	۰,۰۳
	ماهانه	۰/۹۵	۰/۰۰۰۲۷	۷۹۹,۸۴	۰,۴۰	۰/۹۴	۰/۰۲۴	۲۸,۹	۰,۵۸
	جانسون	۰/۸۰	۰/۰۰۷۲	۰,۰۰۱۷	۰,۳۹	۰/۹۲	۰/۰۹	۵۶۶,۸۳	۰,۰۱
ماریره-دورود	فصلی	۰/۶۹	۰/۰۱۵	۸۵۹,۶۸	۰,۴۵	۰/۸۸	۰/۰۱۸	۱۰۴۲,۹۵	۰,۳۵
	خطی	۰/۳۰	۰/۰۱۸	۷۹۹,۶۵	۰,۴۳	۰/۹۴	۰/۰۲۰	۱۰۷۹,۱۳	۰,۱۸
	ماهانه	۰/۲۰	۰/۱۴	۴۶,۹۹	۰,۳۳	۰/۶۷	۰/۰۳	۱,۹۰	۰,۸۵
	جانسون	۰/۳۰	۰/۰۱۸	۷۹۹,۶۵	۰,۴۳	۰/۹۷	۰/۰۱۶	۱۰۵۴,۷۴	۰,۱۵
آبسبزه	فصلی	۰/۸۵	۰/۰۰۰۱۶	۵۳,۵۰	۰,۳۴	۰/۷۲	۰/۰۲۸	۳۲,۵۵	۰,۹۳
	خطی	۰/۸۷	۰/۰۰۵۴	۸۶,۲۷	۰,۴۰	۰/۴۳	۰/۱۲	۶۷,۴۶	۰,۲۴
	ماهانه	۰/۵۶	۰/۰۰۱۴	۳۲,۱۰	۰,۳۷	۰/۸۰	۰/۰۲۴	۱,۳۱	۰,۴۵
	جانسون	۰/۸۷	۰/۰۰۵۴	۸۶,۲۷	۰,۴۰	۰/۴۷	۰/۳۴	۶۶,۷۲	۰,۱۳
ازنا	فصلی	۰/۹۱	۰/۰۰۰۰۰۹۴	۵۲,۱۸	۰,۲۹	۰/۹۸	۰/۰۱۹	۲۶۹,۴۶	۰,۶۰
	خطی	۰/۶۱	۰/۰۰۰۵۷	۱۳۲,۲۴	۰,۵۴	۰/۹۵	۰/۰۲۳	۲۷۸,۱۰	۰,۸۳
	ماهانه	۰/۶۹	۰/۰۰۰۰۰۰۰۱	۵,۵۳	۰,۳۸	۰/۹۸	۰/۰۲۲	۵,۲۹	۰,۶۹
	جانسون	۰/۶۱	۰/۰۰۰۵۷	۱۳۲,۲۴	۰,۵۴	۰/۹۷	۰/۰۲۰	۲۷۱,۵۸	۰,۷۳
آبسرده	فصلی	۰/۳۳	۰/۰۰۵	۲۳,۸۲	۰,۴۶	۰/۶۰	۰/۱۸	۲۴,۹۹	۰,۸۹
	خطی	۰/۶۳	۰/۰۰۶۶	۳۵,۲۹	۰,۳۳	۰/۶۲	۰/۳۱	۲۴,۶۲	۰,۸۰
	ماهانه	۰/۵۴	۰/۰۰۰۰۱۱	۱۱,۰۳	۰,۴۵	۰/۹۷	۰/۰۲۲	۲,۵۶	۰,۸۰
	جانسون	۰/۶۳	۰/۰۰۶۶	۳۵,۲۹	۰,۳۳	۰/۶۰	۰/۳۵	۲۴,۲۰	۰,۲۴
سیلاخور	فصلی	۰/۶۷	۰/۰۱۷	۷۸,۸۷	۰,۳۵	۰/۶۳	۰/۱۵	۳۷,۳۶	۰,۳۷
	خطی	۰/۷۴	۰/۰۰۰۲۴	۱۴,۰۷	۰,۴۰	۰/۷۰	۰/۱	۳۷,۰۸	۰,۶۸
	ماهانه	۰/۱۰	۰/۱۴	۲۱۰,۹۳	۰,۴۰	۰/۲۰	۰/۲۳	۰,۶۳	۰,۲۵
	جانسون	۰/۷۴	۰/۰۰۰۲۴	۱۴,۰۷	۰,۴۰	۰/۷۳	۰/۰۱۲	۳۶,۸۳	۰,۱۸
گله رود	فصلی	۰/۱۳	۰/۰۲۱	۴۴۳,۱۱	۰,۳۹	۰/۴۶	۰/۱۸	۲۷۰,۵۱	۰,۷۹
	خطی	۰/۷۴	۰/۰۰۴۵	۵۱۷,۷۲	۰,۳۳	۰/۴۴	۰/۳۹	۲۷۶,۲۹	۰,۴۷
	ماهانه	۰/۹۵	۰/۰۰۳۹	۳۰,۳۹	۰,۳۱	۰/۸۱	۰/۰۳۷	۱,۲۵	۰,۲۱
	جانسون	۰/۷۴	۰/۰۰۴۵	۶,۹۶	۰,۳۱	۰/۴۵	۰/۳۱	۲,۹۰	۰,۲۶
سراب سفید	فصلی	۰/۵۴	۰/۰۰۵۱	۰,۱۴	۰,۳۰	۰/۶۴	۰/۱۱	۵,۰۳	۰,۴۶
	خطی	۰/۶۵	۰/۰۶۲	۳,۰۲	۰,۳۰	۰/۶۳	۰/۳۱	۳,۶۶	۰,۸۹
	ماهانه	۰/۶۸	۰/۰۴	۲,۳۸	۰,۳۰	۰/۵۰	۰/۱۸	۰,۱۰	۰,۵۷
	جانسون	۰/۶۵	۰/۰۶۲	۳,۰۲	۰,۳۰	۰/۶۴	۰/۲۹	۰,۱۵	۰,۴۵
بیاتون	فصلی	۰/۰۳	۰/۰۰۰۲۷	۴,۹۲	۰,۳۴	۰/۲۸	۰/۲	۶,۹۵	۰,۴۳
	خطی	۰/۱۱	۰/۰۳۷	۶,۷۴	۰,۳۸	۰/۲۴	۰/۱۴	۶,۳۵	۰,۶۸
	ماهانه	۰/۱۱	۰/۱۵	۳,۳۱	۰,۳۱	۰/۱۰	۰/۰۶۹	۰,۰۸	۰,۶۰
	جانسون	۰/۱۱	۰/۰۳۷	۶,۷۴	۰,۳۸	۰/۲۲	۰/۱۳	۶,۲۷	۰,۴۰
ماریره- دره تخت	فصلی	۰/۳۶	۰/۰۱۴	۹,۵۵	۰,۳۳	۰/۴۸	۰/۱۹	۵,۶۹	۰,۱۵
	خطی	۰/۱۵	۰/۰۷۴	۸,۹۹	۰,۳۵	۰/۴۸	۰/۲۱	۵,۲۶	۰,۱۰
	ماهانه	۰/۵۸	۰/۰۰۰۱۷	۱,۴۸	۰,۴۷	۰/۷۱	۰/۰۲۳	۰,۱۴	۰,۱۱
	جانسون	۰/۱۵	۰/۰۷۴	۸,۹۹	۰,۳۵	۰/۴۸	۰/۱۹	۵,۴۱	۰,۱۵



شکل ۳: مقایسه مدل درخت تصمیم و منحنی سنجه ماهانه (مناسبترین نوع داده) شبیه سازی شده با رسوب مشاهده شده

همچنین همان‌طور که در شکل ۲ مشاهده میشود به عنوان نمونه رسوب مشاهده شده در ایستگاه دره تخت با مدل منحنی سنجه ماهانه با ضریب اصلاحی MUVE و درخت تصمیم شبیه سازی شده مورد مقایسه قرار گرفت. منحنی رسوب ماهانه شبیه سازی شده توسط درخت تصمیم در مقایسه با منحنی سنجه ماهانه از تطابق بیشتری با رسوب مشاهده شده برخوردار است.

نتیجه‌گیری

در این پژوهش از دادههای دبی و رسوب ده ایستگاه استان لرستان به صورت ماهانه، فصلی، روزانه و حدوسط دستهها استفاده شد. ابتدا دادهها دبی به عنوان ورودی و رسوب به عنوان خروجی به انواع مدل منحنی سنجه رسوب اعمال شد و با شاخصهای آماری ضریب دقت (P) و میانگین مربعات خطا (RMSE) مورد سنجش قرار گرفت. نتایج منحنی سنجههای رسوب با اعمال ضریبهای اصلاحی بهبود پیدا نمود که مناسبترین ضریب، ضریب اصلاحی MUVE و منحنی سنجی ماهانه به عنوان مناسبترین منحنی سنجه انتخاب شد. در مرحله بعد دادههای فصلی، روزانه، ماهانه و حد وسط دستهها به مدل درخت تصمیم با الگوریتم کارت اعمال شد. نتایج حاصل از مدل درخت تصمیم با انواع مدلهای منحنی سنجه رسوب با اعمال ضرایب اصلاحی MUVE با شاخصهای آماری میانگین مربعات خطا، ضریب دقت، میانگین قدر مطلق خطا و ضریب همبستگی مورد مقایسه قرار گرفت. نتایج حاکی از تفاوت تخمین رسوبات معلق در دو مدل میباشد. در بعضی از ایستگاهها منحنی سنجه مناسبتر و در بعضی دیگر مدل درخت تصمیم به عنوان روش مناسب انتخاب شد. ولی در کل با توجه به مناسبتر بودن منحنی سنجه در بیشتر ایستگاهها، میتوان منحنی سنجه را مناسبتر از مدل درخت تصمیم دانست. اکبری و طالبی (۲۰۱۰) و حجه بخش و همکاران (۲۰۱۱) در مقایسه منحنی سنجه رسوب و درخت تصمیم، درخت تصمیم را به عنوان مدلی مناسب در برآورد رسوب معلق معرفی نموده و کارآیی آن را بیشتر از منحنی سنجه ارزیابی نمودند (لازم به ذکر است تنها از یک ایستگاه در مطالعاتشان استفاده نمودهاند) ولی نتایج بررسی حاضر با نتایج این محققین مغایر می‌باشد، دلیل این مغایرت، تفاوت در نحوه استفاده از منحنی سنجه باشد. بدین معنی که در بررسی‌های قبلی، تنها از شکل متداول منحنی سنجه یعنی منحنی سنجه یک خطی به عنوان مبنای مقایسه استفاده شده است در حالی‌که در تحقیق حاضر، مناسبترین منحنی سنجه از بین منحنی سنجه رسوب فصلی، ماهانه، حد وسط دستهها و روزانه انتخاب و با اعمال ضریب اصلاحی مناسب در مقایسه با مدل درخت تصمیم با الگوریتم کارت مورد بررسی قرار گرفت که این نحوه انتخاب مناسب‌ترین منحنی سنجه میتواند دلیل برتری منحنی سنجه رسوب در برآوردهای این مطالعه باشد.

منابع مورد استفاده

1. Abrahart, RJ., White, SM., 2001. Modeling sediment transfer in Malawi: Comparing backpropagation neural network solutions against a multiple linear regression benchmark using small data set. Phys Chem Earth B.24-19:(1)26pp.
2. Akbari, Z. Talebi, A. 2010, Regression model to estimate the performance of decision tree Ilam dam basin sediments, MS Thesis, School of Natural Resources and the desert, Yazd University.
3. ArabKhedri, M. Hakim Khani, S. 2003, the validity of the extra interpolation to estimate the mean annual suspended sediment hydrometric stations in the country). Journal of Agriculture and Natural Resources, 123-131 :3:11pp.
4. Barzegar, F., 2004, Comparison of methods to estimate suspended sediment (Case Study: Qezel Ozan), MS Thesis, Department of Natural Resources, Tehran University, 65 pp.
5. Bhattacharya, B., Price R.K, and Solomatine D.P., 2007. Machine learning approach to modeling sediment transport, Journal of Hydraulic Engineering. 450-440 ,(4)133pp.

6. Breiman, L., Friedman J., Olshen R., and Stone C.,1984, Classification and Regression Trees, Chapman & Hall/CRC Press, Boca Raton, FL.
7. Dastorani, M.. Huge, Str. Talebi, A. and dedicated CE., 2012,Suspended sediment estimation using Artificial Neural Network) Case Study: Jamyshan watershed in Kermanshah , Watershed management Journal, the third year of 75-61 :6pp.
8. Geissen, V., Kampichler, C,. 2007. Superficial and subterranean soil erosion in Tabasco, tropical Mexico: development of a decision tree modeling approach. Geoderma 287-277 ,139.
9. Hajjahbakhsh, P., Talebi, A., Mahjoub, J., 2011, "Estimation of bed load sediment using decision trees, regression and compared with the experimental method", Student thesis, Department of Natural Resources, Yazd University.
10. Henderson, B.L., E.N. Bui, C.J. Moran, D.A.P. Simon. 2005. Australia-wide predictions of soil properties using decision trees. Geoderma ,124 398–383.
11. Hitoshi Saito, 2009. Daichi Nakayama and Hiroshi Matsuyama. Comparison of landslide susceptibility based on a decision-tree model and actual landslide occurrence: The Akaishi Mountains, Japan, geomorphology, 121-108.
12. Jansson, M.B.1996. Estimating a sediment rating curves of the Reventazon river at Palamo using logged mean loads within discharge classes. Journal of Hydrology . 241-4.227 :183pp.
13. Kantardzyk, M., 2006, Data mining, Translator: Alikhanzadeh, A.,Computer Science Press.
14. Kao, Sh.J., Lee, T.Y., and Milliman, J.D. 2005. Calculating highly fluctuated suspended sediment fluxes from mountainous rivers in Taiwan. TAO. 675-16:3.653pp.
15. Koch, R.W. and G.M.Smillie.1986,Comment on "river loads underestimated by rating curves". Water Resources Research .Vol.22,No.-13,2121 2122p.
16. Mahjoobi, J., Ardalan-samghi, H. 2009. Forecasting for wind parameters waves hind casting of with use from Decision trees approach and Artificial neural network method. Sea Engineer Issue. Five year. Number 9.
17. Mahjoobi, J., Etemad-shahidi, A. 2008. An alternative approach for the prediction of significant wave heights based on classification and regression trees. Applied Ocean Research 177-172, 30.
18. Rusjan , S., Micos, M. 2008. Assessment of hydrolpgical and seasonal controls over the nitrate flushing from a forested watershed using a data mining technique. Hydrol.Earth syst. Sci., -645 ,12 656.
19. -Shabani, M., Shabani, N., 2012, Estimation of Daily Suspended Sediment Yield Using Artificial Neural Network and Sediment Rating Curve in Kharestan Watershed, Iran, Australian Journal of Basic and Applied Sciences, 164-157 :(12)6pp.
20. Thomas, R.B. 1985.Estimating total suspended sediment yield with probability sampling. Water Resources Research.1388-21:1381pp.
21. Zmazek, B., L. Todorovski, S. Dmeroski, J. VaupotiW, I. Kobal.. Application of decision trees to the analysis of soil radon data for earthquake prediction. Applied Radiation and Isotopes ,58 706–697 ,(2003)