# مدل‌سازی بار رسوب کل رودخانه‌ها با استفاده از برنامه‌نویسی ژنتیک

مهناز اسکندری [۱]، امین فلامکی [۲]، میلاد نخعی [۳]
۱-دکتری خاکشناسی و محقق موسسه تحقیقات خاک و آب، ۲-استادیار گروه مهندسی دانشگاه پیام نور تهران،۳-دانشجوی دکتری مهندسی عمران ژئوتکنیک دانشگاه علم و صنعت تهران

**چکیده**
مدلسازی انتقال بار رسوب به دلیل ماهیت چندبعدی و وابستگی غیرخطی متغیرها با یکدیگر، مسأله‌ای چالش‌برانگیز است. برای این منظور از آنالیز رگرسیون در گذشته بسیار استفاده شده، لیکن همواره محدودیتهای بسیاری در توابع بدست آمده وجود داشته است. در این پژوهش از برنامه‌نویسی ژنتیک به منظور مدلسازی بار رسوب کل استفاده شد. داده‌های مورد استفاده شامل ۲۰۰ نمونه آزمایشگاهی بود که ۷۵ درصد آنها به منظور آموزش و ۲۵ درصد برای آزمون مدل استفاده شدند. متغیرهای سرعت جریان، عرض کانال، عمق جریان، شیب خط انرژی و قطر میانه اندازه ذرات به عنوان ورودی مدل و غلظت رسوب کل به عنوان خروجی مدل در نظر گرفته شدند. نتایج نشان داد که مدل طراحی شده قادر است با ضریب تبیین ۹/۰ و RMSE برابر با ۷/۴۲، غلظت رسوب کل را با دقت مطلوب برآورد نماید.

واژه های کلیدی: انتقال رسوب، برنامه‌نویسی ژنتیک، غلظت رسوب کل، مدل هوشمند، مهندسی رودخانه

**مقدمه**
برآورد بار رسوب کل رودخانه‌ها از مسائل مهم و کاربردی در مهندسی عمران و برنامه‌ریزی منابع آب است که اطلاع از آن برای بسیاری از کاربری‌ها مانند طراحی سدها و حوضچه‌های ذخیره آب سطحی، انتقال رسوب و آلودگی در رودخانه‌ها و دریاچه‌ها، طراحی کانال‌ها و نگهداری آن‌ها، حیاتی است. غلظت رسوب می‌تواند به روش‌های مستقیم و یا غیرمستقیم محاسبه شود که معمولاً روش‌های مستقیم پرهزینه و زمان‌بر هستند. همچنین بار رسوب کل می‌تواند به کمک روابط مختلف انتقال رسوب محاسبه شود، لیکن به طور معمول کاربرد این روابط نیاز به شرایط معینی داشته و به علاوه در بیشتر موارد نتایج حاصل از آن‌ها با یکدیگر و با مقادیر اندازه‌گیری شده متفاوت است (فلامکی و همکاران، ۱۳۹۲). در بسیاری از موارد نتایج حاصل از این روابط به دلیل پیچیدگی و عدم شناخت کامل و دقیق مکانیزم انتقال رسوب، با مقادیر اندازه‌گیری شده تطابق نداشته و در مواردی نیز نتایج روابط مختلف تفاوت معنی‌دار با یکدیگر دارند (معیری و همکاران، ۱۳۸۹). این مسأله باعث شده است که در دهه‌های اخیر پژوهشگران مختلف به سمت استفاده از روش‌های هوشمند مانند شبکه‌های عصبی مصنوعی (ANN)[۱۵۱] و برنامه‌نویسی ژنتیک (GP)[۱۵۲] روی آورند. به عنوان مثال Aytek و (۲۰۰۸) Kisi و Kisi و (۲۰۱۲) Shiri برای پیش‌بینی مقدار انتقال بار رسوب معلق از GP و Singh و همکاران (۲۰۰۷)، از ANN-GP استفاده کرده‌اند. بیشتر پژوهش‌های انجام شده، در راستای مدلسازی مقدار بار رسوب معلق انجام گرفته و تنها معدود مواردی مانند Dogan و همکاران (۲۰۰۷)، Yang و همکاران (۲۰۰۹) و فلامکی و همکاران (۱۳۹۲)، به برآورد مقدار رسوب کل رودخانه‌ها پرداخته‌اند.
رسوب کل شامل رسوب معلق و بار بستر است که در بیشتر ایستگاه‌های هیدرولوژی، تنها بار معلق رسوب اندازه‌گیری می‌شود. سپس بار بستر به صورت درصدی از بار معلق تخمین زده می‌شود، لیکن این روش محاسبه قابل اعتماد نیست. اطلاع دقیق از مقدار بار رسوب کل رودخانه‌ها در بسیاری از پژوهش‌های منابع آب به‌ویژه برآورد حجم رسوب‌گذاری در مخازن سدها ضروری است. بنابراین در این پژوهش به منظور مدلسازی برآورد مقدار بار رسوب کل، از برنامه نویسی ژنتیک استفاده شد.

**مواد و روش‌ها**
داده‌های مورد استفاده در این پژوهش، شامل ۲۰۰ نمونه آزمایشگاهی بود که از بانک داده‌های رسوب Molinas و (۲۰۰۳) Wu، تهیه شد. این نمونه‌ها شامل مجموعه‌ای از داده‌های مورد نیاز در انتقال رسوب مانند متوسط سرعت ($V$) برحسب ($ft/s$)، عرض کانال ($W$) برحسب ($ft$)، عمق جریان ($D$) بر حسب ($ft$)، S شیب خط انرژی ($ft/ft$)، قطر میانه اندازه ذرات ($d_{50}$) برحسب ($ft$)، و غلظت رسوب ($Ct$) بر حسب ($ppm$) بود.
متغیرهای سرعت جریان، عرض کانال، عمق جریان، شیب خط انرژی و قطر میانه اندازه ذرات به عنوان ورودی مدل و غلظت رسوب کل به عنوان خروجی مدل در نظر گرفته شدند. پس از مدلسازی به منظور ارزیابی اعتبار آن در برآورد بار رسوب کل، از تحلیل خطاهای باقیمانده و اختلاف بین مقادیر اندازه‌گیری و پیش‌بینی شده استفاده شد. آماره‌های لازم برای این کار، میانگین ریشه دوم

[۱۵۱] Artificial Neural Network
[۱۵۲] Genetic Programming

خطا (RMSE)، کارآیی مدلسازی (EF)، و ضریب باقیمانده (CRM) بودند. تعریف ریاضی این آماره‌ها به صورت زیر است (فلامکی و اسکندری، ۱۳۹۱):

$$RMSE=\left[\frac{\sum_{i=1}^{n}\left(P_i-O_i\right)^2}{n}\right]^{0.5}\frac{100}{\bar{O}} \quad (۱)$$

$$EF=\frac{\sum_{i=1}^{n}(O_i-\bar{O})^2-\sum_{i=1}^{n}\left(P_i-O_i\right)^2}{\sum_{i=1}^{n}(O_i-\bar{O})^2} \quad (۲)$$

$$CRM=\frac{\sum_{i=1}^{n}O_i-\sum_{i=1}^{n}P_i}{\sum_{i=1}^{n}O_i} \quad (۳)$$

$$CD=\frac{\sum_{i=1}^{n}(O_i-\bar{O})^2}{\sum_{i=1}^{n}(p_i-\bar{O})^2} \quad (۴)$$

که در آن‌ها، $Pi$ مقدار پیش‌بینی شده، $Oi$ مقدار اندازه‌گیری شده (مشاهده شده)، $n$ تعداد نمونه‌های به کار رفته و $\bar{O}$ مقدار متوسط مقادیر $O$ است. حداکثر مقدار EF برابر یک است. حداقل مقدار RMSE، ME و CD صفر است.

**برنامه نویسی ژنتیک**

برنامه نویسی ژنتیک توسط کرامر تعریف شد و سپس توسط کزا گسترش یافت. این روش زیر مجموعه‌ای از الگوریتم ژنتیک بوده و از عملگرهای آن از جمله بازترکیبی و جهش استفاده می‌کند. در این روش، درخت‌های عبارت تولید می‌شوند که این درخت‌ها به عنوان مدل یا یک برنامه کامپیوتری استفاده می‌شوند. هر یک از برنامه‌های کامپیوتری یا مدلها، یک عضو از جمعیت می‌باشند. از قابلیت‌های برنامه‌نویسی ژنتیک این است که می‌تواند معادلات ریاضی را مدل کرده و در این راستا، خطوط و منحنی‌های مختلف را شبیه‌سازی کند. تفاوت GP و الگوریتم ژنتیک در نمایش افراد جمعیت است. در برنامه نویسی ژنتیک یک فرد همزمان با ژنوتیپ و فنوتیپ ارائه می‌شود. افراد جمعیت، درختهای کامپیوتری یا گرافهای ساخت‌یافته سلسله مراتبی با اندازه‌ها و شکل های متفاوت هستند. برنامه نویسی ژنتیک روشی است که یک جستجوی غیر مرتبط انجام می‌دهد و هیچ اندازه شیب و پیچیدگی ساختاری را برای راه حل، پیش‌بینی نمی‌کند. به همان اندازه که کاربر از تابع هدف اطلاعات دارد و گروه داده‌های مرتبطی که وارد می شوند، خروجی متناظری دریافت خواهد شد. برنامه نویسی ژنتیک شامل مراحل زیر می‌باشد:

با تولید تصادفی جمعیت، راه حلهای با اندازه N شروع می‌شود.
هر راه حل از جمعیت، ارزیابی شده و مقدار برازندگی آن بدست می‌آید.
یک عملگر ژنتیک به صورت احتمالی انتخاب می‌شود.
مرحله ۳ تا رسیدن جمعیت جدید به N عضو ادامه می‌یابد.

GP از شبکه‌های عصبی مصنوعی توانمندتر است و امکان گرفتارشدن در مینیمم محلی برای آن وجود ندارد. لیکن مانند ANN، برای ساخت مدل باید از ابتدا داده‌ها را به دو دسته آموزش و آزمون قسمت کرد. دسته داده آموزش برای ساخت مدل و داده‌های آزمون برای ارزیابی آن استفاده خواهند شد. در این پژوهش ۷۵% داده‌ها برای آموزش و ۲۵% آنها برای آزمون انتخاب شدند. در GP ساخت مدل به دلیل نیاز به جمعیت زیاد، زمان‌بر است، لیکن مدل تنها یک بار ساخته می‌شود و سپس از آن استفاده خواهد شد، بنابراین زمان‌بر بودن اولیه حین مدلسازی، در کل مشکلی ایجاد نخواهد کرد (سدیدپور و همکاران، ۱۳۸۹).

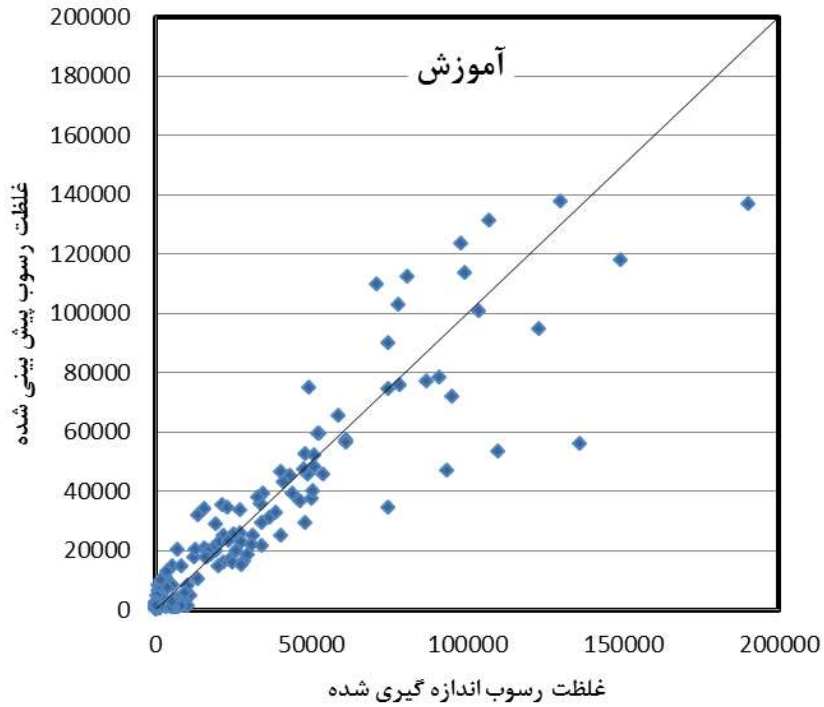
**نتایج و بحث**

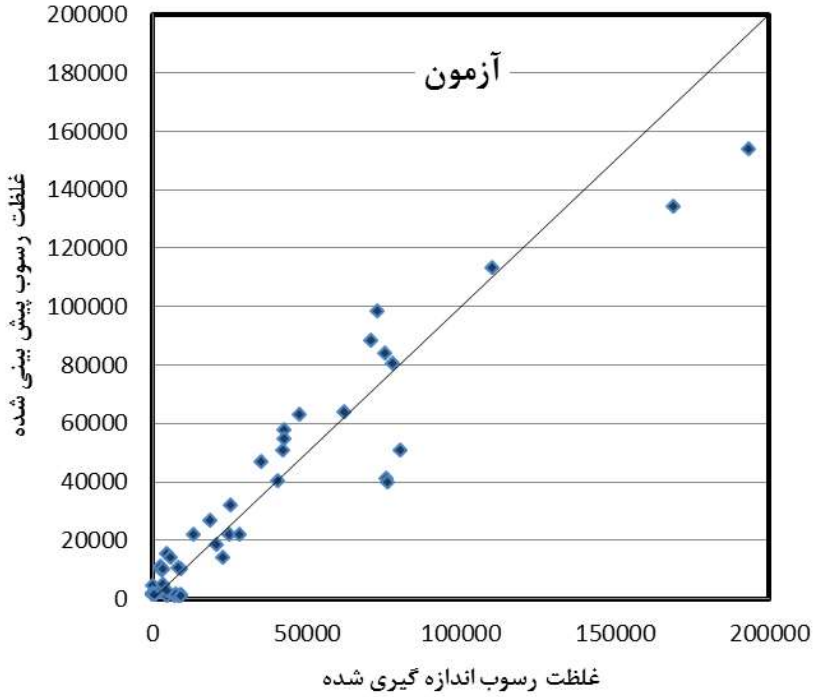
پس از کاربرد داده‌های آموزش، مدل نهایی بدست آمده برای تخمین مقدار غلظت رسوب کل رودخانه، به صورت رابطه ۵ بدست آمد.

$$Ct = \frac{6.45 \times V \times W \times S \times (2-D)}{(d_{50})^{1.5}} \quad (۵)$$

در این رابطه، V متوسط سرعت (ft/s)، W عرض کانال (ft)، S شیب خط انرژی (ft/ft)، D عمق جریان (ft)، $d_{50}$ قطر میانه اندازه ذرات (ft) و Ct غلظت رسوب کل بر حسب ppm است.

شکل ۱، مقدار غلظت رسوب پیش‌بینی شده را که به کمک رابطه ۵ برآورد شده، در مقابل مقدار واقعی آن در دو دسته داده‌های آموزش و آزمون، نشان می‌دهد. در این نمودارها برای هر مقدار اندازه‌گیری شده در محور افقی، متناظر مقدار محاسبه شده آن در محور عمودی نشان داده شده است. همچنین خط تناظر یک به یک نیز نمایان است. هر چه پراکندگی داده‌ها در اطراف این خط بیشتر باشد، توانایی مدل کاربردی در برآورد رسوب کل ضعیف‌تر بوده است. آماره‌های سنجش مدل نیز در جدول ۱ ارائه شده‌اند.






**شکل ۱- مقایسه مقادیر اندازه‌گیری شده مقدار بار کل رسوب با مقادیر پیش‌بینی شده توسط مدل GP در داده‌های آموزش و آزمون**

**جدول ۱- ضرایب معادله‌ی فروندلیچ برای ایزوترم‌های جذب سطحی نیکل**

| | CD | RMSE | EF | CRM |
|---|---|---|---|---|
| آموزش | ۸۳/۰ | ۹۳/۵۰ | ۸۲/۰ | ۰۵/۰ |
| آزمون | ۹۰/۰ | ۶۹/۴۲ | ۹۰/۰ | ۰۳/۰ |

جدول ۱ نشان می‌دهد که مقدار ضریب تبیین در داده‌های آزمون برابر با ۸۳/۰ و در داده‌های آزمون برابر با ۹/۰ بدست آمده است. آماره کارایی مدلسازی در مدل پیشنهادی، برابر با ۹۰/۰ است. هر چقدر این آماره به یک نزدیک‌تر باشد، نشان می‌دهد که مدل طراحی شده، کارایی بهتری دارد. آماره CRM نشانگر بیش‌برآوردی یا کم‌برآوردی مدل در مقایسه با اندازه‌گیری‌هاست و اگر مقدار آن منفی باشد نشان می‌دهد که مدل بیش‌برآورد است. با توجه به مقدار اندک CRM یعنی ۰۳/۰، مشخص است که مقادیر پیش‌بینی شده به مقادیر اندازه‌گیری شده بسیار نزدیک است. بنابراین نتایج این پژوهش نشان می‌دهد که مدل GP می‌تواند به عنوان ابزاری سودمند در برآورد مقدار بار رسوب کل رودخانه‌ها به کار رود.

رابطه شماره ۵، که رابطه بین خروجی با ورودی‌ها را نشان می‌دهد، افزون‌بر اینکه برای داده‌های آموزش و آزمون دارای ضریب تبیین بالایی است؛ از منظر نسبت خروجی با تک تک ورودی‌ها نیز به بهترین شکل ممکن عمل کرده است. همه ورودی‌های مورد نظر به جز $d_{50}$ با خروجی نسبت مستقیم دارند که این مسأله بخوبی در فرمول نهایی بدست آمده است. همچنین، رابطه بدست آمده در مقایسه با بسیاری از روابط ارائه شده پیشین، کوتاه‌تر بوده و بنابراین کاربرد آن نیز بسیار آسان[۱۵۳] است.

## منابع

سدیدپور، س.س. همایونپور، م.م. و فسنقری، م. ۱۳۸۹. بهبود سرعت و دقت در استفاده از برنامه‌نویسی ژنتیک برای تصدیق هویت گوینده. نشریه مهندسی برق و مهندسی کامپیوتر، سال هشتم، شماره چهارم، صفحه‌های ۲۲۳ تا ۲۳۵.

[۱۵۳] User-Friendly

فلامکی، ا.، اسکندری، م. ۱۳۹۱. تخمین ضریب توزیع خاک-آب فلزات سنگین با کاربرد شبکه‌های عصبی مصنوعی. حفاظت منابع آب و خاک، سال دوم، شماره ۱، صفحه‌های ۲۵ تا ۳۶.
فلامکی، ا.، اسکندری، م.، بغلانی، ع.ا. و احمدی، س.ا. ۱۳۹۲. مدل‌سازی بار رسوب کل رودخانه‌ها با استفاده از شبکه‌های عصبی مصنوعی. حفاظت منابع آب و خاک، سال دوم، شماره ۳، صفحه‌های ۱۳ تا ۲۵.
معیری، م.م.، نیک‌پور، م.ر.، حسین‌زاده دلیر، ع. و فرسادی‌زاده، د. ۱۳۸۹. مقایسه روش‌های شبکه‌های عصبی مصنوعی، فازی-عصبی تطبیقی و منحنی سنجه رسوب در برآورد رسوبات معلق رودخانه‌ها (مطالعه موردی: رودخانه آجی‌چای). مجله دانش آب و خاک، ویژه‌نامه ۱/۲۰. صفحه‌های ۷۱ تا ۸۲.
Aytek A. and Kisi, O. ۲۰۰۸. A genetic programming approach to suspended sediment modelling. J. Hydrol, ۳۵۱ (۳-۴): ۲۸۸-۲۹۸.
Dogan, E., Yuksel, I. and Kisi, O. ۲۰۰۷. Estimation of total sediment load concentration obtained by experimental study using artificial neural networks. Environmental Fluid Mechanics, ۷: ۲۷۱-۲۸۸.
Kisi O. and Shiri J. ۲۰۱۲. River suspended sediment estimation by climatic variables implication: Comparative study among soft computing techniques. Comp. Geosci, ۴۳: ۷۳-۸۲.
Molinas, A. and Wu, B. ۲۰۰۳. SedBase (sediment transport database). Hydrau-Tech, Colorado, USA.
Singh K., Deo, M.C. and Sanil Kumar, V. ۲۰۰۷. Neural network-genetic programming for sediment transport. P. I. Civil Eng.- Wat. M, ۱۶۰ (۳): ۱۱۳-۱۱۹.
Yang CT, Marsooli R and Aalami MT. ۲۰۰۹. Evaluation of total load sediment transport formulas using ANN. International Journal of Sediment Research, ۲۴: ۲۷۴-۲۸۶.

**Abstract**

Modeling sediment transport are usually complex problems due to their multi-dimensionality as well as their nonlinear behavior. Although some investigators have been used regression analyses for this purpose, but still many limitations are exist for the fitted functions. In this research, a genetic programming was used to predict the total sediment load. The data used in this study were consisted of ۲۰۰ samples which were divided into training (۷۵%) and testing (۲۵%) data. The input variables were consisted of average flow velocity, water surface slope, average flow depth, mean particle diameter and channel width. The total bed material load was designated as the output of model. Results showed that the proposed model can accurately estimate the total sediment yield with r۲=۰.۹ and RMSE=۴۲.۷.