

بررسی نقش متغیر بارندگی در عملکرد مدل‌سازی بار رسوب معلق روزانه (مطالعه موردی: حوزه آبخیز سعید آباد چای)

عاده علی جانپور شلمانی^{۱*}، علی‌رضا واعظی^۲، محمودرضا طباطبایی^۳

تاریخ دریافت: ۱۳۹۸/۳/۲۰ تاریخ پذیرش: ۱۳۹۸/۴/۳۱

چکیده

آگاهی از مقدار دقیق بار رسوب معلق روزانه می‌تواند در شناخت وضعیت فرسایش و رسوب حوزه‌های آبخیز مورد استفاده قرار گیرد. در این تحقیق از مدل‌های هوشمند شبکه عصبی مصنوعی و برنامه‌ریزی بیان ژن برای برآورد بار رسوب معلق روزانه استفاده شد. همچنین با توجه به اهمیت پاسخ حوزه به متغیرهای ورودی به مدل‌ها، علاوه بر متغیر دبی، متغیر دینامیک بارندگی به دلیل نقش تأثیرگذار در ایجاد فرسایش و تولید رسوب نیز برای ورود به مدل‌ها انتخاب شد. نتایج حاصل از این تحقیق نشان داد که تمام مدل‌هایی که از متغیر بارندگی به همراه دبی استفاده کردند، دارای مقدار آماره‌های NSE و R² بیشتر و مقدار آماره‌های RMSE و MAE کمتر در مقایسه با مدل‌هایی بودند که تنها از متغیر دبی برای برآورد بار رسوب معلق استفاده کردند. همچنین مدل GEP با ترکیب متغیر ورودی دبی لحظه‌ای، دبی متوسط روزانه، دبی متوسط روزانه تا سه روز قبل، بارندگی متوسط روزانه و بارندگی متوسط روزانه تا سه روز قبل، کارآمدترین مدل در برآورد صحیح بار رسوب معلق روزانه با بیشترین مقدار آماره‌های NSE برابر ۰/۹۰ و R² برابر ۰/۹۲ و کمترین مقدار آماره‌های RMSE برابر ۲۲۸۲/۴۲ (ton/day) و MAE برابر ۷۵۰/۳۸ (ton/day) در مقایسه با مدل‌های شبکه عصبی مصنوعی بود. به طور کلی نتایج این تحقیق نشان داد که متغیر دبی به تنهایی نتوانست واریانس رسوب رودخانه را به درستی تبیین نماید و استفاده از متغیر بارندگی به عنوان متغیر ورودی به مدل‌های هوشمند، نقش تأثیرگذار در افزایش دقت برآورد بار رسوب معلق داشت و استفاده از متغیر بارندگی به همراه متغیر دبی در طی فرآیند مدل‌سازی، کارایی مدل‌ها را افزایش داد.

واژه‌های کلیدی: بارندگی، برنامه‌ریزی بیان ژن، شبکه عصبی مصنوعی، نگاهت خود سازمان‌ده

مقدمه

هزینه زیاد و همچنین مشکلات کیفی داده‌ها در ارتباط با عدم دریافت داده رسوب معلق در شرایط سیلابی که بیان‌گر بیشترین مقدار رسوب حمل شده توسط رودخانه می‌باشد، امکان استفاده از روش مستقیم وجود نداشته و بنابراین برآورد بار رسوب معلق روزانه به صورت غیرمستقیم انجام می‌شود (Güven and Kisi, 2011). بدین منظور در این تحقیق، از روش منحنی سنج رسوب (SRC^۴) به عنوان روش سنتی و غیرمستقیم برای برآورد بار رسوب معلق روزانه استفاده شد (Uike et al., 2009). استفاده از روش سنتی SRC که مبتنی بر روش‌های رگرسیونی است، به دلیل سادگی در نحوه استفاده، بسیار مرسوم است. اما در این روش تنها از متغیر دبی برای برآورد SSL استفاده می‌شود. در صورتی که رودریگز و همکاران بیان کردند که متغیر دبی تنها قادر است تا ۱۹ درصد واریانس بار رسوب معلق را بیان کند، بنابراین روش SRC نمی‌تواند برآورد دقیق و مطمئنی از مقدار SSL را داشته باشد (Rodríguez-Blanco et al., 2010). از این رو، برای برآورد دقیق بار رسوب معلق روزانه لازم است تا از

آگاهی از وضعیت رسوب‌دهی حوزه‌های آبخیز، میزان فرسایش، مشکلات مربوط به کیفیت آب‌های سطحی مانند انتقال عناصر و فلزات سنگین، رس و مواد آلی، نیازمند داشتن اطلاعات کافی از مقدار بار رسوب معلق (SSL) رودخانه است (Azamathulla et al., 2013, Kisi and Fedakar, 2014). در واقع بار رسوب معلق شاخصی از وضعیت فرسایش خاک و شرایط اکولوژیک حوضه می‌باشد (Ghavidel and Montaseri, 2014). به دلیل مشکلات کمی در اندازه‌گیری مستقیم بار رسوب معلق مربوط به نبود داده کافی در اثر خرابی دستگاه‌های سنجش رسوب، کمبود نیروی متخصص و

۱- دانشجوی دکتری، گروه علوم خاک، دانشگاه زنجان، زنجان، ایران

۲- استاد، گروه علوم خاک، دانشگاه زنجان، زنجان، ایران

۳- استادیار پژوهشی، پژوهشکده حفاظت خاک و آبخیزداری، سازمان تحقیقات، آموزش و ترویج کشاورزی، تهران، ایران

*- نویسنده مسئول: (Email: adele.alijanpour@gmail.com)

مقایسه و ارزیابی نتایج مدل‌های شبکه عصبی مصنوعی، برنامه‌ریزی بیان ژن و منحنی سنجه رسوب در برآورد بار رسوب معلق روزانه.

مواد و روش‌ها

منطقه مورد مطالعه

حوزه آبخیز سعیدآباد چای، یکی از زیرحوضه‌های حوزه آبخیز آجی‌چای است که بین عرض‌های جغرافیایی شرقی ۲۸° ۴۶ تا ۳۰° ۴۶ و عرض‌های شمالی ۴۷° ۳۷ تا ۱° ۳۸، در استان آذربایجان شرقی و در شمال غرب ایران واقع شده است (شکل ۱). مشخصات فیزیوگرافی این حوضه در جدول ۱ نشان داده شده است. اقلیم غالب منطقه براساس روش دومارتن، نیمه خشک است.

داده‌های مورد استفاده

در این تحقیق، ایستگاه هیدرومتری ونبار، با طول دوره آماری ۳۲ سال (۱۳۹۱-۱۳۶۰) مورد استفاده قرار گرفت. داده‌های مورد استفاده در این تحقیق، از سازمان تماب اخذ شد. این داده‌ها شامل متغیرهای ورودی به مدل‌ها شامل بارندگی متوسط روزانه (P_i)، بارندگی متوسط روزانه یک روز قبل (P_{i-1})، بارندگی متوسط روزانه دو روز قبل (P_{i-2})، بارندگی متوسط روزانه سه روز قبل (P_{i-3})، دبی لحظه‌ای (Q)، دبی متوسط روزانه (Q_i)، دبی متوسط روزانه یک روز قبل (Q_{i-1})، دبی متوسط روزانه دو روز قبل (Q_{i-2}) و دبی متوسط روزانه سه روز قبل (Q_{i-3}) و متغیر خروجی بار رسوب معلق روزانه (SSL) بود. جدول ۲، مشخصات آماری داده‌های مورد استفاده در حوزه آبخیز سعیدآباد چای را نشان می‌دهد. ایستگاه هیدرومتری ونبار از نوع درجه‌ی یک و مجهز به پل تله‌فریک، اشل و لیمنوگراف است.

خوشه‌بندی با روش نگاشت خود سازمان ده

در این تحقیق، به منظور بهبود دقت برآورد بار رسوب معلق روزانه، از روش نگاشت خود سازمان ده که از دقیقترین روش‌ها در شناخت و تحلیل فضاهاى چندبعدی غیرخطی است، استفاده شده است. نگاشت خود سازمان ده در حالت کلی دارای یک ساختار دو لایه با یک لایه ورودی و یک لایه خروجی است. نرون‌های لایه ورودی وظیفه انتقال داده‌ها به شبکه را برعهده داشته و در حالت کلی تعداد آن‌ها با بعد بردارها در فضای ورودی برابر است که به وسیله وزن‌های اتصالی به تمام نرون‌های لایه خروجی متصل می‌باشند. لایه خروجی نیز شامل مجموعه‌ای از نرون‌ها است که معمولاً در یک صفحه مسطح کنار یکدیگر قرار گرفته‌اند. این نرون‌های لایه خروجی با توجه به روابط همسایگی مشخصی که بین آن‌ها تعریف شده و رفتار متقابل روی هم‌دیگر، خروجی شبکه را ایجاد می‌کنند (Wang et al., 2010).

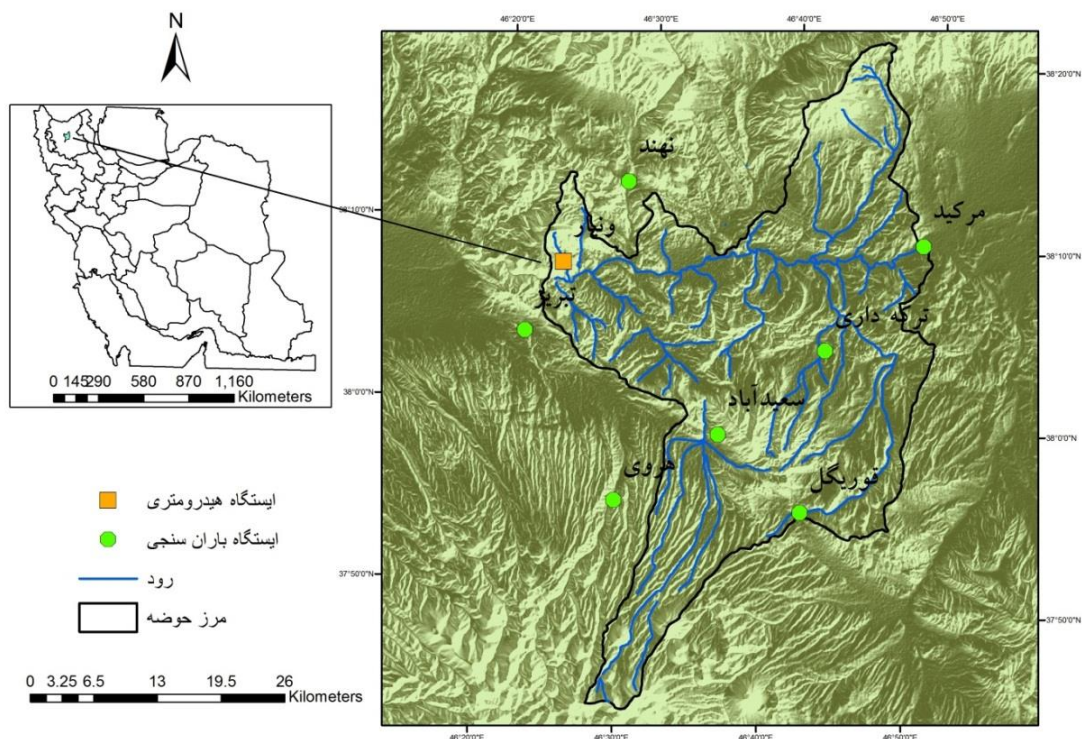
مدل‌های مبتنی بر روش‌های هوش محاسباتی استفاده شود تا بتوان علاوه بر متغیر دبی از متغیر یا متغیرهای دیگر نیز در طی فرآیند مدل‌سازی استفاده کرد. استفاده از این روش‌ها سبب می‌شود تا با تعداد محدودی از متغیرهای دینامیک حوضه مانند متغیر بارندگی، به همراه متغیر دبی بتوان رفتار حوزه آبخیز در تولید رسوب را با دقت بالا برآورد کرد (Kisi and Ozkan, 2017). بنابراین نیازی به اندازه‌گیری و تعیین وزن دقیق فاکتورهای اقلیمی، زمین‌شناسی، توپوگرافی و ... که باعث پیچیدگی در فرآیند مدل‌سازی می‌شود، نیست (Ouillon, 2018). در این تحقیق از روش‌های هوش محاسباتی مانند شبکه عصبی مصنوعی (ANN) و برنامه‌ریزی بیان ژن (GEP) استفاده شد و با استفاده از این مدل‌ها نقش متغیر بارندگی در دقت برآورد SSL مورد بررسی قرار گرفت. در دهه‌ی اخیر در اکثر مطالعات انجام گرفته برای برآورد بار رسوب معلق روزانه، تنها از متغیر دبی به عنوان متغیر ورودی به مدل سنتی منحنی سنجه رسوب، مدل‌های هوشمند و سایر مدل‌ها استفاده شده است و نقش متغیر بارندگی به عنوان متغیر تأثیرگذار در برآورد بار رسوب معلق در نظر گرفته نشده است (Chiang et al., Melesse et al., 2011, Emamgholizadeh and Karimi Demneh, 2014, 2014). این پژوهشگران به منظور برآورد بار رسوب معلق از مدل منحنی سنجه رسوب، مدل‌های هوشمند و مدل رگرسیون استفاده کردند و از داده‌های دبی جریان و رسوب معلق روزانه جهت ورود به مدل‌ها استفاده کردند. نتایج پژوهش‌ها، عملکرد بهتر مدل‌های هوشمند نسبت به سایر مدل‌ها را نشان داد. همچنین یکی از مهم‌ترین مسائل در استفاده از مدل‌هایی مانند مدل‌های رگرسیون، شبکه‌های عصبی مصنوعی، برنامه‌ریزی بیان ژن و ... در مدل‌سازی متغیرهای محیطی، تهیه مجموعه داده‌های همگن و مشابه در سه گروه داده‌های آموزش (Training)، اعتبارسنجی (Cross-Validation) و آزمون (Validation) است. عدم استفاده از داده‌های همگن در این سه گروه باعث کاهش دقت، کارایی و قدرت تعمیم دهی مدل‌ها می‌شود (May et al., 2010). در سوابق تحقیقات انجام شده در مدل‌سازی برآورد SSL با استفاده از روش‌های محاسبات نرم، به ندرت به این مسئله توجه شده است (Bowden et al., 2002, Tayfur, 2012). با توجه به مطالب بیان شده، اهداف این تحقیق را می‌توان به صورت زیر خلاصه کرد:

بررسی نقش استفاده از متغیر دینامیک بارندگی در مقدار کارایی مدل‌های هوشمند،

استفاده از خوشه‌بندی داده‌ها با استفاده از روش نگاشت خود سازمان ده برای افزایش قدرت تعمیم‌دهی مدل‌ها،

1- Artificial Neural Network

2- Gene Expression Programming



شکل ۱- نقشه حوزه آبخیز سعیدآباد چای

جدول ۱- مشخصات فیزیوگرافی حوزه آبخیز سعیدآباد چای

مقدار	مشخصات فیزیوگرافی	مقدار	مشخصات فیزیوگرافی
۱/۴۱	ضریب گراویلیوس	۲۳۸۰	متوسط ارتفاع از سطح دریا (m)
۰/۹۴	ضریب شکل	۲۲۳۰	مساحت (km ²)
۰/۳۷	تراکم آبراهه	۲۳۶	محیط (km)
۴۵/۲۴	طول آبراهه اصلی (km)	۳۴/۵۲	شیب متوسط (%)

جدول ۲- مشخصات آماری داده‌های مورد استفاده در حوزه‌ی آبخیز سعیدآباد چای

ضریب تغییرات	میانگین	بیشینه	کمینه	داده
۱/۴۲	۱۷/۳۴	۱۴۷	۰/۰۱	دبی لحظه‌ای Q (m ³ /s)
۱/۲۴	۲۰/۸۴	۱۴۴	۰/۰۱	دبی متوسط روزانه Q _i (m ³ /s)
۱/۲۶	۲۰/۵۲	۱۳۸	۰/۰۱	دبی متوسط روزانه یک روز قبل Q _{i-1} (m ³ /s)
۱/۲۵	۱۹/۷۷	۱۴۱	۰/۰۱	دبی متوسط روزانه دو روز قبل Q _{i-2} (m ³ /s)
۱/۳۱	۲۰/۰۲	۱۵۰	۰/۰۱	دبی متوسط روزانه سه روز قبل Q _{i-3} (m ³ /s)
۲/۵۴	۱/۲۸	۲۷/۸۲	۰/۰۰	بارندگی متوسط روزانه P _i (mm)
۲/۵۴	۱/۵۱	۳۱/۵۳	۰/۰۰	بارندگی متوسط روزانه یک روز قبل P _{i-1} (mm)
۲/۶۱	۱/۱۵	۲۷/۸۲	۰/۰۰	بارندگی متوسط روزانه دو روز قبل P _{i-2} (mm)
۲/۵۸	۱/۴۴	۳۱/۵۳	۰/۰۰	بارندگی متوسط روزانه سه روز قبل P _{i-3} (mm)
۲/۰۱	۴۷۹۰/۹۵	۵۴۷۷۲/۰۷	۰/۰۷	بار رسوب معلق روزانه SSL (ton/day)

ورودی تطبیق داده به نحوی که در انتهای آموزش، داده‌های ورودی مشابه، همگی به یک نورون خاص در لایه خروجی مربوط می‌شوند و

در عمل خوشه‌بندی و در حین فرآیند آموزش شبکه، با معرفی داده‌های ورودی به شبکه، نورون‌های لایه خروجی، خود را با داده‌های

پیش‌رو^۱ استفاده شد. معماری معمول این شبکه متشکل از سه لایه، شامل لایه ورودی، لایه پنهان و لایه خروجی است. تعداد نورون‌های موجود در لایه ورودی و خروجی، به نوع مسئله بستگی دارد، در صورتی که تعداد نورون‌های لایه پنهان با سعی و خطا به دست می‌آید (ASCE, 2000). شبکه‌های عصبی پیش‌رو که فرآیند یادگیری در آن‌ها به وسیله قاعده کلی دلتا^۲ کنترل می‌شود را شبکه‌های عصبی پس‌انتشار می‌نامند. این الگوریتم، یک تکنیک کاهش شیب است که تابع خطای شبکه (رابطه ۳) را حداقل می‌کند، بدین صورت که هر الگوی ورودی از مجموعه داده‌های آموزش، از لایه ورودی به سمت لایه خروجی عبور داده شده و خروجی شبکه با خروجی هدف مقایسه و مقدار تابع خطا به دست می‌آید. این خطا به صورت برگشتی به داخل شبکه منتشر شده و وزن‌های اتصالی بر اساس رابطه ۴ تنظیم می‌شود (ASCE, 2000).

$$E = \sum_p \sum_q (y_i - t_i)^2 \quad (3)$$

که در آن؛ t_i یک مولفه از بردار هدف T ، y_i خروجی معادل از شبکه عصبی، q ، تعداد گره‌های خروجی و P تعداد الگوهای آموزشی است.

$$\Delta w_{ij}(n) = -\epsilon \times \frac{\partial E}{\partial w_{ij}} + \alpha \times \Delta w_{ij}(n-1) \quad (4)$$

که در آن؛ $\Delta w_{ij}(n)$ و $\Delta w_{ij}(n-1)$ ، افزایش وزن بین گره i و j در طول تکرار n ام و $n-1$ ام، ϵ و α ، به ترتیب نرخ یادگیری^۳ و شتاب (مومتوم) است. در این تحقیق، در طی فرآیند آموزش شبکه عصبی، برای کاهش میزان خطا، از الگوریتم مارکوارت-لونبرگ استفاده شد. همچنین، با سعی و خطا، تعداد یک لایه پنهان در شبکه‌های عصبی استفاده شد. توابع فعال‌سازی برای نورون‌های لایه پنهان لوگ سیگنویید و تانژانت سیگنویید و برای لایه خروجی، خطی در نظر گرفته شد. در این تحقیق برای طراحی مدل شبکه عصبی مصنوعی پروسپترون پیش‌رو از نرم‌افزار MATLAB R2013a استفاده شد.

مدل برنامه‌ریزی بیان ژن

برنامه ریزی بیان ژن (GEP) از مجموعه روش‌های فرا ابتکاری است که در سال ۱۹۹۹ توسط فریرا ابداع شد (۱۰). این روش، توسعه یافته از روش‌های الگوریتم ژنتیک (GA)^۴ و برنامه‌ریزی ژنتیک (GP)^۵ است (Ferreira, 2006). تفاوت اساسی سه روش GA، GP و GEP به ساختار افراد بستگی دارد. روش الگوریتم ژنتیک شامل کروموزم‌های خطی با طول ثابت و بدون بیان پیچیده است که در آن،

بدین صورت، داده‌ها از فضایی با ابعاد بالاتر به فضایی با ابعاد کمتر نگاشت شده، خوشه‌بندی می‌شوند (Tabatabaei et al., 2014). در طی فرآیند آموزش و با تغییر وزن آن‌ها، موقعیت آن‌ها نیز در شبکه اولیه‌شان تغییر نموده به نحوی که به یک یا مجموعه‌ای از داده‌های ورودی که کمترین فاصله اقلیدسی (معادله ۱) را با آن‌ها داشته باشند نزدیک و تشکیل یک خوشه را می‌دهند.

$$D_j = |x - w_j| = \sum_{i=1}^N [(x_i - w_{ji})^2]^{\frac{1}{2}} \quad (1)$$

که در آن؛ D_j فاصله نورون j ام لایه خروجی از بردار ورودی $x = (x_i; i=1,2,3,..,N) \in R^n$ ، N تعداد متغیرهای بردار ورودی، M تعداد نورون‌های لایه خروجی، W_{ij} وزن نورون خروجی $(W_{ji}; i=1,2,..,n; j=1,2,..,M)$ ، نشان‌دهنده فاصله است (Tabatabaei et al., 2014). در این تحقیق برای ارزیابی کیفیت خوشه‌بندی از شاخص دیویس بولدین استفاده شد. این شاخص، میانگین شباهت بین هر خوشه با شبیه‌ترین خوشه به آن را محاسبه کرده و هر چه مقدار محاسبه شده این شاخص کمتر باشد خوشه‌بندی با کیفیت بهتری انجام می‌شود. شاخص دیویس بولدین از طریق رابطه ۲ به دست می‌آید (Theodosiou, 2011):

$$DB = \frac{1}{n_c} \sum_{i=1}^{n_c} R_i \quad (2)$$

که در آن؛ DB شاخص دیویس بولدین، n_c تعداد خوشه‌ها و R_i بالاترین شباهت بین خوشه‌های است. پس از تعیین تعداد بهینه خوشه، نمونه‌گیری از خوشه‌ها با استفاده از روش تخصیص متناسب انجام شد. در این روش، تعداد نمونه‌ها متناسب با اندازه خوشه تغییر می‌کند، به طوری که با افزایش اندازه هر خوشه، تعداد نمونه‌گیری از آن‌ها افزایش می‌یابد (Theodosiou, 2011). در نهایت، در مدل منحنی سنجه‌ی رسوب، ۸۰ درصد داده‌ها برای آموزش و ۲۰ درصد برای آزمون و در مدل‌های هوش مصنوعی ۷۰ درصد داده‌ها برای آموزش، ۱۵ درصد برای اعتبار سنجی و ۱۵ درصد کل داده‌ها برای آزمون در نظر گرفته شد. در این تحقیق برای طراحی این روش از نرم‌افزار MATLAB R2013a استفاده شد.

مدل شبکه عصبی مصنوعی

اساس کار شبکه عصبی مصنوعی، استفاده از دانش نهفته در داده‌ها، تلاش برای استخراج روابط ذاتی بین آن‌ها و تعمیم آن در موقعیت‌های دیگر است (Nivesh and Kumar, 2017). در استفاده از شبکه عصبی مصنوعی سعی می‌شود با آموزش شبکه‌ای پردازشگر، بدون یافتن قانون ریاضی حاکم بر پارامترها، ویژگی‌ها و روابط ذاتی میان متغیرها را به شبکه آموخت و پس از تضمین یادگیری، از آن در تعمیم‌دهی و ایجاد رابطه غیرخطی بین دو فضای ورودی و خروجی بهره برد. در این تحقیق، از شبکه عصبی مصنوعی پروسپترون

- 1- Feed-forward Multi-layer Perceptron
- 2- Delta Rule
- 3- Learning Rate
- 4- Genetic Algorithms
- 5- Genetic Programming

به ترتیب بین $[0/9 \ 0/9]$ و $[0/1 \ 0/9]$ انجام شد:

$$Z = 0.1 + 0.8 * \frac{(X_i - X_{imin})}{(X_{imax} - X_{imin})} \quad (۶)$$

$$Z = \left(1.8 * \frac{(X_i - X_{imin})}{(X_{imax} - X_{imin})} \right) - 0.9 \quad (۷)$$

که در آن‌ها: Z متغیر استاندارد شده، X_i متغیر اولیه، X_{imin} مقدار کمینه و X_{imax} مقدار بیشینه X_i است.

معیارهای ارزیابی مدل‌ها

در این تحقیق، از معیارهای آماری شامل ضریب تبیین (R^2) ، ریشه میانگین مربعات خطا^۳ (RMSE)، میانگین قدر مطلق خطا^۴ (MAE) و معیار نش-ساتکلیف^۵ (NS) برای سنجش کارایی، دقت و خطای مدل‌های هوشمند و منحنی سنج رسوب استفاده شد.

$$R^2 = \left[\frac{\sum_{i=1}^n (s_o - \bar{s}_o)(s_p - \bar{s}_p)}{\sqrt{\sum_{i=1}^n (s_o - \bar{s}_o)^2 \sum_{i=1}^n (s_p - \bar{s}_p)^2}} \right]^2 \quad (۸)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (s_p - s_o)^2} \quad (۹)$$

$$MAE = \frac{\sum_{i=1}^n |s_o - s_p|}{n} \quad (۱۰)$$

$$NS = 1 - \frac{\sum_{i=1}^n (s_p - s_o)^2}{\sum_{i=1}^n (s_o - \bar{s}_o)^2} \quad (۱۱)$$

که در آن‌ها: s_o و s_p به ترتیب بار رسوب معلق مشاهده‌ای و برآورد شده، \bar{s}_o میانگین بار رسوب معلق مشاهده‌ای، \bar{s}_p میانگین بار رسوب معلق برآورد شده و n تعداد داده‌ها است.

نتایج و بحث

نتایج خوشه‌بندی با روش نگاشت خود سازمان ده

در این تحقیق، تعداد بهینه خوشه در داده‌های مورد مطالعه ۴۲ خوشه با استفاده از روش نگاشت خود سازمان ده و شاخص دیویس بولدین تعیین شد (شکل ۲). نتایج پارامترهای آماری در مجموعه داده‌های آموزش (Training)، اعتبارسنجی (Cross-Validation) و آزمون (Validation) با استفاده از خوشه‌بندی داده‌ها با روش تخصیص متناسب در شکل ۳ آورده شده است.

همان‌طور که در شکل ۳ مشاهده می‌شود، خوشه‌بندی داده‌ها با استفاده از روش نگاشت خود سازمان ده باعث توزیع یکنواخت داده‌ها در هر سه گروه شد. در این روش داده‌های قرار گرفته در هر گروه از مجموعه داده‌های آموزش، اعتبارسنجی و آزمون، نماینده‌ای از کلیه داده‌ها در طول دوره آماری مورد مطالعه می‌باشد. استفاده از این داده‌های همگن در طی فرآیند مدل‌سازی بار رسوب معلق روزانه، باعث افزایش قدرت تعمیم‌دهی مدل‌ها خواهد شد (Tabatabaei and Salehpour Jam, 2017).

کروموزوم‌ها طبق برتری فردی و براساس شایستگی حفظ می‌شود. روش برنامه‌ریزی ژنتیک از افرادی با رشته‌های غیرخطی و با اشکال و اندازه‌های متفاوت تشکیل شده است که به شکل درخت تجزیه (ET) بیان می‌شود. در روش برنامه‌ریزی بیان ژن، کروموزوم‌های خطی با طول ثابت مشابه با GA و ساختار درختی با اشکال و اندازه‌های متفاوت، مشابه با GP، با هم ترکیب می‌شوند (Wilson, 2008). این امر سبب می‌شود تا فنوتایپ و ژنوتایپ از هم جدا شده و در نتیجه سیستم قادر خواهد بود تا از تمام مزایای تکاملی بهره‌مند گردد. در واقع در روش GEP به‌سازی در یک ساختار خطی اتفاق می‌افتد و سپس به‌صورت ساختار درختی بیان می‌شود (Ferreira, 2006). فرآیند انجام روش GEP بدین‌صورت است که با تولید تصادفی از جمعیت اولیه آغاز می‌شود. سپس کروموزوم‌های ایجاد شده در GP و GEP به‌صورت درختی بیان می‌شود. در ادامه به کمک تابع هدف میزان برآزندگی و کارایی هر فرد ارزیابی می‌شود. پس از آن اعضاء سازگارتر، انتخاب و اصلاح شده تا فرزندان جدیدتر ایجاد شوند، فرزندان ایجاد شده نیز تحت فرآیند توسعه‌ای قرار می‌گیرند تا پاسخ خوب و مناسب پیدا شود (Lopes and Weinert, 2004). در این روش متغیرهای هدف با استفاده از مجموعه‌ای از توابع و ترمینال‌ها، مدل‌سازی می‌شوند. مجموعه توابع معمولاً شامل توابع اصلی (+, -, ×, /)، توابع مثلثاتی (x², exp, log) یا توابع تعریف شده توسط کاربر است که احتمال می‌دهد برای تفسیر مدل مناسب خواهد بود (Ferreira, 2006). مجموعه‌ی ترمینال شامل متغیرهای مستقل و ثابت‌های عددی است. در این تحقیق، به منظور مدل‌سازی با استفاده از روش GEP از نرم‌افزار GEPXpro Tools5.0 استفاده شد.

منحنی سنج رسوب

منحنی سنج رسوب، یک رابطه رگرسیونی نمایی بین دبی لحظه‌ای (Q) بر حسب (m³/s) و بار رسوب معلق (SSL) متناظر آن بر حسب (ton/day) می‌باشد (رابطه ۵). ضرایب a و b، ضرایب منحنی سنج مربوط به ایستگاه هیدرومتری بوده که به روش حداقل مربعات خطا محاسبه می‌شود (Ulke et al., 2009).

$$SSL = aQ^b \quad (۵)$$

استاندارد سازی داده‌ها

به منظور بی‌بعد کردن داده‌ها در محاسبات فاصله‌ای در عملیات خوشه‌بندی با روش نگاشت خود سازمان ده و یا به‌منظور جلوگیری از کوچک شدن بیش از حد وزن‌های تخصیص یافته به نورون‌ها در مدل‌های شبکه عصبی، استانداردسازی داده‌ها انجام می‌شود. در این تحقیق، به دلیل استفاده از توابع محرک سیگموئید و تانژانت هیپربولیک در شبکه‌های عصبی مصنوعی استانداردسازی داده‌ها

2- Coefficient of Determination

3- Root Mean Square error

4- Mean Absolute Error

5- Nash-Sutcliffe

1- Expression Tree

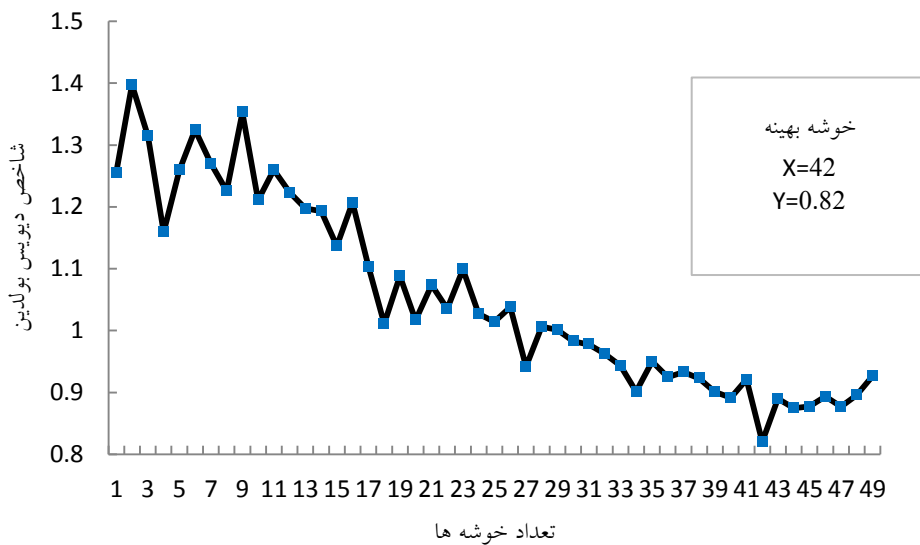
نتایج مدل‌سازی

جدول ۴ نتایج برآورد بار رسوب معلق با استفاده از مدل سنتی منحنی سنجه رسوب (SRC) و مدل‌های هوشمند شبکه عصبی مصنوعی با دو تابع فعال‌سازی لوگ سیگموئید (ANN- log sigmoid) و تانژانت سیگموئید (ANN- tangent sigmoid) و برنامه‌ریزی بیان ژن (GEP) را نشان می‌دهد. در این جدول تنها از ترکیب متغیری دبی با روش آزمون و خطا دبی لحظه‌ای (Q)، دبی متوسط روزانه (Q_i) و دبی متوسط روزانه با زمان تأخیر برای ورود به مدل‌ها استفاده شد. همچنین برای بررسی نقش متغیر بارندگی در میزان برآورد بار رسوب معلق روزانه، از ترکیب متغیرهای دبی به همراه بارندگی با روش آزمون و خطا استفاده شد که جدول ۵ نتایج حاصل از این ترکیبات در مدل‌سازی را نشان می‌دهد.

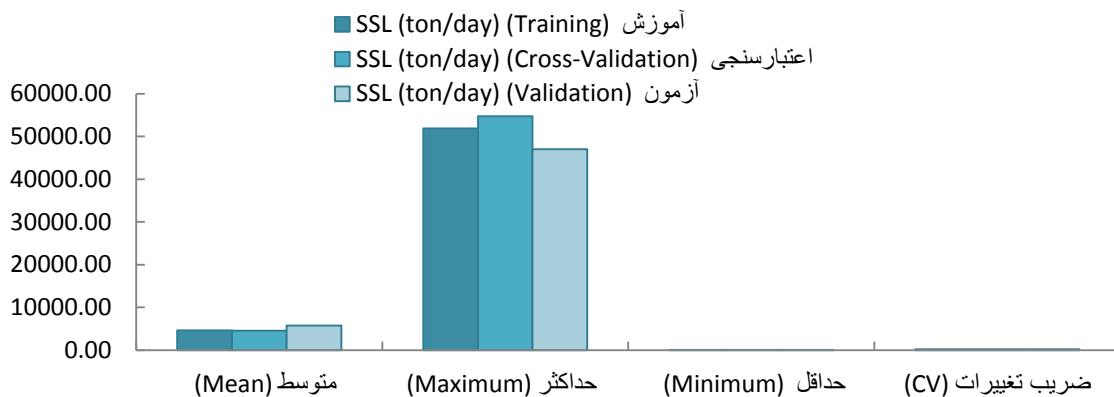
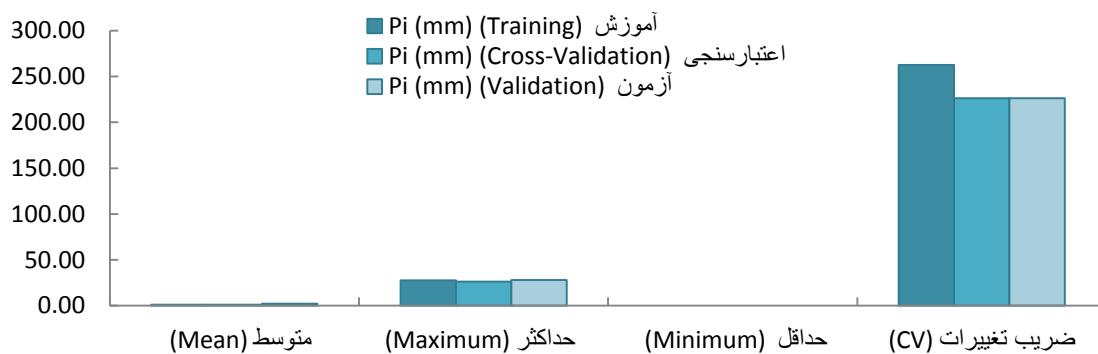
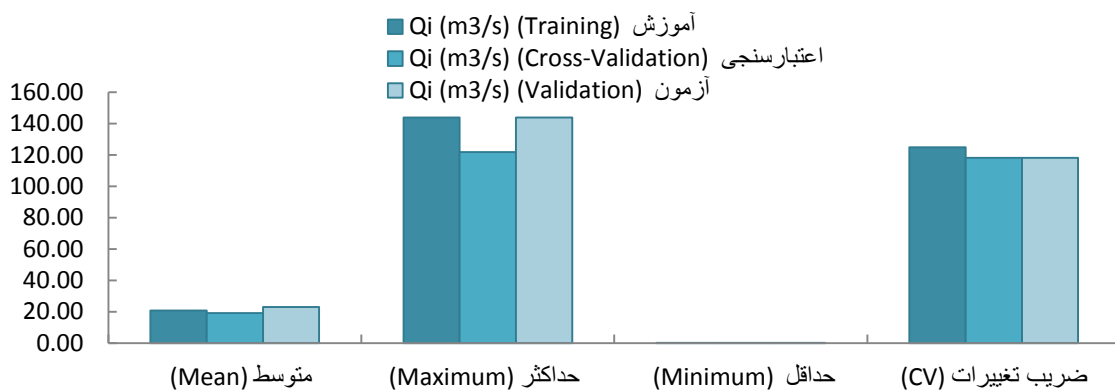
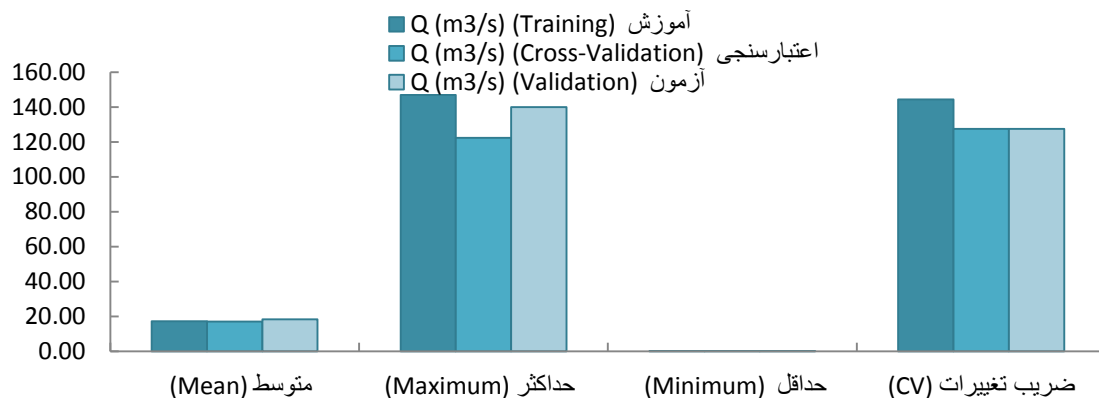
در جدول ۴ مقایسه بین مدل‌ها نشان داد که مدل منحنی سنجه رسوب با کمترین مقدار آماره‌های NSE برابر $0/33$ و R^2 برابر $0/37$ و بیشترین مقدار آماره‌های RMSE برابر $9377/64$ (ton/day) و MAE برابر $3914/24$ (ton/day)، کم‌ترین میزان دقت در برآورد بار رسوب معلق در مقایسه با تمامی مدل‌های هوشمند را داشت. در مدل‌سازی بار رسوب معلق با استفاده از روش شبکه عصبی مصنوعی با تابع فعال‌سازی لوگ سیگموئید، ترکیب متغیری دبی لحظه‌ای، دبی متوسط روزانه و دبی متوسط روزانه یک روز قبل (Q, Q_i, Q_{i-1}) در مقایسه با چهار ترکیب متغیری دیگر در این مدل، با NSE برابر $0/68$

و RMSE برابر $3514/88$ (ton/day) دقیق‌ترین برآورد SSL را داشت. در صورتی که در مدل شبکه عصبی مصنوعی با تابع فعال‌سازی تانژانت سیگموئید و مدل برنامه‌ریزی بیان ژن صحیح‌ترین و دقیق‌ترین برآورد SSL، مربوط به ترکیب متغیری دبی لحظه‌ای (Q) می‌باشد. همچنین، مدل GEP با ترکیب متغیر دبی لحظه‌ای (Q)، قدرت‌مندترین مدل در مقایسه با تمام مدل‌ها در برآورد صحیح SSL با NSE برابر $0/74$ و R^2 برابر $0/74$ و کمترین مقدار آماره‌های RMSE برابر $3412/21$ (ton/day) و MAE برابر $1102/62$ (ton/day) بود.

همان‌طور که نتایج نشان داد به‌طور کلی، کیفیت و کمیت نامناسب داده‌های بار رسوب معلق اندازه‌گیری شده در ایستگاه‌های آب سنتی باعث می‌شود تا دقت برآورد بار رسوب معلق در روش سنتی منحنی سنجه رسوب کاهش یابد. این در حالی است که استفاده از روش‌های محاسباتی مبتنی بر هوش مصنوعی باعث می‌شود تا مدل‌های توسعه یافته از آن‌ها، قدرت تعمیم‌دهی بیشتری داشته و در شرایطی که داده‌ها از شرایط مناسبی برخوردار نیستند هنوز هم کارآمد باشد. در پژوهشی یاداو و همکاران برای برآورد رسوب معلق در حوزه آبخیز رودخانه ماهانادی در هند از شبکه عصبی مصنوعی و مدل سنتی منحنی سنجه رسوب استفاده کردند.



شکل ۲- تعیین تعداد خوشه بهینه خوشه با استفاده از خوشه‌بندی SOM و شاخص دیویس بولدین



شکل ۳- نتایج آماری خوشه‌بندی داده‌ها با استفاده از روش SOM در حوزه آبخیز سعیدآباد چای

جدول ۴- نتایج مدل‌سازی SSL با ترکیب متغیرهای دبی

مدل	ترکیب متغیرهای ورودی	MAE (ton/day)	RMSE (ton/day)	NSE	R ²
ANN- log sigmoid	Q	۱۰۲۱/۰۹	۳۸۱۱/۱۳	-۰/۶۲	-۰/۶۳
	Q, Q _i	۱۳۴۹/۴۶	۴۲۶۹/۴۲	-۰/۵۵	-۰/۵۶
	Q, Q _i , Q _{i-1}	۱۱۲۱/۰۲	۳۵۱۴/۸۸	-۰/۶۸	-۰/۶۹
	Q, Q _i , Q _{i-1} , Q _{i-2}	۱۲۵۱/۰۶	۴۲۹۷/۰۱	-۰/۵۴	-۰/۶۵
ANN- tangent sigmoid	Q, Q _i , Q _{i-1} , Q _{i-2} , Q _{i-3}	۱۴۳۷/۵۰	۴۴۷۱/۱۱	-۰/۵۱	-۰/۵۲
	Q	۱۱۴۳/۹۲	۳۵۶۱/۲۶	-۰/۶۷	-۰/۶۸
	Q, Q _i	۱۳۰۶/۸۷	۴۰۶۷/۵۸	-۰/۵۹	-۰/۵۹
	Q, Q _i , Q _{i-1}	۱۲۴۱/۷۱	۳۷۸۱/۹۲	-۰/۶۴	-۰/۶۴
GEP	Q, Q _i , Q _{i-1} , Q _{i-2}	۱۳۸۳/۵۰	۴۲۵۹/۸۱	-۰/۵۵	-۰/۵۹
	Q, Q _i , Q _{i-1} , Q _{i-2} , Q _{i-3}	۱۰۱۸/۱۵	۳۹۴۰/۲۵	-۰/۶۱	-۰/۶۲
	Q	۱۱۰۲/۶۲	۳۴۱۲/۲۱	-۰/۷۴	-۰/۷۴
	Q, Q _i	۱۲۱۷/۳۷	۳۸۵۸/۷۸	-۰/۶۳	-۰/۶۳
SRC	Q, Q _i , Q _{i-1}	۱۲۰۸/۴۲	۳۷۱۸/۲۷	-۰/۶۵	-۰/۶۵
	Q, Q _i , Q _{i-1} , Q _{i-2}	۱۳۱۲/۵۷	۴۱۸۰/۹۰	-۰/۵۷	-۰/۶۱
	Q, Q _i , Q _{i-1} , Q _{i-2} , Q _{i-3}	۱۲۷۰/۶۲	۳۷۴۲/۱۲	-۰/۶۴	-۰/۶۴
	Q	۳۹۱۴/۲۴	۹۳۷۶/۶۴	-۰/۳۳	-۰/۳۷

نتایج جدول ۵ نشان داد اضافه کردن متغیر بارندگی به متغیر دبی برای ورود به مدل‌های هوشمند، به‌طور چشم‌گیری دقت برآورد بار رسوب معلق روزانه را در مقایسه با ترکیباتی که تنها از متغیر دبی برای ورود به مدل‌ها استفاده کرده‌اند (جدول ۴) افزایش داده است. به‌عنوان مثال در جدول ۴، مدل ANN- log sigmoid با ترکیب متغیر ورودی دبی لحظه‌ای و دبی متوسط روزانه (Q, Q_i) دارای مقدار آماره‌های RMSE=۴۲۶۹/۴۲ (ton/day)، R²=۰/۵۶، NSE=۰/۵۵ و MAE=۱۳۴۹/۴۶ (ton/day) است. در صورتی که در مدل ANN- log sigmoid در جدول ۵ ترکیب متغیر (Q, Q_i, P_i)، دارای مقدار آماره‌های RMSE=۳۲۰۱/۲۹ (ton/day)، R²=۰/۵۶، NSE=۰/۷۷ و MAE=۱۱۰۴/۲۱ (ton/day) می‌باشد.

بنابراین مشاهده شد که اضافه کردن متغیر متوسط بارندگی روزانه به ترکیب متغیری دبی متوسط لحظه‌ای و دبی متوسط روزانه توانست به ترتیب باعث افزایش ۴۰ درصد و ۳۹ درصد در آماره‌های NSE و R² و کاهش ۲۵ درصد و ۱۸ درصد در آماره‌های RMSE و MAE شود. همچنین، ترکیب متغیری دبی لحظه‌ای، دبی متوسط روزانه، دبی متوسط روزانه تا سه روز قبل، بارندگی متوسط روزانه و بارندگی متوسط روزانه تا سه روز قبل (Q, Q_i, Q_{i-1}, Q_{i-2}, Q_{i-3}, P_i, P_{i-1}) در تمامی مدل‌های هوشمند شبکه عصبی مصنوعی با دو تابع فعال‌سازی تانژانت سیگموئید و تانژانت هایپربولیک و برنامه‌ریزی بیان ژن در جدول ۵، در مقایسه با سایر ترکیبات بالاترین میزان دقت و صحت در برآورد SSL را داشت.

متغیرهای مورد استفاده در این پژوهش دبی، بار رسوب معلق و دما با طول دوره آماری ۵ سال (۲۰۰۹-۲۰۰۵) بود. متغیرها به‌صورت تصادفی وارد مدل‌ها شدند. نتایج حاصل از این تحقیق نشان داد که مدل شبکه عصبی مصنوعی با R²=۰/۹۱ و (ton day⁻¹) RMSE=۱۵۷۵ توانست مقدار رسوب معلق را با دقت بالاتری نسبت به مدل منحنی سنجه رسوب با R²=۰/۵۷ و (ton day⁻¹) RMSE=۳۷۱۸ برآورد کند (Yadav et al., 2018). جایتولا و آدیمو، برای برآورد غلظت رسوب معلق در حوزه آبخیز سد اینادا در آفریقای جنوبی از مدل‌های برنامه‌ریزی ژن و منحنی سنجه رسوب استفاده کردند. این پژوهشگران از متغیرهای متوسط ماهانه غلظت رسوب معلق و دبی در طول دوره آماری ۱۴ سال (۲۰۱۳-۱۹۹۰) استفاده کردند. متغیرها به‌صورت تصادفی در دو گروه آموزش و آزمون وارد مدل‌ها شدند. نتایج حاصل از این پژوهش نشان داد که مدل برنامه‌ریزی ژن با R²=۰/۹۹ و RMSE=۰/۰۳۲ دارای قدرت بالاتری نسبت به مدل منحنی سنجه رسوب با R²=۰/۴۸ و RMSE=۸/۳۷ در برآورد غلظت رسوب معلق است (Jaiyeola and Adeyemo, 2019). طباطبایی و صالح‌پور جم برای برآورد بار رسوب معلق در حوزه آبخیز سلمان‌رود، مدل منحنی سنجه رسوب را با استفاده از روش‌های الگوریتم تکاملی کالبره کرده و برای افزایش قدرت تعمیم‌دهی مدل‌ها، از روش خوشه‌بندی با استفاده از شبکه عصبی نگاشت خودسازمان‌ده استفاده کردند. نتایج نشان دهنده تأثیر این روش‌ها در کاهش مقدار RMSE از ۵۷۵۴ به ۱۶۸۱ تن در روز است (Tabatabaei and Salehpour Jam, 2017).

جدول ۵- نتایج مدل‌سازی SSL با ترکیب متغیرهای دبی و بارندگی

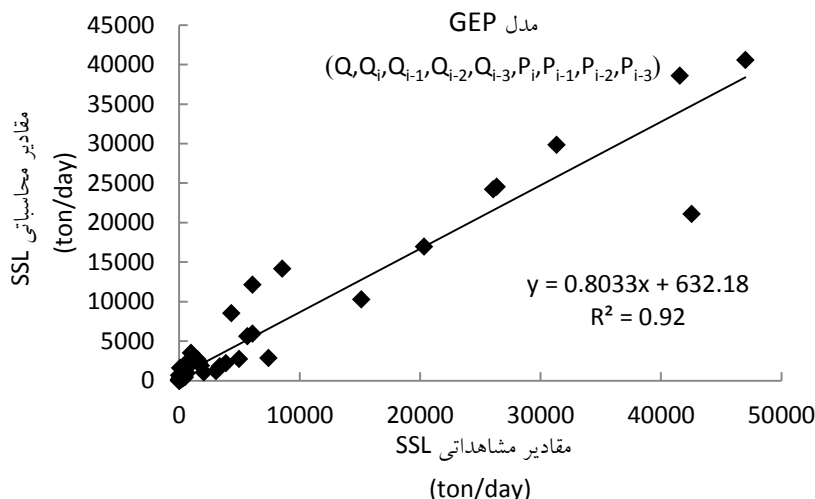
R ²	NSE	RMSE (ton/day)	MAE (ton/day)	ترکیب متغیرهای ورودی	مدل
۰/۷۸	۰/۷۷	۳۲۰۱/۲۹	۱۱۰۴/۲۱	Q, Q _i , P _i	ANN- log sigmoid
۰/۸۵	۰/۸۴	۲۵۰۱/۴۶	۸۴۱/۱۶	Q, Q _i , Q _{i-1} , Q _{i-2} , P _i , P _{i-1} , P _{i-2}	
۰/۷۵	۰/۷۵	۳۴۱۷/۹۰	۱۰۸۴/۲۱	Q, Q _i , Q _{i-1} , Q _{i-2} , Q _{i-3} , P _i	
۰/۸۳	۰/۸۲	۲۷۱۸/۲۲	۸۷۱/۲۳	Q, Q _i , Q _{i-1} , Q _{i-2} , Q _{i-3} , P _i , P _{i-1}	
۰/۸۸	۰/۸۷	۲۴۰۵/۸۱	۸۰۲/۴۱	Q, Q _i , Q _{i-1} , Q _{i-2} , Q _{i-3} , P _i , P _{i-1} , P _{i-2} , P _{i-3}	
۰/۷۵	۰/۷۴	۳۴۵۹/۸۱	۱۱۳۰/۷۸	Q, Q _i , P _i	ANN- tangent sigmoid
۰/۸۷	۰/۸۶	۲۴۴۵/۷۰	۷۲۱/۰۷	Q, Q _i , Q _{i-1} , Q _{i-2} , P _i , P _{i-1} , P _{i-2}	
۰/۷۷	۰/۷۸	۳۲۲۰/۱۵	۱۱۰۱/۸۹	Q, Q _i , Q _{i-1} , Q _{i-2} , Q _{i-3} , P _i	
۰/۸۲	۰/۸۰	۲۹۱۲/۶۸	۹۷۰/۲۸	Q, Q _i , Q _{i-1} , Q _{i-2} , Q _{i-3} , P _i , P _{i-1}	
۰/۸۹	۰/۸۷	۲۴۳۴/۵۲	۷۱۲/۷۶	Q, Q _i , Q _{i-1} , Q _{i-2} , Q _{i-3} , P _i , P _{i-1} , P _{i-2} , P _{i-3}	
۰/۷۰	۰/۷۰	۳۵۴۷/۰۲	۱۱۶۴/۷۸	Q, Q _i , P _i	GEP
۰/۸۵	۰/۸۵	۲۵۱۸/۷۴	۸۰۲/۸۲	Q, Q _i , Q _{i-1} , Q _{i-2} , P _i , P _{i-1} , P _{i-2}	
۰/۷۸	۰/۷۸	۳۲۰۲/۵۷	۹۹۸/۵۹	Q, Q _i , Q _{i-1} , Q _{i-2} , Q _{i-3} , P _i	
۰/۸۴	۰/۸۲	۲۷۵۸/۶۰	۸۵۷/۸۰	Q, Q _i , Q _{i-1} , Q _{i-2} , Q _{i-3} , P _i , P _{i-1}	
۰/۹۲	۰/۹۰	۲۲۸۲/۴۲	۷۵۰/۳۸	Q, Q _i , Q _{i-1} , Q _{i-2} , Q _{i-3} , P _i , P _{i-1} , P _{i-2} , P _{i-3}	

بارش همان روز و بارش با زمان تأخیر یک روز، دبی همان روز و دبی با زمان تأخیر دور روز و غلظت رسوب معلق در طول دوره آماری چهار سال (۲۰۰۰-۱۹۹۷) بود. متغیرها به صورت تصادفی در دو گروه داده‌های آموزش و آزمون قرار گرفتند. نتایج حاصل از این پژوهش نشان داد که مدل هوشمند با ترکیب متغیرهای ورودی بارش همان روز و دبی دو روز قبل با مقادیر $R^2=0/88$ و $1-1$ (mg) $104/85$ RMSE= در مقایسه با سایر مدل‌ها برآورد دقیق تری از غلظت رسوب معلق داشت (Nivesh et al., 2018). شکل ۴ و ۵ به ترتیب نمودار نقطه‌ای و نمودار روند نتایج بهترین برآورد بار رسوب معلق روزانه با استفاده از مدل برنامه‌ریزی بیان ژن را نشان می‌دهد.

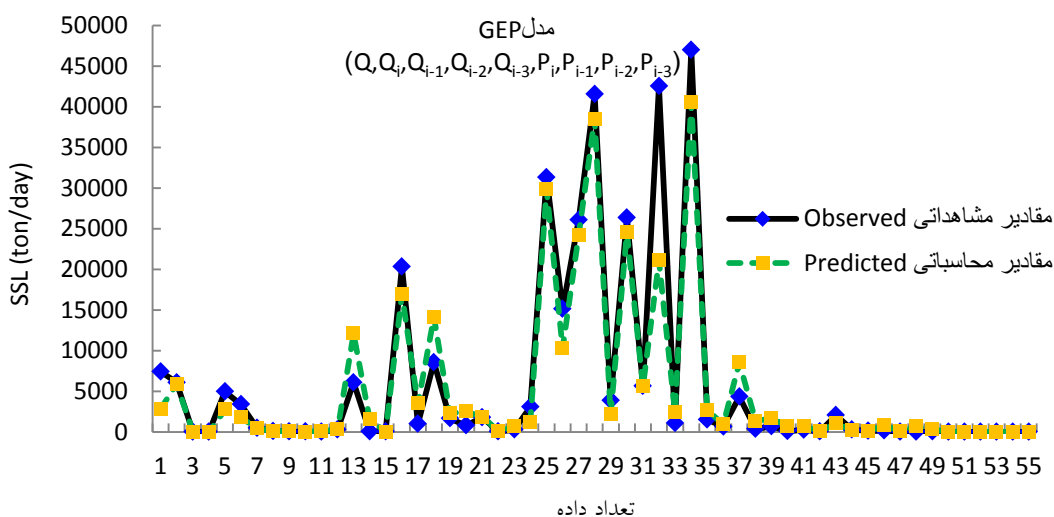
نتیجه‌گیری کلی

در این تحقیق، برای بررسی نقش متغیر بارندگی در میزان دقت برآورد بار رسوب معلق روزانه حوزه آبخیز سعیدآباد چای، از دو سناریوی ترکیب متغیری برای ورود به مدل‌ها استفاده شد. در سناریوی اول تنها از ترکیبات متغیری دبی استفاده شد و سناریوی دوم شامل ترکیبات متغیر ورودی بارندگی به همراه دبی حاصل از روش آزمون و خطا بود که برای ورود به مدل‌های هوشمند مورد استفاده قرار گرفت.

مدل GEP با این ترکیب متغیر ورودی، کارآمدترین مدل در برآورد صحیح بار رسوب معلق روزانه با بیشترین مقدار آماره‌های NSE برابر ۰/۹۰ و R^2 برابر ۰/۹۲ و کمترین مقدار آماره‌های RMSE برابر ۲۲۸۲/۴۲ (ton/day) و MAE برابر ۷۵۰/۳۸ (ton/day) بود. بنابراین، از جداول ۴ و ۵ نتیجه می‌شود که متغیر دبی به تنهایی نمی‌تواند واریانس رسوب رودخانه را به درستی تبیین نماید و استفاده از متغیر بارندگی به عنوان متغیر ورودی به مدل‌های هوشمند، نقش تأثیرگذار در افزایش دقت برآورد بار رسوب معلق دارد و استفاده از متغیر بارندگی به همراه متغیر دبی در طی فرآیند مدل‌سازی، کارایی مدل‌ها را افزایش داد. نتایج حاصل از این تحقیق در مورد استفاده از متغیر بارندگی به همراه دبی در برآورد SSL با نتایج مطالعات ذونعت کرمانی و همکاران مطابقت دارد. این محققان برای برآورد غلظت رسوب معلق روزانه در ایستگاه‌های هیدرومتری ایالت‌های آرکانساز، دلوار و اوهایو در آمریکا، از متغیرهای بارش روزانه و دبی جریان روزانه در طی فرآیند مدل‌سازی استفاده کردند. نتایج نشان داد که تأثیر بارش به همراه دبی در افزایش دقت برآورد ۲۳ درصد بیشتر از زمانی است که دبی به تنهایی استفاده می‌شود (Zounemat-Kermaniet al., 2016). در پژوهشی نیوش و همکاران برای برآورد غلظت رسوب معلق در حوزه آبخیز وامسادهارا در جنوب هند از مدل‌های هوشمند، رگرسیون خطی چندگانه و منحنی سنجه رسوب استفاده کردند. متغیرهای مورد استفاده در این پژوهش شامل متغیر



شکل ۴- نمودار نقطه‌ای نتایج بهترین برآورد SSL با استفاده از مدل GEP



شکل ۵- نمودار روند نتایج بهترین برآورد SSL با استفاده از مدل GEP

برآورد SSL را افزایش خواهد داد. همچنین نتایج این تحقیق نشان داد که مدل منحنی سنج رسوب در مقایسه با مدل‌های هوشمند، دقت بسیار کمی در برآورد بار رسوب معلق داشت. دلیل این امر می‌تواند به دلیل خاصیت رگرسیونی این مدل باشد. بدین معنی که اغلب داده‌های استفاده شده برای مدل‌سازی SRC مربوط به دبی‌های کم بوده و به دلیل این که بیشترین مقدار رسوب در دبی بالا حمل می‌گردد این مدل قادر به معرفی دوره‌های پر رسوب نیست و همچنین تبدیل نتایج این مدل از فضای لگاریتمی به فضای حسابی باعث کم‌برآوردی بار رسوب معلق خواهد شد. در صورتی که مدل‌های هوشمند مانند شبکه عصبی مصنوعی و برنامه‌ریزی بیان ژن حتی زمانی که داده‌ها از کمیت و کیفیت مناسبی برخوردار نیستند، برآورد دقیقی از مقدار SSL را خواهند داشت. از نتایج حاصل از مدل‌سازی

همچنین، برای افزایش قدرت تعمیم‌دهی مدل‌ها، خوشه‌بندی داده‌ها با استفاده از روش نگاشت خودسازمان‌ده انجام شد و تعداد بهینه خوشه با استفاده از شاخص دیویس بولدین تعیین شد. سپس داده‌های حاصل از خوشه‌بندی با استفاده از روش تخصیص متناسب در سه گروه داده‌های آموزش، اعتبارسنجی و آزمون قرار گرفتند و این داده‌ها در مدل‌سازی استفاده شدند. به طور کلی نتایج حاصل از این تحقیق نشان داد که تمام مدل‌های سناریوی دوم که از متغیر بارندگی به همراه دبی استفاده کردند، دارای مقدار آماره‌های NSE و R^2 بیشتر و مقدار آماره‌های RMSE و MAE کمتر در مقایسه با مدل‌های سناریوی اول بودند که تنها از متغیر دبی برای برآورد بار رسوب معلق استفاده کردند. بنابراین، استفاده از متغیر دینامیک بارندگی به همراه متغیر دبی در حوزه آبخیز، عملکرد مدل‌سازی در

Innovation and Development. 1-17.

Khan, M.Y.A., Tian, F., Hasan, F and Chakrapani, G.J. 2019. Artificial neural network simulation for prediction of suspended sediment concentration in the River Ramganga, Ganges Basin, India. *International Journal of Sediment Research*. 34. 2: 95-107.

Kisi, O and Ozkan, C. 2017. A new approach for modeling sediment-discharge relationship: Local weighted linear regression. *Water Resources Management*. 30. 2: 1-23.

Kisi, O and Fedakar, H.I. 2014. Modeling of Suspended Sediment Concentration Carried in Natural Streams Using Fuzzy Genetic Approach. *Computational Intelligence Techniques in Earth and Environmental Sciences*, Springer Netherlands. Pp: 175-196.

Lopes, H.S and Weinert, W.R 2004. EGYPSIS: An enhanced gene expression programming approach for symbolic regression problems. *Journal of Applied Mathematics and Computer Science*. 14. 3: 375-384.

May, R.J., Maier, H.R and Dandy, G.C. 2010. Data splitting for artificial neural networks using SOM-based stratified sampling. *Neural Networks*. 23. 2: 283-294.

Melesse, A.M., Ahmad, S., McClain, M.E., Wang, X and Lim, Y.H. 2011. Suspended sediment load prediction of river systems: An artificial neural network. *Agricultural Water Management*. 98. 5: 855-866.

Nivesh, S., Kumar, P., Sawant, P and Verma, R. 2018. Application of Fuzzy Logic and Statistical Approaches for Estimation of Suspended Sediment Concentration. *International Journal of Current Microbiology and Applied Sciences*. 7. 2: 3716-3733.

Nivesh, S.h and Kumar, P. 2017. Modelling river suspended sediment load using artificial neural network and multiple linear regression: Vamsadhara River Basin, India. *International Journal of Chemical Studies*. 5. 5: 337-344.

Ouillon, S. 2018. Why and how do we study sediment transport? Focus on coastal zones and ongoing methods. *Water*. 10: 4. 390p.

Rodríguez-Blanco, M.L., Taboada-Castro, M.M., Palleiro-Suárez, L and Taboada-Castro, M.T. 2010. Temporal changes in suspended sediment transport in an Atlantic catchment, NW Spain. *Geomorphology*, 123. 1: 181-188.

Tabatabaei, M and Salehpour Jam, A. 2017. Optimization of sediment rating curve coefficients using evolutionary algorithms and unsupervised artificial neural network. *Caspian Journal of Environmental Sciences*. 15. 4: 387-401.

این تحقیق، می‌توان به‌عنوان الگویی برای سایر مطالعات هیدرولوژیک حوزه آبخیز استفاده کرد.

منابع

ASCE. 2000. artificial neural networks in hydrology, I: Preliminary concepts. *Journal of Hydrologic Engineering, ASCE*. 5. 2: 115-123.

Azamathulla, H.M., Cuan, Y.C., Ghani, A.A and Chang, C.K. 2013. Suspended sediment load prediction of river systems: GEP approach. *Arabian Journal of Geosciences*. 6. 9: 3469-3480.

Bowden, G.J., Maier, H.R and Dandy, G.C. 2002. Optimal division of data for neural network models in water resources applications. *Water Resources Research*. 38. 2: 2-10.

Chiang, J.L., Tsai, K.J., Chen, Y.R., Lee, M.H and Sun, J.W. 2014. Suspended Sediment Load Prediction Using Support Vector Machines in the Goodwin Creek Experimental Watershed. In EGU General Assembly Conference Abstracts. 16: 5285.

Emamgholizadeh, S and Karimi Demneh, R. 2018. The comparison of artificial intelligence models for the estimation of daily suspended sediment load: a case study on Telar and Kasilian Rivers in Iran. *Water Science and Technology: Water Supply*. 1-14. <https://doi.org/10.2166/ws.2018.062>.

Ferreira, C. 2006. Automatically defined functions in gene expression programming. In: Nedjah, N. Mourelle, L.M. Abraham, A. (Eds.), *Genetic Systems Programming: Theory and Experiences, Studies in Computational Intelligence*, Springer-Verlag. 13. Pp: 21-56.

Ferreira, C. 2004. Gene expression programming and the evolution of computer programs. In: Castro, L.N. Von Zuben, F.J. (Eds.), *Recent Developments in Biologically Inspired Computing*, Pp: 82-103.

Ferreira, C. 2001. Gene expression programming: a new adaptive algorithm for solving problems. *Complex Systems*. 13. 2: 87-129.

Ghavidel, S.Z.Z and Montaseri, M. 2014. Application of different data-driven methods for the prediction of total dissolved solids in the Zarinroud basin. *Stochastic Environmental Research and Risk Assessment*. Pp: 1-18.

Güven, A and Kişi, O. 2011. Estimation of Suspended Sediment Yield in Natural Rivers Using Machine-coded Linear Genetic Programming. *Water Resources Management*. 25. 2: 691-704.

Jaiyeola, A.T and Adeyemo, J. 2019. Performance comparison between genetic programming and sediment rating curve for suspended sediment prediction. *African Journal of Science, Technology*,

- Wang, P., Xu, L., Zhou, S., Fan, Z., Li, Y and Feng, S. 2010. A Novel Bayesian Learning Method for Information Aggregation in Modular Neural Networks. *Expert Systems with Applications*. 37: 1071–1074.
- Wilson, W.S. 2008. Classifier conditions using gene expression programming. *IlliGAL report*. 4998: 206-217.
- Yadav, A., Chatterjee, S.S.k and Equeenuddin, M.d. 2018. Prediction of suspended sediment yield by artificial neural network and traditional mathematical model in Mahanadi river basin, India. *Sustainable Water Resources Management*. 4: 4: 745-759.
- Zounemat-Kermani, M., Kişi, O., Adamowski, J and Ramezani-Charmahineh, A. 2016. Evaluation of data driven models for river suspended sediment concentration modeling. *Journal of Hydrology*. 16. 2: 1-40.
- Tabatabaei, M., Solaimani, K. Habibnejad Roshan, M and Kavian, A. 2014. Estimation of daily Suspended Sediment Concentration using Artificial Neural Networks and data clustering by Self-Organizing Map (case study: Sira Hydrometry Station- Karaj Dam Watershed). *Journal of Watershed Management Research*, 5. 10: 98-116. (in persian)
- Tayfur, G. 2012. *Soft Computing in Water Resources Engineering - Artificial Neural Networks, Fuzzy Logic and Genetic Algorithms*. WIT Press, Southampton, England, UK. 267p.
- Theodosiou, M. 2011. Disaggregation & Aggregation of Time Series Components: A Hybrid Forecasting Approach Using Generalized Regression Neural Networks and the Theta Method, *Neurocomputing* 74: 896–905.
- Ulke, A., Tayfur, G and Ozkul, S. 2009. Predicting suspended sediment loads and missing data for Gediz River, Turkey. *Journal of Hydrologic Engineering*. 14. 9: 954-965.

Investigating the Role of Precipitation Variable in the Performance of daily Suspended Sediment Load Modeling (case study: Saied Abad Chai watershed)

A. Alijanpour Shalmani^{*1}, A. R. Vaezi², M.R. Tabatabaei³

Received: Jun.10, 2019

Accepted: Jul.22, 2019

Abstract

Knowledge of the exact amount of daily suspended sediment loads can be used to identify the erosion and sedimentation status of the watersheds. In this research, intelligent artificial neural network and Gene Expression Programming was models used to estimate the daily suspended sediment load. Also, due to the importance of the watershed response to the input variables of the models, in addition to the flow discharge variable, the dynamic precipitation variable was also selected for entering the models due to its influential role in creating erosion and sediment production. The results of this study showed that all the models that used the precipitation variable along with the flow discharge had higher NSE and R^2 statistic and lower RMSE and MAE statistics compared to the models that only the flow discharge variable to estimate the suspended sediment load. Also, GEP model with input variable combinations including instantaneous flow, average, average daily flow discharge with a delay time of three days, average daily precipitation and average daily precipitation with a delay time of three days, the most efficient model for estimating the daily suspended sediment load with the highest amount of statistics NSE was 0.90 and R^2 was 0.92 and the lowest value of RMSE was 2282.42 (ton/day) and MAE was 750.38 (ton/day) compared to artificial neural network models. In general, the results of this research showed that the flow discharge variable, alone could not properly explain the variance of river sediment. Using precipitation variable as input variable to intelligent models played a significant role in increasing the precision of estimation of suspended sediment load and using precipitation variables In addition to the flow discharge variable, during the modeling process, the efficiency of the models was increased.

Keywords: Artificial neural network, Gene Expression Programming, Self-organizing map, Precipitation

1- Ph.D. Student, Soil Science. Department, University of Zanjan, Zanjan, Iran

2- Full Professor, Soil Science. Department, University of Zanjan, Zanjan, Iran

3- Assistant Professor, Soil Conservation and Watershed Management Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Tehran, Iran

(*- Corresponding Author Email: adele.alijanpour@gmail.com)