ارزیابی کارایی دو مدل احتمالاتی شبکه عصبی رگرسیون تعمیم یافته (GRNN) و رگرسیون ناپارامتری K- نزدیک‌ترین همسایگی (KNN) برای مدل‌سازی ماهانه فرآیند بارش-رواناب

فرشته مدرسی

استادیار، گروه علوم و مهندسی آب، دانشگاه فردوسی مشهد
Fmodaresi@um.ac.ir

* * *

چکیده

یکی از فرآیندهای اصلی در مدل‌سازی هیدرولوژیکی حوضه‌های آبریز، فرآیند بارش- رواناب است و مدل‌سازی ماهانه آن در برنامه‌ریزی برای بهره‌برداری بلندمدت از منابع آب از اهمیت ویژه‌ای برخوردار است. با توجه به تعدد عوامل اثرگذار بر شکل‌گیری رواناب حاصل از بارش و عدم اندازه‌گیری دقیق مقادیر آنها در بلندمدت، مدل‌سازی بلندمدت این فرآیند، پیچیده و غیرخطی می‌باشد. از این‌رو، در تحقیق حاضر، کارایی دو مدل احتمالاتی شبکه عصبی رگرسیون تعمیم یافته (GRNN) و رگرسیون K- نزدیک‌ترین همسایگی که به ترتیب بر مبنای تابع پایه شعاعی و توابع کرنل می‌باشند، برای مدل‌سازی ماهانه فرآیند بارش- رواناب در حوضه بالادست سد کرخه مورد ارزیابی قرار گرفته است. به منظور تعیین پارامترهای اثرگذار بر هر یک از مدل‌ها، از روش صحت سنجی متقاطع LOOCV استفاده شده است. نتایج بدست آمده نشان می‌دهد که در حالت ارزیابی کلی از تمامی ماه‌ها، هر دو مدل از کارایی مناسبی برخوردارند و کارایی دو مدل، مشابه است ولی در حالت بررسی ماهانه بر اساس رده‌بندی شاخص‌های نش- ساتکلیف، RMSE و ضریب همبستگی، کارایی مدل GRNN بر KNN برتری دارد.

کلید واژه‌ها: مدل‌سازی ماهانه، بارش-رواناب، کرخه، صحت سنجی متقاطع، GRNN، KNN

* * *

مقدمه

یکی از مسائل اساسی در مدیریت بلندمدت حوضه آبریز، آگاهی از میزان منابع آب سطحی موجود می‌باشد. فرآیند بارش- رواناب اصلی‌ترین فرآیند هیدرولوژیکی یک حوضه آبریز است که میزان رواناب حاصل از بارش را تعیین می‌کند. در شکل‌گیری رواناب حاصل از بارش، عوامل بسیار گسترده و متنوعی اثر گذارند؛ از جمله این عوامل می‌توان به شیب زمین، جنس خاک، رطوبت موجود در خاک، میزان نفوذپذیری خاک، تراز آب زیر زمینی، نوع پوشش گیاهی، کاربری اراضی و ... اشاره نمود. تنوع و گستردگی عوامل اثرگذار بر این فرآیند نشان می‌دهد که فرآیند بارش- رواناب یک فرآیند غیرخطی و پیچیده است. از این رو استفاده از مدل‌هایی که قادر به مدل‌سازی شرایط پیچیده و غیر خطی باشند، نظیرانواع مدل‌های شبکه عصبی مصنوعی، ماشین بردار پشتیبان، K- نزدیک‌ترین همسایگی برای مدل‌سازی این فرآیند امری ضروری است. مطالعات صورت گرفته نشان می‌دهد که مدل‌های داده‌مبنا برای مدل‌سازی کوتاه مدت و بلندمدت جریان در

مقایسه با مدل‌های مفهومی از دقت مناسب و بالاتری برخوردارند؛ از آن جمله می‌توان بـه تحقیقـات صـورت گرفته زیر اشاره نمود:

Tokar و Marku (۲۰۰۰) نشان دادند که مدل شبکه عصبی مصنوعی (ANN) دارای کارایی بهتری نسبت به سه مدل مفهومی بیلان آب (Watbal) ، محاسبه رطوبت خاک (SAS) و بارش- رواناب مفهومی ساده (SCRR) برای مدل‌سازی بارش- رواناب در مقیاس‌های روزانه و ماهانه در سه حوضه متفاوت در آمریکـا اسـت. تحقیـق صورت گرفته توسط Nilsson و همکاران (۲۰۰۶) برای پیش‌بینی جریان ماهانه رودخانه در دو حوضه در نروژ نیز نشان داد که به علت بکارگیری داده‌های روزانه توسط مدل HBV، کارایی مدل ANN در مقیاس ماهانه بهتـر از مدل HBV است. همچنین،Nor و همکاران (۲۰۰۷) یک مدل RBF را برای مدل‌سازی رابطه میان بارش و رواناب در دو حوضه در سانجایی بکار بردند. آنها از این روش برای پیش‌بینی هیدروگراف جریان بر اساس رخدادهـای بارش استفاده کردند و برای آموزش مدلRBF، مدل شبکه عصبی رگرسـیون تعمـیم یافتـه (GRNN) را بکـار بردند. ارزیابی نتایج حاصله از این مدل و مقایسه آن با نتایج بدست آمده از مدل مفهومی HEC-HMS بیـانگر کارایی بیشتر و مدل‌سازی صحیح‌تر مدل RBF بود. در تحقیق Ghanbarpour و همکـاران (۲۰۱۲)، کـارایی دو مدل استوکستیک ARIMA و DARMA، مدل داده مبنای شـبکه عصـبی مصـنوعی (ANN)، مـدل فیزیکـی SWRRB و مدل مفهومی IHACRES برای مدل‌سازی فرآیند بارش- رواناب به منظور پیش‌بینی جریان ماهانـه در حوضه کسیلیان مورد ارزیابی قرار گرفته است. نتایج حاصل از این تحقیق نشان می‌دهد کـه کـارایی مـدل شبکه عصبی از سایر مدل‌ها بهتر بوده است. نظام خیاوی و قلمی (۱۳۹۰) کارایی روش شبکه عصبی مصـنوعی را برای مدل‌سازی بارش- رواناب حوضه آبریز قره‌سو در استان اردبیل با روش رگرسیون معمولی مورد مقایسه قرار دادند. نتایج آنها نشان داد که روش شبکه عصبی مصنوعی به علت توانـایی در مـدل‌سـازی فرآینـدهای غیرخطی، نتایج قابل قبول‌تری نسبت به روش رگرسیون ارائه کرده است. در تحقیق دیگری توسـط قربـانی و همکاران (۱۳۹۵) کارایی مدل‌های داده‌مبنای K نزدیک‌ترین همسایگی (KNN) و درخت تصـمیمM5 بـا مـدل مفهومی IHACRES برای شبیه‌سازی و پیش‌بینی جریان ماهانه رودخانه ارازکوسه در اسـتان گلسـتان مـورد مقایسه و ارزیابی قرار گرفته است. نتایج این تحقیق نشان داد که مدل‌های KNN و M5 نتایجی با دقت بالاتر را ارائه کرده‌اند.

با توجه به متفاوت بودن ساختار مدل‌های مبناء، کارایی آنها برای مدل‌سازی در شرایط مختلف، متفاوت است؛ تحقیقات صورت گرفته توسط Cigizuglu (۲۰۰۶)، Kisi و Cigizoglu (۲۰۰۷) و Kisi (۲۰۰۸) نشان می‌دهد که مدل شبکه عصبی GRNN در مقایسه با دو شبکه عصبی انتشار رو به جلو (FFNN) و شعاع مبنا (RBF) برای پیش‌بینی جریان متوسط ماهانه و روزانه از دقت بالاتری برخوردار است. Wu و Chau (۲۰۱۰) کارایی سه مدل ARMA، KNN و ANN را برای پیش‌بینی جریان ماهانه یک ماه بعد در دو رودخانه چین مورد بررسی قرار دادند. نتایج آنها نشان داد در شرایطی که همبستگی میان متغیرهای ورودی و خروجی مدل پایین است، کارایی مدل KNN نسبت به دو مدل دیگر بهتر می‌باشد. دل‌ناز و همکاران (۱۳۹۶) کارایی مدل شبکه عصبی GRNN را در مقایسه با مدل‌های شبکه عصبی ANN و RBF برای تخمین پارامترهای آبخوان محبوس بر اساس داده‌های آزمایش پمپاژ مورد ارزیابی قرار دادند. نتایج آنها نشان داد که مدل GRNN دارای کارایی بهتر از نظر دقت بالای نتایج و زمان اندک برای انجام محاسبات در مقایسه با دو مدل دیگر است. همچنین،

Modaresi و همکاران (۲۰۱۸) کارایی مدل‌های شبکه عصبی ANN و GRNN را با مدل‌های LS-SVR و KNN برای پیش‌بینی جریان براساس متغیرهای اقلیمی در شرایط خطی و غیرخطی مقایسه کردند. نتایج آنها نشان داد که کارایی مدل‌ها در شرایط خطی و غیر خطی متفاوت است و در شرایط خطی کارایی مدل‌های ANN و GRNN بهتر از KNN است ولی در شرایط غیر خطی کارایی مدل‌های KNN و LS-SVR بر مدل‌های ANN و GRNN برتری دارد. Khazaee Poul و همکاران (۲۰۱۹) کارایی چهار مدل MLR، KNN، ANN و ANFIS را برای پیش‌بینی جریان ماهانه با متغیرهای ورودی متفاوت و نیز در شرایط ترکیب شده با تبدیل موجک و بدون آن مورد ارزیابی قرار دادند. نتایج این تحقیق نشان داد در شرایط عدم استفاده از تبدیل موجک، مدل KNN دارای کارایی بهتری نسبت به سایرین است و با بکارگیری تبدیل موجک، کارایی سایر مدل‌ها نسبت به مدل KNN ارتقای بیشتری می‌یابد. Ren و همکاران (۲۰۲۰) اثر بکارگیری هشت روش فیلترکننده انتخاب ویژگی نظیر ضریب همبستگی خطی پیرسون، همبستگی خطی جزئی پیرسون، اطلاعات متقابل، اطلاعات متقابل شرطی، اطلاعات متقابل جزئی و ... را بر مدل‌های KNN، MLR و enELM برای پیش‌بینی جریان ماهانه در یک ماه بعد در سه مورد مطالعاتی مورد ارزیابی قرار دادند. نتایج آنها نشان داد که برای افزایش کارایی مدل‌های KNN و enELM هیچ یک از هشت روش فیلتر کننده ویژگی بر دیگری برتری ندارد.

ارزیابی کارایی مدل‌ها با توجه به تحقیقات فوق نشان دهنده کارایی بالای مدل‌های شبکه عصبی GRNN و رگرسیون K- نزدیک‌ترین همسایگی برای مدل‌سازی فرآیندهایی نظیر پیش‌بینی جریان روزانه و ماهانه و تخمین پارامترهای آبخوان در مقایسه با مدل‌هایی نظیر ANN، RBF، ARMA، ANFIS و MLR می‌باشد. با توجه به تفاوت کارایی مدل‌ها در شرایط مختلف، هدف از انجام این تحقیق، مقایسه کارایی دو مدل GRNN و KNN برای مدل‌سازی ماهانه فرآیند بارش- رواناب می‌باشد. با توجه به کمبود اطلاعات در دسترس از عوامل اثرگذار بر شکل‌گیری رواناب در یک حوضه آبریز در بلندمدت، مزیت استفاده از مدل‌های هوشمند و داده‌مبنا برای مدل‌سازی فرآیند بارش- رواناب در این است که برای استفاده از این مدل‌ها نیازی به استفاده از تمامی عوامل نمی‌باشد؛ در مقابل، هر چه تعداد متغیرهای ورودی کمتر و متغیرها بر خروجی مدل اثرگذارتر باشند، توانایی این مدل‌ها برای شناسایی الگوی موجود میان متغیرها بهتر می‌باشد. از این‌رو در تحقیق حاضر، با توجه به مقیاس زمانی بلندمدت ماهانه و کم شدن اثرگذاری مستقیم سایر عوامل اقلیمی و مورفولوژیکی بر مقدار جریان ماهانه رودخانه، برای ارزیابی کارایی دو مدل مذکور در شناسایی رابطه پیچیده و غیر خطی فرآیند بارش- رواناب ماهانه، تنها از دو متغیر بارش و جریان بدون تأخیر زمانی به عنوان ورودی و خروجی مدل‌ها استفاده شده است. در این تحقیق، بهترین ساختار برای هر یک از مدل‌ها بر اساس روش صحت سنجی متقاطع LOOCV تعیین شده و ارزیابی عملکرد این دو مدل به دو صورت کلی و ماهانه بر اساس رده‌بندی شاخص‌های نش- ساتکلیف، ریشه متوسط مربعات خطا و ضریب همبستگی پیرسون صورت گرفته است.

* * *

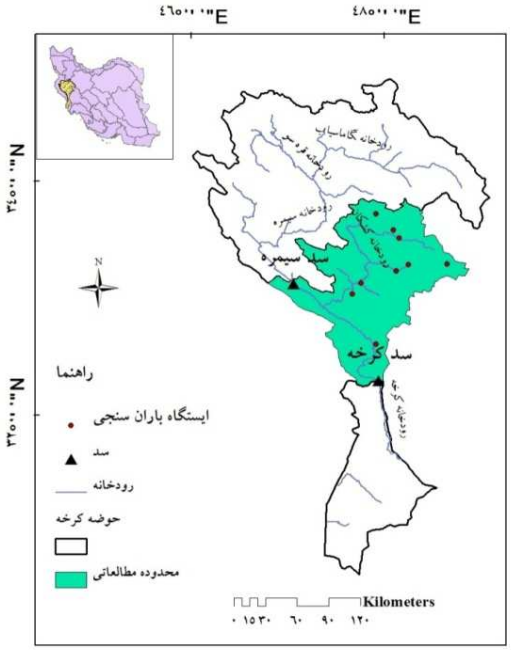
مواد و روش‌ها

معرفی منطقه مطالعاتی

در تحقیق حاضر حوضه کرخه به عنوان منطقه مطالعاتی انتخاب شده است. سد کرخه، سد اصلی حوضه کرخه و در پایین دست سد سیمره در طول جغرافیایی ''۴۸°۰۷'۳۱ و عرض جغرافیایی ''۳۲°۲۹'۲۲ بر روی

رودخانه کرخه قرار دارد. از این‌رو، حوضه کرخه را می‌توان به دو زیرحوضه در بالادست سد سیمره و بالادست سد کرخه تقسیم نمود. با توجه به اهمیت سد کرخه در تأمین آب کشاورزی جلگه خوزستان، در این تحقیق، مدل‌سازی فرآیند بارش- رواناب در زیرحوضه بالادست سد کرخه مورد ارزیابی قرار گرفته است که مساحت آن حدود ۱۴۴۳۰ کیلومتر مربع می‌باشد (شکل ۱).

به منظور انجام این تحقیق از آمار جریان ماهانه در ایستگاه هیدرومتری ورودی به سد کرخه و آمار متوسط بارش ماهانه ۹ ایستگاه باران‌سنجی و سینوپتیک بالادست سد واقع در زیرحوضه مطالعاتی در بازه زمانی ۳۲ ساله (۱۳۹۳-۱۳۶۱) استفاده شد که ۲۲ سال آن (۱۳۸۲-۱۳۶۱) برای واسنجی و ۱۰ سال دیگر (۱۳۹۳-۱۳۸۳) برای صحت سنجی مدل‌ها بکار برده شد. آمارهای بدست آمده از ایستگاه‌ها نشان می‌دهد که متوسط بارش ماهانه این زیرحوضه در طول دوره مطالعاتی برابر با ۵۲/۷ میلی‌متر بوده و بارش اصلی این زیرحوضه در ماه‌های آبان تا اردیبهشت رخ می‌دهد. همچنین، متوسط حجم جریان ماهانه ورودی به سد کرخه در طول دوره مطالعاتی برابر با ۱۴۳/۹ میلیون مترمکعب می‌باشد. در جدول ۱ نام ایستگاه‌ها و در شکل ۱ موقعیت ایستگاه‌های هواشناسی مورد استفاده نشان داده شده است.



شکل ۱- موقعیت زیرحوضه مطالعاتی و ایستگاه‌های هواشناسی آن

مدل‌های بکار رفته در تحقیق

شبکه عصبی مصنوعی رگرسیون تعمیم یافته (GRNN)

مدل شبکه عصبی رگرسیون تعمیم یافته (GRNN) یکی از انواع شبکه‌های عصبی تابع پایه شعاعی (Radial Basis Function) است که برای حل مسائل رگرسیون کاربرد دارد. این مدل که توسط Specht در سال۱۹۹۱ ارائه شد، یک تقریب‌گر همه منظوره برای توابع هموار (Smooth Functions) است و توانایی حل هر نوع از مسائل تقریب تابع هموار با داده‌های کافی را دارد. این مدل همانند مدل شبکه عصبی مصنوعی، یک روش هوش مصنوعی است که برای یادگیری روابط میان متغیرهای مستقل و وابسته از روش آموزش بانظر

بهره می‌گیرد. مزیت این مدل نسبت به سایر شبکه‌های عصبی در داشتن ساختار احتمالاتی است که سبب می‌شود این مدل در فرآیند مدل‌سازی با مشکل مینیمم‌های محلی که سایر شبکه‌های عصبی با آن درگیرند، مواجه نشود (Cigizoglu, 2006).

جدول ۱- مشخصات ایستگاه‌های باران‌سنجی زیرحوضه کرخه

ردیف	نام ایستگاه	طول جغرافیایی		عرض جغرافیایی		سازمان مربوطه
		درجه	دقیقه	درجه	دقیقه	
۱	پل دختر	۴۷	۴۹	۳۳	۱۳	وزارت نیرو
۲	نظر آباد	۴۸	۱	۳۳	۵۵	وزارت نیرو
۳	پل زال	۴۸	۵	۳۲	۴۸	وزارت نیرو
۴	آفرینه	۴۷	۵۴	۳۳	۱۹	وزارت نیرو
۵	چم انجیر	۴۸	۱۵	۳۳	۲۶	وزارت نیرو
۶	دهنو	۴۸	۴۶	۳۳	۳۱	وزارت نیرو
۷	کاکارضا	۴۸	۱۶	۳۳	۴۳	وزارت نیرو
۸	سراب سید علی	۴۶	۴۱	۳۳	۴۵	وزارت نیرو
۹	خرم آباد	۴۸	۲۲	۳۳	۲۹	سازمان هواشناسی

شبکه عصبی رگرسیون تعمیم یافته یک شبکه عصبی سه لایه است که در آن تعداد نورون‌های موجود در لایه اول و آخر مانند سایر شبکه‌های عصبی به ترتیب برابر با ابعاد بردارهای ورودی و خروجی است؛ ولی در این شبکه برخلاف سایر شبکه‌های عصبی، تعداد نورون‌های لایه میانی (مخفی) برابر با تعداد داده‌های مشاهداتی است. در این شبکه یک تابع احتمالاتی نرمال (گوسی) در هر یک از نورون‌های لایه مخفی وجود دارد که به صورت زیر می‌باشد :

$$f(X_r,t) = e^{-\left(\|X_r - X_t\| \times 0.8326/s\right)^2} \quad (1)$$

در رابطه (۱)، X_r: بردار ورودی به شبکه با خروجی نامشخص، X_t: مقادیر ورودی مشاهداتی در زمان t و s: پارامتر spread است که با تغییر آن، میزان جمع و بازشدگی تابع تغییر می‌نماید به نحوی که تابع، بهترین برازش را بر داده‌ها بیابد. پس از محاسبه این تابع احتمالاتی در هر یک از نورون‌های لایه مخفی، خروجی شبکه به ازای ورودی X_r از رابطه (۲) قابل محاسبه است (Araghinejad, 2014):

$$Y_r = \frac{1}{\sum_{t=1}^{n} f(X_r,t)} \sum_{t=1}^{n} \left[f(X_r,t) \times T_t \right] \quad (2)$$

در این رابطه، T_t: مقدار خروجی مشاهداتی متناظر با بردار ورودی t امین داده و n: تعداد داده‌های مشاهداتی است. از آنجایی که مقدار خروجی شبکه عصبی GRNN به پارامتر s وابسته است و این پارامتر برای تنظیم تابع در جهت انجام تخمین‌های دقیق‌تر بکار می‌رود، باید مقدار بهینه آن تعیین شود. به همین دلیل، در تحقیق حاضر با استفاده از کدنویسی در نرم‌افزار MATLAB، از روش صحت سنجی متقاطع (Leave One Out Cross Validation) برای تعیین بهترین مقدار پارامتر s برای مدل GRNN استفاده شده است.

رگرسیون K- نزدیک‌ترین همسایگی (KNN)

رگرسیون K نزدیک‌ترین همسایگی نوعی رگرسیون ناپارامتری می‌باشد که اولین بار توسط Karlsson و Yakowitz در سال ۱۹۸۷ برای پیش‌بینی و تخمین متغیرهای هیدرولوژیکی استفاده شد. در این روش، رابطه پارامتری از پیش تعیین شده‌ای میان متغیرهای ورودی و خروجی برقرار نمی‌شود بلکه در این روش، برای مدل‌سازی یک فرآیند نظیر مدل‌سازی فرآیند بارش- رواناب، از اطلاعات حاصل از داده‌های مشاهداتی بر اساس شباهت میان متغیرهای زمان واقعی مورد نظر و متغیرهای دوره مشاهداتی استفاده می‌شود.

منطق مورد استفاده در این روش، محاسبه احتمال رخداد یک واقعه بر اساس وقایع مشابه تاریخی آن (وقایع مشاهداتی) است. در این روش، برای تعیین میزان مشابهت شرایط کنونی به شرایط تاریخی از تابع احتمالاتی کرنل $f(D_{ri})$ به صورت زیر استفاده می‌شود:

$$f(D_{ri}) = \frac{1/D_{ri}}{\sum_{i=1}^{K} 1/D_{ri}} \qquad (۳)$$

که در آن، D_{ri}:فاصله اقلیدسی بردار شرایط کنونی (X_r) از بردار مشاهداتی تاریخی (X_i) و K: تعداد نزدیک‌ترین همسایگی‌ها به شرایط کنونی است. خروجی حاصل از این مدل رگرسیون (Y_r) به ازای بردار ورودی X_r، بر اساس رابطه کرنل فوق و مقادیر Y_i متناظر با هر D_{ri} از رابطه زیر محاسبه می‌شود:

$$Y_r = \sum_{i=1}^{K} f(D_{ri}) \times Y_i \qquad (۴)$$

در مدل KNN، انتخاب تعداد نزدیک‌ترین همسایگی (K) بر دقت نتایج اثرگذار است؛ به‌طوری که اگر تعداد همسایه‌ها زیاد باشد، نتایج به میانگین‌گیری از داده‌های مشاهداتی نزدیک می‌شود و اگر بسیار کم باشد، امکان زیاد شدن خطا افزایش می‌یابد. بنابراین، تعیین تعداد بهینه این پارامتر در این مدل برای دستیابی به کمترین خطا ضروری است. از این‌رو، در تحقیق حاضر، تعداد بهینه این پارامتر با استفاده از روش صحت سنجی متقاطع LOOCV تعیین شده است.

شاخص‌های ارزیابی مدل‌ها

برای ارزیابی و مقایسه کارایی مدل‌های بکار رفته در این تحقیق از سه شاخص ارزیابی نش- ساتکلیف (NSE)، جذر میانگین مربعات خطا (RMSE) و ضریب همبستگی پیرسون (R) استفاده شده است. رده‌بندی کارایی مدل‌ها بر اساس این سه شاخص در جدول ۲ آورده شده است: (Moriasi et al., 2007).

جدول ۲- طبقه‌بندی کارایی مدل‌ها بر اساس شاخص‌های ارزیابی

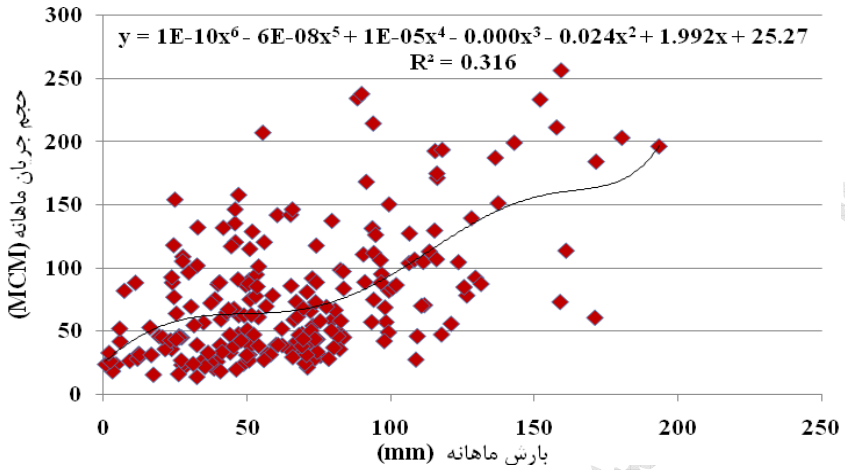
طبقه کارایی مدل‌ها	شاخص‌های ارزیابی مدل‌ها		
	NSE	RMSE	R
بسیار خوب	۰/۷۵ < NSE ≤ ۱	. ≤ RMSE≤ ۰/۵ SD*	۰/۹۳ < R < ۱/۰
خوب	۰/۶۵ < NSE ≤ ۰/۷۵	۰/۵ SD < RMSE ≤ ۰/۶ SD	۰/۸۸ < R < ۰/۹۲
مناسب	۰/۵ < NSE ≤ ۰/۶۵	۰/۶ SD < RMSE ≤ ۰/۷ SD	۰/۸۱ < R < ۰/۸۷
نامناسب	NSE ≤ ۰/۵	RMSE >۰/۷ SD	R< ۰/۸

*SD= انحراف معیار

* * *

نتایج و بحث

برای انجام فرآیند مدل سازی، ابتدا رابطه میان بارش متوسط ماهانه زیرحوضه بالادست سد و جریان ماهانه ورودی به سد کرخه بررسی شد که در شکل ۲ نشان داده شده است.



شکل۲- رابطه میان بارش متوسط ماهانه و حجم جریان ماهانه در زیرحوضه کرخه

همان‌طور که در شکل ۲ مشاهده می‌شود، رابطه میان این دو متغیر غیرخطی است؛ به‌طوری‌که از میان توابع خطی، نمایی، لگاریتمی، توانی و چندجمله‌ای با درجه‌های ۲ تا ۶، بهترین تابع برازش یافته براساس ضریب تبیین، تابع چندجمله‌ای درجه ۶ می‌باشد که ضریب تبیین آن ۰/۳۱۶ است و نشان دهنده شرایط پیچیده غیرخطی می‌باشد. از این‌رو، بکارگیری توابع احتمالاتی برای مدل‌سازی این رابطه ضروری است.

با بکارگیری دو مدل داده‌مبناء و احتمالاتی GRNN و KNN برای مدل‌سازی فرآیند بارش- رواناب در زیرحوضه مطالعاتی و تعیین مقادیر بهینه پارامترهای مدل‌ها به روش صحت سنجی متقاطع LOOCV، نتایج کلی زیر برای تمامی ماه‌ها در مراحل واسنجی و صحت سنجی حاصل شد:

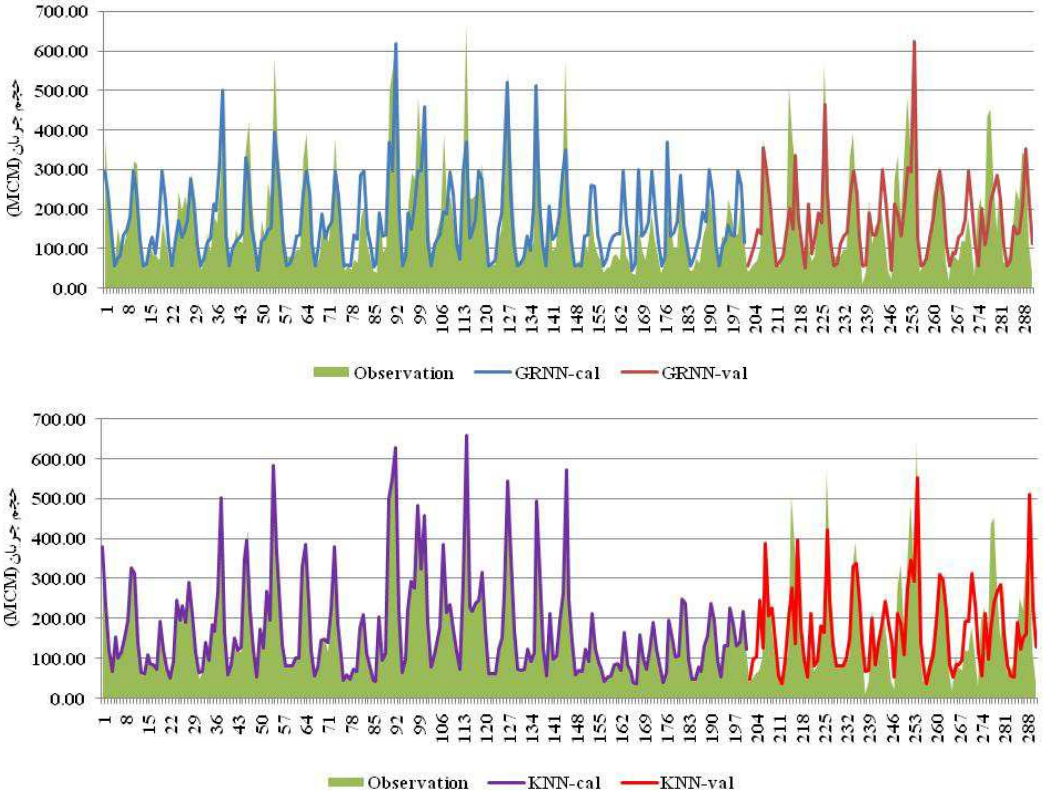
جدول ۳- نتایج ارزیابی مدل‌ها در مراحل واسنجی و صحت سنجی بر اساس شاخص‌های ارزیابی و مقادیر بهینه پارامترها

نام مدل	آموزش (واسنجی)			صحت سنجی			مقادیر بهینه پارامترها
	NSE	RMSE (MCM)	R	NSE	RMSE (MCM)	R	
GRNN	۰/۶۸۵	۷۱/۶۹۰	۰/۸۳۰	۰/۶۱۸	۸۳/۱۶۳	۰/۷۹۳	Spread= ۰/۶۷
KNN	۰/۸۶۲	۴۷/۳۸۹	۰/۹۲۹	۰/۶۰۱	۸۴/۹۷۷	۰/۷۷۷	K= ۳

با توجه به جدول ۳ و براساس طبقه‌بندی شاخص‌های ارزیابی (جدول ۲) مشاهده می‌شود که هر دو مدل در مرحله صحت سنجی از کارایی مناسبی بر اساس دو شاخص NSE و RMSE برخوردار هستند؛ ولی بر اساس معیار ضریب همبستگی، دقت نتایج آنها مناسب نیست. در شکل ۳ نتایج حاصل از مدل‌ها در مقایسه با مقادیر مشاهداتی نشان داده شده است.

با توجه به شکل ۳ مشاهده می‌شود که هر دو مدل برای مدل‌سازی مقادیر متوسط از کارایی بالایی برخوردارند و مقادیر مدل‌سازی شده و مشاهداتی بر هم منطبق می‌باشند ولی در برخی از نقاط اکسترمم (جریان‌های حداکثر و حداقل)، از دقت مناسبی برخوردار نبوده‌اند؛ با این وجود، در مقادیر شبیه‌سازی شده

توسط این دو مدل، مقادیر غیر منطقی نظیر مقدار منفی برای جریان و یا اعداد خیلی بزرگ و غیر واقعی وجود ندارد و این یک مزیت مهم برای آنها می‌باشد.

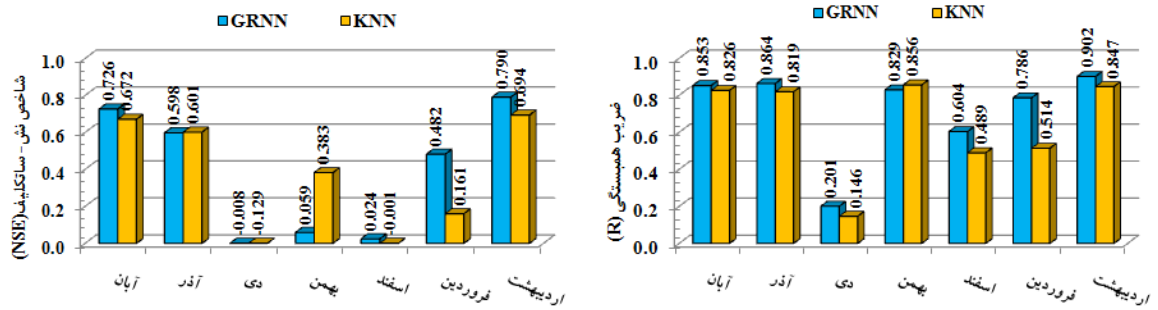


شکل ۳- مقادیر جریان مشاهداتی و شبیه‌سازی شده توسط مدل‌های GRNN و KNN

نظر به اینکه الگوهای بارش- رواناب در ماه‌های مختلف متفاوت است، ارزیابی ماهانه از کارایی مدل‌ها نیز صورت گرفت. نتایج این ارزیابی براساس دو شاخص نش- ساتکلیف و ضریب همبستگی در مرحله صحت سنجی در شکل ۴ آورده شده است.

با توجه به شکل ۴ مشاهده می‌شود که بر اساس هر دو شاخص کارایی نش- ساتکلیف و ضریب همبستگی، دقت نتایج مدل GRNN در تمامی ماه‌ها بجز بهمن ماه از مدل KNN بیشتر است. بر اساس رده‌بندی شاخص نش- ساتکلیف، نتایج مدل GRNN در ماه اردیبهشت در رده بسیار خوب، در آبان در رده خوب و در ماه آذر در رده مناسب قرار دارد. همچنین بر اساس شاخص ضریب همبستگی، نتایج مدل GRNN در ماه‌های آبان، آذر و بهمن در رده مناسب و در ماه اردیبهشت در رده خوب قرار دارد. این در حالی است که نتایج مدل KNN بر اساس شاخص نش- ساتکلیف در دو ماه آبان و اردیبهشت در رده خوب و در ماه آذر در رده مناسب قرار دارد و در ماه‌های آبان، آذر، بهمن و اردیبهشت بر اساس رده بندی شاخص ضریب همبستگی، کارایی این مدل در رده مناسب قرار دارد. با این وجود، بر اساس شاخص نش- ساتکلیف دقت نتایج حاصل از هر دو مدل در ماه‌های دی تا فروردین در رده نامناسب قرار دارد که علت آن رخداد بارش‌های حدی در این ماه‌ها می‌باشد. همچنین، بررسی ضریب همبستگی ماهانه نشان می‌دهد که با وجود قرارگیری ضریب

همبستگی کلی نتایج دو مدل در رده نامناسب، ضریب همبستگی نتایج دو مدل تنها در ماه‌های دی، اسفند و فروردین در رده نامناسب قرار دارد و در سایر ماه‌ها در رده مناسب به بالا قرار گرفته است.



شکل ۴- ارزیابی ماهانه کارایی مدل‌های GRNN و KNN برای مدل‌سازی بارش-رواناب

* * *

نتیجه گیری

در این پژوهش، کارایی دو مدل شبکه عصبی رگرسیون تعمیم یافته (GRNN) و رگرسیون K-نزدیک‌ترین همسایگی (KNN) که هر دو از نوع داده مبنا و احتمالاتی با ساختارهای متفاوت می‌باشند برای مدل‌سازی ماهانه فرآیند بارش رواناب مورد ارزیابی قرار گرفت. این ارزیابی بر اساس مدل‌سازی ماهانه بارش-رواناب در زیرحوضه بالادست سد کرخه صورت پذیرفت و در آن از داده‌های ۳۲ ساله بارش متوسط ماهانه زیرحوضه موردنظر و جریان ماهانه ورودی به سد کرخه استفاده شد. نتایج نشان داد که این دو متغیر دارای رابطه پیچیده و غیرخطی از درجه ۶ به بالا می‌باشند. ساختار بهینه مدل‌ها با استفاده از روش صحت سنجی متقاطع LOOCV تعیین شد و نتایج مدل‌ها بر اساس تمامی ماه‌ها نشان داد که کارایی هر دو مدل در مرحله صحت سنجی مشابه و در رده مناسب بر اساس دو شاخص NSE و RMSE قرار دارد. هر دو مدل برای مدل‌سازی مقادیر حدی اندکی ضعیف هستند ولی در نتایج هیچ یک از آنها، مقادیر منفی و غیرقابل قبول وجود ندارد. با این‌حال، ارزیابی ماهانه نتایج مدل‌ها نشان می‌دهد که با وجود تشابه کلی کارایی دو مدل، دقت نتایج مدل GRNN در تمامی ماه‌ها بجز بهمن ماه از مدل KNN بیشتر می‌باشد و از این جهت، با توجه به بررسی ماهانه، کارایی مدل GRNN بر KNN برتری دارد.

* * *

منابع فارسی

- دل‌ناز، ع.، رخشنده‌رو، غ.، نیکو، م.ر. (۱۳۹۶)، کارایی مدل GRNN در قیاس با مدل‌های ANN و RBF در تخمین پارامترهای آبخوان محبوس، هیدروژئولوژی، ۲(۱)، ۱۰۲-۱۱۷.
- قربانی، خ.، سهرابیان، ا. و سالاری‌جزی، م. (۱۳۹۵)، ارزیابی روش‌های هیدرولوژیکی و داده‌کاوی در شبیه‌سازی و پیش‌بینی دبی جریان ماهانه (مطالعه موردی: ایستگاه هیدرومتری ارازکوسه). پژوهش‌های حفاظت آب و خاک. ۲۳ (۱)، ۲۰۳-۲۱۷.
- نظام خیاوی، خ. و قلمی، و.، ۱۳۹۰، استفاده از شبکه عصبی مصنوعی در مدل‌سازی بارش- رواناب. مطالعه موردی: حوضه آبریز قره‌سو واقع در استان اردبیل. سومین همایش ملی عمران شهری، چهارم آبان ماه، سنندج.

* * *

References

- Araghinejad, S. (2014), Data-Driven Modeling: Using MATLAB in Water Resources and Environmental Engineering. 1st Ed. Water Science and Technology Library, Springer, Volume (67), 265p.
- Cigizuglu, H.K. (2006), Generalized regression neural network in monthly flow forecasting, Civil Engineering and Environmental Systems, 22(2), 71-81.
- Ghanbarpour, M. R., M. Amiri, M. Zarei and Z. Darvari. (2012), Comparison of streamflow predicted in a forest watershed using different modeling procedures: ARMA, ANN, SWRRB, and IHACRES models, River Basin Management, 10 (3), pp 281-292.
- Karlsson, M., and Yakowitz, S. (1987), Nearest-neighbor methods for nonparametric rainfall-runoff forecasting, Water Resources Research, 23(7), pp 1300–1308.
- Khazaee Poul, A., Shourian, M. and Ebrahimi, H. (2019), A Comparative Study of MLR, KNN, ANN and ANFIS Models with Wavelet Transform in Monthly Stream Flow Prediction, Water Resources Management, 33, pp 2907-2923.
- Kisi, O. and Cigizoglu, H.K. (2007), Comparison of different ANN techniques in river flow prediction, Civil Engineering and Environmental Systems, 24(3), pp 211-231.
- Kişi, Ö. (2008), River flow forecasting and estimation using different artificial neural network techniques, Hydrology Research, 39(1), pp 27–40.
- Modaresi, F., Araghinejad, S. and Ebrahimi, K. (2018), A comparative assessment of Artificial Neural Network, Generalized Regression Neural Network, Least-Square Support Vector Regression, and K-Nearest Neighbor Regression for monthly streamflow forecasting in linear and nonlinear conditions, Water Resources Management, 32(1), pp 243-258.
- Moriasi, D. N., J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, T. L. Veith. (2007), Model evaluation guidelines for systemic quantification of accuracy in watershed simulations, Transactions of ASABE, 50(3), pp 885-900.
- Nilsson, P., Uvo, C. B. and Berndtsson, R. (2006), Monthly runoff simulation: comparing and combining conceptual and neural network models, Journal of Hydrology, 321, pp 344–363.
- Nor, N. I. A., Harun, S., and Kassim, A. H. M. (2007), Radial Basis Function modeling of hourly streamflow hydrograph, Journal of Hydrologic Engineering, 12(1), pp 113-123.
- Ren, K., Fang, W., Qu, J., Zhang, X. and Shi, X. (2020), monthly streamflow forecasting – Three case studies on CAMELS data sets, Journal of hydrology, 586, https://doi.org/10.1016/j.jhydrol.2020.124897
- Specht, D.F. (1991), A General Regression Neural Network, Transactions on Neural Network, 2(6), pp 578-576.
- Tokar, A. S. and Markus, M. (2000), Precipitation-runoff modeling using artificial neural networks and conceptual models, Journal of Hydrologic Engineering, 5 (2), pp 156–161.
- Wu, C.L., and Chau, K.W. (2010), Data-driven models for monthly streamflow time series prediction, Engineering Application of Artificial Intelligence, 23(8), pp 1350-1367.