

کاربرد شبکه‌های عصبی بیزین، ماشین بردار پشتیبان و برنامه‌ریزی بیان ژنی در تحلیل بارش - رواناب ماهانه (مطالعه موردی: رودخانه کاکارضا)

محمد علی قربانی^{۱*} و رضا دهقانی^۲

^{۱*} - نویسنده مسئول، دانشیار گروه مهندسی آب، دانشگاه تبریز

^۲ - دانش آموز ارشد منابع آب، دانشگاه تبریز

تاریخ پذیرش: ۹۴/۳/۶

تاریخ دریافت: ۹۲/۹/۱۲

چکیده

شبیه‌سازی فرآیند بارش - رواناب اولین و مهمترین گام برای کنترل سیلاب در مدیریت منابع آب می‌باشد. در این تحقیق فرآیند بارش - رواناب ماهانه رودخانه کاکارضا واقع در استان لرستان، با استفاده از شبکه عصبی بیزین مورد بررسی قرار گرفت و نتایج آن با روش‌های برنامه‌ریزی بیان ژن و ماشین بردار پشتیبان مقایسه گردید. بر این اساس ترکیب‌های مختلفی با استفاده از پارامترهای بارندگی و رواناب، طی دوره آماری (۱۳۹۲-۱۳۴۸) به‌عنوان ورودی برای تخمین رواناب در مقیاس زمانی ماهانه مورد ارزیابی قرار گرفت. معیارهای ضریب همبستگی، ریشه میانگین مربعات خطا، ضریب نش ساتکلیف و بایاس برای ارزیابی و نیز مقایسه عملکرد روش‌ها مورد استفاده قرار گرفت. نتایج حاصله نشان داد ساختارهای ترکیبی با استفاده از روش‌های هوشمند مورد بررسی، در شبیه‌سازی فرآیند بارش - رواناب نتایج قابل قبولی ارائه می‌نماید. مقایسه مدل‌ها نشان داد شبکه عصبی بیزین عملکرد بهتری نسبت به سایر مدل‌ها در شبیه‌سازی فرآیند بارش - رواناب دارد. در مجموع روش شبکه عصبی بیزین توانایی بالایی در تخمین مقادیر کمینه، میانی و مقادیر بیشینه رواناب از خود نشان داده است.

کلمات کلیدی: بارش - رواناب، شبکه عصبی بیزین، برنامه‌ریزی بیان ژن، ماشین بردار پشتیبان.

Application of Bayesian Neural Networks, Support Vector Machines and Gene Expression Programming Analysis of Rainfall - Runoff Monthly (Case Study: Kakarza River)

M. A. Ghorbani^{1*} and R. Dehghani²

^{1*} - Department of Water Engineering, Faculty of Agriculture, University of Tabriz

² - Graduate Student, Department of Water Engineering, Faculty of Agriculture, University of Tabriz

Received: 3 December 2013

Accepted: 27 May 2015

Abstract

Simulation of rainfall - runoff process is one of the most important tasks in water resources management and flood control studies. In this study, the rainfall - runoff process over Kakarza river located at Lorestan province, was simulated using the Bayesian neural network and the results were compared with the gene expression and support vector machine models. In this case, different combinations of monthly rainfall and runoff data in period of 1969-2013 were considered as the input data of the models. Four performance criteria namely, correlation coefficient, root mean square error, Nash-Sutcliffe coefficient and bias were used to evaluate and compare the performance of the models. The results showed that the performance of the models were satisfactory. Results showed that, the Bayesian neural network model is more efficient than the other models in estimation of minimum, mean and peak of runoff .

Keywords: Rainfall-runoff, Bayesian neural network, Gene expression programming, Support vector machine.

مقدمه

برنامه‌ریزی بهره‌برداری از منابع آب‌های سطحی و زیرزمینی، ساماندهی رودخانه‌ها و هشدار سیل نیاز به پیش‌بینی آینده رودخانه و رواناب حوضه آبریز دارد، بنابراین آگاهی از چگونگی ارتباط بین بارندگی و رواناب حوضه‌های آبریز بخش جدانشدنی مطالعات طرح‌های آبی می‌باشد. خاصیت غیرخطی، عدم قطعیت ذاتی فرآیند بارش-رواناب، نیاز به اطلاعات وسیع و پیچیده بودن مدل‌های فیزیکی از جمله دلایلی بوده است که باعث شده محققان برای پیش‌بینی پدیده‌های غیرخطی، روش‌های هوشمند را بکار گیرند، که روش شبکه عصبی بیزین (BNN^۱)، برنامه‌ریزی بیان ژن (GEP^۲) و ماشین بردار پشتیبان (S.V.M^۳) از جمله این روش‌ها است. شبکه‌های عصبی بیزین به‌عنوان یکی از روش‌های مبتنی بر شبکه‌های عصبی در مدل کردن مسائل غیرخطی و پیچیده از طریق الگوریتم‌های خاص و روش‌های آماری است. با استفاده از این روش می‌توان به مدل کردن روابط علت و معلولی یک فرآیند، تحلیل وضعیت موجود و تخمین وضعیت آینده یک سیستم پرداخت (طباطبایی و دشتی زاده، ۱۳۸۷). برنامه‌ریزی ژن یک فن برنامه‌ریزی خودکار است که راه‌حل مسئله را با استفاده از برنامه‌ریزی کامپیوتر ارائه کرده و عضوی از خانواده الگوریتم تکاملی می‌باشد (قربانی و صالحی، ۱۳۹۰). ماشین بردار پشتیبان یک سیستم یادگیری کارآمد بر مبنای تئوری بهینه‌سازی مقید است که از اصل استقرای کمینه‌سازی خطای ساختاری استفاده کرده و منجر به یک جواب بهینه کلی می‌گردد. در سال‌های اخیر استفاده از روش‌های هوشمند در مدل‌سازی فرآیند بارش-رواناب مورد توجه محققین قرار گرفته است، که از جمله می‌توان به موارد ذیل اشاره نمود: سلطانی و همکاران (۱۳۸۹) با کاربرد برنامه‌ریزی ژنتیک در مدل‌سازی فرآیند بارش-رواناب روزانه حوضه آبریز لیقوان، با تعریف دو مجموعه عملگر ریاضی و با محاسبه ضریب همبستگی و میانگین مربعات خطا به این نتیجه رسیدند که مدل حاصل از مجموعه عملگرهای اصلی، مناسب‌تر است. از سوی دیگر دیباک و همکاران (۲۰۰۱) قابلیت مدل ماشین بردار پشتیبان را در خصوص پیش‌بینی‌های هیدرولوژیکی نشان دادند. آنها با استفاده از این روش در زمینه دسته‌بندی داده‌های سنجش از راه دور و مدل‌سازی فرآیند بارش-رواناب استفاده نموده و نتایج حاصل را با روش‌های شبکه عصبی مصنوعی مورد مقایسه قرار دادند، که نتایج حاصل دارای دقت قابل قبولی بود. لیونگ و همکاران (۲۰۰۲) با مطالعه رابطه بارش-رواناب در زمان‌های متفاوت به این نتیجه دست یافتند که استفاده از روش برنامه‌ریزی ژنتیک در پیش‌بینی رفتار بارش-رواناب در حوضه‌های آبریز

سبب بروز خطای کمتری می‌گردد. جایاواردنا و همکاران (۲۰۰۵) با استفاده از برنامه‌ریزی ژنتیک فرآیند بارش-رواناب را با داده‌های روزانه در دو حوضه نسبتاً بزرگ چین مدل‌سازی کردند که نتایج حاصل از برنامه‌ریزی ژنتیک با داده‌های واقعی مطابقت خوبی نشان داد. از سوی دیگر آیتک و همکاران (۲۰۰۸) از شبکه‌های عصبی و برنامه‌ریزی ژنتیک برای مدل‌سازی بارش-رواناب روزانه حوضه رودخانه جونیاتا^۴ در ایالت پنسیلوانیای آمریکا استفاده کرده و نتیجه گرفتند که برنامه‌ریزی ژنتیک با دقت بهتری نسبت به شبکه‌های عصبی مصنوعی فرآیند بارش-رواناب را مدل‌سازی می‌نماید. طاهری و غفوری (۲۰۱۲) دو روش یادگیری فعال^۵ و ماشین بردار پشتیبان را برای شبیه‌سازی جریانات روزانه رودخانه کارون به کار برده و قابلیت بالای ماشین بردار پشتیبان را در شبیه‌سازی رواناب مطرح نمودند. در مجموع با توجه به پژوهش‌های انجام‌شده و ذکر این نکته که رودخانه کاکارضا یکی از مهم‌ترین رودخانه‌های استان لرستان و مهم‌ترین منبع تأمین‌کننده آب بخش‌های مختلف و نواحی مجاور خود می‌باشد، اهمیت شبیه‌سازی فرآیند بارش-رواناب این رودخانه و اقدامات مدیریتی برای بهبود آب آن بیش‌ازپیش ضروری است. لذا هدف این تحقیق بررسی روش‌های جدید و ارزیابی عملکرد و دقت آنها در شبیه‌سازی فرآیند بارش-رواناب می‌باشد.

مواد و روش‌ها

منطقه مورد مطالعه و داده‌ها

منطقه مورد مطالعه رودخانه کاکارضا واقع در استان لرستان می‌باشد. این رودخانه از رودخانه‌های دائمی استان لرستان بوده و از کوه‌های جنوب شرقی شهرستان الشتر و بخش چغلوندی (هرود) سرچشمه گرفته و در محدوده شهرستان الشتر بنام کاکارضا معروف است. این رودخانه بین " ۱۵ ° ۴۸ تا ۴۹ درجه طول شرقی و " ۲۲ ° ۳۲ الی " ۵۲ ° ۳۳ درجه عرض شمالی و در استان لرستان و شرق شهرستان خرم‌آباد واقع شده است و قسمتی از سرشاخه‌های رودخانه کرخه را در زاگرس تشکیل می‌دهد. رودخانه کاکارضا در ارتفاع ۱۵۵۰ متری از سطح آب‌های آزاد قرار دارد. مساحت حوضه کاکارضا ۱۱۴۸ کیلومترمربع و دارای رودخانه‌ای به طول ۸۵ کیلومتر می‌باشد. رودخانه کاکارضا پس از پیوستن به رودخانه‌های کشکان، سیمره و کرخه، در نهایت به خلیج فارس می‌ریزد. که در شکل (۱) موقعیت جغرافیایی منطقه مورد مطالعه نشان داده شده است. در این تحقیق داده‌های بارندگی و رواناب در مقیاس زمانی ماهانه مربوط به ایستگاه هرود (کاکارضا) از سال ۱۳۴۸ تا ۱۳۹۲ که در شرکت آب منطقه‌ای و سازمان هواشناسی استان لرستان در دسترس بود، استفاده گردید.

6- Jayawardena *et al.*

7- Aytak *et al.*

8- Juniata

9- Active Learning Method

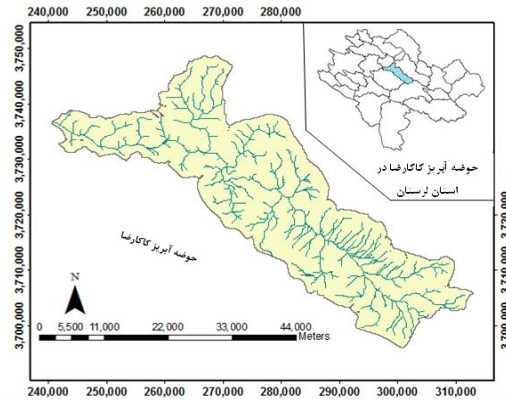
1 - Bayesian Neural Network

2 - Gene Expression Programming

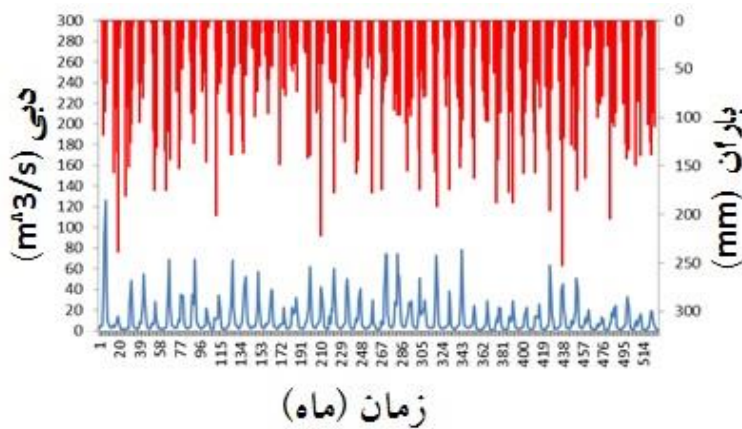
3- Support Vector Machines

4-Dibike *et al.*

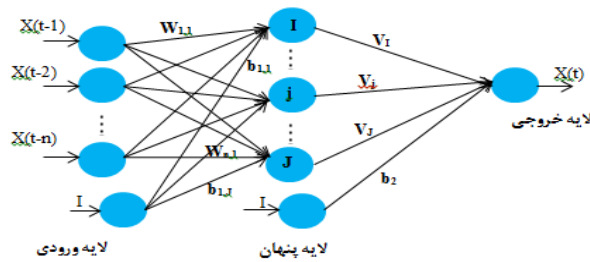
5- Liong *et al.*



شکل ۱- موقعیت محدوده مطالعاتی در استان لرستان



شکل ۲- مقادیر مشاهداتی بارش و رواناب ماهانه کاکارضا (۱۳۹۲-۱۳۹۸)



شکل ۳- نمای کلی یک شبکه عصبی بیزین سه لایه‌ای

جدول ۱- خصوصیات آماری مقادیر بارش-رواناب ماهانه مورد استفاده در بازه زمانی (۱۳۹۲-۱۳۹۸)

پارامتر	واحد	تعداد رکورد	میانگین	حداقل	حداکثر	انحراف معیار	ضریب تغییرات	چولگی
Q(t)	(m ³ /s)	۵۲۸	۱۱/۶۵۰	۰/۶۷۵	۱۲۵/۹۶۷	۱۵/۴۵۱	۱/۳۲۶	۲/۶۹۸
P(t)	(mm)	۵۲۸	۴۴/۹۴۹	۰	۲۵۳	۵۳/۶۹۲	۱/۱۹۵	۱/۱۵۷

شبکه‌های عصبی بیزین

این روش یکی از روش‌های سیستم پشتیبان تصمیم‌گیری می‌باشد که ابزار قدرتمندی در مدل کردن روابط علی و معلولی در قالب شبکه‌ای از احتمالات است. نکته بسیار مهم در مورد روش

در شکل (۲) نمودار تغییرات بارش و رواناب و در جدول (۱) خصوصیات آماری پارامترهای کمی رودخانه کاکارضا در بازه زمانی مذکور نشان داده شده است.

قربانی و دهقانی: کاربرد شبکه‌های عصبی بیزین، ماشین بردار...

$$P(w|D) = \frac{1}{Z_s} \exp(-\beta \sum_n \sum_k^K \{t_{nk} - y_{nk}\}^2 - \frac{\alpha}{2} \sum_j^W w_j^2) \quad (3)$$

$$Z_s(\alpha|\beta) = \int \exp(-\beta \sum_n \sum_k^K \{t_{nk} - y_{nk}\}^2 - \frac{\alpha}{2} \sum_j^W w_j^2) dw = \left(\frac{2\pi}{\beta}\right)^{N/2} \left(\frac{2\pi}{\alpha}\right)^{W/2} \quad (4)$$

اولین ضابطه در تابع احتمال، توان تابع و دومین ضابطه اطلاعات اولیه است. n : شاخص برای الگوی آموزشی، β : داده خطا دار و K : شاخص برای واحد خروجی، α : ضریب اطلاعات اولیه

است. آموزش شبکه با استفاده از الگوریتم بیزین به صورت خودکار و همراه با پیچیدگی بالا است (شیلیدزی، ۲۰۰۷). نمونه ایی از ساختار شبکه سه لایه‌ای متشکل از یک لایه ورودی، یک لایه مخفی و یک لایه خروجی در شکل (۳) نشان داده شده است.

برنامه‌ریزی بیان ژن

روش برنامه‌ریزی بیان ژن، در سال ۱۹۹۹ توسط فریرا ارائه شد (فریرا، ۲۰۰۱). این روش ترکیبی از روش‌های برنامه‌ریزی ژنتیک و الگوریتم ژنتیک بوده که در آن، کروموزوم‌های خطی و ساده با طول ثابت، مشابه با آنچه که در الگوریتم ژنتیک استفاده می‌شود و ساختارهای شاخه‌ای با اندازه‌ها و شکل‌های متفاوت، مشابه با درختان تجزیه در برنامه‌ریزی ژنتیک، ترکیب می‌شوند. از آنجایی که در این روش تمام ساختارهای شاخه‌ای با اندازه شکل‌های متفاوت، در کروموزوم‌های خطی با طول ثابت کدگذاری می‌شوند، معادل این است که در این روش فنوتیپ و ژنوتیپ از هم جدا می‌شوند و سیستم می‌تواند از تمام مزایای تکاملی به سبب وجود آنها بهره‌مند شود. اکنون با وجود اینکه فنوتیپ در برنامه‌ریزی بیان ژن، همان نوع از ساختارهای شاخه‌ای مورد استفاده در برنامه‌ریزی ژنتیک را شامل می‌شود، اما ساختارهای شاخه‌ای که به وسیله برنامه‌ریزی بیان ژن استنتاج می‌شوند (که بیان درختی نیز نامیده می‌شود) مبین تمامی ژنوم‌های مستقل هستند. به طور خلاصه می‌توان گفت در برنامه‌ریزی بیان ژن به سازی‌ها در یک ساختار خطی افتاده و سپس به صورت ساختار درختی بیان می‌شود و این موجب می‌شود تنها ژنوم اصلاح شده به نسل بعد منتقل شده و نیازی به ساختارهای سنگین برای تکثیر و جهش وجود نداشته باشد (فریرا، ۲۰۰۱). در این روش پدیده‌های مختلف با استفاده از مجموعه‌ای از توابع و مجموعه‌ای از ترمینال‌ها، مدل سازی می‌شوند. مجموعه توابع، معمولاً شامل توابع اصلی حسابی $\{+, -, \times, /\}$ ، توابع مثلثاتی یا هر نوع تابع ریاضی دیگر $\{\sqrt{\quad}, x^2, \exp, \log, \sin, \cos, \dots\}$ و یا توابع تعریف شده توسط کاربر است که معتقد است، می‌توانند برای تفسیر مدل مناسب باشند. مجموعه ترمینال‌ها، از مقادیر ثابت و متغیرهای مستقل مسأله تشکیل شده‌اند (فریرا، ۲۰۰۱). برای به کارگیری روش برنامه‌ریزی بیان ژن از نرم افزار GenXproTools 4.0 استفاده گردید.

شبکه عصبی بیزین این است که این روش به اطلاعات دقیق و تاریخچه کامل یک واقعیت نیاز ندارد بلکه می‌تواند با استفاده از اطلاعات ناقص و غیردقیق نیز به نتایج بسیار قانع کننده‌ای در زمینه تخمین وضعیت فعلی یا آینده یک سیستم دست یابد (طباطبایی و دشتی زاده، ۱۳۸۷). همچنین یک روش سازگار و انعطاف پذیر برای مدل کردن موقعیت‌های غیرقطعی و یک مدل گرافیکی مبتنی بر ادراک مستقیم از اندرکنش میان علل و معلول‌های مختلف ارائه می‌نماید. یکی دیگر از روش‌های دستیابی به پارامترهای بهینه در شبکه‌های عصبی، استفاده از الگوریتم بیزین^۱ است که به طور خودکار مقادیری مناسب برای پارامترهای تابع قرار می‌دهد. مزیت این الگوریتم در این است که اندازه شبکه بر روی نتایج آن تأثیر اندکی خواهد داشت. در این قالب فرض می‌شود که وزن‌ها و بایاس‌ها، متغیرهایی تصادفی با توزیع مشخص بوده، و نیز پارامترهای احتمالی مشخص، با داده و مدل ساختار بندی شده باشند. بر این اساس پارامترهای مشخص شده از داده، به صورت شبکه عصبی با گسترش به جلو^۲ هستند. یکی از انواع ویژه شبکه‌های عصبی، شبکه عصبی پرسپترون چند لایه است که گراف‌های ساخته شده با فرضیه‌های احتمالی در مورد داده‌ها را پارامتر گذاری می‌کند. شبکه عصبی پیاده سازی شده در اینجا شامل گره‌های مخفی و لایه خروجی با در نظر گرفتن یک لایه مخفی می‌باشد. روابط مابین خروجی (y) و ورودی (x) به صورت زیر تعریف می‌گردد:

$$y_k = f_{outer} \left(\sum_{j=1}^M w_{kj}^{(2)} f_{inner} \left(\sum_{i=1}^d w_{ji}^{(1)} + w_{j0}^{(1)} \right) + w_{k0}^{(1)} \right) \quad (1)$$

$w_{ji}^{(2)}$ و $w_{ji}^{(1)}$: وزن‌ها در لایه اول و لایه دوم به ترتیب با ورودی I و لایه مخفی Z می‌باشد و $w_{j0}^{(1)}$: بایاس برای واحد مخفی Z است. M : تعداد واحدهای مخفی، d : تعداد واحد ورودی و K : شاخصی برای واحد خروجی است. تابع $f_{outer}(\cdot)$ خطی است و $f_{inner}(\cdot)$ تابع تانژانت هایپربولیک می‌باشد. در روش بیزین برای توزیع وزن‌ها از معادله (۱) استفاده شده و برای مشخص کردن میزان وزن‌های شبکه با قرار دادن در فرمول بیزین که به صورت معادله زیر می‌باشد، محاسبه می‌گردد:

$$P(w|D) = \frac{P(D|w)P(w)}{P(D)} \quad (2)$$

$P(w)$: تابع توزیع احتمال در فضای وزنی با فقدان داده است که به صورت تابع توزیع اولیه می‌باشد. همچنین ماتریس $D = (y_1, \dots, y_N)$ شامل داده خروجی است. کمیت $P(w|D)$: تابع احتمالی وزن‌هاست که به عنوان تابع توزیع احتمالی بعد از آموزش داده‌ها دیده می‌شود. $P(D|w)$: تابع توزیع احتمال و $P(D)$: تابع توزیع احتمال ثانویه است که به صورت معادله زیر می‌باشد: (شیلیدزی، ۲۰۰۷).

- 1-Bayesian
- 2-feed-forward
- 3-Tshildizi

جدول ۲- همبستگی متقابل بین پارامترهای ورودی و خروجی

P(t-3)	P(t-2)	P(t-1)	P(t)	Q(t-1)	Q(t)
۰/۱۴۰	۰/۲۲۰	۰/۱۶۳	۰/۱۰۰	۰/۵۵۰	Q(t)

می‌توان از توابع مختلف کرنل برای ساخت انواع مختلف کرنل‌های ماشین بردار پشتیبان استفاده کرد. انواع توابع کرنل قابل استفاده در مدل ماشین بردار پشتیبان رگرسیونی عبارتند از: کرنل چندجمله‌ای^۱ با سه مشخصه هدف و کرنل توابع پایه شعاعی^۲ با یک مشخصه هدف، طبق به ترتیب طبق روابط زیر محاسبه می‌گردند.

$$k(x_i, x_j) = (x_i \cdot x_j)^d \quad (11)$$

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \quad (12)$$

در مطالعات بارش- رواناب عمدتاً از مدل RBF استفاده می‌شود (اسکندری و نوری، ۱۳۸۹). برای انجام فرآیند محاسبات کدی در محیط متلب نوشته شده، که به راحتی در اختیار سایر محققین قرار می‌گیرد.

معیارهای ارزیابی

در این تحقیق به منظور ارزیابی دقت و کارایی مدل‌ها، از نمایه‌های ضریب همبستگی (R)، ریشه میانگین مربعات خطا (RMSE)، ضریب نش ساتکلیف (NS) و بایاس (Bias) طبق روابط زیر استفاده گردید. بهترین مقدار برای این سه معیار به ترتیب یک، صفر، یک و صفر می‌باشد.

$$R = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}} \quad -1 \leq R \leq 1 \quad (13)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2} \quad (14)$$

$$NS = 1 - \frac{\sum_{i=1}^N (x_i - y_i)^2}{\sum_{i=1}^N (x_i - \bar{y})^2} \quad -\infty \leq NS \leq 1 \quad (15)$$

$$Bias = \frac{1}{N} \sum_{i=1}^N (x_i - y_i) \quad (16)$$

در روابط بالا، x_i و y_i به ترتیب مقادیر مشاهداتی و محاسباتی در گام زمانی i ام، N : تعداد گام‌های زمانی، \bar{x} و \bar{y} : نیز به ترتیب میانگین مقادیر مشاهداتی و محاسباتی می‌باشد. علاوه بر معیارهای فوق از نمودارهای پراکنش و سری زمانی مقادیر مشاهداتی- محاسباتی نسبت به زمان نیز برای مقایسه و تحلیل بیشتر نتایج استفاده شده است.

ماشین بردار پشتیبان

ماشین بردار پشتیبان یک سیستم یادگیری کارآمد بر مبنای تئوری بهینه‌سازی مقید است که از اصل استقرای کمینه‌سازی خطای ساختاری استفاده کرده و منجر به یک جواب بهینه کلی می‌گردد. در مدل رگرسیون ماشین بردار پشتیبان تابعی مرتبط با متغیر وابسته Y که خود تابعی از چند متغیر مستقل x است، برآورد می‌شود. مشابه سایر مسائل رگرسیونی فرض می‌شود رابطه میان متغیرهای مستقل و وابسته با تابع جبری مانند $f(x)$ به علاوه مقداری اغتشاش (خطای مجاز (ε)) مشخص شود:

$$f(x) = W^T \cdot \phi(x) + b \quad (5)$$

$$y = f(x) + \text{noise} \quad (6)$$

چنانچه W بردار ضرایب و b ثابت مشخصه‌های تابع رگرسیونی و ϕ نیز تابع کرنل باشد، آنگاه هدف پیدا کردن فرم تابعی برای $f(x)$ است. این مهم با آموزش مدل ماشین بردار پشتیبان توسط مجموعه‌ای از نمونه‌ها (مجموعه آموزش) محقق می‌شود. برای محاسبه w و b لازم است تابع خطا (معادله ۷) در مدل ماشین بردار پشتیبان با در نظر گرفتن شرایط مندرج (قیود) در معادله (۸) بهینه شود:

$$W^T \cdot \phi(X_i) + b - y_i \leq \varepsilon + \varepsilon_i^*, \quad \frac{1}{2} W^T \cdot W + C \sum_{i=1}^N \varepsilon_i + C \sum_{i=1}^N \varepsilon_i^* \quad (7)$$

$$y_i - W^T \cdot \phi(X_i) - b \leq \varepsilon + \varepsilon_i, \quad \varepsilon_i, \varepsilon_i^* \geq 0, \quad i=1, 2, \dots, N \quad (8)$$

در معادله‌های بالا C : عددی صحیح و مثبت است، که عامل تعیین جرمیه در هنگام رخ دادن خطای آموزش مدل است. ϕ : تابع کرنل، N : تعداد نمونه‌ها و دو مشخصه ε_i و ε_i^* : متغیرهای کمبود هستند. در نهایت تابع ماشین بردار پشتیبان رگرسیونی را می‌توان به شکل زیر بازنویسی کرد:

$$f(x) = \sum_{i=1}^N \bar{\alpha}_i \phi(x_i)^T \cdot \phi(x) + b \quad (9)$$

در معادله (۹) $\bar{\alpha}_i$ میانگین ضرایب لاگرانژ می‌باشد. محاسبه $\phi(x)$ در فضای مشخصه آن ممکن است بسیار پیچیده باشد. برای حل این مشکل روند معمول در مدل ماشین بردار پشتیبان رگرسیونی انتخاب یک تابع کرنل به صورت رابطه زیر است.

$$K(X_i, X) = \phi(X_i)^T \sqrt{b^2 - 4ac} \quad (10)$$

1- Polynomial

2- Radial Basis Functions (RBF)

قربانی و دهقانی: کاربرد شبکه‌های عصبی بیزین، ماشین بردار...

جدول ۳- ترکیب‌های منتخب پارامترهای ورودی مدل‌های شبکه عصبی بیزین، برنامه‌ریزی

شماره مدل	ساختار ورودی	خروجی
۱	P(t)	O(t)
۲	P(t)P(t-1)	Q(t)
۳	P(t) P(t-1)P(t-2)	Q(t)
۴	P(t) P(t-1)P(t-2) P(t-3)	Q(t)
۵	Q(t-1)P(t)	Q(t)
۶	Q(t-1)P(t) P(t-1)	Q(t)
۷	O(t-1) P(t) P(t-1)P(t-2)	O(t)
۸	Q(t-1) P(t) P(t-1)P(t-2) P(t-3)	Q(t)

جدول ۴- مقادیر پارامترهای مورد استفاده با استفاده از شبکه عصبی بیزین

تنظیمات کلی	تنظیمات شبکه بیزین
پارامتر تنظیمات لونبرگ	۰/۰۰۵ β
ضریب کاهش ماتریس ورودی	۰/۱ K
ضریب افزایش ماتریس ورودی	۱۰ γ
حداکثر عدم موفقیت در صحت سنجی	۵
حداقل شیب عملکرد	۱۰ ^{-۱۰}
تعداد دور	۱۰۰۰

جدول ۵- نتایج شبکه عصبی بیزین برای مدل‌های مختلف و شاخص‌های آماری در

مراحل آموزش و صحت سنجی

صحت سنجی			آموزش			
مدل	ساختار	ضریب همبستگی	میانگین مربعات خطا (m ³ /s)	ریشه میانگین مربعات خطا (m ³ /s)	نش	بایاس
یک	۱-۹-۱	۰/۷۱۵	۰/۱۴۷	۰/۵۱۴	۰/۷۳۲	-۰/۰۰۴
دو	۲-۱۲-۱	۰/۷۵۸	۰/۱۳۰	۰/۶۰۲	۰/۷۱۲	-۰/۰۰۱
سه	۳-۵-۱	۰/۸۶۲	۰/۱۲۱	۰/۷۲۴	۰/۸۴۰	-۰/۰۰۱
چهار	۴-۷-۱	۰/۸۷۰	۰/۱۰۷	۰/۷۵۰	۰/۸۷۲	-۰/۰۰۶
پنج	۲-۴-۱	۰/۸۸۷	۰/۰۷۸	۰/۷۹۹	۰/۸۷۹	-۰/۰۰۴
شش	۳-۵-۱	۰/۹۰۲	۰/۰۸۴	۰/۸۲۷	۰/۹۴۷	-۰/۰۰۳
هفت	۴-۱۰-۱	۰/۹۲۷	۰/۰۴۲	۰/۸۰۶	۰/۹۶۲	-۰/۰۰۱
هشت	۵-۸-۱	۰/۹۴۵	۰/۰۵۵	۰/۸۱۲	۰/۹۹۰	-۰/۰۰۱

نتایج و بحث

خروجی مدل در نظر گرفته شده است. با توجه به همبستگی متقابل و معنی‌دار بین داده‌های ورودی و خروجی، ترکیب‌های مختلفی از پارامترهای ورودی به منظور دستیابی به مدل بهینه برای تخمین میزان رواناب ماهانه رودخانه کاکارضا استفاده شده که در جدول (۳) آمده است.

برای تخمین میزان رواناب رودخانه کاکارضا با استفاده از روش‌های شبکه عصبی بیزین، برنامه‌ریزی بیان ژن و ماشین بردار پشتیبان با در دست داشتن داده‌های ایستگاه‌های هواشناسی و هیدرومتری حوضه از ۵۲۸ رکورد ثبت شده که در شرکت آب منطقه‌ای و سازمان هواشناسی استان لرستان در دسترس بود، طی بازه زمانی (۱۳۹۲-۱۳۴۸)، تعداد ۴۲۳ رکورد برای آموزش و تعداد

یکی از مهم‌ترین مراحل در مدل‌سازی، انتخاب ترکیب مناسبی از متغیرهای ورودی است. از این رو ابتدا همبستگی متقابل^۱ بین متغیرهای ورودی و خروجی محاسبه گردید و پارامترهای ورودی به منظور دستیابی به مدل بهینه برای تخمین میزان رواناب ماهانه رودخانه کاکارضا انتخاب و در جدول (۲) ارائه شده است.

در این جدول P(t-2)، P(t-1)، P(t)، Q(t-1) (به ترتیب بارش در زمان t-2، t-1، t و رواناب در زمان t-1 می‌باشند) به عنوان ورودی و Q(t) (رواناب در زمان t) به عنوان

1- Cross-correlation

(۲۰۱۳) و خان و کولیالی^۴ (۲۰۰۶) مطابقت دارد. در تبیین این نتایج می‌توان گفت در شبکه‌های عصبی بیزین با اختصاص وزن مناسب به هر یک از پارامترهای ورودی طبق قانون بیز، باعث کاهش میزان خطای شبکه می‌شود. همچنین استفاده از تقریب گوس نیوتن برای محاسبه پارامترهای تنظیم، سبب تسریع در روند همگرایی شبکه، بهینه نمودن وزن‌ها و کاهش خطای شبکه می‌شود.

نتایج برنامه‌ریزی بیان ژن

استفاده از برنامه‌ریزی بیان ژن به دلیل توان انتخاب متغیرهای مؤثر در مدل و حذف متغیرهایی با تأثیر کمتر و همچنین توانایی ارائه رابطه صریح برای تخمین رواناب رودخانه کاکارضا مدنظر قرار گرفت. لذا هر پنج ورودی برای تعیین متغیرهای معنی‌دار استفاده، و برای بررسی بیشتر علاوه بر مجموعه چهار عملگر اصلی (F1)، حالت‌هایی بر اساس عملگرهای ریاضی پیش‌فرض برنامه (F2) لحاظ گردیده است. نحوه و دلیل انتخاب این نوع عملگرها با توجه به مطالعات (قربانی و همکاران، ۲۰۱۲) صورت گرفته است. پارامترهای مورد استفاده و نرخ آنها در استخراج مدل‌های تخمین رواناب با استفاده از روش برنامه‌ریزی بیان ژن به‌طور خلاصه در جدول (۶) ارائه گردیده است.

$$F1: \{+, -, *, /\} \quad (18)$$

$$F2: \{+, -, *, /, \sqrt{\quad}, \text{Exp}, \text{Ln}, ^2, ^3, \sqrt[3]{\quad}, \text{Sin}, \text{Cos}, \text{Atan}\} \quad (19)$$

نتایج حاصل از مدل برنامه‌ریزی بیان ژن با هر دو عملگر در جدول ۷ بیانگر این است که مدل ترکیبی شماره هشت با عملگر F2 در هر دو مرحله آموزش و صحت‌سنجی با بیشترین ضریب همبستگی ($R=0/802$) و ریشه میانگین مربعات خطا (m^3/s) $1/620$ ($RMSE=1/620$) و ($NS=0/724$) و بایاس $0/047$ نسبت به سایر مدل‌ها از دقت بیشتری برخوردار بوده است. بنابراین برنامه‌ریزی بیان ژن با عملگر F2 شامل چهار عملگر اصلی ریاضی، توانی، لگاریتمی و توابع مثلثاتی، بیشترین دقت را در تخمین رواناب در رودخانه کاکارضا داشته است. نمودارهای پراکنش برنامه‌ریزی بیان ژن مربوط به مرحله صحت‌سنجی در شکل (۵-ب) نشان‌دهنده خط برازش مقادیر محاسباتی با چهار مجموعه عملگر ریاضی، توانی، لگاریتمی و توابع مثلثاتی، با بهترین خط برازش $y=x$ است. همان‌طور که از این شکل مشخص است بیشتر مقادیر تخمینی و مشاهداتی به غیر از چند نقطه روی خط نیمساز قرار گرفته و این امر دلالت بر برابری مقادیر مشاهداتی و محاسباتی بر اساس خط $y=x$ می‌باشد. لازم به ذکر است با توجه به شکل (۵-الف) این مدل در تخمین مقادیر بیشینه رواناب ضعیف عمل کرده، به طوری که این مقادیر بیشتر از مقدار واقعی آنها تخمین زده است.

۱۰۵ رکورد باقی‌مانده برای صحت‌سنجی در نظر گرفته شد. یک نکته مهم در آموزش شبکه‌های عصبی نرمال‌سازی داده‌ها قبل از استفاده در مدل می‌باشد این عمل خصوصاً وقتی دامنه تغییرات ورودی‌ها زیاد باشد کمک شایانی به آموزش بهتر و سریع‌تر مدل می‌کند. اصولاً وارد کردن داده‌ها به‌صورت خام باعث کاهش سرعت و دقت شبکه می‌شود (ژو و همکاران، ۲۰۰۷). برای نرمال‌سازی داده‌های تحقیق از رابطه ۱۷ استفاده شده است:

$$X_n = 0.1 + 0.8 \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (17)$$

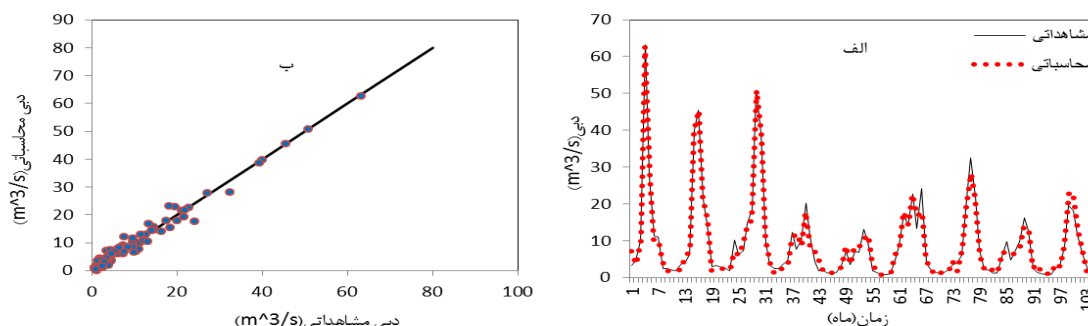
که در آن X_n مقدار نرمال شده ورودی X_i ، X_{\max} و X_{\min} به ترتیب حداکثر و حداقل داده‌ها می‌باشد.

نتایج شبکه عصبی بیزین

به‌منظور تخمین جریان ورودی به سد علویان از روش شبکه عصبی بیزین از نوع شبکه پرسپترون چندلایه با تعداد نرون‌های متفاوت استفاده شد. در این شبکه پارامترهای تابع هدف (پارامترهای تنظیم) با استفاده از تقریب گوس نیوتن محاسبه و نتایج آن در جدول (۴) ارائه شد. با توجه به جدول (۴) خطای شبکه کاهش یافته و به پارامترهای ورودی به شبکه وزن مناسب اختصاص یافت با بهینه نمودن پارامترهای تابع هدف و تأثیر آن در تابع هدف از بیش‌برازشی آموزش شبکه جلوگیری شد (مک کی^۵، ۱۹۹۲). روند کار آموزش شبکه با تعداد نرون‌های کم آغاز و افزودن نرون‌های اضافی تا زمانی ادامه می‌یابد که افزایش نرون‌های بیشتر تأثیری در بهبود خطا نداشته باشد، در این روش با توجه به جدول (۵) مدل ترکیبی شماره هشت به‌صورت شبکه‌ای با چهار نرون در لایه پنهان اول به‌عنوان بهترین مدل در فرآیند شبیه‌سازی حاصل گردید. در جدول (۵) نتایج حاصل از مقایسه عملکرد ساختارهای مختلف به کار رفته به همراه پارامترهای آماری داده‌های بخش آموزش و صحت‌سنجی ارائه شده است. همان‌طور که در جدول (۵) مشاهده می‌شود مدل ترکیبی شماره هشت با بیشترین ضریب همبستگی ($R=0/990$) و کمترین ریشه میانگین مربعات خطا ($RMSE=0/041 m^3/s$) و ($NS=0/826$) و بایاس $0/002$ در مرحله صحت‌سنجی نسبت به سایر مدل‌ها برای شبیه‌سازی فرآیند بارش - رواناب مناسب‌تر خواهد بود. همان‌طور که در شکل (۴-ب) مشخص است تطابق مقادیر محاسباتی رواناب مربوط به مرحله صحت‌سنجی مدل شبکه عصبی بیزین با مقادیر مشاهداتی وجود دارد. در این شکل اختلاف ناچیز برخی از مقادیر با بهترین خط برازش $y=x$ قابل مشاهده می‌باشد. با توجه به نمودار (۴-الف) مقادیر مشاهداتی و محاسباتی نیز قابلیت بالای این مدل در تخمین رواناب رودخانه کاکارضا مشهود است. این نتایج با پژوهش‌های فیتیری ساری و همکاران^۳

- 1 - Xu et al.
- 2- MacKay
- 3 - Fithriasari
- 4- Khan and Coulibaly

قربانی و دهقانی: کاربرد شبکه‌های عصبی بیزین، ماشین بردار...



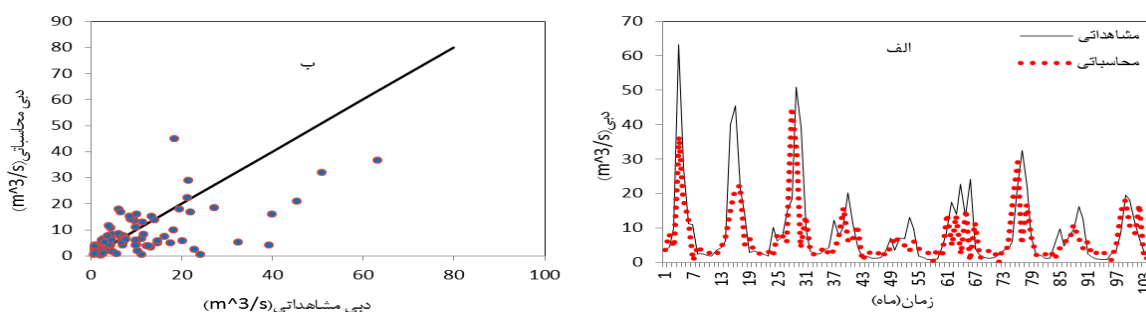
شکل ۴- نمودار حاصل از مقادیر بهینه مدل شبکه عصبی بیزین برای داده‌های ثبت شده مرحله صحت سنجی

ب) نمودار پراکنش میان مقادیر مشاهداتی و محاسباتی

الف) مقادیر محاسباتی و مشاهداتی نسبت به زمان

جدول ۶- مقادیر پارامترهای مورد استفاده با استفاده از روش برنامه‌ریزی بیان ژن

تنظیمات کلی	عملگرهای ژنتیکی	
تعداد کروموزوم‌ها	نرخ جهش	۰/۰۴۴
اندازه رأس	نرخ وارون سازی	۰/۱
تعداد ژن‌ها در هر کروموزوم	نرخ ترانسپوز درج متوالی	۰/۱
تعداد جمعیت تولیدی	نرخ ترانسپوز ریشه درج متوالی	۰/۱
تابع پیوند	نرخ ترکیب تک نقطه‌ای	۰/۳
	جمع	(+)



شکل ۵- نمودار حاصل از مقادیر بهینه مدل برنامه‌ریزی بیان ژن برای داده‌های ثبت شده مرحله صحت سنجی

ب) نمودار پراکنش میان مقادیر مشاهداتی و محاسباتی

الف) مقادیر محاسباتی و مشاهداتی نسبت به زمان

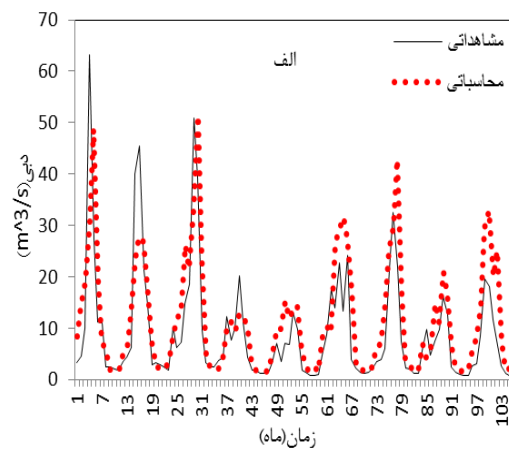
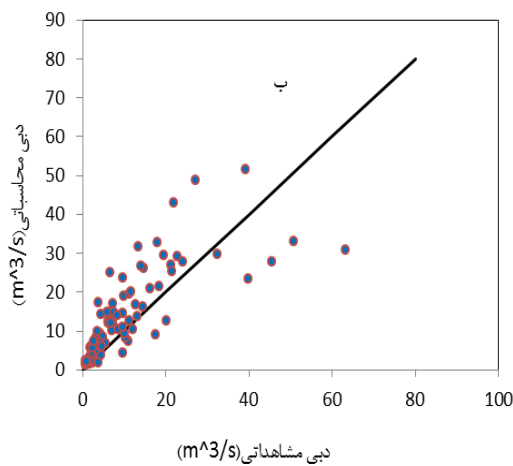
مقادیر تخمینی و مشاهداتی به غیر از یک نقطه روی خط نیمساز قرار گرفته و این امر دلالت بر برابری مقادیر مشاهداتی و محاسباتی بر اساس خط $(y=x)$ می‌باشد. در شکل (۶-الف) نیز تغییرات مقادیر محاسباتی و مشاهداتی نسبت به زمان نشان داده شده است، به طوری که این مدل در تخمین مقادیر بیشینه در شرایط مختلف زمانی دقت کافی را نداشته است که این نتایج با نتایج اسکندری و نوری (۱۳۸۹) مطابقت دارد که می‌توان بیان نمود با توجه به اینکه ماشین بردار پشتیبان بر مبنای استفاده از اصل استقرای کمینه‌سازی خطای ساختاری است بنابراین در شبیه‌سازی، با به کارگیری روش یادگیری با نظارت در توابع پایه شعاعی باعث می‌شود تخمین پارامتر از سرعت بالا و خطای کمتری نسبت به سایر کرنل‌ها داشته باشد و این یکی از ویژگی‌های ممتاز توابع پایه شعاعی محسوب می‌شود.

نتایج ماشین بردار پشتیبان

به منظور شبیه‌سازی فرآیند بارش-رواناب توسط مدل ماشین بردار پشتیبان می‌توان انواع تابع کرنل را مورد بررسی قرارداد، که از بین توابع کرنل، کرنل‌های چندجمله‌ای و توابع پایه شعاعی از انواع رایج مورد استفاده در هیدرولوژی می‌باشند، انتخاب گردید. نتایج حاصل بررسی مدل‌ها در جدول (۸) آورده شده است. با توجه به جدول مذکور مدل ترکیبی شماره هشت با کرنل توابع پایه شعاعی دارای بیشترین میزان ضریب همبستگی ($R=0/890$) و کمترین میزان ریشه مربعات خطا ($RMSE=0/952 \text{ m}^3/\text{s}$) و ($0/706$) و بایاس $-0/008$ در مرحله صحت سنجی جواب بهینه‌تری نسبت به سایر مدل‌ها را به همراه داشته است. در شکل ۶ نمودار بهترین مدل حاصل شده برای داده‌های بخش صحت سنجی نشان داده شده است. همان‌طور که از شکل (۶-ب) مشخص است بیشتر

جدول ۷- نتایج مدل برنامه‌ریزی بیان ژن برای مدل‌های مختلف با استفاده از دو مجموعه عملگر ریاضی منتخب

صحت سنجی				آموزش				عملگر	مدل
بایاس	نش	ریشه میانگین مربعات خطا (m^3/s)	ضریب همبستگی	بایاس	نش	ریشه میانگین مربعات خطا (m^3/s)	ضریب همبستگی		
۰/۵۵۱	۰/۶۵۱	۴/۴۲۶	۰/۶۰۱	۱/۰۵۱	۰/۳۰۶	۸/۵۳۴	۰/۶۴۸	F1	یک
۰/۷۱۲	۰/۶۱۲	۴/۵۱۴	۰/۶۱۰	۰/۷۴۱	۰/۳۲۵	۶/۱۴۴	۰/۶۴۲	F2	
-۱/۴۱۸	۰/۶۵۲	۲/۴۲۸	۰/۶۱۷	۲/۱۵۶	۰/۳۴۴	۶/۳۷۳	۰/۶۶۷	F1	دو
۰/۰۶۴	۰/۶۲۷	۳/۲۷۰	۰/۶۲۲	-۰/۴۶۴	۰/۳۵۴	۵/۳۹۹	۰/۶۵۸	F2	
-۰/۷۲۲	۰/۶۳۷	۳/۳۴۷	۰/۶۲۸	۰/۲۸۵	۰/۴۰۸	۷/۵۴۷	۰/۶۷۴	F1	سه
-۱/۰۴۱	۰/۶۴۸	۲/۳۹۱	۰/۶۳۰	-۰/۵۲۸	۰/۳۹۲	۷/۶۸۰	۰/۶۷۱	F2	
۰/۳۲۵	۰/۶۴۴	۲/۷۵۸	۰/۶۳۲	-۰/۶۸۰	۰/۴۱۹	۶/۷۲۴	۰/۶۸۶	F1	چهار
۱/۷۵۴	۰/۶۳۶	۳/۸۴۲	۰/۶۵۱	-۱/۷۴۴	۰/۳۸۵	۵/۶۲۴	۰/۶۸۱	F2	
۰/۷۵۱	۰/۶۵۲	۳/۲۴۰	۰/۶۵۸	۰/۱۷۸	۰/۴۲۴	۸/۸۷۲	۰/۶۹۱	F1	پنج
-۰/۰۸۷	۰/۶۵۷	۴/۱۰۰	۰/۶۶۴	۰/۶۱۰	۰/۵۶۰	۴/۳۱۴	۰/۶۸۷	F2	
۰/۸۱۲	۰/۶۶۹	۴/۸۱۱	۰/۶۷۵	۱/۲۳۴	۰/۴۸۲	۳/۶۵۲	۰/۷۶۰	F1	شش
۱/۴۶۰	۰/۶۷۴	۲/۷۵۳	۰/۶۸۱	۰/۶۴۲	۰/۴۰۱	۲/۴۱۲	۰/۶۸۹	F2	
۱/۲۴۵	۰/۶۸۱	۲/۸۷۳	۰/۶۸۹	۰/۷۲۸	۰/۵۶۳	۱/۷۴۲	۰/۷۱۸	F1	هفت
-۱/۱۸۷	۰/۶۸۷	۱/۹۱۲	۰/۷۵۰	-۱/۵۱۶	۰/۵۷۵	۱/۷۵۴	۰/۷۳۸	F2	
-۱/۹۲۴	۰/۷۰۵	۱/۹۷۵	۰/۷۵۱	۱/۸۱۷	۰/۵۸۱	۱/۶۱۷	۰/۷۴۲	F1	هشت
۰/۰۴۷	۰/۷۲۴	۱/۶۲۰	۰/۸۰۲	۰/۱۵۲	۰/۶۸۲	۱/۵۲۷	۰/۷۶۵	F2	



شکل ۶- نمودار حاصل از مقادیر بهینه مدل ماشین بردار پشتیبان برای داده‌های ثبت شده مرحله صحت سنجی

ب) نمودار پراکنش میان مقادیر مشاهداتی و محاسباتی

الف) مقادیر محاسباتی و مشاهداتی نسبت به زمان

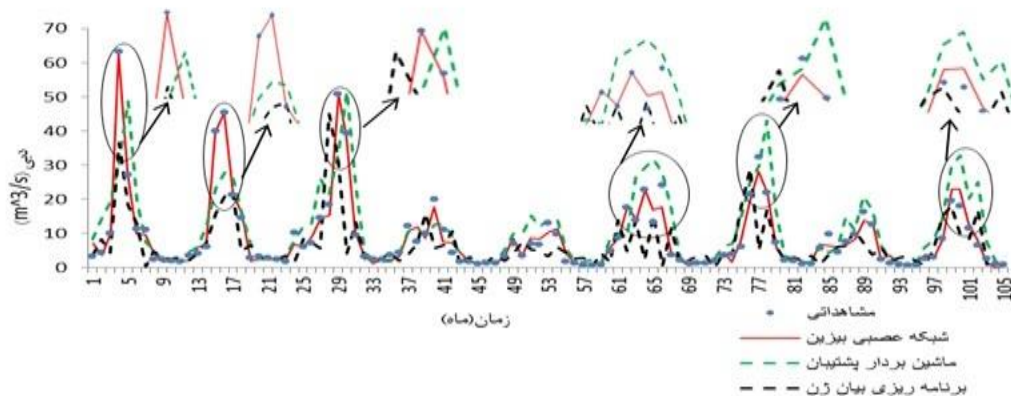
قربانی و دهقانی: کاربرد شبکه‌های عصبی بیزین، ماشین بردار...

جدول ۸- نتایج دو کرنل استفاده شده در روش ماشین بردار پشتیبان برای مدل‌های مختلف

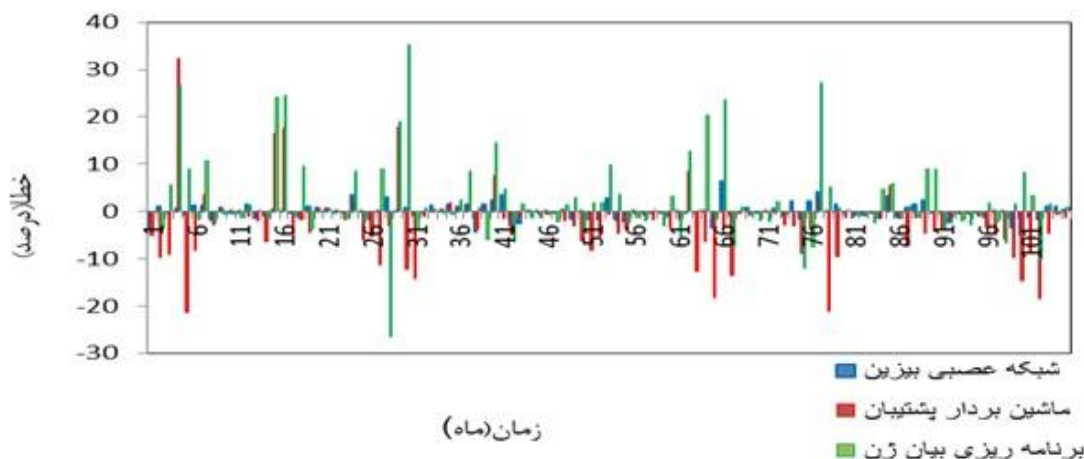
صحت سنجی			آموزش			کرنل	مدل		
بایاس	نش ساتکلیف	ریشه میانگین مربعات خطا (m ³ /s)	ضریب همبستگی	بایاس	نش ساتکلیف			ریشه میانگین مربعات خطا (m ³ /s)	ضریب همبستگی
-۰/۰۵۸	۰/۵۳۲	۲/۶۰۲	۰/۷۰۶	۰/۰۱۴	۰/۶۳۲	۳/۸۴۵	۰/۷۳۶	توابع پایه شعاعی	۱
۰/۰۴۴	۰/۵۶۰	۲/۷۸۱	۰/۷۲۰	-۰/۰۰۱	۰/۶۵۹	۳/۷۵۲	۰/۷۲۸	چندجمله‌ای	
۰/۰۳۲	۰/۵۷۵	۱/۴۱۸	۰/۷۱۴	۰/۰۷۱	۰/۶۴۷	۳/۴۶۶	۰/۷۱۲	توابع پایه شعاعی	۲
-۰/۰۵۲	۰/۵۶۲	۱/۰۴۲	۰/۷۴۴	-۰/۰۰۳	۰/۶۵۸	۳/۴۲۳	۰/۷۴۴	چندجمله‌ای	
-۰/۰۷۷	۰/۵۵۰	۲/۱۱۲	۰/۷۲۱	-۰/۰۰۴	۰/۶۸۷	۲/۸۴۱	۰/۷۶۱	توابع پایه شعاعی	۳
۰/۰۳۵	۰/۵۳۴	۲/۶۲۱	۰/۷۳۸	-۰/۰۰۲	۰/۶۶۶	۳/۴۵۲	۰/۷۵۸	چندجمله‌ای	
-۰/۰۵۱	۰/۵۷۱	۲/۰۵۷	۰/۷۶۹	-۰/۰۰۱	۰/۷۱۴	۳/۳۸۴	۰/۷۵۲	توابع پایه شعاعی	۴
۰/۰۹۷	۰/۵۹۳	۱/۶۵۵	۰/۷۸۸	۰/۰۰۴	۰/۷۰۱	۳/۱۵۰	۰/۷۶۴	چندجمله‌ای	
۰/۰۳۶	۰/۵۸۶	۱/۴۷۸	۰/۸۱۷	-۰/۰۰۷	۰/۷۳۷	۲/۳۱۸	۰/۷۸۵	توابع پایه شعاعی	۵
-۰/۰۴۱	۰/۶۳۱	۱/۵۸۰	۰/۸۵۱	۰/۰۰۵	۰/۷۵۹	۳/۲۵۴	۰/۷۸۰	چندجمله‌ای	
۰/۰۶۴	۰/۶۵۷	۱/۳۴۰	۰/۸۴۲	۰/۰۰۳	۰/۷۴۳	۱/۶۸۴	۰/۸۱۰	توابع پایه شعاعی	۶
-۰/۰۸۶	۰/۶۱۷	۱/۴۶۶	۰/۸۲۶	-۰/۰۰۷	۰/۷۶۸	۱/۳۲۸	۰/۸۰۱	چندجمله‌ای	
-۰/۰۵۸	۰/۶۴۲	۱/۴۵۲	۰/۸۵۴	-۰/۰۰۲	۰/۷۵۸	۱/۴۱۷	۰/۸۱۵	توابع پایه شعاعی	۷
-۰/۰۷۲	۰/۶۰۸	۱/۳۲۶	۰/۸۶۷	-۰/۰۰۱	۰/۷۶۴	۱/۶۷۱	۰/۸۲۸	چندجمله‌ای	
-۰/۰۰۸	۰/۷۰۶	۰/۹۵۲	۰/۸۹۰	۰/۰۰۶	۰/۷۹۹	۱/۲۴۱	۰/۸۵۲	توابع پایه شعاعی	۸
۰/۰۱۷	۰/۶۲۱	۱/۴۲۴	۰/۸۷۱	-۰/۰۰۵	۰/۷۸۰	۱/۳۱۲	۰/۸۴۱	چندجمله‌ای	

جدول ۹- نتایج نهایی حاصل از روش‌های شبکه عصبی بیزین، برنامه‌ریزی بیان ژن و ماشین بردار پشتیبان

مدل	آموزش				صحت سنجی			
	ضریب همبستگی	ریشه میانگین مربعات خطا (m3/s)	نش ساتکلیف	بایاس	ضریب همبستگی	ریشه میانگین مربعات خطا (m3/s)	نش ساتکلیف	بایاس
شبکه عصبی بیزین	۰/۹۴۵	۰/۰۵۵	۰/۸۱۲	-۰/۰۰۱	۰/۹۹۰	۰/۰۴۱	۰/۸۲۶	-۰/۰۰۲
برنامه‌ریزی بیان ژن	۰/۷۶۵	۱/۵۲۷	۰/۶۸۲	-۰/۱۵۲	۰/۸۰۲	۱/۶۲۰	۰/۷۲۴	-۰/۰۴۷
ماشین بردار پشتیبان	۰/۸۵۲	۱/۲۴۱	۰/۷۹۹	-۰/۰۰۶	۰/۸۹۰	۰/۹۵۲	۰/۷۰۶	-۰/۰۰۸



شکل ۷- نمودار پراکنش و مشاهداتی - محاسباتی مقادیر بهینه مدل شبکه عصبی بیزین، برنامه‌ریزی بیان ژن و ماشین بردار پشتیبان برای داده‌های ثبت شده مرحله صحت سنجی



شکل ۸- نمودار خطای بهینه هر سه مدل به صورت درصدی از میانگین مقادیر مشاهداتی

جدول ۱۰- مقادیر دبی حداکثر و خطا مدل شبکه عصبی بیزین، برنامه‌ریزی بیان ژن و ماشین بردار پشتیبان

شماره	مشاهداتی	مقادیر دبی حداکثر		خطاها(درصد)	
		شبکه	ماشین	برنامه‌ریزی	ماشین
یک	۶۳/۲۸۰	عصبی بیزین	بردار پشتیبان	۳۶/۵۱۴	بیان ژن
دو	۴۵/۴۴۸	۴۵/۴۱۶	۲۷/۸۳۲	۲۱/۶۴۵	۲۳/۸۰۳
سه	۵۰/۸۹۰	۵۰/۸۴۴	۳۳/۱۲۸	۳۲/۱۴۴	۱۸/۷۴۶
چهار	۲۲/۷۷۲	۲۲/۶۹۸	۲۹/۱۵۹	۲/۴۲۴	۲۰/۳۴۸
پنج	۳۲/۴۲۹	۲۸/۱۶۹	۲۹/۸۰۱	۵/۲۸۱	۲۷/۱۴۸
شش	۱۹/۵۹۴	۲۲/۹۴۵	۲۹/۴۲۶	۱۸/۱۵۲	۱/۴۴۲

مقایسه عملکرد مدل‌ها

در ادامه با انتخاب جواب بهینه هر کدام از روش‌های هوشمند و مقایسه آنها با یکدیگر مشخص شد، که هر سه روش با دقت خوبی می‌توانند فرآیند بارش - رواناب در حوضه کاکارضا را شبیه‌سازی کنند. همان طور که در جدول (۹) قابل مشاهده است از بین روش‌های به کار رفته، شبکه عصبی بیزین دارای بیشترین دقت ($R=0/990$) و کمترین ریشه میانگین مربعات خطا (m^3/s) ($RMSE=0/041$) و بیشترین ضریب نش-ساتکلیف

($NS=0/826$) و کمترین مقدار بایاس $0/002$ - در مرحله صحت سنجی می‌باشد. مقایسه نتایج روش ماشین بردار پشتیبان با روش برنامه‌ریزی بیان ژن حاکی از نزدیک بودن نتایج این دو روش می‌باشد. در شکل (۷) نتایج مدل انتخاب شده بارش-رواناب در هر سه روش نسبت به مقادیر مشاهداتی در طی زمان نشان داده شده است که هر سه روش در تخمین بیشتر مقادیر خوب عمل کرده، در حالی که روش شبکه عصبی بیزین مقادیر کمیته، بیشینه و میانی را به خوبی پوشش داده است. همان طور که در شکل (۷) قابل مشاهده است دو روش برنامه‌ریزی بیان ژن و ماشین بردار پشتیبان در تخمین مقادیر دبی حداکثر عملکرد قابل قبولی نداشته در حالی که روش شبکه عصبی بیزین در تخمین برخی مقادیر مذکور (دبی حداکثر) قابلیت بالایی داشته و نزدیک به مقدار واقعی تخمین زده است. در جدول ۱۰ نتایج حاصل از مقادیر حداکثر و خطاها هر سه روش نشان داده شده است.

در نهایت اختلاف مقادیر مشاهداتی و محاسباتی بهینه روش‌ها به صورت درصدی از میانگین مقادیر مشاهداتی (مقدار خطا)، محاسبه و نمودار آن نسبت به نمونه داده‌های ثبت شده دوره آماری ترسیم گردید (شکل ۸). همان طور که در این شکل دیده می‌شود برای هر سه مدل بیشتر خطاها در باند ± 10 در صد قرار گرفته است و بیشترین میزان خطای مدل شبکه عصبی بیزین، برنامه‌ریزی بیان ژن و ماشین بردار پشتیبان به ترتیب $6/426$ و $35/132$ و $32/331$ درصد از میانگین مقادیر مشاهداتی می‌باشد که در میان مدل‌های مذکور (مدل‌های استفاده شده)، مدل شبکه عصبی بیزین کمترین میزان خطا را داشته است. در مجموع با توجه به دقت و قابلیت تخمین بالای مدل‌ها، همبستگی بین مقادیر مشاهداتی و محاسباتی مدل‌های مذکور (شبکه عصبی بیزین،

برنامه‌ریزی بیان ژن و ماشین بردار پشتیبان) بررسی شده که به ترتیب $0/990$ ، $0/802$ و $0/890$ می‌باشد. همچنین نتایج معنی‌دار بودن مقادیر مشاهداتی و محاسباتی مدل‌های شبکه عصبی بیزین، برنامه‌ریزی بیان ژن و ماشین بردار پشتیبان در سطوح احتمال پنج و ده درصد نشان داد مدل شبکه عصبی بیزین همبستگی معنی‌دار در هر دو سطح احتمال دارد.

نتیجه گیری

در این تحقیق سعی بر آن شد عملکرد روش‌هایی برای شبیه‌سازی فرآیند بارش-رواناب در رودخانه کاکارضا با استفاده از داده‌های ماهانه این رودخانه مورد ارزیابی قرار گیرد. مدل‌های به کار گرفته شده شامل شبکه عصبی بیزین، برنامه‌ریزی بیان ژن و ماشین بردار پشتیبان می‌باشد. مقادیر رواناب مشاهداتی با رواناب تخمین زده شده در مدل‌های مذکور (شبکه عصبی بیزین، برنامه‌ریزی بیان ژن و ماشین بردار پشتیبان)، با استفاده از معیارهای ارزیابی مورد مقایسه قرار گرفت. نتایج تحقیق را می‌توان به صورت زیر خلاصه نمود:

مدل شبکه عصبی بیزین دقت بسیار بالا و خطای ناچیز در تخمین اکثر مقادیر از خود نشان داد و افزایش تعداد پارامترهای مؤثر در مدل‌های مختلف برای شبیه‌سازی فرآیند بارش - رواناب موجب عملکرد بهتر در تخمین رواناب است. همچنین الگوریتم شبکه بیزین در مقیاس زمانی ماهانه با طول دوره آماری کم نسبت به سایر مدل‌ها به خوبی تعمیم داده شده است. در مجموع نتایج تحقیق نشان داد که روش شبکه عصبی بیزین دارای دقت بالاتری نسبت به سایر مدل‌ها می‌باشد. همان طور که در نتایج تحقیقات فیتری ساری و همکاران^۱ (۲۰۱۳)، جیونیا و همکاران^۲ (۲۰۱۱)، خان و همکاران^۳ (۲۰۰۵)، معروفی و همکاران (۱۳۹۱) و اسکندری و نوری (۱۳۸۹) این امر به اثبات رسیده است. همچنین این تحقیق نشان داد که استفاده از مدل شبکه عصبی بیزین، برنامه‌ریزی بیان ژن و ماشین بردار پشتیبان می‌تواند در زمینه شبیه‌سازی فرآیند بارش - رواناب استفاده شود.

1- Fithriasari et al.

2- Gianola et al.

3- Khan and Coulibaly

منابع

- ۱- اسکندری، ع. و ر. نوری. ۱۳۸۹. توسعه مدلی مناسب بر مبنای شبکه عصبی مصنوعی و ماشین بردار پشتیبان برای پیش بینی به هنگام اکسیژن خواهی بیو شیمیایی ۵ روزه. نشریه محیط شناسی. ۳۸ (۶۱): ۷۶-۷۴.
- ۲-سلطانی، ع.، قربانی، م.، ع.، فاخری فرد، ا.، دربندی، ص. و د. فرسادی زاده. ۱۳۸۹. برنامه ریزی ژنتیک و کاربرد آن در مدل سازی فرآیند بارش- رواناب. مجله دانش آب و خاک. ۲۰(۴): ۶۲-۷۱.
- ۳-طباطبایی، س.، ع. و پ. دشتی زاده. ۱۳۸۷. معرفی شبکه های عصبی Bayesian و کاربرد آنها در نگهداری پل ها. چهاردهمین کنفرانس دانشجویان مهندسی عمران.
- ۴-قربانی، م.، ع. و ا. صالحی. ۱۳۹۰. استفاده از برنامه ریزی بیان ژن در بررسی تغییرات داده های کیفی آب زیرزمینی با نوسانات سطح آب در دشت برخورد اصفهان. ششمین کنگره ملی مهندسی عمران.
- ۵-معروفی، ص.، امیر مرادی، ک. و ن. پارسا فر. ۱۳۹۱. پیش بینی جریان روزانه با استفاده از شبکه عصبی مصنوعی و عصبی-موجکی (مطالعه موردی: رودخانه باراندوزچای). نشریه دانش آب و خاک. ۳۳(۳): ۱۰۳-۹۳.
- 6-Aytek, A., Asce, M. and M. Alp. 2008. An application of artificial intelligence for rainfall-runoff modeling. *Journal Earth System Science*, 117(2): 145-155.
- 7-Dibike, Y. B., Velikov, S., Solomatine, D. and M. B. Abbott. 2001. Model induction with of support vector machines: Introduction and applications. *Journal of Computing in Civil Engineering*, 15(1): 208- 216.
- 8-Ferreira, C. 2001. Gene expression programming: a new adaptive algorithm for solving problems. *Complex Systems*, 13(2): 87-129.
- 9-Fithiasari, K., Brodjol, N., Sutikno, U. and H. Kuswanto. 2013. Prediction of hourly rainfall using bayesian neural network with adjusting procedure. *The Third Basic Science International Conference*. Brawijaya Malang
- 10-Ghorbani, M. A., Khatibi, R., Asadi, H. and P. Yousefi. 2012. Inter- comparison of an evolutionary programming model of suspended sediment time-series whit other local model. *Journal of Hydrology*, 511(1): 1-16.
- 11-Gianola, D., Okut, H., Weigel, K. and G. Rosa. 2011. Predicting complex quantitative traits with bayesian neural networks: a case study with Jersey cows and wheat. *Biomedicine journals Genetics*. 24(3): 77-87.
- 12-Jayawardena, A. W., Muttill, N. and T. Fernando. 2005. Rainfall-runoff modelling using genetic programming.. *International Congress on Modelling and Simulation Society of Australia and New Zealand* December, New Zealand. 1841-1847.
- 13-Khan, M. S. and P. Coulibaly. 2006. Bayesian neural network for rainfall-runoff modeling. *Water Resources Research*, 42 (7): 1-18.
- 14-Liong, S. Y., Gautam, T. R., Khu, S. T., Babovic, V., Keijzer, M. and N. Muttill. 2002. Genetic programming: A new paradigm in rainfall runoff modeling. *Journal of the American Water Resources Association*, 38(3): 705-718.
- 15-MacKay, D. J. C. 1992. Bayesian Interpolation. *Neural Computation*, 4(3): 415-447.
- 16-Taheri, H. and M. Ghafouri. 2012. Comparison between active learning method and support vector machine for runoff modeling. *Journal of Hydrology and Hydromechanics*, 60(1): 16-32.
- 17- Tshilidzi, M. 2007. Bayesian training of neural networks using genetic programming. *Pattern Recognition Letters*, 28(12): 1452-1458.

قربانی و دهقانی: کاربرد شبکه‌های عصبی بیزین، ماشین بردار...

18-Xu, L., Wang, J., Guan, J., and F. Huang .2007. A support vector machine model for mapping of lake water quality from remote-sensed images. *International Journal of Intelligent Computing* , 1(1): 57-66.