



دوره ۳۵، شماره ۱، شماره ۱ پیاپی ۱۳۴، بهار ۱۴۰۱، صفحه‌های ۱۵-۲
شناسه دیجیتال: 10.22092WMRJ.2021.354279.1403

مقاله‌ی پژوهشی



پژوهش‌های آبخیزداری

پیش‌بینی آورد با روش‌های برنامه‌ریزی بیان ژن و وایازی درخت تصمیم با دیدگاه تغییر اقلیم در آبخیز کنگاور استان کرمانشاه

مریم حافظ‌پرست‌مودت

(نویسنده‌ی مسئول)* استادیار گروه علوم و مهندسی آب، دانشگاه رازی

فروزان پای‌فشرده

دانشجوی کارشناسی ارشد منابع آب، دانشگاه رازی

*رایانامه‌ی نویسنده‌ی مسئول: m.hafezparast@razi.ac.ir

تاریخ دریافت: ۱ اردیبهشت ۱۴۰۰ تاریخ پذیرش: ۳۰ خرداد ۱۴۰۰

چکیده

تغییر اقلیم، گرم‌تر شدن زمین، و مدیریت نامناسب اندوخته‌های آب از جمله دشواری‌هایی است که موجب نگرانی جامعه‌ی بشری به‌ویژه در مدیریت آبخیزها شده است. محققان و مهندسان برای جلوگیری از هدررفت، و بهینه‌سازی اندوخته‌های آب در این حوزه‌ها اقلیم را پیش‌بینی می‌کنند. در این پژوهش ایستگاه آران در آبخیز کنگاور برای ارزیابی تغییر بارش، دما، و آورد در آینده با در نظر گرفتن حالت‌های ممکن اقلیمی گزارش پنجم (CMIP5) در نظر گرفته شده است. با در دست داشتن اطلاعات مشاهده‌ی در ۳۳ سال از ۱۹۸۳ تا ۲۰۱۵ از روش‌های برنامه‌ریزی بیان ژن و درخت تصمیم برای آموزش، آزمون، و پیش‌بینی داده‌ها بهره گرفته شد. از میان مدل‌های گوناگون اقلیمی مدل Fgoals-g2 به دلیل مشابهت بیش‌تر ویژگی‌های آماری داده‌های تاریخی با داده‌های بارش و دمای منطقه برای پیش‌بینی در دوره‌ی ۲۰۲۰-۲۰۵۲ برگزیده شد. از حالت ممکن Rcp2.6 برای حالت خوش‌بینانه، و روش ریزمقیاس‌گردانی عامل‌های تغییر، و ورودی‌های مدل GEP و نرم‌افزار ORANGE برای پیش‌بینی ویژگی آورد در آینده بهره گرفته شد. پیشینه‌ی دمای ماهانه به $31/18^{\circ}\text{C}$ می‌رسد. پیشینه‌ی میانگین بارش ماهانه در دوره‌ی آینده نسبت به دوره‌ی پایه حدود ۴٪ افزایش خواهد یافت، و به $169/51$ میلی‌متر می‌رسد. مقدار میانگین سالانه‌ی درازمدت بارش از $423/39$ میلی‌متر به $427/15$ میلی‌متر افزایش خواهد یافت. ضریب هم‌بستگی برای داده‌های آزمون در روش برنامه‌ریزی بیان ژن $0/7$ بود و مقدار میانگین آورد از $3/33 \text{ m}^3/\text{s}$ به $3/14 \text{ m}^3/\text{s}$ تغییر می‌کند، که نشان می‌دهد $5/71$ ٪ کاهش خواهد یافت. در روش وایازی درخت تصمیم ضریب هم‌بستگی $0/995$ بود و میانگین آورد $10/3$ ٪ افزایش خواهد یافت و به $3/69$ متر مکعب بر ثانیه خواهد رسید.

واژگان کلیدی: اورنج، بارش روان‌آب، برنامه‌ریزی بیان ژن، تغییر اقلیم، مدل‌های گزارش پنجم

مقدمه

کارایی مدل وایزی درخت تصمیم در پیش‌بینی خشک‌سالی در ایستگاه هم‌دید یزد در دوره‌ی آماری ۱۳۸۴-۱۳۳۲ در چهار حالت بهره‌گیری از داده‌های خام، داده‌های بهنجار شده، بهره‌گیری از میانگین متحرک ۳ ساله، و بهره‌گیری از میانگین متحرک ۵ ساله اجرا شد. نتیجه‌ها نشان داد که مدل درخت تصمیم وایزی با دقت پذیرفتنی میانگین متحرک ۳ و ۵ ساله‌ی بارش را پیش‌بینی می‌کند (حبیبی پور ۲۰۱۰). فلاحی (۲۰۱۱) مدل وایزی درختی را برای پیش‌بینی بارندگی در سه ایستگاه هواشناسی بندرانزلی، اراک و قم به کار برد، که نتیجه‌ها بیان‌گر توانایی و دقت زیاد آن در پیش‌بینی بود. مدل درخت تصمیم یکی از روش‌های داده‌کاوی برای پیش‌بینی بارش و ارزیابی وضعیت خشک‌سالی در ایستگاه هم‌دید یزد بود، و مشخص کرد که در حالتی که از میانگین متحرک ۵ ساله‌ی داده‌ها بهره گرفته شود، توانایی زیادی در پیش‌بینی اندازه بارش دارد (دستوراتی و همکاران ۲۰۱۲). پیش‌بینی بارش ۳۰ ماه بعد در ایستگاه هم‌دید کرمانشاه با وایزی درخت تصمیم نشان داد که این مدل تقریباً مناسبی برای پیش‌بینی بارش است، و بهره‌گیری از میانگین متحرک منجر به زیاد شدن کارایی مدل نسبت به سایر حالت‌ها می‌شود (کمال امیدوار و همکاران ۲۰۱۴). ارزیابی عمل کرد مدل درخت تصمیم وایزی در پیش‌بینی خشک‌سالی در ایستگاه هم‌دید سندج نشان داد که این روش با ضریب اطمینان ۸۹٪ اندازه‌ی بارش یک‌ساله را به‌خوبی پیش‌بینی می‌کند (مظفری و همکاران ۲۰۱۵). پیش‌بینی آورد رود کشکان در استان لرستان با الگوریتم وایزی درخت تصمیم در مقایسه با الگوریتم شبکه‌های عصبی مصنوعی نشان داد که پیش‌بینی آورد رودخانه بر پایه‌ی اندازه‌های آن در دوره‌های زمانی قبل مناسب‌تر از اتکا به داده‌های بارش به‌تنهایی بوده‌است (محجوبی ۲۰۱۸). کاربرد سه مدل آب‌بهرس و برنامه‌ریزی بیان ژن و وایزی بردار پشتیبان در پیش‌بینی آورد در ایستگاه پیرسلیمان در ورودی سد جامیشان در استان کرمانشاه بر اثر تغییر اقلیم نشان داد که مدل‌های اقلیمی افزایش دمای میان $0/1^{\circ}\text{C}$ تا $1/4^{\circ}\text{C}$ ناشی از دو مدل esm - flo و cm5 - cnrm را نشان می‌دهد. نتیجه‌های بارش پیش‌بینی شده مشخص می‌کند که میانگین درازمدت ماهانه با حالت ممکن RCP8.5 در دوره‌ی آینده به ترتیب $1/1$ و $5/8$ ٪ نسبت به دوره‌ی پایه کاهش داشت (رحیمی و حافظ‌پرست ۲۰۲۰).

در پژوهش‌های قبلی پیش‌بینی آورد رودخانه با مدل ORANGE و روش درخت تصمیم انجام نشده‌است. بررسی هم‌زمان تغییر آورد رود در تاثیر تغییر اقلیم نیز با روش‌های برنامه‌ریزی بیان ژن و درخت تصمیم به شکل هم‌زمان انجام نشده‌است. هدف این پژوهش برپایه‌ی اهمیت تغییر اقلیم و اثر آن در ایستگاه‌های گوناگون، پیش‌بینی متغیرهای اقلیمی در دوره‌ی آینده و پیش‌بینی آورد رود بر اثر تغییر اقلیم با روش برنامه‌ریزی

یکی از مهم‌ترین ویژگی‌های تأثیرگذار بر چرخه‌ی آب‌شناسی تغییر اقلیم است که خیلی از دانشمندان در بخش علوم آب و محیط زیست توجه ویژه‌ی به آن دارند. افزایش گازهای گلخانه‌یی به دلیل آلودگی ناشی از فعالیت‌ها در بخش‌های صنعتی، حمل‌ونقل و مانند آن باعث افزایش دمای کره‌ی زمین و گرم‌تر شدن آن شده‌است. در هر آبخیز تغییر در میزان بارش و دما که از ویژگی‌های اقلیمی است، نقش مهمی در تعیین میزان روان‌آب در دوره‌های آینده دارد. تغییر اقلیم و پی‌رو آن تغییر شرایط محیطی و تأثیر آن بر فرآیندهای آب‌شناختی مانند روان‌آب سطحی برای تغییر اندوخته‌های آبی و برنامه‌ریزی و ایجاد بسترهای مفید برای مدیریت آن مهم است (IPCC ۲۰۱۳). آب‌شناسی با روش‌ها و نرم‌افزارهای گوناگون و اخیراً با روش‌های داده‌کاوی، آموزش و آزمون مدل‌های گوناگون یادگیری ماشینی متغیرهای هواشناسی و آب‌شناسی را پیش‌بینی می‌کند که با آن می‌توان اقلیم منطقه را پیش‌بینی، و دوره‌های ترسالی و خشک‌سالی را مشخص کرد.

پیش‌بینی روان‌آب ساعتی در زمان‌های گوناگون برای افق‌های پیش‌بینی بیشینه‌ی نه‌ساعته در حوزه‌ی آبخیز اورگوال در کشور فرانسه با روش برنامه‌ریزی بیان ژن برای به‌روزرسانیدن خطا و تکمیل کردن مدل بارش-روان‌آب MIKE11/NAM انجام شد. نتیجه‌های آن با اندازه‌های مشاهده‌یی و نتیجه‌های به‌دست‌آمده از روش‌های قدیمی‌تر به‌کاررفته در این حوزه نشان داد که این روش دقتی پذیرفتنی دارد (خو و همکاران ۲۰۰۱). روی داد بارش-روان‌آب در زمان‌های گوناگون در آبخیز سنگاپور با مساحتی حدود 6 km^2 در نظر گرفته شد. شش طوفانی که با شدت و زمان گوناگون برای آموزش و آزمون در نظر گرفته‌شد نشان داد که پیش‌بینی رفتار بارش-روان‌آب با روش برنامه‌ریزی بیان ژن در این آبخیزها موجب خطای کم‌تری می‌شود (لی‌یونگ و همکاران ۲۰۰۲). برنامه‌ریزی بیان ژن برای سه آبخیز در هنگ‌کنگ شامل هوک تاو، شانچیو و شونتشن نشان داد که نتیجه‌ها با داده‌های واقعی مطابقت زیادی دارد (جایاواردنا و همکاران ۲۰۰۵). پژوهش دیگری در ایالت پنسیلوانیای آمریکا که در آن از شبکه‌های عصبی و برنامه‌ریزی بیان ژن برای مدل‌سازی بارش-روان‌آب روزانه‌ی حوزه‌ی رود جونیاتا بهره گرفته شد، نشان داد که برنامه‌ریزی ژنی در قیاس با نتیجه‌های به‌دست‌آمده از شبکه‌های عصبی عمل کرد خیلی خوبی دارد و کاملاً کاربردی است (آیتک و همکاران ۲۰۰۸).

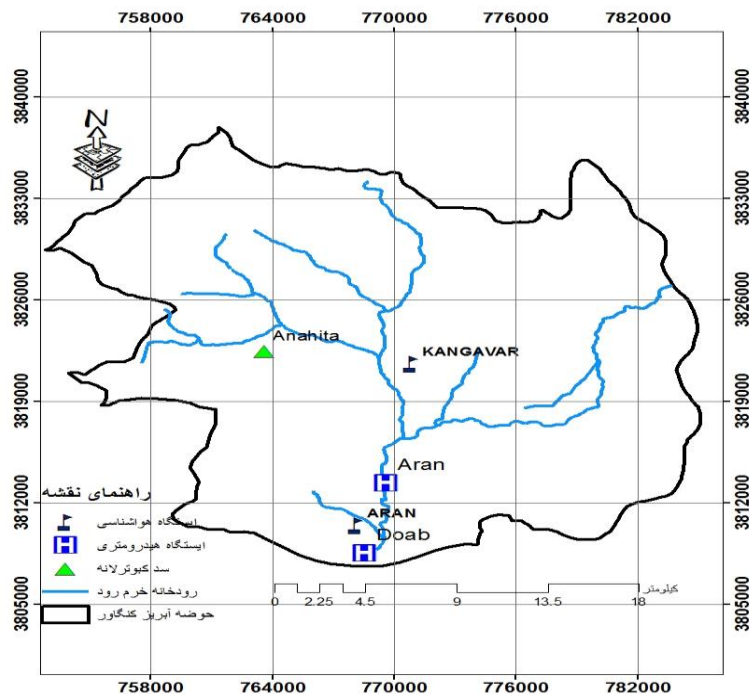
پیش‌بینی جریان روزانه‌ی رود ليقوان نشان داد که مقدار ضریب هم‌بستگی و میانگین مربع‌های خطا با مدل بهینه‌ی برنامه‌ریزی ژنی به ترتیب $0/959$ و $0/029$ ، و با مدل شبکه‌ی عصبی مصنوعی $0/948$ و $0/215$ است (فرودفام و همکاران ۲۰۰۹).

بیان ژن با مدل GEP است، که در پژوهش‌های بسیاری از دانشمندان کارایی زیادی در پیش‌بینی این متغیرها نشان داده است. برای روش وایازی درخت تصمیم و پیش‌بینی با آن از نرم‌افزاری جدید به نام اورنج بهره گرفته شد که با برنامه‌نویسی پایتون طراحی شده است، و عمل کرد آن بر پایه‌ی معیارهای خطا بهتر از وایازی خطی و ماشین بردار پشتیبان است.

مواد و روش‌ها

تغییر بارش و روان‌آب ایستگاه آران در آبخیز کنگاور استان کرمانشاه بر اثر تغییر اقلیم با روش برنامه‌ریزی بیان ژن در مدل GEP و وایازی درخت تصمیم و نرم‌افزار ORANGE

بررسی و پیش‌بینی شد. ایستگاه آب‌سنجی آران در روستای آران در شهر کنگاور، استان کرمانشاه، در غرب ایران با عرض جغرافیایی $34^{\circ} 25'$ ، طول جغرافیایی $47^{\circ} 55'$ ، و بلندی ۱۴۴۰ متر است. مساحت این آبخیز 647 km^2 است. خرم‌رود از بلندی‌های جنوب شرقی ملایر سرچشمه می‌گیرد و پس از آبیاری زمین‌های کشاورزی مسیر خود در استان‌های همدان و کرمانشاه به رود گاماسیاب می‌پیوندد. دشت حاصل‌خیز کنگاور در غرب این بلندی‌ها در ۱۴۵۷ متری از تراز دریا است. در این منطقه عمده‌ی بارش در آذر و دی است. منطقه‌ی پژوهشی در شکل ۱ و ویژگی‌های بارندگی، دما و آب‌دهی در جدول ۱ نشان داده شده است.



شکل ۱- آبخیز کنگاور و ایستگاه آب‌سنجی آران.

سده به حدود ۰/۱ می‌رسد و سپس کاهش می‌یابد و به ۰/۹ وات بر مترمربع در سال ۲۱۰۰ می‌رسد. مدل‌های اقلیمی پرشماری هست، اما در این پژوهش از مدل $FGOALS - g2$ که مدل جوی کاملاً همراه شده است بهره گرفته شد. این مدل چهار زیرمدل دارد، شامل گردش جوی، گردش اقیانوس، فرآیند سطح زمین، و یخ دریا. این مدل‌ها با تبدیل‌کننده‌ی شار به هم متصل می‌شود. $GOALS - g2$ نسخه‌ی روزآمد از $FGOALS - g1$ است (IPCC ۲۰۱۳).

حالت‌های ممکن اقلیمی گزارش پنجم جدید انتشار (IPCC ۲۰۱۳) چهار روند کلیدی دارد که برپایه‌ی واداشت تابشی آن‌ها تا سال ۲۱۰۰ نام‌گذاری شد ($RCP2.6, RCP4.5, RCP6, RCP8.5$). از آن‌جا که آبخیز کنگاور از منطقه‌های صنعتی نیست و جمعیت زیادی ندارد در این پژوهش حالت ممکن خوشبینانه‌ی $RCP2.6$ به کاربرده شد، که در آینده نسبت به حالت‌های ممکن دیگر کم‌ترین اندازه‌ی واداشت تابشی دارد. در این حالت، واداشت تابشی در میانه‌ی این

جدول ۱- اطلاعات هواشناسی ایستگاه هم‌دید کنگاور و آبدهی ایستگاه آران.

ویژگی	ژانویه	فوریه	مارس	آوریل	می	ژوئن	ژوئیه	اوت	سپتامبر	اکتبر	نوامبر	دسامبر
میانگین بارش ماهانه	۵۸	۶۵	۶۹	۶۷	۳۶	۴	۱	۱	۳	۲۸	۶۳	۵۳
بارندگی (mm)	۲۸۳	۳۰۵	۳۳۹	۳۴۷	۲۱۱	۲۱	۸	۱۳	۱۹	۸۳	۱۷۶	۱۰۸
کمینه‌ی میانگین بارش ماهانه	۱۳	۲۷	۲۲	۱۱	۵	۰	۰	۰	۰	۰	۰	۰
میانگین دمای ماهانه	-۰/۰۶	۲/۳۲	۶/۹۷	۱۲/۰۸	۱۶/۸۵	۲۲/۵۳	۲۶/۵۶	۲۵/۸۴	۲۰/۷۶	۱۴/۸۳	۷/۷۸	۲/۹۶
دما (°C)	۴/۴۷	۶/۸۷	۱۱/۰۱	۱۵/۴۱	۲۰/۰۶	۲۵/۸۷	۲۹/۸۵	۲۸/۳۴	۲۳/۸۹	۱۷/۶۲	۱۰/۸۷	۷/۲۸
کمینه‌ی میانگین دمای ماهانه	-۸/۶۵	-۸/۳۰	۱/۴۶	۹/۹۴	۱۳/۷	۱۸/۶۱	۲۳/۳۶	۲۲/۲۱	۱۷/۹۷	۱۲/۸	۵/۵۴	-۱/۲۲
آبدهی متوسط (m ³ /s)	۵/۲۳	۶/۵۳	۸/۱۶	۸/۱۵	۵/۳۱	۱/۸۱	۰/۳	۰/۰۶	۰/۰۵	۰/۵۱	۲/۹۲	۵

شاخه‌یی با اندازه و شکل‌های متفاوت در کروموزوم‌های خطی با طول ثابت شناسه‌گذاری می‌شود، سبب شده که فنوتیپ و ژنوتیپ از هم جدا شود و سامانه بتواند به سبب وجود آن‌ها از همه‌ی مزیت‌های تکاملی بهره‌مند شود. این روش ترکیبی از روش‌های برنامه‌ریزی ژنتیک و الگوریتم ژنتیک است که در آن کروموزوم‌های خطی و ساده با طول ثابت، مشابه با آن‌چه در الگوریتم ژنتیک بهره‌گرفته می‌شود، و ساختارهای شاخه‌یی با اندازه‌ها و شکل‌های متفاوت، مشابه با درختان تجزیه در برنامه‌ریزی ژنتیک، ترکیب می‌شود. اکنون با این‌که فنوتیپ در GEP همان نوع از ساختارهای شاخه‌یی به کار برده شده در GP را شامل می‌شود، ساختارهای شاخه‌یی که با GEP استنتاج می‌شود (بیان درختی) بیان‌کننده‌ی همه‌ی ژنوم‌های مستقل است. در این روش پدیده‌های گوناگون با مجموعه‌یی از تابع‌ها و مجموعه‌یی از پایانه‌ها مدل‌سازی می‌شود. مجموعه‌ی تابع‌ها، معمولاً شامل تابع‌های اصلی حسابی {+, -, ×, /}، تابع‌های مثلثاتی یا هر نوع تابع ریاضی دیگر تعریف شده از سوی کاربر است که برای تفسیر مدل مناسب است. مجموعه‌ی پایانه‌ها، از اندازه‌های ثابت و متغیرهای مستقل مسأله تشکیل شده‌است. خاصیت ناخطی، بی قطعیتی ذاتی فرآیند بارش-روان‌آب، نیاز به اطلاعات وسیع، و پیچیده‌بودن مدل‌های ساختمانی از دلیل‌هایی بوده‌است که باعث شده محققان به سوی روش‌های الهام گرفته‌شده از طبیعت، هم‌چون برنامه‌ریزی ژنتیک رو آورند. برای به کارگیری روش برنامه‌ریزی بیان‌ژن از نرم افزار GenXproTools بهره گرفته شد (قربانی و صالحی ۲۰۱۱). یکی از کارهای مهم در برنامه‌ریزی ژنتیک تعیین تابع برازش است که هدف آن یافتن راه‌حلی است که برای همه‌ی برازش‌ها به اندازه‌ی خطایی معین خوب عمل کند (رابطه‌ی ۵).

برای ریزمقیاس‌سازی از روش عامل‌های تغییر بهره گرفته شد. داده‌های بارش و دمای محاسبه‌شده با مدل AOGCM نسبت به دوره‌ی پایه دریافت، و در نرم‌افزار آرک‌جی‌آی‌اس ریزمقیاس‌سازی مکانی شد. داده‌های با قالب NC به CSV در اکسل تبدیل شد، و بر پایه‌ی خروجی‌های آن دما و بارش دوره‌ی آینده برای متوسط درازمدت همراه با معادله‌های ۱ تا ۴ ریزمقیاس‌گردانی شد (زارع‌زاده‌مهریزی و همکاران ۲۰۱۸).

$$\Delta T_i = \bar{T}_{GCM,Fut,i} - \bar{T}_{GCM,Base,i} \quad 1$$

$$\Delta P_i = \left(\frac{\bar{P}_{AOGCM,Fut,i}}{\bar{P}_{AOGCM,Base,i}} \right) \quad 2$$

ΔT_i و ΔP_i به ترتیب نشان‌دهنده‌ی مقدار تغییر مرتبط به دما و بارندگی برای میانگین درازمدت برای هر ماه، $\bar{T}_{GCM,Fut,i}$ میانگین درازمدت دمای شبیه‌سازی شده با AOGCM در دوره‌ی آینده برای هر ماه، $\bar{T}_{GCM,Base,i}$ میانگین درازمدت دمای شبیه‌سازی شده با AOGCM در دوره‌ی پایه است. برای بارندگی نیز به همین روش کار شد. در مرحله‌ی بعد با ایجاد این تغییر در داده‌های اصلی، داده‌های بارش و دمای ریزمقیاس شده به دست می‌آید (زارع زاده مهریزی و همکاران ۲۰۱۸).

$$T = T_{Obs} + \Delta T \quad 3$$

$$P = P_{Obs} \times \Delta P \quad 4$$

P, T به ترتیب دما و بارش برای دوره‌ی ۲۰۲۰-۲۰۵۳ و T_{Obs} و P_{Obs} دما و بارش مشاهده‌یی (۱۹۸۳ تا ۲۰۱۵) است. روش برنامه‌ریزی بیان ژن (فریرا ۲۰۰۱) که در آن همه‌ی ساختارهای

$$SS(t) = \sum_{i=1}^{N_t} (y_i(t) - \bar{y}(t))^2 \quad 7$$

$$Q(s, t) = SS(t) - SS(t_R) - SS(t_L) \quad 8$$

$SS(t)$ مجموع مربع‌های خطا برای درخت تصمیم، N_t تعداد داده‌ها در گره برگ t ، $y_i(t)$ مقدار خروجی (متغیر هدف در هر گره برگ)، $\bar{y}(t)$ میانگین متغیر هدف برای همه گره‌ها، $SS(t_R)$ و $SS(t_L)$ به ترتیب مقدار $SS(t)$ در شاخه‌ی چپ و راست است. متغیر ورودی s زمانی بهترین متغیر برای ایجاد شاخه در گره t است که مقدار $Q(s, t)$ را بیشینه کند (حبیبی‌پور و همکاران ۲۰۱۱).

اورنج (orange) نرم‌افزاری متن‌باز و رایگان است که با زبان پایتون نوشته شده است، و با نمایش تصویری خود تجزیه و تحلیل‌ها را جذاب‌تر نمایش می‌دهد. اجزای اورنج ویجت نامیده می‌شود. ویجت‌های اورنج امکان ورود داده‌ها به نرم‌افزار، پاک‌سازی و پیش‌پردازش داده‌ها، مصورسازی داده‌ها، مدل‌سازی (رده‌بندی و پیش‌بینی، خوشه‌بندی و کشف قانون‌های انجمنی)، ارزیابی مدل و متن‌کاوی را برای کاربران فراهم می‌کند. می‌توان امکانات اورنج را با افزونه‌های گوناگون افزایش داد. این نرم‌افزار در ۲۰۱۸ دومین ابزار داده‌کاوی برتر انتخاب شد (وایتن و همکاران ۲۰۱۶).

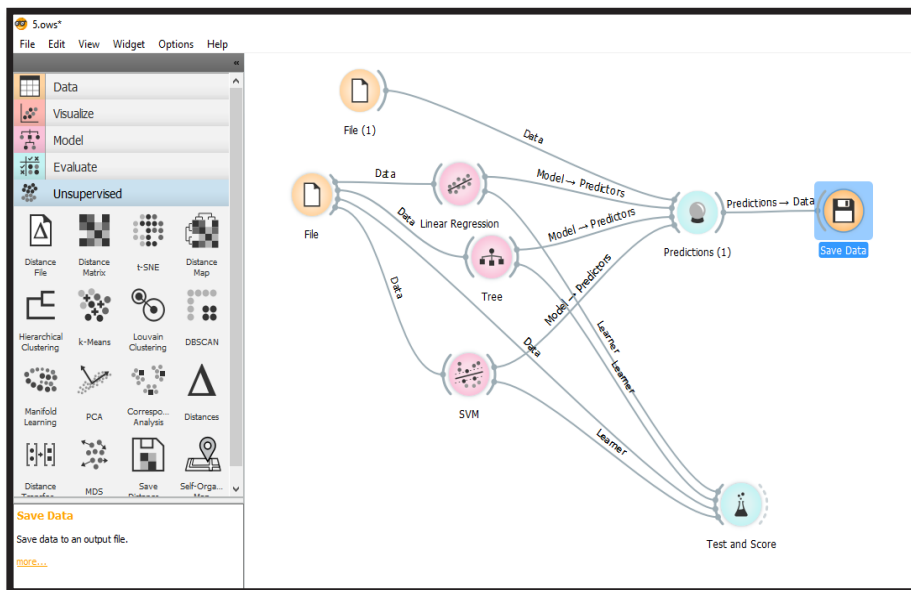
در این پژوهش با این نرم‌افزار داده‌کاوی مدل‌های وایزی گوناگون مانند SVR، Decision Tree، Linear اجرا و ارزیابی شد. با معیارهای خطای به‌دست‌آمده از آن‌ها روش وایزی درخت تصمیم برای پیش‌بینی متغیرها انتخاب شد. نمایش تصویری مدل شبیه‌سازی شده در نرم‌افزار اورنج در شکل ۲ نشان داده شده است.

$$f_i = \sum_{j=1}^n \left(R - \left| 100 \times \frac{P_{ij} - T_j}{T_j} \right| \right) \quad 5$$

R دامنه‌ی انتخابی، P_{ij} مقدار پیش‌بینی شده با برنامه‌ی انفرادی i برای j آزمایش n مورد برآزش، و T_j مقدار هدف برای برآزش j است، و مقدار قدرمطلق متناظر با درصد خطای نسبی نامیده می‌شود. اگر $P_{ij} = T_j$ باشد خطا صفر می‌شود. اگر این حالت پیش‌آید مقدار برآزش بیشینه خواهد بود. ژن‌های برنامه‌نویسی بیان ژن از یک سر و یک دنباله تشکیل شده‌است. سر شامل نمادهایی است که اعضای یک مجموعه‌ی تابع (F و پایانه) عضوهای یک مجموعه‌ی پایانه (T) را بازنمایی می‌کند. دنباله‌ها تنها شامل پایانه‌ها است. در نتیجه دو الفبای متفاوت در ناحیه‌های گوناگون ژن است. در هر مسئله طول سر (h) انتخاب می‌شود؛ در حالی که طول دنباله (t) تابعی از h و تعداد متغیرهای (n) تابعی با بیش‌ترین تعداد متغیر است (رابطه‌ی ۶) (شعیب و همکاران ۲۰۱۵).

$$t = h(n-1) + 1 \quad 6$$

الگوی وایزی درخت تصمیم به شکل بازگشتی دو دویی عمل می‌کند به‌طوری‌که گره‌های مادری را دقیقاً به دو گروه فرزند تقسیم می‌کند، و به‌طور بازگشتی منشعب کردن را تا زمانی که انشعاب دیگری نتواند ساخته شود ادامه می‌دهد (چاتام و لای ۲۰۱۱). معیار بهره‌گیری در درختان تصمیم وایزی، کم‌ترین انحراف مربع‌ها است. این معیار بر پایه‌ی کمیت خطا (رابطه‌ی ۷) به‌دست می‌آید و بهترین $SS(t)$ زمانی است که $Q(s, t)$ (رابطه‌ی ۸) بیش‌ترین مقدار باشد.



شکل ۲- محیط نرم‌افزار اورنج.

$$MAE = \sum_{i=1}^n (Z * (x_i) - Z(x_i))/n \quad 9$$

$$RMSE = (\sum_{i=1}^n (Z * (x_i) - Z(x_i))^2 / n)^{1/2} \quad 10$$

$$MSE = \frac{\sum_{i=1}^n (Z(x_i) - z * (x_i))^2}{n} \quad 11$$

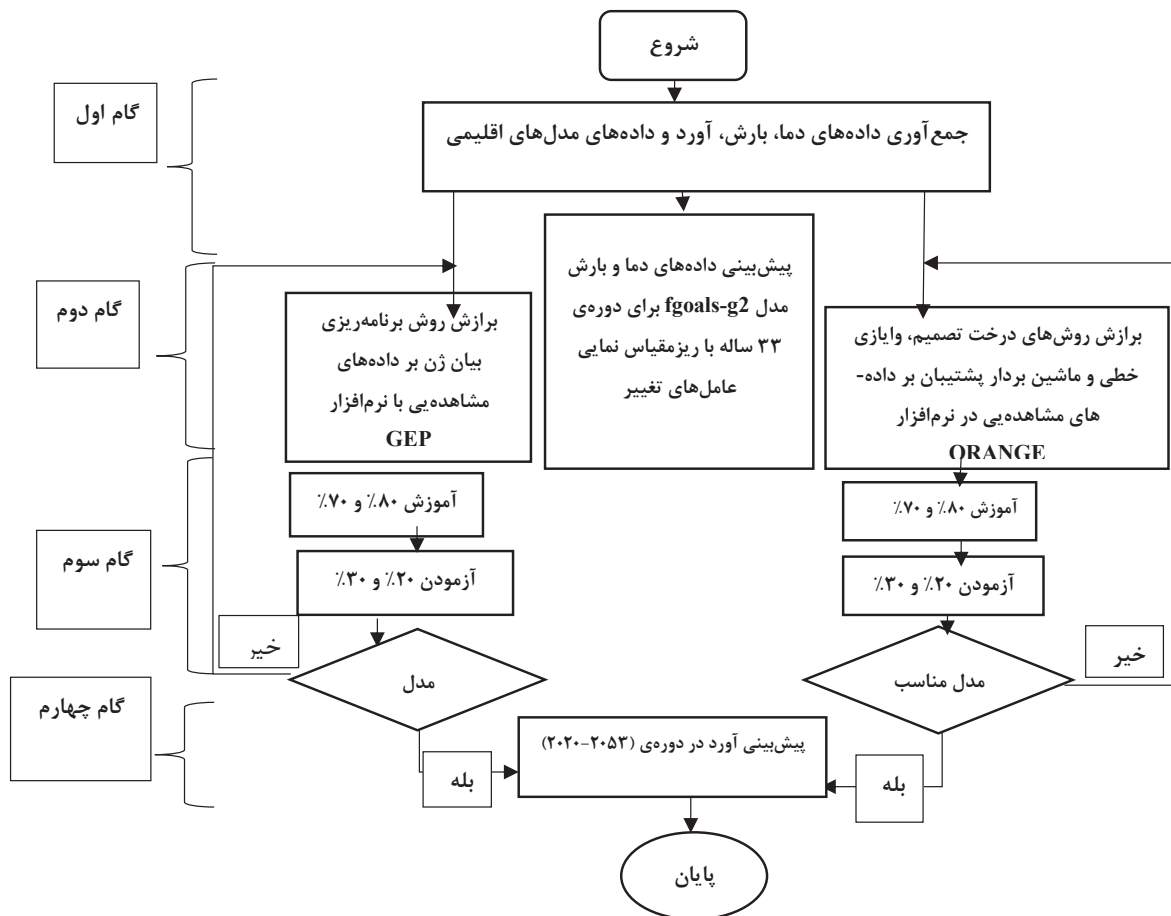
$$R^2 = \frac{(\sum_{i=1}^n [(Z(x_i) - \bar{Z}) \times (z * (x_i) - \bar{z * (x_i)})])^2}{\sum_{i=1}^n (Z(x_i) - \bar{Z})^2 \times \sum_{i=1}^n (z * (x_i) - \bar{z * (x_i)})^2} \quad 12$$

برای بررسی مناسب بودن مدل در این پژوهش از میانگین مربع‌های خطا (MSE) بهره‌گرفته شد، که ممکن است از صفر در عمل کرد عالی تا بی‌نهایت تغییر کند. ریشه‌ی میانگین مربع‌های خطا (RMSE) برای نشان دادن اختلاف میان اندازه‌های شبیه‌سازی شده و اندازه‌های اندازه‌گیری به کار رفت. میانگین مطلق خطا (MAE) برای مقایسه‌ی عبارت خطای نسبی اندازه‌های شبیه‌سازی شده با کمک اندازه‌های اندازه‌گیری شده به کار رفت، و ضریب هم‌بستگی (R) ارتباط خطی میان دو متغیر را اندازه‌گیری کرد (معادله‌های ۹-۱۲) (موریازی و همکاران ۲۰۰۷).

n تعداد داده‌ها $z * (x_i)$ مقدار برآورد در x_i مقدار $Z(x_i)$ مقدار مشاهده‌شده‌ی متغیر اصلی، و $\bar{z * (x_i)}$ متوسط مقدار برآورد در x_i و $Z(x_i)$ متوسط مقدار مشاهده‌ی است. هر چه مقدار سه معیار اول به صفر نزدیک‌تر باشد، و مقدار R^2 به یک نزدیک‌تر باشد بهتر، و داده‌های پیش‌بینی به واقعیت نزدیک‌تر است. مراحل انجام پژوهش در شکل ۳ نشان داده شده است.

نتایج و بحث

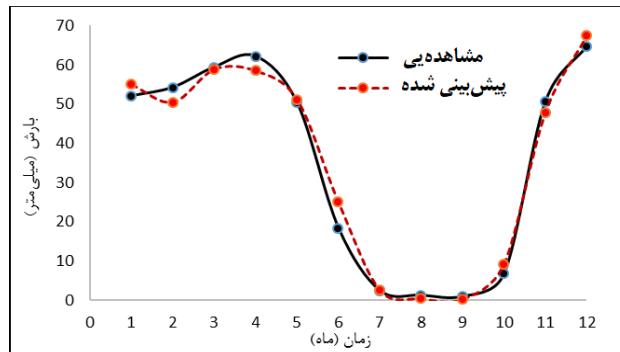
دما و بارش در دوره‌ی ۳۳ ساله (۲۰۲۰-۲۰۵۲) با داده‌های بارش و دمای مشاهده‌ی ایستگاه هم‌دید کنگاور پیش‌بینی شد. اندازه‌های میانگین درازمدت بارش و دما برای دوره مشاهده‌ی (۲۰۱۵-۱۹۸۳) و پیش‌بینی (۲۰۲۰-۲۰۵۲) در شکل‌های ۴ و ۵ نشان داده شده است.



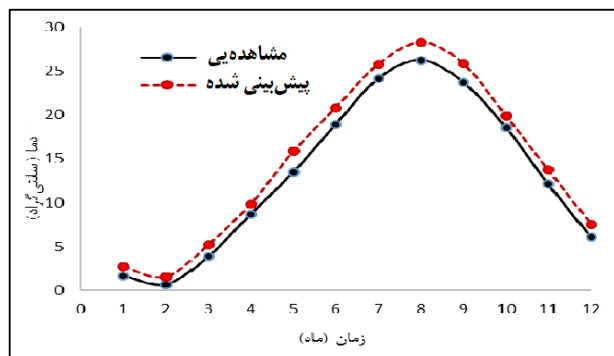
شکل ۳- نمودار جریان‌ی مرحله‌های انجام کار.

آن است (شکل ۴). پیشینه‌ی پیش‌بینی $67/41$ mm در دسامبر است. کمینه‌ی پیش‌بینی $0/3$ mm در سپتامبر است. دما در پژوهش‌های اقلیمی در سراسر دنیا روند افزایش را نشان می‌دهد، و معمولاً بیش‌تر از دمای مشاهده‌ی است. تغییر دما در این پژوهش در شکل ۵ نشان داده شده‌است.

در ژانویه‌ی ۱۹۹۵ بارش پیشینه‌ی مشاهده‌ی $162/6$ میلی‌متر بود که $4/21$ ٪ افزایش یافت و به $169/5$ میلی‌متر در ژانویه ۲۰۳۲ رسید. میانگین سالانه‌ی درازمدت بارش از $423/39$ میلی‌متر در دوره‌ی مشاهده‌ی به $427/15$ میلی‌متر در دوره‌ی آینده افزایش یافت. در ژانویه، می، جولای، اوت، اکتبر، و دسامبر بارش مشاهده‌ی کم‌تر از پیش‌بینی، و در بقیه‌ی ماه‌ها بیش‌تر از



شکل ۴- میانگین ماهانه‌ی بارش مشاهده‌ی و پیش‌بینی.



شکل ۵- میانگین ماهانه‌ی دمای مشاهده‌ی و پیش‌بینی.

پس از برآورد کردن اندازه بارش و دما در دوره‌ی آینده، این اندازه‌ها ورودی مدل‌های GEP و ORANGE گرفته‌شد. داده‌های آموزش و آزمون به نسبت‌های 70% - 30% یا 80% - 20% گرفته‌شد، و ویژگی‌های دیگر مدل برای دست‌یابی به بهترین پاسخ تغییر داده شد. جدول ۲ اندازه‌ی ویژگی‌ها، و جدول ۳ معیارهای خطای مدل برنامه‌ریزی بیان ژن را نشان می‌دهد.

پیشینه‌ی دما در دوره‌ی مشاهده‌ی $29/17^{\circ}\text{C}$ در اوت سال ۱۹۹۰ بود و به $31/18^{\circ}\text{C}$ در همین ماه در ۲۰۲۷ رسید، و میانگین دما از $31/21^{\circ}\text{C}$ به $14/8^{\circ}\text{C}$ رسید. در همه‌ی ماه‌ها دمای مشاهده‌ی کم‌تر از پیش‌بینی است، و در اوت بیش‌ترین مقدار دما برای مشاهده‌ی و پیش‌بینی به ترتیب $26/26^{\circ}\text{C}$ و $28/28^{\circ}\text{C}$ بود. کم‌ترین مقدار در فوریه به ترتیب $0/66^{\circ}\text{C}$ و $1/57^{\circ}\text{C}$ بود.

جدول ۲- ویژگی‌های مدل برگزیده‌ی GEP برای پیش‌بینی آب‌دهی.

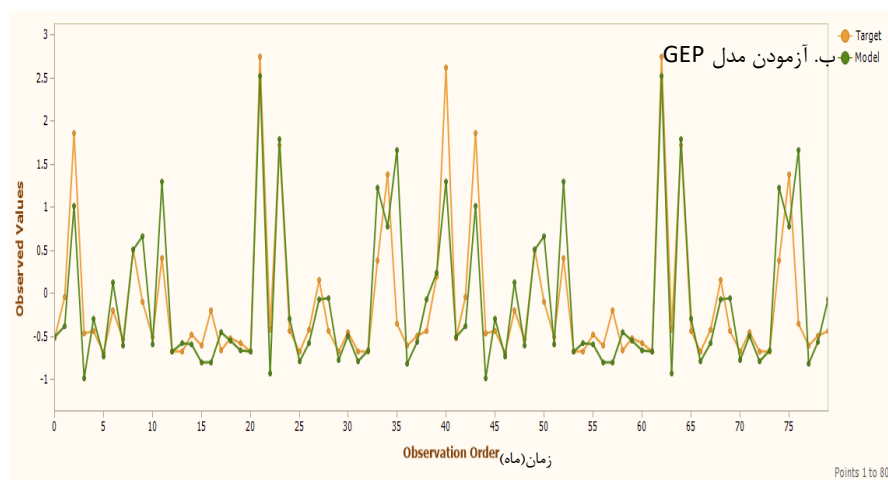
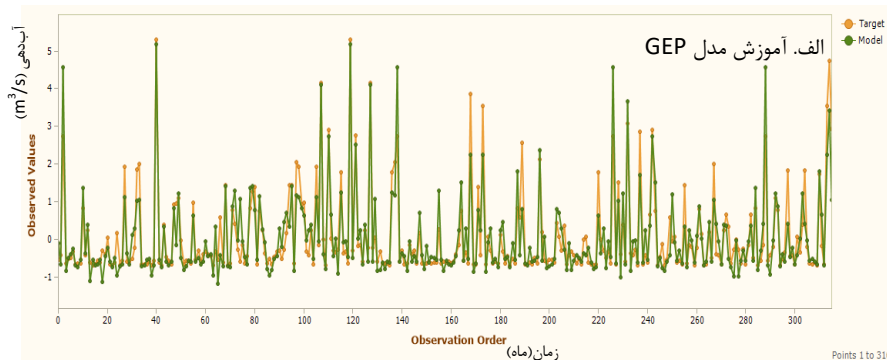
ویژگی	مقدار
تعداد کروموزوم‌ها	۵۰
تعداد ژن‌ها	۳
اندازه‌ی سر	۸
معیار خطای تابع برازش	R

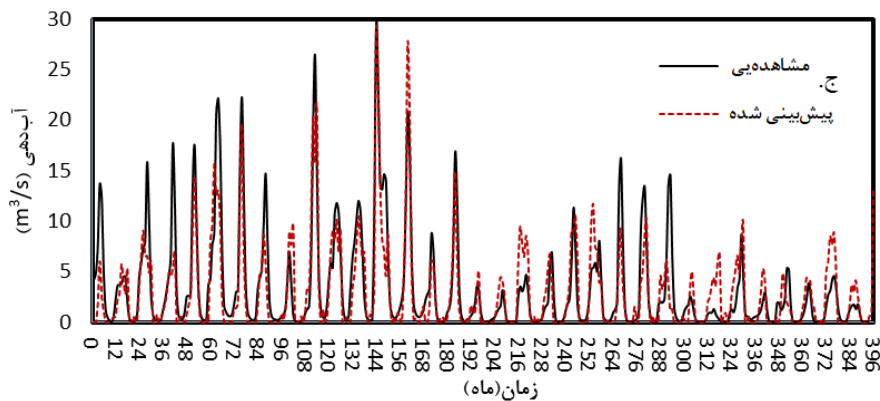
جدول ۳- معیارهای خطا در آموزش و آزمودن مدل برنامه‌ریزی بیان ژن.

مدل	MSE	RMSE	MAE	R ²
GP (آموزش)	۰/۳۱۷	۰/۵۶۳	۰/۳۶	۰/۷۶
GP (آزمون)	۰/۱۹	۰/۴۴	۰/۳۱	۰/۷۰

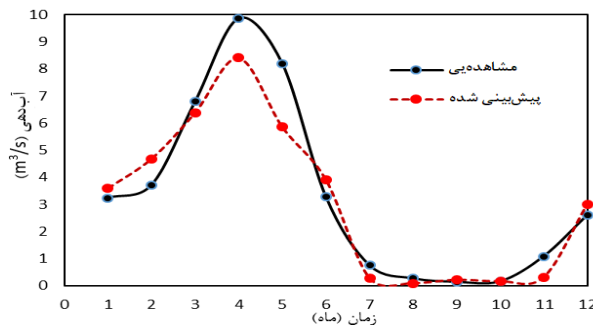
ماهانه به‌دست‌آمده از روش برنامه‌ریزی بیان ژن در شکل‌های ۶ و ۷ نشان داده شده‌است. آورد مشاهده‌ی (۲۰۱۵-۱۹۸۳) و آورد پیش‌بینی‌شده (۲۰۲۰-۲۰۵۲) در دوره‌ی ۳۳ ساله در برابر زمان نشان داده شده‌است.

ضریب هم‌بستگی برای داده‌های آموزش ۰/۷۶ و داده‌های آزمون ۰/۷۰ بود. در پیش‌بینی روان‌آب ضریب هم‌بستگی زیاد ۰/۶، اگر بقیه‌ی معیارهای خطا مناسب باشد پذیرفته می‌شود (بینامان و همکاران ۲۰۰۵؛ سندی و همکاران ۲۰۰۱). اندازه‌های پیش‌بینی شده‌ی آورد ماهانه و میانگین درازمدت





شکل ۶- روان آب مشاهده‌یی و پیش‌بینی شده، بر اثر تغییر اقلیم با برنامه‌ریزی بیان ژن.



شکل ۷- میانگین ماهانه‌ی روان آب مشاهده‌یی و پیش‌بینی شده در اثر تغییر اقلیم با برنامه‌ریزی بیان ژن.

پیش‌بینی شده در آوریل به ترتیب m^3/s ۹/۸۵ و m^3/s ۸/۴ بود که m^3/s ۱۴/۷۲٪ کاهش یافت.

سه روش وایازی خطی، ماشین بردار پشتیبان و درخت تصمیم در نرم‌افزار اورنج اجرا شد و بر پایه‌ی معیارهای خطا (جدول ۴) بهترین نتیجه‌ها، به‌دست آمده از روش درخت تصمیم انتخاب شد.

آورد در ژانویه‌ی ۱۹۹۴ و ۲۰۳۲ بیش‌ترین مقدار را نشان می‌دهد، و به ترتیب m^3/s ۲۹/۳۱ و m^3/s ۲۸/۷۷ است (۸۴/۱٪ کاهش). میانگین آورد از m^3/s ۳/۳۳ در وضع موجود به ۳/۱۴ (کاهش ۵/۷۱٪) در دوره‌ی آینده تغییر خواهد کرد. میانگین درازمدت آورد بیشینه در داده‌های مشاهده‌یی و

جدول ۴- معیارهای خطای به‌دست آمده از سه روش وایازی خطی، ماشین بردار پشتیبان و درخت تصمیم در نرم‌افزار اورنج.

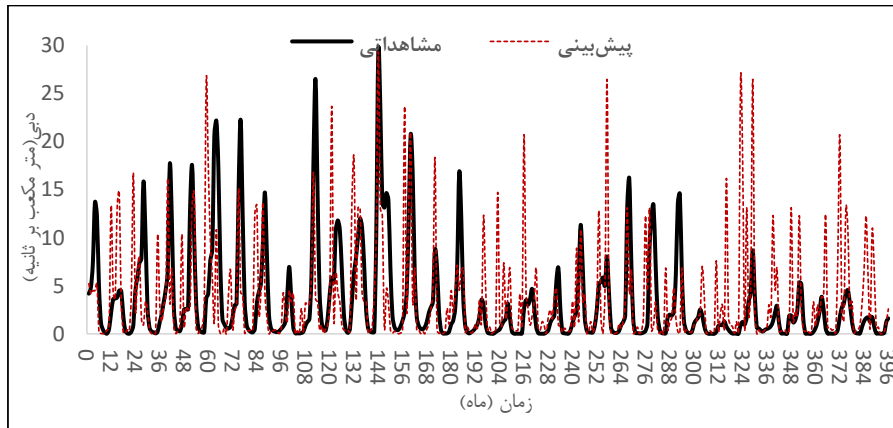
مدل	MSE	RMSE	MAE	R ²
Linear	۱۷/۰۹۱	۴/۱۳۴	۲/۶۳	۰/۲۸۱
SVM	۸/۰۴۰	۲/۸۳۶	۱/۱۸۲	۰/۶۶۲
Decision Tree	۰/۱۳	۰/۳۶۱	۰/۰۷۲	۰/۹۹۵

آورد بیشینه در داده‌های مشاهده‌یی و پیش‌بینی (شکل ۸) مساوی است ($۲۹/۳۱ m^3/s$)، ولی مقدار میانگین داده‌ها از m^3/s ۳/۳۳ به اندازه‌ی m^3/s ۱۰/۳٪ افزایش خواهد یافت و به m^3/s ۳/۶۹ خواهد رسید. میانگین ماهانه‌ی آورد مشاهده‌یی در مارس، آوریل، می، ژوئن و جولای بیش‌تر از آورد پیش‌بینی

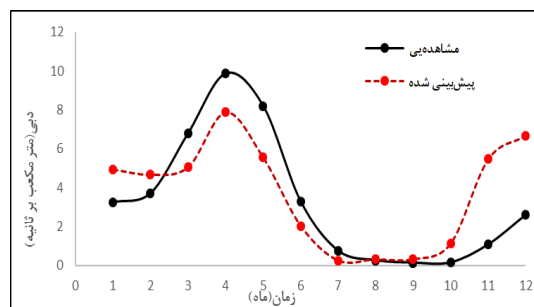
بر پایه‌ی روش وایازی درخت تصمیم که خطای ریشه‌ی مجذور مربع‌های خطا $۰/۳۶۱$ ، میانگین خطای مطلق $۰/۰۷۲$ و میانگین مربع‌های خطای آن $۰/۱۳$ کم‌ترین، و ضریب هم‌بستگی آن بیش‌ترین ($۰/۹۹۵$) بود روش برتر دانسته، و پیش‌بینی آورد در این مدل با این روش انجام شد (شکل ۸ و ۹).

یکی از ویژگی‌های مهم در آب‌شناسی منحنی تداوم جریان یا آورد رده‌بندی شده در رود است که از رسم کردن آورد نسبت به درصد زمان به دست می‌آید. در شکل ۱۰ منحنی تداوم جریان برای روان آب مشاهده‌ی و پیش‌بینی‌شده‌ی هر دو روش برنامه‌ریزی بیان‌زن و وایازی درخت تصمیم رسم شده‌است، تا با کمک آن بتوان تغییر پیش‌بینی جریان در آب‌دهی‌های کم و زیاد را بهتر نشان داد.

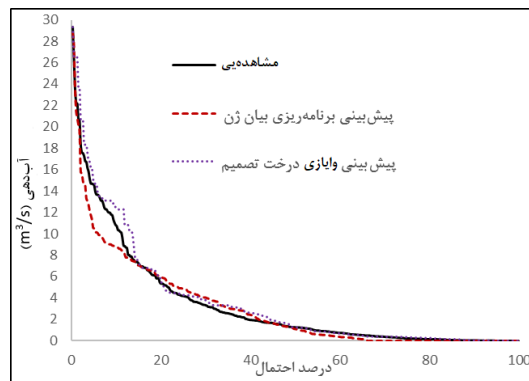
است. در زمان‌های دیگر کمتر از مقدار مشاهده‌ی است، و در آوریل که آورد بیشینه‌ی مشاهده‌ی و پیش‌بینی در آن است به ترتیب $9/85 \text{ m}^3/\text{s}$ و $7/88 \text{ m}^3/\text{s}$ است، و ۲۰٪ کاهش می‌یابد (شکل ۹). افزایش دما و ذوب شدن زود هنگام برف سبب افزایش جریان در ماه‌هایی است که آورد در آن بیش‌تر از مشاهده‌ی است، و در کل افزایش اندازه‌ی تبخیر و تعرق بر اثر افزایش گرما دلیل اصلی کاهش آورد کل است.



شکل ۸- روان آب مشاهده‌ی و پیش‌بینی‌شده بر اثر تغییر اقلیم با روش وایازی درخت تصمیم.



شکل ۹- میانگین ماهانه‌ی روان آب مشاهده‌ی و پیش‌بینی شده بر اثر تغییر اقلیم با روش وایازی درخت تصمیم.



شکل ۱۰- منحنی تداوم جریان روان آب مشاهده‌ی و پیش‌بینی شده بر اثر تغییر اقلیم با برنامه‌ریزی بیان‌زن و درخت تصمیم.

در روش برنامه‌ریزی بیان ژن اندازه‌ی آورد در مجموع ۳۵٪ از زمان‌ها کمتر از مشاهده‌ی است، و تنها در ۲۵٪ از زمان آورد پیش‌بینی بیش‌تر از مشاهده‌ی است، و در حالت‌های دیگر بر هم منطبق است. در روش وایازی درخت تصمیم، آورد پیش‌بینی در ۲۶٪ از زمان بیش‌تر از مشاهده‌ی است، و در ۷۳٪ از زمان آب‌دهی‌ها بر هم منطبق است. انطباق مناسبی در پیش‌بینی آب‌دهی‌های کم و زیاد با آب‌دهی‌های مشاهده‌ی هست (شکل ۱۰).

بحث و نتیجه‌گیری

در دنیای امروز که زمین بر اثر فعالیت انسان و زیاد شدن گازهای گلخانه‌یی در حال گرم‌تر شدن است یکی از نیازهای مهم جامعه اجرا کردن پژوهش‌های مرتبط با تغییر اقلیم است، که نقش اصلی را در ویژگی‌های بارش و دما، و پی‌رو آن تأثیر این عامل‌ها بر اندوخته‌های تامین‌کننده‌ی آب دارد. هدف از این پژوهش پیش‌بینی آورد با روش‌های برنامه‌ریزی بیان ژن و وایازی درخت تصمیم با مدل GEP و نرم‌افزار اورنج بود که با اندازه‌های دما و بارش و آورد مشاهده‌ی در دوره‌ی ۱۹۸۳ تا ۲۰۱۵ اندازه‌ی روان‌آب در دوره‌ی آینده‌ی ۲۰۲۰ تا ۲۰۵۲ با مدل اقلیمی FGOALS-g2 در حالت ممکن RCP2.6 بررسی شد. نتیجه‌ها نشان داد که در دوره‌ی آینده میانگین بارش ۰/۹٪ و دما در حدود ۱۳٪ افزایش می‌یابد. پیش‌بینی آورد با روش برنامه‌ریزی بیان ژن نشان داد که مقدار میانگین آورد از $3/33m^3/s$ به $3/14m^3/s$ تغییر می‌کند و ۵/۷۱٪ کاهش می‌یابد. در روش وایازی درخت تصمیم آورد ۱۰/۳٪ افزایش می‌یابد و به $3/69m^3/s$ می‌رسد. رحیمی و حافظ‌پرست (۲۰۲۰) در منطقه‌ی سد جامیشان استان کرمانشاه دریافتند که با پیش‌بینی آورد در هر سه مدل GEP، SVM، و IHACRES با مدل flo_esm و cnrm_cm5 و حالت ممکن RCP8.5 در دوره‌ی ۲۰۲۱-۲۰۵۰ کاهش روان‌آب روی خواهد داد. بیش‌ترین کاهش روان‌آب در SVM در مدل flo_esm با ۲۸/۹٪،

و کم‌ترین آن در GEP در مدل cnrm_cm5 با ۱۴/۱٪ بود. بنابراین مقایسه‌ی این دو پژوهش کاهش آورد در دوره‌ی آینده به اندازه‌ی ۵/۷۱ و ۱۴/۱٪ با مدل GEP نشان می‌دهد. دستوراتی و همکاران (۲۰۱۲) و امیدوار و همکاران (۲۰۱۴) در پیش‌بینی بارش ایستگاه هم‌دید یزد و کرمانشاه عمل کرد مدل درخت تصمیم را بررسی کردند. نتیجه‌های آنان نشان می‌دهد که این مدل توان زیادی برای پیش‌بینی بارش دارد. در این پژوهش نیز از روش درخت تصمیم برای پیش‌بینی آورد بهره‌گرفته شد و نتیجه‌های رضایت‌بخشی داشت. لی یونگ و همکاران (۲۰۰۲) در آبخیز سنگاپور، و جایاواردنا و همکاران (۲۰۰۵) در هنگ‌کنگ چین نشان دادند که برنامه‌ریزی بیان ژن در پیش‌بینی بارش روان‌آب خطای کم‌تری دارد، در حالی‌که در پژوهش ما نتیجه‌های مدل برنامه‌ریزی بیان ژن با معیارهای خطا، میانگین مربع‌های خطا، ریشه‌ی مجذور مربع‌های خطا، و میانگین خطای مطلق و ضریب هم‌بستگی به ترتیب ۰/۱۹، ۰/۴۴، ۰/۳۱ و ۰/۷۰ دقت مناسبی برای پیش‌بینی نشان داد. از آن‌جا که در این پژوهش ضریب هم‌بستگی و معیارهای خطای روش وایازی درخت تصمیم با اختلاف معنی‌داری بهتر از روش برنامه‌ریزی بیان ژن بود، نتیجه‌های این روش اعتمادپذیرتر است. بر پایه‌ی نتیجه‌های این پژوهش دما در این منطقه در سال‌های آینده حتماً افزایش خواهد یافت. اگر اندازه‌ی بارش زیاد افزایش نیابد، افزایش دما سبب کاهش آورد رودخانه‌ی جامیشان خواهد شد. هنوز نمی‌توان با قطعیت مدل آب‌شناسی مشخصی را برای منطقه پیشنهاد کرد، و بهتر است در هر پژوهش مدل‌های بررسی شده بازتنظیم، و درستی آن‌ها سنجیده شود، و به نتیجه‌های مدلی که معیارهای خطای آن پذیرفتنی باشد استناد شود.

سپاس‌گزاری

با تشکر از شرکت آب منطقه‌یی و سازمان هواشناسی استان کرمانشاه که اطلاعات این پژوهش را در اختیار نویسندگان قرار دادند.

- Aytek A, Asce M, Alp M. 2008. An application of artificial intelligence for rainfall-runoff modeling. *Journal of Earth System Science* 117: 145-155
- Binaman J, Shoemaker CA. 2005. An analysis of high-flow sediment event data for evaluating model performance. *Journal of Hydrological Processes*. 19: 605-620.
- Chattamvelli R. 2011. Data mining algorithms. (1th ed). Oxford: Alpha Science: 274-290
- Daštoorati MT, Habibipour A, Ekhtesasi MR, Talebi A, Mahjoobi J. 2012. Evaluation of the decision Tree Model in precipitation prediction (Case study: Yazd Synoptic Station). *Iran-Water Resources Research*. 8(3): 14-27. (In Persian).
- Farboudfam N, Ghorbani MA, Alami MT. 2009. Forecasting of river flow using genetic programming (Case study: Liqwan River Basin). *Agriculture Science*. 19(4): 107-123 (In Persian).
- Fallahi MR, Varvani H, Glian S. 2011. Flood prediction with decision tree regression model due to flood control. Fifth Conference of watershed management and soil and water resources management. 9th and 10th march, Iran. Kerman (In Persian).
- Ferreira C. 2001. Gene expression programming a new adaptive algorithm for solving problems. *Complex Systems*. 13(2):87-129
- Ghorbani M, Salehi A. 2011. Using genetic programming for investigating ground water quality changes with surface water fluctuations in Bourkhar plain in Esfahan, 6th National Conference in Civil Engineering. 16th Esfand. Mazandaran. (In Persian).
- Habibipour A, Daštoorati MT, Ekhtesasi MR, Afkhami H. 2011. Investigating the effect of decreasing of data range on the efficiency of artificial neural network and decision tree regression models in drought prediction. *Journal of Watershed Management*. 2(3): 63-79. (In Persian).
- Habibipour A. (2010). Prediction of drought in Yazd Synoptic Station using decision tree, M.Sc. Thesis, Yazd University. 130 p. (In Persian).
- IPCC (Intergovernmental Panel on Climate Change). 2013. The physical science basis. In: Stocker TF, Qin Plattner D, Tignor GK, Allen M, Boschung SK, Nauels J, Xia A, Bex Y, Midgley V. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge. United Kingdom and New York, NY, USA, 1535 p.
- Jayawardena AW, Muttill N, Fernando TMKG. 2005. Rainfall-Runoff Modelling using genetic programming. International Congress on Modelling and Simulation Society of Australia and New Zealand. New Zealand: 1841-1847 ISBN: 0-9758400-2-9.
- Khu ST, Liong SY, Babovic V, Madsen H, Muttill N. 2001. Genetic programming and its application in real-time runoff forecasting. *Journal of the American Water Resources Association*. 37(2): 439-451.
- Koza JR, Koza JR. 1992. Genetic programming: on the programming of computers by means of natural selection (Vol. 1). MIT Press.
- Liong SY, Gautam TR, Khu ST, Babovic V, Keijzer M, Muttill N. 2002. Genetic programming: A new paradigm in rainfall runoff modeling. *Journal of the American Water Resources Association*. 38(3): 705-718.
- Mahjoubi E, Mahjoubi J, Kouchaksaraei A. 2018. Comparison of Decision Trees regression (CART Algorithm) and Artificial Neural Network to Predict Kashkan River Flow, 7th National Conference on Water Resources Management. 5-6 Ordibehesht. Yazd Uni-

- versity. (In Persian).
- Moriasi DN, Arnold JG, Van Liew MW, Binger RL, Harmel RD, Veith TL. 2007. American Society of Agricultural and Biological Engineers . 50(3): 885–900.
- Mozafari G, Shafiei SH, Taghizadeh Z. 2015. Evaluating the efficiency of the regression decision tree model in drought forecasting (Case study: Sanandaj Synoptic Station). Dangers of the Natural Environment. 4(6): 1–19. (In Persian).
- Omidvar K, Shafiei S, Taghizadeh M. 2014. Evaluating the efficiency of the decision tree model in predicting precipitation of Kermanshah synoptic station. Journal of Practical Research in Geographical Sciences. 14(34): 89–110. (In Persian).
- Rahimi B, HafezParast Mavadat M. 2020. Comparison of support vector regression models, programming of gene expression and IHACRES in predicting runoff changes due to climate change (Case study: Jamishan Dam). Iranian Soil and Water Research. 15(10): 2483–2499. (In Persian). doi: 10.22059 / ijswr.2020.303779.668640.
- Santhi C, Arnold JG, Williams J, Dugas WA, Hauck L. 2001. Validation of the SWAT model on a large river basin with point and nonpoint sources. The American Water Resources Association, 37 (5): 1169–1188.
- Shoaib M, Shamseldin AY, Melville BW, Khan MM. 2015. Runoff forecasting using hybrid wavelet gene expression programming (WGEP) Approach. Journal of Hydrology, 527: 326–344.
- Soltani A, Ghorbani M, Fakherifard A, Darbandi S, Farsadizadeh D. 2010. Genetic programming and use of it in modeling of precipitation process-runoff. Water and Soil Science. 20(4): 62–71. (In Persian).
- Ian H, Witten, Eibe Frank, Mark A, Hall, Christopher J, Pal. 2016. Data mining, fourth edition: practical machine learning tools and techniques (4th. ed.). Morgan Kaufmann Publishers Inc. San Francisco. CA. USA.
- Zarezade Mehrizi SH, Khorani A, Bazrafshan J, Bazrafshan O. 2018. Assessment of future runoff trends under multiple climate change scenarios in the Gamasiab River Basin, Iranian Journal of Ecohydrology, 5(3): 777–789.



Watershed Management Research

VOL. 35, No.1, Ser. No: 134, Spring 2022, pp. 2-15

DOI: 10.22092WMRJ.2021.354279.1403

Research Paper



Prediction of the Discharge Rate of the Kangavar Watershed, the Province of Kermanshah, Using Gene Expression Programming and the Decision Tree Regression

Maryam Hafezparast Mavadat

(Corresponding Author)* Ph.D. in Water Resources Engineering, Assistant Professor, Department of Water Science and Engineering, Razi University

Foroozan Payfeshorde

M.Sc. Student of Water Resources Engineering, Razi University

Corresponding Author's Email: m.hafezparast@razi.ac.ir

Received: 21 April 2021

Accepted: 20 June 2021

Abstract

A lack of sufficient water resources, climate change, and especially global warming, is causing apprehension in societies, particularly in watershed managers. Scientists and engineers are forecasting climatic data to prevent the waste of water resources and to optimize their use in watersheds. The Aran Station in the Kangavar Watershed was chosen to predict the trend of temperature, precipitation, and runoff using the CMIP5 climate model. Benefiting from 33 years of data (1983 to 2015), the Gene Expression Programming method (GEP) and Decision Tree methods were developed to train, test, and predict the river discharge rate. Different climate models were implemented using the historical data of the study area. The Fgoals-g2 was chosen to predict temperature and precipitation data for the 2020-2052 periods. The RCP2.5 climate scenario was used as an optimistic scenario, and the output of the change factor downscaling method was used as an input for the GEP model and the ORANGE Software to find the best prediction of the discharge parameter in the future. The results indicated that the temperature of the next cycle will increase by 13 degrees and the maximum monthly temperature will reach 31.18 degrees centigrade. The maximum monthly precipitation will increase by 4 percent and reach 169.51mm. The longtime yearly mean precipitation will change from 423.39 mm to 427.15 mm. The correlation coefficient of the test data in the GEP was 0.70. The maximum monthly discharge will decrease 1.84 percent, from 29.31 to 28.77 cubic meters per second (m^3s^{-1}). The mean discharge will decrease 5.71 percent, from 3.33 to 3.14 cubic meters per second. The correlation coefficient of the test data in the decision tree regression method, using the ORANGE software was 0.995. The mean discharge will increase by 10 percent and reach $3.69 m^3s^{-1}$. The maximum yearly discharge will decrease by 6 percent, from 7.62 to $7.12 m^3s^{-1}$.

■ **Keywords:** Gene expression programming, orange, CMIP5 models, rainfall- runoff ■