

Rainfall- Runoff Modeling Using HBV Model and Random Forest Algorithm in Bazoft Watershed

FATEMEH SOHRABI GESHNIGANI¹, RASOUL MIRABBASI NAJAFABADI^{1*}, MOHAMMAD REZA GOLABI²

1. Department of Water Engineering, Faculty of Agriculture, Shahrekord University, Shahrekord, Iran.

2. Department of Hydrology and Water Resources, Faculty of Water and Environmental Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran.

(Received: Feb. 12, 2021- Revised: Apr. 5, 2021- Accepted: Apr. 24, 2021)

ABSTRACT

Estimation of runoff in a catchment area is important from various aspects such as dam reservoir management, water resources management, flood regulation, and erosion control in river banks and bed. In the present study, a conceptual model of HBV and an intelligent model of Random Forest (RF) were used to simulate the rainfall-runoff process in Bazoft watershed at the Landi hydrometric station during the period of 2010 to 2017. In order to evaluate the performance of models, the statistical criteria, including Correlation coefficient (r), Root Mean Squares Error (RMSE), Nash-Sutcliffe efficiency coefficient (NS), Mean Absolute Percentage Error (MAPE), and Mean Absolute Error (MAE) were used. Comparing the results of HBV and RF models revealed that the RF model outperformed the HBV. Thus, the RF model with $r=0.95$, $NS=0.82$, $MAPE=9.59$, $MAE=0.25$, and $RMSE=0.39$ m³/s was selected as the top model which might be used as a new choice to predict runoff in Bazoft watershed.

Keywords: Rainfall, Runoff, Evapotranspiration, Random Forest Model.

* Corresponding Author's E-mail: mirabbasi_r@yahoo.com

مدل سازی بارش-رواناب با استفاده از مدل HBV و الگوریتم جنگل تصادفی در حوضه آبخیز بازفت

فاطمه سهرابی گشنیگانی^۱، رسول میرعباسی نجف آبادی^{۱*}، محمدرضا گلابی^۲

۱. گروه مهندسی آب، دانشکده کشاورزی، دانشگاه شهرکرد، شهرکرد، ایران.

۲. گروه هیدرولوژی و منابع آب، دانشکده مهندسی آب و محیط زیست، دانشگاه شهید چمران اهواز، اهواز، ایران.

(تاریخ دریافت: ۱۳۹۹/۱۱/۲۴ - تاریخ بازنگری: ۱۴۰۰/۱۱/۱۶ - تاریخ تصویب: ۱۴۰۰/۲/۴)

چکیده

برآورد رواناب حاصل از بارندگی در یک حوضه آبخیز از جهات گوناگون از جمله مدیریت مخازن سدها، مدیریت منابع آب، تنظیم سیلاب، کنترل فرسایش کناره و بستر رودخانه حائز اهمیت می باشد. در این مطالعه، از مدل مفهومی HBV و مدل هوش مصنوعی جنگل تصادفی (RF) به منظور شبیه سازی فرایند بارش-رواناب در حوضه آبخیز بازفت در ایستگاه هیدرومتری لندی برای دوره آماری ۲۰۱۰ تا ۲۰۱۷ استفاده شد. برای ارزیابی عملکرد مدل ها، از آماره های ضریب همبستگی (r)، ریشه میانگین مربعات خطا (RMSE)، معیار کارایی نش-ساتکلیف (NS)، میانگین مطلق درصد خطا (MAPE) و میانگین قدرمطلق خطا (MAE) استفاده شد. مقایسه نتایج مدل مفهومی HBV و مدل RF نشان دهنده عملکرد بهتر مدل RF بود. بنابراین، مدل RF با مقادیر $RMSE=0/39 \text{ m}^3/\text{s}$ ، $MAPE=9/59$ ، $MAE=0/25$ ، $r=0/95$ و $NS=0/82$ به عنوان مدل برتر انتخاب گردید و این مدل می تواند برای کاربردهای آینده به عنوان یک گزینه جدید برای پیش بینی رواناب در حوضه بازفت مورد استفاده قرار گیرد.

واژه های کلیدی: بارش، رواناب، تبخیر- تعرق، مدل جنگل تصادفی، حوضه بازفت.

مقدمه

توجه به محدودیت منابع آب شیرین و قابل استحصال، پیش بینی هرچه دقیق تر تغییرات دبی جریان در طول سال از ارکان اساسی برنامه ریزی و مدیریت منابع آب های سطحی می باشد.

در هر حال در طرح های آبی لازم است که به نحوی از میزان تبدیل بارش به رواناب آگاه شد. در این راستا، مدل های زیادی از اوایل دهه ۱۹۶۰ توسعه داده شده است تا فرایند بارش-رواناب را شبیه سازی کنند. مدل ابزاری است که می تواند بخشی از واقعیت را شبیه سازی کند. هدف از مدل سازی، نشان دادن عملکرد و رفتار یک پدیده طبیعی تحت شرایط معین می باشد (Phomcha et al., 2011). مدل ها با توجه به نوع نگرش حاکم در ایجاد رواناب، به دو دسته عمده مدل های فیزیکی و مدل های ریاضی طبقه بندی می شوند (Chow et al., 1998). مدل های مفهومی بارش-رواناب از جمله مدل های ریاضی می باشند که مفاهیم سیستم را در شبیه سازی بارش-رواناب به کار می گیرند. از جمله مدل های بارش-رواناب می توان به مدل های Tank، HEC-HMS، SimHyd و HBV اشاره کرد. با توجه به اینکه مدل های مفهومی اغلب نیازمند شناخت کامل از عوامل اثرگذار بر روی جریان رودخانه هستند و همچنین به دلیل عدم قطعیت ذاتی این فرآیند، پیچیده بودن مدل های مفهومی و نیاز به اطلاعات وسیع باعث شده که محققان به

از مهم ترین مسائل مطرح در مدیریت حوضه های آبخیز، مدل سازی فرآیندهای هیدرولوژیکی می باشد. (Firat and Gungor, 2007). یکی از فرآیندهای حوضه آبخیز که مدل سازی آن نقش بسیار مهمی در ارزیابی و مدیریت منابع آب حوضه دارد، فرایند بارش-رواناب می باشد. برآورد رواناب تولید شده در هر حوضه، با دقت قابل قبول، بخشی مهم و جداناپذیر از اطلاعات لازم برای سیاست گذاری و مدیریت حوضه آبخیز می باشد. اطلاعات حاصل از مدل سازی بارش-رواناب و تخمین جریان آب در رودخانه ها در کارهای عمرانی، ساماندهی رودخانه ها، طراحی و برنامه ریزی منابع آب های سطحی و زیرزمینی و سامانه های هشدار سیل از اهمیت فراوانی برخوردار است. شبیه سازی رواناب ناشی از بارش ها این امکان را برای مدیران و برنامه ریزان فراهم می کند تا خسارات ناشی از وقوع سیلاب را کاهش داده و تدابیر خاصی را برای کنترل آن در نظر بگیرند (Fathabadi et al., 2009; Aronica and Candela, 2007). در این راستا، در دسترس بودن اطلاعات پیوسته بارش و سایر متغیرهای اقلیمی که می توانند برای مدل سازی مقدار رواناب مورد استفاده قرار گیرد، بسیار ضروری می باشد (Lorrai and Sechi, 1995). از سوی دیگر با

آب در دسترس سیستم‌های برق آبی کوچک در جزیره سرام اندونزی و استخراج منحنی تداوم جریان و تعیین دبی طراحی استفاده کرد. او در این زمینه بیان داشت که در صورت در دسترس بودن داده‌های مورد نیاز، HBV مدلی است که با داشتن ساختاری ساده و قابل فهم می‌تواند فرآیند بارش - رواناب را با دقتی قابل قبول شبیه‌سازی کند. (Artimani et al. (2019) به دلیل فقدان داده اندازه‌گیری شده برف در حوضه بوجین واقع در غرب استان همدان، از مدل SRM و HBV برای شبیه‌سازی ذوب برف استفاده نمودند. آنها به منظور پایش زمانی و مکانی سطح پوشش برف در مدل SRM، از داده‌های سنجنده MODIS در فواصل زمانی یک روزه استفاده کردند. با اجرای مدل‌های SRM و HBV شبیه‌سازی ذوب برف انجام شد. نتایج بیانگر دقت خوب مدل‌ها در شبیه‌سازی رواناب بود. همچنین مدل SRM به دلیل استفاده از تصاویر ماهواره‌ای، نسبت به مدل HBV عملکرد بهتری در شبیه‌سازی رواناب حاصل از ذوب برف از خود نشان داد. در زمینه استفاده از مدل HBV در شبیه‌سازی رواناب می‌توان به مطالعات Yaghoubi Nazariipooya et al. and Massah Bavani (2014) Mohammadvand et al. Maneshdavi et al. (2018)، (2015)، (2019)، (2020) Jaiswal et al. (2020) Ahmadi et al. و Moghaddam Nia et al. (2021) اشاره کرد.

در برخی از مطالعات از ترکیب مدل‌های هوشمند و مفهومی برای پیش بینی جریان رودخانه استفاده شده است. از جمله Ren et al. (2018) با استفاده از مدل HBV رواناب حاصل از ذوب برف و یخ در حوضه رودخانه Yarkant در کشور چین را محاسبه نمودند و آن را به عنوان ورودی جدید برای بهبود عملکرد مدل شبکه عصبی بیزین (BNN) استفاده کردند و سپس عملکرد این مدل را با مدل ترکیبی حاصل از ترکیب مدل HBV و ماشین بردار پشتیبان (HLSSVM) مقایسه کردند. نتایج نشان داد که مدل ترکیبی حاصل از مدل HBV و مدل شبکه عصبی بیزین (BNN) دارای عملکرد بهتری نسبت به مدل دیگر می‌باشد و این مدل یک مدل پیش‌بینی قوی جریان برای مناطق کوهستانی با رژیم برفی است. (Seyedian et al. (2019) از مدل‌های جعبه سیاه (نروفازی و ماشین بردار پشتیبان) و مدل‌های جعبه خاکستری (HBV و TOPMODEL) برای شبیه‌سازی فرایند بارش - رواناب در حوضه نوده خاندوز در استان گلستان استفاده کردند. مقایسه نتایج مدل‌های با مقادیر مشاهداتی نشان داد که مدل نروفازی با دبی تا سه گام زمانی قبل و دما با یک گام زمانی قبل عملکرد

استفاده از مدل‌های هوشمند روی آوردند تا با توجه به رفتار گذشته سیستم، رفتار آن را در زمان آینده پیش‌بینی نمایند. به این مدل‌ها، مدل‌های مبتنی بر داده یا آماری گفته می‌شود (El-Shafie et al., 2007) با استفاده از مدل‌های هوشمند می‌توان روابط ذاتی میان داده‌ها را که ریاضیات کلاسیک قادر به حل آنها نمی‌باشد، مدل‌سازی نمود. از جمله مدل‌های هوشمند می‌توان به مدل‌های شبکه‌های عصبی مصنوعی (ANNs)^۱، برنامه ریزی بیان ژن (GEP)^۲، ماشین بردار پشتیبان (SVM)^۳ و مدل‌های جدیدی همچون جنگل تصادفی (RF)^۴ اشاره کرد. تاکنون در مطالعات زیادی از مدل‌های هوشمند برای پیش‌بینی جریان رودخانه استفاده شده است که در ادامه به برخی از آنها اشاره می‌شود.

Pham et al. (2020) برای پیش‌بینی جریان ۸۶ حوضه آبخیز واقع در شمال غربی اقیانوس آرام، از الگوریتم جنگل تصادفی استفاده کردند و عملکرد این الگوریتم را با مدل‌های رگرسیون خطی چندگانه (MLR)^۵ و الگوریتم Naïve مورد مقایسه قرار دادند. نتایج نشان داد که بیانگر عملکرد مدل RF در حوضه‌های آبخیز با رژیم برفی بهتر از دو مدل دیگر بوده است. Hussain and Khan (2020) جهت پیش‌بینی جریان رودخانه هنزه در پاکستان از مدل‌های شبکه عصبی مصنوعی پرسپترون چندلایه^۶ (MLP)، رگرسیون بردار پشتیبان (SVR) و جنگل تصادفی (RF) استفاده نموده‌اند. مجموعه‌ای از پنج پارامتر به اضافه سری زمانی جریان رودخانه با تاخیرهای مختلف بعنوان ورودی مدل‌ها استفاده شد. مقایسه عملکرد مدل‌ها نشان داد که به ترتیب مدل‌های RF، MLP و SCR بالاترین عملکرد را در پیش‌بینی جریان رودخانه هنزه داشتند. همچنین نتایج نشان داد که مدل‌های مبتنی بر الگوریتم‌های یادگیری ماشینی به‌ویژه مدل RF، می‌توانند با دقت بسیار بالایی در پیش‌بینی جریان رودخانه به کار گرفته شوند. همچنین در زمینه استفاده از مدل‌های هوشمند برای پیش‌بینی جریان می‌توان به مطالعات Hassanzadeh et al. (2015)، Nikpour et al. (2017)، Partovyan et al. (2018)، Shafeizadeh, Adnan et al. (2018)، Zeinali et al. (2019)، Davanlou Tajbakhsh et al. (2019)، Mohammadi et al. (2020) و Ghorbani et al. (2020) اشاره کرد.

از سوی دیگر در بسیاری از مطالعات از مدل‌های مفهومی برای پیش‌بینی جریان رودخانه استفاده شده است. برای نمونه، Heidler (2015) از مدل HBV به منظور تجزیه و تحلیل مقدار

4- Random Forest

5- Multiple linear regression

6- Multilayer perceptron

1- Artificial Neural Network

2- Gene Expression Programming

3- Support Vector Machines

شمالی حوضه می‌باشد. متوسط ارتفاع این حوضه حدود ۲۱۱۵ متر از سطح دریا می‌باشد. رودخانه بازفت، از شاخه‌های مهم رودخانه کارون بوده که از دامنه‌های کوه‌های بلند تورک و منار واقع در ۸۰ کیلومتری جنوب غربی فریدن سرچشمه می‌گیرد. طول این رودخانه حدود ۱۶۰ کیلومتر بوده و حوضه آبریز آن در مناطق مرتفع، برفگیر و کوهستانی واقع گردیده است و دارای جریان آب دائمی است. دبی متوسط درازمدت این رودخانه در ایستگاه مرغک حدود ۶۹/۱ متر مکعب بر ثانیه می‌باشد. این حوضه از شمال غرب به حوضه آبریز دز، از شمال، شمال شرق و شرق به حوضه آبریز کوه‌رنگ و از غرب و جنوب غرب به حوضه آبریز کارون منتهی می‌شود. شکل ۱ موقعیت حوضه آبخیز بازفت را نشان می‌دهد.

داده‌های مورد استفاده برای مدل سازی در این پژوهش، شامل داده‌های جریان رودخانه بازفت (روزانه)، باران (روزانه)، دمای میانگین (روزانه) و تبخیر- تعرق پتانسیل (ماهانه و روزانه) در یک دوره شش ساله از ۱ سپتامبر ۲۰۱۰ تا ۲۲ آگوست ۲۰۱۷ می‌باشد. همچنین برای محاسبه تبخیر- تعرق پتانسیل برای منطقه مورد مطالعه به داده‌های حداقل و حداکثر دما (درجه سانتی‌گراد)، رطوبت (درصد)، سرعت باد (متر بر ثانیه) و ساعات آفتابی (ساعت) نیاز است که این اطلاعات از ایستگاه سینوپتیک کوه‌رنگ به دلیل مجاورت و مشابه بودن از نظر اقلیمی و ارتفاعی با حوضه بازفت، مورد استفاده قرار گرفت. در جدول (۱) خلاصه ای از مشخصات آماری داده‌های مورد استفاده در این مطالعه ارائه شده است. در این مطالعه، از نرم افزار HBV 4.0.0.10 برای ایجاد مدل بارش- رواناب و نرم افزار WEKA®3.9.2 برای اجرای مدل جنگل تصادفی استفاده شد.

بهتری نسبت به سایر سناریوها دارد. همچنین نتایج نشان داد که به طور کلی مدل‌های جعبه سیاه، رواناب را در مرحله واسنجی و صحت‌سنجی با دقت بیشتری نسبت به مدل‌های HBV و TOPMODEL شبیه‌سازی کرده‌اند. در این زمینه، Fathabadi *et al.* (2009) و Ahmadpour *et al.* (2017) نیز از ترکیب مدل‌های هوشمند و مفهومی برای پیش بینی جریان رودخانه استفاده کرده‌اند.

هدف از این مطالعه، شبیه‌سازی فرایند بارش- رواناب در حوضه بازفت در استان چهارمحال و بختیاری با استفاده از مدل یکپارچه مفهومی^۱ HBV و مدل جنگل تصادفی و مقایسه عملکرد این دو مدل است. در واقع سؤال اساسی این تحقیق این است که آیا برای شبیه‌سازی فرایند بارش- رواناب، در شرایط کمبود داده و اطلاعات، مدل جنگل تصادفی می‌تواند جایگزین مناسبی برای مدل مفهومی باشد.

مواد و روش‌ها

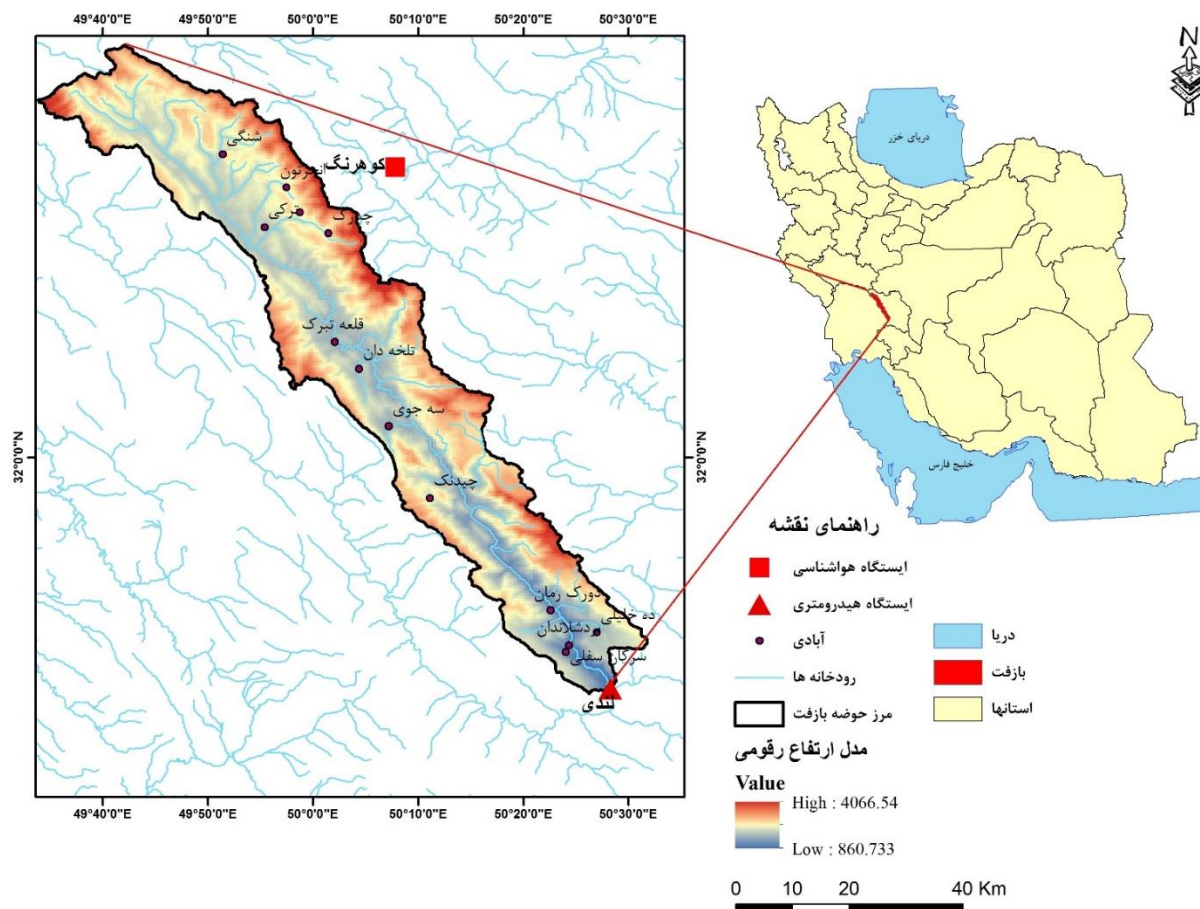
خصوصیات حوضه آبخیز بازفت

حوضه آبخیز بازفت یکی از زیرحوضه‌های آبریز کارون با وسعت ۲۹۷۳ کیلومتر مربع در غرب و شمال غرب استان چهارمحال و بختیاری و در دامنه‌های غربی زاگرس بین طول جغرافیایی ۳۳° ۳۳' تا ۴۹° ۲۹' ۳۰" شرقی و عرض جغرافیایی ۳۷° ۵۶' تا ۳۱° ۳۹' ۱۶" شمالی واقع شده است. این حوضه کاملاً کوهستانی و مرتفع بوده بطوریکه ۲۹۲۴ کیلومتر مربع از وسعت آن را ارتفاعات تشکیل می‌دهد. حداقل ارتفاع در حوضه ۸۶۱ متر و حداکثر آن ۴۰۹۱ متر از سطح دریا مربوط به قسمت‌های

جدول ۱- خلاصه آماری متغیرهای مورد استفاده در این مطالعه

متغیر	میانگین	حداقل	حداکثر	واریانس	چولگی	کشیدگی
دبی (m ³ /s)	۱/۸۲۳	۰/۲۵۰	۵۲/۳۹۷	۵/۸۶۶	۷/۹۴۰	۱۱۵/۶۱۹
دمای حداکثر (°C)	۱۷/۵۶۱	-۵/۲۴۲	۳۵/۶۸۸	۱۰۴/۷۵۵	-۰/۰۳۳	-۱/۳۴۲
دمای حداقل (°C)	۵/۰۹۹	-۱۵/۵۸۰	۲۱/۳۱۹	۷۵/۷۰۷	-۰/۰۷۹	-۱/۲۴۸
دمای متوسط (°C)	۱۳/۶۶۴	-۷/۴۳۳	۳۰/۶۰۲	۹۳/۸۵۷	-۰/۰۱۵	-۱/۳۵۹
سرعت باد (m/s)	۲/۸۳۳	۰/۱۰۰	۱۱/۹۸۰	۱/۳۷۱	۱/۳۰۱	۴/۰۷۱
رطوبت نسبی (%)	۴۱/۳۰۰	۱۹/۲۷۹	۸۷/۰۵۶	۲۵۶/۰۹۶	۰/۶۲۳	-۰/۵۸۵
بارش (mm)	۱/۲۰۷	۰/۰۰	۹۸	۲۸/۸۹۰	۷/۸۱۸	۸۵/۱۶۰
تبخیر- تعرق* (mm)	۴/۳۳۴	۰/۵۵۲	۹/۸۱۷	۶/۵۹۵	۰/۲۹۵	-۱/۳۱۳

* تبخیر- تعرق محاسبه شده با روش هارگریوز- سامانی



شکل ۱- موقعیت جغرافیایی حوضه آبخیز بازفت و ایستگاه هیدرومتری لندی

مدل بارش رواناب HBV

مدل^۱ HBV یک مدل مفهومی است که در سال‌های اخیر در مطالعات بارش-رواناب حوضه‌ها در کشورهای اسکانندیناوی و بسیاری نقاط دیگر جهان مورد استفاده قرار گرفته است. این مدل در موسسه هواشناسی و هیدرولوژی سوئد و در سال‌های آغازین دهه ۱۹۷۰ طرح ریزی و توسعه داده شده است. HBV دارای نسخه‌های مختلفی است که جدیدترین نسخه آن HBV-light در سال ۱۹۹۳ در دانشگاه اوپسالا^۲ تهیه گردیده است. جدیدترین تصحیح مدل HBV-light در دانشگاه زوریخ^۳ ارائه گردیده که قابلیت شبیه سازی با گام‌های زمانی مختلف و نیز استفاده همزمان در چند زیرحوضه را به توانایی‌های مدل افزوده‌اند. همچنین بخشی برای شبیه‌سازی یخچال‌های طبیعی نیز در مدل پیش بینی شده است (Konz and Seibert, 2010). ساختار این مدل طبق شکل ۲ شامل مراحل توزیعی روال برف^۴، روال خاک، مرحله مفهومی عکس العمل حوضه (فرآیند هیدرولوژیکی تولید رواناب) و روندیابی حوضه است (Driessen et al., 2010).

۱- روال برف: در این بخش، بارش در دمای کمتر از دمای آستانه تعیین شده (TT)، به صورت برف لحاظ می‌شود. محاسبه ذوب برف به روش ساده درجه-روز طبق روابط زیر می‌باشد:

(رابطه ۱) $M = P_{CFMAX}(T(t) - TT)$

(رابطه ۲) $R = P_{CFR} \cdot P_{CFMAX}(P_{TT} - T(t))$

در روابط فوق، CFMAX فاکتور درجه-روز، TT آستانه دما، T(t) میانگین دمای روزانه، CFR ضریب تصحیح انجماد آب حاصل از برف ذوب شده است.

۲- روال خاک: در این قسمت تغذیه آب زیرزمینی و تبخیر واقعی به صورت تابعی از ذخیره واقعی آب شبیه‌سازی می‌شود. این قسمت اصلی‌ترین بخش کنترل رواناب است و دارای ۵ پارامتر به صورت زیر است:

$$\frac{F(t)}{I(t)} = \left(\frac{S_{SOIL}(t)}{P_{FC}} \right)^{P_{BETA}} \quad \text{(رابطه ۳)}$$

که در آن، BETA پارامتری است که سهم نسبی در رواناب (تابع عکس العمل یا افزایش رطوبت خاک) از هر میلی‌متر بارش یا ذوب برف را کنترل می‌کند. FC حداکثر ظرفیت ذخیره رطوبت در خاک و LP پارامتر نشانگر محدوده پتانسیل تبخیر است. LP

3- University of Zurich
4- Snow routine

1- Hydrologiska Byrans Vattenavdelning
2- Uppsala University

پارامتر MAXBAS برای روندیابی جریان محاسبه شده در خروجی استفاده می‌کند.

$$Q_{sim}(t) = \sum_{i=1}^{P_{MAXBAS}} C(i) \cdot Q(t - i + 1) \quad (\text{رابطه ۴})$$

که در آن:

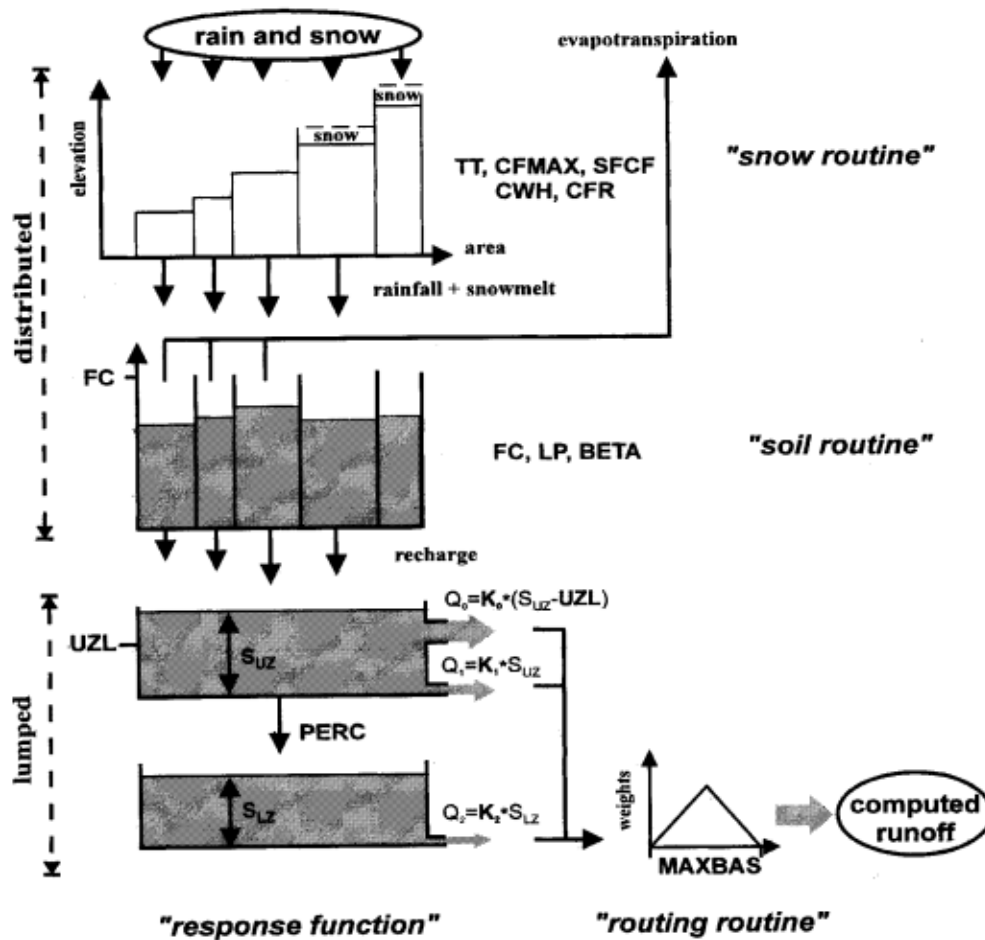
(رابطه ۵)

$$C(i) = \int_{i-1}^i \frac{2}{P_{MAXBAS}} - \left| u - \frac{P_{MAXBAS}}{2} \right| * \frac{4}{P_{MAXBAS}^2} du.$$

مقدار رطوبت خاک است که بالاتر از آن تبخیر-تعرق به مقدار پتانسیل می‌رسد. این پارامتر به صورت درصدی از FC است. همچنین ذخیره واقعی خاک، $I(t)$ مجموع بارش و برف روزانه و $F(t)$ دبی خروجی است.

۳- تابع عکس العمل: در این بخش رواناب محاسبه شده تبدیل به دبی در خروجی زیرحوضه می‌شود.

۴-روال روندیابی: این مدل در نهایت از تابع وزنی مثلثی با



شکل ۲- نمای شماتیک از ساختار کلی مدل (Seibert, 2000) HBV

تصادفی قرار می‌گیرد که هر درخت یک طبقه‌بندی تولید می‌کند و گفته می‌شود این درخت به آن کلاس (رای) می‌دهد. جنگل حاصل از طبقه‌بندی که بیشترین رای را داشته باشد، (بین همه درخت‌های جنگل) انتخاب می‌شود.

الگوریتم جنگل تصادفی (RF) مبتنی بر دسته‌ای از درخت-های تصمیم است و در حال حاضر یکی از بهترین الگوریتم‌های یادگیری است.

هر درخت به صورت زیر تشکیل می‌شود:

۱. اگر N تعداد حالت‌ها در مجموعه داده‌های train (مجموعه کار) باشد، N حالت را به صورت تصادفی با جایگذاری از داده‌های اصلی، نمونه‌گیری می‌شود. این نمونه مجموعه کار

الگوریتم جنگل تصادفی

روش‌های درخت-پایه، روش‌های آماری ناپارامتری (بدون مدل) برای اجرای آنالیز کلاس بندی و آنالیز رگرسیونی با استفاده از الگوریتم افرازهای بازگشتی می‌باشند (Breiman, 1984). جنگل‌های تصادفی یک نوع مدرن از روش‌های درخت-پایه هستند که شامل انبوهی از درخت‌های کلاس بندی و رگرسیونی می‌باشند (Breiman, 2001). مهمترین ویژگی جنگل تصادفی عملکرد بالای آنها در اندازه‌گیری اهمیت متغیرها برای مشخص کردن این است که هر متغیر چه نقشی در پیش‌بینی پاسخ دارد. جنگل تصادفی درخت‌های تصمیم زیادی تولید می‌کند. برای طبقه‌بندی یک شیء جدید، بردار ورودی در انتهای هر یک از درختان جنگل

استفاده می‌شود. در هر درخت، ابتدا اندازه ناخالصی روی مشاهدات خارج از کیسه محاسبه می‌شود. سپس مقادیر متغیر X_i مشاهدات خارج از کیسه به طور تصادفی جابجا می‌شوند و اندازه ناخالصی درخت روی مقادیر جابجا شده محاسبه می‌شود. اندازه اهمیت متغیر X_i در هر درخت، اختلاف بین این دو اندازه ناخالصی است و میانگین این مقادیر شاخص اهمیت جایگشتی است. انگیزه این روش این است که اگر X_i متغیر مهمی باشد، جابجا شدن مقادیر آن به طور تصادفی منجر به افزایش ناخالصی درخت می‌شود، درحالی که اگر متغیر تأثیرگذاری نباشد، تغییری در ناخالصی ایجاد نمی‌شود.

مزیت‌های جنگل تصادفی

- در میان الگوریتم‌های فعلی از نظر دقت کم‌نظیر است.
- روی داده‌های بسیار بزرگ قابل اجراست.
- می‌تواند هزاران متغیر را بدون حذف متغیرها مدیریت کند.
- برآوردی از مهم‌ترین متغیرها در طبقه بندی می‌دهد.
- راهکارهایی برای برآورد داده‌های گم شده دارد.

معیارهای ارزیابی

برای ارزیابی مدل‌ها و مقایسه کارایی آنها، در این پژوهش از معیارهای ضریب همبستگی (r)، جذر میانگین مربعات خطا^۶ ($RMSE$)، میانگین مطلق خطا^۷ (MAE)، درصد میانگین مطلق خطاهای پیش‌بینی^۸ ($MAPE$) و ضریب کارایی نش- ساتکلیف^۹ (NS) استفاده می‌شود. در ادامه روابط این معیارها ارائه شده است:

$$r = \frac{\left(\sum_{i=1}^N \Sigma(Q_{obs} - \bar{Q}_{obs})(Q_{pre} - \bar{Q}_{pre}) \right)^2}{\sum_{i=1}^N (Q_{obs} - \bar{Q}_{obs})^2 \sum_{i=1}^N (Q_{pre} - \bar{Q}_{pre})^2} \quad (\text{رابطه ۶})$$

$$RMSE = \sqrt{\frac{\sum (Q_{obs} - Q_{pre})^2}{n}} \quad (\text{رابطه ۷})$$

$$NS = 1 - \frac{\sum_{i=1}^n (Q_{obs} - Q_{pre})^2}{\sum_{i=1}^n (Q_{obs} - \bar{Q})^2} \quad (\text{رابطه ۸})$$

$$MAE = \frac{\sum_{i=1}^n |Q_{obs} - Q_{pre}|}{n} \quad (\text{رابطه ۹})$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Q_{obs} - Q_{pre}}{Q_{obs}} \right| * 100 \quad (\text{رابطه ۱۰})$$

در روابط فوق، n تعداد داده‌ها، Q_{obs} داده‌های مشاهداتی، \bar{Q} میانگین داده‌های مشاهداتی و Q_{pre} داده‌های محاسباتی می‌باشد. هر چقدر مقدار R به عدد یک نزدیک‌تر باشد، برآورد و برآزش بهتر را نشان می‌دهد. ضریب کارایی نش- ساتکلیف می

برای این درخت می‌باشد.

۲. اگر M متغیر داشته باشیم و m را کوچکتر از M در نظر بگیریم، به طوری که در هر گره، m متغیر به صورت تصادفی از M انتخاب می‌شوند و بهترین جداسازی روی این m متغیر برای جداسازی گره استفاده می‌شود. مقدار m در طول ساخت جنگل ثابت در نظر گرفته می‌شود.

۳. هر درخت به اندازه ممکن بزرگ می‌شود. هیچ هرسی وجود ندارد.

نرخ خطای جنگل به دو مورد زیر بستگی دارد:

الف- همبستگی بین هر دو درخت در جنگل: افزایش همبستگی نرخ خطای جنگل را افزایش می‌دهد.

ب- قدرت هر یک از درختان در جنگل: هر درخت با نرخ خطای کم، یک طبقه‌بند قوی است. افزایش قدرت هر یک از درختان نرخ خطای جنگل را کاهش می‌دهد.

یک جنگل تصادفی آنقدر بزرگ است که تفسیر آن کار بسیار دشواری است، لذا نیاز است که اطلاعات آن با استفاده از شاخص‌های کمی خلاصه شود. یکی از این شاخص‌ها، شاخص اهمیت متغیر^۱ (VI) است.

شاخص اهمیت متغیر، شاخصی برای رتبه‌بندی متغیرها بر حسب اهمیت آن‌ها در اثرگذاری روی پاسخ است. معروف‌ترین شاخص‌های اهمیت متغیر، شاخص اهمیت جینی^۲ و شاخص اهمیت جایگشتی^۳ است.

شاخص اهمیت جینی: در طی ساخت درخت‌های جنگل تصادفی برای تعیین اینکه گره بر اساس کدام متغیر افزاز شود، از شاخص ناخالصی جینی استفاده می‌شود (Breiman, 1984)

اهمیت متغیر X_i در یک درخت، مجموع کاهش در شاخص ناخالصی جینی روی تمام گره‌هایی است که بر اساس X_i افزاز شده اند. میانگین اندازه اهمیت متغیر X_i روی تمام درخت‌های جنگل، اندازه شاخص اهمیت جینی است. شاخص اهمیت جایگشتی: برای محاسبه این شاخص، الگوریتم جنگل تصادفی از تمام مشاهدات نمونه، برای ساخت درخت استفاده نمی‌کند، بلکه یک نمونه تصادفی با جایگذاری به حجم $n1$ (معمولاً دوسوم داده‌ها) برابر از مشاهدات انتخاب می‌شود. به مشاهدات انتخاب شده نمونه آموزشی^۴ (LS) گفته می‌شود و به بقیه آن‌ها نمونه خارج کیسه^۵ گفته می‌شود. درخت‌ها با مشاهدات نمونه آموزشی ساخته می‌شود و از نمونه خارج از کیسه برای اندازه‌گیری ناخالصی درخت

6- Root Mean Squared Error (RMSE)

7- Mean Absolute Error (MAE)

8- Mean Absolute Percentage Error (MAPE)

9- Nash-Sutcliffe efficiency coefficient

1- Variable importance

2- Gini importance index

3- Permutation importance index

4- Learning Sample (LS)

5- off-bag (OOB)

در این مطالعه، جریان روزانه رودخانه بازفت در یک دوره ۷ ساله از ۱ سپتامبر ۲۰۱۰ تا ۲۲ آگوست ۲۰۱۷ در نظر گرفته شد. برای آماده سازی مدل، داده‌ها از ۱ سپتامبر ۲۰۱۰ تا ۳۱ دسامبر ۲۰۱۱ مورد استفاده قرار گرفت. سپس مدل طی دوره آماری ۲۰۱۲/۱/۱ تا ۲۰۱۶/۷/۶ واسنجی شد و داده‌ها از ۲۰۱۷/۱/۱ تا ۲۰۱۷/۸/۲۲ برای مرحله صحت‌سنجی مدل مورد استفاده قرار گرفت. پارامترهای بهینه مدل بارش-رواناب HBV برای منطقه مورد مطالعه و نتایج مدل در مرحله صحت‌سنجی به ترتیب در جداول ۲ و ۳ ارائه شده است.

همانگونه که در جدول (۳) مشاهده می‌شود مقدار ضریب همبستگی (r) مدل HBV در مرحله صحت‌سنجی برابر با ۰/۸۷۲ بوده و که نشان‌دهنده دقت قابل قبول این مدل در برآورد رواناب می‌باشد. همچنین مقدار RMSE بین رواناب مشاهده‌ای و محاسباتی توسط مدل HBV برابر با ۰/۵ متر مکعب بر ثانیه است که مقدار قابل قبولی است. همچنین مقدار معیار NS برابر با ۰/۶۳۱۴ بدست آمده که نشان دهنده کارایی خوب این مدل در برآورد رواناب می‌باشد. بر اساس مطالعات (Santhi et al., 2001) و همچنین (Binaman and Shoemake, 2005). هنگامی شبیه-سازی توسط مدل HBV رضایت بخش است که ضریب همبستگی (r) بیشتر از ۰/۸ باشد. همچنین (Radchenko et al., 2014) بیان کردند که مقدار ضریب نش-ساتکلیف (NS) بیشتر از ۰/۵ برای مدل HBV قابل قبول است. بنابراین، با در نظر گرفتن ضرایب آماری ارائه شده در جدول (۳)، مدل HBV استفاده شده در این مطالعه عملکرد خوبی در شبیه‌سازی رواناب در حوضه بازفت داشته است.

تواند از $-\infty$ تا ۱ تغییر کند، هرچه مقادیر به عدد یک نزدیک‌تر باشد، مدل برازش بهتری دارد و عدد یک برازش عالی را نشان می‌دهد. $RMSE$ جذر میانگین مربع خطاهای داده‌های محاسباتی و مشاهداتی را بیان می‌کند. واضح است که هر چه مقدار این عدد کمتر باشد، عملکرد مدل در شبیه‌سازی داده‌ها بهتر است.

نتایج و بحث

محاسبه تبخیر-تعرق پتانسیل با استفاده از روش هارگریوز-سامانی اصلاح شده (HS)

در این مطالعه، تبخیر-تعرق با روش هارگریوز-سامانی (HS) محاسبه و به عنوان ورودی مدل‌ها مورد استفاده قرار گرفت. بدین منظور، پارامترهای معادله هارگریوز-سامانی برای منطقه مورد مطالعه بر اساس مقادیر تبخیر-تعرق پتانسیل محاسبه شده توسط روش فانو-پنمن -مانتیت که به عنوان یک روش مرجع است، اصلاح گردید و رابطه اصلاحی به صورت زیر بدست آمد:

(رابطه ۱)

$$ET_0 = 0.0031 R_a (T_{mean} + 15.68)(T_{max} - T_{min})^{0.5}$$

نتایج مدل‌سازی رواناب با استفاده از مدل HBV-Light

در این مطالعه، برای ارزیابی عملکرد مدل HBV در مدل‌سازی بارش-رواناب، داده‌های بارندگی روزانه، جریان روزانه رودخانه، دمای متوسط روزانه و تبخیر-تعرق ماهانه وارد مدل گردید. مقادیر تبخیر-تعرق ماهانه محاسبه شده توسط روش هارگریوز-سامانی اصلاحی برای منطقه مورد مطالعه همراه با بقیه داده‌ها، وارد مدل شد و نتایج مدل HBV که با استفاده از الگوریتم ژنتیک بهینه شده بود، مورد بررسی قرار گرفت.

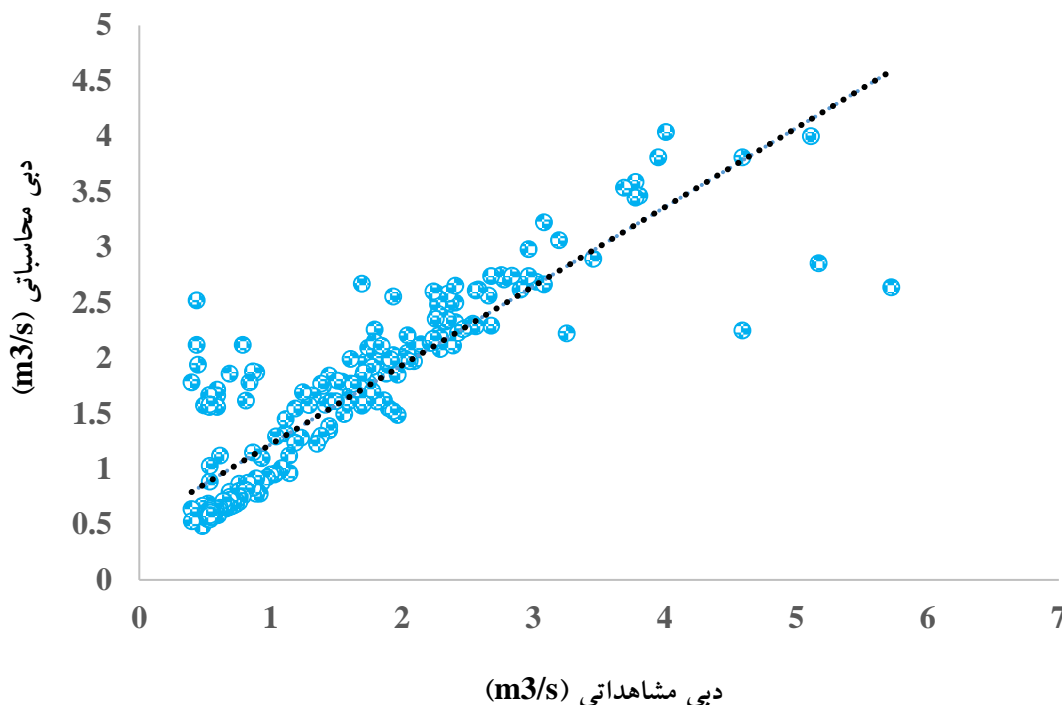
جدول ۲- پارامترهای واسنجی شده مدل HBV-Light در طول دوره آماری ۲۰۱۶-۲۰۱۲

پارامتر	توضیحات	واحد	حد بالا	حد پایین	ET-HS
TT	آستانه دمای	°C	۰/۵	-۲/۵	-۳
CFMAX	فاکتور درجه-روز	d ⁻¹ mm°C	۴	۰/۵	۱
SFCF	فاکتور تصحیح برف	-	۰/۹	۰/۵	۰/۰۱
CFR	ضریب انجماد مجدد	-	۰/۰۵	۰/۰۵	۰/۰۵
CWH	ضریب ذخیره آب	-	۰/۱	۰/۰۰۰۱	۰
FC	بیشینه ذخیره در خاک	mm	۵۰۰	۰	۶۸۰
LP	آستانه کاهش تبخیر	-	۱	۰/۱	۰/۴۶
BETA	ضریب ذخیره و نفوذ عمقی خاک	-	۵	۱	۲/۲
K0	ضریب فروکش (بالای مخزن پایینی)	d ⁻¹	۰/۵	۰/۱	۰/۶
K1	ضریب فروکش (پایین مخزن بالایی)	d ⁻¹	۰/۲	۰/۰۱	۰/۰۹
K2	ضریب فروکش (مخزن پایینی)	d ⁻¹	۰/۱	۵	۰/۰۵
UZL	آستانه خروجی از مخزن پایینی	mm	۷۰	۰	۱۴
PERC	جریان بیشینه از لایه بالایی به پایین	mmd ⁻¹	۴	۰	۵/۴
MAXBAS	روندپایی، طول تابع وزنی	d	۲/۵	۱	۱

جدول ۳- نتایج ارزیابی مدل HBV-Light در مرحله صحت‌سنجی

r	RMSE (m ³ /s)	NS
۰/۸۷۲	۰/۵۰۲۶	۰/۶۳۱۴

در شکل (۳) نمودار پراکنش مقادیر مشاهداتی در ایستگاه لندی در برابر مقادیر شبیه‌سازی شده دبی‌های روزانه توسط مدل HBV برای داده‌های بخش صحت‌سنجی نشان داده شد.



شکل ۳- نمودار پراکنش مقادیر مشاهداتی و شبیه‌سازی شده دبی‌های روزانه در ایستگاه هیدرومتری لندی با استفاده از مدل HBV

گرفته و این امر دلالت بر تطابق خوب مقادیر مشاهداتی و شبیه‌سازی شده بر اساس خط $(Y=X)$ است. مقادیر پنج معیار ارزیابی برای ارزیابی کمی عملکرد مدل RF در ایستگاه هیدرومتری لندی نیز در جدول (۵) ارائه شده است. با توجه به نتایج جدول (۵) مشخص می‌شود که مدل با دقت خوبی توانسته دبی جریان روزانه را در ایستگاه لندی شبیه‌سازی کند و می‌توان گفت که مدل RF عملکرد خوبی در برآورد جریان روزانه در ایستگاه هیدرومتری لندی داشته است.

جدول ۵- نتایج مرحله صحت‌سنجی مدل جنگل تصادفی برای دبی‌های

روزانه ایستگاه هیدرومتری لندی				
MAE	MAPE (%)	RMSE (m ³ /s)	r	NS
۰/۲۵۰۸	۹/۵۹۷	۰/۳۹۹۲	۰/۹۵۹	۰/۸۲۳

در شکل (۵) تغییرات مقادیر شبیه‌سازی شده و مشاهداتی دبی‌های روزانه نسبت به زمان برای مدل جنگل تصادفی (RF) نشان داده شده است. همانطور که مشاهده می‌گردد اگر چه در کل تطابق مناسبی میان مقادیر مشاهده‌ای و شبیه‌سازی شده جریان وجود دارد، ولی عملکرد این مدل در برآورد مقادیر اوج جریان نسبتاً ضعیف بوده است.

نتایج مدل‌سازی رواناب با استفاده از مدل جنگل تصادفی (RF) در این مطالعه، برای مدل‌سازی رواناب در حوضه بازفت با استفاده از مدل RF، داده‌های بارندگی روزانه، جریان روزانه رودخانه، دمای متوسط روزانه و تبخیر-تعرق روزانه برآورد شده با استفاده از روش HS مورد استفاده قرار گرفت. بدین منظور، پارامترهای مدل جنگل تصادفی در نرم افزار وکا (WEKA) به صورت جدول (۴) تنظیم شد.

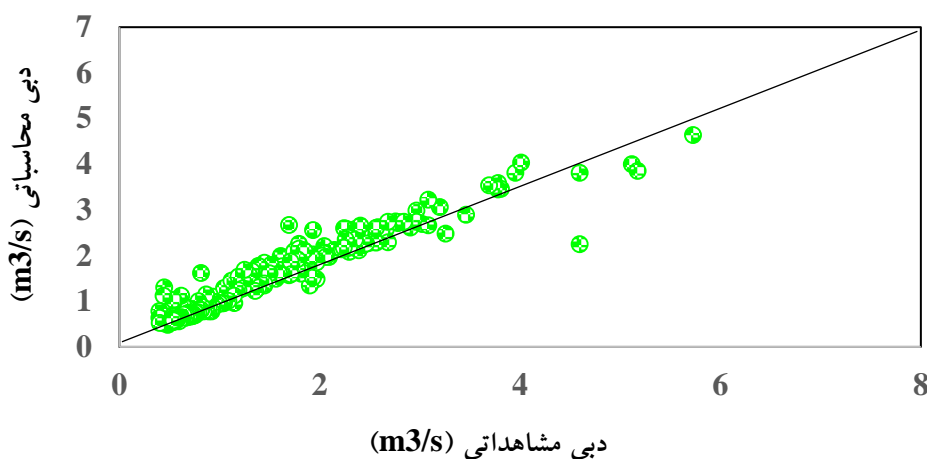
جدول ۴- مقادیر پارامترهای مورد استفاده در مدل جنگل تصادفی

پارامتر	مقدار بهینه
Bag Size Percent	۱۰۰
Batch Size	۱۰۰
Maximum Depth of Tree	۰
Number of Execution Slots	۱
Seed	۱

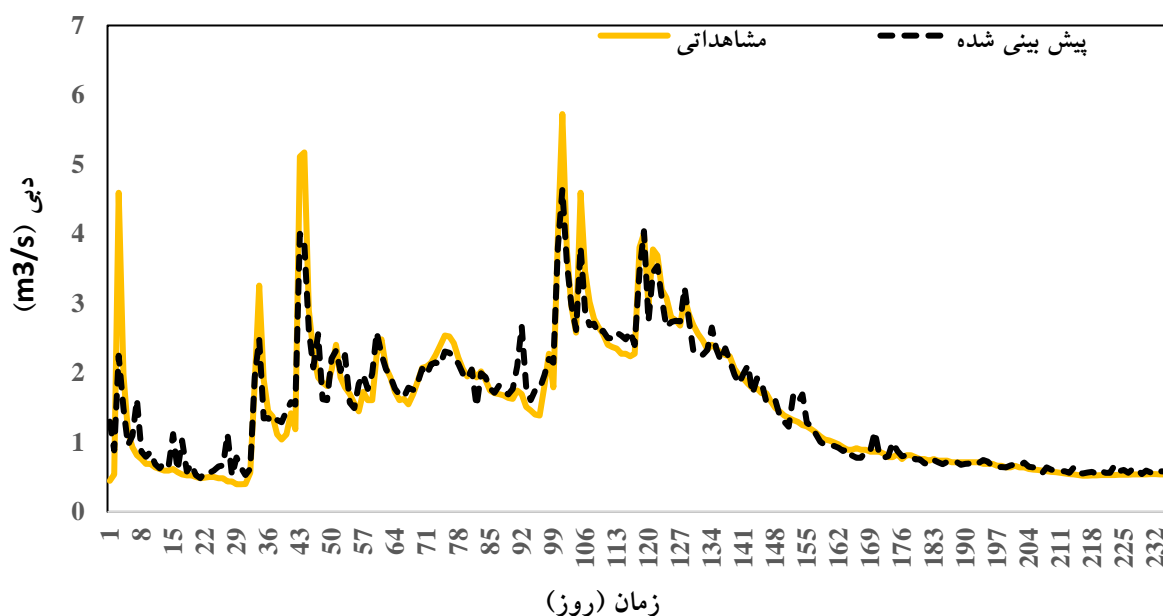
در شکل (۴) نمودار پراکنش مقادیر مشاهداتی در برابر مقادیر شبیه‌سازی شده دبی‌های روزانه توسط مدل جنگل تصادفی برای داده‌های بخش صحت‌سنجی نشان داده شد. همانطور که از شکل (۴) مشخص است، بیشتر مقادیر شبیه‌سازی شده و مشاهداتی، به غیر از چند نقطه روی خط نیمساز قرار

مدل جنگل تصادفی بیشتر خطاها در باند $\pm 0/5$ (میلی متر بر ثانیه) قرار گرفته است و بیشترین مقدار خطا برابر $2/97$ (میلی متر بر ثانیه) می باشد که نشان دهنده کارایی مناسب این مدل در برآورد رواناب روزانه می باشد.

در شکل (۶) اختلاف مقادیر مشاهداتی و شبیه سازی شده دبی روزانه (مقدار خطا) در مرحله صحت سنجی توسط مدل جنگل تصادفی محاسبه و نمودار آن نسبت به دوره آماری ترسیم شده است. همانطور که در این شکل نشان داده شده است، برای



شکل ۴- نمودار پراکنش مقادیر مشاهداتی و شبیه سازی شده دبی های روزانه در ایستگاه هیدرومتری لندی با استفاده از مدل جنگل تصادفی.

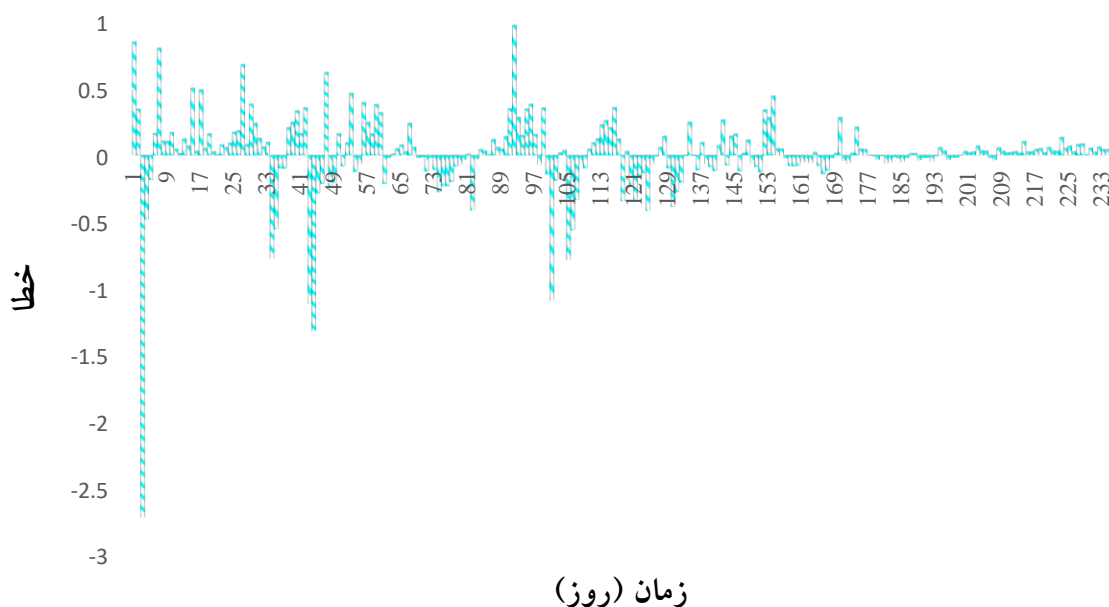


شکل ۵- سری های زمانی رواناب مشاهداتی و شبیه سازی شده توسط مدل جنگل تصادفی در ایستگاه هیدرومتری لندی.

نتایج مقایسه مدل ها

در شبیه سازی جریان روزانه بهتر از مدل HBV بوده است، بطوریکه آماره های NS و r مدل درخت تصادفی بالاتر از مقادیر متناظر برای مدل HBV بوده اند. همچنین مقادیر خطای مدل RF کمتر از مدل HBV بوده است.

در جدول (۶) مقایسه ای بین عملکرد مدل HBV و مدل RF در برآورد جریان روزانه در ایستگاه هیدرومتری لندی بر اساس معیارهای آماری (r , MAE, RMSE, NS و MAPE) انجام گرفته است. همانگونه که در این جدول مشخص است، عملکرد مدل RF



شکل ۶- نمودار خطای مدل جنگل تصادفی در مرحله صحت‌سنجی در ایستگاه هیدرومتری لندی.

جدول ۶- مقایسه نتایج مدل‌های HBV و جنگل تصادفی در برآورد رواناب در ایستگاه لندی

مدل	NS	RMSE (m ³ /s)	MAE	r	MAPE (%)
HBV-HS	۰/۶۳۱۴	۰/۵۰۲۶	۰/۳۵۴۹	۰/۸۷۲	۱۸/۰۸
RF	۰/۸۲۳۰	۰/۳۲۹۲	۰/۲۵۰۸	۰/۹۵۹	۹/۵۹۷

نتیجه‌گیری

با مدل RF حدود ۲۱ درصد بیشتر از مدل HBV بود. همچنین نتایج نشان داد که اگر چه دقت مدل جنگل‌های تصادفی در برآورد مقادیر دبی پایه و متوسط رواناب روزانه خوب بود، ولی عملکرد این مدل در برآورد مقادیر اوج جریان نسبتاً ضعیف بوده است که با توجه به رژیم برفی حوضه، بنظر می‌رسد به دلیل اثر رواناب حاصل از ذوب برف در دربی‌های اوج باشد که در مدل جنگل تصادفی لحاظ نشده است. لذا پیشنهاد می‌شود در مطالعات آتی، مقادیر رواناب حاصل از ذوب برف نیز در مدل‌سازی رواناب در نظر گرفته شود.

"هیچ‌گونه تعارض منافع توسط نویسندگان وجود ندارد"

در این مطالعه، جهت پیش‌بینی جریان روزانه رودخانه بازفت در ایستگاه هیدرومتری لندی، از مدل مفهومی هیدرولوژیکی HBV و مدل هوش مصنوعی جنگل تصادفی (RF) استفاده گردید. مقدار تبخیر- تعرق با روش هارگریوز- سامانی اصلاح شده محاسبه و بعنوان ورودی به این مدل‌ها وارد شد. سپس مقادیر رواناب روزانه مشاهداتی با رواناب روزانه تخمین زده شده توسط هر کدام از مدل‌ها با استفاده از معیارهای ارزیابی مورد بررسی گرفت. نتایج نشان داد که هر دو مدل از عملکرد قابل قبولی در برآورد رواناب روزانه برخوردار بودند. مقایسه نتایج دو مدل نشان داد که مدل RF عملکرد بهتری نسبت به مدل HBV داشت و دقت مدل‌سازی

REFERENCES

- Adnan, R.M., Yuan, X., Kisi, O., Adnan, F., and Mehmood, A. (2018). Stream flow forecasting of poorly gauged mountainous watershed by Least Square Support Vector Machine, Fuzzy Genetic Algorithm and M5 Model Tree using climatic data from Nearby Station. *Water Resources Management*, 32: 4469- 4486.
- Ahmadi, M., Dadashi Roudbari, A., and Deyrmajai, A. (2020). Runoff estimation using IHACRES model based on CHIRPS satellite data and CMIP5 models (Case study: Gorganroud Basin- Aq Qala Area. *Iranian Journal of Soil and Water Research*, 51(3): 659-671. (In Farsi)
- Ahmadpour, A., Mirhashemi, S., and Haghghatjou, P. (2020). Evaluation of classical, conceptual IHACRES and hybrid ARMA-ANN models in simulation and prediction of daily discharge of Maroun River. *Iranian Journal of Soil and Water Research*, 51(3): 727-736. (In Farsi)
- Aronica, G.T. and Candela, A. (2007). Derivation of flood frequency curves in poorly gauged Mediterranean catchments using a simple stochastic hydrological rainfall-runoff model. *Journal of Hydrology*, 347: 132-142.

- Artimani, M., Zeinivand, H., and Tahmasebipour, N. (2019). Performance evaluation of SRM and HBV model in simulation of snowmelt runoff in Bujin Basin. *Iran-Water Resources Research*, 15(2): 228-241. (In Farsi)
- Binaman J., and Shoemaker C.A. (2005). An analysis of high-flow sediment event data for evaluating model performance. *Journal of Hydrological Processes*, 19: 605-620.
- Breiman, L. (1984). Classification and regression trees CA, Wadsworth International Groups. *Handbook of data mining and knowledge discovery*. (PP. 248-276).
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1): 5-32.
- Chang, T.K., Talei, A., Alaghmand, S., and Ooi, M. P.-L. (2017). Choice of rainfall inputs for event-based rainfall-runoff modeling in a catchment with multiple rainfall stations using data-driven techniques. *Journal of Hydrology*, 545: 100-108
- Chow, V.T., Maidment, D.R., and Mays, L.W. (1988). *Applied Hydrology*. McGraw-Hill Series in Water Resources and Environmental Engineering. McGraw-Hill: New York. PP 572.
- Davanlou Tajbakhsh, A., Nourani, V., and Molajou, A. (2019). Hybrid Wavelet-M5 modeling in rainfall-runoff process forecast. *Iran-Water Resources Research*, 15(2): 1-10 (In Farsi)
- Driessen, T.L.A., Hurkmans, R.T.W.L., Terink, W., Hazenberg, P., Torfs, P.J.J.F., and Uijlenhoet, R. (2010). The hydrological response of the Ourthe catchment to climate change as modelled by the HBV model, *Hydrology and Earth System Sciences*, 14: 651-665,
- El-Shafie, A., Taha, M.R., and Noureldin, A. (2007). A neuro-fuzzy model for inflow forecasting of the Nile River at Aswan high dam. *Water Resource Management*, 21(3): 533-556.
- Fathabadi, A., Salajegheh, A., and Mahdavi, M. (2009). Streamflow forecasting using neuro-fuzzy and time series methods. *Iranian Journal of Watershed Management Science and Engineering*, 2(5): 21-30. (In Farsi)
- Ghorbani, M.A., Deo, R.C., Kim, S., Hasanpour Kashani, M., Karimi, V., and Izadkhah, M. (2020). Development and evaluation of the cascade correlation neural network and the random forest models for river stage and river flow prediction in Australia. *Soft Computing*, 24: 12079- 12090.
- Hassanzadeh, Y., Abdi Kordani, A., Shafiei Najd, M., and Khoshtinat, S. (2015). Daily streamflow forecasting of Nooranchay River using the hybrid model of artificial neural networks- principal component analysis. *Water and Soil Science*, 25(3): 53-63. (In Farsi)
- Heidler, L.M. (2015). Evaluation of Different *Hydrological Models in Data Scarce Regions on the Island of Ceram, Indonesia*. Technische Universitat Munchen. Faculty of Civil, Geo and Environmental Engineering. Master of Science Program Environmental Engineering.
- Hussain, D., and Khan, A.A. (2020). Machine learning techniques for monthly river flow forecasting of Hunza River, Pakistan. *Earth Science Informatics*, 13: 939- 949.
- Jaiswal, R.K., Ali, S., and Bharti, B. (2020). Comparative evaluation of conceptual and physical rainfall- runoff models. *Applied Water Science*, 10, 48. doi:10.1007/s13201-019-1122-6
- Konz, M., and Seibert, J. (2010). On the value of glacier mass balances for hydrological model calibration. *Journal of Hydrology*, 385(1-4): 238- 246.
- Lorrai, M., and Sechi, M.G. (1995). Neural nets for modeling rainfall-runoff transformation, *Water Resources Management*, 9: 299 - 313.
- Maneshdavi, A., Nikbakht Shahbazi, A., Fathian, H. (2018). Rainfall-runoff continuous simulation in Abolabbas watershed using SMA by HEC-HMS. *Iranian Journal of Soil and Water Research*, 49(2): 317-327. (In Farsi)
- Moghaddam Nia, A., Almasi, P., Khalighi Sigaroodi, S., Salajeghe, A., and Soltani Koopaei, S. (2021). Performance evaluation of WetSpa hydrological model for runoff simulation in semi-arid climatic conditions (Case study: Menderjan Basin). *Iranian Journal of Soil and Water Research*, doi: 10.22059/ijswr.2021.315031.668827. (In Farsi)
- Mohammadivand, M., Araghinejad, S., Ebrahimi, K., and Modaresi, F. (2019). Performance evaluation of AWBM, Sacramento and SimHyd models in runoff simulation of the Amameh Watershed using automatic calibration optimization method of Genetic Algorithm. *Iranian Journal of Soil and Water Research*, 50(7): 1759-1769. (In Farsi)
- Mohammadi, M., Vagharfard, H., Mahdavi Najafabadi, R., Daneshkar Arasteh, P., and Nazemosadat, M. (2021). Rainfall-runoff modelling of coastal watersheds near the Strait of Hormuz using data mining. *Iranian Journal of Soil and Water Research*, doi: 10.22059/ijswr.2021.309641.668732. (In Farsi)
- Nazaripoooya, H., Kardavani, P., Farajirad, A. (2015). Calibration and evaluation of hydrological models, IHACRES and SWAT models in runoff simulation. *Journal of Spatial Analysis Environmental Hazards*, 2(2): 99-112. (In Farsi)
- Nikpour, M., Sanikhani, H., Mahmodi Babelan, S., and Mohammadi, A. (2017). Application of LS-SVM, ANN, WNN and GEP in rainfall-runoff modeling of Kiyav-Chay River. *Iranian Journal of Ecohydrology*, 4(2) : 627-639. (In Farsi)
- Partovyan, A., Nourani, V., and Aalami, M. (2018). Noise injection- denoising techniques to improve artificial intelligence-based rainfall- runoff modeling. *Water Engineering*, 36: 81-94. (In Farsi)
- Pham, L.T., Luo, L., and Finley, A.O. (2020). Evaluation of Random Forest for short-term daily streamflow forecast in rainfall and snowmelt driven watersheds. *Hydrology and Earth System Sciences Discussions*, <https://doi.org/10.5194/hess-2020-305>.
- Phomcha, P., Wirojanagud, P., Vangpaisal, T., and

- Thaveevouthti, T. (2011). Suitability of SWAT model for simulating of monthly streamflow in Lam Sonthi Watershed. *The Journal of Industrial Technology*, 7(2): 49- 56.
- Radchenko, I., Breuer, L., Forkutsa, I., and Frede, H.G. (2014). Simulating water resource availability under data scarcity- a case study for the Ferghana Valley (Central Asia). *Water*, 6(11): 3270–3299.
- Ren ,W.W., Yang, T., Huang, C.S., Xu, C., and Shao, Q.X. (2018). Improving monthly streamflow prediction in alpine regions: integrating HBV model with Bayesian Neural Network. *Stochastic Environmental Research and Risk Assessment*, 32: 3381–3396.
- Santhi, C., Arnold, J.G., Williams, J., Dugas, W.A., and Hauck L. (2001). Validation of the SWAT model on a large river basin with point and nonpoint sources. *Journal of the American Water Resources Association*, 37(5): 1169–1188.
- Seibert, J. (2000). Multi-criteria calibration of a conceptual runoff model using a genetic algorithm. *Hydrology and Earth System Sciences Discussions*, 2: 215- 224.
- Seyedian, S., Bagherpour, M., Fathabadi, A., and Mohammadi, A. (2019). Runoff prediction using black and gray box models. *Iran-Water Resources Research*, 14(5): 204-219. (In Farsi)
- Shafeizadeh, M., Fathian, H., and Nikbakht Shahbazi, A. (2019). Continuous rainfall-runoff simulation by artificial neural networks based on efficient input variables selection using partial mutual information (PMI) algorithm. *Iran-Water Resources Research*, 15(2): 144-161. (In Farsi)
- Yaghoubi, M., and Massah Bavani, A. (2014). Sensitivity analysis and comparison of capability of three conceptual models HEC-HMS, HBV and IHACRES in simulating continuous rainfall-runoff in SEMI-ARID Basins. *Journal of the Earth and Space Physics*, 40(2): 153-172. (In Farsi)
- Zeinali, M., Golabi, M., Sharifi, M., and Hafezparast, M., (2020). Evaluation of artificial intelligence models in river flow modeling (Case study: Gamasiab River). *Watershed Engineering and Management*, 11(4): 941-954. (In Farsi)