

مقدمه

پیش‌بینی بلندمدت جریان رودخانه بالاخص، جریان ورودی به مخازن سدها نقش کلیدی در مدیریت بلندمدت منابع آب ایفا می‌کند. از این‌رو، افزایش دقت این نوع پیش‌بینی از اهمیت ویژه‌ای برخوردار است. از سال ۲۰۰۰ تاکنون، در راستای افزایش دقت پیش‌بینی‌ها، تکنیک ترکیب مدل‌ها در پیش‌بینی‌های کوتاه مدت و بلند مدت زمانی، مکانی و زمانی-مکانی در علوم مربوط به منابع طبیعی و محیط زیست نظیر پیش‌بینی پدیده مه [۱۳]، پیش‌بینی زمانی رطوبت خاک [۹]، پیش‌بینی کوتاه مدت و بلندمدت دینامیکی اقیانوس‌ها [۱۰] و پیش‌بینی سرعت و پروفیل باد [۱۲] مورد استفاده قرار گرفت. این تکنیک در زمینه پیش‌بینی‌های کوتاه مدت و بلندمدت هیدرولوژیکی نیز در تعدادی از تحقیقات برای افزایش دقت پیش‌بینی‌ها بکاربرده شده است. [۱] و ۳ و ۴ و ۷ و ۱۵ و ۱۶ و ۱۷ و ۱۸ و ۲۱ و ۲۲.

تکنیک ترکیب مدل‌ها زیر مجموعه‌ای از تکنیک ترکیب اطلاعات می‌باشد به گونه‌ای که نتیجه حاصل از آمیختن داده‌های به دست آمده از منابع مختلف، جامع‌تر، دقیق‌تر و مطمئن‌تر از نتایج حاصل از هر یک از منابع اولیه باشد [۸]. ترکیب مدل‌ها در قالب دیدگاه‌های موازی، سری و ترکیب آنها صورت می‌گیرد [۶]. در دیدگاه ترکیب موازی مدل‌ها، نتایج حاصل از چندین مدل منفرد با روش‌هایی نظیر وزدن دهی^۴ [۴ و ۷ و ۱۸ و ۲۲] و خودراه اندازی^۵ [۲۱] با یکدیگر ترکیب می‌شوند تا نتایج دقیق‌تری حاصل شود؛ در حالیکه در دیدگاه ترکیب سری مدل‌ها، نتایج حاصل از یک مدل به عنوان ورودی به مدل دیگر وارد می‌شود [۱۷]. در دیدگاه ترکیبی که اغلب در تحقیقات هیدرولوژیکی مورد استفاده قرار می‌گیرد، نتایج حاصل از چندین مدل منفرد به عنوان ورودی به یک مدل دیگر وارد می‌شود [۱ و ۳ و ۴ و ۷ و ۱۵ و ۱۶]. در تحقیقات مذکور، مدل شبکه عصبی مصنوعی به عنوان بهترین مدل ترکیب‌کننده شناخته شده و سبب ارتقای نتایج نسبت به بهترین مدل منفرد شده است.

مدل‌سازی با روش‌هایی نظیر شبکه عصبی مصنوعی از نوع مدل‌سازی باناظر است و نیاز به آموزش دارد. این امر سبب می‌شود که برای فرآیند ترکیب مدل‌ها، خطای موجود در نتایج مدل‌های منفرد بر آموزش این مدل اثر بگذارد و از دقت نتایج آن بکاهد. بنابراین اگر بتوان برای ترکیب مدل‌ها بجای دیدگاه ترکیبی از دیدگاه موازی (روش‌های وزندهی) استفاده نمود که نیازی به مرحله

ارزیابی استراتژی‌های میانگین‌گیری وزنی رتبه‌ای در ترکیب مدل‌های پیش‌بینی‌کننده جریان مطالعه موردی: رودخانه کرخه

فرشته مدرسی^۱، شهاب عراقی‌نژاد^۲ و کیومرث ابراهیمی^۳

تاریخ دریافت: ۱۳۹۴/۰۳/۰۳ تاریخ پذیرش: ۱۳۹۵/۰۴/۱۵

چکیده

پیش‌بینی ماهانه جریان رودخانه در برنامه‌ریزی‌های بلندمدت منابع آب نقش کلیدی ایفا می‌کند. در مقاله حاضر برای افزایش دقت پیش‌بینی جریان ماهانه رودخانه کرخه در محل ورودی به سد کرخه در فصل زمستان از تکنیک ترکیب مدل‌ها استفاده شده است. بدین منظور، پنج مدل شبکه عصبی مصنوعی، شبکه عصبی رگرسیون تعمیم یافته، رگرسیون بردار پشتیبان، K نزدیک‌ترین همسایگی و رگرسیون خطی با ساختار بهینه به عنوان مدل‌های منفرد مورد استفاده قرار گرفتند. برای ترکیب مدل‌های مذکور از دو استراتژی وزندهی ثابت و متغیر به مدل‌های منفرد بر اساس روش میانگین‌گیری وزنی رتبه‌بندی شده (OWA) استفاده شده است و در آنها، روش Orlike برای تعیین وزن‌ها بکار رفته است. نتایج نشان می‌دهد که استراتژی وزندهی متغیر دارای قابلیت بیشتری برای ارتقای نتایج پیش‌بینی نسبت به وزندهی ثابت است. همچنین، مقایسه نتایج این دو استراتژی با دو استراتژی ترکیب مدل‌ها با شبکه عصبی مصنوعی و انتخاب بهترین مدل منفرد نشان می‌دهد که استراتژی وزندهی متغیر به‌طور قابل توجهی سبب ارتقای دقت نتایج نسبت به هر دو استراتژی مذکور می‌شود به نحوی که این استراتژی دقت نتایج را نسبت به شبکه عصبی ۵۱/۸، ۳۸/۱ و ۴۴/۵ درصد و نسبت به بهترین مدل منفرد ۷/۶، ۱۳۲ و ۵۲/۹ درصد به ترتیب در دی، بهمن و اسفندماه بهبود داده است.

کلید واژه‌ها: ترکیب مدل‌ها، میانگین‌گیری وزنی رتبه‌بندی شده، وزندهی متغیر، روش Orlike، کرخه.

۱- دکتری مهندسی منابع آب - گروه مهندسی آبیاری و آبادانی - دانشگاه تهران
Email: fmodaresi@ut.ac.ir

۲- دانشیار گروه مهندسی آبیاری و آبادانی - دانشگاه تهران araghinejad@ut.ac.ir

۳- نویسنده مسئول و دانشیار گروه مهندسی آبیاری و آبادانی - دانشگاه تهران.

EbrahimiK@ut.ac.ir

4- Weighting

5- Bootstrap

در مقایسه با استراتژی‌های بکارگیری شبکه عصبی مصنوعی و نیز انتخاب بهترین مدل منفرد ارزیابی شده است.

مواد و روش

منطقه مطالعات و داده‌ها

در تحقیق حاضر، حوضه کرخه به عنوان منطقه مطالعاتی انتخاب شده است. با توجه به اهمیت تأمین آب کشاورزی جلگه خوزستان توسط آب موجود در شبکه سد کرخه برای کشت بهاره و تأمین این آب توسط جریان زمستانه، در این تحقیق، پیش‌بینی جریان ماهانه ورودی به سد کرخه در فصل زمستان مورد ارزیابی قرار گرفته است. سد کرخه، سد اصلی حوضه کرخه می‌باشد و در پایین دست سد سیمره در طول جغرافیایی $48^{\circ}07'31''$ و عرض جغرافیایی $33^{\circ}29'22''$ بر روی رودخانه کرخه قرار دارد. رودخانه کرخه در ابتدا از اتصال دو شاخه اصلی به نام‌های گاماسیاب و قره‌سو شکل می‌گیرد. رودخانه گاماسیاب از بخش شمال شرقی و رودخانه قره‌سو از قسمت شمال غربی حوضه از سرشاخه‌های فرعی زیادی شکل می‌گیرند. این دو رودخانه در انتهای دشت کرمانشاه به هم پیوسته و رودخانه سیمره را تشکیل می‌دهند. از تلاقی رودخانه سیمره و رودخانه کشکان که از بخش شرقی حوضه مذکور سرچشمه می‌گیرد، رودخانه کرخه در انتهای دشت دره‌شهر شکل می‌گیرد. رودخانه کرخه در منتهی الیه حوضه به سمت شمال غرب تغییر جهت داده و به هورالعظیم می‌ریزد. در شکل ۱ موقعیت زیرحوضه مطالعاتی در حوضه کرخه نشان داده شده است.

برای پیش‌بینی جریان ماهانه ورودی به سد کرخه در فصل زمستان، اطلاعات ۳۲ ساله متغیرهای هواشناسی و هیدرولوژیکی شامل بارش متوسط ماهانه ۹ ایستگاه بارش واقع در حوضه بالادست سد کرخه با تأخیر زمانی یک تا پنج ماه (موقعیت ایستگاه‌های بارش در شکل ۱ نشان داده شده است)، میزان سطح پوشش برفی (SAE^5) در حوضه بالادست سد در سه ماه قبل از ماه مورد پیش‌بینی و جریان ماهانه ورودی به سد کرخه در پنج ماه قبل از ماه مورد پیش‌بینی در بازه زمانی ۱۳۹۲-۱۳۶۱ مورد بررسی قرار گرفتند و پیش‌بینی‌کننده‌های مناسب برای هر ماه بر اساس دو معیار ضریب همبستگی (R^2) بزرگتر از ۰/۵ و مقادیر بزرگتر شاخص اطلاعات مشترک (MI)^۶ که به ترتیب نشان دهنده میزان همبستگی خطی و غیرخطی میان پیش‌بینی‌کننده‌ها و پیش‌بینی شونده‌ها هستند، استخراج شده و مورد استفاده قرار گرفتند.

روش انجام تحقیق

تحقیق حاضر در سه مرحله کلی انجام شده است که به صورت زیر می‌باشد:

مرحله اول: پیش‌بینی جریان ماهانه با استفاده از پنج مدل پیش‌بینی

آموزش نداشته باشند، می‌توان دقت نتایج را افزایش بیشتری داد. در تحقیقاتی که روش‌های وزن‌دهی بکار رفته است، برای وزن‌دهی به نتایج مدل‌های منفرد اغلب از روش‌های میانگین‌گیری ساده (وزن‌های یکسان برای تمامی مدل‌ها) استفاده شده است [۱ و ۴ و ۷ و ۱۸ و ۲۲]. همچنین، در تحقیقات معدودی که روش وزن‌دهی به مدل‌های منفرد با وزن‌های متفاوت بکار رفته است، برای تعیین وزن‌ها از روش‌هایی نظیر روش حداقل مربعات [۷ و ۲۲] و روش استکینگ^۴ [۴] استفاده شده است. به دلیل آنکه مبنای تعیین وزن مدل‌های منفرد در این دو روش، حداقل نمودن مجموع خطاهای مقادیر شبیه‌سازی شده توسط مدل‌ها از مقادیر مشاهداتی می‌باشد، در این روش‌ها، وزن غالب (نزدیک به یک) به بهترین مدل منفرد داده می‌شود و سایر مدل‌ها از وزنی نزدیک به صفر برخوردار خواهند شد. این امر سبب می‌شود که نتایج سایر مدل‌ها تقریباً نادیده گرفته شود و نتیجه حاصل از روش وزن‌دهی به نتیجه حاصل از بهترین مدل منفرد شباهت داشته باشد و بهبودی در نتایج نسبت به نتیجه بهترین مدل منفرد حاصل نشود.

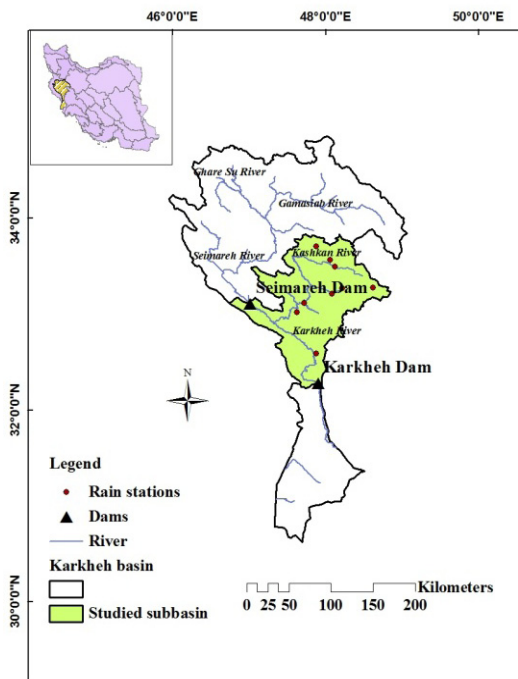
حال آنکه اگر فرآیند تعیین وزن‌ها مستقل از میزان خطای مدل‌ها صورت گیرد، می‌توان از قابلیت تمامی مدل‌ها در فرآیند ترکیب مدل‌ها استفاده نمود. همچنین، در تمامی روش‌های ترکیب مدل‌ها (اعم از وزن‌دهی یا مدل‌سازی با روش‌هایی نظیر شبکه عصبی)، به هر مدل منفرد یک وزن مشخص اختصاص داده می‌شود؛ ولی اگر بتوان از سیستم وزن‌دهی متغیر به مدل‌های منفرد استفاده نمود به نحوی که به ازای هر مقدار تخمین زده شده در هر بازه زمانی، وزن مدل‌ها تغییر نماید و بیشترین وزن به بهترین مقدار تخمین زده شده اختصاص یابد، می‌توان دقت پیش‌بینی‌ها را افزایش داد.

روش میانگین‌گیری وزنی رتبه‌بندی شده (OWA)^۲ که اولین بار توسط یاگر در سال ۱۹۸۸ ارائه شد [۲۶]، روشی است که در آن دامنه وسیعی از وزن‌ها به صورت مستقل از عوامل وزن‌گیرنده ایجاد می‌شود و با توجه به اهمیت عوامل مختلف، هر وزن می‌تواند به مورد خاصی اختصاص یابد. بدین دلیل، کاربرد اصلی این روش در تحقیقات صورت گرفته، در زمینه فرآیندهای تصمیم‌گیری^۴ بوده است [۵ و ۱۱ و ۱۹ و ۲۰ و ۲۳ و ۲۴ و ۲۷]. مدرسی و همکاران [۱۴] از این تکنیک برای ترکیب مدل‌های پیش‌بینی جریان استفاده نمودند و بر اساس وزن‌های بدست آمده از روش Orness، دقت پیش‌بینی آینده ورودی به سد سیمره را در ماه‌های فصل بهار افزایش دادند.

در تحقیق حاضر برای ارتقای دقت نتایج پیش‌بینی جریان رودخانه از استراتژی‌های وزن‌دهی ثابت و متغیر برای ترکیب مدل‌ها بر اساس روش میانگین‌گیری وزنی رتبه‌بندی شده (OWA) استفاده شده است. برای تعیین وزن مدل‌های منفرد، روش Orlike مورد استفاده قرار گرفته و بر اساس این روش، کارایی دو استراتژی مذکور

- 1- Ordinary Least Square Error
- 2- Staking
- 3- Ordered Weighted Averaging
- 4- Decision Making

5- Snow Area Extent
6- Mutual Information

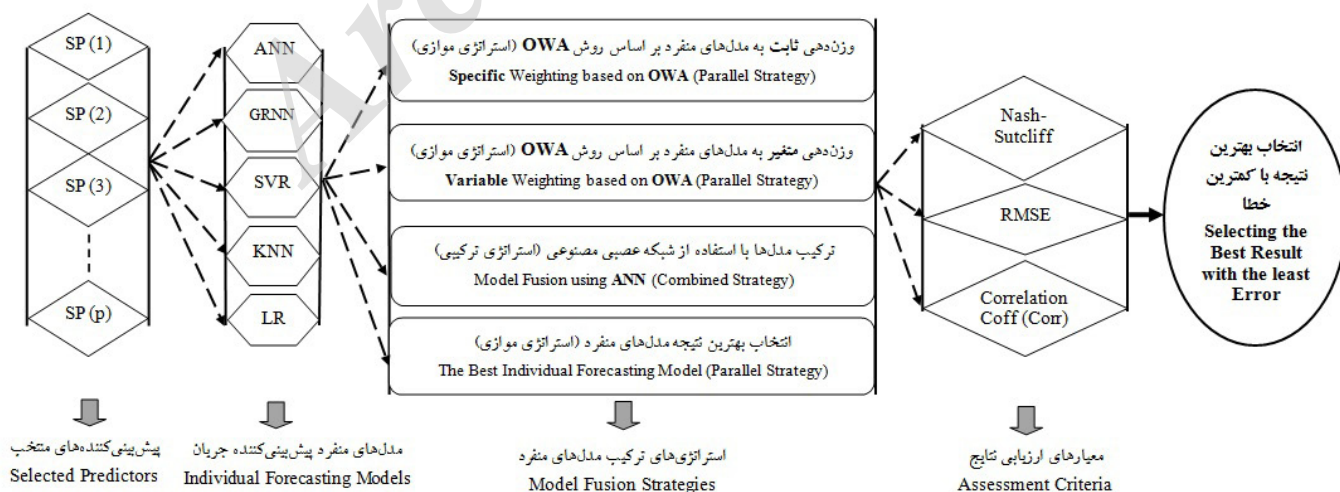


شکل ۱- موقعیت زیرحوضه مطالعاتی در حوضه کرخه
Fig 1. Location Map of the Studied Subbasin in Karkheh Basin

داده‌ها آموزش داده شده و برای تخمین آن داده، صحت سنجی می‌شود. در نهایت، مقداری از پارامتر مدل موردنظر به عنوان مقدار بهینه انتخاب می‌شود که سبب ایجاد حداقل میانگین خطا در مراحل صحت سنجی شود. در جدول ۱، نام و نام اختصاری مدل‌ها و نام پارامترهای بهینه شده برای هر مدل بیان شده است. جزئیات بیشتر درباره ساختار مدل‌های منفرد در [۲] موجود می‌باشد.

کننده شناخته شده متفاوت به عنوان مدل‌های منفرد پیش‌بینی کننده جریان و براساس پیش‌بینی کننده‌های منتخب. مرحله دوم: ترکیب نتایج حاصل از مدل‌های منفرد پیش‌بینی کننده جریان با استفاده از چهار استراتژی متفاوت شامل: الف) ترکیب مدل‌ها بر مبنای روش OWA و بر اساس استراتژی وزن‌دهی ثابت به مدل‌های منفرد ب) ترکیب مدل‌ها بر مبنای روش OWA و بر اساس استراتژی وزن‌دهی متغیر به مدل‌های منفرد ج) ترکیب مدل‌ها با استفاده از روش شبکه عصبی مصنوعی د) انتخاب بهترین مدل منفرد با کمترین خطا (این روش معادل با دادن وزن ۱ به بهترین مدل و وزن صفر به سایر مدل‌ها است) مرحله سوم: ارزیابی نتایج حاصل از چهار استراتژی فوق بر اساس سه معیار ارزیابی شناخته شده و انتخاب بهترین نتیجه الگوریتم مراحل انجام تحقیق در شکل ۲ نشان داده شده است. در ادامه، روش انجام هر مرحله و هر استراتژی و مدل‌های بکار رفته در هر بخش شرح داده شده است:

مدل‌های منفرد پیش‌بینی جریان در تحقیق حاضر، پنج مدل داده مبناء شناخته شده با ساختاری بهینه به عنوان مدل‌های منفرد برای پیش‌بینی جریان مورد استفاده قرار گرفته‌اند. برای اجرای مدل‌های منفرد از داده‌های ۲۲ و ۱۰ سال به ترتیب برای آموزش (واسنجی) و آزمایش (صحت سنجی) مدل‌ها استفاده شده است. ساختار بهینه مدل‌ها با استفاده از روش صحت سنجی متقاطع (Cross-Validation) تعیین شده است. در این روش، به ازای هر مقدار از پارامتر مدل مورد نظر، مدل به تعداد داده‌های مشاهداتی موجود، آموزش داده و صحت سنجی می‌شود؛ به‌طوری‌که در هر بار از این فرآیند، یک داده کنار گذاشته شده و مدل با سایر



شکل ۲- الگوریتم مراحل تحقیق
Fig 2. Algorithm of This Paper

Table 1. Name and Properties of Individual Models

نام کامل Complete Name	نام اختصاری Acronyms	نوع پارامترهای بهینه شده Optimized Parameters
شبکه عصبی رگرسیون تعمیم یافته Generalized Regression Neural Network	GRNN	پارامتر تنظیم مدل (پارامتر سیگما) Regularization Parameters (Sigma Parameter)
K نزدیک‌ترین همسایگی K-Nearest Neighbor Regression	KNN	وزن متغیرهای پیش‌بینی‌کننده و تعداد همسایگی‌ها (پارامتر K) Weight of Predictors and Number of Neighbors (K)
شبکه عصبی مصنوعی Artificial Neural Network	ANN	وزن‌های اولیه و تعداد نورون‌های لایه میانی Initial Weights and Number of Neurons in Middle Layer
رگرسیون بردار پشتیبان Support Vector Regression	SVR	نوع تابع کرنل، پارامترهای تابع کرنل و پارامتر تنظیم مدل (پارامتر گاما) Kernel type, Kernel Parameters, and Regularization Parameter (Gama)
رگرسیون خطی Linear Regression	LR	ضریب متغیرهای پیش‌بینی‌کننده Coefficient of Predictors

استراتژی‌های ترکیب مدل‌های منفرد

الف) وزن‌دهی بر اساس روش OWA

روش میانگین‌گیری وزنی رتبه‌بندی شده (OWA) با ابعاد n نگاهی از R^n به R است که در آن، n تعداد معیارها (تعداد مدل‌های منفرد) است و برای هر معیار، وزنی به صورت w_i ($i=1, n$) وجود دارد به طوریکه [۲۶]:

$$\sum_{i=1}^n w_i = 1, \quad w_i \in [0, 1] \quad (1)$$

برای بکارگیری عملگر OWA برای پیش‌بینی جریان، فرض کنید مقدار جریان پیش‌بینی شده توسط مدل منفرد i ام برابر با a_i باشد. آنگاه بردار نتایج حاصل از n مدل منفرد برابر با $A = \{a_1, a_2, \dots, a_n\}$ خواهد بود. در این صورت، بردار مرتب شده نزولی حاصل از بردار A برابر با $B = \{b_1, b_2, \dots, b_n\}$ و بردار وزن‌های متناظر با بردار B به صورت $W = \{w_1, w_2, \dots, w_n\}$ خواهد بود. در نهایت عملگر OWA به صورت زیر بر روی نتایج اعمال می‌شود:

$$OWA(a_1, a_2, \dots, a_n) = \sum_{i=1}^n w_i b_i \quad (2)$$

در استراتژی وزن‌دهی ثابت، مدل‌های منفرد بر اساس دقت نتایجشان به صورت نزولی مرتب می‌شوند و به هر مدل منفرد وزن ثابتی برای تمامی مقادیر تخمین زده شده اختصاص می‌یابد؛ در حالیکه در استراتژی وزن‌دهی متغیر، مقادیر تخمین زده شده توسط مدل‌ها در هر گام زمانی به صورت نزولی مرتب می‌شوند و در نتیجه وزن مدل‌های منفرد به ازای هر یک از مقادیر تخمین زده شده، متغیر می‌باشد.

در این روش، برای تعیین نحوه پراکندگی وزن‌ها حول مقدار میانه آنها، یاگر [۲۶] پارامتر Orness را به صورت زیر تعریف نمود:

$$Orness(w) = \frac{1}{n-1} \sum_{i=1}^n (n-i) w_i \quad 0 \leq Orness(w) \leq 1 \quad (3)$$

که در این رابطه، n : تعداد عوامل وزن گیرنده (مدل منفرد)، i : شمارنده عوامل (مدل‌ها) و w_i : وزن اختصاص داده شده به هر عامل (مدل منفرد) می‌باشد. مقادیر صفر و یک این پارامتر به ترتیب نشان دهنده وزن برابر با یک به بدترین و بهترین حالت (خروجی مدل) هستند. در این شرایط، وزن سایر مدل‌ها برابر با صفر خواهد بود. مقدار $0/5$ برای این پارامتر نشان می‌دهند که یا همگی وزن‌ها برابر هستند و یا اینکه وزن‌ها به صورت توزیع نرمال حول میانه قرار دارند. بنابراین مقادیر میان $0/5$ و یک برای این پارامتر نشان می‌دهد که توزیع وزن‌ها دارای چولگی مثبت است و این بدین معناست که وزن بیشتری به مدل بهتر اختصاص یافته است.

تعیین وزن‌ها در روش OWA به شیوه‌های متفاوتی صورت می‌گیرد. یکی از معروف‌ترین آنها، خانواده‌ای از وزن‌های OWA با عنوان S-OWA است که توسط یاگر و فیلو [۲۵] در سال ۱۹۹۴ ارائه شد. این روش شامل دو نوع وزن‌دهی Andlike و Orlike است. در نوع Orlike، بیشترین وزن به بدترین حالت و در نوع Andlike، بیشترین وزن به بهترین حالت اختصاص داده می‌شود. بنابراین، در تحقیق حاضر، از روش Orlike استفاده شده است و بیشترین وزن به بهترین مدل منفرد (استراتژی وزن‌دهی ثابت) و بهترین جواب (استراتژی وزن‌دهی متغیر) داده شده است. در روش Orlike، مقادیر وزن‌ها و عملگر OWA به ترتیب از روابط ۴ و ۵ محاسبه می‌شوند:

$$w_i = \begin{cases} \frac{1}{n}(1-\alpha) + \alpha, & i = 1 \\ \frac{1}{n}(1-\alpha) & i = 2, \dots, n \end{cases}, \quad \alpha \in [0, 1] \quad (4)$$

می‌باشند و برای آموزش و آزمایش این مدل، از خروجی مدل‌های منفرد در دوره‌های واسنجی و صحت سنجی استفاده شده است. (ج) انتخاب بهترین نتیجه مدل‌های منفرد در این استراتژی، وزن یک، به بهترین مدل منفرد و وزن صفر به سایر مدل‌های منفرد اختصاص داده می‌شود.

روش‌های ارزیابی نتایج

در تحقیق حاضر، برای ارزیابی نتایج پیش‌بینی‌های صورت گرفته و طبقه‌بندی مدل‌های منفرد، از سه معیار ارزیابی شامل شاخص ارزیابی مدل نش- ساتکلیف (E)، ریشه متوسط مربعات خطا (RMSE) و ضریب همبستگی (R^2) که به ترتیب در روابط ۶، ۷ و ۸ نشان داده شده‌اند، استفاده شده است.

$$E = 1 - \frac{\sum_{t=1}^n (Y_t - \hat{Y}_t)^2}{\sum_{t=1}^n (Y_t - \bar{Y})^2} \quad (6)$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (Y_t - \hat{Y}_t)^2}{n}} \quad (7)$$

$$R = \frac{\sum_{t=1}^n (Y_t - \bar{Y})(\hat{Y}_t - \bar{\hat{Y}})}{\sqrt{\sum_{t=1}^n (Y_t - \bar{Y})^2} \cdot \sqrt{\sum_{t=1}^n (\hat{Y}_t - \bar{\hat{Y}})^2}} \quad (8)$$

در روابط فوق، Y_t و \hat{Y}_t به ترتیب، مقادیر جریان مشاهداتی و تخمین زده شده و \bar{Y} و $\bar{\hat{Y}}$ به ترتیب، میانگین مقادیر مشاهداتی و تخمین زده شده و n تعداد مشاهدات می‌باشد.

$$OWA(a_1, a_2, \dots, a_n) = \alpha \text{Max}(a_i) + \frac{1}{n}(1 - \alpha) \sum_{i=1}^n a_i \quad (9)$$

در رابطه ۹، به ازای مقادیر مختلف پارامتر α در بازه صفر و یک می‌توان به ترکیب متفاوتی از وزن‌ها دست یافت و به ازای مقادیر صفر و یک برای پارامتر α ، پارامتر اورنس (Omness) به ترتیب برابر با ۰/۵ و یک خواهد شد؛ این امر نشان می‌دهد که در این روش به ازای تمامی مقادیر α ، وزن بیشتری به مدل اول اختصاص می‌یابد. بر اساس وزن‌های محاسبه شده از رابطه ۹، عملگر OWA از رابطه ۵ قابل محاسبه است؛ به طوری که وزن بهترین مدل ($\text{Max}(a_i)$) که برای آن، $i=1$ است، برابر با α و وزن سایر مدل‌ها برابر با $\frac{1}{n}(1 - \alpha)$ می‌باشد.

مزیت تعیین وزن در روش OWA به سایر روش‌های تعیین وزن نظیر روش استکینگ و یا روش حداقل مربعات در این است که این روش به کاربر این امکان را می‌دهد که با توجه به مقدار پارامتر α ، به تنوع وسیعی از ترکیب وزن‌ها دست یابد و با توجه به اهمیت مدل‌ها، به هر مدل وزن مناسب را اختصاص دهد. همچنین، در این روش مقادیر وزن مدل‌ها به صورت مستقل از میزان خطای مدل‌ها تعیین می‌شود.

ب) ترکیب مدل‌ها با استفاده از روش شبکه عصبی مصنوعی برای ترکیب مدل‌های منفرد با استفاده از روش شبکه عصبی، یک مدل شبکه عصبی مصنوعی سه لایه (MLP) با نوع آموزش پیشخور و برگشت به عقب (FFBP) استفاده شده است. در این مدل، تعداد نورون‌های لایه میانی و وزن‌های اولیه هر نورون برای ترکیب مدل‌های پیش‌بینی کننده جریان هرماه با استفاده از روش صحت سنجی متقاطع (cross-validation) تعیین شده است. ورودی‌های مدل شبکه عصبی: نتایج حاصل از پیش‌بینی جریان توسط مدل‌های منفرد و خروجی‌های آن: مقادیر واقعی جریان رودخانه در هر ماه

جدول ۲- نتایج پیش‌بینی جریان توسط مدل‌های منفرد در مرحله صحت سنجی

Table 2. Results of Streamflow Forecasting by Individual Models in Validation Phase

مدل‌های منفرد Individual Models					شاخص‌های ارزیابی مدل‌ها Assessment Criteria	ماه پیش‌بینی آبدهی Month
LR	SVR	ANN	KNN	GRNN		
0.760	0.841	0.678	0.597	0.699	Nash-Sutcliff	دی
0.648	0.527	0.750	0.840	0.726	RMSE	January
0.895	0.926	0.886	0.833	0.925	Correlation Coefficient	
0.178	0.067	0.081	0.121	-0.034	Nash-Sutcliff	بهمن
1.135	1.209	1.200	1.173	1.273	RMSE	February
0.574	0.572	0.574	0.662	0.531	Correlation Coefficient	
0.080	0.231	0.382	0.356	0.123	Nash-Sutcliff	اسفند
0.796	0.728	0.652	0.666	0.777	RMSE	March
0.676	0.688	0.656	0.645	0.635	Correlation Coefficient	

نتایج و بحث

نتایج حاصل از اجرای مدل‌های منفرد

همان‌طور که در بخش مواد و روش‌ها بیان شد، برای اجرای مدل‌های منفرد از داده‌های ۲۲ و ۱۰ سال به ترتیب برای آموزش (واسنجی) و آزمایش (صحت سنجی) مدل‌ها استفاده شده است. از آنجایی که هدف این تحقیق، افزایش دقت نتایج در مرحله صحت سنجی است، نتایج حاصل از اجرای مدل‌های منفرد برای پیش‌بینی جریان در مرحله صحت سنجی در جدول ۲ آورده شده است.

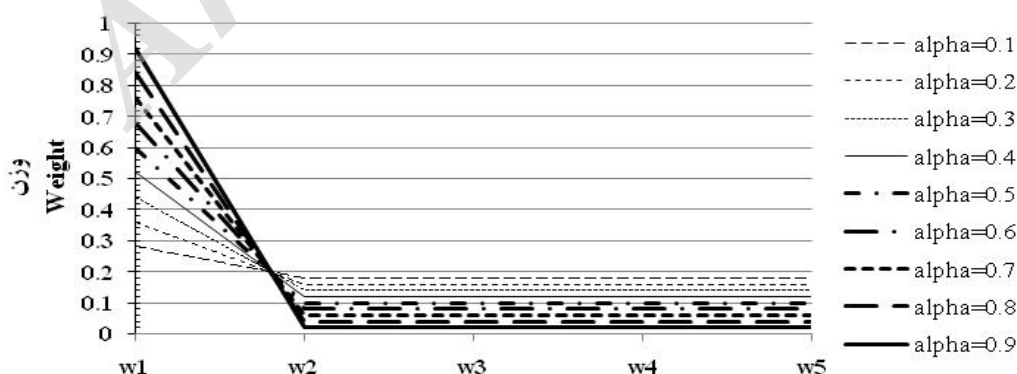
همان‌طور که در جدول ۲ مشاهده می‌شود، تمامی مدل‌ها برای پیش‌بینی جریان دی‌ماه از کارایی خوبی برخوردار هستند به‌طوری‌که مقادیر شاخص نش برای همه آنها بزرگتر از ۰/۵، RMSE کوچکتر از ۰/۹ و ضریب همبستگی مقادیر مشاهداتی و شبیه‌سازی شده بزرگتر از ۰/۸ می‌باشد. در این ماه، مدل SVR از بیشترین کارایی برای پیش‌بینی جریان نسبت به سایر مدل‌ها برخوردار است. برای پیش‌بینی آبدی بهمن‌ماه، اگرچه ضریب همبستگی میان مقادیر مشاهداتی و شبیه‌سازی شده بزرگتر از ۰/۵ است، ولی شاخص نش برای تمامی مدل‌ها کوچکتر از ۰/۲ و شاخص RMSE بزرگتر از یک می‌باشد و این امر نشان می‌دهد که پیش‌بینی‌های صورت گرفته توسط مدل‌های منفرد از دقت بالایی برخوردار نیست. در این ماه، مدل LR بهترین نتایج را ارائه داده است. دقت نتایج پیش‌بینی آبدی اسفندماه توسط مدل‌های منفرد از بهمن ماه بیشتر و از دی ماه کمتر است. ضریب همبستگی میان نتایج حاصل برای این ماه و مقادیر مشاهداتی بزرگتر از ۰/۶، شاخص RMSE کوچکتر از یک و شاخص نش کوچکتر از ۰/۴ می‌باشد. در این ماه، بهترین نتایج متعلق به مدل ANN است.

نتایج حاصل از پیش‌بینی آبدی برای این سه ماه نشان می‌دهد که یک مدل منفرد به تنهایی توانایی ارائه بهترین پیش‌بینی را در همه ماه‌ها ندارد و نتایج حاصل از بهترین مدل‌های هر ماه نیز کاملاً دقیق نیستند. از این‌رو لزوم بکارگیری چندین مدل منفرد و استراتژی‌های ترکیب مدل‌ها برای افزایش دقت پیش‌بینی‌ها وجود دارد.

نتایج حاصل از استراتژی‌های ترکیب مدل‌ها استخراج وزن‌های مدل‌های منفرد با روش Orlike در این بخش، وزن‌های حاصل از روش Orlike برای پنج مدل منفرد به ازای مقادیر α از ۰/۱ تا ۰/۹ از رابطه ۳ استخراج و به صورت نمودار به تفکیک مقادیر α در شکل شماره ۳ نشان داده شده است. در شکل ۳، محور افقی نماینده وزن مدل‌ها است که از w_1 تا w_5 به ترتیب به بهترین تا بدترین مدل منفرد (از نظر میزان خطا) اعمال می‌شود و محور عمودی نشان دهنده مقدار وزن‌ها می‌باشد. در این شکل مشاهده می‌شود که در روش Orlike، به ازای تمامی مقادیر α ، وزن مدل اول از سایر مدل‌ها بیشتر و وزن سایر مدل‌ها با هم برابر است. در این روش، وزن بهترین مدل از ۰/۲۸ به ازای α برابر با ۰/۱ برای α تا ۰/۹۲ به ازای α برابر با ۰/۹ تغییر می‌نماید.

استراتژی وزن‌دهی ثابت به مدل‌های منفرد بر مبنای عملگر OWA و مقایسه آن با دو استراتژی شبکه عصبی و انتخاب بهترین مدل منفرد در استراتژی وزن‌دهی ثابت به مدل‌های منفرد بر مبنای عملگر OWA، مدل‌های منفرد بر مبنای نتایج در مرحله صحت سنجی و بر اساس دو شاخص نش-ساتکلیف و RMSE که نشان‌دهنده میزان خطای پیش‌بینی هستند، به صورت نزولی از بهترین به بدترین مدل مرتب شده و با استفاده از وزن‌های حاصل از روش Orlike، وزن‌دهی و ترکیب شدند. نتایج حاصل از این استراتژی (استراتژی ۱) در جدول ۳ آورده شده و با نتایج حاصل از دو استراتژی ترکیب مدل‌ها با شبکه عصبی (استراتژی ۳) و انتخاب بهترین مدل منفرد (استراتژی ۴) در مرحله صحت سنجی مقایسه شده است.

در جدول ۳ مشاهده می‌شود که برای پیش‌بینی آبدی دی‌ماه، در استراتژی ۱ به ازای مقادیر α از ۰/۴ تا ۰/۹ برای پارامتر α ، دقت نتایج بر اساس هر سه معیار ارزشیابی نش-ساتکلیف، RMSE و ضریب همبستگی نسبت به نتایج حاصل از بهترین مدل منفرد (استراتژی ۴) افزایش یافته است و بیشترین افزایش دقت در نتایج به ازای مقدار ۰/۷ برای پارامتر α رخ داده است که در این حالت ۲/۴ درصد افزایش در معیار ارزیابی نش صورت گرفته است. این در حالی است



شکل ۳- نمودار وزن‌های ایجاد شده بر اساس روش Orlike

Fig 3. The Graph of the Weights Produced Based on Orlike Method

جدول ۳- نتایج ترکیب مدل‌ها با استراتژی وزن‌دهی ثابت بر اساس روش Orlike

Table3. Results of Model Fusion Using Specific Weighting Based on Orlike Method

استراتژی ۴ Strategy 4	استراتژی ۳ Strategy 3	استراتژی ۱: وزن‌دهی ثابت با روش OWA Strategy 1. Specific Weighting Using OWA									شاخص‌های ارزیابی نتایج Assessment Criteria	ماه پیش‌بینی آبدهی Month		
		$\alpha=0.9$	$\alpha=0.8$	$\alpha=0.7$	$\alpha=0.6$	$\alpha=0.5$	$\alpha=0.4$	$\alpha=0.3$	$\alpha=0.2$	$\alpha=0.1$				
انتخاب بهترین مدل منفرد Best IFM Selection	ترکیب مدل‌ها با شبکه عصبی Model Fusion Using ANN										Nash-Sutcliff	دی January		
		0.841	0.596	0.851	0.857	0.861	0.860	0.857	0.850	0.840	0.826		0.809	
		0.527	0.841	0.511	0.499	0.494	0.494	0.501	0.513	0.530	0.552		0.578	RMSE
		0.926	0.832	0.928	0.930	0.932	0.934	0.935	0.936	0.937	0.937	0.937	Correlation	
		0.178	0.299	0.184	0.189	0.193	0.196	0.197	0.197	0.197	0.195	0.192	Nash-Sutcliff	بهمن February
		1.135	1.048	1.131	1.127	1.125	1.123	1.122	1.121	1.122	1.123	1.125	RMSE	
		0.574	0.718	0.584	0.594	0.603	0.612	0.621	0.629	0.636	0.643	0.649	Correlation	
		0.382	0.404	0.379	0.374	0.367	0.360	0.352	0.343	0.332	0.320	0.308	Nash-Sutcliff	اسفند March
		0.652	0.641	0.654	0.657	0.660	0.664	0.668	0.673	0.678	0.684	0.691	RMSE	
		0.656	0.700	0.660	0.663	0.666	0.669	0.672	0.674	0.676	0.678	0.680	Correlation	

دارد که وزن‌دهی مناسب به نتایج این مدل‌ها، سبب افزایش دقت نتایج نهایی پیش‌بینی جریان می‌شود.

مقایسه نتایج حاصل از استراتژی ۱ و ۳ نشان می‌دهد که در ماه‌های بهمن و اسفند، عملکرد استراتژی ۳ بهتر از استراتژی ۱ است در حالیکه در دی‌ماه، نتایج استراتژی ۱ به ازای تمامی مقادیر پارامتر α بهتر از استراتژی ۳ می‌باشد.

استراتژی وزن‌دهی متغیر به مدل‌های منفرد بر مبنای عملگر OWA و مقایسه آن با دو استراتژی شبکه عصبی و انتخاب بهترین مدل منفرد

در این استراتژی، مقادیر جریان پیش‌بینی شده توسط مدل‌های منفرد در هر گام زمانی بر اساس میزان خطا نسبت به مقادیر واقعی جریان، به صورت صعودی (از کم‌ترین خطا به بیشترین خطا) منظم شده و توسط وزن‌های حاصل از روش Orlike وزن‌دهی و با یکدیگر ترکیب شدند. نتایج حاصل از این استراتژی (استراتژی ۲) در جدول ۴ آورده شده و با نتایج حاصل از دو استراتژی ترکیب مدل‌ها با شبکه عصبی (استراتژی ۳) و انتخاب بهترین مدل منفرد (استراتژی ۴) در مرحله صحت سنجی مقایسه شده است.

همان‌طور که در جدول ۴ مشاهده می‌شود، برای پیش‌بینی آبدهی دی‌ماه، در استراتژی ۲ به ازای مقادیر $\alpha=0.3$ تا $\alpha=0.9$ برای پارامتر α ، دقت نتایج بر اساس هر سه شاخص ارزیابی مدل‌ها نسبت به بهترین مدل منفرد افزایش یافته است و بیشترین افزایش دقت به ازای مقدار $\alpha=0.8$ برای پارامتر α رخ داده است و در این حالت دقت نتایج بر اساس معیار نش-ساتکلیف، $7/8$ درصد نسبت به نتایج بهترین مدل

که بکارگیری شبکه عصبی برای ترکیب نتایج (استراتژی ۳) دقت نتایج بهترین مدل منفرد را افزایش نداده است.

برای پیش‌بینی آبدهی بهمن ماه، بکارگیری استراتژی ۱ به ازای تمامی مقادیر α و نیز استراتژی ۳ (بکارگیری شبکه عصبی) دقت نتایج را نسبت به نتایج بهترین مدل منفرد (استراتژی ۴) افزایش داده است. برای این ماه، نتایج استراتژی ۳ از دقت بالاتری نسبت به استراتژی ۱ برخوردار است؛ به طوری‌که در استراتژی ۳، دقت نتایج به ازای معیار نش، 68 درصد نسبت به بهترین مدل منفرد افزایش یافته است. در حالیکه، در استراتژی ۱، بهترین نتایج متعلق به مقدار $\alpha=0.3$ برای پارامتر α است و به ازای آن، معیار نش، $10/7$ درصد نسبت به بهترین مدل منفرد افزایش پیدا کرده است. در این حالت، ضریب همبستگی میان مقادیر واقعی و مدل شده بیشتر از دو حالت 0.5 و 0.4 برای پارامتر α است.

همچنین، جدول ۳ نشان می‌دهد که برای پیش‌بینی آبدهی اسفند ماه، بکارگیری استراتژی ۱ به ازای تمامی مقادیر α ارتقایی در نتایج نسبت به نتایج بهترین مدل منفرد ایجاد نکرده است؛ در حالیکه، بکارگیری استراتژی ۳ سبب ارتقای نتایج به میزان $5/8$ درصد بر حسب شاخص نش نسبت به بهترین مدل منفرد شده است.

با مقایسه بهترین نتایج در استراتژی ۱ می‌توان دریافت که اگرچه با زیاد شدن مقدار پارامتر α ، وزن بهترین مدل افزایش می‌یابد، ولی بهترین نتایج به ازای مقادیر $\alpha=0.7$ و $\alpha=0.3$ برای پارامتر α به ترتیب برای دی‌ماه و بهمن‌ماه رخ داده است. این بدان معناست که در سایر مدل‌های منفرد (غیر از بهترین مدل منفرد) نیز نتایج مطلوبی وجود

جدول ۴- نتایج ترکیب مدل‌ها با استراتژی وزن‌دهی متغیر بر اساس روش Orlike

Table 4. Results of Model Fusion Using Specific Weighting Based on Orlike Method

استراتژی ۴ Strategy 4	استراتژی ۳ Strategy 3	استراتژی ۲: وزن‌دهی متغیر با روش OWA Strategy 2. Specific Weighting Using OWA									شاخص‌های ارزیابی نتایج Assessment Criteria	ماه پیش‌بینی آبدهی Month
		$\alpha=0.9$	$\alpha=0.8$	$\alpha=0.7$	$\alpha=0.6$	$\alpha=0.5$	$\alpha=0.4$	$\alpha=0.3$	$\alpha=0.2$	$\alpha=0.1$		
انتخاب بهترین مدل منفرد Best IFM Selection	ترکیب مدل‌ها با شبکه عصبی Model Fusion Using ANN										Nash- Sutcliff RMSE	ژانویه دی
0.841	0.596	0.905	0.907	0.904	0.899	0.889	0.876	0.860	0.840	0.816	Correlation	فروردین
0.527	0.841	0.407	0.404	0.409	0.421	0.440	0.465	0.495	0.529	0.567	Nash- Sutcliff RMSE	
0.926	0.832	0.954	0.954	0.953	0.952	0.951	0.949	0.947	0.944	0.941	Correlation	
0.178	0.299	0.413	0.401	0.385	0.366	0.344	0.319	0.291	0.259	0.225	Nash- Sutcliff RMSE	فروردین
1.135	1.048	0.959	0.969	0.982	0.997	1.014	1.033	1.054	1.077	1.102	Correlation	اسفند
0.574	0.718	0.716	0.713	0.709	0.705	0.699	0.693	0.685	0.677	0.667	Nash- Sutcliff RMSE	
0.382	0.404	0.584	0.563	0.539	0.513	0.483	0.451	0.416	0.378	0.337	Correlation	
0.652	0.641	0.536	0.549	0.563	0.579	0.597	0.615	0.634	0.655	0.676	Nash- Sutcliff RMSE	اسفند
0.656	0.700	0.816	0.807	0.797	0.786	0.774	0.759	0.743	0.724	0.704	Correlation	

مقایسه نتایج حاصل از استراتژی ۲ با استراتژی‌های ۳ و ۴ برای هر سه ماه نشان می‌دهد که این استراتژی توانسته است در هر سه ماه دقت نتایج را به‌طور چشمگیری بهبود بخشد و از آنجایی که در این استراتژی، در هر گام زمانی، نتایج مدل‌های منفرد از بهترین به بدترین مرتب و وزن‌دهی می‌شود، بهترین نتایج در مقادیر α برابر با ۰/۸ و ۰/۹ رخ می‌دهد. در نمودارهای شکل ۴، بهترین نتایج استراتژی‌های ۲ (Orlike(1)) و (Orlike(2)) با نتایج استراتژی‌های ۳ و ۴ برای ماه‌های دی، بهمن و اسفند مقایسه شده است.

در شکل ۴ مقادیر آبدهی نرمال استاندارد شده (روش استانداردسازی به شیوه توزیع نرمال انجام شده است) واقعی در مقابل مقادیر پیش‌بینی شده توسط استراتژی‌های ترکیب مدل‌ها به تفکیک سه ماه در دوره صحت سنجی ترسیم شده است. در این شکل مشاهده می‌شود که در هر سه ماه، نتایج حاصل از استراتژی ۲ (Orlike (2)) نسبت به سایر استراتژی‌ها به مقادیر مشاهداتی نزدیک‌تر و دارای خطای کمتر می‌باشد. برای پیش‌بینی آبدهی بهمن ماه، نتایج در تمامی سال‌ها بجز سال ۱۳۸۴ از دقت مناسبی برخوردار است به نحوی که با حذف نتیجه این سال از سایر نتایج، دقت نتایج به‌طور قابل توجهی افزایش می‌یابد و شاخص نش- ساتکلِف به مقدار ۰/۷۷۸، ۰/۹۱۸، ۰/۶۹۲ و ۰/۷۵۷ به ترتیب برای استراتژی‌های یک تا چهار می‌رسد.

منفرد افزایش یافته است؛ ولی همان‌طور که در بخش قبل نیز مشاهده شد، بکارگیری استراتژی ۳ (ترکیب مدل‌ها با شبکه عصبی) سبب ارتقای دقت نتایج بهترین مدل منفرد نشده است.

همچنین، برای پیش‌بینی آبدهی بهمن‌ماه، با بکارگیری استراتژی ۲ به ازای تمامی مقادیر پارامتر α ، دقت نتایج بر حسب هر سه معیار ارزیابی نتایج نسبت به بهترین مدل منفرد افزایش یافته است. در این استراتژی، بهترین نتیجه متعلق به α برابر با ۰/۹ است که در این حالت، دقت نتایج بر حسب معیار نش- ساتکلِف، ۱۳۲ درصد نسبت به بهترین مدل منفرد افزایش یافته است. مقایسه نتایج حاصل از استراتژی ۲ و ۳ برای پیش‌بینی جریان بهمن‌ماه نشان می‌دهد که بر اساس معیارهای نش- ساتکلِف و RMSE، دقت نتایج حاصل از استراتژی ۲ به ازای مقادیر ۰/۴ تا ۰/۹ برای پارامتر α ، بیشتر از دقت نتایج حاصل از استراتژی ۳ می‌باشد؛ به‌طوری‌که به ازای α برابر با ۰/۹، دقت نتایج استراتژی ۲، ۳۸ درصد بیشتر از استراتژی ۳ است. نتایج پیش‌بینی آبدهی اسفندماه نیز نشان می‌دهد که در استراتژی ۲ به ازای مقادیر ۰/۳ تا ۰/۹ برای پارامتر α ، دقت نتایج نسبت به استراتژی‌های ۳ و ۴ به ازای هر سه معیار ارزیابی افزایش یافته است و بهترین نتیجه به ازای α برابر با ۰/۹ رخ داده است که در این حالت، دقت نتایج، ۴۴/۵ درصد نسبت به استراتژی ۳ و ۵۳ درصد نسبت به استراتژی ۴ بهبود یافته است.

نتیجه گیری

پیش‌بینی بلندمدت جریان رودخانه بالاخص، جریان ورودی به مخازن سدها نقش کلیدی در مدیریت بلندمدت منابع آب ایفا می‌کند. از این‌رو، افزایش دقت این نوع پیش‌بینی از اهمیت ویژه‌ای برخوردار است. با توجه به اهمیت تأمین آب کشاورزی جلگه خوزستان توسط آب موجود در شبکه سد کرخه برای کشت بهاره و تأمین این آب توسط جریان زمستانه، در این مقاله، پیش‌بینی جریان ماهانه ورودی به سد کرخه در فصل زمستان مورد ارزیابی قرار گرفته است.

در تحقیق حاضر برای ارتقای دقت نتایج حاصل از پیش‌بینی جریان ماهانه از دیدگاه ترکیب مدل‌های پیش‌بینی کننده استفاده شده است. برای این منظور، پنج مدل داده مبناء شامل شبکه عصبی مصنوعی، شبکه عصبی رگرسیون تعمیم یافته، رگرسیون بردار پشتیبان، k نزدیک‌ترین همسایگی و رگرسیون خطی با ساختار بهینه به عنوان مدل‌های منفرد پیش‌بینی جریان برای هر ماه مورد استفاده قرار گرفتند و بر اساس پیش‌بینی‌کننده‌های منتخب برای هر ماه اجرا شدند.

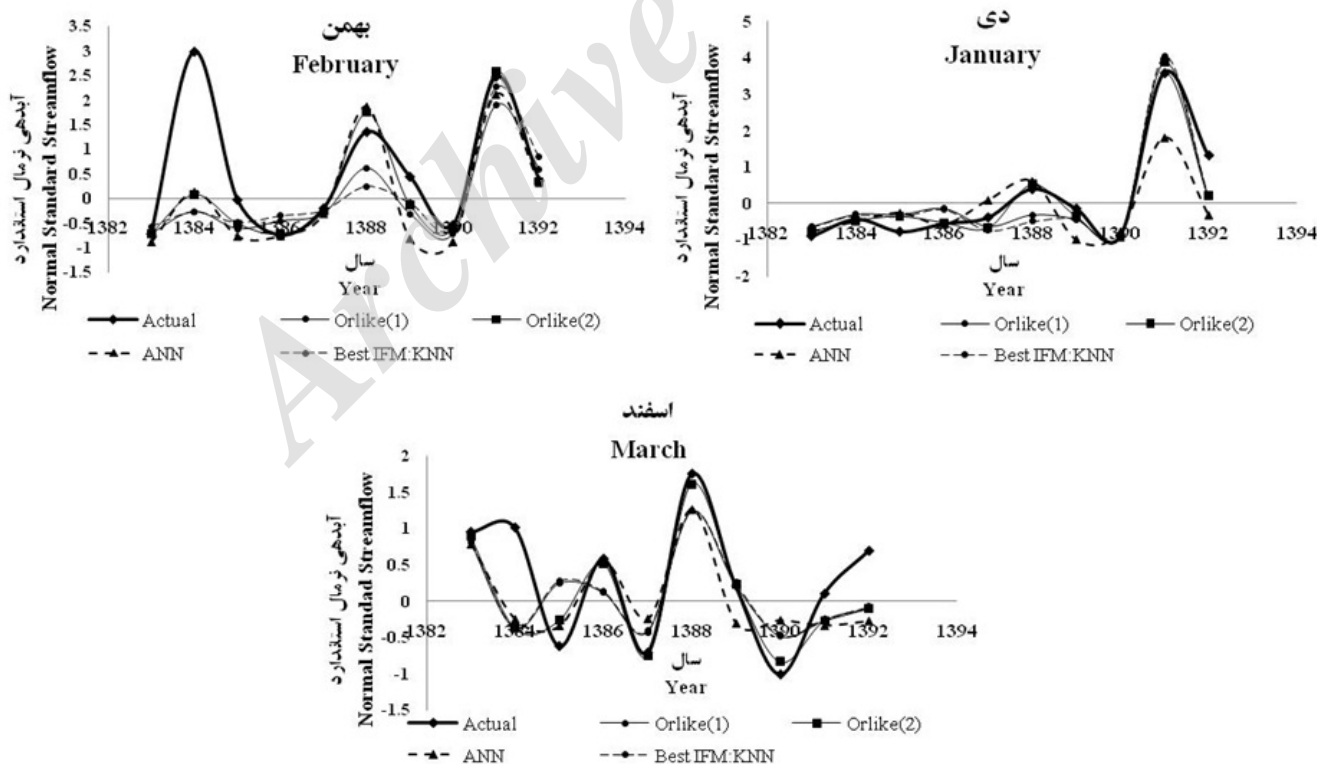
برای ترکیب کردن مدل‌ها، دو استراتژی وزن‌دهی جدید بر مبنای روش میانگین‌گیری وزنی رتبه‌بندی شده (OWA) شامل استراتژی‌های وزن‌دهی ثابت و متغیر برای ترکیب نتایج مدل‌های منفرد استفاده شده و نتایج آنها با دو استراتژی ترکیب مدل‌ها با شبکه

عصبی و استراتژی انتخاب بهترین نتیجه مدل‌های منفرد مقایسه شده است.

در استراتژی‌های وزن‌دهی ثابت و متغیر برای تعیین وزن‌ها از روش Orlike استفاده شده است که در آن، وزن‌ها به صورت مستقل از میزان خطای مدل‌های منفرد تعیین می‌شود.

نتایج حاصل از این تحقیق نشان می‌دهد که:

- مدل‌های رگرسیون بردار پشتیبان (SVR)، رگرسیون خطی (LR) و شبکه عصبی مصنوعی (ANN) به عنوان بهترین مدل‌های منفرد برای پیش‌بینی جریان به ترتیب در ماه‌های دی، بهمن و اسفند شناخته شدند.
- استراتژی ۱ سبب ارتقای نتایج پیش‌بینی جریان در ماه‌های دی و بهمن به ترتیب به میزان $2/4$ و $10/7$ درصد نسبت به نتایج بهترین مدل منفرد بر حسب شاخص نش شده است ولی در اسفند ماه ارتقای در نتایج ایجاد نکرده است. این استراتژی فقط در دی‌ماه توانسته است، نتایج حاصل از شبکه عصبی را بهبود بخشد.
- استراتژی ۳ (شبکه عصبی) در دی‌ماه ارتقای در نتایج ایجاد نکرده است در حالیکه در بهمن ماه و اسفندماه به ترتیب به میزان 68 و $5/8$ درصد سبب ارتقای نتایج بهترین مدل منفرد بر اساس شاخص نش شده است.
- استراتژی ۲ سبب ارتقای قابل ملاحظه در نتایج پیش‌بینی جریان در تمامی ماه‌ها نسبت به هر دو استراتژی ۳ و ۴ شده است به



شکل ۴- مقایسه بهترین نتایج حاصل از استراتژی‌های ترکیب مدل‌ها برای پیش‌بینی آبدی ماه‌های دی، بهمن و اسفند

Fig 4. Comparison of the Best Results of Model Fusion Strategies for Streamflow Forecast in January, February, and March

Application of artificial neural network ensembles in probabilistic hydrological forecasting. *Hydrology*. 407: 94-104.

4. Azmi, M. Araghinejad, S. and Kholghi, M. 2010. Multi model data fusion for hydrological forecasting using K- nearest neighbor method. *Iranian Journal of Science and Technology*. 34 (B1): 81-92.

5. Chiclana, F. Herrera-Viedma, E. Herrera, F. and Alonso, S. 2007. Some induced ordered weighted averaging operators and their use for solving group decision-making problems based on fuzzy preference relations. *European Journal of Operational Research*. 182: 383-399.

6. Dasarathy, B. V. 1994. *Decision Fusion*. IEEE Computer Society Press, Los Alamitos, California., ISBN: 0818644524.

7. Goswami, M. and O'Connor, K. M. 2007. Real-time flow forecasting in the absence of quantitative precipitation forecasts: a multi-model approach. *Hydrology*. 334: 125-140.

8. Hall, D. L. and Llinas, J. 1997. An introduction to multisensor data fusion. In *proceedings of the IEEE*. 85: 6-10.

9. Khalili, A. Gill, M. K. and McKee, M. 2005. New applications for information fusion and soil moisture forecasting. 8th International Conference on Information Fusion. Philadelphia, 25-28 July.

10. Leslie, W. G. Robinson, A. R. Haley J, P. J. Logutov, O. Moreno, P. A. Lermusiaux, P. F. J. and Coelho, E. 2008. Verification and training of real-time forecasting of multi-scale ocean dynamics for maritime rapid environmental assessment. *Marine Systems – Maritime Rapid Environmental Assessment – New Trends in Operational Oceanography*. 69: 3-16.

11. Makropoulos, C. K. and Butler, D. 2006. Spatial ordered weighted averaging: Incorporating spatially variable attitude towards risk in spatial multicriteria decision-making. *Environmental Modelling & Software*. 21(1): 69-84.

12. Mandic, D.P. Javidi, S. Goh, S. L. Kuh, A. and Aihara, K. 2009. Complex-valued prediction of wind profile using augmented complex statistics. *Renewable Energy*. 34: 196-201.

13. Mitsukura, Y. Ito, S. Fukumi, M. and Akamatsu, N. 2002. Genetic fog occurrence forecasting system using a LVQ network. in: M.H. Hamza (Ed.), *Applied Informatics*.

نحوی که این استراتژی دقت نتایج را بر حسب شاخص نش نسبت به شبکه عصبی ۵۱/۸، ۳۸/۱ و ۴۴/۵ درصد و نسبت به بهترین مدل منفرد ۷/۶، ۱۳۲ و ۵۲/۹ درصد به ترتیب در دی، بهمن و اسفندماه بهبود داده است.

• مقایسه نتایج استراتژی ۱ و ۲ نیز نشان می دهد که استراتژی ۲ در هر سه ماه نتایج بهتری نسبت به استراتژی ۱ ایجاد نموده است به نحوی که این استراتژی دقت نتایج را نسبت به استراتژی ۱ بر حسب شاخص نش، ۵/۱۱، ۱۰۹/۶۴ و ۵۴/۱ درصد به ترتیب در دی، بهمن و اسفندماه ارتقا داده است.

با توجه به نتایج به دست آمده در تحقیق حاضر، می توان گفت که استراتژی ترکیب مدل ها بر اساس وزن دهی متغیر به مدل های منفرد می تواند به طور قابل توجهی دقت نتایج پیش بینی را بهبود بخشد. علت برتری این استراتژی نسبت به استراتژی شبکه عصبی در این است که شبکه عصبی بوسیله خروجی های مدل های منفرد آموزش می بیند و این امر سبب می شود که خطای موجود در نتایج مدل های منفرد بر آموزش شبکه عصبی اثر گذارد؛ ولی در روش ارائه شده در این مقاله، وزن ها مستقل از میزان خطای مدل ها محاسبه می شوند و وزن مدل های منفرد در هر گام زمانی بر حسب میزان خطا، تغییر می نماید. این امر سبب می شود که دقت نتایج استراتژی ۲ از دقت نتایج حاصل از شبکه عصبی بیشتر شود.

همچنین، مزیت تعیین وزن در روش OWA به سایر روش های تعیین وزن نظیر روش استکینگ و یا روش حداقل مربعات در این است که این روش به کاربر امکان می دهد که با توجه به مقدار پارامتر α ، به تنوع وسیعی از ترکیب وزن ها دست یابد و با توجه به اهمیت مدل ها، به هر مدل وزن مناسب را اختصاص دهد. علاوه بر این، در این روش مقادیر وزن مدل های منفرد به صورت مستقل از میزان خطای مدل ها تعیین می شود و این امر سبب می شود که قابلیت های موجود در نتایج تمامی مدل های منفرد بکار گرفته شود.

تشکر و قدردانی

بدین وسیله از دانشگاه تهران و شرکت مدیریت منابع آب ایران به دلیل تأمین داده ها و امکانات لازم جهت انجام این تحقیق و تهیه مقالات آن تشکر و قدردانی می نمایم.

منابع

1. Abrahart, R.J. and See, L. 2002. Multi-model data fusion for river flow forecasting: an evaluation of six alternative methods based on two contrasting catchments. *Hydrology and Earth System Sciences*. 6: 655-670.
2. Araghinejad, S. 2014. *Data-derived modeling: using MATLAB in water resources and environmental engineering*. Springer, New York, 292p.
3. Araghinejad, S. Azmi, M. and Kholghi, M. 2011.

20. Somlikova, R. and Wachowiak, M.P. 2001. Aggregation operators for selection problems. *Fuzzy Sets and Systems*. 131: 23-34.
21. Srinivas, V.V. and Srinivasan, K. 2001. Post-blackening approach for modeling periodic streamflows. *Hydrology*. 241: 221-269.
22. Xiong, L. Shamseldin, A. Y. and O'Connor, K. M. 2001. A non-linear combination of the forecasts of rainfall-runoff models by the first-order Takagi-Sugeno fuzzy system. *Hydrology*. 245: 196-217.
23. Xu, Z. 2006. Induced uncertain linguistic OWA operators applied to group decision making. *Information Fusion*. 7: 231-238.
24. Yager, R. R. and Xu, Z. 2006. The continuous ordered weighted geometric operator and its application to decision making. *Fuzzy Sets and Systems*. 157: 1393-1402.
25. Yager, R. R. and Filev, D. P. 1994. Parameterized "andlike" and "orlike" OWA operators. *International. Gen. Systems*. 22: 297-316.
26. Yager, R. R. 1988. On ordered weighted averaging aggregation operators in multi criteria decision making. *IEEE Trans Syst Man Cybern*. 18:183-190.
27. Zhou, L. G. and Chen, H. Y. 2011. Continuous generalized OWA operator and its application to decision making. *Fuzzy Sets and Systems*. 168: 18-34.
- Track, 285- 351.
14. Modaresi, F. Araghinejad, S. and Ebrahimi, K. 2015. Promotion of the accuracy of monthly streamflow forecast using Orness method for model fusion. (Case study: forecast of inflow to Seimareh Reservoir). In *Proceeding of the first National Congress on Iran's Irrigation and Drainage*, 13 and 14 May, Mashhad, Iran. (In Persian)
15. See, L. and Abraham, R.J. 2001. Multi-model data fusion for hydrological forecasting. *Computers & Geosciences*. 27: 987-994.
16. Shamseldin, A. Y. O'Connor, K.M. and Nasr, A. E. 2007. A comparative study of three neural network forecast combination methods for simulated river flows of different rainfall-runoff models. *Hydrological Sciences*. 52(5): 896-916.
17. Shamseldin, A. Y. and O'Connor, K.M. 2001. A non-linear neural network technique for updating of river flow forecasts. *Hydro Earth Syst Sci*. 5: 577-597.
18. Shamseldin, A. Y. O'Connor, K.M. and Liang, G.C. 1997. Methods for combining the output of different rainfall-runoff models. *Hydrology*. 197: 203-229.
19. Smith, P.N. 2006. Flexible aggregation in multiple attribute decision making: application to the Kuranda Range road upgrade. *Cybernetics and Systems: An International Journal*. 37: 1-22.

*Abstract***Assessment of Ordered Weighted Averaging Strategies in Combination of Streamflow Forecasting Models Case study: Karkheh River**F. Modaresi¹, S Araghinejad² and K. Ebrahimi³

Received: 2015/05/24 Accepted: 2016/07/05

Monthly streamflow forecasting plays an important role in long-lead water resources planning and management. In the current paper, model fusion technique has been used in order to increase the accuracy of monthly streamflow forecast of Karkheh River at the entrance of Karkheh reservoir in winter. For this purpose, five models including: Artificial Neural Network (ANN), Generalized Regression Neural Network (GRNN), Support Vector Regression (SVR), K-Nearest Neighbor (KNN), and Linear Regression (LR) with optimized structure have been applied as individual forecasting models (IFMs). In order to combine the IFM models, two model fusion strategies including constant and variable weighting based on ordered weighted averaging (OWA) have been performed, where the Orlike method has been applied to determine the weights of IFMs. The results show that variables weighting strategy is more performable than constant weighting strategy in order to promote the accuracy of the forecast results. In addition, the comparison of the two strategies including model fusion with artificial neural network and selecting the best IFM reveals that variable weighting strategy can significantly promote the accuracy of the forecast results than the latest strategies; such that this strategy increases the accuracy of the results 51.8, 38.1, and 44.5 percent as compared to ANN model fusion, and 7.6, 132, and 52.9 percent as compared to the best IFM for January, February, and March, respectively.

Keywords: *Model fusion, Ordered Weighted Averaging, Variable Weighting, Orlike Method, Karkheh.*

1. PhD of Water Resources Engineering, University of Tehran, Iran. Email: fmodaresi@ut.ac.ir

2. Assistant Professor, University of Tehran, Iran. Email: araghinejad@ut.ac.

3. Associate Professor, University of Tehran, Iran. Corresponding Author Email: EbrahimiK@ut.ac.ir