



## ارتقای دقت پیش‌بینی جریان ماهانه با استفاده از Orness Method برای ترکیب مدل‌ها (مطالعه موردی: پیش‌بینی جریان ورودی به سد سیمره)

فرشته مدرسی\* - دانشجوی دکتری مهندسی منابع آب دانشگاه تهران

شهاب عراقی نژاد - استادیار دانشگاه تهران

کیومرث ابراهیمی - دانشیار دانشگاه تهران

\*تلفن نویسنده مسئول: ۰۹۱۲۵۱۴۲۳۷۰ - پست الکترونیکی: fmodaresi@ut.ac.ir

### چکیده

یکی از مسائل اساسی در مدیریت منابع آب، پیش‌بینی بلند مدت جریان است. در تحقیق حاضر برای افزایش دقت پیش‌بینی‌های ماهانه جریان بهاره ورودی به سد سیمره از دیدگاه ترکیب مدل‌ها استفاده شده است. برای این منظور، پنج مدل داده مبنا 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 برای پیش‌بینی جریان مورد استفاده قرار گرفتند. در این تحقیق، برای ترکیب نتایج مدل‌های مذکور از استراتژی وزن‌دهی به مدل‌های منفرد بر اساس Orness Method استفاده شده است. در این روش، وزن‌ها بر اساس حل یک مدل بهینه‌سازی غیرخطی با روش الگوریتم ژنتیک محاسبه می‌شوند. نتایج این روش با دو استراتژی دیگر شامل استراتژی ترکیب مدل‌ها با شبکه عصبی مصنوعی و استراتژی انتخاب بهترین مدل منفرد مقایسه شد. نتایج نشان می‌دهند که بکارگیری Orness Method سبب افزایش قابل ملاحظه‌ای در دقت نتایج پیش‌بینی در مقایسه با دو استراتژی دیگر می‌شود؛ به طوری که دقت نتایج پیش‌بینی جریان برای فروردین، اردیبهشت و خردادماه به ترتیب ۸۰٪، ۳۴٪ و ۲۴٪ نسبت به استراتژی انتخاب بهترین نتایج مدل‌های منفرد و ۶۵٪، ۷۵٪ و ۲۴٪ نسبت به استراتژی بکارگیری شبکه عصبی افزایش یافته است.

**کلید واژه:** ترکیب مدل‌ها، پیش‌بینی بلندمدت جریان، Orness Method، مدل‌های شبکه عصبی، ماشین بردار پشتیبان، K نزدیک‌ترین همسایگی

# ارتقای دقت پیش‌بینی جریان ماهانه با استفاده از Orness Method برای ترکیب مدل‌ها (مطالعه موردی: پیش‌بینی جریان ورودی به سد سیمره)

فرشته مدرسی - دانشجوی دکتری مهندسی منابع آب دانشگاه تهران\*

شهاب عراقی نژاد - استادیار دانشگاه تهران

کیومرث ابراهیمی - دانشیار دانشگاه تهران

\*تلفن نویسنده اصلی: ۰۹۱۲۵۱۴۲۳۷۰، شماره: ۸۸۳۳۹۵۹۲ (۰۲۱)، پست الکترونیک: fmodaresi@ut.ac.ir

## چکیده

یکی از مسائل اساسی در مدیریت منابع آب، پیش‌بینی بلند مدت جریان است. در تحقیق حاضر برای افزایش دقت پیش‌بینی‌های ماهانه جریان بهاره ورودی به سد سیمره از دیدگاه ترکیب مدل‌ها استفاده شده است. برای این منظور، پنج مدل داده مبنا 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 برای پیش‌بینی جریان مورد استفاده قرار گرفتند. در این تحقیق، برای ترکیب نتایج مدل‌های مذکور از استراتژی وزن‌دهی به مدل‌های منفرد بر اساس روش Orness Method استفاده شده است. در این روش، وزن‌ها بر اساس حل یک مدل بهینه‌سازی غیرخطی با روش الگوریتم ژنتیک محاسبه می‌شوند. نتایج این روش با دو استراتژی دیگر شامل استراتژی ترکیب مدل‌ها با شبکه عصبی مصنوعی و استراتژی انتخاب بهترین مدل منفرد مقایسه شد. نتایج نشان می‌دهند که بکارگیری روش Orness Method سبب افزایش قابل ملاحظه‌ای در دقت نتایج پیش‌بینی در مقایسه با دو استراتژی دیگر می‌شود؛ به طوری که دقت نتایج پیش‌بینی جریان برای فروردین، اردیبهشت و خردادماه به ترتیب  $۸۰\%$ ،  $۳۴\%$  و  $۲۴\%$  نسبت به استراتژی انتخاب بهترین نتایج مدل‌های منفرد و  $۶۵\%$ ،  $۷۵\%$  و  $۲۴\%$  نسبت به استراتژی بکارگیری شبکه عصبی افزایش یافته است.

**کلید واژه‌ها:** ترکیب مدل‌ها، پیش‌بینی بلندمدت جریان، Orness Method، مدل‌های شبکه عصبی، ماشین بردار پشتیبان،  $K$  نزدیک‌ترین همسایگی

## ۱- مقدمه

یکی از مسائل اساسی در مدیریت بلندمدت مخازن سدها، اطلاع از میزان آورد رودخانه‌ها به مخازن می‌باشد. از این رو، پیش‌بینی بلندمدت جریان رودخانه‌ها در مقیاس‌های زمانی نظیر ماهانه، فصلی و سالانه از اهمیت ویژه‌ای برخوردار است. با این وجود، برای برنامه‌ریزی تخصیص آب به مصارف مختلف نظیر تأمین آب کشاورزی، شهری و صنعت، اسکیل زمانی

ماهانه نسبت به سایرین ارجحیت دارد زیرا در این صورت می توان در هر ماه برای ماه بعد برنامه ریزی های لازم را انجام داد و بدین ترتیب برنامه ریزی های دقیق تر و قابل اعتمادتری را بر منابع آب اتخاذ نمود. باتوجه به اهمیت پیش بینی جریان در مقیاس ماهانه، ضروری است که دقت این نوع پیش بینی تا حد ممکن افزایش یابد.

در این زمینه، محققین تلاش کرده اند تا با بکارگیری مدل های پیش بینی کننده کارا تر نظیر شبکه های عصبی مصنوعی (NN<sup>۱</sup>) [۳ و ۲]، ماشین بردار پشتیبان (SVM<sup>۲</sup>) [۴ و ۵]، K نزدیک ترین همسایگی (KNN<sup>۳</sup>) [۷ و ۹] و غیره بتوانند دقت پیش بینی های ماهانه را افزایش دهند. با این وجود، پیش بینی های صورت گرفته بوسیله هر یک از این روش ها با درصدی خطا همراه است که اجتناب ناپذیر می باشد.

تکنیک ترکیب مدل ها تکنیک نسبتا جدیدی است که از سال ۲۰۰۰ در پیش بینی های کوتاه مدت و بلند مدت زمانی، مکانی و زمانی-مکانی بخصوص در علوم مربوط به منابع طبیعی و محیط زیست نظیر به پیش بینی پدیده مه [۱۰]، پیش بینی زمانی رطوبت خاک [۱۱]، پیش بینی کوتاه مدت و بلندمدت دینامیکی اقیانوس ها [۱۲] و پیش بینی سرعت و پروفیل باد [۱۳] در راستای افزایش دقت پیش بینی ها مورد استفاده قرار گرفت. ترکیب مدل ها زیر مجموعه ای از ترکیب اطلاعات می باشد به گونه ای که نتیجه حاصل از آمیختن داده های به دست آمده از منابع مختلف، جامع تر، دقیق تر و مطمئن تر از نتایج حاصل از هر یک از منابع اولیه باشد [۱۴]. در زمینه پیش بینی های کوتاه مدت و بلندمدت هیدرولوژیکی نیز مفاهیم ترکیب مدل ها برای افزایش دقت پیش بینی ها بکاربرده شده است [۱۵ و ۱۶ و ۱۷ و ۱۸ و ۱۹ و ۲۰ و ۲۱ و ۲۲ و ۲۳ و ۲۴].

تکنیک ترکیب مدل ها در قالب دیدگاه های موازی، سری و یا ترکیب آنها صورت می گیرد [۲۵]. در دیدگاه ترکیب مدل های موازی، نتایج حاصل از چندین مدل منفرد با روش هایی نظیر وزدن دهی<sup>۴</sup> [۱۵ و ۱۷ و ۲۰ و ۲۱] و خودراه اندازی<sup>۵</sup> [۲۲] با یکدیگر ترکیب می شوند تا نتایج دقیق تری حاصل شود در حالیکه در دیدگاه ترکیب مدل های سری، نتایج حاصل از یک مدل به عنوان ورودی به مدل دیگر وارد می شود [۲۳]. در دیدگاه ترکیبی که اغلب در تحقیقات هیدرولوژیکی مورد استفاده قرار می گیرد، نتایج حاصل از چندین مدل منفرد به عنوان ورودی به یک مدل دیگر وارد می شود [۱۶ و ۱۷ و ۱۸ و ۱۹ و ۲۰ و ۲۴]. در تحقیقات مذکور، مدل شبکه عصبی مصنوعی به عنوان بهترین مدل ترکیب کننده شناخته شده است که سبب ارتقای نتایج نسبت به بهترین مدل منفرد شده است.

از آنجایی که مدل سازی با شبکه عصبی مصنوعی از نوع مدل سازی باناظر است که نیاز به آموزش دارد، این امر سبب می شود که خطای موجود در نتایج مدل های منفرد بر آموزش این مدل اثر بگذارند و دقت آن را کاهش دهند. بنابراین اگر برای ترکیب مدل ها بتوان از روش های وزن دهی مناسبی (دیدگاه موازی) استفاده نمود که نیازی به مرحله آموزش نداشته باشند، می توان دقت نتایج را افزایش بیشتری داد. در تحقیقاتی که از روش های وزن دهی استفاده شده است، برای وزن دهی به نتایج مدل های منفرد اغلب از روش های میانگین گیری ساده (وزن های یکسان برای تمامی مدل ها) استفاده شده است [۱۵ و ۱۶ و ۱۷ و ۲۰ و ۲۱]. در تحقیقات معدودی که روش وزن دهی به مدل های منفرد با وزن های متفاوت بکار رفته است، برای تعیین وزن ها از روش هایی نظیر روش حداقل مربعات<sup>۶</sup> [۱۵ و ۱۷] و روش استکینگ<sup>۷</sup> [۲۰] استفاده شده است. به دلیل آنکه مبنای تعیین وزن

<sup>1</sup> Neural Network

<sup>2</sup> Support Vector Machine

<sup>3</sup> K-Nearest Neighbor

<sup>4</sup> Weighting

<sup>5</sup> Bootstrap

<sup>6</sup> Ordinary Least Square Error

<sup>7</sup> Staking

مدل‌های منفرد در این دو روش، حداقل نمودن مجموع خطاهای مقادیر شبیه‌سازی شده توسط مدل‌ها از مقادیر مشاهداتی می‌باشد، در این روش‌ها، وزن زیادی (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. در حالیکه اگر تعیین وزن‌ها مستقل از میزان خطای مدل‌ها صورت گیرد، می‌توان به نتایج مناسبی دست یافت به طوری که دقت نتایج حاصل از روش ترکیب مدل‌ها بیشتر از دقت نتایج بهترین مدل منفرد شود.

با توجه به این مسئله، ضرورت یافتن یک روش وزن‌دهی مناسب به نحوی که سبب ارتقای نتایج حاصل از ترکیب مدل‌ها شود، وجود دارد. در این راستا در تحقیق حاضر روش وزن‌دهی جدیدی با نام Orness Method برای ترکیب نتایج حاصل از پیش‌بینی جریان توسط مدل‌های منفرد مورد استفاده و ارزیابی قرار گرفته است. Orness Method روش اصلی تکنیک وزن‌دهی مرتبه‌ای<sup>۱</sup> است که به طور اختصار با عنوان OWA اطلاق می‌شود. روش وزن‌دهی OWA اولین بار توسط Yager در سال ۱۹۸۸ ارائه شد [۲۶] و کاربرد اصلی آن در تحقیقات صورت گرفته در زمینه فرآیندهای تصمیم‌گیری<sup>۲</sup> بوده است [۲۷ و ۲۸ و ۲۹]. زیرا در این روش، دامنه وسیعی از وزن‌ها ایجاد می‌شود که با توجه به اهمیت منابع مختلف، هر وزن می‌تواند به منبع خاصی اختصاص یابد. در این تحقیق، روش مذکور برای وزن‌دهی به مدل‌های منفرد پیش‌بینی‌کننده جریان مورد استفاده قرار گرفته است و کارایی آن برای پیش‌بینی جریان ماهانه ورودی به مخزن سد سیمره در فصل بهار در مقایسه با شبکه عصبی مورد ارزیابی قرار گرفته است.

## ۲- معرفی منطقه مطالعاتی و داده‌های مورد استفاده

در تحقیق حاضر، پیش‌بینی جریان ماهانه ورودی به سد سیمره در فصل بهار مورد ارزیابی قرار گرفته است. سد سیمره یک از سدهای اصلی حوضه کرخه می‌باشد که در بالادست سد کرخه در طول جغرافیایی ۴۷° ۱۲' ۴۷" و عرض جغرافیایی ۳۳° ۱۶' ۴۱" بر روی رودخانه سیمره قرار دارد. رودخانه سیمره 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 توسط سرشاخه‌های فرعی زیادی شکل می‌گیرند. این دو رود در انتهای دشت کرمانشاه به هم پیوسته و رودخانه سیمره را تشکیل می‌دهند [۳۰]. در شکل ۱ موقعیت سد سیمره در حوضه کرخه نشان داده شده است.

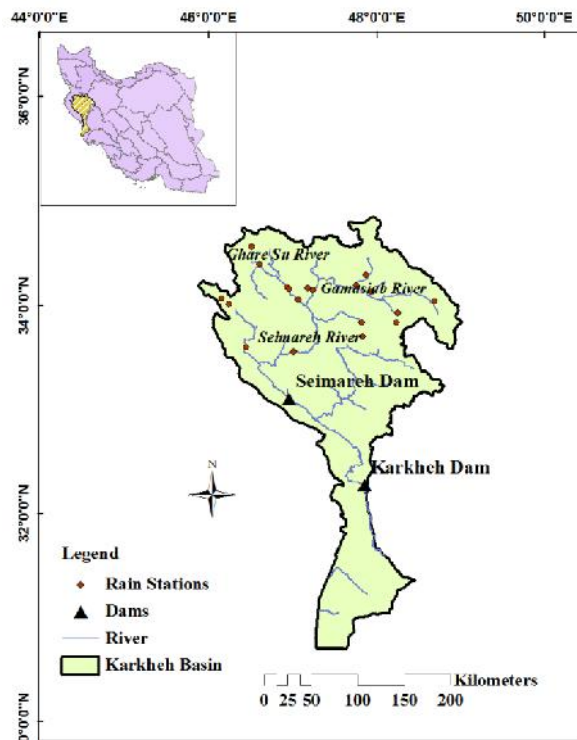
برای پیش‌بینی جریان ماهانه ورودی به سد سیمره در فصل بهار، اطلاعات ۳۲ ساله متغیرهای هواشناسی و هیدرولوژیکی شامل بارش متوسط ماهانه ۱۹ ایستگاه بارش واقع در حوضه بالادست سد سیمره با تأخیر زمانی یک تا پنج ماه (موقعیت ایستگاه‌های بارش در شکل ۱ نشان داده شده است)، میزان سطح پوشش برفی (SAE<sup>۳</sup>) در حوضه بالادست سد در سه ماه قبل از ماه مورد پیش‌بینی و جریان ماهانه ورودی به سد سیمره در پنج ماه قبل از ماه مورد پیش‌بینی در بازه زمانی ۱۳۶۱-۱۳۹۲ 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<sup>۴</sup> (MI) که به ترتیب نشان دهنده میزان همبستگی خطی و غیرخطی میان پیش‌بینی‌کننده‌ها و پیش‌بینی شونده‌ها هستند، استخراج شده و مورد استفاده قرار گرفتند.

<sup>1</sup> Ordered Weighted Aggregation

<sup>2</sup> Decision Making

<sup>3</sup> Snow Area Extent

<sup>4</sup> Mutual Information



شکل شماره ۱- موقعیت قرارگیری سد سیمره در حوضه کرخه

### ۳- روش انجام تحقیق

تحقیق حاضر در سه مرحله کلی صورت گرفته است که به صورت زیر می باشد:

**مرحله اول:** پیش بینی جریان ماهانه بر اساس پیش بینی کننده های منتخب با استفاده از پنج مدل پیش بینی کننده شناخته شده

متفاوت به عنوان مدل های منفرد پیش بینی کننده جریان

**مرحله دوم:** ترکیب نتایج حاصل از مدل های منفرد پیش بینی کننده جریان با استفاده از سه استراتژی متفاوت شامل:

**الف)** ترکیب مدل ها بر اساس روش وزن دهی Orness Method

**ب)** ترکیب مدل ها با استفاده از روش شبکه عصبی مصنوعی

**ج)** انتخاب بهترین مدل منفرد با کمترین خطا (این روش معادل با دادن وزن ۱ به بهترین مدل و وزن صفر به سایر

مدل ها است)

**مرحله سوم:** ارزیابی نتایج حاصل از سه استراتژی فوق بر اساس سه معیار ارزیابی شناخته شده و انتخاب بهترین نتیجه

الگوریتم مراحل انجام تحقیق در شکل شماره ۲ نشان داده شده است. در ادامه، روش انجام هر مرحله و مدل های بکار رفته در

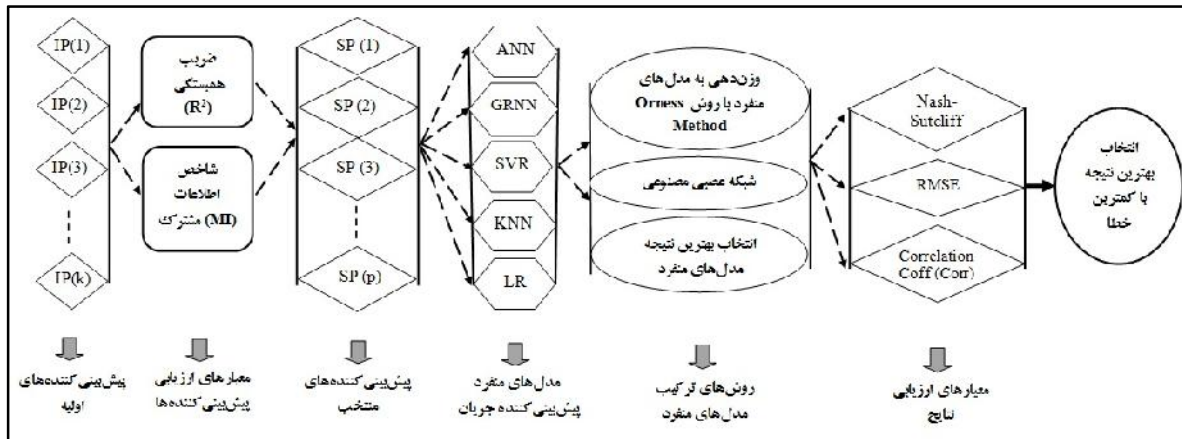
هر بخش تشریح شده است:

#### ۳-۱- مدل های منفرد پیش بینی جریان

در تحقیق حاضر، پنج مدل داده مبناء شناخته شده با ساختاری بهینه به عنوان مدل های منفرد برای پیش بینی جریان

مورد استفاده قرار گرفته اند. این ساختار بهینه به نحوی تعیین شده است که نتایج مرحله صحت سنجی دارای کمترین خطا باشد.

در جدول شماره ۱، نام و نام اختصاری مدل‌ها و نام پارامترهای بهینه شده برای هر مدل بیان شده است. جزئیات بیشتر درباره ساختار مدل‌های در [۳۱] موجود می‌باشد.



شکل شماره ۲- الگوریتم مراحل انجام تحقیق

جدول شماره ۱- نام و خصوصیات مدل‌های منفرد

نام کامل	نام اختصاری	نوع پارامترهای بهینه شده
شبکه عصبی مصنوعی <sup>۱</sup>	ANN	وزن‌های اولیه و تعداد نورون‌های لایه میانی
شبکه عصبی رگرسیون تعمیم یافته <sup>۲</sup>	GRNN	پارامتر تنظیم مدل (پارامتر سیگما)
رگرسیون بردار پشتیبان <sup>۳</sup>	SVR	نوع تابع کرنل، پارامترهای تابع کرنل و پارامتر تنظیم مدل (پارامتر گاما)
K نزدیک‌ترین همسایگی <sup>۴</sup>	KNN	وزن متغیرهای پیش‌بینی کننده و تعداد همسایگی‌ها (پارامتر K)
رگرسیون خطی <sup>۵</sup>	LR	ضریب متغیرهای پیش‌بینی کننده

### ۲-۳ روش ترکیب مدل‌های Orness Method

روش Orness Method یکی از انواع روش‌های عملگر وزن‌دهی مرتبه‌ای (OWA<sup>۶</sup>) است. عملگر OWA با ابعاد  $m$  نگاشتی 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]:

$$\sum_{i=1}^n w_i = 1 \text{ و } w_i \in [0,1]$$

به طور کلی برای اعمال عملگر OWA بر روی یک بردار نتایج حاصل از مدل‌ها یا حسگرهای مختلف، از سه گام زیر

باید پیروی شود [۳۲]:

(الف) مرتب کردن نزولی بردار نتایج حاصل از مدل‌ها و یا حسگرها

(ب) استخراج اوزان مربوط به هر مقدار بردار نتایج نزولی شده با استفاده از روش OWA مورد نظر کاربر

<sup>1</sup> Artificial Neural Network

<sup>2</sup> Generalized Regression Neural Network

<sup>3</sup> Support Vector Regression

<sup>4</sup> K-Nearest Neighbor Regression

<sup>5</sup> Linear Regression

<sup>6</sup> Ordered Weighted Aggregation

ج) اعمال عملگر OWA برای محاسبه مقدار برآیند بردار نتایج

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:

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

مقادیر عملگر OWA در بازه‌ای میان کمینه و بیشینه مقادیر  $a_i$  قرار می‌گیرند. هاگان [۳۳] برای تعیین بهترین وزن‌ها از مدل بهینه‌سازی زیر استفاده نمود که در آن به ازای مقدار مشخصی از پارامتر Orness، بیشینه مقدار برای پارامتر پراکندگی (Disp) ایجاد شود.

$$Maximize : -\sum_{i=1}^n w_i \ln w_i \quad (2)$$

$$Subject\ to: \frac{1}{n-1} \sum_{i=1}^n (n-i) w_i = \alpha \quad 0 \leq \alpha \leq 1 \quad (3)$$

$$\sum_{i=1}^n w_i = 1, \quad 0 \leq w_i \leq 1, \quad i = 1, 2, \dots, n$$

در رابطه ۳، مقدار پارامتر  $\alpha$  باید با توجه به دقت موردنظر در پیش‌بینی، توسط کاربر تعیین شود. با توجه به اینکه برای مقادیر  $\alpha$  کوچکتر از ۰/۵، وزن‌های متغیرها به صورت صعودی و برای بزرگتر از ۰/۵، به صورت نزولی هستند، برای تطابق میان وزن‌ها و متغیرها از مقادیر  $\alpha$  بزرگتر از ۰/۵ برای اجرای فرآیند بهینه‌سازی و تعیین وزن‌ها استفاده شده است. در تحقیق حاضر از روش بهینه‌سازی فراکاوشی الگوریتم ژنتیک برای محاسبه وزن‌های بهینه استفاده شده است.

مزیت تعیین وزن در این روش به سایر روش‌های تعیین وزن نظیر روش استکینگ و یا روش حداقل مربعات در این است که روش Orness Method به کاربر این امکان را می‌دهد که با توجه به مقدار پارامتر  $\alpha$ ، 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. بنابراین بجای اختصاص یک وزن خاص به هر مدل برای تمامی مقادیر پیش‌بینی شده، وزن مدل‌ها برای هر یک از مقادیر پیش‌بینی شده قابل تغییر است و این امر سبب ارتقای دقت نتایج حاصله خواهد شد.

### ۳-۳- 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 پیشخور با ساختاری بهینه از نظر تعداد نوروها و وزن‌های اولیه استفاده شده است و ساختار بهینه برای هر ماه به نحوی تعیین شده است که نتایج مرحله صحت سنجی از بیشترین دقت برخوردار باشند. برای آموزش این مدل نیز از خروجی مدل‌های منفرد در دوره واسنجی استفاده شده است.

### ۳-۴- روش‌های ارزیابی نتایج

در تحقیق حاضر، برای ارزیابی نتایج پیش‌بینی‌های صورت گرفته، علاوه بر دو معیار ریشه متوسط مربعات خطا<sup>۱</sup> (RMSE) و ضریب همبستگی<sup>۲</sup> ( $R^2$ ) میان مقادیر آبدهی واقعی و مدل‌سازی شده، از شاخص دیگری با عنوان شاخص ارزیابی مدل نش- ساتکلیف<sup>۳</sup> استفاده شده است. مقدار این شاخص از ۱ تا  $-\infty$  تغییر می‌نماید. اگر مقدار آن برابر با یک شود بدین معناست که نتایج حاصل از مدل کاملاً با مقادیر مشاهداتی برابر است. جزئیات بیشتر درباره این شاخص در [۳۴] موجود است.

### ۴- نتایج و بحث

#### ۴-۱- نتایج حاصل از اجرای مدل‌های منفرد

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 شماره ۲ آورده شده است. همان‌طور که در جدول ۲ مشاهده می‌شود، بجز مدل LR که مقادیر شاخص Nash آن برای ماه‌های فروردین و اردیبهشت منفی است، سایر مدل‌ها کارایی نسبتاً مناسبی برای پیش‌بینی جریان ماهانه داشته‌اند و مقادیر شاخص Nash آنها مثبت و مقادیر ضریب همبستگی آنها بزرگتر از ۰/۵ و مقادیر شاخص RMSE آنها کوچکتر از یک است. با این وجود، بر اساس شاخص‌های ارزیابی می‌توان گفت که نتایج حاصله از دقت بالایی برخوردار نیستند. بهترین نتایج حاصله برای ماه‌های فروردین و خرداد متعلق به مدل KNN و برای اردیبهشت ماه متعلق به مدل GRNN است. مقایسه بهترین نتایج بیانگر دقت پایین‌تر پیش‌بینی آبدهی فروردین ماه نسبت به دو ماه دیگر است. به علت پایین بودن دقت نتایج، لزوم بکارگیری روش‌های ترکیب مدل‌ها وجود دارد.

جدول شماره ۲- نتایج مرحله صحت سنجی پیش‌بینی توسط مدل‌های منفرد

مدل‌های منفرد					شاخص‌های ارزیابی مدل‌ها	ماه پیش‌بینی آبدهی
LR	KNN	SVR	GRNN	ANN		
-۰/۰۵۵	۰/۲۸۵	۰/۲۲۲	۰/۲۱۳	۰/۲۳۵	Nash-Sutcliffe	فروردین
۰/۹۸۷	۰/۸۱۲	۰/۸۴۷	۰/۸۵۲	۰/۸۴۰	RMSE	
۰/۴۹۳	۰/۶۹۳	۰/۵۶۱	۰/۶۱۲	۰/۵۲۵	Correlation coefficient	
-۰/۲۳۸	۰/۳۷۱	۰/۴۸۱	۰/۴۸۶	۰/۴۲۶	Nash-Sutcliffe	اردیبهشت
۰/۶۹۱	۰/۴۹۲	۰/۴۴۷	۰/۴۴۵	۰/۴۷	RMSE	
۰/۵۷۳	۰/۶۴۸	۰/۶۹۴	۰/۷۰۰	۰/۷۴۰	Correlation coefficient	
۰/۱۸۰	۰/۵۵۸	۰/۲۲۱	۰/۳۴۱	۰/۲۲۱	Nash-Sutcliffe	خرداد
۰/۹۸۱	۰/۷۲۰	۰/۹۵۶	۰/۸۷۹	۰/۹۵۶	RMSE	
۰/۴۳۰	۰/۷۵۴	۰/۴۹۳	۰/۵۹۴	۰/۴۷۹	Correlation coefficient	

<sup>۱</sup> Root Mean Square Error

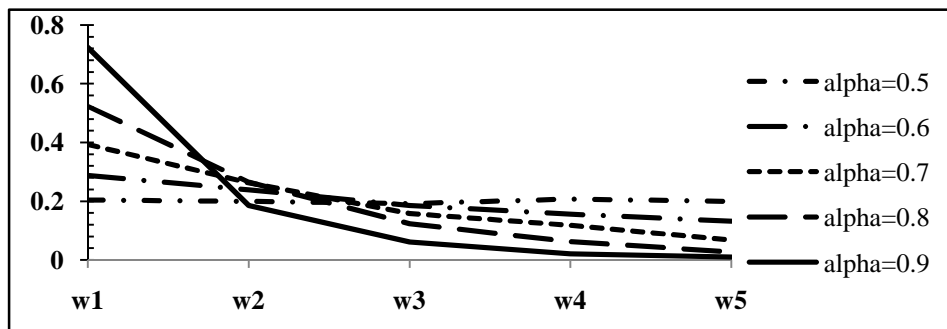
<sup>۲</sup> Correlation Coefficient

<sup>۳</sup> Nash-Sutcliffe model efficiency coefficient



#### ۲-۴- نتایج حاصل از استراتژی ترکیب مدل‌ها با روش Orness Method و مقایسه آن با دو استراتژی دیگر

برای بکارگیری روش Orness، وزن‌های لازم برای ۵ مدل منفرد، از روش بهینه‌سازی الگوریتم ژنتیک به ازای مقادیر مختلف  $\alpha$  استخراج شدند که به صورت نمودار در شکل شماره ۳ نشان داده شده است. همان‌طور که در شکل شماره ۳ مشاهده می‌شود، به ازای مقدار ۰/۵ برای  $\alpha$ ، وزن تمامی مدل‌ها تقریباً با هم برابر است. با زیاد شدن مقدار پارامتر  $\alpha$ ، وزن تعیین شده برای جواب‌های بهتر، بیشتر شده و از وزن جواب‌های نامناسب کاسته می‌شود.



شکل شماره ۳-وزن‌های حاصل از بکارگیری الگوریتم بهینه‌سازی ژنتیک برای Orness Method

بر اساس وزن‌های محاسبه شده، فرآیند ترکیب مدل‌های منفرد برای هر ماه صورت گرفته است و نتایج حاصل از آن با نتایج حاصل از دو استراتژی دیگر شامل استراتژی ترکیب مدل‌ها با شبکه عصبی و استراتژی انتخاب بهترین مدل منفرد (وزن ۱ به بهترین مدل و وزن صفر به سایر مدل‌ها) مقایسه شده است. نتایج حاصل، برای سه ماه فروردین، اردیبهشت و خرداد در جدول شماره ۳ آورده شده است.

همان‌طور که در جدول ۳ مشاهده می‌شود، نتایج حاصل از استراتژی ۱ در هر سه ماه به ازای مقدار ۰/۵ برای پارامتر  $\alpha$ ، که در آن، وزن داده شده به تمامی نتایج با هم برابر است (معادل با مینگین گیری حسابی) دارای دقت کمتری نسبت به نتایج حاصل از استراتژی ۳ است. ولی بر اساس هر سه شاخص ارزیابی مدل‌ها، در دو ماه فروردین و اردیبهشت، نتایج حاصل از Orness Method به ازای مقادیر  $\alpha$  از ۰/۶ تا ۰/۹ دارای دقت بیشتری نسبت به نتایج حاصل از بهترین مدل منفرد است. از میان مقادیر مختلف  $\alpha$ ، مقدار ۰/۹ سبب افزایش قابل ملاحظه‌ای در دقت نتایج شده است؛ به طوری که بر اساس شاخص Nash، با بکارگیری وزن‌های حاصل از این روش، دقت نتایج حاصل از پیش‌بینی جریان ماهانه به ترتیب برای فروردین و اردیبهشت، ۸۰٪ و ۳۴٪ نسبت به نتایج حاصل از بهترین مدل منفرد (استراتژی ۳) افزایش یافته است. افزایش چشمگیر دقت نتایج حاصل برای فروردین ماه که بر اساس نتایج بهترین مدل‌های منفرد از کمترین دقت برخوردار بود، بسیار ارزشمند و قابل ملاحظه است.

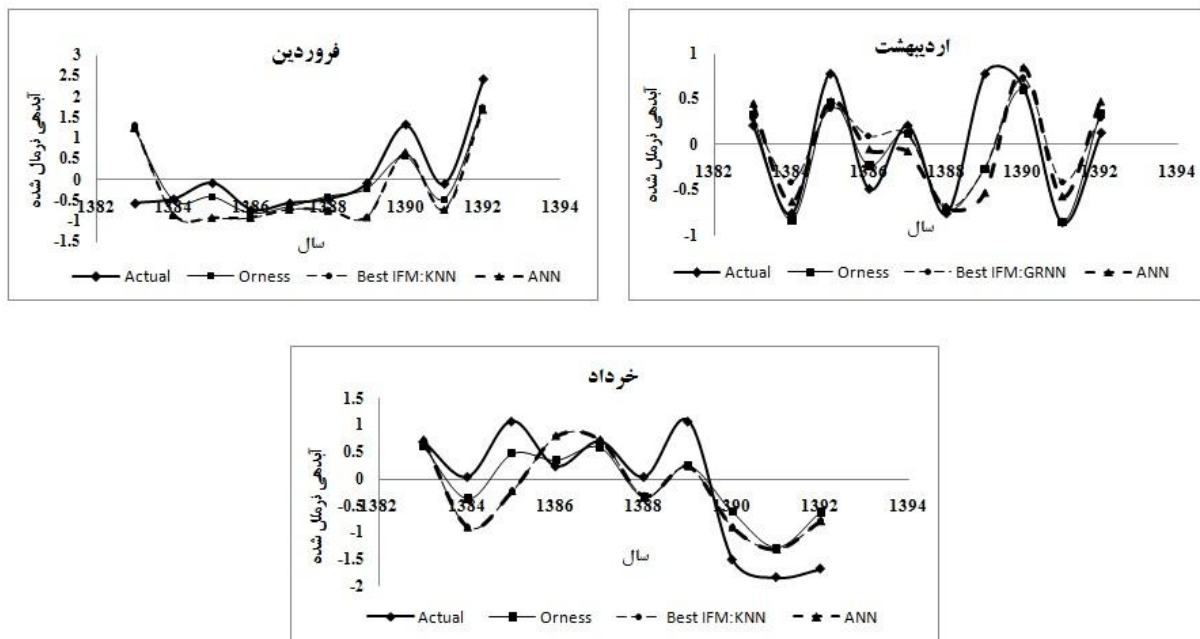
بر اساس سه شاخص ارزیابی مدل‌ها برای خردادماه نیز، بکارگیری Orness Method با مقادیر ۰/۸ و ۰/۹ برای  $\alpha$  سبب ارتقای نتایج پیش‌بینی نسبت به نتایج حاصل از بهترین مدل منفرد شده است؛ به طوری که با بکارگیری وزن‌های ایجاد شده به ازای  $\alpha$  برابر با ۰/۹ برای ترکیب مدل‌ها، شاخص Nash برای نتایج حاصل از استراتژی ۱، ۲۴٪ نسبت به این شاخص برای بهترین مدل منفرد افزایش یافته است.

نتایج حاصل از استراتژی ۲ (شبکه عصبی) نیز نشان می‌دهد که بکارگیری شبکه عصبی فقط برای ماه فروردین توانسته است دقت نتایج حاصل از بهترین مدل منفرد را ۸٪ افزایش دهد. درحالی که برای اردیبهشت ماه، دقت نتایج کاهش یافته و برای خردادماه، تغییری در دقت نتایج حاصل نشده است.

جدول شماره ۳- نتایج استراتژی‌های ترکیب مدل‌ها

استراتژی ۳	استراتژی ۲	Orness Method					شاخص‌های ارزیابی نتایج	ماه پیش‌بینی آبدهی	
		$\alpha=0/9$	$\alpha=0/8$	$\alpha=0/7$	$\alpha=0/6$	$\alpha=0/5$			
انتخاب بهترین مدل منفرد	ترکیب مدل‌ها با شبکه عصبی مصنوعی	0/285	0/309	0/511	0/461	0/399	0/325	0/244	فروردین
		0/812	0/799	0/672	0/705	0/745	0/789	0/835	
		0/693	0/700	0/728	0/701	0/669	0/634	0/598	
اردیبهشت		0/486	0/371	0/651	0/623	0/580	0/518	0/441	اردیبهشت
		0/445	0/492	0/367	0/381	0/402	0/431	0/464	
		0/700	0/647	0/821	0/801	0/771	0/732	0/688	
خرداد		0/558	0/558	0/689	0/610	0/533	0/454	0/374	خرداد
		0/720	0/720	0/604	0/676	0/740	0/800	0/857	
		0/754	0/754	0/895	0/861	0/814	0/750	0/672	

مقایسه نتایج استراتژی ۱ و ۲ نشان می‌دهد که نتایج Orness Method به ازای مقادیر  $\alpha$  از ۰/۶ تا ۰/۹ برای فروردین و اردیبهشت ماه دارای دقت بیشتری نسبت به نتایج شبکه عصبی است. همچنین، برای خردادماه به ازای مقادیر ۰/۸ و ۰/۹ برای  $\alpha$  نتایج Orness Method دارای دقت بیشتری نسبت به نتایج شبکه عصبی است. در شکل ۴ مقادیر جریان ماهانه پیش‌بینی شده حاصل از سه استراتژی بکار رفته در این تحقیق، با مقادیر واقعی جریان مقایسه شده است.



شکل شماره ۴-مقایسه مقادیر حاصل از استراتژی‌های ترکیب مدل‌ها با مقادیر واقعی در مرحله صحت سنجی

همان‌طور که در شکل ۴ مشاهده می‌شود، در هر سه ماه، مقادیر جریان پیش‌بینی شده توسط استراتژی ترکیب مدل‌ها با شبکه عصبی بسیار نزدیک به نتایج حاصل از بهترین مدل منفرد است. در حالیکه، نتایج حاصل از استراتژی ترکیب مدل‌ها با Orness Method به مقادیر واقعی نزدیک‌تر است که این امر تأیید کننده نتایج حاصل از معیارهای ارزیابی در جدول ۳ می‌باشد.

## ۵- جمع‌بندی و نتیجه‌گیری

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 با ساختار بهینه برای هر ماه مورد استفاده قرار گرفتند و بر اساس پیش‌بینی‌کننده‌ها منتخب برای هر ماه اجرا شدند. برای ترکیب مدل‌ها، استراتژی وزن‌دهی جدیدی با نام Orness Method برای ترکیب نتایج مدل‌های منفرد استفاده شده است و نتایج آن با دو استراتژی ترکیب مدل‌ها با شبکه عصبی و استراتژی انتخاب بهترین نتیجه مدل‌های منفرد مقایسه شده است. برای روش Orness Method، وزن‌ها به صورت مستقل از میزان خطای مدل‌های منفرد بر اساس یک فرآیند بهینه سازی غیرخطی با استفاده از الگوریتم ژنتیک تعیین شده‌اند.

نتایج حاصل از بکارگیری این استراتژی برای ترکیب مدل‌ها نشان می‌دهد که وزن‌دهی به مدل‌ها با این روش سبب ارتقای قابل ملاحظه‌ای در دقت نتایج نسبت به دو استراتژی دیگر می‌شود به طوری که دقت نتایج پیش‌بینی جریان برای فروردین، اردیبهشت و خردادماه به ترتیب  $80\%$ ،  $34\%$  و  $24\%$  نسبت به استراتژی انتخاب بهترین نتایج مدل‌های منفرد و  $65\%$ ،  $75\%$  و  $24\%$  نسبت به استراتژی بکارگیری شبکه عصبی افزایش یافته است. دلیل برتری این روش به روش شبکه عصبی در این است که شبکه عصبی بوسیله خروجی‌های مدل‌های منفرد که دارای خطا هستند، آموزش می‌بیند ولی در روش ارائه شده در این مقاله، وزن‌ها مستقل از میزان خطای مدل‌ها محاسبه می‌شوند. این امر سبب می‌شود که دقت نتایج روش پیشنهادی از دقت نتایج حاصل از شبکه عصبی بیشتر باشد.

## ۶- مراجع

- [1]. نوری، ر.، فرخ‌نیا، ا.، مرید، س. و ریاحی مدوار، ح. (۱۳۸۸). تأثیر پیش پردازش متغیرهای ورودی به شبکه عصبی برای پیش‌بینی جریان ماهانه با آنالیز مؤلفه‌های اصلی و موجک. فصلنامه علمی- پژوهشی آب و فاضلاب، دوره ۲۰، شماره ۱، ص ۱۳-۲۲.
- [2]. عراقی‌نژاد، ش. و کارآموز، م. (۱۳۸۴). پیش‌بینی بلندمدت رواناب با استفاده از شبکه‌های عصبی مصنوعی و سیستم استنتاج فازی. مجله علمی- پژوهشی تحقیقات منابع آب ایران، دوره ۱، شماره ۲، ص ۲۹-۴۱.
- [3]. Cigizoglu, H. K. (2005). Generalized regression neural network in monthly flow forecasting. *J. Civil Engineering and Environmental Systems.*, 22(2), 71-81.
- [4]. Wang, W.C., Chau K. W., Cheng, C. T. and Qiu, L. (2009). A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *J. Hydrology.*, 374, 294-306.
- [5]. Guo, J., Zhou, J., Qin, H., Zou, Q. and Li, Q. (2011). Monthly streamflow forecasting based on improved support vector machine model. *Expert System with Applications.* 38 (10), 13073-13081.
- [6]. نیک بخت شهبازی، ع. (۱۳۸۸). کاربرد ماشین بردار پشتیبان در پیش‌بینی جریان رودخانه. هشتمین کنفرانس هیدرولیک ایران، تهران، دانشگاه تهران.

- [7]. Wu, C.L. and Chau, K.W. (2010). Data-driven models for monthly streamflow time series prediction. *Engineering Applications of Artificial Intelligence.*, 23(8), 1350-1367.
- [8]. Prairie, J., Rajagopalan, B., Fulp, T., and Zagona, E. (2006). Modified K-NN Model for Stochastic Streamflow Simulation. *J. Hydrol. Eng.*, 11(4), 371-378.

- [۹]. نقدی بانسوله، ک. و موسوی، س. ج. (۱۳۹۲). پیش‌بینی جریان ورودی به مخزن سد زاینده‌رود با استفاده از الگوریتم پیش‌بینی K نزدیک‌ترین همسایگی (KNN). پنجمین کنفرانس مدیریت منابع آب ایران، تهران، دانشگاه شهید بهشتی.
- [10]. 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., Track, 285- 351.
- [11]. Khalili, A., Gill, M. K., and McKee, M. (2005). New applications for information fusion and soil moisture forecasting. 8<sup>th</sup> International Conference on Information Fusion., Philadelphia, 25-28 July.
- [12]. Leslie, W. G., Robinson, A. R., Haley Jr, 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. J. Marine Systems – Maritime Rapid Environmental Assessment – New Trends in Operational Oceanography., 69, 3–16.
- [13]. 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.
- [14]. Hall, D. L., and Llinas, J. (1997). An introduction to multisensor data fusion. in proceedings of the IEEE., 85, 6-10.
- [15]. 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. J. Hydrology., 245, 196-217.
- [16]. 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.
- [17]. Goswami, M., and O'Connor, K. M. (2007). Real-time flow forecasting in the absence of quantitative precipitation forecasts: a multi-model approach. J. Hydrology., 334, 125–140.
- [18]. 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 J., 52(5), 896-916.
- [19]. Araghinejad, S., Azmi, M., and Kholghi, M. (2011). Application of artificial neural network ensembles in probabilistic hydrological forecasting. J. Hydrology., 407, 94-104.
- [20]. 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.
- [21]. Shamseldin, A.Y., O'Connor, K.M., and Liang, G.C. (1997). Methods for combining the output of different rainfall-runoff models. J. Hydrology., 197, 203-229.
- [22]. Srinivas, V.V., and Srinivasan, K. (2001). Post-blackening approach for modeling periodic streamflows. J. Hydrology., 241, 221-269.
- [23]. Shamseldin, A.Y., and O'Connor, K.M. (2001). A non-linear neural network technique for updating of river flow forecasts. Hydrol. Earth Syst. Sci., 5, 577–597.
- [24]. See, L., and Abrahart, R.J. (2001). Multi-model data fusion for hydrological forecasting. Computers & Geosciences., 27, 987–994.
- [25]. Dasarathy, B. V. (1994). Decision Fusion. IEEE Computer Society Press, Los Alamitos, California., ISBN: 0818644524.
- [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]. Somlikova, R., and Wachowiak, M.P. (2001). Aggregation operators for selection problems. Fuzzy Sets and Systems, 131: 23-34.
- [28]. 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.
- [29]. 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.
- [۳۰]. شرکت مهندسين مشاور جاماب. (۱۳۷۸). طرح جامع آب کشور: حوضه آبریز کرخه.
- [31]. Araghinejad, S. (2014). Data-derived modeling: using MATLAB in water resources and environmental engineering. Springer, New York.
- [32]. Yager, R. R. (1992). Applications and extensions of OWA aggregation. Int J Man Mach Stud 37, 103–132.
- [33]. O'Hagan, M. (1988). Aggregating template rule antecedents in real-time expert systems with fuzzy set logic. In: Proc 22nd Annual IEEE Asilomar Conference on Signals, Systems and Computers. Pacific Grove, CA: IEEE and Maple Press; pp 681–689.
- [34]. Nash J. E., and Sutcliffe J. V. (1970). River flow forecasting through conceptual models part I- A discussion of principles. J. Hydrology, 10(3): 282-290.