

## پیش بینی خشک سالی با بهره گیری از مدل ترکیبی ماشین بردار پشتیبان موجکی و شاخص SPI (مطالعه موردی: حوضه دریاچه ارومیه-ایران)

مهدی کماسی<sup>\*۱</sup>

[komasi@abru.ac.ir](mailto:komasi@abru.ac.ir)

سروش شرقی<sup>۲</sup>

تاریخ پذیرش: ۹۶/۱۰/۶

تاریخ دریافت: ۹۶/۶/۱

### چکیده

**زمینه و هدف:** خشک سالی تهدیدی جدی برای انسان و محیط زیست بوده از این رو یافتن شاخصی جهت پیش بینی این پدیده از اهمیت به سزایی برخوردار است. شاخص بارش استاندارد شده (SPI) یک شاخص جامع جهت طبقه بندی شدت خشک سالی به حساب می آید. مدل های هوش مصنوعی کلاسیک از متداول ترین مدل هایی هستند که جهت پیش بینی شاخص SPI مورد استفاده قرار گرفته اند. از آن جایی که این مدل ها بر پایه ی ویژگی خودهم بستگی استوار هستند، بنابراین توانایی رصد نمودن سری های زمانی درازمدت و فصلی را دارا نمی باشند. در این پژوهش برای پیش بینی خشک سالی از مدل ترکیبی ماشین بردار پشتیبان موجکی و شاخص SPI استفاده شده است. **روش بررسی:** برای این منظور سری زمانی شاخص SPI مربوط به حوضه ارومیه توسط آنالیز موجک به چندین زیر سری با مقیاس های زمانی مختلف تبدیل شده و این زیر سری های زمانی به عنوان ورودی مدل ماشین بردار پشتیبان برای پیش بینی خشک سالی در نظر گرفته می شوند.

**یافته ها:** نتایج حاصل از صحت سنجی مدل ها بیان گر آن است که بیش ترین مقدار ضریب تبیین و کم ترین مقدار جذر میانگین مربع خطا برای مدل منفرد ماشین بردار پشتیبان به ترتیب ۰/۸۶۵ و ۰/۲۳۷ و برای مدل ترکیبی ماشین بردار پشتیبان موجکی به ترتیب ۰/۹۵۴ و ۰/۰۵۶ می باشد.

**بحث و نتیجه گیری:** بنابراین مدل ترکیبی ماشین بردار پشتیبان موجکی در مقایسه با مدل منفرد ماشین بردار پشتیبان توانایی به سزایی جهت پیش بینی سری زمانی شاخص SPI و نیز رصد نمودن نقاط بیشینه این سری زمانی به سبب در نظر گرفتن تغییرات فصلی دارا می باشد. از سوئی نشان داده شد که این مدل ترکیبی در مقایسه با سایر مدل های خودهم بسته کلاسیک هم چون شبکه عصبی مصنوعی از دقت و کارایی بالاتری برخوردار است.

**واژه های کلیدی:** مدل ماشین بردار پشتیبان، تبدیل موجک، شاخص SPI، پیش بینی خشک سالی، حوضه ارومیه ..

۱- استادیار گروه مهندسی عمران، دانشگاه آیت اله العظمی بروجردی (ره)، بروجرد. \* (مسئول مکاتبات)

۲- دانشجوی کارشناسی ارشد، مهندسی آب و سازه های هیدرولیکی، دانشگاه آیت اله العظمی بروجردی (ره)، بروجرد.

## **Drought Forecasting Using Wavelet - Support Vector Machine and Standardized Precipitation Index (Case Study: Urmia Lake-Iran)**

**Mehdi Komasi<sup>1</sup> \***

[komasi@abru.ac.ir](mailto:komasi@abru.ac.ir)

**Soroush Sharghi<sup>2</sup>**

Admission Date: December 27, 2017

Date Received: August 23, 2017

### **Abstract**

**Background and Objectives:** Drought is regarded as a serious threat for people and environment. As a result, finding some indices to forecast the drought is an important issue that needs to be addressed urgently. The appropriate and flexible index for drought classification is the Standardized Precipitation Index (SPI). Artificial intelligence models were commonly used to forecast SPI time series. These models are based on auto regressive property. So, they are not able to monitor the seasonal and long-term patterns in time series. In this study, the Wavelet-Support Vector Machine (WSVM) approach was used for the drought forecasting through employing SPI.

**Method:** In this way, the SPI time series of Urmia Lake watershed was decomposed to multiple frequent time series by wavelet transform; then, these time series were imposed as input data to the Support Vector Machine (SVM) model to forecast the drought.

**Findings:** The results showed that, the maximum value of  $R^2$  and minimum value of RMSE indexes for SVM model are 0.865 and 0.237 and for WSVM model are 0.954 and 0.056 respectively in verification step.

**Discussion and Conclusion:** So, the propounded hybrid model has superior ability in forecasting SPI time series comparing with the single SVM model and also it can accurately assess the extreme data in SPI time series by considering the seasonality effects. Finally, it was concluded that, the proposed hybrid model is relatively more appropriate than classical autoregressive models such as ANN.

**Keywords:** SVM Model, Wavelet Transform, SPI, Drought Forecasting, Urmia Lake Watershed.

---

1- Assistant Professor, Faculty of Civil Engineering, Ayatollah Boroujerdi University, Boroujerd, Iran.  
\*(Corresponding Author)

2- MS Student, Hydraulic Structures, Ayatollah Boroujerdi University, Boroujerd, Iran

## مقدمه

و پیش‌بینی جریان ماهانه ورودی به مخزن سد دز با استفاده از شبکه عصبی مصنوعی استاتیکی و دینامیکی پرداختند. نسل بعدی از روش‌های هوش مصنوعی، ماشین بردار پشتیبان (SVM)<sup>۶</sup> نام دارد که نسبت به روش شبکه عصبی مصنوعی سریع‌تر آموزش دیده و نتایج دقیق‌تری را ارائه می‌دهد. از طرفی این مدل کارایی مناسبی در کار با داده‌های غیرخطی نسبت به سایر مدل‌ها از خود نشان داده است (۹-۱۱). در سال‌های اخیر کاربردهای مدل ماشین بردار پشتیبان در مباحث هیدرولوژیکی و هواشناسی در مقایسه با مدل شبکه عصبی مصنوعی موردسنجش قرار گرفته است (۱۷-۱۲). در این زمینه صحرایی و مشفق (۱۸) به پیش‌بینی جریان رودخانه با استفاده از روش ماشین بردار پشتیبان پرداختند. نتایج حاصل از محاسبه و بررسی شاخص‌های آماری و مقادیر خطا، حاکی از دقت بالای این مدل در مقایسه با نتایج محققین قبلی بود. نیک‌بخت شهبازی و همکاران (۱۹) با استفاده از مدل ماشین بردار پشتیبان، شاخص SPI را در چهار حوضه تامین آب تهران پیش‌بینی کردند. آن‌ها به این نتایج دست یافتند که مدل ماشین بردار پشتیبان در مقایسه با مدل شبکه عصبی مصنوعی دقت بالایی در مدل‌سازی داده‌های زمانی بلندمدت داراست. هم‌چنین زهرایی و ناصری (۲۰) از مدل ماشین بردار پشتیبان جهت پیش‌بینی SPI فصلی استفاده نمودند و به این نتیجه دست یافتند که داده‌های SPI دو تا پنج ماه آینده را می‌توان با دقت بالایی به‌وسیله‌ی این مدل پیش‌بینی کرد، به صورتی که این پیش‌بینی نقش به‌سزایی را در مدیریت منابع آب ایفا می‌کند.

هرچند مدل ماشین بردار پشتیبان دارای انعطاف و دقت مناسبی جهت مدل‌سازی سری زمانی ایستا است، اما برای مدل‌سازی سری‌های زمانی نا ایستا که گستره‌ی وسیعی از مقیاس‌های زمانی هم‌چون دهه، سال و ماه را دربر می‌گیرند، دقت مناسبی ندارد. در چنین وضعیتی اگر داده‌های ورودی به مدل ماشین بردار پشتیبان توسط یک تابع تبدیل پیش‌پردازش

برخلاف سایر مخاطرات و بلایای طبیعی، خشک‌سالی پدیده‌ای است که به‌صورت آرام و کاملاً غیر محسوس رخ می‌دهد. بنابراین مدل‌سازی و پیش‌بینی مشخصه‌های مختلف خشک‌سالی مانند شروع، تداوم و شدت آن در یک منطقه نقش به‌سزایی را در مدیریت منابع آب و صنعت کشاورزی ایفا می‌کند. یکی از متداول‌ترین روش‌ها که جهت مشخص نمودن موقعیت و شدت خشک‌سالی در یک منطقه به‌کار می‌رود، استفاده از شاخص‌های خشک‌سالی است. از معروف‌ترین شاخص‌های خشک‌سالی می‌توان به شاخص شدت خشک‌سالی پالم (PDSI)<sup>۱</sup> و شاخص بارش استاندارد شده (SPI)<sup>۲</sup> اشاره کرد. لازم به ذکر است که محدودیت‌هایی در استفاده از شاخص PDSI وجود دارد. برای نمونه این شاخص برخلاف شاخص SPI فقط برای حوضه‌هایی با توپوگرافی یکنواخت کاربرد دارد. از سویی یکی از ویژگی‌های مثبت شاخص SPI توانایی توصیف خشک‌سالی در مقیاس‌های زمانی مختلف است (۱ و ۲ و ۳). یکی دیگر از برتری‌های شاخص SPI در مقایسه با شاخص PDSI، دارا بودن ساختاری ساده برای بیان موقعیت خشک‌سالی می‌باشد (۲). بنابراین در این پژوهش شاخص SPI مورداستفاده قرار می‌گیرد.

جهت پیش‌بینی سری زمانی داده‌های SPI نیازمند به یک ماشین یادگیری می‌باشد. امروزه مدل‌های هوش مصنوعی (AI)<sup>۳</sup> مانند مدل شبکه عصبی مصنوعی (ANN)<sup>۴</sup> و سیستم استنتاج فازی عصبی تطبیقی (ANFIS)<sup>۵</sup> در پژوهش‌های متعددی برای پیش‌بینی فرآیندهای هیدرولوژیکی به‌کار گرفته شده‌اند (۳-۶). به‌عنوان مثال ازدری مقدم و همکاران (۴) از مدل ANFIS جهت پیش‌بینی خشک‌سالی در زاهدان استفاده کردند. هم‌چنین این مدل توسط شفیع و همکاران (۷) برای پیش‌بینی جریان ورودی رود نیل به سد اسوان مورد استفاده قرار گرفت. بنی‌حبیب و همکاران (۸) نیز به شبیه‌سازی

1- Palmer Drought Severity Index

2- Standardized Precipitation Index

3- Artificial Intelligence

4- Artificial Neural Network

5- Adaptive Neural-Fuzzy Inference System

6- Support Vector Machine

شرایط رطوبت خاک می‌باشد. هم‌چنین این شاخص از توپوگرافی تاثیر نمی‌پذیرد. مشخصه اصلی SPI انعطاف‌پذیری آن در اندازه‌گیری خشک‌سالی در مقیاس‌های زمانی متفاوت است، زیرا خشک‌سالی از لحاظ دوره دارای دامنه بسیار گسترده‌ای می‌باشد. بنابراین تشخیص و پایش آن با انواع مقیاس‌های زمانی مهم می‌باشد. محاسبه SPI برای هر منطقه بر پایه آمار بلندمدت بارندگی استوار می‌باشد که این آمار به یک تابع توزیع احتمال، برازش یافته و تابع حاصل می‌تواند برای پیدا کردن احتمال تجمعی بارندگی برای یک ایستگاه و یک مقیاس زمانی مشخص مورد استفاده قرار گیرد. این احتمال سپس به متغیر تصادفی نرمال تبدیل می‌شود و بیان‌گر شاخص SPI است که دامنه تغییرات و طبقه‌بندی آن برای بیان شدت خشک‌سالی در جدول (۱) آورده شده است (۲۲).

بنابراین مقدار SPI انحراف استاندارد یک رویداد مشخص را از شرایط نرمال نشان می‌دهد. به‌طور کلی این شاخص بر اساس توضیحات گفته شده برای بیان خشک‌سالی چنین تعریف می‌شود: دوره‌ای که SPI دائماً منفی است و به مقدار ۱- و کم‌تر از آن می‌رسد و زمانی پایان می‌پذیرد که مقدار آن مثبت شود. بنابراین مقادیر مثبت نمایان‌گر بارشی بیش از مقدار متوسط و مقادیر منفی نشان‌گر بارشی کم‌تر از بارش متوسط می‌باشد (۲۳).

### محاسبه شاخص SPI

داده‌های اصلی شاخص SPI را داده‌های بارندگی ماهانه ایستگاه‌های باران‌سنجی تشکیل می‌دهند. پس از اطمینان از همگن و تصادفی بودن داده‌های ماهانه، سری زمانی در بازه‌های ۳، ۶، ۱۲، ۲۴ و ... ماهه تشکیل می‌شود و سری زمانی آن با توزیع گاما برازش داده می‌شود که تابع دانسیته احتمال یا فراوانی آن به‌صورت رابطه (۱) تعیین می‌گردد.

$$g(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} X^{(\alpha-1)} e^{-x/\beta} \quad X > 0 \quad (1)$$

نشوند دقت و کارایی مدل پایین خواهد آمد (۲۱). عمل پیش‌پردازش داده‌ها می‌تواند با استفاده از تابع تبدیل موجک انجام شود به این صورت که سری زمانی داده‌های ورودی به زیرسری‌هایی با درجات تجزیه مختلف، از یک‌دیگر مجزا شده و جداگانه وارد مدل ماشین بردار پشتیبان می‌شوند و مدل شناخت دقیق‌تری را از داده‌های ورودی پیدا می‌کند تا عمل پیش‌بینی در نهایت دقت صورت پذیرد. در نتیجه در این مقاله ایده و نوآوری مدل ترکیبی ماشین بردار پشتیبان موجکی (WSVM)<sup>۱</sup> که از داده‌های با مقیاس زمانی مختلف به‌عنوان ورودی استفاده می‌کند، جهت پیش‌بینی و مدل‌سازی پیشنهاد گردیده است. در این راستا در پژوهش حاضر از مدل ترکیبی ماشین بردار پشتیبان موجکی و شاخص SPI برای مدل‌سازی خشک‌سالی حوضه دریاچه ارومیه بهره‌گیری شده است. بنابراین هدف، اثبات برتری این روش نسبت به سایر روش‌های کلاسیک است.

در ادامه این مقاله، در بخش دوم مفاهیم مربوط به شاخص SPI، مدل ماشین بردار پشتیبان، مدل شبکه عصبی مصنوعی و تابع موجک مرور و سپس ساختار مدل ترکیبی ماشین بردار پشتیبان موجکی ارائه شده است. در بخش سوم، حوضه‌ی مورد مطالعه معرفی گردیده و در بخش چهارم کارایی‌های مدل منفرد ماشین بردار پشتیبان و مدل ترکیبی ماشین بردار پشتیبان موجکی مورد بحث و بررسی قرار گرفته است و در نهایت نتیجه‌گیری و جمع‌بندی در قسمت پایانی ارائه شده است.

### مواد و روش‌ها

#### شاخص بارش استاندارد شده (SPI)

شاخص SPI یک ابزار قوی در پردازش داده‌های بارندگی می‌باشد و هدف آن اختصاص ارزش عددی به بارندگی است که از این طریق بتوان نواحی با آب‌وهوای متفاوت را بایک‌دیگر مقایسه نمود. از مزیت‌های شاخص SPI استوار بودن محاسبه SPI بر پایه داده‌های بارندگی و عدم وابستگی این شاخص به

جدول ۱- طبقه‌بندی خشک‌سالی بر اساس شاخص SPI (۲۱)

TABLE 1. Drought classification based on SPI

| احتمال وقوع | دامنه SPI     | شدت خشک‌سالی               |
|-------------|---------------|----------------------------|
| ۰/۰۲۲۸      | بزرگ‌تر از ۲  | شدیداً مرطوب (مرطوب حدی)   |
| ۰/۰۴        | ۱/۹۹ ~ ۱/۵    | خیلی مرطوب                 |
| ۰/۰۹۱۹      | ۱/۴۹ ~ ۱      | مرطوب متوسط                |
| ۰/۳۴۱۳      | ۰/۹۹ ~ ۰      | مرطوب خفیف (تقریباً نرمال) |
| ۰/۳۴۱۳      | ۰ ~ -۰/۹۹     | خشک خفیف (تقریباً نرمال)   |
| ۰/۰۹۱۹      | -۱ ~ -۱/۴۹    | خشک متوسط                  |
| ۰/۰۴        | -۱/۹۹ ~ -۱/۵  | خیلی خشک                   |
| ۰/۰۲۲۸      | کوچک‌تر از -۲ | شدیداً خشک (خشک حدی)       |

محاسبه شده برای یافتن احتمال تجمعی بارندگی برای مقیاس زمانی مشخص برای هر یک از ایستگاه‌ها به کار گرفته می‌شود که این احتمال را با فرض  $t=X/\beta$  می‌توان بر طبق رابطه (۶) تبدیل به تابع گامای ناقص نمود.

$$G(X) \int_0^x g(x) dx = \frac{1}{\Gamma(\alpha)} \int_0^x t^{(\alpha-1)} e^{-t} dt \quad (۶)$$

از آنجایی که تابع گاما برای  $X=0$  تعریف نشده است، هرگاه توزیع بارندگی دارای مقدار صفر باشد احتمال تجمعی به صورت زیر محاسبه می‌شود.

$$H(X) = q + (1-q)G(X) \quad (۷)$$

در معادله فوق  $q$  احتمال بارندگی صفر و  $m$  تعداد صفرهای موجود در سری‌های زمانی بارندگی می‌باشد که  $q$  را به صورت حاصل تقسیم  $m$  بر تعداد کل داده‌ها ( $n$ ) تخمین می‌زند، سپس با داشتن  $H(x)$  و روابط (۸) تا (۱۱)  $SPI$  محاسبه می‌گردد:

$$SPI = - \left[ t - \frac{c_0 + c_1 t + c_2 t}{1 + d_1 t + d_2 t^2 + d_3 t^3} \right] \quad 0 < H(X) \leq 0.5 \quad (۸)$$

$$SPI = + \left[ t - \frac{c_0 + c_1 t + c_2 t}{1 + d_1 t + d_2 t^2 + d_3 t^3} \right] \quad 0.5 < H(X) < 1 \quad (۹)$$

$$t = \sqrt{\ln \left( \frac{1}{H(X)^2} \right)} \quad 0 < H(X) \leq 0.5 \quad (۱۰)$$

$\alpha$  و  $\beta$  به ترتیب پارامتر شکل و پارامتر مقیاس می‌باشند و  $\Gamma(\alpha)$  تابع گاما بوده که به صورت زیر تعریف می‌شود:

$$\Gamma(\alpha) = \int_0^{\infty} y^{(\alpha-1)} e^{-y} dy \quad (۲)$$

پارامترهای  $\alpha$  و  $\beta$  مربوط به تابع چگالی گاما برای هر ایستگاه و برای هر مقیاس زمانی و هر ماه از سال برآورد می‌شود. مک‌کی و همکاران با استفاده از روش حداکثر درست‌نمایی مقادیر بهینه ضرایب  $\alpha$  و  $\beta$  را بر اساس معادلات زیر برآورد نمودند.

(۲۴)

$$\alpha = \frac{1}{4A} \left[ 1 + \sqrt{1 + \frac{4A}{3}} \right] \quad (۳)$$

$$A = \ln(\bar{X}) - \frac{\sum \ln(X)}{n} \quad (۴)$$

$$\beta = \frac{\bar{X}}{\alpha} \quad (۵)$$

در روابط بالا  $n$  تعداد مشاهدات بارندگی و  $\bar{X}$  میانگین بارندگی برای یک دوره چندماهه مشخص می‌باشد. سپس پارامترهای

$$t = \sqrt{\ln\left(\frac{1}{(1-H(X))^2}\right)} \quad 0.5 < H(X) < 1 \quad (11)$$

$c_0, c_1, c_2, d_1, d_2, d_3$  ضرایب ثابت به این شرح

می باشند:

$$c_2=0.010328, c_1=0.802853, c_0=2.515517, d_3=0.001308, d_2=0.189269, d_1=1.432788$$

### ماشین بردار پشتیبان (SVM)

فرض مقادیر ورودی و خروجی  $x_i$  و  $y_i$ ، هدف یافتن تابعی است که کمترین انحراف  $\varepsilon$  را از  $y_i$  ها داشته باشد. اگر داده‌های آموزشی به تعداد  $K$  سری به صورت  $(x_1, y_1), \dots, (x_k, y_k)$  موردنظر باشند، تابع خطی سازی عبارت است از:

$$f(x) = (w, x) + b \quad \text{with} \quad (w, x) \in R^N, b \in R \quad (12)$$

گاهی ممکن است نتوان میزان انحراف  $\varepsilon$  را لحاظ کرد و باید مقدار انحراف بیش‌تر از  $\varepsilon$  را نیز قابل قبول فرض کرد. این میزان انحراف اضافه با لحاظ کردن متغیرهای  $(\xi_i, \xi'_i)$  در رابطه‌ی بهینه‌سازی (۱۳) اعمال می‌شود که منجر به رابطه‌ی بهینه‌سازی (۱۴) می‌گردد (۲۵).

ماشین‌های بردار پشتیبان همانند دیگر روش‌های هوش مصنوعی بر مبنای الگوریتم داده‌کاوی عمل می‌کنند. مهم‌ترین وظایف مدل ماشین‌های بردار پشتیبان، طبقه‌بندی و خطی‌سازی یا رگرسیون داده‌ها است. خطی‌سازی داده‌ها به وسیله‌ی الگوریتم خطی‌سازی  $\varepsilon$ -SVR و با معرفی تابع  $\varepsilon$ -sensitive انجام می‌پذیرد. در یک الگوریتم خطی‌سازی  $\varepsilon$ -SVR، با

در تعیین تابع  $f(x)$  باید همواره مقادیر خروجی به‌گونه‌ای تعیین شوند که به‌طور هم‌زمان، کمترین میزان انحراف  $\varepsilon$  تعیین‌شده و تابع مقدار مناسب  $w$  را اختیار کند. با حل رابطه‌ی بهینه‌سازی (۱۳) دست‌یابی به چنین شرایطی محقق می‌شود. (۱۳)

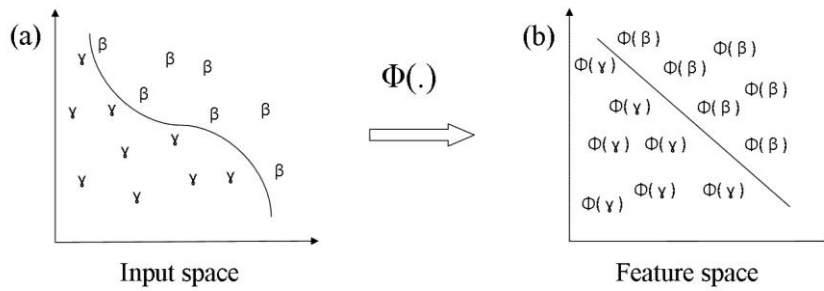
$$\text{Minimize } \frac{1}{2} w^2 \quad \text{subject to} \begin{cases} y_i - (w, x_i) - b \leq \varepsilon \\ (w, x_i) + b - y_i \leq \varepsilon \end{cases}$$

$$\text{Minimize } \frac{1}{2} w^2 + C \sum_{i=1} (\xi_i + \xi'_i) \quad \text{subject to} \begin{cases} y_i - (w, x_i) - b \leq \varepsilon + \xi_i \\ (w, x_i) + b - y_i \leq \varepsilon + \xi'_i \\ \xi_i, \xi'_i \geq 0 - \text{for all } i = 1, 2, \dots, K \end{cases} \quad (14)$$

در رابطه با تابع هسته لازم به ذکر است که قبل از خطی‌سازی داده‌ها یا همان تقسیم خطی، جهت توانایی ماشین بردار پشتیبان برای پیش‌بینی داده‌های با پیچیدگی بالا، داده‌ها به وسیله‌ی تابع هسته  $(\Phi$  یا  $K)$  به فضای با ابعاد بسیار بالاتر منتقل می‌شوند تا تقسیم بندی خطی داده‌ها میسر شود (شکل ۱). بنابراین کافی است در مسایل غیرخطی از هسته یا همان کرنل مقادیر ورودی به‌جای خود تابع استفاده شود.

در رابطه‌ی (۱۴)،  $C$  خطای تولید نام دارد که در واقع رابطه‌ی جایگزین را برای تابع  $f(x)$  مشخص کرده و مقدار آن تا مقدار مجاز انحراف  $\varepsilon$  افزایش می‌یابد. در نهایت با استفاده از توابع کرنل  $k(x_i, x)$  و پارامترهای لاگرانژین  $\lambda_i$  و  $\lambda'_i$  رابطه‌ی تابع خطی‌سازی، به شکل رابطه‌ی (۱۵) بازنویسی خواهد شد.

$$f(x, b) = \sum_{i=1}^k (\lambda'_i - \lambda_i) k(x_i, x) + b \quad (15)$$



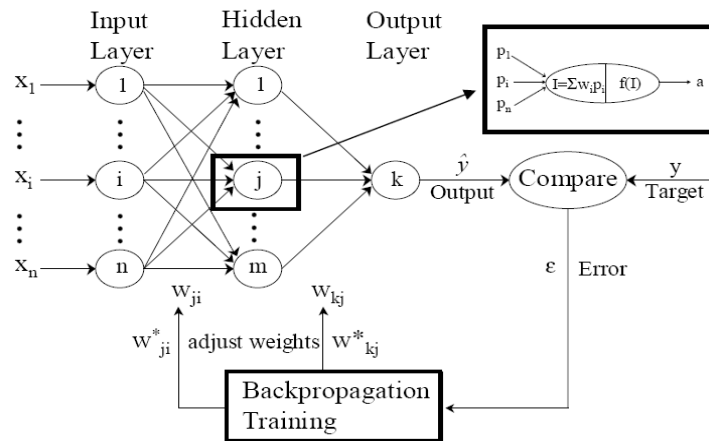
شکل ۱- مفهوم تبدیل از فضای ورودی به فضای مشخصه (۱۶)

Figure 1. The converting of the input space to feature space

معماری شبکه عصبی - مصنوعی

t, 2, 3, t-1, ... می‌تواند وابسته شود ۲, ۴, ۶ و ... نرون در نظر گرفته شده است. هم‌چنین تعداد نرون‌های لایه میانی متغیر بوده و برای حالت‌های مختلف بین ۳ تا ۳۰ نرون بررسی صورت گرفته است. لایه خروجی نیز شامل یک نرون می‌باشد که مقدار SPI در زمان t+1 می‌باشد.

یکی از انواع مهم شبکه‌های عصبی، شبکه پرسپترون چندلایه (Multi-Layer Perception) است که شامل یک لایه ورودی، یک یا چندلایه مخفی و یک لایه خروجی می‌باشد. در این مقاله از یک شبکه سه لایه مطابق شکل (۲) با تابع محرک سیگموئید استفاده شده که تعداد نرون‌های لایه اول در شبکه با رویداشت به این نکته که SPI در زمان t+1 به SPI در زمان‌های t, t-1



شکل ۲- طرح و آرایش شبکه عصبی مصنوعی

Figure 2. Design and arrangement of artificial neural network

$$\hat{y}_k = f_0 \left[ \sum_{j=1}^M w_{kj} \cdot f_h \left( \sum_{i=1}^N w_{ji} x_i + w_{j0} \right) + w_{ko} \right] \quad (16)$$

در این تحقیق از الگوریتم LM به منظور شناسایی بهترین روش با بالاترین بازده برای آموزش شبکه استفاده شده است. این الگوریتم تغییر یافته‌ی الگوریتم کلاسیک نیوتن است که

در آرایش معماری شبکه عصبی شکل (۲)،  $X_i$  متغیرهای ورودی،  $y$  متغیر خروجی،  $w$  ضرایب وزن در شبکه،  $n, m$  و  $k$  به ترتیب تعداد نرون در لایه ورودی، لایه میانی و لایه خروجی می‌باشد که این آرایش را می‌توان به فرم رابطه (۱۶) بیان داشت.

**تبدیل موجک**

تابع تبدیل موجک (شکل ۳) قابلیت تجزیه سری زمانی به چندین زیر سری زمانی با مقیاس‌های مختلف را دارا است و با مطالعه زیر سری‌های زمانی حاصل از سری زمانی کلی، رفتار کوچک مقیاس و بزرگ مقیاس یک فرایند هیدرولوژیکی را مورد آنالیز قرار می‌دهد.

برای پیدا کردن راه حلی مناسب برای مسائلی که نیاز به کمینه‌سازی دارند با فرمول بندی زیر به کار می‌رود.

$$X_{k+1} = X_k - [J^T J + \mu I]^{-1} J^T + e \quad (17)$$

در این رابطه،  $X$  وزن‌های شبکه عصبی،  $J$  ژاکوبین ماتریس معیار اجرای شبکه که باید حداقل شود و  $\mu$  عددیست که فرایند آموزش را کنترل می‌کند و  $e$  بردار خطای باقی‌مانده است (۲۶).



شکل ۳- طرح شماتیک تبدیل موجک

Figure 2. Schematic of the wavelet transform

استفاده در تحلیل، با دو عمل ریاضی انتقال و مقیاس در طول سیگنال مورد تحلیل، تغییر اندازه و تغییر محل قرار می‌گیرد.

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) \quad (19)$$

در نهایت ضرایب موجک در هر نقطه از سیگنال (b) و برای هر مقدار از مقیاس (a) با رابطه (۲۰) قابل محاسبه است (۲۷):

$$CWT(a,b) = Wf(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x) \psi\left(\frac{x-b}{a}\right) dx = \int_{-\infty}^{+\infty} f(x) \psi_{a,b}(x) dx \quad (20)$$

باشند که در شکل (۴) نمودار تابع چهار نمونه از موجک‌ها نشان داده شده است (۲۷).

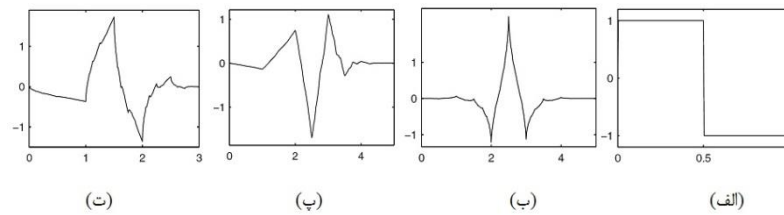
تابع موجک، تابعی است که دارای دو ویژگی مهم نوسانی و کوتاه‌مدت بودن می‌باشد.  $\Psi(X)$ ، تابع موجک است اگر و فقط اگر تبدیل فوریه آن  $\Psi(\omega)$ ، شرط زیر را برآورده کند.

$$\int_{-\infty}^{+\infty} \frac{|\Psi(\omega)|}{|\omega|^2} d\omega < +\infty \quad (18)$$

این شرط با عنوان شرط پذیرفتگی برای موجک  $\Psi(X)$  شناخته می‌شود.  $\Psi(X)$  تابع موجک مادر است که در آن توابع مورد

توابع موجک دارای انواع بسیاری هستند که مهم‌ترین و پرکاربردترین آن‌ها شامل تابع موجک Haar, Daubechies, Mexican hat, Morle, Symlet, Coiflet, Meyer می‌-



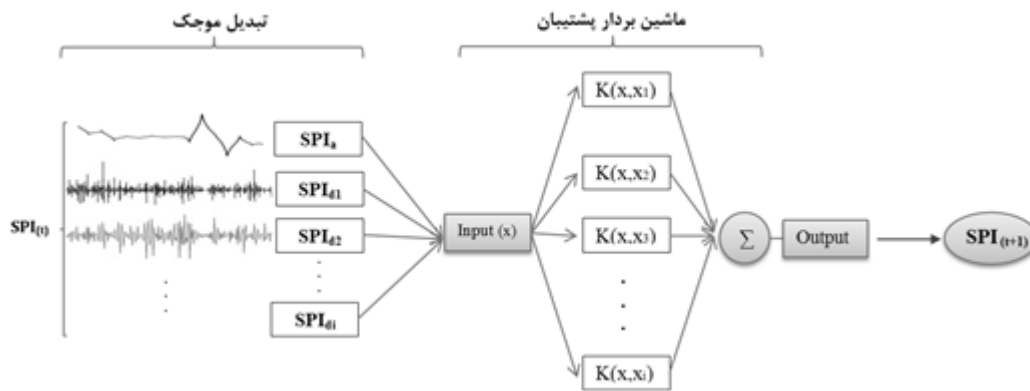


شکل ۴- نمودار الف) تابع موجک Haar ب) تابع موجک Coif1 پ) تابع موجک Sym3 و ت) تابع موجک Db2  
 Figure. 4. (a) Haar wavelet function, (b) Coif1 wavelet function, (c) Sym3 wavelet function, (d) Db2 wavelet function

مدل ترکیبی ماشین بردار پشتیبان موجکی (WSVM)

که تبدیل موجک نقش پیش‌پردازش داده‌های خام و مدل ماشین بردار پشتیبان نقش مدل‌سازی را ایفا می‌کند (شکل ۵).

معماری مدل ترکیبی ماشین بردار پشتیبان موجکی از دو بخش تبدیل موجک و ماشین بردار پشتیبان تشکیل شده است



شکل ۵- ساختار مدل ترکیبی ماشین بردار پشتیبان موجکی

Figure 5. The architecture of the hybrid WSVM model

تقسیم می‌شوند: ۷۵٪ کل داده‌ها جهت واسنجی و ۲۵٪ باقی‌مانده جهت صحت‌سنجی مدل ترکیبی مورد استفاده قرار می‌گیرند.

$$R^2 = 1 - \frac{\sum_{i=1}^N (SPI_{obs_i} - SPI_{com_i})^2}{\sum_{i=1}^N (SPI_{obs_i} - \overline{SPI_{obs}})^2} \quad (21)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (SPI_{obs_i} - SPI_{com_i})^2}{N}} \quad (22)$$

به‌طوری‌که  $R^2$ ،  $RMSE$ ،  $N$ ،  $SPI_{obs_i}$ ،  $SPI_{com_i}$  و  $\overline{SPI_{obs}}$  به ترتیب ضریب تبیین، جذر میانگین مربع خطا، تعداد داده‌های مشاهداتی  $SPI$ ، داده‌های مشاهداتی  $SPI$ ، پیش‌بینی‌شده (محاسباتی) و میانگین داده‌های مشاهداتی است.

در روش پیشنهاد شده ابتدا سری داده‌های  $SPI$  به زیر سری‌هایی با درجه مختلف تجزیه شده و در مرحله بعد داده‌های تجزیه شده جهت مدل‌سازی به‌عنوان ورودی مدل ماشین بردار پشتیبان به کار گرفته می‌شوند. در این پژوهش برای انجام مدل‌سازی از برنامه‌نویسی Matlab استفاده شده است. در این مدل‌سازی جهت سنجش میزان صحت و دقت از روابط آماری ضریب تبیین  $R^2$  و جذر میانگین مربع خطا  $(RMSE)$  در مرحله‌ی واسنجی<sup>۱</sup> و صحت‌سنجی<sup>۲</sup> به شرح روابط (۲۱) و (۲۲) استفاده می‌شود. برای این منظور داده‌ها به دو گروه

- 1- Determination Coefficient
- 2- Root Mean Square Error
- 3- Calibration
- 4- Verification

انتخاب بهترین تابع موجک و بهترین درجه تجزیه مورد صحت سنجی قرار می‌گیرد.

#### مطالعه موردی و داده‌های مطالعاتی

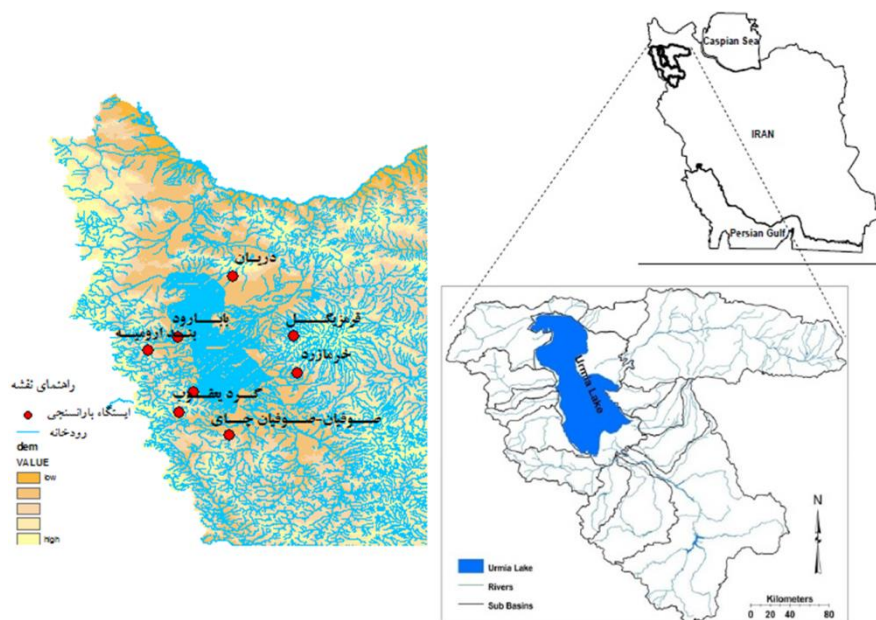
دریاچه ارومیه یک دریاچه نمکی محسوب می‌شود و در شمال غرب ایران بین دو استان آذربایجان شرقی و غربی واقع شده است. این دریاچه بزرگ‌ترین دریاچه در خاورمیانه به حساب می‌آید. طول این دریاچه ۱۴۰ کیلومتر، عرض آن ۴۰ تا ۵۵ کیلومتر و بیش‌ترین عمق آن ۱۶ متر می‌باشد (شکل ۶). سطح آب در این دریاچه به حد بحرانی خود رسیده است و ارتفاع آب زیرزمینی در قسمتی از حوضه آن تا ۱۶ متر کاهش یافته است. دمای میانگین در اطراف دریاچه و در مناطق کوهستانی آن به ترتیب ۱۱ و ۲/۵ درجه سانتی‌گراد می‌باشد. بیش‌تر جریان ورودی به دریاچه مربوط به سه رودخانه زربینه، سیمینه و تلخینه می‌باشد. در این پژوهش از میانگین داده‌های بارش سالانه در هشت ایست‌گاه باران‌سنجی (شکل ۶) طی سال‌های ۱۳۴۸ تا ۱۳۸۸ استفاده شده و با استفاده از این داده‌ها مقادیر سری زمانی SPI در بازه‌ی زمانی ۱۲ ماهه محاسبه شده‌اند (شکل ۷).

هم‌چنین از رابطه (۲۳) جهت ارزیابی قابلیت مدل در پیش‌بینی نقاط بیشینه یا کمینه استفاده می‌شود.

$$R^2_{\text{peak}} = 1 - \frac{\sum_{i=1}^N (SPI_{PC_i} - SPI_{PO_i})^2}{\sum_{i=1}^N (SPI_{PO_i} - \overline{SPI}_{PO})^2} \quad (23)$$

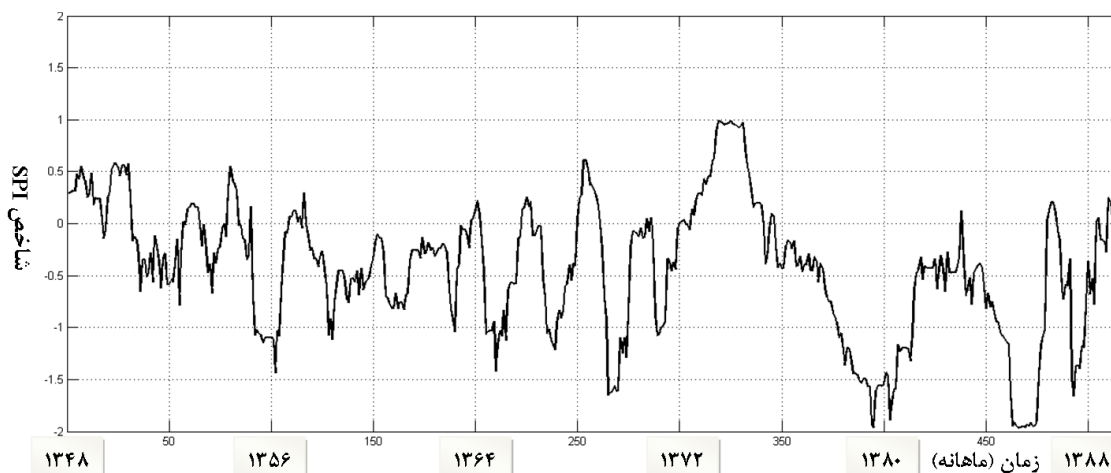
به‌طوری‌که  $R^2_{\text{peak}}$  ضریب تبیین برای مقادیر بیشینه و کمینه و نیز N تعداد داده‌های بیشینه و کمینه و  $SPI_{pci}$ ،  $SPI_{poi}$  و  $\overline{SPI}_{po}$  به ترتیب داده‌های مشاهداتی، داده‌های محاسباتی و میانگین داده‌های مشاهداتی برای مقادیر بیشینه و کمینه می‌باشد.

مدل ماشین بردار پشتیبان موجکی از پارامترهای ساختاری مختلفی هم‌چون تابع کرنل، تابع موجک و درجه تجزیه تشکیل شده است که باید به‌درستی انتخاب و کالیبره شوند. بنابراین در گام نخست نتایج پژوهش، مدل ماشین بردار پشتیبان منفرد در نظر گرفته شده و کارآمدترین تابع کرنل انتخاب می‌گردد تا در مدل ترکیبی به کار گرفته شود. سپس این مدل با آنالیز موجک ترکیب شده و نتایج حاصل از آن مانند



شکل ۶- موقعیت جغرافیایی دریاچه ارومیه و ایست‌گاه‌های باران‌سنجی

Figure 6. Location of Lake Urmia and drainage sub-basins



شکل ۷- سری زمانی داده‌های SPI حوضه دریاچه ارومیه

Figure 7. SPI time series of Urmia Lake watershed

## مدل منفرد ماشین بردار پشتیبان (SVM)

این منظور همان گونه که در جدول (۲) نشان داده شده، مدل‌سازی ماشین بردار پشتیبان توسط سه نوع تابع هسته بر روی داده‌های SPI در زمان  $t$  ( $SPI_{(t)}$ ) برای حوضه‌ی ارومیه انجام گرفته و از  $R^2$  و RMSE به عنوان معیاری برای تعیین کارایی مدل در واسنجی و صحت سنجی استفاده گردیده است. به دلیل بی‌بعد بودن داده‌های SPI مقادیر RMSE نیز بی‌بعد می‌باشند.

تابع هسته یا کرنل، ساختاری مهم در الگوریتم ماشین بردار پشتیبان محسوب می‌شود و دارای انواع مختلفی می‌باشد. در این پژوهش چهار نوع تابع هسته مورد ارزیابی قرار گرفته است که شامل RBF-kernel، Poly-kernel، MLP-kernel و Lin-kernel می‌باشند. در گام نخست هدف انتخاب بهترین تابع هسته از نظر کارایی در مدل‌سازی در مقایسه با سایر توابع هسته می‌باشد. برای

## جدول ۲- نتایج مدل‌سازی ماشین بردار پشتیبان جهت انتخاب بهترین تابع هسته

Table 2. Result of SVM model for finding the best kernel function

| تابع کرنل   | واسنجی |       | صحت سنجی |       |
|-------------|--------|-------|----------|-------|
|             | $R^2$  | RMSE  | $R^2$    | RMSE  |
| RBF-kernel  | ۰/۸۸۳  | ۰/۱۵۴ | ۰/۸۶۵    | ۰/۲۳۷ |
| Poly-kernel | ۰/۷۶۵  | ۰/۳۰۲ | ۰/۷۶۱    | ۰/۳۴۱ |
| MLP-kernel  | ۰/۷۹۱  | ۰/۲۱۰ | ۰/۷۷۱    | ۰/۳۰۲ |
| Lin-kernel  | ۰/۷۸۴  | ۰/۲۹۵ | ۰/۶۹۸    | ۰/۳۸۶ |

کارایی بالاتری را نسبت به سایر توابع از خود نشان داده است. بنابراین تابع RBF-kernel برای مدل‌سازی در سایر گام‌ها به-

در جدول (۲) با توجه به مقادیر  $R^2$  و RMSE در مرحله‌ی صحت سنجی یا شبیه‌سازی، تابع RBF-kernel دقت و

داده‌های ورودی (Comb(6) و Comb(5)) موجب بیش-برازش<sup>۲</sup> در مرحله واسنجی شده و این امر باعث کاهش دقت مدل سازی در مرحله صحت سنجی می‌شود. در این راستا شکل (۸) نشان‌دهنده‌ی مقایسه داده‌های مشاهداتی و محاسباتی توسط مدل ماشین بردار پشتیبان با به‌کارگیری تابع هسته RBF-kernel و ترکیب ورودی‌های Comb(2) می‌باشد.

#### مدل ترکیبی ماشین بردار پشتیبان موجکی (WSVM)

در این مرحله به منظور سنجش اثر آنالیز موجک بر فرآیند مدل سازی، داده‌های پیش‌پردازش شده به‌عنوان ورودی‌های مدل ماشین بردار پشتیبان مورد استفاده قرار می‌گیرند. برای این منظور از تبدیل موجک به‌صورت مجزا استفاده می‌گردد. تبدیل موجک داده‌ها را در مقیاس‌های زمانی مختلف (درجه تجزیه) مورد ارزیابی قرار می‌دهد. بنابراین سری‌های زمانی کوتاه و بلند مقیاس طی این فرآیند از هم مجزا می‌گردند. در این پژوهش سعی شده تا اثر استفاده از توابع موجک مختلف با درجه تجزیه مختلف بر روی کارایی مدل مشخص گردد. بنابراین سری‌های زمانی با درجه تجزیه ۱ تا ۷ به‌وسیله‌ی چهار نوع تابع موجک تجزیه می‌شوند. تابع موجک Sym3، تابع موجک Db2، تابع موجک Haar و تابع موجک Coife1 توابع موجکی هستند که در این پژوهش از آن‌ها استفاده شده است. در مدل ترکیبی روش کار بدین شرح است که سری‌های تجزیه‌شده SPI توسط تابع موجک وارد مدل ماشین بردار پشتیبان شده و پس از واسنجی، SPI را در یک روز جلوتر (SPI<sub>(t+1)</sub>) به‌عنوان خروجی پیش‌بینی می‌کند و سپس صحت سنجی داده‌های محاسباتی به‌وسیله‌ی داده‌های مشاهداتی صورت می‌گیرد. نتایج حاصل از مدل سازی بر اساس بهترین تابع هسته (RBF-kernel) برای انتخاب بهترین تابع موجک و درجه‌ی تجزیه در جدول (۴) آورده شده است.

کار گرفته می‌شود. در مدل ماشین بردار پشتیبان داده‌های SPI با الگوی زمانی مختلف به‌عنوان ورودی وارد الگوریتم ماشین بردار پشتیبان شده تا SPI در زمان t+1 (SPI<sub>(t+1)</sub>) پیش‌بینی شود. به بیانی دیگر الگوریتم ماشین بردار پشتیبان در مدل سازی داده‌های SPI عمل کردی خود هم‌بسته<sup>۱</sup> دارد، به طوری که برای پیش‌بینی داده‌های SPI آینده (SPI<sub>(t+1)</sub>) از داده‌های زمان t و t-1 و... (SPI<sub>(t)</sub>, SPI<sub>(t-1)</sub>, ...) استفاده می‌گردد (رابطه (۲۴)).

$$SPI_{(t+1)} = f_{svm}(SPI_{(t)}, SPI_{(t-1)}, SPI_{(t-2)}, \dots) \quad (24)$$

در این پژوهش از الگوریتم مارکوف برای تعیین وابستگی زمانی داده‌های SPI استفاده شده است. بدین ترتیب که برای پیش-بینی SPI در زمان t+1 (SPI<sub>(t+1)</sub>) از ترکیب ورودی‌های SPI به شرح ترکیب‌های زیر استفاده گردیده است.

Comb. (1): SPI<sub>(t)</sub>

Comb. (2): SPI<sub>(t)</sub>, SPI<sub>(t-1)</sub>

Comb. (3): SPI<sub>(t)</sub>, SPI<sub>(t-1)</sub>, SPI<sub>(t-2)</sub>

Comb. (4): SPI<sub>(t)</sub>, SPI<sub>(t-1)</sub>, SPI<sub>(t-2)</sub>, SPI<sub>(t-3)</sub>

Comb. (5): SPI<sub>(t)</sub>, SPI<sub>(t-1)</sub>, SPI<sub>(t-2)</sub>, SPI<sub>(t-3)</sub>, SPI<sub>(t-4)</sub>

Comb. (6): SPI<sub>(t)</sub>, SPI<sub>(t-1)</sub>, SPI<sub>(t-2)</sub>, SPI<sub>(t-3)</sub>, SPI<sub>(t-4)</sub>,

SPI<sub>(t-5)</sub>

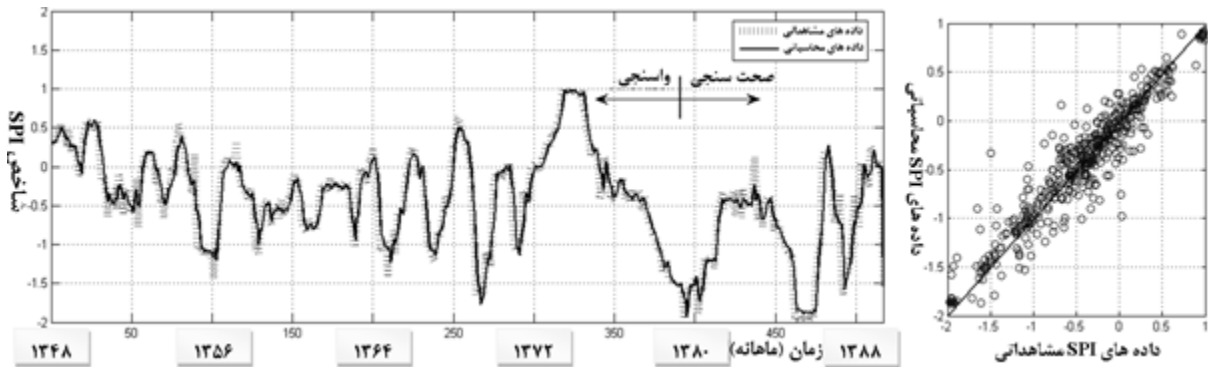
مانند گام قبلی هر کدام از ترکیب ورودی‌های SPI توسط مدل ماشین بردار پشتیبان مدل سازی شده و دقت آن‌ها با توجه به مقادیر R<sup>2</sup> و RMSE ارزیابی می‌شوند. جدول (۳) نتایج این مدل سازی را نشان می‌دهد.

طبق نتایج به‌دست آمده از جدول (۳) ترکیب ورودی‌های Comb(2) نتایج مطلوب‌تری را در مقایسه با سایر ترکیب ورودی‌ها نشان داده است. این نتیجه در تطابق با نتایج سایر پژوهش‌گران می‌باشد هنگامی که از قانون خودهم‌بستگی استفاده می‌کنند (۲۶ و ۲۸). هم‌چنین درجات بالاتر ترکیب

جدول ۳- نتایج مدل سازی تحت ترکیب ورودی های مختلف به وسیله ی تابع هسته ی RBF-kernel

Table 3. Result of SVM model with different combs of input data by RBF-kernel function

| الگوی ترکیبی | واسنجی         |       | صحت سنجی       |       |
|--------------|----------------|-------|----------------|-------|
|              | R <sup>2</sup> | RMSE  | R <sup>2</sup> | RMSE  |
| Comb. (1)    | ۰/۷۸۱          | ۰/۳۲۳ | ۰/۷۶۳          | ۰/۳۴۱ |
| Comb. (2)    | ۰/۸۸۳          | ۰/۱۹۰ | ۰/۸۶۵          | ۰/۲۳۷ |
| Comb. (3)    | ۰/۸۲۰          | ۰/۲۰۱ | ۰/۸۵۶          | ۰/۲۴۱ |
| Comb. (4)    | ۰/۷۸۶          | ۰/۳۰۵ | ۰/۷۶۱          | ۰/۳۴۵ |
| Comb. (5)    | ۰/۸۸۷          | ۰/۱۵۶ | ۰/۶۷۰          | ۰/۴۴۷ |
| Comb. (6)    | ۰/۸۸۸          | ۰/۱۴۴ | ۰/۶۴۱          | ۰/۴۵۵ |



شکل ۸- سری زمانی داده های SPI مشاهده ای و محاسباتی توسط ترکیب داده های ورودی (Comb(2) و تابع هسته

RBF-kernel

Figure.8. Simulated and actual SPI time series for comb (2) as input data and RBF-kernel function

جدول ۴- نتایج مدل سازی مدل ماشین بردار پشتیبان موجکی توسط توابع موجک و درجات تجزیه مختلف

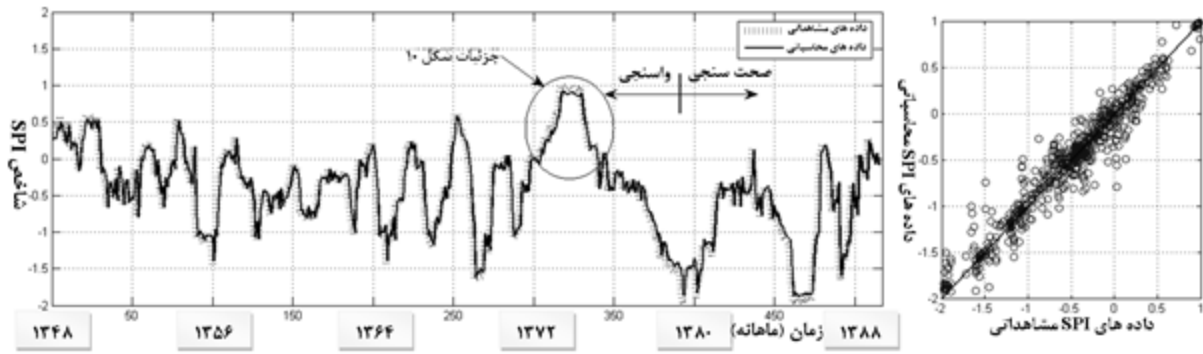
Table 4. Result of the WSVM model by RBF-kernel function with different mother wavelets and decomposition levels

| تابع موجک | درجه تجزیه | واسنجی         |       | صحت سنجی       |       |
|-----------|------------|----------------|-------|----------------|-------|
|           |            | R <sup>2</sup> | RMSE  | R <sup>2</sup> | RMSE  |
| Sym3      | ۱          | ۰/۹۲۶          | ۰/۰۵۸ | ۰/۸۹۰          | ۰/۰۷۲ |
| Sym3      | ۲          | ۰/۹۲۶          | ۰/۰۵۸ | ۰/۸۹۰          | ۰/۰۷۲ |
| Sym3      | ۳          | ۰/۹۲۶          | ۰/۰۵۸ | ۰/۸۹۲          | ۰/۰۷۱ |
| Sym3      | ۴          | ۰/۹۲۶          | ۰/۰۵۸ | ۰/۸۹۲          | ۰/۰۷۱ |
| Sym3      | ۵          | ۰/۹۲۴          | ۰/۰۵۹ | ۰/۸۹۰          | ۰/۰۷۲ |
| Sym3      | ۶          | ۰/۹۳۲          | ۰/۰۵۷ | ۰/۳۷۷          | ۰/۱۷۲ |

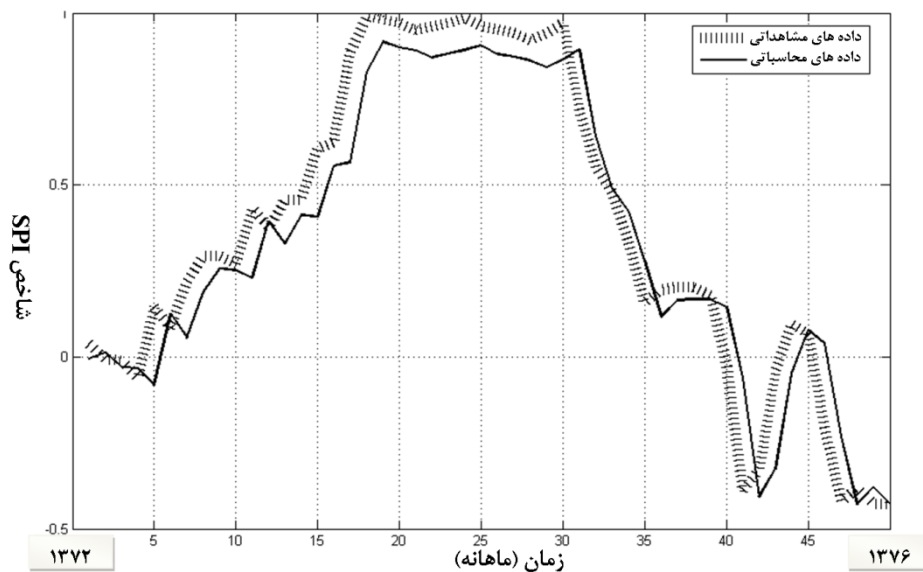
|       |   |       |       |       |       |
|-------|---|-------|-------|-------|-------|
| Sym3  | ۷ | ۰/۹۳۳ | ۰/۰۵۷ | ۰/۳۷۵ | ۰/۱۷۴ |
| Db2   | ۱ | ۰/۹۳۶ | ۰/۰۵۳ | ۰/۸۹۰ | ۰/۰۷۲ |
| Db2   | ۲ | ۰/۹۴۰ | ۰/۰۵۲ | ۰/۹۰۹ | ۰/۰۶۵ |
| Db2   | ۳ | ۰/۹۴۱ | ۰/۰۵۲ | ۰/۹۰۷ | ۰/۰۶۶ |
| Db2   | ۴ | ۰/۹۴۲ | ۰/۰۵۱ | ۰/۹۰۹ | ۰/۰۶۵ |
| Db2   | ۵ | ۰/۹۴۲ | ۰/۰۵۱ | ۰/۹۱۰ | ۰/۰۶۵ |
| Db2   | ۶ | ۰/۹۵۳ | ۰/۰۵۱ | ۰/۰۲۰ | ۰/۲۱۶ |
| Db2   | ۷ | ۰/۹۵۵ | ۰/۰۵۱ | ۰/۰۱۹ | ۰/۲۱۸ |
| Haar  | ۱ | ۰/۹۲۶ | ۰/۰۵۸ | ۰/۸۹۰ | ۰/۰۷۲ |
| Haar  | ۲ | ۰/۹۲۶ | ۰/۰۵۸ | ۰/۸۹۰ | ۰/۰۷۲ |
| Haar  | ۳ | ۰/۹۲۶ | ۰/۰۵۸ | ۰/۸۹۲ | ۰/۰۷۱ |
| Haar  | ۴ | ۰/۹۲۶ | ۰/۰۵۸ | ۰/۸۹۲ | ۰/۰۷۱ |
| Haar  | ۵ | ۰/۹۲۴ | ۰/۰۵۹ | ۰/۸۹۰ | ۰/۰۷۲ |
| Haar  | ۶ | ۰/۹۳۶ | ۰/۰۵۳ | ۰/۳۷۷ | ۰/۱۷۲ |
| Haar  | ۷ | ۰/۹۳۸ | ۰/۰۵۲ | ۰/۳۷۵ | ۰/۱۷۳ |
| Coif1 | ۱ | ۰/۹۲۸ | ۰/۰۵۷ | ۰/۹۲۹ | ۰/۰۵۸ |
| Coif1 | ۲ | ۰/۹۳۱ | ۰/۰۵۶ | ۰/۹۳۰ | ۰/۰۵۷ |
| Coif1 | ۳ | ۰/۹۳۲ | ۰/۰۵۵ | ۰/۹۳۰ | ۰/۰۵۷ |
| Coif1 | ۴ | ۰/۹۳۲ | ۰/۰۵۵ | ۰/۹۳۱ | ۰/۰۵۷ |
| Coif1 | ۵ | ۰/۹۴۳ | ۰/۰۵۵ | ۰/۹۵۴ | ۰/۰۵۶ |
| Coif1 | ۶ | ۰/۹۸۳ | ۰/۰۴۳ | ۰/۷۸۰ | ۰/۱۰۲ |
| Coif1 | ۷ | ۰/۹۸۴ | ۰/۰۴۲ | ۰/۷۷۱ | ۰/۱۰۱ |

بر طبق این جدول درجات بالای تفکیک داده‌ها (درجه ۶ و ۷) موجب بیش‌برازش داده‌ها در مرحله‌ی واسنجی شده و دقت مدل‌سازی را در مرحله شبیه‌سازی (صحت‌سنجی) کاهش می‌دهد. سری زمانی مربوط به داده‌های مشاهداتی و محاسباتی در مدل‌سازی به روش ترکیبی ماشین بردار پشتیبان موجکی برای بهترین تابع موجک (Coife1) و درجه‌ی تجزیه ۵ در اشکال (۹ و ۱۰) نشان داده شده است.

با توجه به جدول (۴) تابع موجک Coife1 و درجه تجزیه ۵ بهترین نتایج مدل‌سازی را از نظر دقت شبیه‌سازی نشان می‌دهند. دلیل برتری تابع موجک Coife1 را می‌توان در شباهت سیگنال این تابع موجک با شکل و روند تغییرات سری زمانی SPI به خصوص در نقاط بیشینه دانست (شکل ۴-ب). از سویی، دلیل کارایی مناسب درجه تجزیه ۵ نسبت به سایر درجات تجزیه موجک این است که سطح تجزیه ۵ شامل الگوی ۳<sup>۵</sup> روزه است که به الگوی SPI ماهانه بسیار نزدیک می‌باشد.



شکل ۹- مقایسه داده های محاسباتی و مشاهداتی مربوط به تابع موجک Coif1 و درجه تجزیه ۵  
 Figure 9. Simulated and actual SPI series by Coif1 wavelet function and five decomposition level



شکل ۱۰- مقایسه داده های محاسباتی و مشاهداتی در ناحیه مشخص شکل ۹  
 Figure 10. Simulated and actual SPI series specified in Fig. 9

بردار پشتیبان موجکی نسبت به مدل های خود هم پسته مانند مدل ماشین بردار پشتیبان توانایی بالاتری در درک و شبیه سازی نقاط اکسترمی دارند. هم چنین یک نکته ی مهم در مدل سازی ماشین بردار پشتیبان موجکی عدم نیاز به کنترل ترکیبات مختلف به شیوه ی مارکوف می باشد، زیرا مدل ترکیبی آنالیز موجک یک تحلیل چند مقایسه ای است که الگوها و روندهای حاکم بر پدیده را به خوبی رصد می نماید. از مقایسه جداول (۲ و ۳) با جدول (۴) این نتیجه حاصل می شود که عملیات پیش پردازش داده ها به وسیله تابع موجک باعث افزایش دقت مدل سازی و پیش بینی در الگوریتم ماشین بردار پشتیبان شده است به طوری که بیشترین دقت مدل سازی در صحت-سنجی بر روی داده های خام (پیش پردازش نشده) در جداول (۲)

نقاط بیشینه و کمینه در سری زمانی داده های SPI نشان دهنده ی پتانسیل ترسالی یا خشک سالی در حوضه می باشد و نقش موثری را در مدیریت خشک سالی ایفا می کند. در این ارتباط نکته ی مهمی که از شکل های (۸ و ۹) استنباط می شود، توانایی بالای مدل ترکیبی ماشین بردار پشتیبان موجکی در محاسبه ی نقاط بیشینه و کمینه ی سری زمانی در قیاس با مدل ماشین بردار پشتیبان است، به طوری که با توجه به رابطه (۲۳) مقادیر  $R^2_{paek}$  برای مدل ماشین بردار پشتیبان و ماشین بردار پشتیبان موجکی به ترتیب برابر با ۰/۶۲ و ۰/۹۳ می باشد. در نتیجه مدل های فصلی<sup>۱</sup> مانند مدل ترکیبی ماشین

1- Seasonal models

ترتیب منجر به محاسبه ی ضریب تبیین ( $R^2$ ) معادل ۰/۶۳، ۰/۷۴، ۰/۹۲ و ۰/۸۷ در مرحله ی صحت سنجی گردید. از مقایسه ی نتایج این تحقیق با مقاله ی حاضر این نتیجه حاصل می شود که دقت داده کاوی در روش ماشین بردار پشتیبان منفرد از روش های دیگری مانند ARIMA و ANN بیش تر است و هم چنین مدل ترکیبی ماشین بردار پشتیبان موجکی نتیجه ی مطلوب تری را نسبت به مدل های ترکیبی دیگری مانند WANN و MMLAN ارائه می دهد. در پژوهش دیگری که توسط نورانی و همکاران (۲۶) صورت گرفت دقت مدل سازی برنامه نویسی ژنتیک (GP) در شبیه سازی فرآیند بارش و رواناب روزانه حوضه آقچای با ضریب تبیینی به اندازه ی ۰/۷۱۳ برآورده شد. همانطور که پیداست دقت مدل سازی روش مطرح در این مقاله بر روش GP نیز مقدم است.

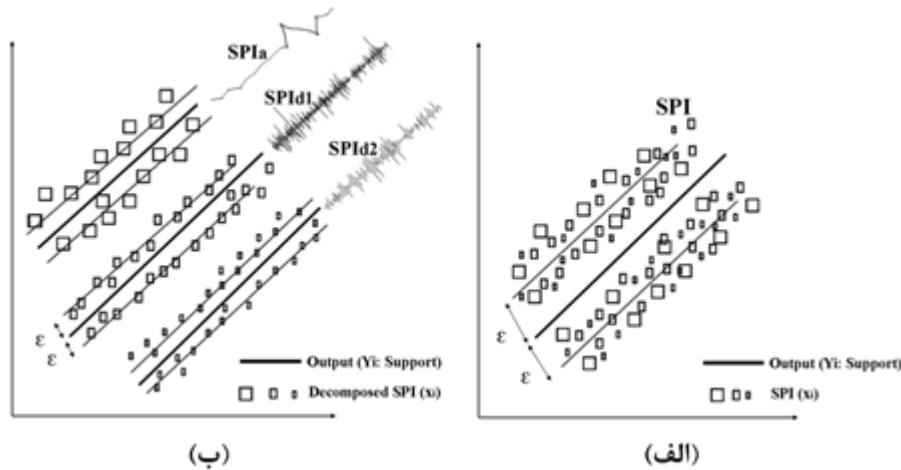
و ۳) به صورت میانگین ۰/۸۶ بوده در حالی که پس از انجام عملیات تبدیل موجک بر روی داده های ورودی و تفکیک و پردازش آن ها بیشینه دقت مدل سازی در صحت سنجی به ۰/۹۳۴ رسیده است. این امر بیان گر اثرات مثبت تبدیل موجک بر روی مدل سازی داده های SPI می باشد. شکل (۱۱) به صورت شماتیک اثر مثبت تبدیل موجک بر روی داده ها را قبل و بعد از پیش پردازش و تفکیک سری زمانی داده ها به وسیله تبدیل موجک نشان می دهد. همان طور که مشخص است پس از انجام عملیات پیش پردازش داده ها، خطای مدل سازی (E) به طور قابل توجهی کم تر می شود.

در گام آخر در این پژوهش، مدل ترکیبی ماشین بردار پشتیبان موجکی با سایر مدل های ترکیبی مانند مدل شبکه عصبی مصنوعی موجکی (WANN)<sup>۱</sup> مقایسه گردیده است. نتایج به شرح جدول (۵) می باشد.

بر طبق مقادیر  $R^2$  و RMSE مدل ترکیبی ماشین بردار پشتیبان موجکی نتایج دقیق ترین را نسبت به مدل ترکیبی ماشین بردار پشتیبان موجکی از خود نشان داده است. از سویی، ترکیب آنالیز موجک در هر کدام از مدل ها موجب بهبود نتایج نسبت به مدل منفرد آن شده است. به عنوان نمونه از میان کارهای مشابه انجام گرفته در خصوص مدل سازی به روش های داده کاوی می توان به نتایج حاصل از پژوهش نورانی و همکاران (۲۹) اشاره کرد. در این پژوهش مدل سازی بارش و رواناب برای مدل های ARIMA<sup>۲</sup>، ANN، WANN و MMLAN<sup>۳</sup> به

1- Wavelet Artificial Neural Network  
2- Auto Regressive Integrated Moving Average  
3- Modified Multi-Layer Artificial Network





شکل ۱۱- مدل سازی توسط ماشین بردار پشتیبان قبل (الف) و بعد (ب) از پیش پردازش داده ها تحت تبدیل موجک

Figure 11. Support vector machine modeling before (A) and after (B) of the data processing by wavelet transform

جدول ۵- مقایسه نتایج مدل سازی SVM، WSVM، ANN و WANN

Table 5. Comparison of SVM, WSVM, ANN and WANN modeling

| مدل  | نوع مدل | واسنجی $R^2$ | صحت سنجی $R^2$ |
|------|---------|--------------|----------------|
| ANN  | منفرد   | ۰/۵۳۵        | ۰/۵۵۸          |
| WAAN | ترکیبی  | ۰/۸۹۶        | ۰/۸۹۸          |
| SVM  | منفرد   | ۰/۸۸۳        | ۰/۸۶۵          |
| WSVM | ترکیبی  | ۰/۹۴۳        | ۰/۹۵۴          |

توجه: در این جدول بهترین نتایج از هر مدل نمایش داده شده است.

نتیجه گیری و جمع بندی

در این پژوهش مدل ترکیبی ماشین بردار پشتیبان موجکی (WSVM) به منظور مدل سازی سری زمانی داده های شاخص SPI ارایه گردید. برای این منظور سری زمانی داده های SPI مربوط به حوضه ی دریاچه ارومیه به وسیله ی تبدیل موجک به زیر سری هایی تجزیه شد و سپس این زیر سری های تجزیه شده به عنوان داده های ورودی در مدل ماشین بردار پشتیبان جهت پیش بینی SPI در یک گام زمانی جلوتر مورد استفاده قرار گرفت. نتایج نشان داد که تابع هسته RBF-Kernel نسبت به سایر توابع از دقت و کارایی بالاتری برخوردار می باشد. نکته ی مهم این پژوهش بهره گیری از توابع تبدیل موجک جهت رفع مشکل استفاده از ترکیب متعدد و ویژگی خودهم بستگی مدل ماشین بردار پشتیبان است. بدین منظور از چهار نوع تابع

موجک (Coife1 و Haar، Db2، Sym3) با مدل مورد نظر و درجات آزادی ۱ تا ۷ جهت مدل سازی استفاده گردید. نتایج مدل سازی نشان داد که تابع موجک Coife1 و درجه تجزیه ۵ به عنوان مناسب ترین گزینه ها از لحاظ کارایی در مدل ترکیبی می باشند. هم چنین نشان داده شد که یکی از توانایی های ویژه مدل ترکیبی ماشین بردار پشتیبان موجکی نسبت به مدل منفرد، محاسبه ی نقاط بیشینه و کمینه ی سری زمانی SPI است. در نهایت دقت بالای این مدل در مقایسه با مدل ترکیبی شبکه عصبی مصنوعی موجکی (WANN) به اثبات رسید. به دلیل قابلیت ویژه مدل ترکیبی پیشنهاد می گردد که از این مدل جهت مدل سازی فرآیندهای هیدرولوژیکی مانند رسوب گذاری نیز استفاده شود.

- Autoregressive Static and Artificial Dynamic Neural Network for the Forecasting of Monthly Inflow of Dez Reservoir. *Journal of Environmental Science and Technology*, Vol. 13(4), pp. 1-14.
9. Lin G., Chen G., Huang P., Chou Y., 2009. Support vector machine-based models for hourly reservoir inflow forecasting during typhoon-warning periods. *Journal of hydrology*. Vol. 372, pp.17-29.
  10. Han, D., Chan, L., Zhu, N., 2007. Flood forecasting using support vector machines. *Journal of Hydroinformatics*, Vol. 9(4), pp. 267-276.
  11. Tripathi, S.h., Srinivas, V.V., Nanjundiah R.S., 2006. Downscaling of precipitation for climate change scenarios: A support vector machine approach, *Journal of Hydrology*, Vol. 330, pp. 621- 640.
  12. Wang, W., Men, C., Lu, W., 2008. Online prediction model based on support vector machine. *Neuro Computing*, Vol. 71, pp. 550-558.
  13. Behzad, M.K., Asghari, M., Eazi, M., Palhang M., 2009. Generalization performance of support vector machines and neural networks runoff modeling. *Expert Systems with Applications*, Vol. 36, pp. 7624-7629.
  14. Asefa, T., Kemblowski, M.W., Urroz, G., McKee, M. Khalil, A., 2004. Support vectors-based groundwater head observation networks design. *Water Resources Research*, Vol. 40 (11).
  15. Asefa, T., Kemblowski, M.W., McKee, M., Khalil A., 2006. Multi-time scale stream flow prediction: The support vector machines approach.

## Reference

1. Cacciamani, C., Morgillo, A., Marchesi, S., Pavan, V.,2007. Monitoring and forecasting drought on a regional scale: Emilia-Romagna Region. *Water Science and Technology*. Lib, Vol. 62, pp. 29-48.
2. Tsakiris, G., Vangelis, H.,2004. Towards a drought watch system based on spatial SPI. *Water Resources Management*, Vol. 18. pp. 11-12.
3. Mishra, A.K., Desai, V.R.,2006. Drought forecasting using feed-forward recursive neural network. *Ecological Modelling*, Vol. 198(1-2), pp. 127-138.
4. Azhdari Moghadam, M., Khosravi, M., Hosseinpour Niknam, H., Jafari Nodoshan, E.,2012. Drought forecasting using neuro-fuzzy model. Climate indices and time series of precipitation and drought. *Applied Geography and. Development*, Vol. 26, pp. 17-20.
5. Marj, A.F., Meijerink, A.M.J.,2011. Agricultural drought forecasting using satellite images, climate indices and artificial neural network. *International Journal of Remote Sensing*, Vol. 32(24), pp. 9707-9719.
6. Ebrahimpour, R., 2014. Using Artificial Neural Networks to Estimate the Return Sludge Rate, A Case Study of Torbat Heydarieh Wastewater Treatment Plant. *Journal of Water and Wastewater*, Vol. 25(4), pp. 99-107. (In Persian)
7. Shafie, A., Taha, M.R., Noureldin, A., 2007. A neuro-fuzzy model for inflow forecasting of the Nile river at Aswan high dam. *Journal of Water Resources Management*, Vol. 21, pp. 533-556.
8. Banihabib, M., Valipoor, M., Behbahani, S. 2011. Comparison of

- Hydrologic Engineering*, Vol. 6, pp. 319–328.
23. Vermes, L., 2001. How to work out a Drought mitigation strategy. European Regional conference of ICID.
  24. McKee, T.B., Doesken, N.J., Kleist, J., 1993. The relationship of drought frequency and duration to time scales. in Proceedings of the 8th Conference on Applied Climatology, Anaheim, Calif, USA.
  25. Vapnik, V.N., Cortes, C., 1995. Support Vector Networks. *Machine Learning*, Vol. 20, pp. 273–297.
  26. Nourani, V., Komasi, M., Taghi Alami, M., 2012. Hybrid wavelet–genetic programming approach to optimize artificial neural network modeling of rainfall–runoff process. *Journal of Hydrology Engineering*, Vol. 17(6), pp. 724–741.
  27. Mallat, S.G., 1998. A Wavelet Tour of Signal Processing. Second ed. Academic Press. San Diego.
  28. Nourani V., Kisi Ö., Komasi M., 2011. Two hybrid artificial intelligence approaches for modeling rainfall–runoff process. *Journal of Hydrology*, Vol. 402, pp. 41–59.
  29. Nourani V., Komasi M., Mano A., 2009. A multivariate ANN-wavelet approach for rainfall–runoff modeling. *Water Resources Management*, Vol. 23, pp. 2877–2894.
  - Journal of Hydrology*, Vol. 318, pp. 7–16.
  16. Dibike, Y.B., Velickov, S., Solomatine, D., Abbott, M.B., 2001. Model induction with support vector machines: introduction and applications. *Journal of Computing in Civil Engineering*, Vol. 15 (3), pp. 208–216.
  17. Liong, S.Y., Sivapragasam, C., 2002. Flood stage forecasting with support vector machines. *The Journal of the American Water Resources Association*, Vol. 38 (1), pp. 173–196.
  18. Sahraie, S., Zaker Moshfegh, M., 2013. River Flow Prediction Using Case Study Support Vector Machine, 7th National Congress of Civil Engineering, Zahedan, Sistan and Baluchestan University. (In Persian)
  19. Nikbakht Shahbazi, A., Zahraei, B., Sadghi, H., Manshouri, M., Nasserli, M., 2011. Seasonality meteorological drought prediction using support vector machine. *World Applied Sciences Journal*, Vol. 13(6), pp. 1387-1397.
  20. Zahraei, B., Nasserli, M., 2014. Basin scale meteorological drought forecasting using support vector machine. International Conference on Drought Management Strategies in Arid and Semi-Arid Regions.
  21. Cannas, B., Fanni, A., See, L., Sias, G., 2006. Data preprocessing for river flow forecasting using neural networks: wavelet transforms and data partitioning. *Physics and Chemistry of the Earth*, Vol. 31(18), pp. 1164–1171.
  22. Kim, T., Valdes, J.B., 2003. Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. *Journal of*