

## پیش‌بینی نسبت جذب سدیم آب با استفاده از مدل تلفیقی شبکه عصبی و تبدیل موجک (مطالعه موردی: ایستگاه رودبار رودخانه سفیدرود)

ظاهر رجایی<sup>1\*</sup>، حمیده جعفری<sup>2</sup>

تاریخ دریافت: 92/12/25 تاریخ پذیرش: 95/02/20

<sup>1</sup> دانشجویار گروه مهندسی عمران، دانشکده فنی و مهندسی، دانشگاه قم

<sup>2</sup> دانشجوی دکتری، گروه مهندسی عمران، دانشکده فنی و مهندسی، دانشگاه قم

\* مسئول مکاتبات، پست الکترونیکی: [Taher\\_rajae@yahoo.com](mailto:Taher_rajae@yahoo.com)

### چکیده

از مهم‌ترین عوامل مؤثر بر توسعه پایدار حوزه‌های آبریز، فراهم بودن کمی و کیفی منابع آب است. در این مطالعه مدل‌های شبکه عصبی، رگرسیون خطی چند متغیره، برنامه‌ریزی ژنتیک و ترکیب موجک-عصبی برای مدل‌سازی نسبت جذب سدیم (SAR) ماهانه در ایستگاه رودبار رودخانه سفیدرود بررسی شده و تأثیر پیش‌پردازش داده‌ها روی عملکرد مدل شبکه عصبی به‌وسیله تبدیل موجک ناپیوسته موردتحقیق قرار گرفت. بدین منظور سری‌های زمانی دبی و نسبت جذب سدیم به‌وسیله تبدیل موجک ناپیوسته مختلف در سطوح متفاوتی به زیرسری‌ها تجزیه شدند و سپس به‌عنوان ورودی شبکه عصبی مورد استفاده قرار گرفتند. نتایج نشان داد که مدل ترکیبی موجک-عصبی به‌خصوص برای پیش‌بینی مقادیر بیشینه SAR در مقایسه با روش‌های رگرسیون خطی، شبکه عصبی و برنامه‌ریزی ژنتیک مناسب‌تر هست. همچنین مدل ترکیبی توانست پدیده پس‌ماند را برای نسبت SAR به‌خوبی شبیه‌سازی نماید، درحالی‌که مدل رگرسیون در تشخیص آن ناتوان بود.

واژه‌های کلیدی: برنامه‌ریزی ژنتیک، تبدیل موجک، رودخانه سفیدرود، شبکه عصبی، نسبت جذب سدیم

## Prediction of Water Sodium Absorption Ratio (SAR) using ANN and Wavelet Conjunction Model (Case Study: Rudbar Station of Sefidrud River)

T Rajae<sup>1\*</sup>, H Jafari<sup>2</sup>

Received: 16 March 2014 Accepted: 09 May 2016

<sup>1</sup> Associate Prof., Civil Eng. Dept, Univ. of Qom.

<sup>2</sup> Ph.D. Student, Civil Eng. Dept, Univ. of Qom.

\* Corresponding Author, Email: [taher\\_rajae@yahoo.com](mailto:taher_rajae@yahoo.com)

### Abstract

One of the most important factors of sustainable development of watersheds is qualitative and quantitative availability of suitable water resources. In this study, the artificial neural network (ANN), multi - variable linear regression (MLR), genetic programming (GP) and hybrid wavelet -ANN (WANN) models were considered for modeling the monthly sodium absorption ratio (SAR) in Sefidrud River - Roudbar Station, and the effect of data preprocessing on model performance was investigated using the discrete wavelet Transform method. For this purpose, observed time series of river discharge and SAR were decomposed into several sub - time series at different scales by discrete wavelet transform. Then these sub - time series were introduced as inputs to the ANN model. The results showed that the hybrid wavelet - neural network model was more suitable for predicting maximum SAR values than the MLR, ANN, and GP models. Furthermore, the hybrid model could simulate the hysteresis phenomenon for SAR modeling rigorously, while multi linear regression method was incapable of detecting it.

**Keywords:** Artificial neural network, Genetic programming, Sefidrud River, Sodium absorption ratio, Wavelet transform

### مقدمه

در چند سال اخیر جایگزین روش‌های سنتی تخمین و پیش‌بینی پارامترها شده، استفاده از شبکه‌های عصبی مصنوعی هست. امروزه این مدل‌ها در طیف وسیعی از تحقیقات استفاده می‌شود. به‌عنوان نمونه پالانی و همکاران (2008) از شبکه‌های عصبی برای پیش‌بینی مشخصات کیفی آب‌های ساحلی سنگاپور استفاده کردند. رجایی و میرباقری (1388) نیز برای پیش‌بینی بار معلق رودخانه‌ها از شبکه عصبی استفاده کرده‌اند. آن‌ها در تحقیق خود با استفاده از سری زمانی دبی جریان و غلظت بار معلق در زمان‌های قبل، توانستند غلظت بار معلق در گام زمانی آینده را پیش‌بینی نمایند.

آگاهی از کیفیت منابع آب یکی از نیازمندی‌های مهم در برنامه‌ریزی و توسعه منابع آب و حفاظت آن هست. حفظ کیفیت آب در اولویت سیاست‌گذاری‌ها قرار دارد، لذا مطالعه در جهت پیش‌بینی روند تغییرات کیفیت آب و استفاده از روش‌های کارا (با دقت مناسب)، برای مدیریت بهتر رودخانه‌ها ضروری به نظر می‌رسد. نسبت جذب سدیم در آب (SAR)<sup>1</sup> یکی از پارامترهای موردتوجه در ارزیابی کیفی آب به شمار می‌رود. از آنجاکه سری زمانی این پارامتر ماهیت دینامیکی دارد، پیش‌بینی آن به‌وسیله روش‌های خطی با خطا همراه است. یکی از روش‌هایی که

<sup>1</sup>Sodium absorption ratio

دو روش دیگر بالاتر است. نوشادی و همکاران (1386) برخی پارامترهای کیفی آب شامل هدایت الکتریکی، کل جامدات محلول، اسیدیته، بی‌کربنات و کلراید را در رودخانه زاینده‌رود با شبکه عصبی شبیه‌سازی کردند. این مدل توانست نتایج قابل‌اعتمادی را برای پیش‌بینی این پارامترها ارائه دهد. نیکو و همکاران (1390) با استفاده از شبکه‌های عصبی به تعیین کیفیت آب در طول مسیر رودخانه کارون پرداختند. نتایج این تحقیق نشان داد که مدل شبکه عصبی نسبت به مدل‌های آماری رگرسیون دقت بیشتری دارد. اسداللهی فرد و همکاران (2013) از شبکه‌های عصبی برای پیش‌بینی SAR رودخانه چهل‌گزی کردستان استفاده کردند. آنان برای پیش‌بینی SAR از یک تابع برازشی با ورودی‌های سولفات، سدیم، کلسیم، کلرید، منیزیم و بی‌کربنات استفاده کردند. نتایج تحقیق نشان داد که مدل شبکه عصبی دقت قابل قبولی برای پیش‌بینی SAR ارائه می‌دهد. بانژاد و همکاران (1392) از مدل‌های ترکیبی موجک-عصبی برای تخمین برخی از پارامترهای کیفی آب (TDS, EC, SAR) رودخانه‌های جاجرود و قره‌سو استفاده کردند. نتایج تحقیق نشان داد که مدل توسعه‌یافته موجک-عصبی با دقت بالا و خطای کمتری نسبت به مدل شبکه عصبی پارامترهای TDS و EC را پیش‌بینی کرده است.

در مقاله حاضر از ترکیب شبکه عصبی خودبازگشتی غیرخطی<sup>6</sup> و موجک برای پیش‌بینی سری زمانی پارامتر کیفی SAR استفاده شده است. بدین منظور، پس از تحلیل نتایج همبستگی بین پارامترهای SAR با دبی و تأخیرهای 1 تا 4 روزه دبی و نیز بین SAR با تأخیرهای 1 تا 4 روزه SAR، ترکیب‌های مختلف ورودی برای پیش-بینی سدیم جذبی ماه آینده ایجاد می‌شود. سپس نتایج با استفاده از چهار نوع تابع موجک با سطوح تجزیه 1 تا 5 موردبررسی قرار می‌گیرد.

ماهیت غیر ایستای<sup>2</sup> پدیده‌های هیدرولوژیکی، دقت مدل شبکه عصبی را پایین می‌آورد، لذا پیش‌پردازش ورودی مدل‌های شبکه عصبی می‌تواند به بهبود نتایج بیانجامد. از پیش‌پردازش‌هایی که در سری زمانی مورد استفاده قرار گرفته، می‌توان به تبدیل موجک<sup>3</sup> اشاره کرد. استفاده از تبدیل موجک برای تحلیل و پیش‌بینی سری‌های زمانی با استفاده از شبکه عصبی روشی مناسب بوده که هم‌اکنون نیز در حال گسترش است. رجایی (2011) مدل جدید موجک-عصبی را برای رسوب معلق روزانه در رودخانه‌ها ارائه داد. به‌منظور ارزیابی دقت مدل، مدل پیشنهادی با نتایج مدل‌های شبکه عصبی، رگرسیون چندخطی<sup>4</sup> (MLR) و منحنی سنج رسوب<sup>5</sup> (SRC) مورد مقایسه قرار گرفت. مقایسه دقت پیش‌بینی مدل‌ها نشانگر دقت بالای مدل ترکیبی بود. نورانی و همکاران (2009) ترکیب دو قابلیت شبکه‌های عصبی و تحلیل موجک را با کارایی بهتر در پیش‌بینی فرآیندهای هیدرولوژیکی به نام شبکه‌های موجک-عصبی ارائه دادند. رجایی و همکاران (2011) با مدل ترکیبی موجک-عصبی معلق روزانه رودخانه را پیش‌بینی کردند. با مقایسه نتایج توانایی بالای ترکیب موجک و شبکه عصبی نسبت به دیگر روش‌ها برای محاسبه بار معلق رسوب رودخانه مشخص گردید. سینگ (2012) مدل ترکیبی موجک-عصبی را برای پیش‌بینی سیلاب بکار برد. نتایج نشان‌دهنده توانایی بالای مدل ترکیبی موجک-عصبی نسبت به روش‌های آماری بود. لانگ‌کین و همکاران (2013) مدلی را برای پیش‌بینی کیفی آب بر اساس موجک و شبکه عصبی ارائه دادند. این پژوهش نشان داد که استفاده از موجک دقت مدل را نسبت به مدل شبکه عصبی بالا می‌برد. جین لیو و همکاران (2013) در تحقیقی بار رسوب رودخانه‌ای در چین را به‌وسیله روش ترکیبی موجک-عصبی مدل‌سازی کردند. نتایج نشان داد که دقت روش ترکیبی موجک-عصبی از

<sup>6</sup> Sediment rating curve<sup>7</sup> Nonlinear auto regressive<sup>2</sup> Non stationary<sup>3</sup> Wavelet transform<sup>4</sup> Multi linear regression

## مواد و روش‌ها

## منطقه مورد مطالعه

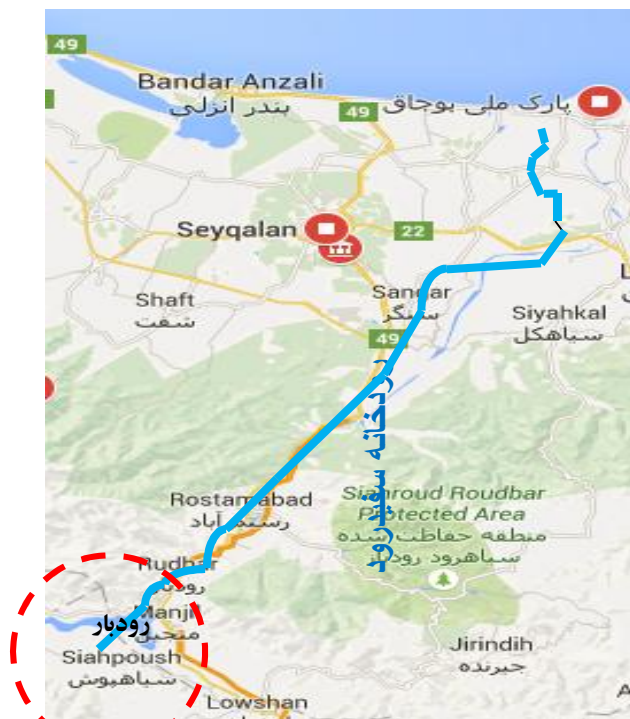
رودخانه سفیدرود با وسعتی حدود 67000 کیلومترمربع یکی از بزرگ‌ترین و پرآب‌ترین رودخانه‌های کشور بوده و از مهم‌ترین منابع تجدیدشونده آب شیرین جهت استفاده در کشاورزی، شرب و صنعت به شمار می‌رود. منطقه مورد مطالعه در جلگه ساحلی جنوب دریای خزر و در دامنه شمالی رشته‌کوه‌های البرز و تالش از محل منجیل تا دریای خزر بین 49 درجه و 15 دقیقه تا 50 درجه و 15 دقیقه طول شرقی و 36 درجه و 45 دقیقه تا 37 درجه و 30 دقیقه عرض شمالی واقع شده است. این حوضه بخش‌هایی از استان‌های گیلان، اردبیل، آذربایجان شرقی، زنجان، قزوین، کردستان، تهران، همدان و آذربایجان غربی را شامل می‌شود. این رودخانه شامل یک شاخه اصلی و حدود 27 شاخه فرعی مهم است (باقرزاده 1377).

داده‌های استفاده‌شده در این تحقیق شامل مقادیر SAR و میزان دبی در یک دوره 120 ماهه است که این آمار به صورت ماهانه از فروردین 79 تا شهریور 89 در محل ایستگاه آب‌سنجی رودبار توسط شرکت آب منطقه‌ای استان گیلان به ثبت رسیده است. پس از بررسی داده‌ها، از دوره 120 ماهه داده‌های موجود، معادل 75% کل داده‌ها برای آموزش و داده‌های 30 ماه باقی‌مانده معادل 25% داده‌ها به منظور اعتبار سنجی و آزمون مدل‌ها مورد استفاده قرار گرفت. شکل 2 سری زمانی داده‌های نسبت سدیم جذبی و دبی را در دوره آزمون نشان می‌دهد.

## تحلیل آماری داده‌ها

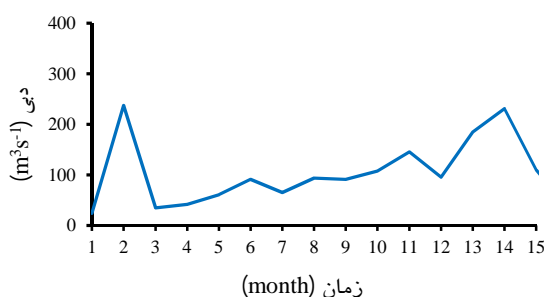
در جدول 1 تحلیل آماری داده‌های روزانه دبی و نسبت جذب سدیم برای کل دوره مورد مطالعه، داده‌های دوره آموزش و داده‌های دوره آزمون به تفکیک ارائه شده است. این تحلیل با استفاده از نرم‌افزار SPSS انجام یافته است. تحلیل ارائه‌شده شامل میانگین، کمینه، بیشینه، انحراف معیار، ضریب چولگی، ضریب خودهمبستگی<sup>1</sup> دبی با تأخیر یک‌ماهه سری زمانی دبی ( $r_1$ )، ضریب خودهمبستگی دبی با تأخیر دو ماهه سری زمانی دبی ( $r_2$ ).

ضریب خودهمبستگی دبی با تأخیر سه‌ماهه سری زمانی دبی ( $r_3$ )، ضریب خودهمبستگی دبی با تأخیر چهارماهه سری زمانی دبی ( $r_4$ ) و به همین ترتیب ضرایب خودهمبستگی SAR با تأخیرهای یک تا 4 ماهه سری زمانی SAR ( $r_1$  تا  $r_4$ ) هست.



شکل 1- موقعیت ایستگاه رودبار در رودخانه

سفیدرود.



شکل 2- سری زمانی دبی و نسبت جذب سدیم ماهانه دوره

آزمون ایستگاه مورد مطالعه.

برای مدل‌سازی مناسب هست، زیرا ضریب چولگی تأثیر نسبتاً قابل‌ملاحظه‌ای روی عملکرد شبکه عصبی دارد. بررسی ضرایب همبستگی نشان می‌دهد که خودهمبستگی پارامتر SAR به ازای تأخیرهای 1 تا 4 روزه SAR و پارامتر دبی به ازای تأخیرهای 1 و 2 روزه دبی، نسبتاً بالا هست (در جدول 1 داده‌های اعتبارسنجی جزء داده‌های آموزش لحاظ شده است).

به‌منظور انتخاب ترکیب‌های مناسب ورودی مدل-های رگرسیون خطی، شبکه عصبی و شبکه عصبی موجک، همبستگی بین سری‌های زمانی نسبت جذب سدیم ( $SAR_t$ ) و دبی جریان زمان حال و تأخیرهای دبی در جدول 2 محاسبه شده است.

هر چه ایستایی داده‌ها بالاتر باشد و تناوب و روند بالاتری در سری زمانی وجود داشته باشد پیش‌بینی با دقت بالاتری انجام می‌شود. باید توجه شود که مدل‌های شبکه عصبی زمانی بهتر عمل می‌کنند که خصوصیات آماری داده‌های آموزش و آزمون مشابهت داشته باشند. بررسی جدول 1 نشان می‌دهد که شباهت خصوصیات آماری داده‌های دوره آموزش و آزمون در حد قابل قبولی هست. میانگین مجموعه آزمون و آزمایش برای هر یک از سری‌های زمانی دبی و نسبت جذب سدیم به یکدیگر نزدیک هست. مقادیر بیشینه و کمینه برای هر دو پارامتر در مجموعه آزمایش قرار گرفته که این مسئله بر عملکرد بهتر شبکه عصبی تأثیر مثبت دارد. ضریب چولگی داده-های SAR در مقایسه با داده‌های دبی نسبتاً پایین بوده و

جدول 1- تحلیل آماری سری‌های زمانی داده‌های نسبت جذب سدیم و دبی.

مجموعه آزمایش		مجموعه آزمون		همه داده‌ها		پارامترهای آماری
Q ( $m^3s^{-1}$ )	SAR	Q ( $m^3s^{-1}$ )	SAR	Q ( $m^3s^{-1}$ )	SAR	
86/05	2/56	68	1/98	74/30	2/87	میانگین
67/80	1/87	121	1/23	98	1/40	انحراف معیار
0/98	0/31	1/20	0/23	0/65	0/10	ضریب چولگی
1/26	0/14	3/87	0/54	1/26	1/42	کمینه
398/30	10/10	100/30	3/81	398/3	10/10	بیشینه
0/43	0/68	0/38	0/76	0/41	0/66	$r_1$
0/40	0/44	0/36	0/69	0/38	0/56	$r_2$
0/23	0/35	0/13	0/54	0/17	0/48	$r_3$
-0/01	0/09	-0/01	0/34	0/10	0/20	$r_4$

جدول 2- ضرایب همبستگی  $SAR_t$  با  $Q_t$  و تأخیرات 1 تا 4 روزه  $Q$ .

کل داده‌ها	مجموعه آموزش	مجموعه آزمون	سری‌های زمانی
0/38	0/33	0/43	$Q_t$
0/26	0/18	0/32	$Q_{t-1}$
0/13	0/21	0/25	$Q_{t-2}$
0/09	-0/01	0/11	$Q_{t-3}$
0/001	-0/05	0/03	$Q_{t-4}$

مناسب، منطبق با شکل و رفتار تابع محرک سیگموئید (که مورد استفاده در شبکه عصبی هست) منتقل شوند. در این تحقیق از رابطه 3 برای استانداردسازی داده‌ها در محدوده [0 و 1] استفاده شده است (رجایی و همکاران: 2011):

$$X_{norm} = \frac{(X_i - X_{min})}{(X_{max} - X_{min})} \quad [3]$$

#### مدل‌سازی

با در نظر گرفتن نتایج تحلیل آماری ارائه شده در جداول 1 و 2 ترکیب‌های مختلف ورودی برای پیش‌بینی میزان SAR روز آینده در زمان t+1 به صورت زیر در نظر گرفته شد:

- 1) SAR<sub>t</sub>
- 2) SAR<sub>t</sub>, SAR<sub>t-1</sub>
- 3) SAR<sub>t</sub>, SAR<sub>t-1</sub>, SAR<sub>t-2</sub>
- 4) SAR<sub>t</sub>, Q<sub>t</sub>
- 5) SAR<sub>t</sub>, SAR<sub>t-1</sub>, Q<sub>t</sub>
- 6) SAR<sub>t</sub>, SAR<sub>t-1</sub>, Q<sub>t</sub>, Q<sub>t-1</sub>

در این تحقیق به منظور آموزش و آزمون از مدل شبکه عصبی سه لایه با روش پیش‌خور پس‌انتشار خطا<sup>9</sup> (FFBP) با الگوریتم آموزشی لوبنبرگ-ماکوارت<sup>10</sup> استفاده شد. از هر دو شبکه به صورت سه لایه پیش‌خور با تابع فعالیت سیگموئید<sup>11</sup> در لایه مخفی و تابع انتقال خطی<sup>12</sup> در لایه خروجی استفاده شده است. تعداد نرون لایه ورودی در ترکیب یک برابر 1، در ترکیب دو برابر 2، در ترکیب سه برابر 3، در ترکیب چهار برابر 2، در ترکیب پنج برابر 3 و در ترکیب شش برابر 3 هست. یکی از مواردی که برای ترکیب‌های فوق باید در نظر گرفت، مسئله هم‌خطی چندگانه هست. در واقع مسئله هم‌خطی چندگانه صرفاً در مورد روش‌های آماری، همچون

رابطه 1 برای محاسبه ضرایب همبستگی بین SAR و Q به کار گرفته شد:

$$r = \frac{\sum (Q_i - \bar{Q})(SAR_i - \bar{SAR})}{\sqrt{\sum (Q_i - \bar{Q})^2 \sum (SAR_i - \bar{SAR})^2}} \quad [1]$$

هر چه ضریب همبستگی به عدد 1 نزدیک‌تر باشد، نشان‌دهنده سازگاری مناسب بین داده‌هاست. بررسی ضرایب همبستگی بین SAR و Q نشان می‌دهد که همبستگی میان SAR با Q<sub>t</sub> و Q<sub>t-1</sub> بالاتر از سایر موارد است.

مطابق رابطه 2 پارامتر SAR وابسته به تغییرات سه یون Na, Ca, Mg است که تغییرات هر کدام از این یون‌ها با مقدار دبی متفاوت است. غلظت برخی از یون‌ها در ماه‌های با جریان کم (به علت تغییر دما و تبخیر) و برخی دیگر در ماه‌های با جریان بالا (به علت پدیده آبشویی) افزایش می‌یابد، بنابراین در مورد نحوه تغییرات SAR با روند دبی نمی‌توان به طور قطعی اظهار نظر نمود.

$$SAR = \frac{Na^+}{\sqrt{\frac{Ca^{2+} + Mg^{2+}}{2}}} \quad [2]$$

Mg<sup>2+</sup> یون منیزیم بر حسب میلی‌گرم بر لیتر، Ca<sup>2+</sup> یون کلسیم بر حسب میلی‌گرم بر لیتر، Na<sup>+</sup> یون سدیم بر حسب میلی‌گرم بر لیتر هست.

برای بررسی نرمال بودن توزیع داده‌های دبی و SAR از آزمون کولموگوروف - اسمیرنوف<sup>8</sup> استفاده شد. نتایج این آزمون نشان داد سطح معنی‌داری در هر دو متغیر دبی و SAR بزرگ‌تر از 0/05 است که نمایانگر طبیعی بودن توزیع داده‌ها است (P>0/05). در صورتی که داده‌ها به صورت خام وارد شبکه شوند، به علت تغییرات زیاد داده‌ها تأثیر متفاوتی بر روی شبکه می‌گذارند به طوری که برخی از نرون‌ها خیلی زود به حد فعالیت رسیده، در حالی که برخی دیگر حتی به آستانه فعالیت نیز نرسیده‌اند، لذا لازم است متغیرهای ورودی به بازه‌ای

<sup>11</sup> Logsig

<sup>12</sup> Linear transfer function

<sup>8</sup> Kolmogorov-Smirnov

<sup>9</sup> Feed-Forward Back propagation

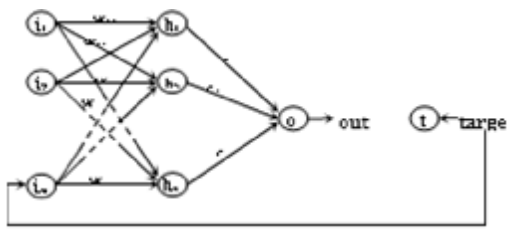
<sup>10</sup> Levenberg - Marquardt

ترکیب‌های ورودی شبکه عصبی استفاده می‌شود، این مدل به کار می‌رود، به عبارت دیگر:

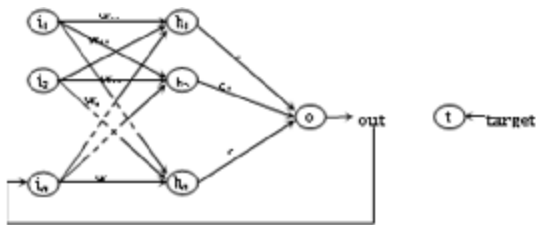
$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-d}) \quad [4]$$

$$y_t = f(x_{t-1}, x_{t-2}, \dots, x_{t-d}, y_{t-1}, y_{t-2}, \dots, y_{t-d}) \quad [5]$$

در این روابط  $d$  تعداد تأخیر است. به دلیل اینکه در هنگام آموزش شبکه خروجی صحیح در دسترس است، می‌توان معماری حلقه پس‌خور باز شکل 3 را به کار برد. اگر حلقه پس‌خور باز باشد تنها می‌توان یک گام جلو را پیش‌بینی کرد. پس از آموزش شبکه می‌توان حلقه پس‌خور را بست (شکل 4) و چند گام زمانی آینده را پیش‌بینی کرد. در هنگام آموزش شبکه عصبی باید به مسئله بیش‌برازش<sup>18</sup> دقت شود. در این حالت حین آموزش شبکه خطای مجموعه آموزشی به مقدار بسیار ناچیزی می‌رسد، ولی هنگام معرفی داده‌های اعتبارسنجی و آزمون خطای مدل زیاد می‌شود. در واقع شبکه در مرحله آموزش، روند داده‌ها را یاد نگرفته بلکه با به حافظه سپردن داده‌ها از بهینه کردن وزن ممانعت کرده و نتایج قابل تعمیم به مرحله آزمون نیست. بنابراین بهترین عملکرد شبکه عصبی زمانی است که خطای هر سه مرحله آموزش، اعتبارسنجی و آزمون کمینه گردد.



شکل 3- معماری شبکه عصبی در حالت حلقه پس‌خور باز.



شکل 4- معماری شبکه عصبی در حالت حلقه پس‌خور بسته.

معادله خط رگرسیون بوده و در مورد روش موجک-عصبی تأثیر منفی نمی‌گذارد. هم‌خطی بودن متغیرها یکی از مشکلاتی است که در سری زمانی ممکن است به رگرسیون کاذب بیانجامد. در این مقاله آزمون هم‌خطی در نرم‌افزار SPSS برای هر یک از ترکیب‌های شش‌گانه انجام گرفت. نتایج این آزمون با معیارهای ضریب تولورانس<sup>13</sup>، ضریب VIF<sup>14</sup>، مقدار ویژه<sup>15</sup> و شاخص وضعیت<sup>16</sup> قابل ارزیابی هست. ارزیابی تمامی این شاخص‌ها نشان داد که احتمالاً بین متغیرهای مستقل هم-بستگی داخلی وجود دارد. بررسی مقادیر ویژه برای این ترکیب‌ها نشان داد که به علت کوچک بودن این مقادیر (از 0/003 تا 0/008) امکان همبستگی داخلی (هم‌خطی بودن) بین متغیرهای مستقل وجود دارد. همچنین ضریب VIF برای این آزمون بین 0/9 تا 1/004 و ضریب تولورانس بین 0/974 تا 0/998 برای این شش ترکیب به دست آمد. در آزمون هم‌خطی اگر شاخص وضعیت بیشتر از 15 باشد، نشان‌دهنده احتمال هم‌خطی بین متغیرهای مستقل بوده و مقدار بیشتر از 30 بیانگر مشکل جدی در استفاده از رگرسیون در وضعیت موجود آن هست. در تحقیق حاضر، در تمامی ترکیب‌ها به جز ترکیب سه با شاخص 30، مقدار این شاخص کمتر از 30 بود. بنابراین در این مقاله برای تحلیل رگرسیون، با اعتماد بر نتایج آزمون فوق، از شش ترکیب معرفی شده برای مدل‌سازی استفاده شد.

### شبکه‌های عصبی NAR و NARX

در مدل شبکه عصبی NAR یا خودبازگشتی غیرخطی<sup>17</sup>، تنها داده‌های یک سری زمانی مورد استفاده قرار می‌گیرد، درحالی‌که در مدل شبکه NARX از داده-های سری زمانی دیگری نیز می‌توان استفاده کرد. مثلاً در تحقیق حاضر که از سری‌های زمانی دبی نیز در

<sup>13</sup> Condition index

<sup>14</sup> Nonlinear auto regressive

<sup>15</sup> Over training

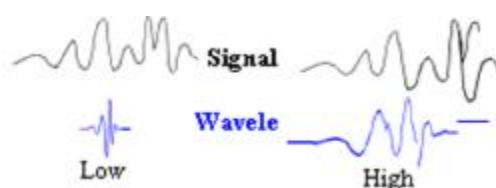
<sup>16</sup> Tolerance coefficients

<sup>17</sup> Variance inflation factor

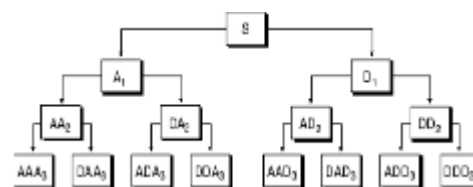
<sup>18</sup> Eigen value

## تحلیل موجک

روش متداول برای کاهش اغتشاش، تبدیل فوریه است که در تحلیل و پردازش سیگنال‌ها کاربرد دارد، اما این تبدیل دارای محدودیت‌هایی هست. اطلاعات زمان یا مکان ممکن است در طول انجام تبدیل خراب شوند و بعضی از خصوصیات سیگنال کاسته شود و در بعضی مواقع تشخیص زمان و مکان یک رویداد خاص غیرممکن باشد (لانگکین و همکاران 2013). نظریه موجک به عنوان ایده‌ای تکمیلی در جهت رفع مشکلات تبدیل‌های فوریه قادر هست بسیاری از منظرهای پنهان داده‌هایی که دیگر روش‌های تحلیل سیگنال در شناسایی آن‌ها ناتوان هستند را شناسایی کند. موجک، موجی کوچک و نوسان کننده است که انرژی آن در ناحیه کوچکی متمرکز شده و می‌تواند ابزاری مناسب برای بررسی پدیده‌های نایستا باشد. در تبدیل موجک گسسته<sup>19</sup> (DWT)، سیگنال از یک سری از فیلترهای بالاگذر و پایین‌گذر عبور داده می‌شود. بخش حاصل از عبور سیگنال از فیلتر بالاگذر که شامل اطلاعات فرکانس بالا از جمله اغتشاش هست، جزئیات<sup>20</sup> نام دارد. بخش حاصل از عبور سیگنال از فیلتر پایین‌گذر که شامل اطلاعات فرکانس پایین و دربرگیرنده مشخصات هویتی در سیگنال است و تغییرات آرام را نشان می‌دهد، تقریبات<sup>21</sup> نامیده می‌شود (شکل 5).



شکل 5- موجک با مقیاس بالا و پایین.



شکل 6- درخت تجزیه سری زمانی.

فرآیند تجزیه سیگنال به چندین جزء با قدرت تفکیک پائین شکسته می‌شود، این حالت درخت تجزیه موجک<sup>22</sup> نامیده می‌شود (شکل 6).

در این رابطه  $\Psi(t)_{a,b}$  موجک‌های پی‌درپی،  $a$  مقیاس یا فاکتور فرکانس،  $b$  فاکتور زمان و  $R$  دامنه اعداد حقیقی هستند. انتقال موجک گسسته تابع  $f(t)$  به صورت زیر نوشته می‌شود:

$$W_{\psi}f(j, k) = a_0^{-\frac{j}{2}} \int f(t) \bar{\psi}(a_0^{-j}t - kb_0) dt \quad [6]$$

با جایگذاری مقادیر  $a_0$  و  $b_0$  برای سری زمانی گسسته  $f(t)$ ، انتقال موجک گسسته به صورت ذیل تعریف می‌شود (پالانی و همکاران 2008):

$$W_{\psi}f(j, k) = 2^{-\frac{j}{2}} \sum_{i=0}^{N-1} f(t) \bar{\psi}\left(2^{-\frac{j}{2}}t - k\right) dt \quad [7]$$

رگرسیون خطی چند متغیره (MLR)

تحلیل رگرسیون خطی چند متغیره تکنیکی برای مدل کردن رابطه خطی بین یک متغیر وابسته و یک یا چند متغیر مستقل است. شکل معمول معادله MLR به صورت زیر نشان داده می‌شود که  $a_0$  عرض از مبدأ،  $a_i$  ضریب رگرسیون متغیرهای مستقل  $x_i$  و  $y$  مقدار پیش‌بینی شده است (رضوی پاریزی 1390):

$$y = a_0 + \sum a_i x_i \quad [8]$$

برنامه‌ریزی ژنتیک (GP)<sup>23</sup>

مبنای روش شبیه‌سازی GP بر اساس نظریه تکاملی داروین استوار بوده و یکی از شاخه‌های الگوریتم ژنتیک (GA)<sup>24</sup> هست. در GA متغیرهای تصمیم در قالب ژن‌ها در فرایند جستجو وارد می‌شوند، اما در مسائل بهینه‌سازی این امکان وجود دارد که تنها اعداد به عنوان متغیرهای تصمیم معرفی نشده و عملگرهای ریاضی یا منطقی نیز به عنوان متغیرهای تصمیم در فرایند بهینه‌سازی شرکت جویند. فرایند شبیه‌سازی در GP به این صورت است که ابتدا مجموعه‌هایی که می‌توانند جهت

<sup>22</sup> Packet decomposition tree

<sup>1</sup> Genetic programming

<sup>24</sup> Genetic algorithm

<sup>19</sup> Discrete wavelet transforms

<sup>20</sup> Details

<sup>21</sup> Approximation



### معیارهای ارزیابی مدل

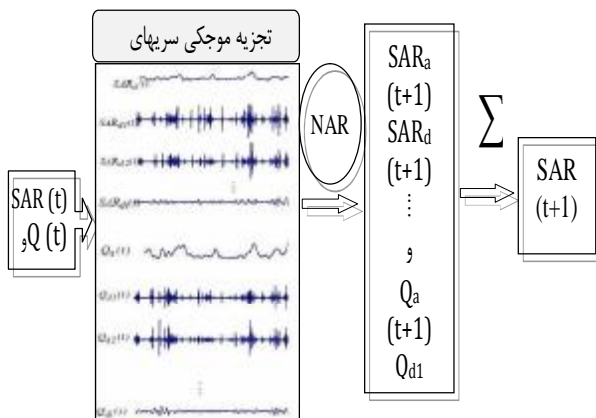
به منظور ارزیابی عملکرد مدل‌ها سه شاخص معیاری متفاوت شامل ریشه میانگین مربعات خطای بین داده‌های پیش‌بینی و مشاهده شده<sup>26</sup> (RMSE)، میانگین قدر مطلق خطا<sup>27</sup> (MAE)، ضریب کارایی نش-ساتکلیف<sup>28</sup> (E) به عنوان ارزیابی توان پیش‌بینی مدل، روابط (8 تا 10) مورد استفاده قرار گرفته است. به طور کلی هرچه ضریب تبیین به 1 و خطا به صفر نزدیک شود، مدل عملکرد بهتری دارد.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_i - Y_i)^2}{n}} \quad [9]$$

$$MAE = \frac{\sum_{i=1}^n |X_i - Y_i|}{n} \quad [10]$$

$$E = 1 - \frac{\sum_{i=1}^n (Y_i - X_i)^2}{\sum_{i=1}^n (\bar{Y}_i - X_i)^2} \quad [11]$$

در این روابط،  $X_i$  و  $Y_i$  به ترتیب آامین داده پارامتر برآوردی و اندازه‌گیری،  $\bar{Y}$  متوسط داده‌های  $Y_i$  و  $n$  تعداد نمونه‌های مورد ارزیابی است.  $E=1$  به این معنی است که داده‌های پیش‌بینی شده با داده‌های مشاهده شده همخوانی کامل دارد.  $E=0$  یعنی پیش‌بینی مدل منطبق بر مقادیر میانگین داده‌ها هست؛ درحالی‌که  $E<0$  نشان‌دهنده این است که میانگین داده‌های مشاهداتی برآورد بهتری نسبت به مدل پیش‌بینی شده دارد.



شکل 7- ساختار مدل پیشنهادی ترکیب شبکه عصبی - موجک.

انتخاب متغیرها و عملگرها در فرایند جستجو به کار روند، معرفی می‌شوند. سپس کروموزوم‌ها به وسیله انتخاب یک مجموعه جواب اولیه تصادفی از مجموعه‌های اتصال و توابع ایجاد می‌شوند، سپس تابع هدف متناظر با هر کروموزوم محاسبه می‌شود. سپس مجموع تفاوت مقادیر محاسباتی و مشاهداتی کمینه می‌شود. در مرحله بعد عملگرهای ژنتیک (تزیوج و جهش) اعمال می‌شوند. در ادامه، فرایند توسعه‌ای تکراری در مورد فرزندان انجام می‌شود تا در نهایت پس از تعداد تکرار مشخصی که روند تغییرات تابع هدف در آن تقریباً ثابت گردید، مقدار بهینه یا نزدیک به بهینه به عنوان جواب مسئله استخراج گردد.

در این مقاله برای پیش‌بینی به روش GP از نرم-افزار Discipulus محصول مؤسسه انستیتو فنی رجیستر<sup>25</sup> استفاده شده است. برای معیار توقف در این نرم‌افزار میزان حداکثر تعداد اجرا و میزان تولید بدون بهبود 300 در نظر گرفته شده است. در این نرم‌افزار از 75% داده‌ها برای آموزش، 25% برای صحت‌سنجی و آزمون استفاده شد.

### مدل پیشنهادی ترکیبی موجک-عصبی (WANN)

در مدل ترکیبی شبکه عصبی-موجک، زیرسری-هایی که توسط تبدیل‌های موجک تولید شده‌اند، به عنوان ورودی شبکه عصبی استفاده می‌شوند. مدل موجک-عصبی که سیگنال‌های چند مقیاسه را به عنوان داده‌های ورودی مورد استفاده قرار می‌دهد، نسبت به ورودی تک الگو پیش‌بینی قابل‌اعتمادتری را ارائه خواهد داد. توابع موجک انواع زیادی دارد که در این تحقیق برای تجزیه سری زمانی داده‌ها 4 نوع تابع متفاوت در سطوح تجزیه 1 تا 5 به کار گرفته شده است. ساختار مدل ترکیبی پیشنهادی در این تحقیق در شکل 7 نشان داده شده است.

<sup>26</sup> Nash-Sutcliffe model efficiency coefficient

<sup>27</sup> Register Machine Learning Technologies Inc.

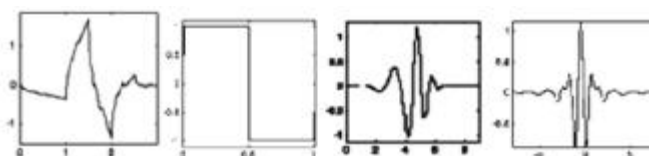
<sup>28</sup> Root mean squared error

<sup>29</sup> Mean absolute error

## نتایج و بحث

برای به دست آوردن بهترین نتیجه به منظور پیش‌بینی نسبت جذب سدیم، با در نظر گرفتن شدت همبستگی بین پارامترها، مدل‌های رگرسیون خطی چندمتغیره، شبکه عصبی و برنامه‌ریزی ژنتیک با استفاده از ترکیب‌های مختلف ورودی تشکیل یافت. سپس بهترین ترکیب با ضریب کارایی بالا و خطای کمتر توسط مدل موجک-عصبی شبیه‌سازی شد. در این مدل-ها پیش‌بینی 15 ماهه (دوره آزمون) انجام شد و مقادیر  $E$ ،  $RMSE$  و  $MAE$  محاسبه شد (جدول 3). برای مدل رگرسیون خطی چند متغیره، تحلیل رگرسیون خطی برای هر شش ترکیب مذکور انجام شد. در اجرای مدل رگرسیونی، نسبت جذب سدیم ماه فعلی به‌عنوان متغیر وابسته و سایر پارامترها در ترکیبات ورودی مربوط به ماه‌های قبل به‌عنوان متغیر مستقل در نظر گرفته شدند. بهترین نتیجه برای این مدل مربوط به ترکیب‌های ورودی 4 و 6 با ضریب کارایی 0/33 و 0/35 هست (جدول 3). مدل‌های شبکه عصبی بر اساس تعداد لایه‌های مخفی، تعداد نرون‌های لایه مخفی و تعداد تأخیرها (1 تا 4 ماه قبل) آموزش داده شد و مورد آزمون قرار گرفت. در اجرای مدل شبکه عصبی  $NAR$  و  $NARX$ ، مطابق جدول 4 بهترین نتیجه در ترکیب 4 با معماری شبکه 1-8-2 (یعنی شبکه یک لایه با دو نرون در لایه ورودی، هشت نرون در لایه مخفی و یک نرون در لایه خروجی) حاصل شد. ترکیب 4 نشان می‌دهد که مقادیر  $SAR$  در ماه بعدی به مقدار دبی و  $SAR$  در همان زمان وابسته است. از بالا بودن مقدار همبستگی یک‌ماهه  $SAR$  (برای مجموعه آموزش) طبق جداول 1 و 2 انتظار می‌رفت که افزودن دبی و نسبت جذب سدیم با یک ماه تأخیر باعث بهبود نتایج گردد. همان‌گونه که از جدول 3 مشخص است، مدل

رگرسیون خطی ضعیف بوده و قادر به شبیه‌سازی نوسانات شدید و پی‌درپی و پیش‌بینی مناسب  $SAR$  نیست. در مدل برنامه‌ریزی ژنتیک بهترین نتیجه برای الگوی ورودی دوم به دست آمد. برای مدل ترکیبی موجک-عصبی ابتدا سری‌های زمانی نسبت جذب سدیم و دبی با تأخیرهای مطابق ترکیب‌های ورودی توسط موجک‌های مادر تجزیه شدند. برای این کار از موجک‌های مادر دابیز 2، دابیز 5، هار و میر با سطوح تجزیه 1 تا 5 استفاده شد. شکل 8 این موجک‌ها را نشان می‌دهد. بانژاد و همکاران (1392) نیز از موجک مورلت برای تجزیه سیگنال‌های اصلی پارامتر  $SAR$  استفاده کرده‌اند. در تحقیق حاضر در ترکیب‌های یک تا سه فقط از پارامتر  $SAR$  برای پیش‌بینی استفاده شده و دبی جریان در ورودی‌های مدل لحاظ نشده، ولی در سه ترکیب بعدی پارامتر دبی نیز لحاظ شده است. با توجه به ضرایب همبستگی و خودهمبستگی جداول 1 و 2، وجود دبی و  $SAR$  تنها با یک ماه تأخیر دقت بالاتری داشته و در مدل شبکه عصبی نیز ترکیب چهار نتایج بهتری را نسبت به سایر ترکیب‌ها ارائه کرده است. بنابراین با توجه به نتایج بهتر ترکیب چهار در مدل شبکه عصبی و همبستگی بالای  $SAR$  به ازای تأخیرهای یک‌ماهه، این ترکیب به‌عنوان ورودی برای مدل پیشنهادی موجک-عصبی انتخاب شد. زیرسری‌های تقریب و جزئیات به‌دست‌آمده از ناپیوستگی موجکی سیگنالی  $SAR_t$  و  $Q_t$  به‌صورت جداگانه به شبکه عصبی اعمال شدند. نتیجه نهایی مدل ترکیبی، جمع جبری مقادیر پیش‌بینی‌شده هر یک از زیرسری‌هاست. به نظر می‌رسد با افزایش سطح تجزیه و ظاهر شدن نمونه‌های فرکانس پایین، عملکرد مدل موجک-عصبی مناسب‌تر باشد ولی مطابق جدول 4 بهترین نتیجه در سطح تجزیه 3 با موجک دابیز 2 به‌دست‌آمده است.



شکل 8- موجک‌های ماد: (a) دابیز 2 (b) هار (c) دابیز 5 (d) میر.

جدول 3- نتایج مدل‌های رگرسیون، شبکه عصبی و برنامه‌ریزی ژنتیک.

ANN						MLR						GP			انواع مدل‌ها
6	5	*4	3	2	1	*6	5	4	3	2	1	3	*2	1	ترکیب‌ها
9	6	8	6	3	11	-	-	-	-	-	-	-	-	-	نرون‌های لایه پنهان
-	-	-	-	-	-	-	-	-	-	-	-	100	100	100	تعداد اجرا
0/71	0/48	0/76	0/53	0/68	0/64	0/35	0/27	0/33	0/20	0/24	0/21	0/48	0/55	0/44	E
0/87	2/54	0/55	2/15	1/45	1/87	2/54	2/67	2/54	2/88	2/80	2/88	2/32	1/95	2/60	MAE
2/18	5/40	2/21	5/10	3/98	4/15	5/43	5/60	5/32	5/80	2/82	5/88	5/20	4/55	5/65	RMSE

\* بهترین نتیجه به دست آمده برای هر یک از مدل‌ها

هر یک از عوامل فوق را که به صورت ضمنی در بین داده‌ها قرار گرفته است، در تاریخچه سری زمانی SAR رمزگشایی کنند. عملکرد این 3 سطح به گونه‌ای است که در هر سطح قسمت‌هایی از سیگنال SAR که با فرکانس‌های مورد نیاز در طبقه‌بندی سیگنال SAR همبستگی بالایی داشتند، در ضرایب موجک میرنگه داشته شده‌اند. البته تأثیر پارامتر دبی به صورت مستقیم در ترکیب‌های ورودی پیشنهادی لحاظ شده است. زیرسری‌های تجزیه موجک توانسته‌اند اثر اغتشاش‌هایی که آلودگی‌های نقطه‌ای و غیر نقطه‌ای (در سری زمانی مشاهداتی SAR) وارد کرده‌اند را در پیش‌بینی این پارامتر لحاظ کنند. با توجه به اطلاعات جغرافیایی ذکر شده ایستگاه رودبار، توزیع رودها و جویبارهای موجود در این حوضه به گونه‌ای است که رواناب‌های سطحی از جهات مختلف جغرافیایی که از مسیرهای اصلی جریان آب به سمت رودخانه هستند، به این سمت سرازیر می‌شوند. هر یک از این شاخه‌ها دربرگیرنده پساب کارخانه‌ها، کشتارگاه‌ها، پساب کشاورزی حاصل از کودها و سموم دفع آفات هست. این عوامل علت بخش دیگری از نویزهای موجود در سری زمانی می‌باشند که به صورت افزایش و کاهش‌های متناوب در سری زمانی SAR دیده می‌شوند.

جدول 4 به وضوح برتری مدل ترکیبی موجک-عصبی را نسبت به مدل‌های شبکه عصبی، برنامه‌ریزی ژنتیک و رگرسیون خطی نشان می‌دهد. بهترین نتایج مدل ترکیبی مربوط به دو موجک میر و دابیچز2 با ضرایب کارایی 0/80 و 0/81 هست. به نظر می‌رسد موجک میر و دابیچز2 به لحاظ دارا بودن شکل هندسی مناسب برای سری‌های زمانی SAR جواب‌های مناسب‌تری را نسبت به دو موجک دیگر نتیجه داده‌اند. با توجه به شکل 1 سری زمانی اغلب دارای ویژگی نوسان پی‌درپی و روند کاهش و افزایشی متناوب بوده و نقاط بیشینه و کمینه زیادی دارد که این روند شباهت نسبی به موجک دابیچز2 دارد. با توجه به رابطه SAR با یون‌های موجود در آب [رابطه 2] و تغییرات این یون‌ها با دبی، مهم‌ترین عواملی که می‌توانند در فرکانس نسبت سدیم جذبی در آب تأثیرگذار باشند، کاهش و افزایش ناگهانی دبی رودخانه، فرسایش و آبخوئی و انتقال رسوبات، رس، لای، مواد آلی و غیر آلی، عملیات مکانیکی در بستر فعال و غیرفعال رودخانه و افزایش سطح تماس با سایر کانی‌ها می‌باشند. با توجه به اینکه کانی‌های کربناته نقش مهمی در قلیابیت دارند، سطح تماس با آن‌ها باعث افزایش SAR در آب می‌شود. جدول 4 نشان می‌دهد که زیرسری‌های تجزیه‌شده موجک دابیچز2 در 3 سطح بهتر از سطوح دیگر توانسته‌اند تأثیر

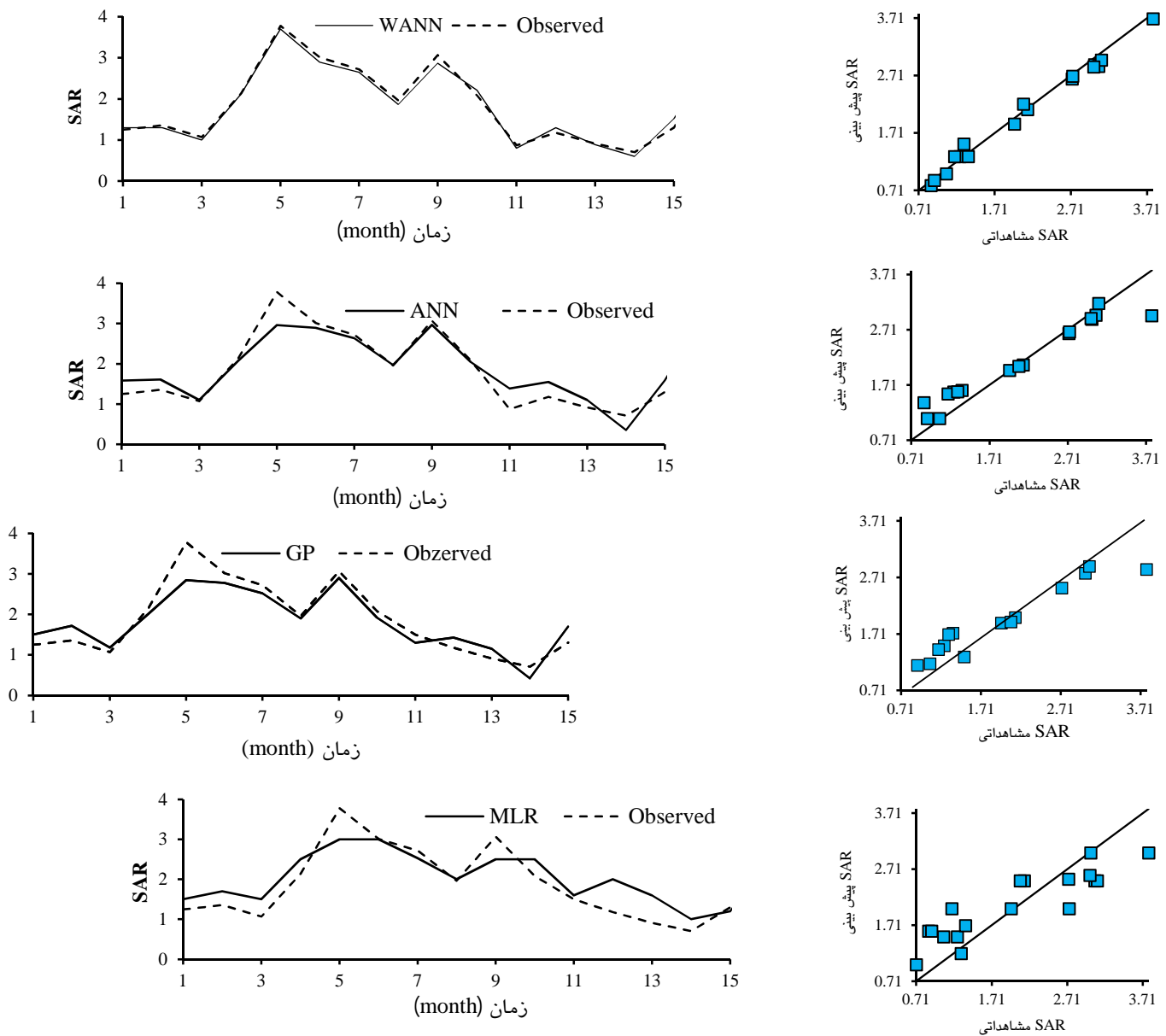
جدول 4- نتایج به دست آمده از مدل ترکیبی شبکه عصبی-موجک.

MAE	RMSE	E	ساختار شبکه عصبی	سطوح تجزیه	انواع موجک مادر
1/85	4/16	0/65	2-3-1	1	دایچز 2
0/88	2/16	0/71	2-5-1	2	
0/61	1/87	*0/81	2-7-1	3	
2/24	4/98	0/54	2-9-1	4	
2/61	5/45	0/43	2-7-1	5	
2/20	4/95	0/54	2-5-1	1	دایچز 5
2/30	5/30	0/50	2-9-1	2	
1/48	4/08	0/65	2-11-1	3	
2/54	5/32	0/32	2-11-1	4	
2/88	5/88	0/22	2-2-1	5	
2/16	5/12	0/51	2-3-1	1	میر
1/44	3/98	0/67	2-7-1	2	
0/68	1/86	0/78	2-11-1	3	
0/85	2/10	0/73	2-9-1	4	
0/60	1/87	*0/80	2-7-1	5	
2/55	5/31	0/34	2-5-1	1	هار
2/78	5/81	0/23	2-11-1	2	
2/21	4/93	0/55	2-9-1	3	
3/56	6/05	0/11	2-7-1	4	
3/51	5/90	0/09	2-5-1	5	

\* بهترین نتیجه به دست آمده برای هر یک از مدل‌ها

زمان‌ها دست بالا پیش‌بینی کرده است. این روش قادر به تشخیص روابط پیچیده بین داده‌های سری زمانی نبوده و پیش‌بینی با خطای نسبتاً بالایی همراه است. در صورتی که شبکه عصبی قادر به انطباق با راه‌حلهایی برای جبران تغییرات به وجود آمده در طول زمان بوده است، به خصوص مقادیر بیشینه و کمینه را نسبتاً به خوبی شناسایی کرده است. این حالت در نمودار مربوط به موجک به دقت بیشتری می‌رسد. مدل ترکیبی موجک-عصبی نه تنها قادر به تشخیص تغییرات سری زمانی بوده بلکه نقاط شکستگی، بیشینه و کمینه را به خوبی پیش‌بینی کرده است. لانگ‌کین و همکاران (2013) نیز به این نتیجه رسیدند که مدل موجک-عصبی نقاط شکستگی، بیشینه و کمینه را در تخمین پارامترهای کیفی آب با دقت بیشتری پیش‌بینی می‌کند.

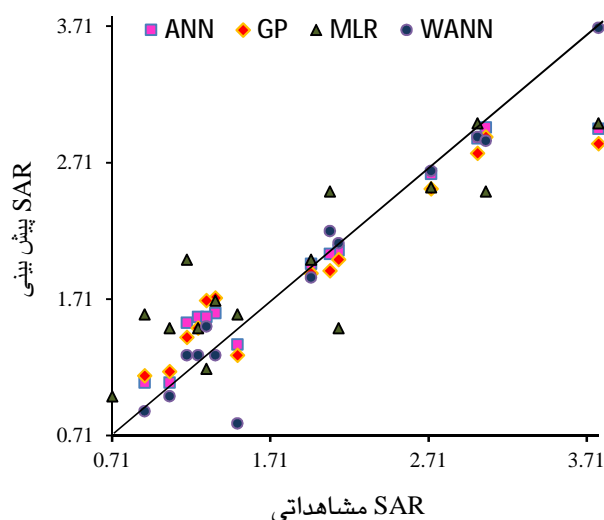
شکل 9، SAR پیش‌بینی شده توسط چهار مدل در دوره آزمون را در مقایسه با مقادیر SAR مشاهده شده نشان می‌دهد. با توجه به نمودارها، نمودار مدل موجک-عصبی با نمودار مقادیر مشاهده شده تطابق بیشتری دارد. بانژاد و همکاران (1392) نیز بر بالا بودن صحت مدل موجک-عصبی نسبت به مدل‌های شبکه عصبی و رگرسیون در پیش‌بینی SAR تأکید کرده‌اند. عملکرد ضعیف روش رگرسیون خطی در شکل 9 با پراکندگی داده‌ها در اطراف خط نیم‌ساز دیده می‌شود. هرچه این نقاط به نیم‌ساز نزدیک‌تر باشند، پیش‌بینی بهتری انجام شده است. به نظر می‌رسد روش رگرسیون مدل مقادیر میانگینی را ارائه داده و در هیچ‌یک از نوسان‌ها و نقاط حداکثری با مقادیر مشاهداتی تطابق ندارد. در نقاط حداکثری مقادیر را کمتر از مقدار واقعی و در سایر



شکل 9- مقایسه SAR اندازه‌گیری و پیش‌بینی‌شده برای هر چهار مدل.

ضریب E از 0/76، 0/35 و 0/65 در مدل‌های شبکه عصبی، رگرسیون خطی و GP به مقدار 0/81 در مدل ترکیبی موجک-عصبی بهبود می‌یابد.

در شکل 10 نتایج مدل‌های ANN، MLR، WANN و GP در برابر داده‌های مشاهداتی به صورت هم‌زمان رسم شده تا نتایج به خوبی مشهود باشند. بررسی کارایی در بهترین ترکیب‌های هر چهار مدل فوق نشان می‌دهد که



شکل 10- نمایش هم‌زمان عملکرد مدل‌های به‌کار گرفته‌شده در تحقیق.

مطابق جدول 5، مقدار RMSE محاسبه‌شده در روش برنامه‌ریزی ژنتیک کمتر از مدل رگرسیون هست. همچنین مدل ترکیبی موجک-عصبی با خطای کمتری مقادیر بیشینه را پیش‌بینی کرده است.

با توجه به شکل 9، یکی از ویژگی‌های موجک توانایی تخمین و شبیه‌سازی نقاط حداکثری و کمینه نسبت به سه روش دیگر هست. جدول 5 توان پیش‌بینی مقادیر بیشینه و کمینه در 15 ماهه آخر بازه زمانی را توسط چهار مدل نشان می‌دهد.

جدول 5- پیش‌بینی مقادیر بیشینه SAR توسط چهار مدل.

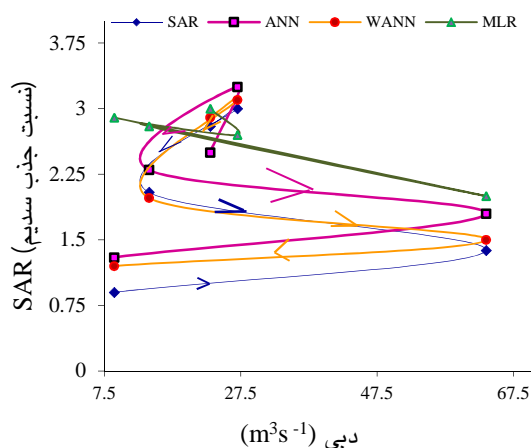
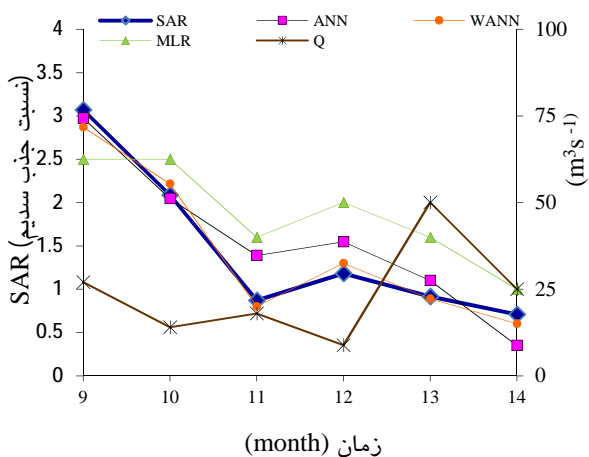
ANN	WANN	GP	MLR	اندازه‌گیری شده	تاریخ	ردیف
3/12	3/76	3/17	3/09	3/81	آذر 88	1
2/70	2/73	2/55	2/45	2/78	دی 88	2
2/54	2/66	2/23	2/12	2/67	فروردین 89	3
2/89	2/61	2/91	2/54	2/65	اردیبهشت 89	4
2/87	2/90	2/77	2/70	3/10	خرداد 89	5
14/12	14/66	13/63	12/90	15/01		جمع مقادیر خطا (%)
%6	%2/30	%9	%15			RMSE
0/68	0/21	0/76	1/78			

#### حلقه پس‌ماند<sup>1</sup>

دبی رخ دهد، احتمال به وجود آمدن این پدیده وجود دارد. در شکل 11 پدیده فوق در بازه زمانی آزمون، بین ماه‌های مهر تا بهمن رسم شده است. ویلیامز (1989) پنج کلاس معمولی از حلقه‌های پس‌ماند را شامل: ساعتگرد، پادساعت‌گرد، شکل 8 مانند، خطی<sup>2</sup> و خطی به‌اضافه یک حلقه<sup>3</sup> معرفی نمود. با بررسی و تحلیل این منحنی‌ها می‌توان از منابع آلاینده و نحوه انتقال آن

اطلاع از تغییرات میزان یون‌های معلق در مقابل نوسانات ناگهانی و تدریجی دبی رودخانه‌ها از مسائل مهم در مدیریت حوضه‌های آبریز است. در برخی موارد این تغییرات موجب پدیده پس‌ماند می‌شود. به‌عبارت‌دیگر تغییرات SAR وابسته به تغییرات دبی بوده و اگر تغییرات دوره‌ای SAR، همراه با تأخیر در پی تغییرات

سری زمانی پارامتر و اطلاعات مهم و کلیدی از حافظه آن استخراج کند.



شکل 11- نمودار SAR مشاهده شده و پیش‌بینی شده، رابطه پس‌ماند بین دبی و SAR در یک بازه از دوره آزمایش.

بین تعداد داده‌های سری زمانی و سطح تجزیه موجکی رابطه معناداری وجود دارد. به طوری که خروجی مطلوب مربوط به هر پارامتر در سطح تجزیه 3 حاصل شد. زیرا افزایش سطح تجزیه به منزله وزن‌دهی به

آگاهی یافت. برای دو مدل شبکه عصبی و موجک-عصبی، 8 مانند هست. در این بررسی به علت شباهت عملکرد مدل شبکه عصبی و مدل برنامه‌ریزی ژنتیک و نزدیک بودن پیش‌بینی‌های این دو روش، نمودار برای مقادیر برنامه‌ریزی ژنتیک ترسیم نشده است. همان‌طور که از شکل 11 مشخص است، روش رگرسیون این پدیده را تشخیص نداده و رابطه‌ای کاملاً مستقیم و خطی از دو پارامتر دبی و SAR ارائه کرده است. طبق این روش با کاهش دبی، میزان SAR سیری کاملاً نزولی داشته و اثر تناوب‌ها و نوسانات دیده نمی‌شود. در حالی که با توجه به نمودارهای ارائه شده برای دو مدل شبکه عصبی و موجک-عصبی مشخص می‌شود که چرخه تغییرات SAR و دبی به صورت غیرخطی و مشابه الگوی کلاس 8 مانند هست.

### نتیجه‌گیری کلی

در این تحقیق کاربرد مدل‌های شبکه عصبی، برنامه‌ریزی ژنتیک، موجک-عصبی و رگرسیون خطی چند متغیره در پیش‌بینی نسبت جذب سدیم آب رودخانه سفیدرود در ایستگاه رودبار بررسی شد. با بررسی سری زمانی SAR، مشخص شد که سری زمانی غیر ایستا بوده، ماهیت نوسانی داشته و تنها لحاظ کردن تغییرات زمانی آن راهگشا نیست، بلکه باید اثر فرآیندی که ساختار پارامتر را در طول زمان به وجود می‌آورد نیز لحاظ شود. نتایج مدل‌سازی‌های تحقیق نشان داد که ترکیب موجک و شبکه عصبی در بازه‌های زمانی پراگتاش با دقت خوبی عمل می‌کند. به عبارت دیگر، در ابتدا به وسیله ناپیوستگی موجکی اختلالات از روند اصلی جدا می‌شوند و سپس شبکه عصبی با تشخیص تعداد تأخیرها و همچنین دور یا نزدیک بودن زمان وقوع اغتشاش‌ها نسبت به گام زمانی پیش‌بینی، تخمین واقعی-تری از پارامتر مورد نظر دارد. تحلیل موجک با لحاظ مکانیسم سری زمانی قادر است تفسیری از ساختار

<sup>۲۱</sup> Linear valued plus a loop

<sup>۲۲</sup> Hysteresis loop

<sup>۲۳</sup> Linear valued

کارآیی E را بهبود بخشد. استفاده از موجک‌های مادری که از نظر شکل هندسی به سیگنال سری زمانی SAR شبیه‌تر بودند، نتایج بهتری را در برداشت. مزیت مهم مدل‌های به کار گرفته شده در تحقیق حاضر این بود که تنها با استفاده از سری‌های زمانی دبی و خود پارامتر SAR و بدون استفاده از پارامترهای دیگر (که هزینه و زمان‌بر هستند) قادر به پیش‌بینی بود. مدل ترکیبی توانست رابطه پس‌ماند بین دو پارامتر دبی و SAR را به خوبی مدل نماید، درحالی‌که مدل رگرسیونی این پدیده را تشخیص نداد. در این تحقیق انتخاب ورودی‌ها بر اساس تحلیل همبستگی سری‌های زمانی صورت گرفت. پیشنهاد می‌شود در تحقیقات بعدی از آنتروپی موجک برای انتخاب ورودی‌ها استفاده شود. همچنین از سری‌های زمانی روزانه یا هفتگی در بازه زمانی طولانی‌تری استفاده شود. این کار باعث می‌شود اغتشاشات سیگنال و فیلترهای پایین‌گذر کاهش یابد و با ورود به شبکه عصبی، حاصلی هموارتر و منطبق بر داده‌های پارامتر اندازه‌گیری شده وجود داشته باشد.

داده‌هایی است که تعداد نمونه‌های کمتری از آن سیگنال در سری اصلی وجود دارد و در بازسازی سیگنال اصلی کم اثر می‌باشند. بنابراین دقت تخمین واریانس نویز کاهش یافته و نتیجه معکوس روی پیش‌بینی خواهد داشت. همچنین مشخص شد مدل MLR ضمن عملکرد آن حول میانگین داده‌ها، نتایج ضعیفی از خود به نمایش می‌گذارد. در مواردی تنها روند کلی سری زمانی را پیش‌بینی می‌کند و شبیه‌سازی نقاط بیشینه و کمینه تفاوت زیادی با مقادیر مشاهده‌ای دارد. درحالی‌که در مدل‌های هوشمند عصبی و عصبی موجک، ماهیت خروجی‌ها به هیچ وجه تصادفی نبوده و هر پیش‌بینی برگرفته از چندین الگوی پیش پرداز شده در ورودی هست. در واقع در این روش‌ها مدل هوشمند با تشخیص روند افزایشی و کاهش، مقادیر خود را با مقادیر واقعی انطباق می‌دهند. در تمامی شکل‌های مربوط به شبکه عصبی حرکت مدل به سمت انطباق با مقادیر واقعی دیده می‌شود. مدل ترکیبی موجک-عصبی توانست مقدار خطای پیش‌بینی را نسبت به مدل‌های شبکه عصبی، برنامه‌ریزی ژنتیک و رگرسیون کاهش دهد و ضریب

#### منابع مورد استفاده

- باقرزاده آ، 1377. مدیریت زیست‌محیطی رودخانه سفیدرود قبل و بعد از بازشدن دریاچه‌های تخلیه رسوبات سد منجیل، پایان‌نامه کارشناسی ارشد، دانشگاه آزاد اسلامی واحد علوم و تحقیقات تهران.
- بانژاد، ح، کمالی م، امیرمرادی ک و علیائی ا، 1392. تخمین برخی پارامترهای کیفی رودخانه‌ها با استفاده از مدل هیبرید شبکه‌های عصبی-موجکی (منطقه مطالعاتی: رودخانه جاجرود تهران و قره‌سو کرمانشاه)، مجله سلامت و محیط، فصلنامه علمی پژوهشی انجمن علمی بهداشت محیط ایران، دوره 6، شماره 3، صفحه‌های 277 تا 294.
- رجائی ط و میر باقری ا، 1388. مدل بار معلق رودخانه‌ها با استفاده از شبکه‌های عصبی مصنوعی. نشریه دانشکده مهندسی دانشگاه فردوسی مشهد، جلد 21، شماره 1، صفحه‌های 27 تا 43.
- رضوی پاریزی ا، 1390. مقدمه‌ای بر تحلیل رگرسیون خطی (ترجمه)، انتشارات دانشگاه شهید باهنر کرمان.
- نوشادی م، سالمی ح و احمدزاده م، 1386. شبیه‌سازی و پیش‌بینی برخی از پارامترهای کیفیت آب رودخانه زاینده رود با استفاده از شبکه عصبی مصنوعی. نشریه آب و فاضلاب، شماره 64، صفحه‌های 49 تا 65.
- نیکو م، نیکو مه، بابایی نژاد ت، امیری آ و رستم پور ق، 1390. تعیین کیفیت آب در طول مسیر رودخانه با استفاده از شبکه‌های عصبی مصنوعی تکاملی (مطالعه موردی رودخانه کارون بازه شهید عباسپور-عرب اسد). فصلنامه تخصصی علوم و مهندسی آب، سال 1، شماره 1، صفحه‌های 45 تا 58.
- Asadollahifari G, Hemati A, Moradinejad S and Asadollahifardi R, 2013. Sodium adsorption ratio (SAR) prediction of the chalgazi river using artificial neural network (ANN) iran. Current World Environment 8(2): 169-178.



- Longqin Xu and Shuangyin Liu, 2013. Study of short-term water quality prediction model based on wavelet neural network. Elsevier, Journal of Mathematical and Computer Modeling 58: 807-813.
- Nourani V, Alami MT and Aminfar MH, 2009. A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation. Engineering Application of Artificial Intelligence 22: 146-477.
- Palani S, Shie-Yui Liong and Tkalich P, 2008. An ANN application for water quality forecasting. Marine Pollution Bulletin 56: 1586-1597.
- Qian-JinL, Zhi-HueS, Nu-FangF, Hua-DeZ and Lei A, 2013. Modeling the daily suspended sediment concentration in a hyperconcentrated river on the Loess Plateau, China, using the Wavelet-ANN approach. Geomorphology 186: 181-190.
- Rajae T, 2011. Wavelet - ANN combination model for prediction of daily suspended sediment load in rivers. Science of the Total Environment 409: 2917-2928.
- Rajae T, Nourani V, Zounemat-Kermani M and Kisi O, 2011. River suspended sediment load prediction: Application of ANN and Wavelet conjunction. Journal of Hydrologic Engineering 16: 613-627.
- Singh, R, 2012. Wavelet-ANN model for flood events. Advances in Intelligent and Soft Computing 131: 165-175.
- Williams G, 1989. Sediment concentration versus water discharge during single hydrologic events in rivers. Journal of Hydrologic Engineering 111: 89-106.