

## پیش‌بینی پارامترهای کیفی ( $NO_3^-$ , CL) رودخانه کرج با استفاده از مدل‌های ترکیبی شبکه

### عصبی-موجکی، ANN و MLR

طاهر رجایی<sup>1</sup> - رقیه رحیمی بنماران<sup>2\*</sup>

تاریخ دریافت: 1393/01/20

تاریخ پذیرش: 1394/02/13

#### چکیده

کنترل کیفیت آب رودخانه کرج، به‌عنوان یکی از مهم‌ترین منابع تأمین‌کننده آب از اهمیت بسزایی برخوردار است. پیش‌بینی پارامترهای کیفی آب، ابزاری سودمند در جهت مدیریت منابع آب می‌باشد. در این تحقیق، عملکرد مدل‌های شبکه عصبی مصنوعی (ANN)، مدل ترکیبی شبکه عصبی-موجکی (WANN) و رگرسیون خطی چند متغیره (MLR)، در پیش‌بینی یک ماه آینده یون نیترات و کلراید ایستگاه ورودی آبگیر بیلقان واقع در رودخانه کرج، مورد ارزیابی قرار گرفت. برای این منظور از یک دوره آماری جهت ورودی مدل‌ها استفاده شد. در مدل ترکیبی WANN سری‌های زمانی واقعی دبی و پارامتر کیفی مورد نظر (نیترات و کلراید) توسط آنالیز موجکی در سطوح مختلف تجزیه شده و به‌عنوان ورودی ANN به کار گرفته شد. کارایی مدل‌ها با ضریب تبیین (E) و ریشه میانگین مربعات خطا (RMSE) ارزیابی شدند. نتایج حاکی از دقت بالای مدل ترکیبی شبکه عصبی-موجکی نسبت به دو مدل دیگر بوده است؛ به طوری که مدل ترکیبی شبکه عصبی-موجکی قادر بود میزان RMSE را برای یون نیترات در مقایسه با مدل شبکه عصبی و رگرسیون خطی چند متغیره به ترتیب به مقدار 30/13 درصد و 71/89 درصد و برای یون کلراید، به اندازه 31/3 درصد و 57/1 درصد بهبود بخشد. در ادامه، توانایی هر سه مدل، برای پیش‌بینی نقاط پیک سری زمانی بررسی شد که مدل ترکیبی WANN پیش‌بینی بهتری را در مقایسه با دو مدل دیگر در برداشت.

**واژه‌های کلیدی:** تبدیل موجکی، رگرسیون خطی، رودخانه کرج، شبکه عصبی، یون نیترات و کلراید

سروکار دارد (10).

#### مقدمه

نیترات ( $NO_3^-$ ) و کلراید ( $Cl^-$ ) از مهم‌ترین پارامترهای کیفی آب به شمار می‌آیند. نیترات ماده‌ای است که آب‌های سطحی را به طور گسترده‌ای آلوده می‌کند. تجمع نیترات در محیط ناشی از هرزآب‌های کشاورزی (که به علت مصرف بیش از حد کودهای نیتراتی حاوی این ماده هستند)، به طور منتشر آب را آلوده می‌کنند یا آلودگی‌های نقطه‌ای از فاضلاب‌های انسانی می‌باشد. نیترات شاخص مناسبی برای نشان دادن آلودگی از نظر بهداشت آب آشامیدنی است که به‌عنوان یک ماده مغذی بوجود آورنده یوتریفیکاسیون نیز حائز اهمیت است (27). یون کلراید از جمله آلودگی‌های رودخانه به‌شمار می‌آید که ناشی از فعالیت‌های بشری است؛ افزایش آن در آب‌های سطحی و زیرزمینی آلودگی محسوب می‌شود (8). حداکثر مقدار مجاز آلاینده، برای نیترات در حد 10 پی‌پی‌ام نیتروژن یا 44 پی‌پی‌ام نیترات و استاندارد کلراید آب معادل 250 میلی‌گرم بر لیتر تعیین شده است. پیش‌بینی یون‌های نیترات و کلراید نقش تعیین‌کننده‌ای در

رودخانه‌ها، به‌عنوان یکی از منابع تأمین‌کننده نیاز آبی جوامع بشری از اهمیت خاصی برخوردار بوده و در حقیقت شریان‌های هر کشور محسوب می‌گردند. علاوه بر کمیت و میزان آورد رودخانه، باید کیفیت آب نیز جز پارامترهای مهم، مورد توجه قرار گیرد؛ لذا بررسی و پیش‌بینی پارامترهای کیفی در طول یک رودخانه، یکی از اهداف مدیران و برنامه‌ریزان منابع آب می‌باشد (13). در این راستا شناخت کافی از شرایط منطقه و مشخصات شیمیایی و بیولوژیکی آب رودخانه و نیز تغییرات پارامترهای کیفی به‌عنوان گام اول برنامه‌های مدیریت کیفیت منابع آب در نظر گرفته می‌شود؛ به عبارتی مدیریت کیفیت با شناخت کیفیت موجود آب، آلودگی‌ها و روش‌های حفظ کیفیت آب

1 و 2- دانشیار و کارشناس ارشد سازه‌های هیدرولیکی، گروه مهندسی عمران، دانشکده فنی و مهندسی، دانشگاه قم

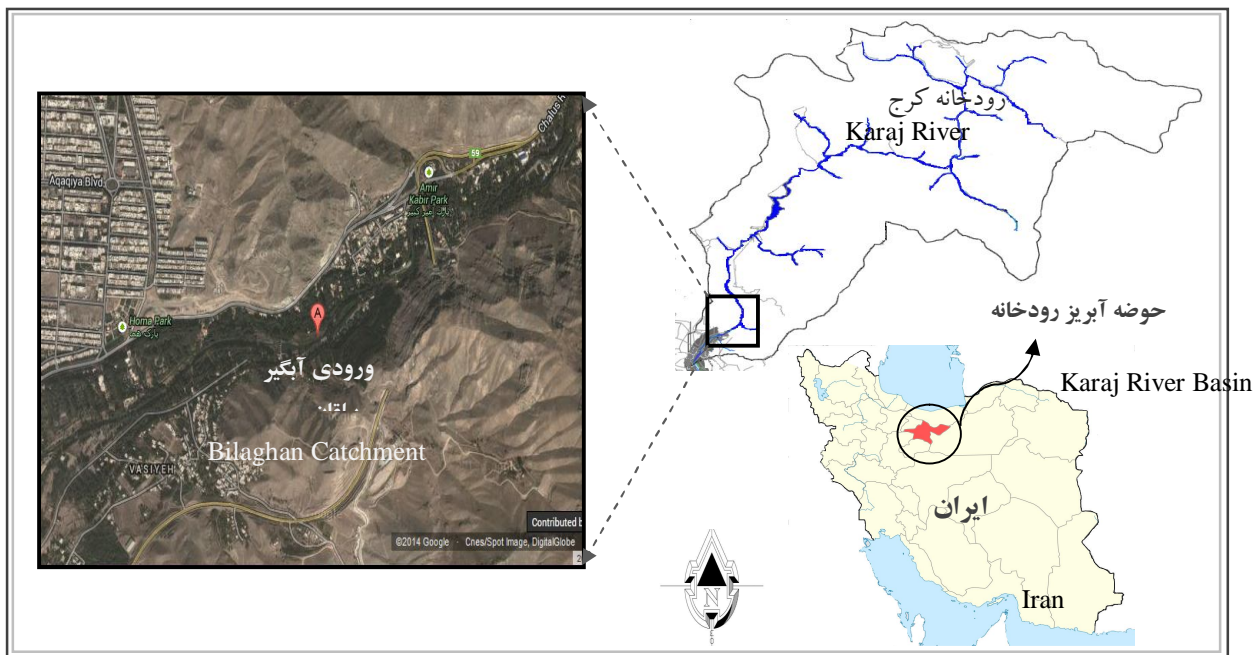
\* - نویسنده مسئول: (Email: r.rahimi\_b@yahoo.com)

در حالتی که به طور مجزا در مسائل مربوط به آب و محیط زیست به کار گرفته شده‌اند، عملکرد مناسبی داشته‌اند. اخیراً تمایل به استفاده از مدل‌های ترکیبی شبکه عصبی-موجک به طور روز افزون افزایش یافته است (22). کیم و والدز (11) برای پیش‌بینی خشکسالی و کاناس و همکاران (3) برای مدل‌سازی بارش رواناب ماهانه در تحقیقات مجزایی مدل ترکیبی شبکه عصبی-موجکی را توسعه دادند. این تحقیقات نشان داده که شبکه عصبی آموزش داده شده با داده‌های پیش پردازش شده توسط موجک، در مقایسه با شبکه‌های عصبی که با داده‌های خام و نویزدار آموزش داده شده‌اند، کارایی بهتری داشته‌اند. نورانی و کوماسی (19) از مدل تلفیقی شبکه عصبی-موجکی برای پیش‌بینی دبی رواناب یک روز بعد، استفاده نموده و به این نتیجه رسیدند که توانایی مدل شبکه عصبی-موجکی در پیش‌بینی نقاط بیشینه نمودار سری زمانی، نسبت به سایر روش‌ها بیشتر است. ادموسکی و سون (2) از مدل ترکیبی تبدیلات موجکی گسسته و شبکه‌های عصبی، برای پیش‌بینی جریان یک و سه روزه در رودخانه‌های غیردائمی حوضه‌های نیمه خشک کارگوتیس استفاده کردند؛ نتایج تحقیق نشان داده که شبکه‌های عصبی-موجکی روش‌مؤثرتری برای پیش‌بینی جریان کوتاه مدت در این‌گونه حوضه‌ها می‌باشند. رجایی و همکاران (22) از مدل شبکه عصبی موجکی برای پیش‌بینی بار معلق روزانه رودخانه Iowa آمریکا استفاده کردند. محققین در این تحقیق توانایی بالای ترکیب موجک و شبکه عصبی را نسبت به دیگر روش‌ها (مدل‌های رگرسیون، منحنی‌های سنج) نتیجه‌گیری کردند. رجایی (23) مدل ترکیبی جدیدی را برای پیش‌بینی بار معلق روزانه رودخانه‌ها ارائه داد و توانایی بالای مدل ترکیبی را نسبت به دیگر روش‌ها نتیجه‌گیری کرد. سینگ (26) از مدل تلفیقی شبکه عصبی و روش‌های آماری برای مدل‌سازی سیلاب استفاده کرد و نشان داد که قابلیت مدل ترکیبی نسبت به روش‌های آماری بالاتر است.

این مطالعه به لحاظ تکنیک ترکیب شبکه عصبی اتورگرسیو غیرخطی با موجک، برای پیش‌بینی سری زمانی یون‌های نیترات و کلراید ایستگاه ورودی آبگیر بیلقان واقع در رودخانه کرج تحقیقی جدید محسوب می‌شود. هدف از تحقیق حاضر، بکارگیری شبکه‌های عصبی مصنوعی، شبکه‌های عصبی-موجکی و رگرسیون خطی چند متغیره، برای برآورد یون‌های نیترات و کلراید، مقایسه عملکرد و کارایی این سه روش در شبیه‌سازی پارامترهای کیفی مذکور رودخانه کرج می‌باشد.

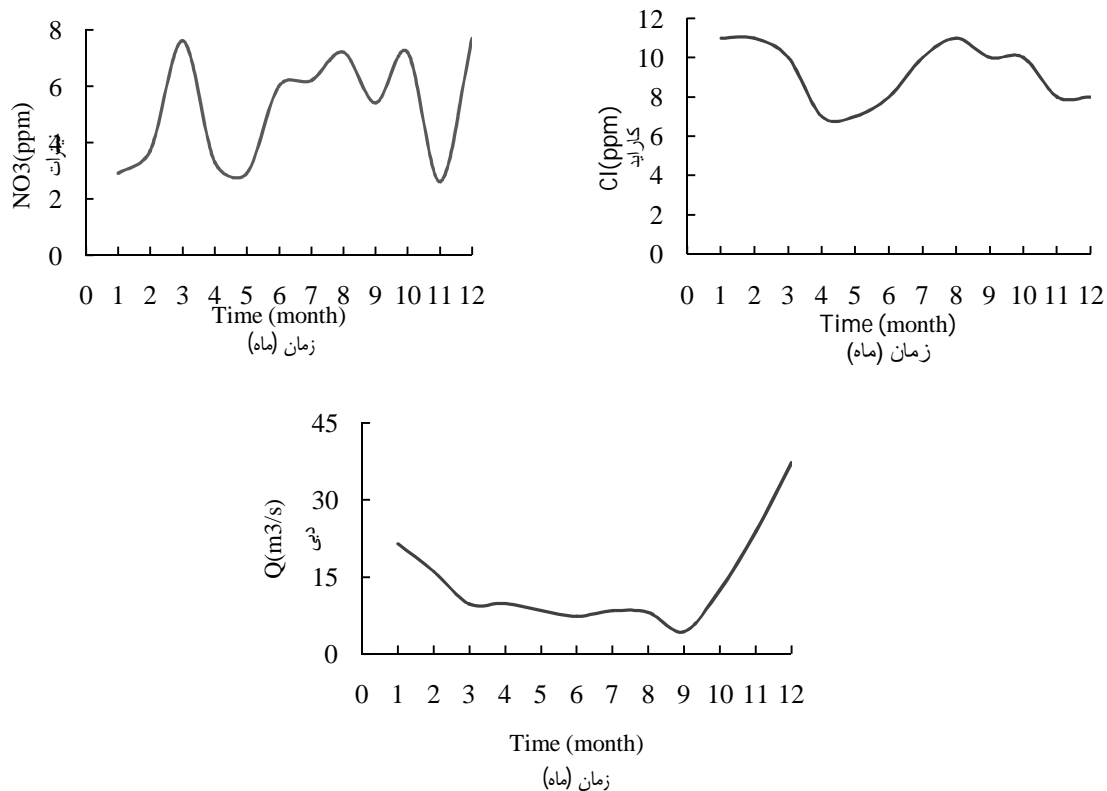
مدیریت کیفی و سیاست‌گذاری‌های زیست محیطی منابع آب دارد. به لحاظ غیر خطی بودن برخی فرآیندهای فیزیکی و عدم قطعیت در مشخصه‌های تأثیر گذار می‌توان از تکنیک مدل‌سازی جهت شبیه‌سازی و پیش‌بینی فرآیندهای غیرخطی و غیر قطعی، استفاده کرد (17 و 24). در همین راستا به‌علت ناتوانی روش‌های تحلیلی و همچنین پیچیدگی ذاتی مسائل هیدرولوژیکی، استفاده از روش‌های هوشمند جهت شبیه‌سازی و پیش‌بینی پدیده‌ها، در سال‌های اخیر بسیار متداول شده که در این بین مدل‌های شبکه عصبی مصنوعی به دلیل قابلیت یادگیری ارتباطات پیچیده و مدل‌سازی غیرخطی بسیار کاربردی می‌باشند (12). در این راستا پالانی و همکاران (21)، پارامترهای کیفی آب‌های ساحلی سنگاپور را با استفاده از شبکه عصبی مدل کردند. در این مدل پارامترهای شوری، دما، اکسیژن محلول و کلروفیل a، به‌عنوان داده‌های ورودی در نظر گرفته شد. می و سیواکومار (14) شبکه عصبی و مدل‌های رگرسیونی را برای پیش‌بینی کیفیت آب رودخانه‌های شهری مورد استفاده قرار دادند. نتایج تحقیق آن‌ها نشان داد که مدل رگرسیونی نسبت به مدل شبکه عصبی دقت بیشتری دارد. سینگو همکاران (25) مدل شبکه عصبی را برای تخمین میزان غلظت اکسیژن محلول (DO)، اکسیژن خواهی بیوشیمیایی (BOD)، رودخانه گومتی در هند توسعه دادند. در این تحقیق، از 11 پارامتر کیفی آب که طی 10 سال به‌طور ماهانه در 8 ایستگاه مختلف اندازه‌گیری شده بود به‌عنوان ورودی مدل استفاده شده است. نجاه و همکاران (15) کیفیت آب رودخانه جوهر و ون و همکاران (29) اکسیژن محلول رودخانه Heihe چین را در تحقیق جداگانه‌ای، با استفاده از شبکه عصبی پیش‌بینی نمودند. نتایج آن‌ها حاکی از دقت و توانایی بالای مدل برای پیش‌بینی و شبیه‌سازی این پارامترها بود. نجاه و همکاران (16) تکنیک‌های هوشمند عصبی مختلفی را برای پیش‌بینی پارامترهای کیفی DO، BOD و اکسیژن خواهی شیمیایی (COD) به کار بردند. هدف آن‌ها ارائه مدلی برای پیش‌بینی پارامترهای کیفی با هدف صرفه جویی در هزینه و زمان لازم برای اندازه‌گیری این پارامترها بود. محققین قابلیت بالای مدل شبکه عصبی پرسپترون با یک لایه مخفی را نسبت به رگرسیون خطی چند متغیره نتیجه گرفتند.

یکی از روش‌هایی که طی دهه‌های اخیر در مسائل آب و هیدرولوژیکی مورد توجه قرار گرفته است، استفاده از تئوری موجک<sup>1</sup>، به‌عنوان روشی نوین با قابلیت بالا، در زمینه تحلیل نوسانات، تناوب و سری‌های زمانی ناپایست و گذرا است. در این راستا طوفانی و همکاران (28) از توانایی موجک در تحلیل سری‌های زمانی غیرخطی برای پیش‌بینی بارندگی ایستگاه زرینگل در استان گلستان بهره گرفتند. مطالعات انجام شده یافته نشان داده که، شبکه عصبی و آنالیز موجک



شکل 1- منطقه مورد مطالعه

Figure 1- Location of the study area



شکل 2- سری‌های زمانی نیترات (NO<sub>3</sub>)، کلراید (CL) و دبی (Q) در دوره آزمون

Figure 2- Time series of Nitrate (NO<sub>3</sub>), Chloride (CL) and discharge (Q) for testing period

## مواد و روش‌ها

### منطقه مورد مطالعه

رودخانه کرج، یکی از مهم‌ترین رودخانه‌های حوضه آبریزی مرکزی بوده که در سمت شمال شرقی این حوضه واقع شده است؛ از ضلع جنوبی کوهستان البرز و 40 کیلومتری غرب تهران، از کانون آبیگر خرسنگ کوه، سرچشمه می‌گیرد.

طول رودخانه کرج، حدود 245 کیلومتر و مساحت حوضه آبریز آن، حدود 5000 کیلومتر مربع برآورد شده است. رودخانه کرج به‌عنوان یکی از مهم‌ترین منابع تأمین کننده آب مورد نیاز شهر تهران می‌باشد و کنترل کیفیت آب آن، از اهمیت بسزایی برخوردار است. شکل 1 موقعیت ایستگاه ورودی آبیگر بیلقان را به عرض جغرافیایی "35°49'46.8" و طول جغرافیایی "51°2'18.63" و ارتفاع 1812 به‌عنوان ایستگاه مورد مطالعه، نشان می‌دهد.

### داده‌ها و تحلیل آماری

آمار مورد استفاده در این تحقیق، غلظت یون‌های نیترات، کلراید و میزان دبی در یک دوره ده ساله را می‌باشد که این داده‌ها به‌صورت ماهانه از خرداد 1379 تا اردیبهشت 1389 توسط شرکت آب و فاضلاب استان تهران به ثبت رسیده است. نمودار تغییرات سری زمانی دبی (Q) و غلظت یون‌های نیترات (NO<sub>3</sub><sup>-</sup>) و کلراید (CL) ایستگاه مورد مطالعه در طول دوره آماری دوره آزمون در شکل 2 نشان داده شده است. پس از بررسی داده‌ها، دوره‌های مربوط به آموزش و آزمون مدل‌ها انتخاب گردیده است. در جدول 2 تحلیل آماری هر سه پارامتر ایستگاه، برای مجموعه‌های آموزشی، آزمون و کل داده‌ها ارائه شده است.

هر سری زمانی ویژگی‌های رفتاری نهفته‌ای همچون خودهمبستگی<sup>1</sup>، خودشبهی، خودنسبتی<sup>2</sup> و روند در بازه‌ی تغییرات مقیاسی خود دارد که به این وسیله می‌توان بر تاریخچه فرآیند آن اشراف پیدا کرده و مقدمات پیش‌بینی آن را در آینده فراهم ساخت. این جدول شامل مقادیر بیشینه، کمینه، میانگین، انحراف معیار (S<sub>d</sub>)، ضریب چولگی (C<sub>s</sub>)، ضریب خودهمبستگی یک ماهه (R<sub>1</sub>)، ضریب خودهمبستگی دو ماهه (R<sub>2</sub>)، ضریب خودهمبستگی سه ماهه (R<sub>3</sub>) و ضریب خودهمبستگی چهار ماهه (R<sub>4</sub>) می‌باشد. مدل‌های شبکه عصبی، زمانی بهتر عمل می‌کنند که خصوصیات آماری داده‌های آموزش و آزمون، تا حد زیادی مشابهت داشته باشند (23). طبق جدول 2، ضریب چولگی نسبتاً پایین است؛ که این مسئله برای

مدل‌سازی مناسب می‌باشد. زیرا ضریب چولگی بالا اثرات منفی قابل ملاحظه‌ای بر عملکرد شبکه‌های عصبی دارد (23). مطابق این جدول، ضریب همبستگی برای تأخیرهای 1 و 2 ماهه دوره آموزش و تأخیر 4 ماهه دوره آزمون نیترات، نسبتاً بالاست؛ برای دبی و کلر تأخیر یک ماهه مجموعه آموزش و آزمون و تأخیرهای 2 و 3 ماهه دوره آموزش، همبستگی بالایی را نشان می‌دهد. برای بدست آوردن ترکیبات ورودی مناسب برای مدل‌های ANN، MLR و WANN ضریب همبستگی<sup>4</sup> بین سری‌های زمانی نیترات و دبی و سری‌های زمانی کلراید و دبی محاسبه شد. نتایج آن در جدول 2 ارائه شده است. ضریب همبستگی (R)، بین دبی (Q) و پارامتر مورد نظر (NO<sub>3</sub>، CL) برای n جفت در دسترس از رابطه 1 بدست آمد.

$$R = \frac{\sum_{i=1}^n (Q_i - \bar{Q})(N_i - \bar{N})}{\sqrt{\sum_{i=1}^n (Q_i - \bar{Q})^2 \sum_{i=1}^n (N_i - \bar{N})^2}} \quad (1)$$

این ضریب در محدوده 0 تا 1 تعریف می‌شود که هر چه به 1 نزدیک‌تر باشد نشان‌دهنده همبستگی بالای دو پارامتر مورد ارزیابی است. جهت دستیابی به ترکیبات ورودی مناسب بایستی خودهمبستگی و همبستگی متقابل پارامترها مورد ارزیابی قرار گیرد (6). مطابق جدول 2 ضریب همبستگی NO<sub>3</sub>(t) با Q(t-4) و CI(t) با Q(t-4) نسبت به دیگر ضرایب عدد بالایی را نشان می‌دهد؛ به‌عبارت دیگر هر یک از پارامترهای نیترات و کلراید با دبی 4 ماه گذشته همبستگی بالایی دارد.

استفاده از یک روش خطی برای پیش‌پردازش داده‌ها روشی معمول است که باعث بهبود عملکرد آموزش شبکه می‌شود (9). در این پژوهش، جهت حذف بعد پارامترها، داده‌ها با استفاده از رابطه 2 عادی‌سازی<sup>5</sup> شده است (در این صورت، مقادیر بین 0 تا 1 قرار می‌گیرند). در این رابطه N<sub>normal</sub> معرف داده عادی شده، N داده واقعی، N<sub>max</sub> و N<sub>min</sub> به ترتیب، مقادیر کمینه و بیشینه سری زمانی است.

$$N_{normal} = \frac{(N - N_{min})}{(N_{max} - N_{min})} \quad (2)$$

### شبکه‌های عصبی مصنوعی (ANN)

شبکه عصبی مصنوعی<sup>6</sup>، الهام گرفته از فرآیند یادگیری مغز انسان انسان است که شامل مجموعه‌ای از واحدهای عصبی به نام نرون<sup>7</sup> می‌باشد که توسط ارتباطاتی موسوم به آکسون<sup>8</sup>، به هم متصل هستند.

4- Correlations Coefficient  
5- Normalize  
6- Artificial Neural Network  
7- Neuron  
8-Axon

1- Autocorrelations Coefficient  
2-Self Similarity  
3-Self Affinity

جدول 1- تحلیل آماری داده‌های کل، آموزش و آزمون  
Table1- Statistics analysis for train, test and all data sets

| آمارهای پارامتر<br>(Statistical Parameters) | NO3(ppm)         |                  |                 | Q (m3/s)         |                  |                 | Cl(ppm)          |                  |                 |
|---|------------------|------------------|-----------------|------------------|------------------|-----------------|------------------|------------------|-----------------|
|   | کل داده<br>(All) | آموزش<br>(Train) | آزمون<br>(Test) | کل داده<br>(All) | آموزش<br>(Train) | آزمون<br>(Test) | کل داده<br>(All) | آموزش<br>(Train) | آزمون<br>(Test) |
| (n) داده تعداد                              | 120              | 108              | 12              | 120              | 108              | 12              | 120              | 108              | 12              |
| Mean  | 3.65             | 3.47             | 5.23            | 14.96            | 15.08            | 13.92           | 10.15            | 10.25            | 9.25            |
| Max   | 7.70             | 7.40             | 7.70            | 70.13            | 70.13            | 37.33           | 23.52            | 23.52            | 11.0            |
| Min   | 1.70             | 1.70             | 2.60            | 1.849            | 1.849            | 4.309           | 3.86             | 3.86             | 7.0             |
| Sd  | 1.345            | 1.133            | 2.020           | 10.332           | 10.464           | 9.407           | 3.346            | 3.479            | 1.545           |
| Cs  | 1.475            | 1.597            | -0.098          | 2.140            | 2.182            | 1.408           | 1.101            | 1.023            | -0.280          |
| R1  | 0.424            | 0.539            | -0.253          | 0.690            | 0.696            | 0.722           | 0.714            | 0.715            | 0.575           |
| R2  | 0.356            | 0.342            | -0.056          | 0.341            | 0.363            | -0.047          | 0.530            | 0.540            | -0.218          |

**تحلیل رگرسیون خطی چند متغیره (MLR)**

تحلیل رگرسیون خطی چند متغیره<sup>3</sup>، روشی در پیش‌بینی‌های سری زمانی، برای مدل کردن رابطه خطی بین یک متغیر وابسته و یک یا چند متغیر مستقل است. شکل معمول معادله به صورت زیر نشان داده می‌شود:

$$y = a_0 + \sum a_i x_i \quad (3)$$

که در آن y مقدار پیش‌بینی شده، a<sub>0</sub> ثابت رگرسیون، a<sub>i</sub> ضریب رگرسیون متغیرهای مستقل x<sub>i</sub> است. هدف این روش، تعیین پارامترهای معادله رگرسیون است به گونه‌ای که مقادیر تخمین زده شده کارا و سازگار باشند (23، 26).

**تبدیل موجک**

تبدیل موجک، تابعی است که نسخه بهبود یافته تبدیل فوریه را تداعی می‌نماید. تحلیل فوریه ابزاری ریاضی است که سیگنال را از حال زمانی به حالت فرکانسی تبدیل می‌کند. در حالی که تبدیل‌های موجک یک بیان زمان-مقیاس از سری‌های زمانی (سیگنال) ارائه می‌دهند. روش متداول برای کاهش نویز و تجزیه سیگنال به فرکانس‌های سازنده آن تبدیل فوریه است که در آنالیز و پردازش سیگنال‌ها کاربرد دارد اما این تبدیل دارای محدودیت‌هایی است (7). نقص و مشکل جدی که تحلیل فوریه در انتقال به حوضه فرکانسی دارد این است که اطلاعات زمانی سیگنال گم می‌شود. در صورتی که موجک‌ها، توابع ریاضی هستند که یک بیان زمان-مقیاس از سری‌های زمانی و نسبت‌های آن ارائه می‌دهند؛ این بیان برای تحلیل‌های سری زمانی نایست<sup>4</sup> مفید است (1 و 23). به وسیله این

**جدول 2- ضریب همبستگی بین مقادیر مشاهده‌ای Q با NO<sub>3</sub> و CL**

Table 2- Correlation coefficients observed NO<sub>3</sub> and CL with Q

| سری زمانی<br>(Time Series) | آزمون (Test) | آموزش (Train) | کل داده (All) |
|----------------------------|--------------|---------------|---------------|
| NO3(t) Q(t)                | -0.024       | 0.122         | 0.077         |
| Q(t-1)                     | -0.296       | 0.242         | 0.144         |
| Q(t-2)                     | 0.354        | 0.238         | 0.201         |
| Q(t-3)                     | 0.254        | 0.158         | 0.120         |
| Q(t-4)                     | 0.629        | 0.078         | 0.094         |
| Cl(t) Q(t)                 | -0.131       | -0.534        | -0.508        |
| Q(t-1)                     | -0.107       | -0.412        | -0.391        |
| Q(t-2)                     | 0.236        | -0.218        | -0.190        |
| Q(t-3)                     | 0.360        | -0.051        | -0.026        |
| Q(t-4)                     | 0.432        | 0.041         | 0.058         |

در این مدل سعی بر این است، ساختاری مشابه ساختار بیولوژیکی مغز انسان و شبکه عصبی بدن، طراحی شود و با معرفی تاریخچه عملکرد سیستم دینامیکی، مدل آموزش<sup>1</sup> یافته؛ نحوه عملکرد سیستم در حافظه ذخیره شود؛ تا در مواردی که مدل قبلاً با آن مواجه نشده است، مورد استفاده قرار گیرد (20). در واقع شبکه عصبی یک مدل ریاضی است که توانایی مدل‌سازی و ایجاد روابط ریاضی غیرخطی برای درون‌یابی را دارد (4). برای آموزش شبکه الگوریتم‌های مختلفی وجود دارد؛ بر اساس تحقیقات موجود، الگوریتم لونیبرگ-مارکوات<sup>2</sup> به دلیل سرعت و دقت و اطمینان بالای آن در پیش‌بینی‌های هیدرولوژیکی، الگوریتم مناسبی است (9).

3-Multi Linear Regression

4-Nonstationary Time Series

1-Train

2-Levenberg-Marquardt

عبارتی  $S = a_4 + d_4 + d_3 + d_2 + d_1$

### کارایی مدل‌ها

در این تحقیق، کارایی مدل‌های ANN، MLR و WANN با استفاده از ضریب نش ساتکلیف<sup>3</sup> (E) و ریشه میانگین مربعات خطا (RMSE) ارزیابی می‌شود. نش و ساتکلیف ضریب کارایی مدل را که برای سنجش توان پیش‌بینی هیدرولوژیکی به کار می‌رود؛ در محدوده 1 تا منفی بی‌نهایت به صورت رابطه 7 تعریف کردند و رابطه 8 معرف RMSE (ریشه میانگین مربعات خطا) بین داده‌های مشاهده شده و پیش‌بینی شده می‌باشد.

$$E = 1 - \frac{\sum_{i=1}^n (N_{i(o)} - N_{i(p)})^2}{\sum_{i=1}^n (N_{i(o)} - N_m)^2} \quad (9)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (N_{i(o)} - N_{i(p)})^2}{n}} \quad (10)$$

در روابط بالا  $N_{i(o)}$ ،  $N_{i(p)}$ ،  $N_m$  به ترتیب نمایانگر مقادیر مشاهده شده، پیش‌بینی شده و میانگین غلظت هریک از یون‌های نیترات و کلراید و  $n$  تعداد داده‌ها است. اگر  $E$  و  $RMSE$  به ترتیب نزدیک به 1 و 0 شود؛ عملکرد مدل قابل قبول خواهد بود.

### توسعه مدل‌ها

شبکه استفاده شده در این تحقیق، شبکه عصبی NARX<sup>4</sup> است. این مدل، برای پیش‌بینی مقادیر آینده  $N(t)$ ، از مقادیر گذشته همان سری زمانی به صورت پس‌خور<sup>5</sup> و مقادیر گذشته سری زمانی دوم  $Q(t)$ ، که به عنوان ورودی به مدل داده می‌شود، توسعه می‌یابد. در واقع مقادیر آینده سری زمانی پارامتر مورد نظر (نیترات و کلراید) به صورت تابعی از مقادیر گذشته همان سری زمانی و مقادیر گذشته سری زمانی دیگر، پیش‌بینی می‌شوند. به عبارت دیگر:

$$N_t = f(N_{t-1}, N_{t-2}, \dots, N_{t-d}, Q_{t-1}, Q_{t-2}, \dots, Q_{t-d}) \quad (11)$$

در رابطه 11،  $N(t)$  سری زمانی پارامتر مورد نظر برای پیش‌بینی (نیترات و کلراید)،  $Q(t)$  سری زمانی دبی و  $d$  تعداد تأخیر است. بدین ترتیب، شبکه عصبی NARX با حلقه پس‌خور باز<sup>6</sup> (استفاده از مقادیر واقعی برای پیش‌بینی) با سه لایه پیش‌خور، با تابع فعالیت سیگموئید در لایه مخفی و تابع فعالیت خطی در لایه خروجی، توسعه داده شد.

تبدیل‌ها، سری‌های زمانی به موجک‌هایی تجزیه می‌شوند که هریک نسخه‌ای جایجا شده و مقیاس شده از موجک مادر می‌باشند. تابع موجک به صورت زیر تعریف می‌شود:

$$\int_{-\infty}^{+\infty} \psi(t) dt = 0 \quad (4)$$

$\psi(t)$  موجک مادری است که از سری‌های موجک  $\Psi_{a,b}(t)$  با تعریف زیر بدست می‌آید:

$$\Psi_{a,b}(t) = |a|^{-1/2} \psi\left(\frac{t-b}{a}\right) \quad a, b \in \mathbb{R}; a \neq 0 \quad (5)$$

در رابطه 5،  $a$  فاکتور مقیاس،  $b$  فاکتور انتقال و  $\mathbb{R}$  دامنه اعداد حقیقی است (13). تبدیل موجک در دونوع پیوسته و گسسته در زمان به کار می‌رود. اختلاف انواع مختلف تبدیل موجک، مربوط به روشی است که مقیاس و انتقال را پیاده‌سازی می‌کند. برای یک سری زمانی با تابع  $f(t)$  تبدیل موجک پیوسته آن به صورت زیر است (13).

$$W_{a,b} = |a|^{-1/2} \int_{-\infty}^{+\infty} f(t) \bar{\Psi}\left(\frac{t-b}{a}\right) dt \quad (6)$$

$$f(t) = \frac{1}{W_{\Psi}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} W_{a,b} \psi_{a,b}(t) \frac{da db}{a^2} \quad (7)$$

تبدیل موجک پیوسته در تمام مقیاس‌ها می‌تواند عمل کند، ولی در عین حال به مقدار زیادی داده و زمان محاسباتی نیاز دارد. همچنین برای بدست آوردن تبدیل موجک گسسته با جایگذاری مقیاس‌ادیر  $a = kb_0 a_0^j$ ،  $b \in \mathbb{R}$ ،  $a_0 > 1$ ،  $b = kb_0 a_0^j$ ،  $K, j \in \mathbb{Z}$ ،  $a_0$  در رابطه 8 حاصل می‌شود:

$$DWT(j, k) = W_{\Psi} f(j, k) = a_0^{-j/2} \int_{-\infty}^{+\infty} f(t) \bar{\Psi}(a_0^{-j} t - kb_0) dt \quad (8)$$

در اکثر تحقیقات از تبدیل موجک گسسته، به دلیل سادگی و سرعت بالای محاسبات با توجه به دقت مورد نیاز، استفاده شده است. در تبدیل موجک، سیگنال از یک سری از فیلترهای بالا گذر و پایین گذر عبور داده می‌شود. بخش حاصل از عبور سیگنال از فیلتر بالا گذر که شامل اطلاعات فرکانس بالا از جمله نویز می‌باشد، جزئیات<sup>1</sup> نام دارد. بخش حاصل از عبور سیگنال از فیلتر پایین‌گذر که شامل اطلاعات فرکانس پایین و تغییرات آرام را نشان می‌دهد، تقریب<sup>2</sup> نامیده می‌شود. در این تحقیق نیز تجزیه سیگنال سری زمانی غلظت نیترات، کلراید و دبی به سیگنال‌های پایه سازنده آن‌ها، توسط موجک مادر با سطوح تجزیه مختلف انجام می‌شود. به طوری که برای مثال موجک db2 با سطح تجزیه 4، موجک مادر (S) را به یک موجک تقریب (a4) و چهار موجک جزئیات ( $d_4, d_3, d_2, d_1$ ) تجزیه می‌کند به

3-Nash-Sutcliffe

4-Nonlinear Auto Regressive with eXogenous input

5-Feedback

6-open-loop

1- Details

2- Approximation

ترکیبات ورودی به‌عنوان متغیر مستقل لحاظ شدند.

### نتایج و بحث

برای بدست آوردن بهترین نتایج، ترکیبات گوناگون مدل شبکه عصبی بر اساس تعداد لایه‌های مخفی، تعداد نرون‌های لایه مخفی آموزش داده شده و موردآزمون قرار گرفت. تعداد لایه‌های مخفی و تعداد نرون‌های لایه مخفی هر شبکه، با استفاده از داده‌های موجود و روش سعی و خطا بهینه شد. باید توجه داشت که تعداد لایه‌های مخفی بیشتر و تعداد نرون بیشتر در لایه مخفی، اگرچه ممکن است دقت محاسبات در آموزش شبکه را بالا ببرد؛ ولی الزاماً دقت مدل را در پیش‌بینی داده‌های اعتبارسنجی و آزمون افزایش نخواهد داد (23). بنابراین از بین دو معماری شبکه با دقت برابر، موردی که سادگی بیشتری داشته باشد ارجح است. بهترین عملکرد، مدل زمانی است که خطای هر سه مرحله آموزش، اعتبارسنجی و آزمون را حداقل گرداند؛ چون در صورت بروز یادگیری اضافی خطای مجموعه آموزشی حداقل می‌شود؛ اما برای مرحله آزمون افزایش می‌یابد؛ لذا توجه به خطای هر سه مرحله امری ضروری است.

نتایج پیش‌بینی یک ماه آینده مدل شبکه عصبی و رگرسیون خطی چند متغیره، برای نیترات و کلراید به‌ترتیب در جدول 3 و جدول 4 ارائه شده است. طبق جداول بهترین نتیجه در مدل شبکه عصبی برای پارامتر کیفی نیترات، ترکیب ورودی 16 (N<sub>t</sub>, Q<sub>t</sub>, Q<sub>t-1</sub>, Q<sub>t-2</sub>) و 6 نرون در لایه مخفی و مقادیر E و RMSE به‌ترتیب 0/727 و 1/009 پی‌پی‌ام؛ و برای پارامتر کلر، ترکیب 15 (N<sub>t</sub>, N<sub>t-1</sub>, Q<sub>t</sub>, Q<sub>t-1</sub>) و 7 نرون در لایه مخفی و مقادیر E و RMSE 0/704 و 0/805 پی‌پی‌ام حاصل شد. برای مدل رگرسیون خطی چند متغیره، آنالیز رگرسیون خطی بین پارامتر کیفی موردنظر (نیترات یا کلراید) در ماه فعلی و ترکیبات 16 گانه صورت گرفت؛ سپس با یافتن ثابت‌های رگرسیونی، پیش‌بینی یک ماه آینده پارامتر کیفی در دوره آموزش، انجام شد. مطابق جدول 3 و جدول 4 پیش‌بینی نیترات با روش MLR غیر قابل قبول بوده در صورتی که پیش‌بینی کلراید با ضریب نش ساتکلیف و RMSE به‌ترتیب 0/239 و 0/289 ppm با ترکیب ورودی 6 (N<sub>t</sub>, N<sub>t-3</sub>) همراه بود. روش MLR تکنیکی خطی جهت پیش‌بینی سری‌های زمانی محسوب می‌شود؛ چون یون نیترات ماهیت نوسانی شدیدی در کل دوره آماری داشته؛ لذا این روش در پیش‌بینی این پارامتر ناتوان بوده است. به‌طور کلی نتایج حاکی از دقت بالای مدل شبکه عصبی در مقایسه با رگرسیون خطی چند متغیره در پیش‌بینی دو پارامتر مذکور می‌باشد.

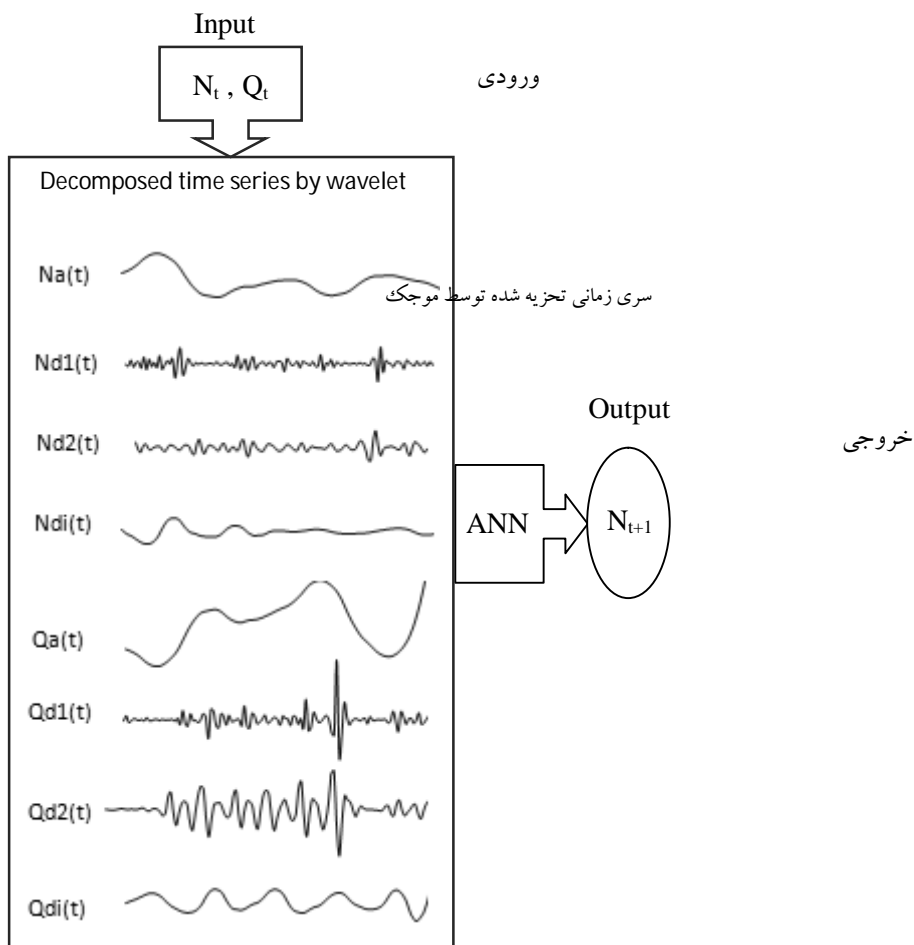
هدف از بکارگیری مدل حلقه باز، لحاظ نمودن داده‌های مشاهداتی که برداشت‌های کاملاً واقعی از پدیده هستند، به‌جای داده‌های محاسبه شده توسط مدل که ناشی از عملکرد محاسباتی مدل هستند، می‌باشد.

با استفاده از نتایج تحلیل آماری جدول 1 و جدول 2، ترکیبات زیر که شامل مقادیر مختلف پارامتر موردنظر و دبی هستند؛ به‌عنوان بردار ورودی جهت پیش‌بینی یک ماه آینده نیترات یا کلراید در گام زمانی (N<sub>t+1</sub>) در نظر گرفته شدند. در ترکیبات زیر نمایانگر نیترات یا کلراید و Q معرف دبی است.

$$\begin{aligned}
 & N_t \\
 & N_t, N_{t-1} \\
 & N_t, N_{t-1}, N_{t-2} \\
 & N_t, N_{t-1}, N_{t-2}, N_{t-3} \\
 & N_t, N_{t-2} \\
 & N_t, N_{t-3} \\
 & N_t, N_{t-2}, N_{t-3} \\
 & N_t, N_{t-1}, N_{t-3} \\
 & N_t, Q_t \\
 & N_t, N_{t-1}, Q_t \\
 & N_t, N_{t-1}, Q_t, Q_{t-1} \\
 & N_t, Q_t, Q_{t-1} \\
 & N_t, N_{t-1}, N_{t-2}, Q_t \\
 & N_t, N_{t-1}, N_{t-2}, Q_t, Q_{t-1} \\
 & N_t, N_{t-1}, Q_t, Q_{t-1}, Q_{t-2} \\
 & N_t, Q_t, Q_{t-1}, Q_{t-2}
 \end{aligned}$$

در دو مدل جداگانه برای پیش‌بینی یک ماه آینده یون‌های نیترات و کلراید از آمار 10 ساله مربوط به ایستگاه ورودی آبیگر بیلقان استفاده شد. بدین ترتیب 8 سال ابتدایی داده‌ها، از خرداد 1379 تا اردیبهشت 1387 برای آموزش شبکه، یک سال بعدی، از خرداد 1387 تا اردیبهشت 1388، برای اعتبارسنجی و یک سال انتهایی دوره از خرداد 1388 تا اردیبهشت 1389، برای آزمون شبکه در نظر گرفته شد.

در اجرای مدل WANN سری زمانی Q و پارامتر مورد نظر (نیترات یا کلراید)، با استفاده از تبدیل موجک تجزیه شده؛ به‌عنوان ورودی شبکه عصبی، برای پیش‌بینی یک ماه آینده پارامتر، به کار گرفته می‌شود. ساختار پیشنهادی مدل ترکیبی WANN در شکل 3 نمایش داده شده است. تعداد گره‌ها در لایه ورودی از رابطه  $(i+1)*2$  پیروی می‌کند؛ زیرا در مدل ترکیبی WANN دو متغیر ورودی (دبی و پارامتر نیترات یا کلراید)، دخیل هستند؛ به طوری که هر سری زمانی در سطح تجزیه  $i$ ، به تعداد  $i$  زیرسری جزئیات و یک زیرسری تقریب تجزیه می‌شود (23). تمامی عملیات کدنویسی مدل شبکه عصبی مصنوعی و هیبرید شبکه عصبی موجکی در محیط نرم‌افزار Matlab2013 انجام گرفت. در اجرای مدل رگرسیون خطی چند متغیره در محیط نرم‌افزار IBM SPSS Statistics، پارامتر مورد نظر برای پیش‌بینی در ماه فعلی، به‌عنوان متغیر وابسته و هریک از



شکل 3- ساختار مدل ترکیبی شبکه عصبی موجک

Figure3- Structure of the proposed WANN combination model

جدول 3- نتایج پیش‌بینی نیترات مدل ANN و MLR در دوره آزمون

Table3- The results prediction Nitrate by ANN and MLR models in testing period

|                                 |                  | NO3  |      |      |       |       |       |       |      |       |       |       |       |       |       |       |       |
|---------------------------------|------------------|------|------|------|-------|-------|-------|-------|------|-------|-------|-------|-------|-------|-------|-------|-------|
| ترکیبات ورودی<br>(Combinations) |                  | 1    | 2    | 3    | 4     | 5     | 6     | 7     | 8    | 9     | 10    | 11    | 12    | 13    | 14    | 15    | 16    |
| ANN                             | نرون<br>(Neuron) | 4    | 3    | 4    | 6     | 5     | 7     | 3     | 3    | 7     | 6     | 6     | 8     | 8     | 7     | 4     | 6     |
|                                 | E                | 0.62 | 0.61 | 0.56 | 0.58  | 0.54  | 0.49  | 0.54  | 0.52 | 0.68  | 0.6   | 0.561 | 0.661 | 0.594 | 0.684 | 0.498 | 0.73  |
|                                 | RMSE<br>(ppm)    | 1.19 | 1.2  | 1.28 | 1.26  | 1.31  | 1.38  | 1.32  | 1.33 | 1.09  | 1.23  | 1.28  | 1.26  | 1.23  | 1.09  | 1.37  | 1.01  |
| MLR                             | E                | -0.6 | -0.8 | -0.8 | -0.79 | -0.85 | -0.87 | -0.87 | -0.8 | -0.73 | -0.72 | -0.71 | -0.78 | -0.78 | -0.75 | -0.68 | -0.77 |
|                                 | RMSE<br>(ppm)    | 2.64 | 2.56 | 2.59 | 2.59  | 2.63  | 2.64  | 2.65  | 2.61 | 2.54  | 2.53  | 2.53  | 2.58  | 2.58  | 2.56  | 2.51  | 2.58  |



جدول 4- نتایج پیش‌بینی کلراید مدل ANN و MLR در دوره آزمون

Table4- The results prediction of Chloride by ANN and MLR models in testing period

|                              |               | Cl    |       |       |       |       |       |       |       |      |      |      |      |      |      |      |      |
|------------------------------|---------------|-------|-------|-------|-------|-------|-------|-------|-------|------|------|------|------|------|------|------|------|
| ترکیبات ورودی (sCombination) |               | 16    | 15    | 14    | 13    | 12    | 11    | 10    | 9     | 8    | 7    | 6    | 5    | 4    | 3    | 2    | 1    |
| ANN                          | نرون (Neuron) | 5     | 7     | 4     | 3     | 2     | 9     | 6     | 5     | 4    | 4    | 6    | 3    | 7    | 7    | 5    | 6    |
|                              | E             | 0.65  | 0.7   | 0.51  | 0.46  | 0.66  | 0.5   | 0.54  | 0.63  | 0.62 | 0.69 | 0.63 | 0.64 | 0.56 | 0.66 | 0.68 | 0.54 |
|                              | RMSE (ppm)    | 0.88  | 0.81  | 1.03  | 1.09  | 0.86  | 1.05  | 1.01  | 0.9   | 0.91 | 0.83 | 0.9  | 0.89 | 0.98 | 0.86 | 0.84 | 1    |
| MLR                          | E             | -0.19 | -0.18 | -0.15 | -0.09 | -0.15 | -0.16 | -0.09 | -0.04 | 0.18 | 0.23 | 0.24 | 0.24 | 0.19 | 0.22 | 0.2  | 0.23 |
|                              | RMSE (ppm)    | 1.61  | 1.6   | 1.58  | 1.55  | 1.59  | 1.59  | 1.55  | 1.51  | 1.34 | 1.3  | 1.29 | 1.29 | 1.33 | 1.31 | 1.32 | 1.3  |

جدول 5- نتایج پیش‌بینی نیترات با مدل WANN در دوره آزمون

Table5- The results prediction Nitrate by WANN model in testing period

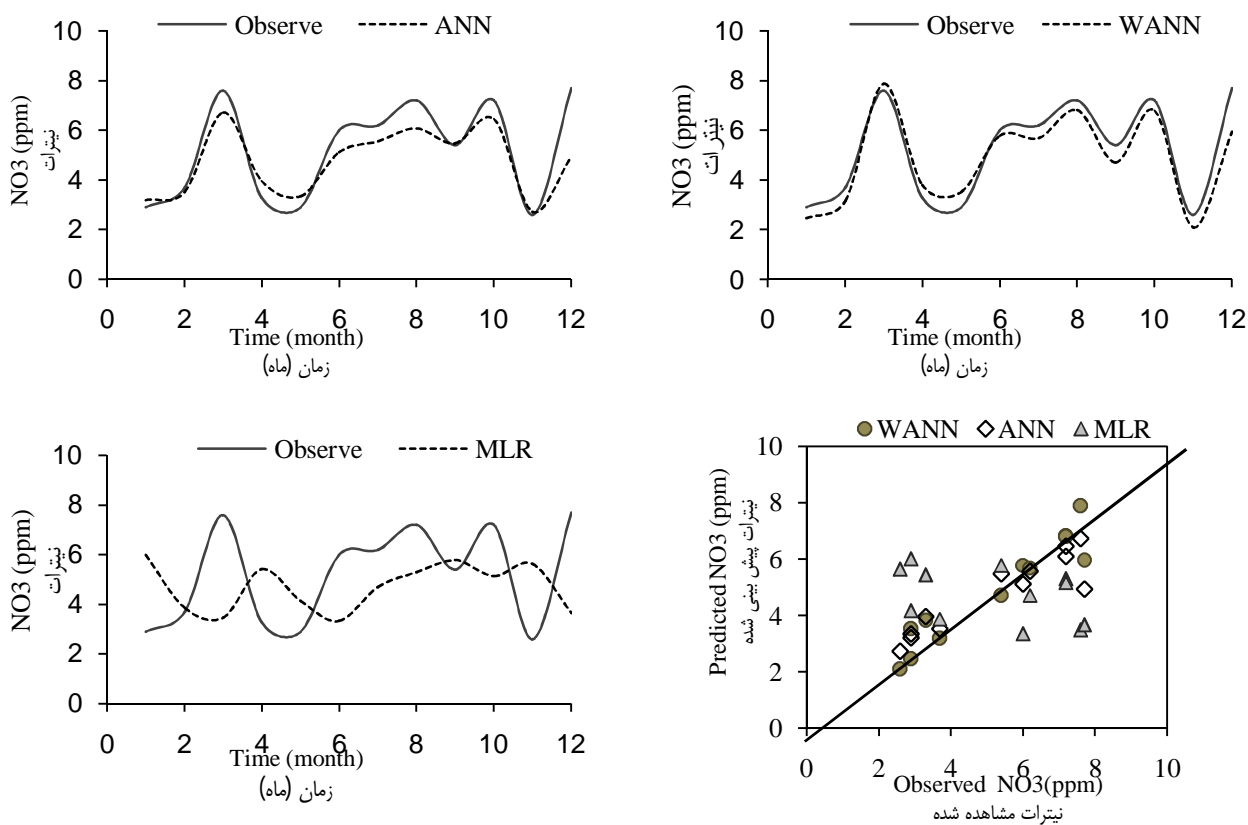
| موجک مادر (Mother wavelet) | سطح تجزیه (Decomposition level) | RMSE (ppm) | E     | ساختار (Structure) |  |
|----------------------------|---------------------------------|------------|-------|--------------------|--|
|                            |                                 |            |       | ANN                |  |
| Dmey                       | 1                               | 1.147      | 0.648 | 4- 7- 1            |  |
|                            | 2                               | 0.832      | 0.815 | 6- 6- 1            |  |
|                            | 3                               | 0.705      | 0.867 | 8- 2- 1            |  |
|                            | 4                               | 0.746      | 0.851 | 10-1- 1            |  |
| Db5                        | 1                               | 1.284      | 0.559 | 4- 5- 1            |  |
|                            | 2                               | 0.771      | 0.841 | 6- 1- 1            |  |
|                            | 3                               | 0.742      | 0.852 | 8- 1- 1            |  |
|                            | 4                               | 0.961      | 0.753 | 10-1- 1            |  |
| Sym3                       | 1                               | 1.304      | 0.545 | 4- 2- 1            |  |
|                            | 2                               | 1.216      | 0.604 | 6- 5- 1            |  |
|                            | 3                               | 0.878      | 0.794 | 8- 4- 1            |  |
|                            | 4                               | 1.332      | 0.525 | 10-2- 1            |  |
| Coif2                      | 1                               | 1.457      | 0.432 | 4- 6- 1            |  |
|                            | 2                               | 1.051      | 0.704 | 6- 3- 1            |  |
|                            | 3                               | 0.884      | 0.791 | 8- 2- 1            |  |
|                            | 4                               | 0.946      | 0.761 | 10-1- 1            |  |

به دلیل دارا بودن پریود کوتاه‌تر در پیش‌بینی یک ماه بعد  $N_{t+1}$  دارد. لازم به ذکر است که افزایش سطوح تجزیه تا یک حدی باعث بهبود عملکرد مدل می‌شود. زیرا در سطوح تجزیه بالاتر، تعداد پارامترهای ورودی به مدل (زیرسری‌هایی که هر کدام ماهیت نوسانی غیرخطی و پیچیده‌ای دارند) افزایش می‌یابد، لذا هر یک از پارامترها خطایی تولید می‌کنند که در نهایت باعث انباشته شدن خطای کل و در نتیجه کاهش عملکرد مدل می‌شود. مطابق جدول 5 و جدول 6 با افزایش سطوح از دقت و عملکرد مدل کاسته شده و بهترین نتیجه برای پیش‌بینی نیترات در سطح تجزیه 3 و موجک مادر Dmey با مقادیر E و RMSE به ترتیب 0/867 و 0/705 ppm حاصل شده است. پیش‌بینی کلراید در سطح تجزیه 2 با موجک مادر Sym3 و مقادیر E و RMSE به ترتیب 0/860 و 0/553 ppm بهترین خروجی را داشته است.

در اجرای مدل ترکیبی شبکه عصبی - موجکی ابتدا سری‌های زمانی Q و پارامتر موردنظر برای پیش‌بینی، توسط موجک‌های مادر تجزیه شدند. برای این کار از موجک‌های مادر Db5، Sym3، Coif2 و Dmey با سطوح تجزیه شده مختلف استفاده شد. زیرسری‌ها به‌عنوان ورودی شبکه عصبی مورد استفاده قرار گرفتند و پارامتر کیفی مورد نظر در گام زمانی t+1 به عنوان خروجی مدل حاصل شد. عملکرد مدل ترکیبی WANN، در جدول 5 و 6 ارائه شده است. زمانی که چند زیرسری به‌عنوان بردارهای ورودی، وارد مدل ترکیبی WANN می‌شود؛ وزن‌های اختصاص یافته در شبکه عصبی (در سطوح مختلف) متفاوت خواهد بود، به‌طوری که وزن‌های بالا به سطوح تجزیه بالا، اعمال می‌شود (23). برای مثال در تجزیه سطح 2، برای هر یک از سری‌های زمانی Q و N، 3 زیرسری تولید می‌شود که زیرسری‌های جزئیات دوم  $Qd_2(t)$ ، نسبت به  $Qd_1(t)$ ، نقش موثرتری

جدول 6- نتایج پیش‌بینی کلراید با مدل WANN در دوره آزمون  
**Table6- The results prediction Chloride by WANN model in testing period**

| موجک مادر<br>(Mother wavelet) | سطح<br>تجزیه<br>(Decomposition level) | RMSE<br>(ppm) | ساختار (structure) |         |
|-------------------------------|---------------------------------------|---------------|--------------------|---------|
|                               |                                       |               | E                  | ANN     |
| Dmey                          | 1                                     | 1.072         | 0.474              | 4- 3- 1 |
|                               | 2                                     | 0.730         | 0.756              | 6- 1- 1 |
|                               | 3                                     | 0.772         | 0.727              | 8- 2- 1 |
|                               | 4                                     | 0.881         | 0.645              | 10-4- 1 |
| Db5                           | 1                                     | 0.947         | 0.590              | 4- 2- 1 |
|                               | 2                                     | 0.681         | 0.788              | 6- 2- 1 |
|                               | 3                                     | 0.795         | 0.711              | 8- 7- 1 |
|                               | 4                                     | 0.873         | 0.652              | 10-1- 1 |
| Sym3                          | 1                                     | 1.018         | 0.526              | 4- 4- 1 |
|                               | 2                                     | 0.553         | 0.860              | 6- 1- 1 |
|                               | 3                                     | 0.779         | 0.722              | 8- 1- 1 |
|                               | 4                                     | 0.784         | 0.719              | 10-1- 1 |
| Coif2                         | 1                                     | 0.992         | 0.550              | 4- 2- 1 |
|                               | 2                                     | 0.764         | 0.733              | 6- 1- 1 |
|                               | 3                                     | 0.791         | 0.714              | 8- 1- 1 |
|                               | 4                                     | 0.909         | 0.622              | 10-4- 1 |



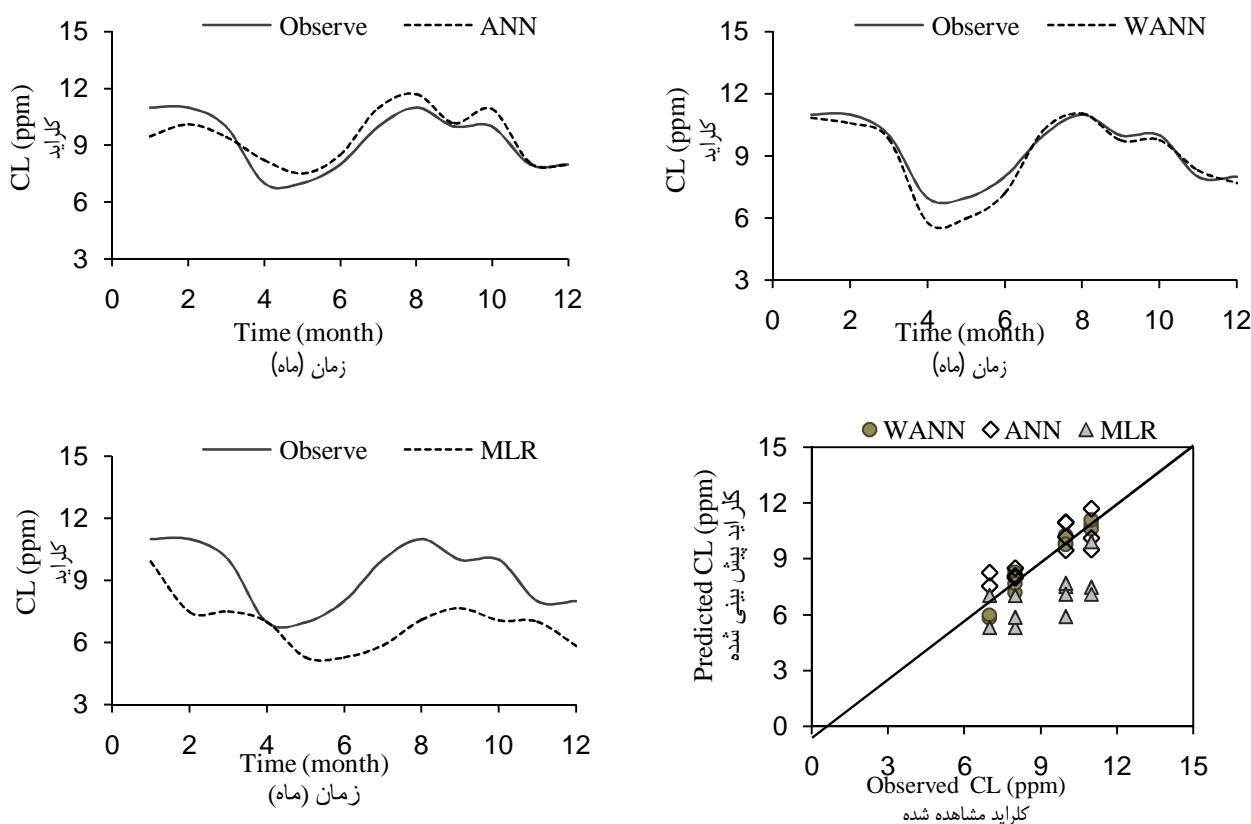
شکل 4- مقایسه نیترات مشاهده شده و پیش‌بینی شده در دوره آزمون

Figure4- Comparisons between the observed and predicted Nitrate based on the testing data

سه مدل برای دوره آزمون در مقایسه با غلظت نیترات و کلراید مشاهده شده نشان می‌دهند. به نظر می‌رسد روش MLR در پیش‌بینی کلراید، توانسته است روند کلی دوره آزمون را پیش‌بینی کند اما انطباقی با مقادیر مشاهده‌ای ندارند. به طور کلی نمودارها، برتری و قابلیت بالای مدل ترکیبی را نسبت به مدل شبکه عصبی و رگرسیون خطی به وضوح نشان می‌دهند. مقادیر پیش‌بینی شده توسط مدل ترکیبی موجک عصبی تطابق بیشتری با مقادیر مشاهده شده داشته و رابطه تقریباً 1:1 برقرار است. در واقع WANN جهت ارزیابی هوشمند مدل شبکه عصبی مصنوعی، استخراج ویژگی‌های مهم ضرایب موجک گسسته و بدست آوردن مولفه‌های موثر در پیش‌بینی پارامتر کیفی به کار گرفته می‌شود که نتایج حاکی از عملکرد مناسب مدل ترکیبی توسعه یافته می‌باشد. از ویژگی‌های یک مدل هیدرولوژیکی خوب، توانایی آن در تخمین و شبیه‌سازی نقاط پیک سری زمانی می‌باشد. زیرا بسیاری از کارهای طراحی و اصلاحات لازم در کیفیت آب‌های سطحی معمولاً براساس نقاط ماکزیمم انجام می‌شود.

نکته مهم در انتخاب نوع موجک مادر، طبیعت رخداد پدیده و نوع سری زمانی مشاهده شده است (18). سری زمانی یون نیترات و کلراید برخلاف پدیده‌های بارش - رواناب ماهیت تک قله ای ندارند بلکه اغلب دارای ویژگی نوسانی پی در پی هستند که ایستگاه این تحقیق نیز از این امر مستثنی نیستند (شکل 2). بنابراین موجک مادری که از لحاظ شکل هندسی به خوبی بر منحنی تغییرات سری زمانی منطبق شوند، بهتر می‌توانند عمل نگاشت را انجام دهند.

بر طبق جدول‌های جدول 3، جدول 4، جدول 5 و جدول 6 بهترین مدل ترکیبی موجک عصبی قادر بوده میزان RMSE را برای یون نیترات در مقایسه با مدل شبکه عصبی و رگرسیون خطی چند متغیره به ترتیب به میزان 30/13 درصد و 71/89 درصد برای یون کلراید، نسبت به مدل شبکه عصبی و رگرسیون خطی چندمتغیره به ترتیب به اندازه 31/30 درصد و 57/10 درصد بهبود بخشد. این برتری به این دلیل است که مدل ترکیبی به واسطه سیگنال‌های تجزیه شده سری زمانی توانسته است خصوصیات غیرخطی و پیچیده سری زمانی را دقیق‌تر از ANN و MLR شبیه‌سازی کند. شکل 4 و شکل 5، به ترتیب یون‌های نیترات و کلراید پیش‌بینی شده را توسط



شکل 5-مقایسه کلراید مشاهده شده و پیش‌بینی شده در دوره آزمون

Figure5- Comparisons between the observed and predicted Chloride based on the testing data

درصد خطا و RMSE کمتر نسبت به دو مدل دیگر توانسته است نقاط پیک سری زمانی را پیش‌بینی کند. شکل 6، به مقایسه مقادیر نیترات و کلراید مشاهده و پیش‌بینی شده توسط سه مدل برای مقادیر بیشتر از میانگین هر پارامتر در دوره آزمون پرداخته است. همان‌طور که از این نمودارها مشخص است، برای هر پارامتر کیفی، مدل موجک عصبی نسبت به دو مدل دیگر، رابطه‌ای نزدیک به 1:1 (حول نیمساز رسم شده) بین مقادیر مشاهده شده و مقادیر پیش‌بینی شده برقرار کرده است.

به لحاظ اینکه مقدار میانگین یون نیترات و کلراید در دوره آزمون، به ترتیب 5/23 و 9/25 پی‌پی‌ام می‌باشد؛ لذا در این قسمت از پژوهش، به ارزیابی مقادیر بالاتر از میانگین هر پارامتر، و ماه‌هایی از دوره آزمون که تغییر ناگهانی داشته اند؛ پرداخته شده است، چون یافتن علت تغییرات ناگهانی پارامترهای کیفی، از مهم‌ترین اهداف مدیریت کیفیت منابع آب محسوب می‌شود. جدول 7 و جدول 8، به ترتیب عملکرد هر سه مدل برای پیش‌بینی نقاط پیک سری زمانی نیترات و کلراید را نشان می‌دهد. مطابق این جداول در پیش‌بینی هر دو پارامتر مدل WANN با

جدول 7- ارزیابی مدل‌ها برای پیش‌بینی مقادیر نیترات بیشینه در دوره آزمون

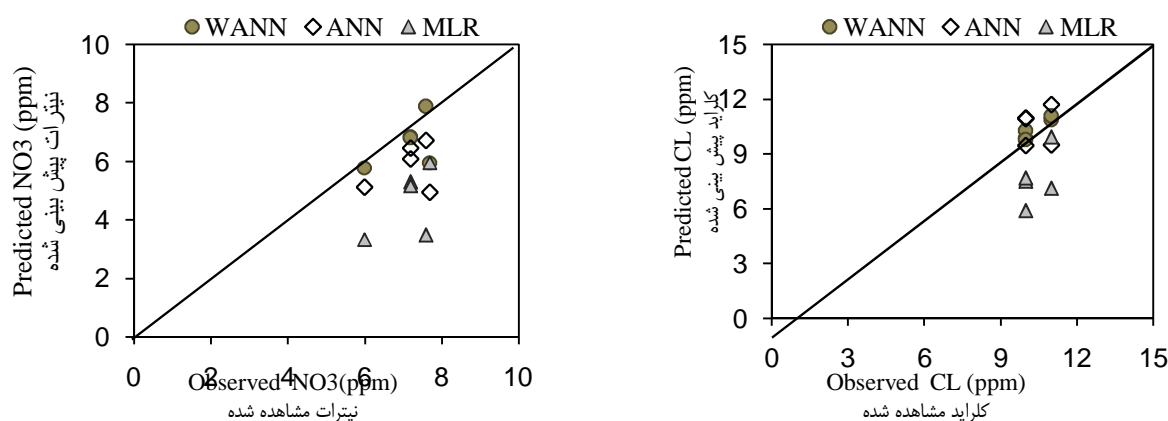
Table 7- Evaluation of models for predictions of Nitrate values max in testing period

| پیش‌بینی (Prediction) | پیش‌بینی (Prediction) | پیش‌بینی (Prediction) | ردیف (Row) تاریخ (Date) |
|-----------------------|-----------------------|-----------------------|-------------------------|
| WANN (ppm)            | MLR (ppm)             | ANN (ppm)             | pr                      |
| 7.878                 | 3.473                 | 6.714                 | 1 مرداد 88              |
| 5.781                 | 3.334                 | 5.114                 | 2 آبان 88               |
| 6.815                 | 5.296                 | 6.079                 | 3 دی 88                 |
| 6.795                 | 5.160                 | 6.452                 | 4 اسفند 88              |
| 5.956                 | 5.956                 | 4.931                 | 5 اردیبهشت 88           |
| 33.225                | 23.219                | 29.290                | جمع (Sum)               |
| -7.50                 | -53.75                | -21.88                | درصد خطا (Error)        |
| 0.834                 | 2.645                 | 1.487                 | RMSE (ppm)              |

جدول 8- ارزیابی مدل‌ها برای پیش‌بینی مقادیر کلراید بیشینه در دوره آزمون

Table 8- Evaluation of models for predictions of max Chloride values in testing period

| پیش‌بینی (Prediction) | پیش‌بینی (Prediction) | پیش‌بینی (Prediction) | ردیف (Row) تاریخ (Date) |
|-----------------------|-----------------------|-----------------------|-------------------------|
| WANN (ppm)            | MLR (ppm)             | ANN (ppm)             | pr                      |
| 10.845                | 9.909                 | 9.479                 | 1 خرداد 88              |
| 9.834                 | 7.496                 | 9.443                 | 2 مرداد 88              |
| 10.257                | 5.885                 | ####                  | 3 آذر 88                |
| 11.055                | 7.099                 | ####                  | 4 دی 88                 |
| 9.675                 | 7.660                 | ####                  | 5 اسفند 88              |
| 51.756                | 38.049                | ####                  | جمع (Sum)               |
| -0.47                 | -36.67                | 0.98                  | درصد خطا (Error)        |
| 0.188                 | 3.003                 | 0.990                 | RMSE (ppm)              |



شکل 6- مقایسه نیترات و کلراید مشاهده شده و پیش‌بینی شده برای مقادیر بیشینه پارامترها در دوره آزمون  
Figure 6- Observed and predicted Nitrate and Chloride in verification period for max parameters values

### نتیجه‌گیری کلی

در این تحقیق، کاربرد مدل‌های شبکه عصبی، رگرسیون خطی چند متغیره و مدل ترکیبی موجک عصبی برای مدل‌سازی سری زمانی یون نیترات و کلراید که از جمله سری‌های ناپایستا بوده و ماهیت نوسانی دارند؛ مورد بررسی قرار گرفته است. در این پژوهش این نتیجه حاصل شده که مدل‌های هوشمندی که دارای چندین پارامتر قابل تنظیم هستند در مقایسه با مدل‌های خطی بهتر می‌توانند خصوصیات غیرخطی و پیچیده پارامترها را شبیه‌سازی کنند. مدل ترکیبی شبکه عصبی موجکی توانست، در هر دو پارامتر به مقدار قابل توجهی میزان E و RMSE را بهبود بخشد. پیش‌بینی نقاط پیک سری زمانی توسط سه مدل مربوطه ارزیابی شد و قابلیت بالای مدل ترکیبی موجک عصبی نسبت به دو مدل دیگر حاصل

شد. بنابراین، مزیت مدل ترکیبی موجک عصبی به کار گرفته شده در این تحقیق برای پیش‌بینی یون نیترات و کلراید این است که، می‌توان با استفاده از اطلاعات سری زمانی پارامترهای کیفی شرایط تأثیرگذار مانند دبی، روشی مناسب و سریع برای پیش‌بینی پارامترهای مذکور و در نتیجه مدیریت برتر کیفیت منابع آب به شمار آید. پیشنهاد می‌شود تحقیقات بعدی جهت بررسی کارایی این مدل‌ها برای پیش‌بینی پارامترهای کیفی مختلف در دیگر حوضه‌های آبخیز و یا گام‌های زمانی مختلف (روزانه، هفتگی، سالانه و ...) انجام شود. همچنین استفاده از سری‌های زمانی تأثیر گذار متفاوتی مانند دما، اسیدیته و... تجزیه با موجک و توسعه مدل ترکیبی شبکه عصبی - موجک کارآمد در پیش‌بینی کیفیت رودخانه می‌تواند مورد بررسی قرار گیرد.

### منابع

- 1- Adamowski J., and Chan F.H. 2011. A wavelet neural network conjunction model for groundwater level forecasting. *Journal of Hydrology*, 407(1- 4): 28-40.
- 2- Adamowski J., and Sun K. 2010. Development of a coupled wavelet transform and neural network method for flow forecasting of non- perennial rivers in semi- arid watersheds. *Journal of Hydrology*, 390 (1-2): 85-91.
- 3- Cannas B., Fanni A., Sias G., Tronei S., and Zedda M.K. 2005. River flow forecasting using neural networks and wavelet analysis. *European Geosciences Union, Vienna: Austria*, 7: 24-29.
- 4- Dastorani M.T., Azimi Fashi K.H., Talebi A., and Ekhtesasi M.R. 2012. Estimation of Suspended Sediment Using Artificial Neural Network (Case Study: Jamishan Watershed in Kermanshah). *Journal of the Watershed Management Research*, 3(6):61-74. (in Persian with English Abstract)
- 5- Eynard J., Grieu S.P., and Polit M. 2011. Wavelet-based multi-resolution analysis and artificial neural networks for forecasting temperature and thermal power consumption. *Engineering Applications of Artificial Intelligence*, 24:501-516.
- 6- Fernando A.K., and Kerr T. 2003. Runoff forecasting with artificial neural network model. The 3rd Pacific Conference on stormwater and aquatic resource protection, 14-16 May 2003, Auckland, NZ.
- 7- Han J.G., Ren W.X., and Sun Z.S. 2005. Wavelet packet based damage identification of Beam structures. *International. Journal of Solids and Structures*, 4: 6610-6627.
- 8- Heydarizad M., and Mohammadzadeh H. 2012. Investigation of Seasonal and Spatial Variation of Hydrochemical

- Parameters in Karde River (North of Mashhad). *Journal of Water and Soil*, 26(5):1161-1170 (in Persian with English Abstract).
- 9- Imrie C.E., Durucan S., and Korre A. 2000. River flow prediction using artificial neural networks: generalisation beyond the calibration range. *Journal of Hydrology*, 233(1-4):138-153.
  - 10- Kashefiolasl M., and Zaeimdar M. 2009. Necessity of quality management Jajrood River. *Journal of Environmental Science and Technology*, 11(2):120-129. (in Persian)
  - 11- Kim T.W., and Valdes J.B. 2003. Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. *Journal of Hydrology Engineering*, ASCE, 8(6):319-28.
  - 12- Kisi O. 2004. Multi-layer perceptrons with Levenberg-Marquardt training algorithm for suspended sediment concentration prediction and estimation. *Hydrological Sciences Journal*, 49(6):1025-1040.
  - 13- Makhdoum M. 2005. Living in the environment. Tehran: Tehran University Publishings. (in Persian)
  - 14- May D., and Sivakumar M. 2008. Comparison of artificial neural network and regression models in the prediction of urban stormwater quality. *Water Environment Research; Water Environment Research*, 80(1):4-9.
  - 15- Najah A., El-shafie A., and Karim O.A. 2009. Prediction of Johor River Water Quality Parameters Using Artificial Neural Networks. *European Journal of Scientific Research*, 422-435.
  - 16- Najah A., El-Shafie A., Karim O.A., Jaafar O., and El-Shafie A.H. 2011. An application of different artificial intelligences techniques for water quality prediction. *Journal of the Physical Sciences*, 6(22): 5298-5308.
  - 17- Nayak P.C., SatyjiRao Y.R., and Sudheer K.P. 2006. Groundwater level forecasting in a shallow aquifer using artificial neural network approach. *Water Resources Management*, 20:77-90.
  - 18- Nourani V., Hasanzade Y., Komasi M., and Sharafi A. 2008. Modeling of rainfall - runoff by using wavelet - neural network model. 4th National Congress on Civil Engineering, University of Tehran (in Persian).
  - 19- Nourani V., Komasi M., and Mano A. 2009. A multivariate ANN-Wavelet approach for rainfall runoff modeling. *Water Resources Management*, 23: 2877-2894.
  - 20- Noshadi A., Salemi H., and Ahmadzade M. 2007. Simulation and prediction of some water quality parameters in the Zayanderoud River using artificial neural networks. *Journal of the Water and the Wastwater*, 64(18):49-65. (in Persian)
  - 21- Palani S., Liong Sh., and Tkalich P. 2008. An ANN application for water quality forecasting. *Marine Pollution Bulletin*, 56:1586-1597.
  - 22- Rajae T., Nourani V., Zounemat K.M., and Kisi O. 2011. River suspended sediment load prediction: application of ANN and wavelet conjunction model. *Journal of Hydrology Engineering*, ASCE, 16(8):613-627.
  - 23- Rajae T. 2011. Wavelet and ANN combination model for prediction of daily suspended sediment load in rivers. *Science of the Total Environment*, 409:2917-2928.
  - 24- Safavi H.R. 2010. Prediction of river water quality by adaptive neuro fuzzy inference system (ANFIS). *Journal of the Environmental Studies*, 36(53):1-10. (in Persian)
  - 25- Singh K. P., Basant A., Malik A., and Jain, G. 2009. Artificial Neural Network modelling of the river water quality-A case study. *Journal of Ecological Modeling*, 220:888-895.
  - 26- Singh R.M. 2012. Wavelet-ANN model for flood events, *Advances in Intelligent and Soft Computing*, 131: 165-175.
  - 27- Taebi A., and Vashtani M. 2000. Prediction the quality of urban runoff: Methods and Selected Model. *Journal of the College of the Engineering*, 11(2):41-48. (in Persian)
  - 28- Toufani P., Mosaedi A., and Fakheri Fard A. 2011. Prediction of Precipitation Applying Wavelet Network model (Case study: Zarringol stati on, Golestan province, Iran). *Journal of Water and Soil*, 25(5):1217-1226. (in Persian with English abstract)
  - 29- Wen X., Fang J., Diao M., and Zhang C. 2012. Artificial Neural Network modeling of Dissolved Oxygen in the Heihe River, Northwestern China. *Environ Monit Assess*, 185(5):4361-4371.

## Prediction of Water Quality Parameters (NO<sub>3</sub>, CL) in Karaj River by Using a Combination of Wavelet Neural Network, ANN and MLR Models

T. Rajaei<sup>1</sup> - R. Rahimi Benmaran<sup>2\*</sup>

Received: 09-04-2014

Accepted: 03-05-2015

**Introduction** The water quality is an issue of ongoing concern. Evaluation of the quantity and quality of running waters is considerable in hydro-environmental management. The prediction and control of the quality of Karaj river water, as one of the important needed water supply sources of Tehran, possesses great importance. In this study, Performance of Artificial Neural Network (ANN), Wavelet Neural Network combination (WANN) and multi linear regression (MLR) models, to predict next month the Nitrate (NO<sub>3</sub>) and Chloride (CL) ions of "gate of Bylaqan sluice" station located in Karaj River has been evaluated.

**Materials and Methods** In this research two separate ANN models for prediction of NO<sub>3</sub> and CL has been expanded. Each one of the parameters for prediction (NO<sub>3</sub> / CL) has been put related to the past amounts of the same time series (NO<sub>3</sub> / CL) and its amounts of Q in past months. From a statistical period of 10 years was used for the input of the models. Hence 80% of entire data from (96 initial months of data) as training set, next 10% of data (12 months) and 10% of the end of time series (terminal 12 months) were considered as for validation and test of the models, respectively. In WANN combination model, the real monthly observed time series of river discharge (Q) and mentioned quality parameters (NO<sub>3</sub> / CL) were decomposed to some sub-time series at different levels by wavelet analysis. Then the decomposed quality parameters to predict and Q time series were used at different levels as inputs to the ANN technique for predicting one-step-ahead Nitrate and Chloride. These time series play various roles in the original time series and the behavior of each is distinct, so the contribution to the original time series varies from each other. In addition, prediction of high NO<sub>3</sub> and CL values greater than mean of data that have great importance were investigated by the models. The capability of the models was evaluated by Coefficient of Efficiency (E) and the Root Mean Square Error (RMSE). An efficiency of one corresponds to an accurate match of forecasted data to the observed data. RMSE indicates the discrepancy between the observed and predicted values

**Results Discussion** The results indicates that the accuracy and the ability of hybrid model of wavelet neural network had been better than the other two modes; so that hybrid model of Wavelet artificial neural network was able to improve the rate of RMSE for Nitrate ions in comparison with ANN and MLR models respectively, amounting to 30.13% and 71.89%, for chloride ion as much as 31.3% and 57.1%. In the WANN model increasing the decomposition level, in level 1 to Level 3, increases the model's performance, but increasing the decomposition level, in levels over Level 3, decreases the model's efficiency, because high decomposition levels lead to a large number of parameters with complex nonlinear relationships in the ANN technique. The WANN model needed 1 to 7 neurons in the hidden layer for the best performance result. In prediction of high NO<sub>3</sub> values the amount RMSE for ANN, MLR and WANN models are 1.487, 2.645 and 0.834 ppm, respectively. Also, for CL values the mentioned statistical parameter is 0.990, 3.003 and 0.188 ppm, respectively for models. The results exhibits that the combined model of WANN the forecast was better than the other two models.

**Conclusion** Wavelet transforms provide useful decompositions of original time series, so that wavelet-transformed data improve the ability of a predicting model by capturing useful information on various resolution levels. The main advantage of this study is that only from the Q and slightly quality of parameter time series are used until the same quality of parameter in one month ahead is predicted. The purpose of entering Q time series with quality of parameter as inputs of models is analysis the efficacy of Q in the accuracy of prediction. owing of the high capability wavelet neural network in the prediction of quality parameters of river's water, this model can be convenient and fast way to be proposed for management of water quality resources and assurance from water quality monitoring results and reduction its costs.

**Keywords:** Karaj River, Linear Regression, Neural Network, Nitrate and Chloride ions, Wavelet Analysis

1, 2 - Associate Professor and MSc Hydraulic Structures, Department of Civil Engineering, University of Qom  
(\*-Corresponding Author Email: r.rahimi\_b@yahoo.com)