

پیش بینی پارامترهای کیفی (DO ، NO_3) رودخانه کرج با استفاده از مدل‌های
 ANN ، MLR و تلفیق شبکه عصبی - موجکی بر پایه نويززداییطاهر رجایی^۱، رقيه رحيمي بنماران^{۲*}، حميده جعفری^۳

تاریخ دریافت: ۹۳/۰۴/۳۱ تاریخ پذیرش: ۹۳/۰۷/۳۰

چکیده

زمینه و هدف: پیش بینی و کنترل کیفیت آب رودخانه کرج، به عنوان یکی از مهمترین منابع تامین کننده آب مورد نیاز شهر تهران، از اهمیت بسزایی برخوردار است. در این تحقیق، عملکرد مدل‌های شبکه عصبی (ANN)، مدل ترکیبی شبکه عصبی - موجکی ($WANN$) و رگرسیون خطی چند متغیره (MLR)، برای پیش بینی یک ماه آینده یون نیترات و اکسیژن محلول ایستگاه پل خواب واقع در رودخانه کرج، مورد ارزیابی قرار گرفت. روش بررسی: از یک دوره آماری ۱۱ ساله جهت ورودی مدل‌ها استفاده شد. در مدل ترکیبی $WANN$ سری‌های زمانی واقعی دبی و پارامتر کیفی مورد نظر (نیترات و اکسیژن محلول) توسط آنالیز موجکی تجزیه شدند؛ سپس مجموع سری‌های زمانی موثر آنها به عنوان ورودی ANN به کار گرفته شد. در ادامه، توانایی هر سه مدل، برای پیش بینی نقاط پیک سری زمانی که دارای اهمیت بسزایی هستند، بررسی شد. کارایی مدل‌ها با ضریب تبیین یا E و ریشه میانگین خطای مربع ($RMSE$) ارزیابی شدند.

یافته‌ها: نتایج، حاکی از دقت و توانایی بالای مدل هیبرید شبکه عصبی - موجکی با رویکرد حذف نویزهای سری زمانی نسبت به دو مدل دیگر بوده است؛ بطوریکه 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 $۳۵/۶۰\%$ و $۷۵/۹۳\%$ و برای یون اکسیژن محلول، به اندازه $۴۰/۵۷\%$ و $۶۰/۱۳\%$ بهبود بخشد. **نتیجه‌گیری:** به لحاظ قابلیت بالای شبکه عصبی موجکی و حذف نویزهای سری‌های زمانی در پیش بینی پارامترهای کیفی آب رودخانه، این مدل می‌تواند، راهکاری مناسب و سریع در مدیریت برتر کیفیت منابع آب و اطمینان از نتایج پایش کیفی و کاهش هزینه‌های آن مطرح شود.

واژگان کلیدی: رودخانه کرج، شبکه عصبی، تبدیل موجکی، یون نیترات و اکسیژن محلول، نويززدایی.

r.rahimi_b@yahoo.com

۱- دکتری آب- محیط زیست، استادیار گروه مهندسی عمران، دانشگاه قم
 ۲- (نویسنده مسئول): کارشناس ارشد سازه‌های هیدرولیکی، گروه مهندسی عمران، دانشگاه قم
 ۳- کارشناس ارشد سازه‌های هیدرولیکی، گروه مهندسی عمران، دانشگاه قم

مقدمه

رودخانه ها و آب های سطحی در تولید بخشی از مواد غذایی مورد نیاز جامعه اهمیت زیادی دارند و از این جهت مدیریت آگاهانه بر آن ضروری است. از آنجایی که امروزه در دنیا آب و منابع آب یکی از پایه های اصلی توسعه پایدار به شمار می روند، علاوه بر کمیت و میزان آورد رودخانه باید کیفیت آب نیز جز پارامترهای مهم، مورد توجه قرار گیرد؛ لذا بررسی و پیش بینی پارامترهای کیفی در طول یک رودخانه، یکی از اهداف مدیران و برنامه ریزان منابع آب است (۱). در این راستا شناخت کافی از شرایط منطقه و مشخصات شیمیایی و بیولوژیکی آب رودخانه و نیز تغییرات پارامترهای کیفی به عنوان گام اول برنامه های مدیریت کیفیت منابع آب در نظر گرفته می شود؛ به عبارتی مدیریت کیفیت با شناخت کیفیت موجود آب، آلودگی ها و روش های حفظ کیفیت آب سروکار دارد (۲).

نیتрат (NO₃⁻) و اکسیژن محلول (DO) از مهمترین پارامترهای کیفی آب به شمار می آیند. نیترات ماده ای است که آب های سطحی را به طور گسترده ای آلوده می کند. از مهمترین منابع آلودگی نیترات استفاده از کودهای ازته است که در کشاورزی، برای حاصلخیزی زمین استفاده می شود. این کودها در اثر تجزیه شدن و انحلال در آب های باران و کشاورزی در زمین جاری شده و نهایتاً به قسمت عمقی زمین فرو می رود. بنابراین می تواند آب های سطحی و زیرزمینی را تحت تاثیر قرار دهند (۳). تجمع نیترات در محیط، ناشی از هرزآب های کشاورزی هستند که، به طور منتشر آب را آلوده می کنند یا از آلودگی های نقطه ای مانند فاضلاب های انسانی به درستی دفع و فراوری نشده است، بوجود می آیند. نیترات شاخص مناسبی برای نشان دادن آلودگی از نظر بهداشت آب آشامیدنی است که به عنوان یک ماده مغذی بوجود آورنده بوتریفیکاسیون نیز حائز اهمیت است (۴). حداکثر مقدار مجاز آلاینده، برای نیترات در حد ۱۰ ppm نیتروژن یا ۴۴ ppm نیترات تعیین شده است. اکسیژن به مقدار کم در آب محلول بوده و حلالیت آن با فشار اتمسفر و درجه حرارت متغیر است. مقدار اکسیژن محلول در آبی که ماهی در آن پرورش می یابد، نباید کمتر از ۵ ppm باشد. کمبود

حلالیت اکسیژن در آب یکی از فاکتورهای اصلی است که ظرفیت تصفیه طبیعی آب را کاهش می دهد. لذا تصفیه فاضلاب ها قبل از ورود به رودخانه ها ضروری خواهد بود. با اکسیژن محلول رودخانه ها نیز می توان آلودگی آنها را کنترل نمود. پیش بینی یون های نیترات و اکسیژن محلول نقش تعیین کننده ای در مدیریت کیفی و سیاست گذاری های زیست محیطی منابع آب دارد. به لحاظ غیرخطی بودن برخی فرایندهای فیزیکی و عدم قطعیت در مشخصه های تاثیرگذار می توان از تکنیک مدل سازی جهت شبیه سازی و پیش بینی فرایندهای غیرخطی و غیرقطعی، استفاده کرد (۵ و ۶). در این راستا، Palani و همکاران (۲۰۰۸) پارامترهای کیفی آب های ساحلی سنگاپور را با استفاده از شبکه عصبی مصنوعی (Artificial Neural Network) مدل کردند؛ در این مدل پارامترهای شوری، دما، اکسیژن محلول و کلروفیل a، به عنوان داده های ورودی در نظر گرفته شد (۷). May و همکار (۲۰۰۸) شبکه عصبی مصنوعی و مدل های رگرسیونی را برای پیش بینی کیفیت آب رودخانه های شهری مورد استفاده قرار دادند (۸). نتایج تحقیق آنها و همچنین مطالعات Kisi و همکاران (۲۰۰۴) نشان داد، مدل های شبکه عصبی مصنوعی به دلیل قابلیت یادگیری ارتباطات پیچیده و مدل سازی غیرخطی، عملکرد بهتری دارند (۸ و ۹). Singh و همکاران (۲۰۰۹) مدل شبکه عصبی مصنوعی را برای تخمین میزان غلظت اکسیژن محلول، اکسیژن خواهی بیوشیمیایی، رودخانه Gomti در هند توسعه دادند؛ از ۱۱ پارامتر کیفی آب که طی ۱۰ سال بطور ماهانه در ۸ ایستگاه مختلف اندازه گیری شده بود به عنوان ورودی مدل استفاده کردند (۱۰). Najah و همکاران (۲۰۰۹) کیفیت آب رودخانه Johor (۱۱) و Wen و همکاران (۲۰۱۲) اکسیژن محلول رودخانه Heihe چین (۱۲) را در تحقیق جداگانه ای، با استفاده از شبکه عصبی مصنوعی پیش بینی نمودند؛ نتایج آنها حاکی از دقت و توانایی بالای مدل پیشنهادی برای پیش بینی و شبیه سازی این پارامترها بود. Najah و همکاران (۲۰۱۱) تکنیک های هوشمند عصبی مختلفی را برای پیش بینی پارامترهای کیفی اکسیژن محلول

استفاده کردند؛ در تحقیق خود توانایی بالای ترکیب موجک و شبکه عصبی نسبت به دیگر روش ها (مدل های رگرسیون، منحنی های سنجه) برای محاسبه بار معلق رسوب رودخانه را نتیجه گیری کردند (۱۴). Rajaee (۲۰۱۱) مدل ترکیبی جدیدی را برای پیش بینی بار معلق روزانه رودخانه ها ارائه داد و توانایی بالای مدل ترکیبی را نسبت به دیگر روش ها را نتیجه گیری کرد (۱۹). Banejad و همکاران (۲۰۱۲) از دو مدل شبکه عصبی مصنوعی و تلفیقی شبکه عصبی موجکی برای تخمین کل جامدات محلول (TDS)، هدایت الکتریکی (EC) و نسبت جذبی سدیم (SAR)، رودخانه های جاجرود و قره سو کرمانشاه استفاده کردند و قابلیت بالای مدل ترکیبی را نسبت به ANN نشان دادند (۲۰).

هدف از تحقیق حاضر، توسعه شبکه های عصبی مصنوعی، شبکه های عصبی-موجکی با ایده ای نو بر پایه حذف نویزهای سری زمانی و رگرسیون خطی چند متغیره، برای برآورد یون های نیترا و اکسیژن محلول و مقایسه عملکرد و کارایی این سه روش، در شبیه سازی پارامترهای کیفی مذکور ایستگاه پل خواب رودخانه کرج است. این مطالعه، به لحاظ تکنیک ترکیب شبکه عصبی اتورگرسیو غیرخطی با موجک و حذف نویزهای سری زمانی جهت بهبود عملکرد مدل، برای پیش بینی سری زمانی یون های نیترا و اکسیژن محلول، تحقیقی کاربردی محسوب می شود.

مواد و روش ها

منطقه مورد مطالعه

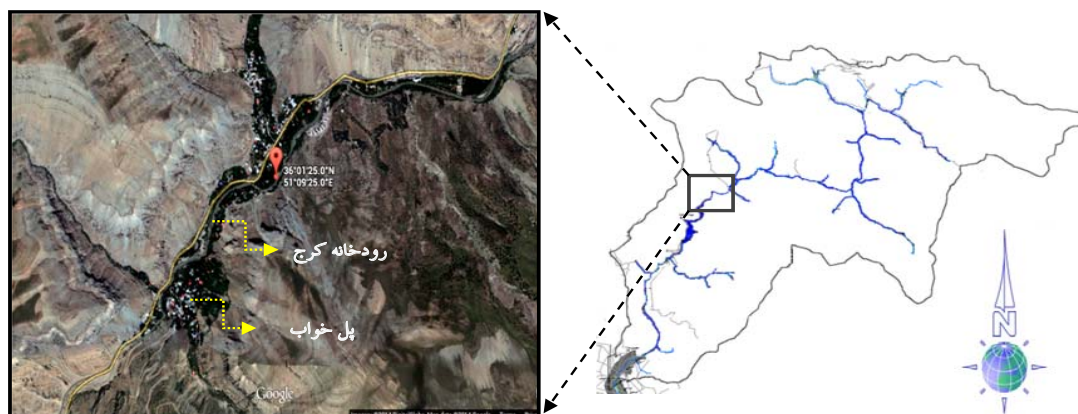
رودخانه کرج، یکی از مهمترین رودخانه های حوضه آبریزی مرکزی بوده که در سمت شمال شرقی این حوضه واقع شده است. از ضلع جنوبی کوهستان البرز و ۴۰ km غرب تهران، از کانون آبریز خرسنگ کوه، سرچشمه می گیرد. طول رودخانه کرج، حدود ۲۴۵ km و مساحت حوضه آبریز آن، حدود ۵۰۰۰ km² برآورد شده است. نظر به اهمیت منابع آب این حوضه و کیفیت در شرف تهدید آن، آگاهی از میزان و نوع آلاینده ها، به منظور اتخاذ تصمیمات مدیریتی امری ضروری

(DO)، اکسیژن خواهی بیولوژیکی (BOD) و اکسیژن خواهی بیوشیمیایی (COD) به کار بردند. هدف آنها ارائه مدلی برای پیش بینی پارامترهای کیفی با هدف صرفه جویی در هزینه و زمان لازم برای اندازه گیری این پارامترها بود. قابلیت بالای مدل شبکه عصبی پرسپترون با یک لایه مخفی را نسبت به رگرسیون خطی چند متغیره نتیجه گرفتند (۱۳).

یکی از روش های هوشمندی که طی دهه های اخیر در مسائل آب و هیدرولوژیکی مورد توجه قرار گرفته است، استفاده از تئوری موجک (Wavelet Transforms) به عنوان روشی نوین با قابلیت بالا، در زمینه تحلیل نوسانات، تناوب و سری های زمانی نایستا و گذرا است. شبکه عصبی و آنالیز موجک در حالتی که به طور مجزا در مسائل مربوط به مدیریت آب و محیط زیست به کار گرفته شده اند عملکرد مناسبی ارائه داده اند. اخیرا تمایل به استفاده از مدل های ترکیبی مانند هیبرید شبکه عصبی-موجک به طور روز افزون افزایش یافته است (۱۴). در این راستا، Kim و همکار (۲۰۰۳) برای پیش بینی خشکسالی (۱۵) و Cannas و همکاران (۲۰۰۵) برای مدل سازی بارش رواناب ماهانه (۱۶) در تحقیقات مجزایی مدل ترکیبی شبکه عصبی-موجکی را توسعه دادند. هر یک از این محققین نتیجه گرفتند؛ شبکه عصبی آموزش داده شده با داده های پیش پردازش شده، در مقایسه با شبکه های عصبی که با داده های خام و نویزدار آموزش داده شده اند، کارایی بهتری داشته اند. Nourani و همکاران (۲۰۰۹) از مدل تلفیقی شبکه عصبی-موجکی برای پیش بینی دبی رواناب یک روز بعد، استفاده نموده و به این نتیجه رسیدند که توانایی شبکه عصبی-موجکی در پیش بینی نقاط بیشینه نمودار سری زمانی، نسبت به سایر روش ها بیشتر است (۱۷). Adamowski و همکاران (۲۰۱۰) از مدل ترکیبی تبدیلات موجکی گسسته و شبکه های عصبی، برای پیش بینی جریان یک و سه روزه در رودخانه های غیردائمی حوضه های نیمه خشک Kargotis، استفاده کردند؛ نتایج نشان داد، شبکه های عصبی-موجکی روش مؤثری برای پیش بینی جریان کوتاه مدت، در اینگونه حوضه ها هستند (۱۸). Rajaee و همکاران (۲۰۱۱) از مدل شبکه عصبی موجکی برای پیش بینی بار معلق روزانه رودخانه Iowa آمریکا

ایستگاه پل خواب را به عرض جغرافیایی ۳۶° ۰۱' ۲۵" و طول جغرافیایی ۵۱° ۰۹' ۲۵" و ارتفاع ۱۸۱۹ به عنوان ایستگاه مورد مطالعه، نشان می دهد.

به نظر می رسد. رودخانه کرج به عنوان یکی از مهمترین منابع تامین کننده آب مورد نیاز شهر تهران بوده و کنترل کیفیت آب آن، از اهمیت بسزایی برخوردار است. شکل ۱ موقعیت



شکل ۱. منطقه مورد مطالعه

داده ها و تحلیل آماری

کل داده ها) از خرداد ۱۳۸۹ تا اردیبهشت ۱۳۹۰ برای آزمون شبکه در نظر گرفته شد. در جدول ۱ تحلیل آماری هر سه پارامتر ایستگاه، برای مجموعه های آموزشی (مجموعه اعتبار سنجی جزء داده های آموزش لحاظ شده است)، آزمون و کل داده ها ارائه شده است. این جدول شامل مقادیر بیشینه، کمینه، میانگین، انحراف معیار (S_d), ضریب چولگی (C_s), ضریب خودهمبستگی یک ماهه (P_1), ضریب خودهمبستگی دو ماهه (P_2), ضریب خودهمبستگی سه ماهه (P_3) و ضریب خودهمبستگی چهار ماهه (P_4) است.

آمار مورد استفاده در این تحقیق، غلظت یون های نترات، اکسیژن محلول و میزان دبی در یک دوره یازده ساله است که این داده ها بصورت ماهانه از خرداد ۱۳۷۹ تا اردیبهشت ۱۳۹۰ توسط شرکت آب و فاضلاب استان تهران به ثبت رسیده است. پس از بررسی داده ها، دوره های مربوط به آموزش و آزمون مدلها انتخاب گردید. بدین ترتیب ۱۰ سال ابتدایی داده ها (۹۰٪ از کل داده ها) از خرداد ۱۳۷۹ تا اردیبهشت ۱۳۸۸ برای آموزش شبکه، یک سال بعدی، از خرداد ۱۳۸۸ تا اردیبهشت ۱۳۸۹ برای اعتبار سنجی و یک سال انتهایی دوره (۱۰٪ از

جدول ۱. تحلیل آماری داده های کل، آموزش و آزمون پارامترهای ایستگاه مورد مطالعه

پارامتر آماری	NO ₃ (ppm)			Q (m ³ /s)			DO(ppm)		
	کل داده	آموزش	آزمون	کل داده	آموزش	آزمون	کل داده	آموزش	آزمون
تعداد داده ها	۱۳۲	۱۲۰	۱۲	۱۳۲	۱۲۰	۱۲	۱۳۲	۱۲۰	۱۲
میانگین	۳/۵۵	۳/۳۹	۵/۱۳۰	۱۵/۰۹	۱۴/۹۱	۱۶/۹۱	۸/۵۹	۸/۵۶	۸/۸۴
Max	۹/۱۰	۹/۱۰	۶/۷۰	۱۳۷/۱۷	۱۳۷/۱۷	۵۶/۸۱	۱۰/۹	۱۰/۷	۱۰/۹
Min	۱/۸۰	۱/۸۰	۲/۲۹	۰/۹۰۶	۰/۹۰۶	۳/۷۵۵	۶/۶۰	۶/۶۰	۶/۹۰
S_d	۱/۴۷۱	۱/۳۷۲	۱/۵۴۸	۱۷/۳۵۱	۱۷/۳۴۵	۱۸/۰۶۹	۰/۹۲۶	۰/۹۰۶	۱/۱۲۲
C_s	۱/۴۵۳	۱/۸۵۰	-۰/۷۳۷	۲/۹۶۸	۳/۱۲۸	۱/۴۳۷	-۰/۲۴۹	-۰/۲۳۹	۰/۱۵۵
R1	۰/۴۴۳	۰/۴۳۶	-۰/۰۱۴	-۰/۵۰۷	۰/۴۹۰	-۰/۷۶۹	۰/۶۴۸	۰/۶۴۶	۰/۶۳۶
R2	۰/۴۳۳	۰/۴۳۵	-۰/۱۴۷	-۰/۱۷۸	۰/۱۹۰	-۰/۱۳۶	۰/۳۷۶	۰/۳۶۲	۰/۴۴۸
R3	۰/۳۳۷	۰/۲۹۶	-۰/۳۴۰	-۰/۰۳۶	-۰/۰۱۵	-۰/۳۵۹	۰/۰۳۲	۰/۰۱۴	۰/۱۷۰
R4	۰/۲۷۴	۰/۱۹۰	۰/۰۴۶	-۰/۱۹۳	-۰/۱۷۵	-۰/۴۳۲	-۰/۳۲۰	-۰/۳۱۹	-۰/۴۸۲

(Regression)، تکنیکی در پیش بینی های 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 می شود:

$$y = a_0 + \sum a_i x_i \quad (1)$$

که در آن y مقدار پیش بینی شده، a_0 ثابت رگرسیون، a_i ضریب رگرسیون متغیرهای مستقل X_i است. هدف این روش، تعیین پارامترهای معادله رگرسیون است به گونه ای که مقادیر تخمین زده شده، کارا و سازگار باشند (۱۹ و ۱۲).

تبدیل موجک

موجک ها، توابع ریاضی هستند که یک بیان زمان-مقیاس از سری های زمانی و نسبت های آن ارایه می دهند، این بیان برای تحلیل های سری زمانی نایستا (Nonstationary time series) مفید است (۲۴ و ۱۹). به وسیله این تبدیل ها، سری های زمانی به موجک هایی تجزیه می شوند که هر یک نسخه ای جابجا شده و مقیاس شده از موجک مادر هستند. تبدیل های موجک، همانند تبدیل های فوریه عمل می کنند (۱۴). تبدیل موجک در دو نوع پیوسته و گسسته در زمان به کار می رود. اختلاف انواع مختلف تبدیل موجک، مربوط به روشی است که مقیاس و انتقال را پیاده سازی می کند. برای یک سری زمانی با تابع $f(t)$ تبدیل موجک پیوسته آن به صورت زیر است.

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

که در آن تابع $\Psi(t)$ تابعی مرکب از $\Psi(t)$ است. تبدیل موجک پیوسته، در تمام مقیاس ها می تواند عمل کند؛ ولی در عین حال به مقدار زیادی داده و زمان محاسباتی نیاز دارد؛ همچنین برای بدست آوردن تبدیل موجک گسسته با جایگذاری مقادیر در رابطه ۲ رابطه ۳ حاصل می شود.

$$CWT(j, k) = W_{\psi} f(j, k) = a_0^{-1/2} \int_{-\infty}^{+\infty} f(t) \bar{\psi}(a_0^{-1}t - kb_0) dt \quad (3)$$

در اکثر تحقیقات از تبدیل موجک گسسته، به دلیل سادگی و سرعت بالای محاسبات با توجه به دقت مورد نیاز، استفاده

مدل های شبکه عصبی، زمانی بهتر عمل می کنند که خصوصیات آماری داده های آموزش و آزمون، تا حد زیادی مشابهت داشته باشند (۱۹ و ۲۱). طبق این جدول، ضریب چولگی خصوصا برای اکسیژن محلول نسبتا پایین است؛ زیرا ضریب چولگی بالا اثرات منفی قابل ملاحظه ای بر عملکرد شبکه های عصبی دارد (۱۹). قرار گرفتن مقادیر بیشینه و حدی در مجموعه آموزشی در بهبود عملکرد مدل ها بسیار موثر است. مطابق این جدول، خودهمبستگی برای تاخیرهای ۱ و ۲ و ۳ ماهه دوره آموزش نیترا، نسبتا بالاست؛ برای دی و اکسیژن محلول تاخیر ۱ و ۲ ماهه مجموعه آموزش و آزمون، خودهمبستگی بالایی را نشان می دهد. بنابراین جدول ۱ خصوصیات آماری رضایت بخشی را بین مجموعه آموزش و آزمون از لحاظ میانگین، بیشینه، کمینه، انحراف معیار و ضرایب همبستگی نشان می دهد.

شبکه های عصبی مصنوعی (ANN)

شبکه عصبی مصنوعی (Artificial Neural Network) الهام گرفته از فرآیند یادگیری مغز انسان است؛ 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 (Train) یافته؛ نحوه عملکرد سیستم در حافظه ذخیره شود؛ تا در مواردی که مدل قبلا با آن مواجه نشده است، مورد استفاده قرار گیرد (۲۲). در واقع شبکه عصبی مصنوعی یک مدل ریاضی است که توانایی مدل سازی و ایجاد روابط ریاضی غیرخطی برای درون یابی را دارد (۲۳). برای آموزش شبکه الگوریتم های مختلفی وجود دارد؛ براساس تحقیقات موجود، الگوریتم لونیگ-مارکوات (Levenberg-Marquardt) به دلیل سرعت، دقت و اطمینان بالای آن در پیش بینی های هیدرولوژیکی، الگوریتم مناسبی است (۲۲).

تحلیل رگرسیون خطی چند متغیره (MLR)

تحلیل رگرسیون خطی چند متغیره (Multi Linear

شده است. در این تحقیق نیز تجزیه سیگنال سری زمانی غلظت نیترات و اکسیژن محلول به سیگنال های پایه سازنده آنها، توسط موجک مادر با سطح تجزیه ۱۰ انجام می شود. به طوریکه برای مثال موجک db2 با سطح تجزیه ۴، موجک مادر (S) را به یک موجک تقریب (a₄) و چهار موجک جزئیات (d₄, d₃, d₂, d₁) تجزیه می کند به عبارتی $S = a_4 + d_4 + d_3 + d_2 + d_1$.

کارایی مدل ها

در این تحقیق، کارایی مدل های ANN، MLR و WANN با استفاده از ضریب تبیین یا نش (Coefficient of Efficiency) و ریشه خطای مربع متوسط (RMSE) ارزیابی می شود. نش و ساتکلیف (Nash-Sutcliffe) ضریب کارایی مدل را که برای سنجش توان پیش بینی هیدرولوژیکی به کار می رود؛ در محدوده ۱ تا منفی بی نهایت به صورت رابطه ۴ تعریف کردند؛ رابطه ۵ نیز معرف RMSE ریشه میانگین خطای مربع، بین داده های مشاهده شده و پیش بینی شده است.

$$E = 1 - \frac{\sum_{i=1}^n (N_{i(o)} - N_{i(p)})^2}{\sum_{i=1}^n (N_{i(o)} - N_m)^2} \quad (4)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (N_{i(o)} - N_{i(p)})^2}{n}} \quad (5)$$

در روابط بالا به ترتیب نمایانگر مقادیر مشاهده شده، پیش بینی شده و میانگین غلظت یون نیترات و اکسیژن محلول و n تعداد داده ها است. اگر E و RMSE به ترتیب نزدیک به ۱ و ۰ شود؛ عملکرد مدل قابل قبول خواهد بود.

یافته ها

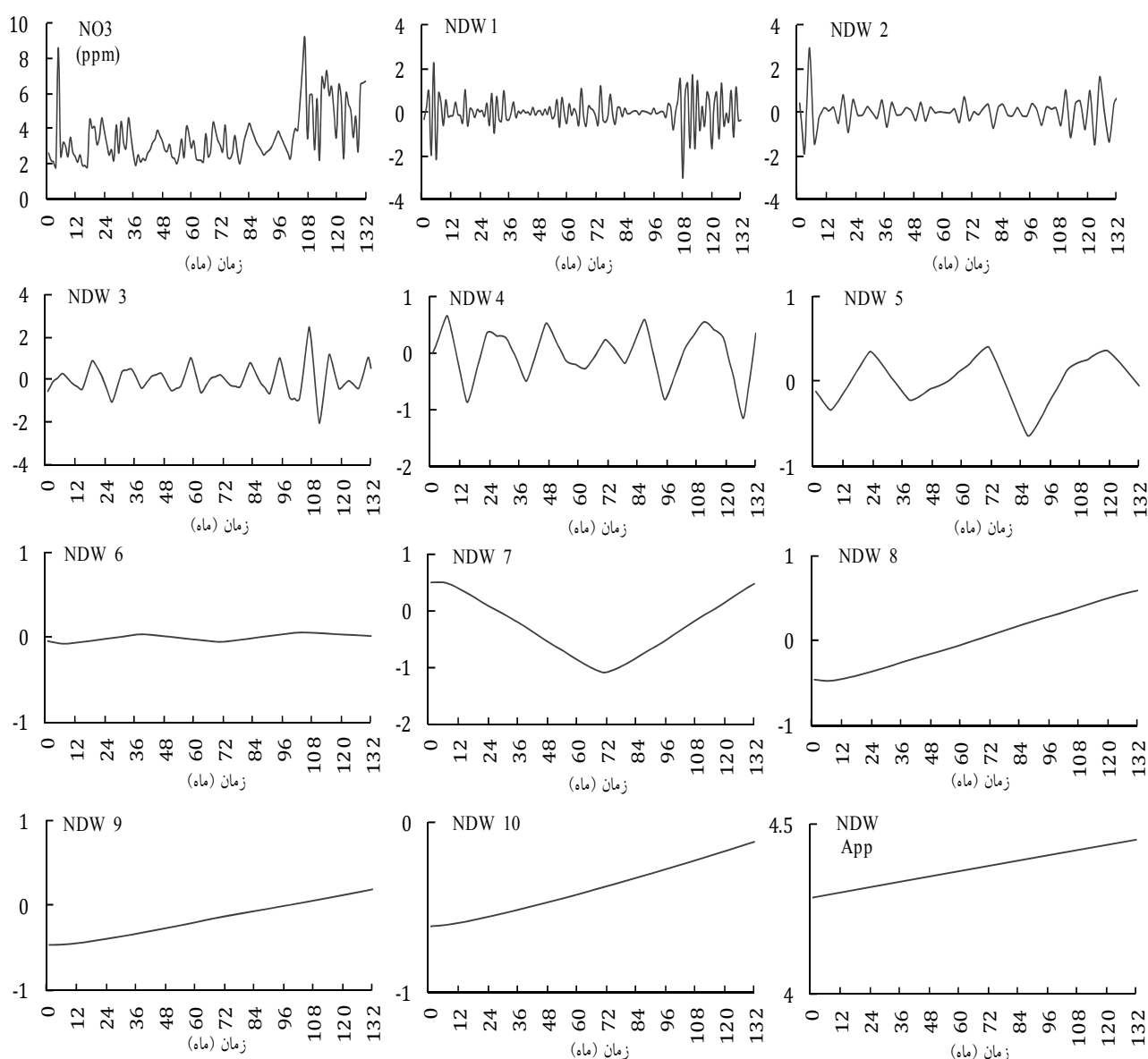
شبکه عصبی استفاده شده در این تحقیق، شبکه عصبی NARX (Nonlinear AutoRegressive with eXogenous input) است. این مدل، برای پیش بینی مقادیر آینده N(t)، از مقادیر گذشته همان سری زمانی به صورت پس خور (Feedback) و مقادیر گذشته سری زمانی دوم Q(t)، که به عنوان ورودی به مدل داده می شود، توسعه می یابد. در واقع در دو مدل جداگانه برای پیش بینی نیترات و اکسیژن محلول، مقادیر آینده سری زمانی پارامتر مورد نظر (نیترات و اکسیژن محلول)، به صورت

تابعی از مقادیر گذشته همان سری زمانی و مقادیر گذشته سری زمانی دیگر (دبی)، پیش بینی می شوند. به عبارت دیگر:

$$N_t = f(N_{t-1}, N_{t-2}, \dots, N_{t-d}, Q_{t-1}, Q_{t-2}, \dots, Q_{t-d}) \quad (6)$$

در رابطه ۶، N(t) سری زمانی پارامتر مورد نظر برای پیش بینی (نیترات و اکسیژن محلول)، Q(t) سری زمانی دبی و d تعداد تاخیر است. بدین ترتیب، شبکه عصبی NARX با حلقه پسخور باز (استفاده از مقادیر واقعی برای پیش بینی) با سه لایه پیشخور (یک لایه ورودی، یک لایه پنهان و یک لایه خروجی)، با تابع فعالیت سیگموئید در لایه پنهان و تابع فعالیت خطی در لایه خروجی، توسعه داده شد. هدف از به کارگیری مدل حلقه باز این است که، داده های مشاهداتی (برداشت های کاملاً واقعی از پدیده)، به جای داده های محاسبه شده توسط مدل که ناشی از عملکرد محاسباتی مدل هستند، لحاظ شوند. برای توسعه مدل هیبرید عصبی موجک، ابتدا هر یک از سری های زمانی، با استفاده از تبدیل موجک گسسته (DWT) به زیرسری های $N_{di}(t), \dots, N_{d2}(t), N_{d1}(t), N_a(t)$ و $D_{di}(t), \dots, D_{d2}(t), D_{d1}(t), D_a(t)$ و $Q_{di}(t), \dots, Q_{d2}(t), Q_{d1}(t), Q_a(t)$ تجزیه می شوند؛ که $N_a(t), D_a(t)$ و $Q_a(t)$ به ترتیب نشان دهنده زیرسری های تقریب نیترات، اکسیژن محلول و دبی،

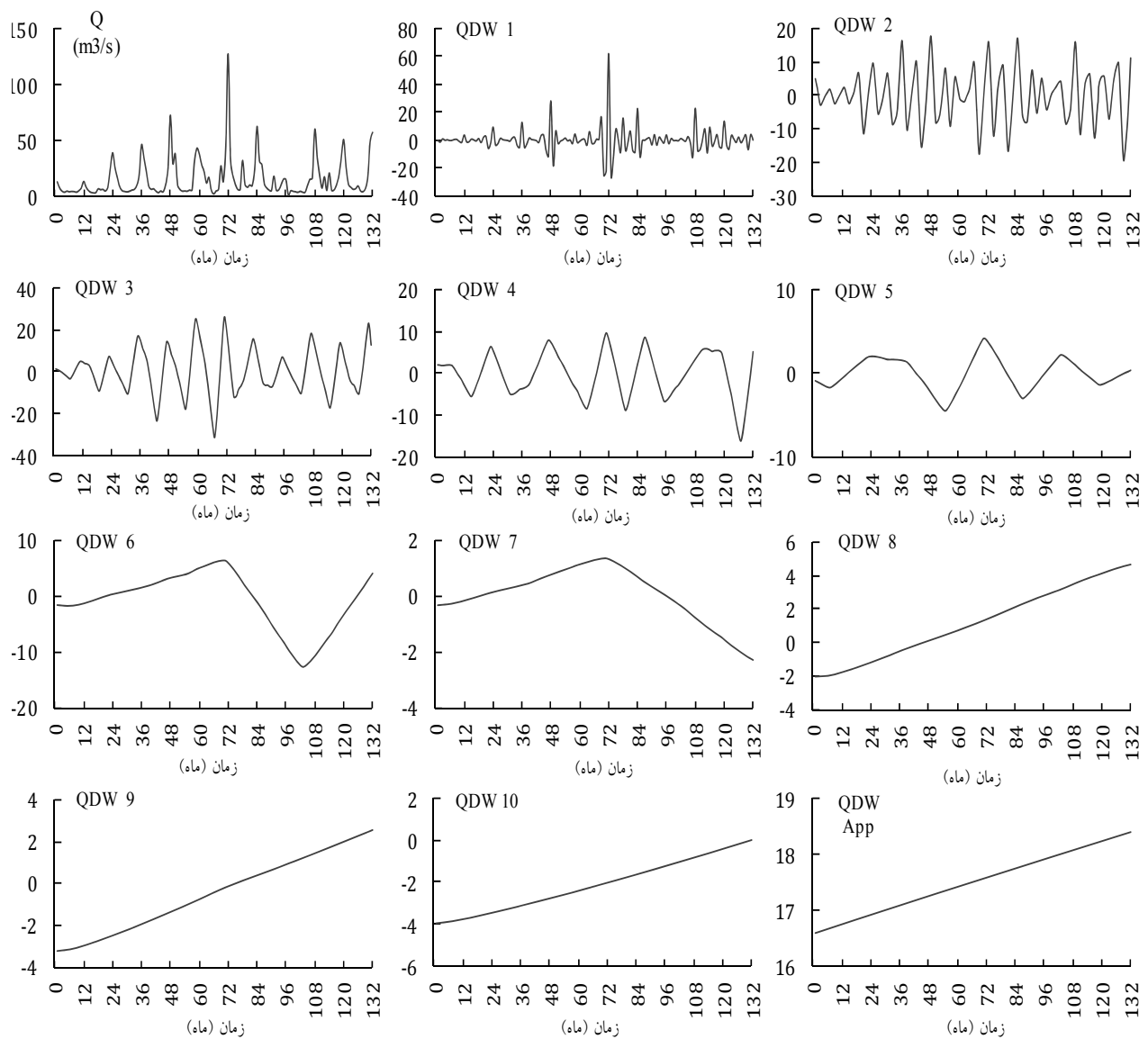
$N_{di}(t), \dots, N_{d2}(t), N_{d1}(t)$ و $D_{di}(t), \dots, D_{d2}(t), D_{d1}(t)$ و $Q_{di}(t), \dots, Q_{d2}(t), Q_{d1}(t)$ به ترتیب نشان دهنده زیرسری های جزئیات نیترات، اکسیژن محلول و دبی و di معرف سطح تجزیه هستند. این زیرسری ها نقش های مختلفی در سری زمانی اصلی ایفا می کنند، لذا رفتار هر یک متمایز است. نکته مهم در انتخاب نوع موجک مادر، طبیعت رخداد پدیده و نوع سری زمانی مشاهده شده است (۲۵). در این مطالعه به لحاظ ماهیت سری های زمانی نیترات و اکسیژن محلول، که ساختاری غیرخطی و نامنظم دارند؛ از موجک coif2 به عنوان موجک مادر و سطح تجزیه ۱۰ جهت تجزیه استفاده شده است. NDW1 اولین زیر سری جزئیات نیترات، به همین ترتیب NDW10 دهمین زیرسری جزئیات نیترات و NDWApp زیرسری تقریب نیترات در شکل ۲ نشان داده شده است.



شکل ۲. سری زمانی مشاهده ای و زیرسری های تجزیه شده نترات

پیش بینی پارامتر موردنظر به کارگیرد؛ از مجموع زیرسری های موثر حاصل از ضرایب موجک در سطوح مختلف که منجر به تولید سری زمانی با تغییرپذیری کم می شود، به عنوان ورودی مدل ترکیبی استفاده شود.

شکل های ۳ و ۴ زیرسری های اکسیژن محلول (DDW) و زیرسری دبی (QDW) را نمایش می دهند. هدف پژوهش حاضر در نحوه ورود بردارهای ورودی مدل ترکیبی شبکه عصبی - موجکی است. به عبارتی، به جای آنکه، سری زمانی اصلی با نویز زیاد (داده های مشاهده ای ایستگاه آبرسنجی) برای



شکل ۳. سری زمانی مشاهده‌ای و زیرسری‌های تجزیه شده دبی

شود. ضریب همبستگی (P)، بین دبی (Q) و پارامتر موردنظر (NO_3 , DO) برای n جفت در دسترس از رابطه ۷ بدست می‌آید.

$$P = \frac{\sum_{i=1}^n (Q_i - \bar{Q})(N_i - \bar{N})}{\sqrt{\sum_{i=1}^n (Q_i - \bar{Q})^2 \sum_{i=1}^n (N_i - \bar{N})^2}} \quad (7)$$

در این مرحله به منظور انتخاب زیرسری‌های موثر، ضریب همبستگی هر یک از اجزای موجک با سری زمانی اصلی محاسبه می‌شود. چون بیشتر سری‌های زمانی تجزیه شده تغییرپذیری پایینی دارند؛ لذا انتظار می‌رود باعث بهبود نتایج

همبستگی سری زمانی اکسیژن محلول با زیرسری های اکسیژن محلول و زیرسری های دبی، با تاخیرهای یک و دو ماهه در جدول ۲ ارائه شده است.

این ضریب در محدوده ۰ تا ۱ تعریف می شود که هر چه به ۱ نزدیکتر باشد؛ نشان دهنده همبستگی بالای دو پارامتر مورد ارزیابی است. ضرایب همبستگی سری زمانی نیترا با زیرسری های نیترا و زیرسری های دبی همچنین ضرایب

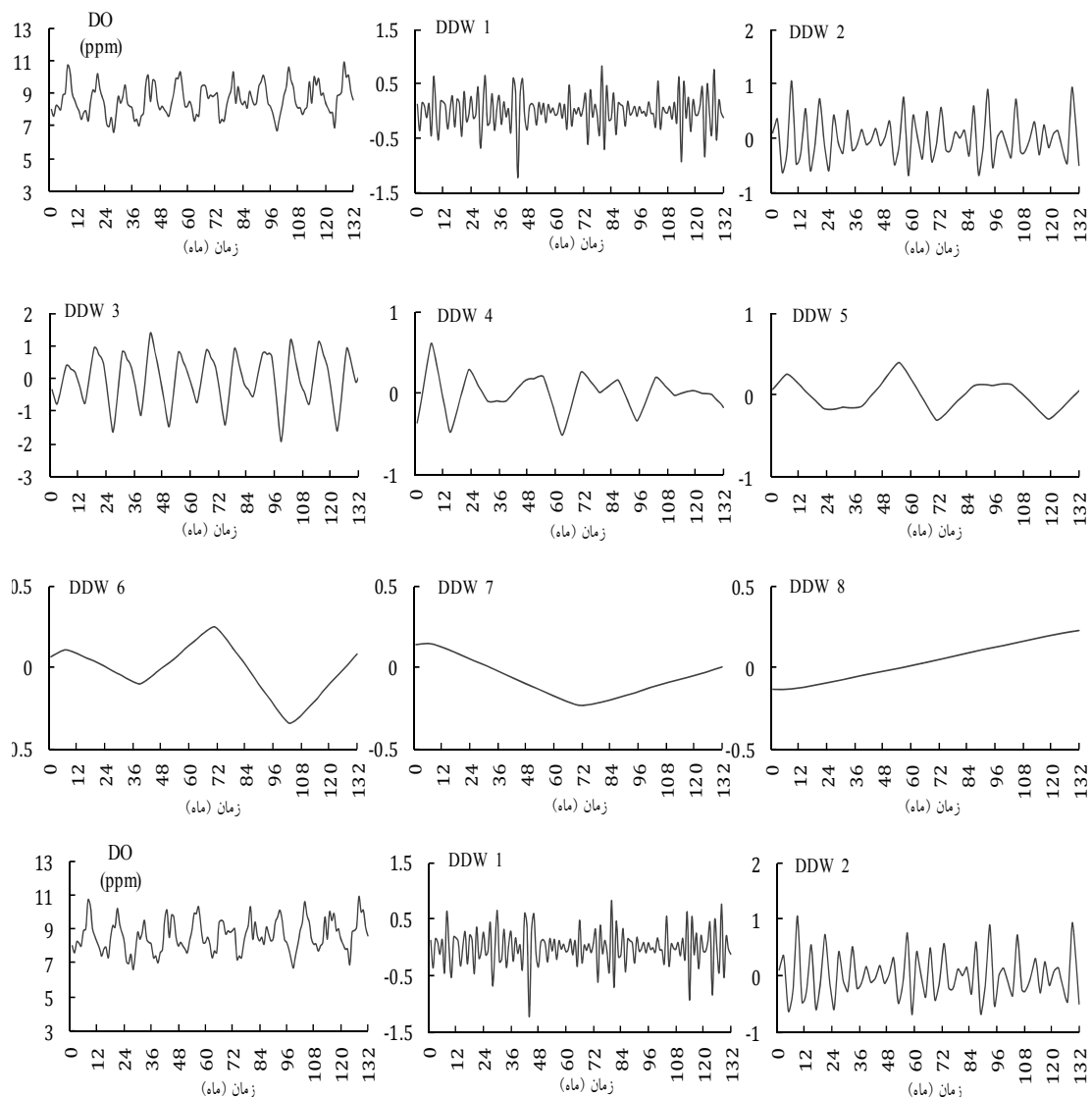
جدول ۲. ضرایب همبستگی زیرسری های هر پارامتر با سری زمانی اصلی نیترا و اکسیژن محلول

DO						NO3					
QDW t-2	QDW t-1		DDW t-2	DDW t-1		QDW t-2	QDW t-1		NDW t-2	NDW t-1	
۰/۰۰۸	-۰/۰۰۳	QDW 1	-۰/۰۱۳	-۰/۲۲۳	DDW 1	۰/۰۷۹	-۰/۰۰۴	QDW 1	۰/۰۵۶	-۰/۳۸۱	NDW 1
-۰/۲۲۴	-۰/۲۱۴	QDW 2	۰/۰۹۷	۰/۳۲۹	DDW 2	۰/۰۷۱	۰/۰۱۹	QDW 2	-۰/۱۹۱	۰/۱۴۲	NDW 2
۰/۵۰۸	۰/۲۰۱	QDW 3	۰/۳۲۷	۰/۶۳۸	DDW 3	۰/۲۳۱	۰/۲۳۶	QDW 3	۰/۱۲۳	۰/۳۳۶	NDW 3
۰/۲۰۱	۰/۱۳۵	QDW 4	۰/۱۰۶	۰/۱۸۸	DDW 4	۰/۱۶۱	۰/۱۷۸	QDW 4	۰/۲۱۶	۰/۲۳۱	NDW 4
-۰/۱۲۴	-۰/۱۳۶	QDW 5	۰/۰۷۳	۰/۱۰۲	DDW 5	۰/۰۰۶	۰/۰۳۲	QDW 5	۰/۲۸۵	۰/۲۹۴	NDW 5
-۰/۰۱۲	-۰/۰۱۹	QDW 6	۰/۰۲۸	۰/۰۱۹	DDW 6	-۰/۲۸۲	-۰/۲۹۴	QDW 6	۰/۳۳۸	۰/۳۴۹	NDW 6
-۰/۱۱۹	-۰/۱۲۵	QDW 7	-۰/۰۵۰	-۰/۰۵۱	DDW 7	-۰/۵۱۰	-۰/۵۲۳	QDW 7	۰/۲۶۰	۰/۲۵۸	NDW 7
۰/۱۶۶	۰/۱۷۲	QDW 8	۰/۱۶۷	۰/۱۷۲	DDW 8	۰/۴۴۶	۰/۴۶۱	QDW 8	۰/۴۴۵	۰/۴۶۱	NDW 8
۰/۱۶۹	۰/۱۷۴	QDW 9	۰/۱۷۰	۰/۱۷۵	DDW 9	۰/۴۳۶	۰/۴۵۲	QDW 9	۰/۴۴۱	۰/۴۵۷	NDW 9
۰/۱۷۱	۰/۱۷۶	QDW 10	۰/۱۷۱	۰/۱۷۷	DDW 10	۰/۴۵۴	۰/۴۷۰	QDW10	۰/۴۵۷	۰/۴۷۳	NDW 10
۰/۱۶۷	۰/۱۷۳	QDW App.	۰/۱۶۷	۰/۱۷۳	DDW App.	۰/۴۳۸	۰/۴۵۳	QDW App.	۰/۴۳۸	۰/۴۵۴	NDW App.

QDWApp و QDW4, QDW8, QDW9, QDW10 همبستگی بالایی با سری زمانی نیترا و اکسیژن محلول را نشان می دهند، لذا مجموع آنها سری زمانی TQDW را تولید می کند.

در شکل ۵ سری های زمانی TNDW, TDDW و TQDW نشان داده شده است. ساختار شماتیک مدل تلفیقی شبکه عصبی موجکی در شکل ۶ نمایش داده شده است. مطابق این ساختار، ورودی های اصلی به برنامه، شامل دبی (Q_t) و پارامتر مورد نظر (N_t) جهت پیش بینی است. این سیگنال ها با استفاده از موجک به زیرسری های جزئیات تجزیه شده؛ سپس مجموع زیرسری های موثر (TQDW و NDWT) به عنوان ورودی مدل استفاده می شوند.

طبق این جدول، تنها NDW1 همبستگی پایینی با نیترا از خود نشان می دهد، بنابراین زیرسری های NDW2, NDW3, NDW4, NDW5, NDW6, NDW7, NDW8, NDW9, NDW10 و NDWApp به عنوان اجزای موجک موثر انتخاب شدند و مجموع آنها سری زمانی TNDW با نویز کمتر نسبت به سری زمانی نیترا را تولید کرد. به همین ترتیب DDW1, DDW6 و DDW7 همبستگی پایینی با سری زمانی اکسیژن محلول داشته، بنابراین، DDW2, DDW3, DDW4, DDW5, DDW8, DDW9, DDW10 و DDWApp از زیرسری های موثر انتخاب شدند که با جمع کردن آنها، سری زمانی TDDW با نویز کمتر در مقایسه با سری زمانی اکسیژن محلول حاصل شد. مطابق این جدول، زیر سری های QDW3



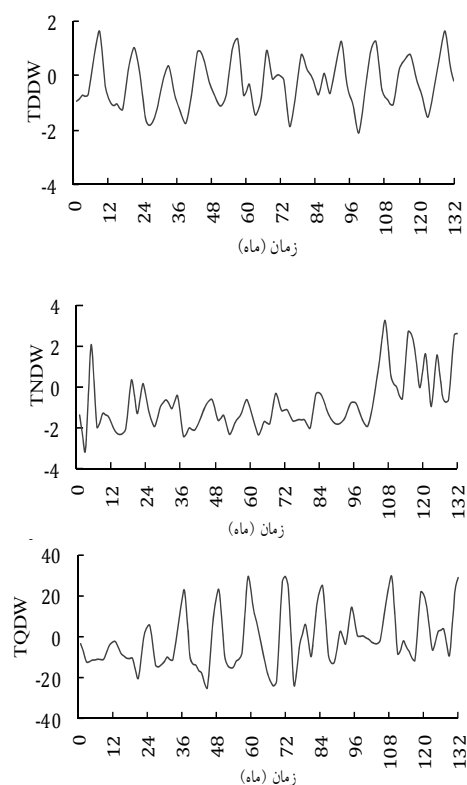
شکل ۴. سری زمانی مشاهده ای و زیرسری های تجزیه شده اکسیژن محلول

مقادیر بالایی را نشان می دهد؛ همچنین نیترات با TNDW و اکسیژن محلول با TDDW نسبت به خودهمبستگی نیترات و اکسیژن محلول (جدول ۱) همبستگی بهتری را در بردارد. از نتایج ضرایب همبستگی بدست آمده در جدول های ۳ و ۱ انتظار می رود، تبدیل موجک ابزاری مفید در جهت بهبود پیش بینی پارامترها باشد. ترکیبات مختلف ورودی مدل شبکه عصبی و مدل ترکیبی شبکه عصبی-موجکی برای هریک از پارامترها جهت پیش بینی یک گام زمانی نیترات (N_{t+1}) و اکسیژن محلول (D_{t+1}) در جدول ۴ ارائه شده است.

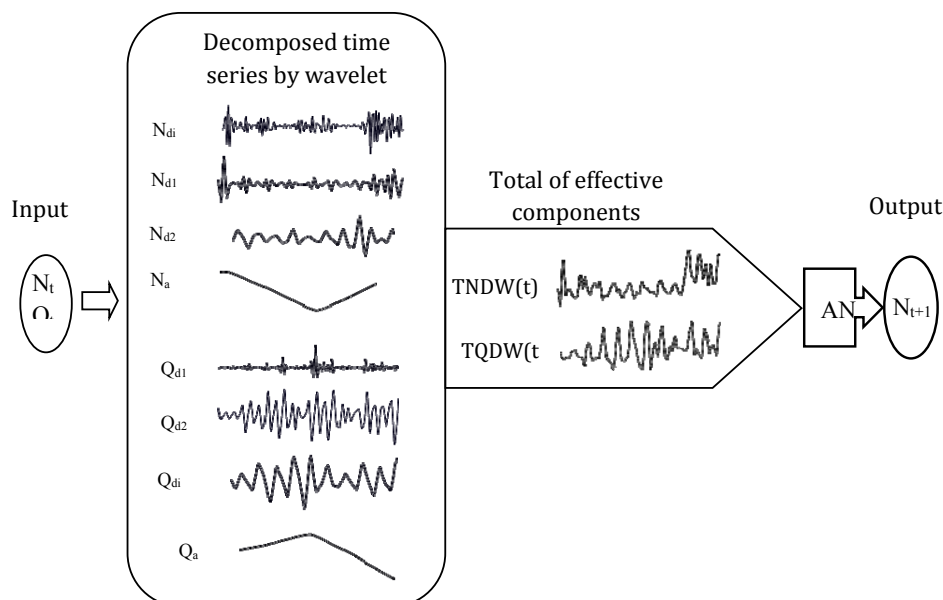
جهت دستیابی به ترکیبات ورودی مناسب مدل های شبکه عصبی، ترکیبی شبکه عصبی موجکی و رگرسیون خطی چند متغیره بایستی خودهمبستگی و همبستگی متقابل پارامترها مورد ارزیابی قرار گیرد (۱۴ و ۲۴). به همین منظور ضرایب همبستگی نیترات با دبی، TSDW و TQDW، اکسیژن محلول با دبی، TDDW و TQDW محاسبه شده و در جدول ۳ آمده است. طبق این جدول، ضرایب همبستگی نیترات و اکسیژن محلول با TQDW در مقایسه با ضرایب همبستگی نیترات و اکسیژن محلول با Q در تاخیرهای یک و دو ماهه

جدول ۳. ضرایب همبستگی NO_3 با Q ، $TNDW$ و $TQDW$ و DO با Q ، $TNDW$ و $TQDW$

ضریب همبستگی با DO		ضریب همبستگی با NO_3	
-۰/۱۱۶	$Q(t)$	-۰/۱۵۹	$Q(t)$
-۰/۰۹۸	$Q(t-1)$	-۰/۲۳۳	$Q(t-1)$
-۰/۳۰۵	$Q(t-2)$	-۰/۲۸۵	$Q(t-2)$
-۰/۴۴۷	$Q(t-3)$	-۰/۱۹۸	$Q(t-3)$
-۰/۵۰۹	$Q(t-4)$	-۰/۰۹۰	$Q(t-4)$
-۰/۹۱۴	$TDDW(t)$	-۰/۸۷۰	$TNDW(t)$
-۰/۷۶۴	$TDDW(t-1)$	-۰/۷۲۱	$TNDW(t-1)$
-۰/۳۹۹	$TDDW(t-2)$	-۰/۴۶۵	$TNDW(t-2)$
-۰/۰۱۰	$TDDW(t-3)$	-۰/۳۵۱	$TNDW(t-3)$
-۰/۳۳۴	$TDDW(t-4)$	-۰/۳۳۴	$TNDW(t-4)$
-۰/۰۱۹	$TQDW(t)$	-۰/۴۱۲	$TQDW(t)$
-۰/۲۸۶	$TQDW(t-1)$	-۰/۴۵۲	$TQDW(t-1)$
-۰/۵۵۳	$TQDW(t-2)$	-۰/۴۳۲	$TQDW(t-2)$
-۰/۶۹۲	$TQDW(t-3)$	-۰/۳۸۲	$TQDW(t-3)$
-۰/۶۵۴	$TQDW(t-4)$	-۰/۳۲۴	$TQDW(t-4)$



شکل ۵. مجموع زیرسری های موثر نترات ($TNDW$)، اکسیژن محلول ($TDDW$) و دی ($TQDW$)



شکل ۶. ساختار مدل ترکیبی شبکه عصبی - موجک

جدول ۴. ترکیبات ورودی مدل های شبکه عصبی، شبکه عصبی-موجکی و رگرسیون خطی چند متغیره

شماره	ترکیبات ورودی NO ₃		ترکیبات ورودی DO	
	ANN / MLR		WANN	
1	Nt	TNDWt	Dt	TDDWt
2	Nt , Nt-1	TNDWt , TNDWt-1	Dt , Dt-1	TDDWt , TDDWt-1
3	Nt , Nt-1 , Nt-2	TNDWt , TNDWt-1 , TNDWt-2	Dt , Dt-1 , Dt-2	TDDWt , TDDWt-1 , TDDWt-2
4	Nt , Nt-1 , Nt-2 , Nt-3	TNDWt , TNDWt-1 , TNDWt-2 , TNDWt-3	Dt-1 , Dt-2	TDDWt-1 , TDDWt-2
5	Nt , Nt-2	TNDWt , TNDWt-2	Dt , Qt-2	TDDWt , TQDWt-2
6	Nt , Qt	TNDWt , TQDWt	Dt , Qt-3	TDDWt , TQDWt-3
7	Nt , Nt-1 , Qt	TNDWt , TNDWt-1 , TQDWt	Dt , Qt-4	TDDWt , TQDWt-4
8	Nt , Nt-1 , Qt , Qt-1	TNDWt , TNDWt-1 , TQDWt , TQDWt-1	Dt , Dt-1 , Qt-2	TDDWt , TDDWt-1 , TQDWt-2
9	Nt , Qt , Qt-1	TNDWt , TQDWt , TQDWt-1	Dt , Dt-1 , Qt-3	TDDWt , TDDWt-1 , TQDWt-3
10	Nt , Nt-1 , Nt-2 , Qt	TNDWt , TNDWt-1 , TNDWt-2 , TQDWt	Dt , Dt-1 , Qt-3	TDDWt , TDDWt-1 , TQDWt-4

$$N_{normal} = \frac{(N - N_{min})}{(N_{max} - N_{min})} \quad (8)$$

برای بدست آوردن بهترین نتایج، ترکیبات گوناگون مدل شبکه عصبی بر اساس تعداد لایه های مخفی، تعداد نرون های لایه مخفی، آموزش داده شد و مورد آزمون قرار گرفت. تعداد لایه های مخفی و تعداد نرون های لایه مخفی هر شبکه، با استفاده از داده های موجود و روش سعی و خطا بهینه شد. نرون های لایه ورودی برابر تعداد ورودی در ترکیبات مختلف و تعداد نرون در لایه خروجی همواره یک نرون در نظر گرفته شد. باید توجه داشت که لایه های مخفی زیادتر و تعداد نرون بیشتر در لایه مخفی، اگرچه ممکن است دقت محاسبات در آموزش شبکه را بالا ببرد؛ ولی الزاما دقت مدل را در پیش بینی داده های اعتبارسنجی و آزمون افزایش نخواهد داد (۱۹). 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 (Over Training) خطای مجموعه آموزشی حداقل می شود؛ اما برای مرحله آزمون افزایش می یابد؛ لذا توجه به خطای هر سه مرحله امری ضروری است. روند آموزش شبکه تا زمانی که خطای مربع متوسط (MSE) در مراحل آموزش و

در دو مدل جداگانه برای پیش بینی یک ماه آینده نیترا (N_{t+1}) و اکسیژن محلول (D_{t+1}) از آمار ۱۱ ساله مربوط به ایستگاه پل خواب استفاده شد. تمامی عملیات کدنویسی مدل شبکه عصبی مصنوعی و هیبرید شبکه عصبی موجکی در محیط نرم افزار Matlab2013 انجام گرفت. در اجرای مدل رگرسیون خطی چند متغیره در محیط نرم افزار IBM SPSS Statistics، پارامتر مورد نظر برای پیش بینی در ماه فعلی، به عنوان متغیر وابسته و هریک از ترکیبات ورودی به عنوان متغیر مستقل لحاظ شدند.

آماده سازی داده های ورودی به شبکه یکی از نکات مهم در استفاده از مدل های هوشمند است. استفاده از یک روش خطی برای پیش پردازش داده ها روشی معمول است که باعث بهبود عملکرد آموزش شبکه می شود (۲۶). این پیش پردازش، شامل تبدیل داده ها از دامنه تغییرات واقعی خود به دامنه ای است که شبکه عصبی در آن دامنه، کارایی بهتری دارد (۲۶ و ۲۷). در این تحقیق، جهت حذف بعد پارامترها، داده ها را قبل از ورود به شبکه با استفاده از رابطه ۸ نرمالایز کرده، در این صورت، مقادیر بین ۰ تا ۱ را در برمی گیرند. در این رابطه N_{normal} معرف داده نرمالیزه شده، N داده واقعی، N_{min} و N_{max} به ترتیب، مقادیر کمینه و بیشینه سری زمانی است.

اعتبارسنجی حداقل گردد، ادامه می یابد.

نتایج پیش بینی یک ماه آینده مدل شبکه عصبی، تلفیق شبکه عصبی موجکی و رگرسیون خطی چند متغیره، برای نیترات و اکسیژن محلول به ترتیب در جدول های ۵ و ۶ ارائه شده است. طبق این جداول، بهترین نتیجه در اجرای مدل شبکه عصبی، برای پارامتر کیفی نیترات، ترکیب ورودی ۱ (N_t) و ساختار

۱-۶-۱ (یک نرون در لایه ورودی، ۶ نرون در لایه مخفی و یک نرون در لایه خروجی) و مقادیر E و RMSE به ترتیب ۰/۷۸۲ و ۰/۶۹۱ ppm؛ و برای پارامتر اکسیژن محلول، ترکیب ۱۰ (D_t, D_{t-1}, Q_{t-4}) با ساختار ۱-۴-۳ (سه نرون در لایه ورودی، ۴ نرون در لایه پنهان و یک نرون در لایه خروجی) و مقادیر E و RMSE ۰/۷۵۶ و ۰/۵۳ ppm حاصل شد.

جدول ۵. نتایج پیش بینی نیترات مدل های شبکه عصبی، ترکیبی شبکه عصبی موجکی و رگرسیون خطی چند متغیره برای دوره آزمون

DO										پارامتر
۱۰	۹	۸	۷	۶	۵	۴	۳	۲	۱	ترکیبات ورودی
۴	۶	۳	۷	۵	۳	۲	۵	۷	۸	نرون
۰/۹۱۴	۰/۸۶۱	۰/۸۴۰	۰/۸۸۲	۰/۸۳۴	۰/۸۸۷	۰/۸۲۲	۰/۸۶۷	۰/۸۲۷	۰/۸۷۶	E WANN
۰/۳۱۵	۰/۴۲۹	۰/۴۶۸	۰/۳۸۶	۰/۴۶۹	۰/۳۶۰	۰/۴۷۳	۰/۳۹۱	۰/۴۴۶	۰/۳۷۸	RMSE (ppm)
۴	۳	۶	۴	۵	۵	۳	۷	۲	۵	نرون
۰/۷۵۶	۰/۷۳۴	۰/۷۴۵	۰/۷۳۵	۰/۷۰۳	۰/۷۴۰	۰/۶۵۸	۰/۷۲۹	۰/۷۱۹	۰/۷۳۷	E ANN
۰/۵۳۰	۰/۵۵۴	۰/۵۴۳	۰/۵۵۳	۰/۵۸۶	۰/۵۴۸	۰/۶۲۸	۰/۵۵۹	۰/۵۶۹	۰/۵۵۱	RMSE (ppm)
۰/۴۱۰	۰/۴۲۵	۰/۴۵۰	۰/۴۱۶	۰/۴۲۹	۰/۴۵۹	۰/۲۰۱	۰/۴۶۲	۰/۳۹۱	۰/۴۰۲	E MLR
۰/۸۲۵	۰/۸۱۴	۰/۷۹۷	۰/۸۲۱	۰/۸۱۲	۰/۷۹۰	۰/۹۶۰	۰/۷۸۸	۰/۸۳۸	۰/۸۳۱	RMSE (ppm)

جدول ۶. نتایج پیش بینی اکسیژن محلول مدل های شبکه عصبی، ترکیبی شبکه عصبی موجکی و رگرسیون خطی چند متغیره برای دوره آزمون

NO ₃										پارامتر
۱۰	۹	۸	۷	۶	۵	۴	۳	۲	۱	ترکیبات ورودی
۸	۶	۳	۴	۴	۳	۷	۴	۴	۶	نرون
۰/۸۵۸	۰/۸۷۵	۰/۸۶۷	۰/۸۴۴	۰/۸۱۸	۰/۸۵۹	۰/۸۷۶	۰/۸۸۳	۰/۸۹۳	۰/۹۱۰	E WANN
۰/۵۵۹	۰/۵۲۳	۰/۵۳۵	۰/۵۸۵	۰/۶۳۲	۰/۵۵۷	۰/۵۲۳	۰/۵۰۸	۰/۴۸۳	۰/۴۴۵	RMSE (ppm)
۴	۳	۵	۶	۳	۵	۵	۳	۵	۶	نرون
۰/۶۵۳	۰/۷۰۶	۰/۷۲۶	۰/۶۷۶	۰/۶۵۱	۰/۶۸۶	۰/۷۱۴	۰/۷۱۸	۰/۷۴۴	۰/۷۸۲	E ANN
۰/۸۷۳	۰/۸۰۴	۰/۷۷۶	۰/۸۴۴	۰/۸۷۶	۰/۸۳۱	۰/۷۹۲	۰/۷۸۸	۰/۷۴۹	۰/۶۹۱	RMSE (ppm)
-۰/۶۳۸	-۰/۸۶۰	-۰/۶۹۳	-۰/۷۰۳	-۰/۸۵۸	-۰/۶۲۲	-۰/۶۱۲	-۰/۵۵۵	-۰/۶۲۰	-۰/۸۰۸	E MLR
۱/۸۹۷	۲/۰۲۲	۱/۹۲۹	۱/۹۳۴	۲/۰۲۱	۱/۸۸۷	۱/۸۸۲	۱/۸۴۸	۱/۸۸۶	۱/۹۹۳	RMSE (ppm)

پیش بینی نیترات با روش MLR غیر قابل قبول بوده، در صورتی که پیش بینی اکسیژن محلول با ضریب نش و RMSE به ترتیب ۰/۴۵۹ و ۰/۷۹۰ ppm با ترکیب ورودی ۵ (D_t, Q_{t-1}) همراه بود. در اجرای مدل ترکیبی شبکه عصبی - موجکی از مجموع

برای مدل رگرسیون خطی چند متغیره، آنالیز رگرسیون خطی بین پارامتر کیفی موردنظر (نیترات و اکسیژن محلول) در ماه فعلی و ترکیبات ۱۰ گانه صورت گرفت؛ سپس با یافتن ثابت های رگرسیونی، پیش بینی یک ماه آینده پارامتر کیفی در دوره آزمون، انجام شد.

جدول ۷. ارزیابی مدل ها برای پیش بینی مقادیر نترات بیشتر از ۵/۱۳ ppm مربوط به دوره آزمون

ردیف	تاریخ	مقادیر واقعی	پیش بینی ANN (ppm)	پیش بینی MLR (ppm)	پیش بینی WANN (ppm)
۱	خرداد ۸۹	۶/۵	۵/۴۸۲	۴/۰۵۲	۶/۵۳۸
۲	تیر ۸۹	۵/۷	۵/۱۵۰	۴/۴۰۸	۵/۷۲۰
۳	شهریور ۸۹	۶/۰	۴/۹۹۷	۳/۹۳۳	۵/۹۶۴
۴	اسفند ۸۹	۶/۵	۵/۸۴۷	۳/۶۱۲	۶/۵۴۰
۶	فروردین ۹۰	۶/۶	۵/۷۸۹	۴/۴۶۵	۶/۶۰۹
۷	اردیبهشت ۹۰	۶/۷	۵/۹۸۳	۵/۱۹۳	۶/۷۸۸
جمع		۳۸/۰	۳۳/۲۴۸	۲۵/۳۶۳	۳۸/۱۵۹
درصدخطا			-۱۴/۳۰	-۴۹/۸۰	-۰/۴۲۰
RMSE (ppm)			۰/۸۱۱	۲/۱۷۹	۰/۰۴۶

جدول ۸. ارزیابی مدل ها برای پیش بینی مقادیر اکسیژن محلول بیشتر از ۸/۸۴ ppm مربوط به دوره آزمون

ردیف	تاریخ	مقادیر واقعی	پیش بینی ANN (ppm)	پیش بینی MLR (ppm)	پیش بینی WANN (ppm)
۱	آبان ۸۹	۸/۹	۸/۹۵۱	۸/۷۵۶	۹/۰۶۰
۲	آذر ۸۹	۹/۱	۹/۰۸۳	۹/۳۴۱	۹/۱۴۰
۳	دی ۸۹	۱۰/۹	۹/۵۰۹	۸/۸۹۸	۱۰/۷۱۴
۴	بهمن ۸۹	۱۰/۰	۹/۳۵۶	۱۰/۱۱۱	۹/۵۴۵
۶	اسفند ۸۹	۱۰/۱	۹/۴۲۲	۹/۶۶۴	۹/۵۴۱
۷	فروردین ۹۰	۹/۱	۸/۹۰۷	۹/۰۶۱	۹/۱۲۹
جمع		۵۸/۱	۵۵/۲۲۸	۵۵/۹۲۱	۵۷/۱۲۹
درصدخطا			-۵/۲۰	-۳/۷۰	-۱/۷۰
RMSE (ppm)			۰/۶۸۹	۰/۸۱۰	۰/۳۱۱

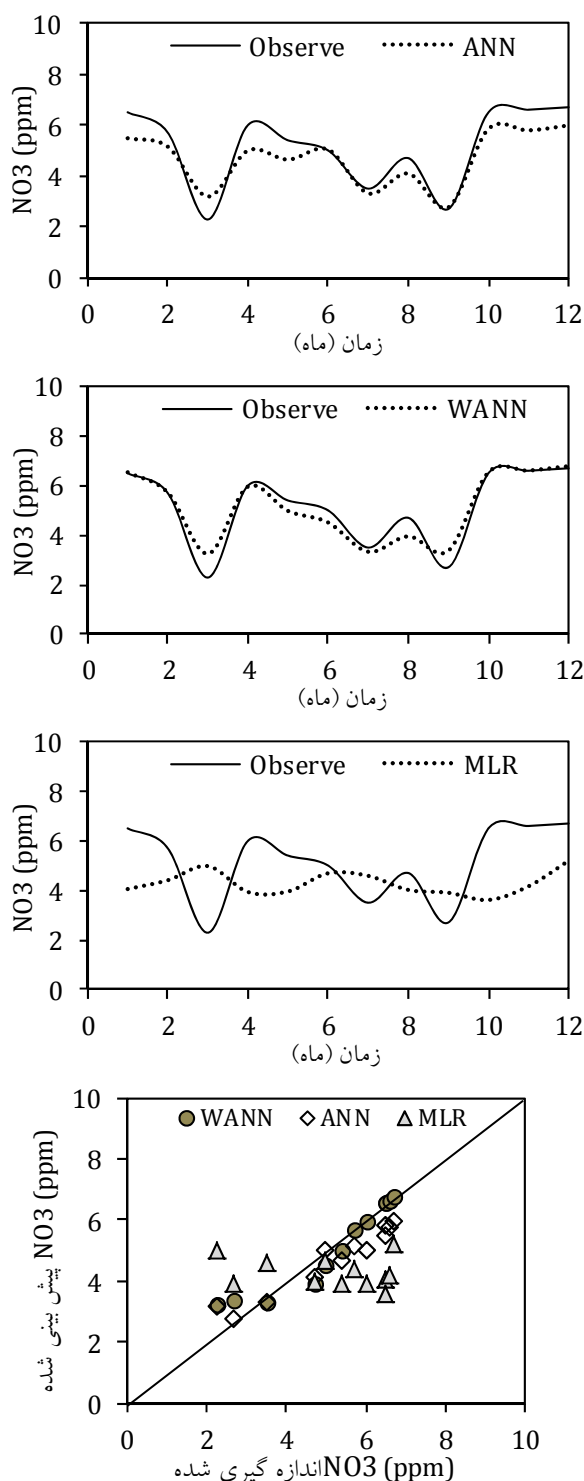
زیرسری های موثر با ترکیبات ورودی مختلف (جدول ۴) به عنوان بردار ورودی استفاده شد و پارامتر کیفی مورد نظر در گام زمانی $t+1$ به عنوان خروجی مدل حاصل شد. مطابق جداول ۵ و ۶ نتیجه مطلوب در اجرای این مدل برای پیش بینی پارامتر نترات ترکیب ورودی $(TNDW_t)$ و ساختار ۱-۶-۱ (یک نرون در لایه ورودی، ۶ نرون در لایه مخفی و یک نرون در لایه خروجی) و مقادیر E و RMSE به ترتیب ۰/۹۱ و ۰/۴۴۵ ppm؛ و برای پارامتر اکسیژن محلول، ترکیب ۱۰-۴-۱ (سه نرون در لایه ورودی، ۴ نرون در لایه پنهان و یک نرون در لایه خروجی) و مقادیر E و RMSE ۰/۹۱۴ و ۰/۳۱۵ ppm بدست آمد. به طور کلی نتایج حاکی از قابلیت بالای مدل ترکیبی شبکه عصبی موجکی در مقایسه با مدل شبکه عصبی و رگرسیون خطی چند متغیره در پیش بینی دو پارامتر مذکور است. همچنین عملکرد بهتر شبکه عصبی مصنوعی نسبت به رگرسیون خطی چند متغیره به وضوح مشخص است.

ارزیابی مدل ها در پیش بینی نقاط پیک سری زمانی
از ویژگی های یک مدل هیدرولوژیکی خوب، توانایی آن در تخمین و شبیه سازی نقاط پیک سری زمانی است. زیرا برخی کارهای طراحی و اصلاحات لازم در کیفیت آب های سطحی معمولاً براساس نقاط ماکزیمم انجام می شود (۲۸). به لحاظ اینکه مقدار میانگین یون نترات و اکسیژن محلول در دوره آزمون، به ترتیب ۵/۱۳ و ۸/۸۴ ppm هستند؛ لذا در این قسمت از پژوهش، به ارزیابی مقادیر بالاتر از میانگین هر پارامتر، و ماه هایی 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 می شود. جداول های ۷ و ۸، به ترتیب عملکرد هر سه مدل برای پیش بینی نقاط پیک سری زمانی نترات و اکسیژن محلول را نشان می دهند. مطابق این جداول در پیش بینی هر دو پارامتر، مدل WANN با درصد خطا و RMSE کمتر نسبت به دو مدل دیگر توانسته است نقاط پیک سری زمانی را پیش بینی کند.

بحث

در این تحقیق به توسعه شبکه های عصبی مصنوعی، هیبرید شبکه عصبی موجکی و رگرسیون خطی چند متغیره برای پیش بینی یک ماه آینده پارامتر نیترا و اکسیژن محلول ایستگاه پل خواب رودخانه کرج پرداخته شد. در اجرای شبکه های عصبی مصنوعی مربوط به تخمین نیترا و اکسیژن محلول، به ترتیب ترکیبات ورودی ۱ و ۱۰ بهترین خروجی را داشتند. به عبارتی پیش بینی نیترا یک ماه آینده تنها به نیترا ماه فعلی وابسته است. اکسیژن محلول ماه بعد، علاوه بر میزان اکسیژن محلول ماه فعلی و یک ماه قبل، به دبی چهار ماه قبل ایستگاه نیز بستگی دارد.

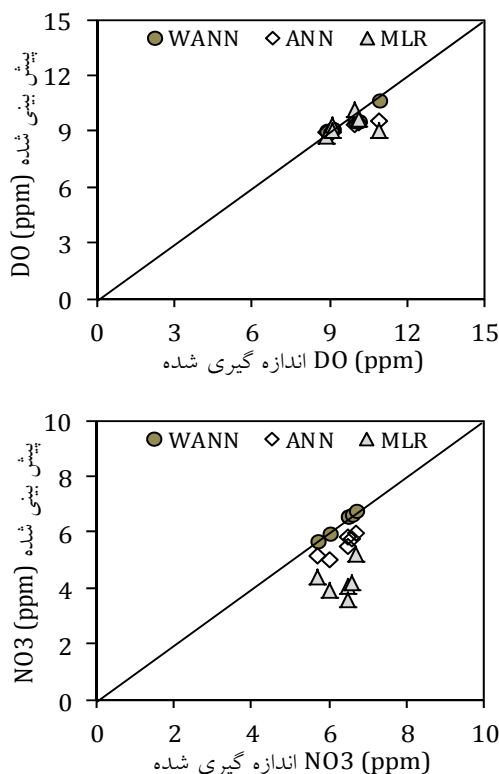
روش MLR تکنیکی خطی جهت پیش بینی سری های زمانی محسوب می شود؛ چون یون نیترا ماهیت نوسانی شدیدتری دارد (شکل ۲)؛ لذا این روش در پیش بینی این پارامتر ناتوان بوده است در صورتی که خروجی بهینه رگرسیون خطی چند متغیره در برآورد میزان DO یک ماه بعد، در ترکیب ورودی ۵ حاصل شد. در واقع اکسیژن محلول ماه آینده وابسته به اکسیژن محلول ماه فعلی و دبی دو ماه قبل بوده است. با مقایسه نتایج عملکرد سه مدل در جداول ۵ و ۶ می توان نتیجه گرفت، شبکه عصبی موجکی از کارایی مطلوب تری در تخمین این پارامترها برخوردار بوده است. این برتری به این دلیل است که مدل ترکیبی به واسطه سیگنال های تجزیه شده سری زمانی توسط موجک و جمع زیرسری های موثر توانسته است خصوصیات غیرخطی و پیچیده سری زمانی را دقیق تر از ANN و MLR شبیه سازی کند (۱۴ و ۱۹). در واقع حذف کردن برخی از زیرسری ها (که باعث ایجاد نویز زیاد در سیگنال اصلی می شدند)، در افزایش دقت و عملکرد مدل برای پیش بینی پارامترها، بسیار موثر واقع شد. شکل های ۷ و ۸، به ترتیب نیترا پیش بینی شده و اکسیژن محلول پیش بینی شده را توسط سه مدل برای دوره آزمون در مقایسه با مقادیر اندازه گیری شده هر یک از پارامترها نشان می دهند. نمودارها نمایانگر برتری و قابلیت بالای مدل ترکیبی نسبت به مدل شبکه عصبی مصنوعی و رگرسیون خطی چند متغیره است.



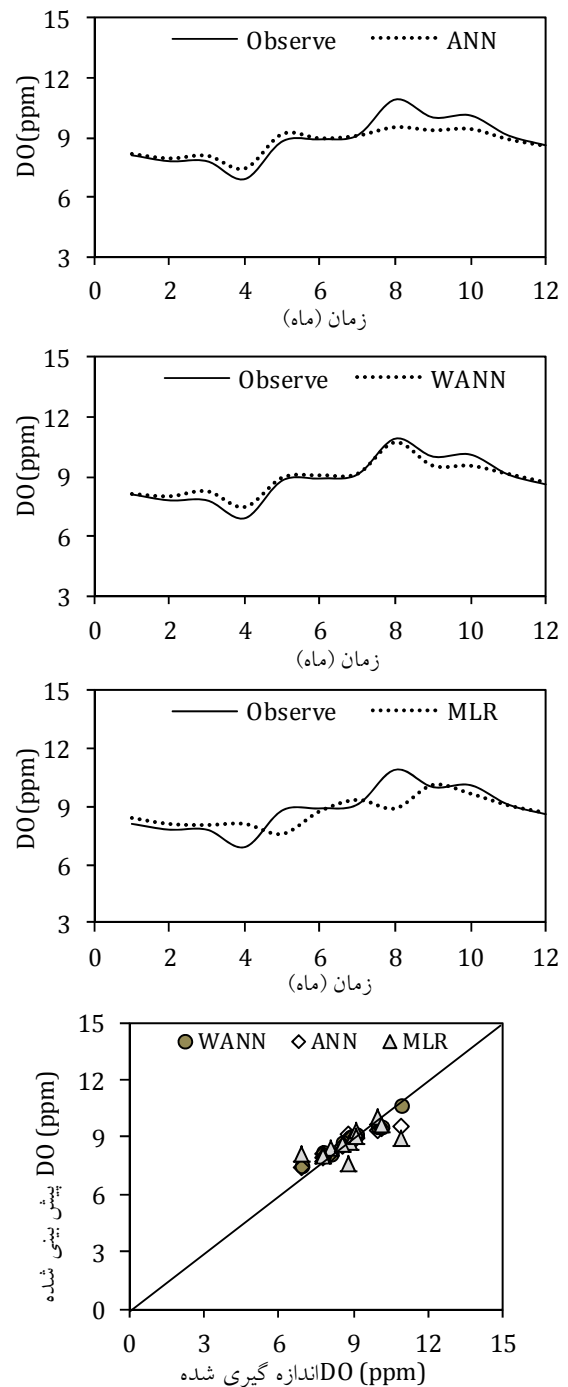
شکل ۷. مقایسه نیترا اندازه گیری شده و پیش بینی شده با ANN، WANN و MLR دوره آزمون

عصبی مصنوعی، و استخراج ویژگی های مهم ضرایب موجک گسسته، و بدست آوردن مولفه های موثر در پیش بینی پارامتر کیفی، به کار گرفته می شود؛ که نتایج حاکی از عملکرد مناسب مدل ترکیبی توسعه یافته است.

در ادامه پژوهش، عملکرد سه مدل در پیش بینی نقاط پیک سری زمانی ارزیابی شدند. شکل ۹ به مقایسه مقادیر نترات و اکسیژن محلول مشاهده و پیش بینی 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 ترکیبی شبکه عصبی-موجکی نسبت به دو مدل دیگر، رابطه ای نزدیک به ۱:۱ (حول نیمساز رسم شده) بین مقادیر اندازه گیری شده و مقادیر پیش بینی شده برقرار کرده است.



شکل ۹. مقایسه نترات و اکسیژن محلول اندازه گیری شده و پیش بینی شده با ANN، WANN و MLR برای مقادیر بیشتر از میانگین هر پارامتر در دوره آزمون



شکل ۸. مقایسه اکسیژن محلول اندازه گیری شده و پیش بینی شده با ANN، WANN و MLR در دوره آزمون

مقادیر پیش بینی شده توسط مدل ترکیبی شبکه عصبی-موجکی تطابق بیشتری با مقادیر واقعی داشته و رابطه تقریباً ۱:۱ برقرار است. در واقع WANN جهت ارزیابی هوشمند مدل شبکه

نتیجه گیری

تبدیلات موجک در حال حاضر به خوبی جایگاه خود را در مدیریت جامع حوضه های آبریز پیدا کرده است. با تلفیق شبکه عصبی مصنوعی و تئوری موجک، می توان تا حد قابل توجهی زمان محاسبات را کاهش داده و دقت پیش بینی شبکه را افزایش داد. در این تحقیق، کاربرد مدل های شبکه عصبی، رگرسیون خطی چند متغیره و مدل ترکیبی شبکه عصبی-موجکی برای مدل سازی سری زمانی یون نیترات و اکسیژن محلول که از جمله سری های نایبنا بوده و ماهیت نوسانی دارند؛ مورد بررسی قرار گرفت. مدل های هوشمندی که دارای چندین پارامتر قابل تنظیم هستند در مقایسه با مدل های خطی بهتر می توانند خصوصیات غیر خطی و پیچیده پارامترها را شبیه سازی کنند. مدل ترکیبی شبکه عصبی موجکی توانست، در هر دو پارامتر به مقدار قابل توجهی میزان E و RMSE را بهبود بخشد. پیش بینی نقاط پیک سری زمانی توسط سه مدل مربوطه ارزیابی شد و قابلیت بالای مدل ترکیبی شبکه عصبی موجکی نسبت به دو مدل دیگر حاصل شد. بنابراین، مزیت مدل ترکیبی شبکه عصبی-موجکی به کار گرفته شده در این تحقیق برای پیش بینی یون اکسیژن محلول و نیترات این است که، می توان با استفاده از اطلاعات سری زمانی پارامترهای کیفی و شرایط تأثیرگذار مانند دبی، روشی مناسب و سریع برای پیش بینی پارامترهای مذکور و در نتیجه مدیریت برتر کیفیت

منابع آب دست یافت. پیشنهاد می شود، تحقیقات بعدی جهت بررسی کارایی این مدل ها برای پیش بینی پارامترهای کیفی مختلف در دیگر حوضه های آبخیز و یا گام های زمانی مختلف (روزانه، هفتگی، سالانه و ...) انجام شود. استفاده از سری های زمانی تأثیرگذار متفاوتی مانند دما، اسیدیته و ... می تواند در توسعه مدل ترکیبی شبکه عصبی-موجک کارآمد در پیش بینی کیفیت رودخانه موثر باشد. با توجه به انعطاف پذیری بالای مدل های هوش مصنوعی و همچنین امکان طرح مدل های دیگر در جهت انتخاب معماری شبکه های عصبی در بحث پیش بینی سری های زمانی، توصیه می شود رویکردهای دیگری نیز بررسی گردد که از این بین می توان به استفاده از منطق فازی برای بررسی تأثیر آن بر نتایج مدل هیبرید شبکه عصبی موجکی و استفاده از موجک های مادر دیگری جهت تجزیه سیگنال ها در زمینه پیش بینی کیفیت آب رودخانه ها و مخازن سدها و بررسی معماری های مختلف شبکه عصبی، اشاره کرد.

تشکر و قدرانی

این پژوهش حاصل پایان نامه دوره کارشناسی ارشد با عنوان پیش بینی سری های زمانی پارامترهای کیفی رودخانه با استفاده از مدل تلفیقی شبکه عصبی و موجک (مطالعه موردی رودخانه کرج) در سال ۹۳ بوده که در دانشگاه قم به تصویب رسیده است.

منابع

- 1- Makhdoum M. Living in the Environment. Tehran: Tehran University Press; 2005 (in Persian).
- 2- Kashefiolasl M, Zaeimdar M. The need for quality management Jajrood River. Journal of Environmental Science and Technology. 2009;11(2):119-29 (in Persian).
- 3- Jahed Khaniki GR, Mahdavi M, Ghasri A, Saeednia S. Investigation of nitrate concentrations in some bottled water available in Tehran. Iranian Journal of Health and Environment. 2008;1(1):45-50 (in Persian.)
- 4- Taebi A, Vashtani M. Prediction the quality of urban runoff: Methods and selected model. Journal of the College of the Engineering. 2000;11(2):41-48 (in Persian).
- 5- Safavi HR. Prediction of river water quality by adaptive neuro fuzzy inference system (ANFIS). Journal of the Environmental Studies. 2010;36(53):1-10 (in Persian).
- 6- Nayak PC, Rao YS, Sudheer K. Groundwater level forecasting in a shallow aquifer using artificial neural network approach. Water Resources Management. 2006;20(1):77-90.
- 7- Palani S, Liong S-Y, Tkalich P. An ANN application for water quality forecasting. Marine Pollution Bulletin. 2008;56(9):1586-97.
- 8- May D, Sivakumar M. Comparison of artificial neural network and regression models in the prediction of urban stormwater quality. Water Environment Research. 2008;80(1):4-9.
- 9- Kisi Ö. Multi-layer perceptrons with Levenberg-Marquardt training algorithm for suspended sediment concentration prediction and estimation. Hydrological Sciences Journal. 2004;49(6):1025-40.
- 10- Singh KP, Basant A, Malik A, Jain G. Artificial neural network modeling of the river water quality—a case study. Ecological Modelling. 2009;220(6):888-95.
- 11- Najah A, Elshafie A, Karim OA, Jaffar O. Prediction of Johor River water quality parameters using artificial neural networks. European Journal of Scientific Research. 2009;28(3):422-35.
- 12- Wen X, Fang J, Diao M, Zhang C. Artificial neural network modeling of dissolved oxygen in the Heihe River, Northwestern China. Environmental monitoring and assessment. 2013;185(5):4361-71.
- 13- Najah A, El-Shafie A, Karim O, Jaafar O, El-Shafie AH. An application of different artificial intelligences techniques for water quality prediction. Journal of the Physical Sciences. 2011;6(22):5298-308.
- 14- Rajaei T, Nourani V, Zounemat-Kermani M, Kisi O. River suspended sediment load prediction: Application of ANN and wavelet conjunction model. Journal of Hydrologic Engineering. 2010;16(8):613-27.
- 15- Kim T-W, Valdés JB. Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. Journal of Hydrologic Engineering. 2003;8(6):319-28.
- 16- Cannas B, Fanni A, Sias G, Tronei S, Zedda MK. River flow forecasting using neural networks and wavelet analysis. Journal of European Geosciences. 2005;7:24-29.
- 17- Nourani V, Komasi M, Mano A. A multivariate ANN-wavelet approach for rainfall-runoff modeling. Water Resources Management. 2009;23(14):2877-94.
- 18- Adamowski J, Sun K. Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. Journal of Hydrology. 2010;390:85-91.
- 19- Rajaei T. Wavelet and ANN combination model for prediction of daily suspended sediment load in rivers. Science of the Total Environment. 2011;409:2917-28.
- 20- Banejad H, Kamali M, Amirmoradi K, Olyaie E. Forecasting some of the qualitative parameters of rivers using wavelet artificial neural network hybrid (W-ANN) model (Case of study: Jajroud River of Tehran and Gharaso River of Kermanshah). Iranian Journal of Health and Environment. 2014;6(3):277-94 (in Persian).
- 21- Noshadi A, Salemi H, Ahmadzade M. Simulation and prediction of some water quality parameters in the Zayanderoud river using artificial neural networks. Journal of Water and the Wastewater. 2007;64(18):49-65 (in Persian).
- 22- Dastourani MT, Azimi KH, Talebi A, Ekhtesasi MR. Estimation of suspended sediment using arti-

- ficial neural network (Case study: Kermanshah basin watershed Jamyshan). Journal of the Watershed Management Resaerch. 2013;3(6):61-74 (in Persian).
- 23- Adamowski J, Chan HF. A wavelet neural network conjunction model for groundwater level forecasting. Journal of Hydrology. 2011;407(1):28-40.
- 24- Fernando AK, Kerr T. Runoff forecasting with an artificial neural network model. Proceedings of 3rd Pacific Conference on Stormwater and Aquatic Resource Protection; 2003; Auckland, New Zealand.
- 25- Nourani V, Hasanzade Y, Komasi M, Sharafi A. Modeling of rainfall - runoff by using wavelet – neural network model. Proceedings of 4th National Congress on Civil Engineering; 2008; University of Tehran, Tehran (in Persian).
- 26- Imrie C, Durucan S, Korre A. River flow prediction using artificial neural networks: generalisation beyond the calibration range. Journal of Hydrology. 2000;233(1):138-53.
- 27- Masters T. Practical Neural Network Recipes in C++. San Diego: Academic Press; 1993.
- 28- Ott WR. Environmental Indices: Theory and Practice. USA: Ann Arbor Science; 1979.

Prediction of quality parameters (NO₃, DO) of Karaj River using ANN, MLR, and Denoising-based combined wavelet-neural network based on Models

T. Rajaei¹, R. Rahimi Benmaran^{1*}, H. Jafari¹

¹ Department of Civil Engineering, Faculty of Engineering, University of Qom, Qom, Iran

Received: 22July 2014 ; Accepted: 22October 2014

ABSTRACT

Background & Objectives: The prediction and quality control of the 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), combined wavelet-neural network (WANN), and multi linear regression (MLR) models were evaluated to predict next month nitrate and dissolved oxygen of “Pole Khab” station located in Karaj River.

Materials and Methods: A statistical period of 11 years was used for the input of the models. In combined WANN model, the real monthly-observed time series of river discharge (Q) and the quality parameters (nitrate and dissolved oxygen) were analyzed using wavelet analyzer. Then, their completely effective time series were used as ANN input. In addition, the ability of all three models were investigated in order to predict the peak points of time-series that have great importance. The capability of the models was evaluated by coefficient of efficiency (E) and the root mean square error (RMSE).

Results: The research findings indicated that the accuracy and the ability of hybrid model of wavelet neural network with the attitude of eliminations of time series noise 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 neural network and multiple linear regression models respectively, amounting to 35.6% and 75.92%, for Dissolved Oxygen ion as much as 40.57% and 60.13%.

Conclusion: owing to the high capability wavelet neural network and the elimination of the time series noises 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, Neural Network, Wavelet analysis, Nitrate and Dissolved Oxygen ions, Denoising.

*Corresponding Author: r.rahimi_b@yahoo.com

Mob: +989359530798