

## مدل ترکیبی موجک و شبکه عصبی برای پیش‌بینی غلظت سولفات و کلراید رودخانه‌ها (مطالعه موردی: رودخانه دز)

طاهر رجایی\*<sup>۱</sup>، حمیده جعفری<sup>۲</sup>

تاریخ دریافت: ۱۳۹۳/۲/۲ تاریخ پذیرش: ۱۳۹۳/۶/۵

### چکیده

از مهم‌ترین عوامل مؤثر در توسعه پایدار یک منطقه، فراهم آوردن منابع آب مناسب از نظر کمی و کیفی می‌باشد. در این مطالعه مدل‌های شبکه عصبی مصنوعی، رگرسیون خطی چند متغیره و ترکیب شبکه عصبی-موجک برای پیش‌بینی غلظت سولفات و کلراید در ایستگاه بامدژ رودخانه دز بررسی شد و تأثیر پیش‌پردازش داده‌ها روی عملکرد مدل‌ها به وسیله تجزیه موجک مورد تحقیق قرار گرفت. بدین منظور سری‌های زمانی دبی، سولفات و کلراید مشاهده شده رودخانه، در سطوح مختلفی به زیرسری‌ها تجزیه شدند. سپس حاصل جمع جبری زیرسری‌های مؤثر دبی، سولفات و کلراید به عنوان ورودی مدل شبکه عصبی برای پیش‌بینی یک ماه بعدی سولفات و کلراید در نظر گرفته شد. نتایج نشان داد که عملکرد پیش‌بینی مدل ترکیبی موجک نسبت به شبکه عصبی و رگرسیون بهتر می‌باشد. همچنین مدل ترکیبی موجک عصبی برای مقادیر بیشینه، پیش‌بینی قابل قبولی را ارائه داد. این مدل، میانگین حداقل خطا را برای مدل‌های رگرسیون و شبکه عصبی در مدل‌سازی کلراید، به ترتیب از ۰/۸۴ و ۰/۶۴ به ۰/۵۲ و در مدل‌سازی سولفات به ترتیب از ۱/۷ و ۰/۹۵ به ۰/۶۳ کاهش داد. همچنین این مدل توانست پدیده هیستریزیس را برای مدل سولفات به خوبی شبیه‌سازی کند، درحالی که مدل رگرسیون در تشخیص آن ناتوان بود.

**واژه‌های کلیدی:** تبدیل موجک، رودخانه دز، سولفات، شبکه عصبی، کلراید

### مقدمه

سنتی پیش‌بینی‌های هیدرولوژیکی شده، استفاده از شبکه‌های عصبی مصنوعی می‌باشد. امروزه این مدل‌ها در طیف وسیعی از تحقیقات استفاده می‌شوند. ساندهو و فینچ توانایی شبکه‌های عصبی را برای پیش‌بینی میزان شوری، میزان آنیون‌هایی مانند سولفات و کلراید، مقدار هدایت الکتریکی و میزان جامدات محلول در آب این حوضه را تأیید کردند (Sandhu and Finch., 1995). نوشادی و همکاران (۱۳۸۶) برخی پارامترهای کیفیت آب شامل سولفات و کلراید را در رودخانه زاینده‌رود با استفاده از شبکه‌های عصبی شبیه‌سازی نمودند. نتایج نشان داد، پیش‌بینی‌های شبکه عصبی با دقت خوبی صورت گرفته است. گلابی و همکاران (۱۳۸۵) به شبیه‌سازی پارامترهای دبی، کربنات، سولفات، کلراید، هدایت الکتریکی، نسبت جذب سدیم چند ایستگاه مهم رودخانه کارون پرداختند. نجاح و همکاران توانایی شبکه‌های عصبی را در پیش‌بینی شاخص‌های کیفی آب رودخانه جوهر مالزی و توانایی برآورد پارامترهای کیفیت در این حوضه را مورد تأکید قرار دادند (Najah et al., 2009). نتایج نشان داد که مدل عصبی در پیش‌بینی این پارامترها دقت قابل قبولی دارد. ودیعتی و همکاران (۱۳۹۲) برخی از فراسنج‌های طبیعی رودخانه کارون، مانند اکسیژن محلول، نیترات، کلراید، هدایت الکتریکی و اکسیژن خواهی

رودخانه‌ها از مهم‌ترین منابع تجدیدشونده و حیاتی آب شیرین جهت استفاده در امور کشاورزی، شرب و صنعت به شمار می‌روند. از آنجا که حفاظت کیفی آب رودخانه‌ها، نیاز به سرمایه‌گذاری‌های اضافی جهت تصفیه پساب‌ها یا سیستم‌های جمع‌آوری و کنترل زه‌آب‌ها دارد، می‌تواند اثرات اقتصادی قابل توجهی داشته باشد. سولفات و کلراید به عنوان پارامترهایی مهم و مورد توجه در ارزیابی کیفی آب از آنیون‌هایی هستند که در تمام آب‌های سطحی یافت می‌شوند. در صورتی که غلظت کلراید در آب زیاد باشد، آب برای رشد گیاه سمی خواهد بود. در صورت وجود یون سولفات در آب، سازه‌های آبی می‌بایست با مصالح ضد سولفات ساخته شود. یون کلراید به دلیل تغییرپذیری شدید، انحلال‌پذیری بالا و واکنش‌ناپذیری در فرآیندهای شیمیایی در اغلب مطالعات مورد توجه است.

یکی از روش‌هایی که در چند دهه‌ی اخیر جایگزین روش‌های

۱- استادیار گروه مهندسی عمران و نویسنده مسئول، دانشگاه قم

۲- دانش‌آموخته کارشناسی ارشد سازه‌های هیدرولیکی، گروه عمران، دانشگاه قم

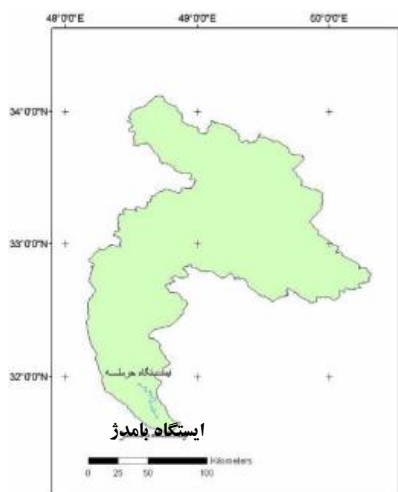
\* - نویسنده مسئول: (Email: Taher\_rajae@yahoo.com)

## مواد و روش‌ها

### منطقه مورد مطالعه

رودخانه دز یکی از شاخه‌های اصلی تشکیل دهنده رودخانه کارون، دارای پتانسیل‌های مناسبی جهت توسعه و بهره‌برداری از منابع آبی می‌باشد. حوزه آبریز رودخانه دز ۲۳۲۵۰ کیلومتر مربع و طول رودخانه از سرچشمه تا محل تلاقی با کارون حدود ۵۲۰ کیلومتر است. این حوضه از لحاظ موقعیت جغرافیایی بین ۴۸ درجه و ۱۰ دقیقه تا ۵۰ درجه و ۲۱ دقیقه طول شرقی و ۳۱ درجه و ۳۴ دقیقه تا ۳۴ درجه و ۷ دقیقه شمالی محدود شده است. رودخانه دز از دو شاخه اصلی به نام‌های دز بختیاری و دز سزار تشکیل شده است (شکل ۱). برای انجام پژوهش حاضر از داده‌های شرکت مدیریت منابع آب ایران در ایستگاه بامدژ رودخانه دز استفاده شده است.

در این تحقیق از دوره ۲۴۶ ماهه‌ی استفاده شده (فروردین ۷۰ تا شهریور ۹۰)، ۸۰٪ داده‌ها برای آموزش و ۲۰٪ داده‌ها به‌منظور آزمون مدل‌ها استفاده شد.



شکل ۱ - موقعیت ایستگاه بامدژ در حوزه رودخانه دز

### تحلیل آماری داده‌ها

هر سری زمانی ویژگی‌های رفتاری نهفته‌ای مانند خودهمبستگی، خودشباهتی، خودنسبتی و روند در بازه‌ی تغییرات مقیاسی خود دارد که به این وسیله می‌توان بر تاریخچه‌ی فرآیند آن تسلط پیدا کرده و مقدمات پیش‌بینی رفتار آن را در آینده فراهم ساخت. در جدول ۱ تحلیل آماری داده‌های ماهانه دبی، سولفات و کلراید برای کل دوره و نیز به تفکیک، برای داده‌های دوره آموزش و آزمون ارائه شده است. آنالیز ارائه شده شامل میانگین، کمینه، بیشینه، انحراف معیار ( $S_d$ )، ضریب چولگی ( $C_s$ )، ضریب خودهمبستگی ۱ ماهه ( $R_1$ )، ضریب خودهمبستگی ۲ ماهه ( $R_2$ )، ضریب خودهمبستگی ۳ ماهه ( $R_3$ ) و

بیوشیمیایی را با استفاده از روش استنتاج فازی شبیه‌سازی کردند. نتایج نشان داد که این روش نسبت به روش‌های سنتی خطای کمتری دارد.

پیش‌پردازش روی مدل‌های شبکه عصبی می‌تواند دقت پیش‌بینی را بهبود بخشد. استفاده از تبدیل موجک در پیش‌بینی‌های هیدرولوژیکی کارا بوده و هم‌اکنون نیز در حال گسترش است. نورانی و همکاران ترکیب دو قابلیت شبکه‌های عصبی و آنالیز موجک را در پیش‌بینی فرآیندهای هیدرولوژیکی به نام شبکه‌های عصبی-موجکی، ارائه دادند (Nourani et al., 2009). رجایی مدل جدید شبکه عصبی-موجک را برای بار رسوب معلق روزانه در رودخانه‌ها ارائه داد. نتایج نشان داد که مدل ترکیبی در پیش‌بینی بار رسوب از دقت بیش‌تری نسبت به مدل‌های دیگر برخوردار است (Rajaei et al., 2011). رجایی و همکاران مدل ترکیبی موجک و شبکه عصبی را برای بار معلق روزانه رودخانه پیشنهاد کردند. با مقایسه نتایج، توانایی بالای ترکیب موجک و شبکه عصبی مشخص گردید (Rajaei et al., 2011). سینگ مدل ترکیبی شبکه عصبی-موجک را برای پیش‌بینی سیلاب به کار برد. نتایج نشان دهنده توانایی بالای مدل ترکیبی شبکه عصبی-موجک نسبت به روش‌های آماری بود (Singh et al., 2012). لانگین و شوانگ‌ین مدلی را برای پیش‌بینی کیفی آب بر اساس موجک و شبکه عصبی ارائه دادند. این پژوهش نشان داد که استفاده از موجک دقت مدل را نسبت به مدل شبکه عصبی بالا برده است (Longqin and Shuangyin., 2013). کیان جین و همکاران بار رسوب رودخانه‌ای در چین را به وسیله روش ترکیبی عصبی-موجک مدل‌سازی کردند (Qian-Jin et al., 2013). نتایج نشان داد که دقت روش ترکیبی موجک از دو روش دیگر بالاتر است. نورانی و همکاران از داده‌های ماهواره‌ای بارش روزانه ۴ باران سنج برای پیش‌بینی رواناب روز بعد استفاده کردند. نتایج آن‌ها حاکی از برتری مدل ترکیبی موجک-عصبی بود (Nourani et al., 2013).

مقاله‌ی حاضر تحقیقی است که از ترکیب شبکه عصبی خودبازگشتی غیرخطی<sup>۱</sup> و موجک، برای پیش‌بینی سری زمانی دو پارامتر کیفی سولفات و کلراید استفاده می‌کند. بدین منظور پس از تحلیل نتایج همبستگی بین پارامترهای سولفات و کلراید با دبی، به کمک خاصیت خودهمبستگی سری‌های زمانی و زیرسری‌ها و انتخاب زیرسری‌های بهینه و جمع جبری آن‌ها، ترکیبات ورودی مختلفی برای پیش‌بینی هر دو پارامتر ایجاد شد. در این مقاله از تابع موجک Db2 استفاده شده است. در انتها توانایی تشخیص پدیده هیستریزس در سه مدل بررسی می‌شود.

هر چه ضریب همبستگی به عدد ۱ نزدیکتر باشد، نشان‌دهنده سازگاری مناسب بین داده‌هاست. ضرایب همبستگی بین دو پارامتر سولفات و کلراید با تأخیرهای مختلف سری‌های زمانی دبی در جدول ۲ نشان داده شده است. بررسی ضرایب همبستگی بین سولفات و دبی نشان می‌دهد که همبستگی بین  $SO_4$  و  $Cl_t$  با  $Q_{t-1}$  بالاتر از سایر موارد است. منفی بودن ضرایب، همبستگی معکوس پارامترها سولفات و کلراید با پارامتر دبی را نشان می‌دهد.

در صورتیکه داده‌ها به صورت خام وارد شبکه شوند، به علت تغییرات زیاد داده‌ها تأثیر متفاوتی بر روی شبکه می‌گذارند، لذا لازم است متغیرهای ورودی به بازه‌ای مناسب، منطبق با شکل و رفتار تابع محرک سیگموئید منتقل شوند. در این تحقیق از رابطه ۲ برای نرمال‌سازی داده‌ها در محدوده [۰، ۱] استفاده شده است (Singh, 2012).

$$X_{norm} = \frac{X_i - X_{min}}{X_{max} - X_{min}} \quad (2)$$

ضریب خودهمبستگی ۴ ماهه ( $R_4$ ) می‌باشد. باید توجه داشت که مدل‌های شبکه عصبی زمانی بهتر عمل می‌کنند که خصوصیات آماری داده‌های آموزش و آزمون مشابهت داشته باشند. بررسی جدول ۱ نشان می‌دهد که شباهت خصوصیات آماری داده‌های دوره آموزش و آزمون در حد قابل قبولی می‌باشد. ضریب چولگی داده‌های دبی در مقایسه با دو پارامتر دیگر نسبتاً بالا بوده ولی این مقدار پایین بوده و برای مدل‌سازی مناسب است. بررسی‌ها نشان می‌دهد که همبستگی هر دو پارامتر  $SO_4$  و  $Cl_t$ ، به ازای تأخیرهای ۱ تا ۳ ماهه و پارامتر دبی به ازای تأخیرهای ۱ و ۲ ماهه دبی، مقادیر نسبتاً بالایی دارد.

به منظور انتخاب ترکیبات مناسب ورودی مدل‌های رگرسیون خطی و شبکه عصبی، همبستگی بین سری‌های زمانی سولفات ( $SO_4$ ) و کلراید ( $Cl_t$ ) با دبی جریان زمان حال و تأخیرهای آن محاسبه شد (جدول ۲). رابطه ۱ برای محاسبه ضرایب همبستگی بین سولفات و کلراید به طور جداگانه با دبی به کار گرفته شد.

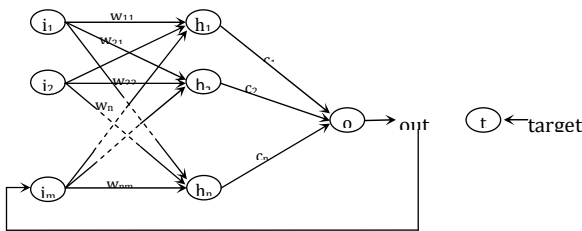
$$R^2 = \frac{\sum(Q_t - \bar{Q})(R_t - \bar{R})}{\sqrt{\sum(Q_t - \bar{Q})^2 \sum(R_t - \bar{R})^2}} \quad (1)$$

جدول ۱ - تحلیل آماری سری‌های زمانی داده‌های  $SO_4$ ،  $Cl$  و  $Q$

کل داده‌ها			مجموعه آزمون			مجموعه آزمایش			پارامترهای آماری
Q ( $m^3/s$ )	$SO_4$ (mEq/l)	Cl (mEq/l)	Q ( $m^3/s$ )	$SO_4$ (mEq/l)	Cl (mEq/l)	Q ( $m^3/s$ )	$SO_4$ (mEq/l)	Cl (mEq/l)	
۱۱۶	۴/۹	۳/۷۸	۱۰۶/۱۶	۷/۱۱	۵/۳	۱۴۶	۴/۴۳	۳/۴	میانگین
۷۱	۲/۲۵	۱/۴۸	۶۵	۲/۴۴	۱/۴۳	۹۵	۲/۳۳	۱/۵۷	انحراف معیار ( $S_d$ )
۱/۱۱	-۰/۲۳	-۰/۲۵	-۰/۴۵	-۰/۵۷	-۰/۰۴	۱/۲۶	-۰/۳۵	-۰/۲	چولگی ( $C_{sx}$ )
۱۳/۶۴	-۰/۹۵	-۰/۵	۱۵/۶	۳/۱۵	۳/۵۲	۱۳/۶۴	-۰/۹۵	-۰/۵	مینیمم
۱/۲۷	۱۲/۵۸	۱۴/۶۵	۴۵۰	۱۲/۵	۸/۲۶	۱۰۰۲۷	۱۲/۵۸	۱۴/۶۵	ماکزیمم
-۰/۶۵	-۰/۵۱	-۰/۶۵	-۰/۷۴	-۰/۵۳	-۰/۵۲	-۰/۵۸	-۰/۵	-۰/۴۳	$R_1$
-۰/۵۳	-۰/۴۵	-۰/۵۱	-۰/۶۱	-۰/۳۱	-۰/۵	-۰/۴۲	-۰/۳۸	-۰/۳۶	$R_2$
-۰/۴۳	-۰/۳۴	-۰/۳۷	-۰/۳	-۰/۲۵	-۰/۳۱	-۰/۳۲	-۰/۲۳	-۰/۱۳	$R_3$
-۰/۳۱	-۰/۰۷	-۰/۲۱	-۰/۲۱	-۰/۱۷	-۰/۱	-۰/۲۵	-۰/۱	-۰/۰۴	$R_4$

جدول ۲ - ضرایب همبستگی  $SO_4$ ،  $Cl_t$  با  $Q_t$  و تأخیرات ۱ تا ۴ روزه  $Q$

کل داده‌ها	Cl (mEq/l)		$SO_4$ (mEq/l)		سری‌های زمانی
	آزمایشی	آزمون	کل داده‌ها	آزمایشی	
-۰/۴۳	-۰/۳۸	-۰/۴۵	-۰/۱۹	-۰/۱۸	$Q_t$
-۰/۳۷	-۰/۳۵	-۰/۴۰	-۰/۱۷	-۰/۱۲	$Q_{t-1}$
-۰/۲۷	-۰/۲۳	-۰/۳۴	-۰/۰۹	-۰/۰۱	$Q_{t-2}$
-۰/۱۴	-۰/۱۱	-۰/۲۹	-۰/۰۴	-۰/۰۴۱	$Q_{t-3}$
-۰/۰۲۳	-۰/۰۵	-۰/۰۴	-۰/۰۳	-۰/۰۰۳	$Q_{t-4}$



شکل ۳ - معماری شبکه عصبی در حالت حلقه پس‌خور بسته

### آنالیز موجک<sup>۵</sup>

روش متداول برای کاهش اغتشاش و تجزیه سیگنال به فرکانس‌های سازنده آن، تبدیل فوریه است که در آنالیز و پردازش سیگنال‌ها کاربرد دارد، اما این تبدیل دارای محدودیت‌هایی می‌باشد (Han et al., 2005). نظریه موجک به عنوان ایده‌ای تکمیلی در جهت رفع مشکلات تبدیل‌های فوریه قادر می‌باشد بسیاری از منظرهای پنهان داده‌هایی که دیگر روش‌های تحلیل سیگنال، در شناسایی آن‌ها ناتوان هستند را شناسایی کند. موجک، موجی کوچک و نوسان‌کننده است که انرژی آن در ناحیه کوچکی متمرکز شده است. در تبدیل موجک، سیگنال از یک سری از فیلترهای بالاگذر و پایین‌گذر عبور داده می‌شود. بخش حاصل از عبور سیگنال از فیلتر بالاگذر که شامل اطلاعات فرکانس بالا از جمله نویز می‌باشد، جزئیات<sup>۶</sup> نام دارد. بخش حاصل از عبور سیگنال از فیلتر پایین‌گذر که شامل اطلاعات فرکانس پایین و دربرگیرنده مشخصات هویتی در سیگنال است و تغییرات آرام را نشان می‌دهد، تقریباً<sup>۷</sup> نامیده می‌شود. فرآیند تجزیه سیگنال به چندین جزء با ریزنمایی پائین شکسته می‌شود. این حالت درخت تجزیه موجک<sup>۸</sup> نامیده می‌شود که در شکل ۴ نشان داده شده است. تابع موجک که موجک مادر نامیده می‌شود به صورت زیر تعریف می‌شود:

$$\int_{-\infty}^{+\infty} \psi(t) dt = 0 \quad (5)$$

$\Psi_{a,b}(t)$  با فشرده‌سازی و کشیدن  $\Psi(t)$  به دست می‌آید و به صورت زیر تعریف می‌شود:

$$\psi_{a,b}(t) = |a|^{-1/2} \psi\left(\frac{t-b}{a}\right) \quad a, b \in \mathbb{R}, a \neq 0 \quad (6)$$

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

$$W_{\psi} f(y, t) = \int_{-\infty}^{+\infty} f(\xi) \psi(a_0^{-1}t - k_0 b_0) d\xi \quad (7)$$

عموماً برای پارامترهای  $a_0$  و  $b_0$  به ترتیب مقادیر ۲ و ۱ انتخاب

### شبکه‌های عصبی مصنوعی<sup>۱</sup>

شبکه‌های عصبی مصنوعی سیستم‌هایی با پردازش اطلاعات موازی هستند که قادرند روابط پیچیده و غیرخطی بین ورودی و خروجی‌های یک سیستم فیزیکی را توسط شبکه‌ای از گره‌های متصل به هم تعیین نمایند (Haykin., 1994). ورودی‌های شبکه عصبی به سه دسته داده‌های آموزش، اعتبارسنجی و آزمون تقسیم می‌شوند. روند آموزش شبکه تا زمانی که خطای مربع متوسط در مراحل آموزش، اعتبارسنجی و آزمون، حداقل گردد ادامه می‌یابد.

### شبکه‌های عصبی NAR و NARX

در مدل شبکه عصبی NAR یا خودبازگشتی غیرخطی<sup>۲</sup>، تنها داده‌های یک سری زمانی مورد استفاده قرار می‌گیرد و مقادیر آینده سری زمانی به صورت تابعی از مقادیر گذشته همان سری زمانی پیش‌بینی می‌شوند. به عبارت دیگر:

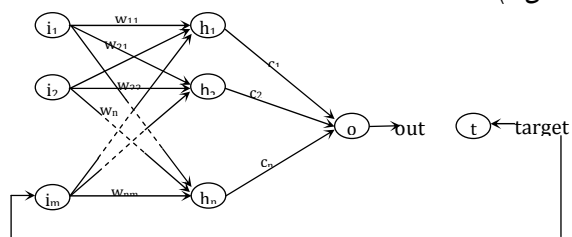
$$x_t = f(x_{t-1}, x_{t-2}, \dots, x_{t-d}) \quad (3)$$

در مدل شبکه NARX، از داده‌های سری زمانی دیگری هم می‌توان استفاده کرد، مثلاً در تحقیق حاضر که از سری‌های زمانی دبی نیز در ترکیب‌های ورودی شبکه عصبی استفاده می‌شود، این مدل به کار می‌رود. به عبارت بهتر:

$$x_t = f(x_{t-1}, x_{t-2}, \dots, x_{t-d}, x_{t-1}^*, x_{t-2}^*, \dots, x_{t-d}^*) \quad (4)$$

در این روابط،  $d$  تعداد تأخیر است. برای آموزش مؤثرتر شبکه این حلقه پس‌خور می‌تواند باز باشد یا حلقه پس‌خور بسته<sup>۳</sup> باشد. در هنگام آموزش که شبکه خروجی صحیح در دسترس است، می‌توان معماری حلقه پس‌خور باز<sup>۴</sup> (شکل ۲) را به کار برد و از خروجی پس‌خور تضمینی استفاده کرد.

اگر حلقه پس‌خور باز باشد تنها می‌توان یک گام جلو را پیش‌بینی کرد (شکل ۳).



شکل ۲ - معماری شبکه عصبی در حالت حلقه پس‌خور باز

- 5- Wavelet analysis
- 6- Details
- 7- Approximation
- 8- Packet decomposition tree

- 1- Artificial Neural Networks
- 2- Nonlinear Auto Regressive
- 3- Close-loop
- 4- Open-loop

لایه‌های مخفی و خروجی به ترتیب از توابع انتقال سیگموئید<sup>۲</sup> و خطی استفاده شد. تعداد نورون‌های لایه مخفی به روش سعی و خطا انتخاب گردید.

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

فرآیندهای هیدرولوژیکی اغلب به صورت الگوهای پیچیده‌ای از همبستگی‌ها که مقیاس‌های زمانی تودرتو و متعددی را شامل می‌شوند، بیان می‌گردند. بنابراین یک مقدار در یک گروه زمانی در یک زمان خاص نه تنها به مقادیری که بلافاصله بعد از آن حاضر می‌شوند، بستگی دارد، بلکه به تغییرات گذشته و آینده‌ی دور نیز بستگی دارد. میزان همبستگی سری‌های زمانی به طور معنی‌داری با انجام پیش‌پردازش موجکی آن‌ها افزایش می‌یابد. که منجر به افزایش قابلیت پیش‌بینی شدن سری‌های زمانی می‌شود.

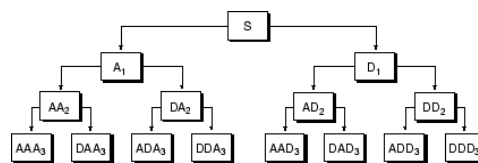
ترکیب موجک و شبکه عصبی، الگوی جدیدی از هوش مصنوعی با نام شبکه عصبی-موجکی را تشکیل می‌دهد. در مدل ترکیبی شبکه عصبی-موجک، زیرسری‌هایی که توسط تبدیل‌های موجک تولید شده اند، به عنوان ورودی شبکه عصبی استفاده می‌شوند. استفاده از تبدیل موجک روی داده‌های سری زمانی قبل از ورود به شبکه این امکان را فراهم می‌کند که به‌توان از یک آنالیز که تأثیرات کوتاه‌مدت و بلندمدت را دربر می‌گیرد بهره‌مند شد. باید توجه داشت که استفاده از تبدیل‌های موجک و تجزیه سری‌های زمانی الزاماً دقت شبکه عصبی را افزایش نخواهد داد، و این نکته مهم در انتخاب موجک مادر است.

در این تحقیق از میان توابع موجک مختلف، موجک Db2 به لحاظ دارا بودن شکل هندسی مناسب برای سری‌های زمانی پارامترهای سولفات، دی و کلراید استفاده شده است. شکل ۵ سری زمانی اصلی سولفات، کلراید و دی را نشان می‌دهد. در مدل پیشنهادی، پس از تجزیه موجک به زیرسری‌ها در ده سطح، ضرایب همبستگی بین زیرسری‌های تولید شده محاسبه می‌شود. زیرسری‌هایی که همبستگی بالایی دارند انتخاب شده و مجموع آن‌ها به عنوان سری زمانی جدید در نظر گرفته می‌شود. با به کار بردن روش همبستگی سری‌های زمانی می‌توان تعداد و تأخیر زمانی مؤثر عناصر ورودی در شبکه‌های عصبی را تعیین کرد، به‌طوری که با این روش، زیرسری‌های اضافی و غیرمؤثر حذف و مجموعه داده‌های ورودی بهینه می‌گردد. به این منظور، ضریب همبستگی هر زیرسری با سیگنال‌های اصلی سولفات، کلراید و دی محاسبه می‌شود. به‌منظور استفاده از خواص خودهمبستگی زیرسری‌ها، ضرایب همبستگی بین زیرسری‌های به دست آمده با تأخیر گام‌های زمانی یک و دو ماهه نیز محاسبه شد (جدول ۳).

می‌گردند. با جایگذاری مقادیر و ، انتقال موجک گسسته به صورت زیر تعریف می‌شود:

$$W_{\psi}f(t, k) = 2^{-t/2} \sum_{k=-\infty}^{\infty} f(\tau) \psi(2^{-t/2} \tau - k) d\tau \quad (8)$$

در این رابطه  $W_{\psi}f(t)$  ضریب موجک گسسته با مقیاس  $a =$  و  $b = 1$  است (Han, et al. 2005).



شکل ۴ - موجک با مقیاس بالا و پایین

### رگرسیون خطی چندمتغیره (MLR)

تحلیل رگرسیون خطی چندمتغیره<sup>۱</sup> تکنیکی برای مدل کردن رابطه خطی بین یک متغیر وابسته و یک یا چند متغیر مستقل است. شکل معمول معادله MLR به این صورت نشان داده می‌شود:

$$y = a_0 + \sum a_i x_i \quad (9)$$

که عرض از مبدأ، ضریب رگرسیون متغیرهای مستقل و مقدار پیش‌بینی شده است (Qian-Jin et al., 2013).

### مدل سازی

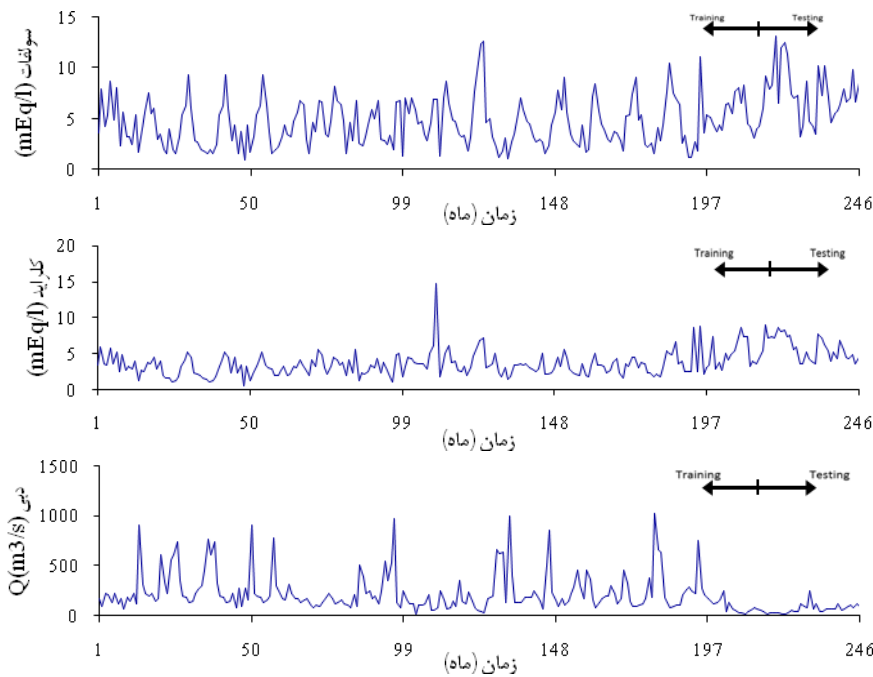
با در نظر گرفتن نتایج تحلیل آماری ارائه شده در جداول ۱ و ۲ ترکیبات مختلف ورودی برای پیش‌بینی میزان  $SO_4$  و  $Cl$  ماه آینده در زمان  $t+1$  به صورت زیر در نظر گرفته شد.

1.  $SO_4t$
2.  $SO_4t, SO_4t-1$
3.  $SO_4t, SO_4t-1, SO_4t-2$
4.  $SO_4t, Q_t$
5.  $SO_4t, SO_4t-1, Q_t$
6.  $SO_4t, SO_4t-1, Q_t, Q_t-1$
1.  $Cl_t$
2.  $Cl_t, Cl_{t-1}$
3.  $Cl_t, Cl_{t-1}, Cl_{t-2}$
4.  $Cl_t, Q_t$
5.  $Cl_t, Cl_{t-1}, Q_t$
6.  $Cl_t, Cl_{t-1}, Q_t, Q_{t-1}$

در این تحقیق به‌منظور آموزش و آزمون از مدل شبکه عصبی سه لایه با روش پیش‌خور پس انتشار خطا<sup>۲</sup> (FFBP) با الگوریتم آموزشی لونیبرگ - مارکوارت استفاده شد. بر اساس تحقیقات، الگوریتم آموزشی لونیبرگ-مارکوارت، به دلیل سرعت، دقت و اطمینان بالایی آن در پیش‌بینی‌های هیدرولوژیکی مناسب می‌باشد. هم‌چنین برای

مؤثرترین زیرسری‌های تجزیه شده از سیگنال کلراید نیز مربوط به CDW6, CDWApp, CDW9, CDW3, CDW2, CDW1 و CDW7 بود که در این حالت نیز، این زیرسری‌های مؤثر با هم جمع شده و سیگنال اصلی با عنوان  $TCDW_t$  و  $W_{t-1}$  بازسازی می‌شود. با توجه به ارتباط درجه سطوح با تناوب رخداد زمانی و بازه زمانی اصلی که ۲۴۶ ماه است، انتظار می‌رفت که زیرسری‌های DDW8, DDW9 و DDW10 با سری زمانی اصلی همبستگی بالایی نداشته باشند.

با توجه به جدول ۳ بالاترین ضریب همبستگی بین زیرسری‌های دبی با یک و دو تأخیر و سولفات مربوط به DDW1, DDW2, DDW3, DDW5, DDW7 و دبی با یک و دو تأخیر و کلراید مربوط به DDW3, DDW5, DDW6, DDWApp, DDW5, DDW3 می‌باشد. بنابراین جمع این زیرسری‌های منتخب، حالت  $W_{t-1}, TDDW_t$  را ایجاد می‌کند. از بین زیرسری‌های سولفات با یک و دو تأخیر SDW1, SDW2, SDW3, SDW5, SDW6, SDWApp و SDWApp به عنوان زیرسری‌های مؤثر انتخاب شدند که حاصل جمع این زیرسری‌های منتخب با عنوان  $TSDW_t$  و  $W_{t-1}$  در نظر گرفته شده است.



شکل ۵- سری‌های زمانی یون سولفات، کلراید و دبی ماهانه ایستگاه مورد مطالعه

جدول ۳- ضرایب همبستگی زیرسری‌های تجزیه شده با سولفات و کلراید

زیرسری‌های سری زمانی دبی	همبستگی بین زیرسری‌های دبی (با یک گام تأخیر یا کلراید و سولفات)		همبستگی بین زیرسری‌های دبی (با دو گام تأخیر یا کلراید و سولفات)		زیرسری‌های سری زمانی سولفات	همبستگی بین زیرسری‌های سولفات (با یک گام تأخیر یا سیگنال اصلی سولفات)		زیرسری‌های سری زمانی کلراید	همبستگی بین زیرسری‌های سولفات (با یک گام تأخیر یا سیگنال اصلی کلراید)	
	کلراید	سولفات	کلراید	سولفات		سولفات	سولفات		کلراید	سولفات
DDW 1	-۰/۱۸	-۰/۰۸	-۰/۱۶	-۰/۰۷	SDW 1	-۰/۲۸	-۰/۱۶	CDW 1	-۰/۱۴	-۰/۱
DDW 2	-۰/۱۳	-۰/۰۷	-۰/۱۵	-۰/۰۱	SDW 2	-۰/۲۱	-۰/۱۸	CDW 2	-۰/۱۸	-۰/۱۲
DDW 3	-۰/۱۵	-۰/۲۳	-۰/۱۱	-۰/۱۳	SDW 3	-۰/۲۱	-۰/۱۴	CDW 3	-۰/۲۶	-۰/۱۷
DDW 4	-۰/۰۱۷	-۰/۰۳	-۰/۰۵۱	-۰/۰۷	SDW 4	-۰/۰۱	-۰/۰۵	CDW 4	-۰/۱۱	-۰/۰۵
DDW 5	-۰/۱۷	-۰/۲۱	-۰/۱۱	-۰/۱۷	SDW 5	-۰/۱۴	-۰/۱	CDW 5	-۰/۰۸	-۰/۰۱
DDW 6	-۰/۱	-۰/۲	-۰/۰۹	-۰/۱	SDW 6	-۰/۱۷	-۰/۱۶	CDW 6	-۰/۳۳	-۰/۱۷
DDW 7	-۰/۱۴	-۰/۰۹	-۰/۱۱	-۰/۰۸	SDW 7	-۰/۰۲	-۰/۰۴	CDW 7	-۰/۱۷	-۰/۱۵
DDW 8	-۰/۰۱	-۰/۰۲	-۰/۰۲	-۰/۰۲	SDW 8	-۰/۰۱۸	-۰/۰۱۴	CDW 8	-۰/۰۹	-۰/۰۱
DDW 9	-۰/۱	-۰/۰۱	-۰/۰۹	-۰/۱	SDW 9	-۰/۰۹۱	-۰/۱	CDW 9	-۰/۰۲۵	-۰/۰۲۶
DDW 10	-۰/۰۳	-۰/۰۷	-۰/۰۱	-۰/۰۷	SDW 10	-۰/۰۱	-۰/۰۹	CDW 10	-۰/۰۱	-۰/۰۳
DDW App.	-۰/۰۷	-۰/۱۶	-۰/۰۶	-۰/۱۷	SDW App.	-۰/۱۳	-۰/۱۸	CDW App.	-۰/۱۷	-۰/۱۸

جدول ۴- مقایسه ضرایب همبستگی پیش و پس از نویززدایی سیگنال‌ها

سری‌های زمانی اصلی	ضریب همبستگی با سیگنال اصلی سولفات	ضریب همبستگی با سیگنال اصلی کلراید
Qt	-۰/۲۲	-۰/۴۵
Qt-1	-۰/۱۹	-۰/۴
Qt-2	-۰/۱۵	-۰/۳۴
TDDW <sub>t</sub>	-۰/۲۵	-۰/۴۸
TDDW <sub>t-1</sub>	-۰/۱۹	-۰/۴۳
TDDW <sub>t-2</sub>	-۰/۱۳۴	-۰/۳۱
TS(C)DW <sub>t</sub>	-۰/۶۳	-۰/۶۱
TS(C)DW <sub>t-1</sub>	-۰/۴۸	۰/۵۴
TS(C)DW <sub>t-2</sub>	۰/۳۴	۰/۴۴

خطا به ۰ نزدیک شود، مدل پیش‌بینی بهتری دارد.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_i - Y_i)^2}{n}} \quad (10)$$

$$MAE = \frac{\sum_{i=1}^n |X_i - Y_i|}{n} \quad (11)$$

$$E = 1 - \frac{\sum_{i=1}^n (Y_i - X_i)^2}{\sum_{i=1}^n (Y_i - \bar{X}_i)^2} \quad (12)$$

در این روابط،  $X_i$  و  $Y_i$  به ترتیب آمین داده پارامتر برآوردی و اندازه گیری، متوسط داده‌های  $X_i$  و  $n$  تعداد نمونه‌های مورد ارزیابی است.  $E=1$  به این معنی است که داده‌های پیش‌بینی شده با داده‌های مشاهده شده هم‌خوانی کامل دارد.  $E=0$  یعنی پیش‌بینی مدل منطبق بر مقادیر میانگین داده‌ها می‌باشد؛ در حالی که  $E<0$  نشان‌دهنده این است که میانگین داده‌های مشاهداتی نسبت به مدل پیش‌بینی شده دقت بیش‌تری دارد (Zhou et al., 2008).

### نتایج و بحث

برای به دست آوردن بهترین نتیجه به منظور پیش‌بینی پارامترهای سولفات و کلراید، با در نظر گرفتن شدت همبستگی بین پارامترها، ترکیبات مختلف ورودی ایجاد شد. در این مدل‌ها پیش‌بینی ۲۵ ماهه (دوره آزمون) انجام شد و مقادیر  $E$ ،  $RMSE$  و  $MAE$  محاسبه شد (جدول ۱ تا ۴). در اجرای مدل رگرسیون خطی چندمتغیره، سولفات و کلراید ماه فعلی به عنوان متغیر وابسته و داده‌های مربوط به ماه‌های قبل به عنوان متغیر مستقل در نظر گرفته شدند. روابط مربوط به مدل رگرسیون برای ترکیبات ورودی شش‌گانه در جدول ۵ و نتایج در جدول ۶ نمایش داده شده است. بهترین نتیجه برای هر دو پارامتر سولفات و کلراید در این مدل مربوط به ترکیب ورودی ۱ با ضریب تبیین ۰/۲۳ و ۰/۳۸ می‌باشد. مدل بعدی که در این تحقیق مورد استفاده قرار گرفت، مدل شبکه عصبی بود. این مدل‌ها بر اساس تعداد لایه‌های مخفی، تعداد نرون‌های لایه مخفی و تعداد تأخیرها (۱ تا ۴ ماه قبل) آموزش داده شد و مورد آزمون قرار گرفتند.

به‌منظور ایجاد ترکیب‌های ورودی مناسب به مدل شبکه عصبی، ضرایب همبستگی بین سیگنال حاصل از جمع جبری زیرسری‌ها با تأخیرات مختلف و سری زمانی اصلی سولفات و دبی مطابق جدول ۴ محاسبه می‌شود. همانطور که در این جدول مشخص است، ضرایب همبستگی بین دبی و سیگنال بدون پیش‌پردازش سولفات ۰/۲۲- و کلراید ۰/۴۵- می‌باشد. در حالی که حذف نویز به وسیله موجک این ضریب را به ۰/۲۵- و ۰/۴۸- بهبود داده است. همچنین همبستگی بین تأخیرات یک ماهه دبی در پارامتر کلراید از ۰/۴- به ۰/۴۳- افزایش یافته است. با مقایسه ضرایب خودهمبستگی سولفات و کلراید در جدول ۴ مشخص می‌شود که پیش‌پردازش آنالیز موجک، در تمامی حالات این ضرایب را افزایش داده است.

با استفاده از نتایج تحلیل آماری در جدول ۴ ترکیب‌های زیر برای هر یک از دو پارامتر سولفات و کلراید به عنوان ورودی شبکه عصبی - موجک در نظر گرفته می‌شود، تا بهترین پیش‌بینی حاصل شود.

$$\begin{aligned} &TSDW_{t,1} \\ &TSDW_{t-1}, TSDW_{t,2} \\ &, TSDW_{t-1}, TSDW_t, TSDW_{t-2,2} \\ &TDDW_t, TSDW_{t,4} \\ &, TDDW_t, TSDW_t, TSDW_{t-1,5} \\ &TDDW_t, TSDW_t, TSDW_{t-1,6} \\ &TCDW_{t,1} \\ &TCDW_{t-1}, TCDW_t, TCDW_{t,2} \\ &, TCDW_{t-1}, TCDW_t, TCDW_{t-2,2} \\ &TDDW_t, TCDW_{t,4} \\ &, TDDW_t, TCDW_t, TCDW_{t-1,5} \\ &TDDW_t, TCDW_t, TCDW_{t-1}, TDDW_t,6 \end{aligned}$$

### معیارهای ارزیابی مدل

در این تحقیق سه شاخص معیاری شامل ریشه میانگین مربعات خطای بین داده‌های پیش‌بینی و مشاهده شده ( $RMSE$ )، میانگین قدر مطلق خطا ( $MAE$ )، ضریب کارایی نش-ساتکلیف یا ضریب تبیین ( $E$ ) به عنوان ارزیابی توان پیش‌بینی مدل‌ها، روابط ۱۰ تا ۱۲ مورد استفاده قرار گرفته است. به طور کلی هرچه ضریب تبیین به ۱ و

جدول ۵ - روابط مربوط به مدل رگرسیون خطی چند متغیره

ترکیبات	MLR(SO4)	MLR(CL)
۱	$y_i = 2.35 + 0.4684y_{t-1}$	$y_i = 2.64 + 0.2143y_{t-1}$
۲	$y_i = 2.29 + 0.455y_{t-1} + 0.027y_{t-2}$	$y_i = 2.18 + 0.1736y_{t-1} + 0.177y_{t-2}$
۳	$y_i = 2.82 + 0.46y_{t-1} + 0.137y_{t-2} - 0.235y_{t-3}$	$y_i = 2.44 + 0.194y_{t-1} + 0.1968y_{t-2} - 0.116y_{t-3}$
۴	$y_i = 4.218 + 0.2358y_{t-1} - 0.006x_{t-1}$	$y_i = 1.87 + 0.448y_{t-1} - 0.067x_{t-1}$
۵	$y_i = 3.34 + 0.16y_{t-1} + 0.32y_{t-2} - 0.09x_{t-1}$	$y_i = 2.15 + 0.32y_{t-1} + 0.12y_{t-2} + 0.054x_{t-1}$
۶	$y_i = 3.78 + 0.188y_{t-1} + 0.1y_{t-2} + 0.109x_{t-1} - 0.14x_{t-2}$	$y_i = 2.23 + 0.268y_{t-1} - 0.08y_{t-2} + 0.109x_{t-1} - 0.14x_{t-2}$

جدول ۶ - نتایج به دست آمده از مدل رگرسیون خطی چندمتغیره برای ترکیبات ۱ تا ۶

مدل‌ها	MLR (SO4)						MLR (Cl)					
	*۱	۲	۳	۴	۵	۶	*۱	۲	۳	۴	۵	۶
E	۰/۲۳	۰/۱۳	۰/۱۵	-۰/۱۲	-۰/۲۵	۰/۰۵	۰/۳۸	-۰/۲۶	۰/۳۳	۰/۳۵	۰/۲۳	-۰/۰۹
MAE(mEq/l)	۱/۷	۱/۸۳	۱/۹۱	-	-	-	۰/۸۴	۱/۲۱	۰/۹۳	۰/۸۸	۱/۲	-
RMSE(mEq/l)	۱/۰۹	۱/۳۹	۱/۴۳	-	-	-	۰/۶۸	۱/۲۳	۱/۱	۰/۷۶	۱/۲۶	-

آزمون حداقل گردد. همانطور که از نتایج جدول ۷ مشخص است، افزودن دبی جریان، نتایج را بهبود نمی‌دهد. بهترین پیش‌بینی برای کلراید از ترکیب ۴ (وابستگی کلراید به کلراید و دبی زمان حال) با معماری ۲-۳-۴ و الگوریتم آموزش لونیبرگ-مارکوارت حاصل شد. همان‌گونه که از نتایج مشخص است، مدل شبکه عصبی نتایج دقیقتری نسبت به مدل رگرسیون خطی دارد. رگرسیون خطی ضعیف بوده و قادر به شبیه‌سازی نوسانات شدید و پی در پی و پیش‌بینی مناسب SO<sub>4</sub> و Cl نیست.

در مدل پیشنهادی عصبی-موجک زیرسری‌های تقریب و جزئیات به دست آمده از تجزیه موجکی سیگنالی SO<sub>4</sub>، Cl و Q در ده سطح مورد بررسی قرار گرفتند. هر یک از زیرسری‌ها نماینده بی‌نظمی‌ها، تغییرات و جهش‌های سریع در هر یک از گروه‌های زمانی سولفات، کلراید و دبی می‌باشد. گروه‌های زمانی پس از حذف اختلالات دارای همبستگی بالاتری می‌باشند که به علت افزایش حافظه سری زمانی می‌باشد.

تعداد لایه‌های مخفی و تعداد نرون‌های لایه‌ی مخفی هر شبکه با استفاده از تعداد داده‌های موجود و روش سعی و خطا بهینه شد. باید توجه داشت که تعداد لایه‌های مخفی بیش‌تر و تعداد نرون بیش‌تر در لایه‌ی 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. بنابراین از بین دو معماری شبکه با دقت برابر، موردی که سادگی بیش‌تری داشته باشد مناسب‌تر است. در هنگام آموزش شبکه می‌بایست به مسئله یادگیری اضافی<sup>۱</sup> دقت شود. در این حالت حین آموزش شبکه خطای مجموعه آموزشی به مقدار بسیار ناچیزی می‌رسد، ولی هنگام معرفی داده‌های جدید (اعتبارسنجی و آزمون) خطای مدل زیاد می‌شود؛ در واقع شبکه در مرحله آموزش، روند داده‌ها را یاد نگرفته بلکه با به حافظه سپردن داده‌ها از بهینه کردن وزن ممانعت می‌کند و از این رو نتایج قابل تعمیم به مرحله آزمون نمی‌باشد. بنابراین بهترین عملکرد شبکه عصبی زمانی است که خطای هر سه مرحله آموزش، اعتبارسنجی و

جدول ۷ - نتایج به دست آمده از مدل شبکه عصبی برای ترکیبات ۱ تا ۶

مدل‌ها	ANN (SO4)						ANN (Cl)					
	۱	*۲	۳	۴	۵	۶	۱	۲	۳	*۴	۵	۶
تعداد نرون‌های لایه پنهان	۵	۸	۶	۹	۸	۵	۶	۶	۹	۸	۵	۹
E	۰/۶۵	۰/۷۵	۰/۷۳	۰/۶۱	۰/۶۶	۰/۵۹	۰/۶۹	۰/۷۱	۰/۷۴	۰/۷۹	۰/۷۴	۰/۶۸
MAE(mEq/l)	۰/۲۳	۰/۹۵	۰/۹۷	۱/۳۱	۱/۱۸	۱/۳۴	۱/۱۳	۰/۹۳	۰/۸۷	۰/۶۴	۰/۸۶	۱/۱۸
RMSE(mEq/l)	۰/۴۶	۰/۲۷	۰/۳۵	۰/۵۸	۰/۴۱	۰/۶۱	۰/۸۵	۰/۸۲	۰/۷۸	۰/۳۶	۰/۶۷	۸



به این ترتیب سیگنالی بهینه شده از جمع زیرسری‌های مؤثر و با همبستگی بالا به دست آمد و تحت شش ترکیب ارائه شده فوق به شبکه عصبی اعمال شدند. در این مدل از تابع موجکی Db2 به علت شکل هندسی مشابه با سری‌های زمانی سولفات و کلراید استفاده شد. مطابق جدول ۸ بهترین نتیجه برای پارامتر سولفات در ترکیب ۲ و پارامتر کلراید در ترکیب ۴ با ضریب تبیین ۰/۸۱ و ۰/۸۴ به دست آمد. این نتیجه در مقایسه با دو روش قبلی رگرسیون و عصبی به ترتیب برای سولفات ۰/۰۶ و ۰/۵۸ و برای کلراید ۰/۰۵ و ۰/۴۶ افزایش را نشان می‌دهد. جدول ۸ به وضوح برتری مدل ترکیبی شبکه عصبی-موجک را نسبت به مدل‌های شبکه عصبی و رگرسیون خطی نشان می‌دهد. تغییرات فرکانسی پارامترهای سولفات و کلراید به عوامل گوناگونی بستگی دارد. برخی از این عوامل مانند تبخیر، فرسایش و آبخویی و انتقال رسوبات، ساختار زمین‌شناسی و سازندهای منطقه، افزایش سطح تماس با سایر کانی‌ها، دخالت‌های انسانی، عبور از لایه‌های گچی و نم‌زار اثر افزایشی و برخی دیگر مانند افزایش ناگهانی دبی رودخانه و سیلاب اثر کاهش‌دهنده دارد. البته در صورتیکه افزایش دبی منجر به افزایش فرسایش شود، این رابطه افزایشی خواهد بود. تأثیر هر یک از این عوامل در طول زمان، به صورت ضمنی در بین داده‌ها قرار گرفته است.

آنالیز رگرسیون کلیات روند تغییرات سری زمانی را نشان می‌دهد و نتوانسته جزئیات را شبیه‌سازی کند، در حالیکه مدل شبکه عصبی با بهینه کردن وزن‌ها تا حد قابل قبولی توانسته سولفات و کلراید را پیش‌بینی نماید. مدل ترکیبی موجک-عصبی با تجزیه سری زمانی با ماهیت کاملاً نوسانی، به زیرسری‌هایی ایستا عملکرد بهتری از خود نشان داده است. آلودگی شیمیایی رودخانه دز به دلیل ورود پساب کارخانجات اطراف به خصوص کارخانجات کاغذسازی پارس، کشت و صنعت نیشکر هفت تپه و شیر پاستوریزه شوش و هم‌چنین ورود رواناب و زه‌آب زمین‌های کشاورزی اطراف می‌باشد که به دلیل برهم‌کنش بالقوه‌ای که بین آلاینده‌ها وجود دارد، بسیار پیچیده می‌باشد. منطقی است که با این روابط غیرخطی، نتیجه مناسبی از طریق آنالیز رگرسیون حاصل نمی‌شود. در این میان، روابط پیچیده

زمانی و مکانی آلاینده‌ها از طریق جداسازی نشانه‌ها از حافظه سری زمانی این دو پارامتر قابل پیش‌بینی می‌شود. هر یک از ده سطح تجزیه سیگنالی آنالیز موجک ارتباط‌های تناوبی و همبستگی بین پارامترهای سولفات، کلراید و دبی را به صورت ماهانه رمزگشایی کردند. تأثیرات دبی در این روابط به صورت مستقیم اعمال شد. با توجه به رابطه معکوس دبی و غلظت آلاینده‌ها، مقادیر بیشینه برای هر دو پارامتر در فصول خشک و ماه‌های گرم مشاهده شد. شدت تبخیر در این منطقه به این رابطه کمک می‌کند. نتایج نشان‌دهنده نقش مؤثر دبی در پیش‌بینی پارامتر کلراید می‌باشد، در حالیکه لحاظ کردن مستقیم زیرسری‌های تجزیه‌شده دبی از سیگنال اصلی آن در ترکیبات پیش‌بینی سولفات راه‌گشا نبوده است. با بررسی خطای RMSE در جدول ۸، مدل ترکیبی Db2 توانست خطا را در دو مدل دیگر رگرسیون و شبکه عصبی برای سولفات به مقدار قابل توجهی کاهش دهد.

با توجه به شکل ۶ و ۷، نمودار مدل ترکیبی موجک-عصبی با نمودار مقادیر مشاهده شده تطابق بیشتری دارد. عملکرد ضعیف روش رگرسیون خطی در شکل ۴ با پراکندگی داده‌ها در اطراف خط نیمساز دیده می‌شود. نقاط این پراکندگی اکثراً در سمت راست خط نیمساز قرار دارد. هرچه این نقاط به نیمساز نزدیک‌تر باشند، پیش‌بینی بهتر انجام شده است. به نظر می‌رسد روش رگرسیون بر مبنای بازه آموزش که ۱۹۷ ماه اول دوره را در بر می‌گرفت، مدل مقادیر میانگینی را ارائه داده و در هیچ یک از نوسان‌ها و پیک‌ها با مقادیر مشاهداتی تطابق ندارد. در واقع مدل رگرسیون شکل کلی روند تغییرات را با یک گام تأخیر دنبال کرده، که این موضوع به ماهیت خطی این مدل بر می‌گردد. این روش قادر به تشخیص روابط پیچیده بین داده‌های سری زمانی نبوده و پیش‌بینی با خطای نسبتاً بالایی همراه است، در صورتیکه با توجه به شکل ۶ شبکه عصبی بهتر توانسته ساز و کارهایی را برای جبران تغییرات به وجود آمده در طول زمان، ارائه دهد. به خصوص مقادیر بیشینه و کمینه را به خوبی شناسایی کرده ولی در بیشینه‌ها مقادیر را اندکی کم‌تر از مقادیر واقعی پیش‌بینی کرده است.

جدول ۸ - نتایج به دست آمده از مدل ترکیبی شبکه عصبی- موجک برای ترکیبات ۱ تا ۶

مدل‌ها	WANN (SO4)						WANN (Cl)					
	۱	۲*	۳	۴	۵	۶	۱	۲	۳	۴*	۵	۶
تعداد نرون‌های لایه پنهان	۷	۵	۴	۹	۷	۵	۱۱	۹	۹	۵	۷	۹
E	۰/۶۳	۰/۸۱	۰/۷۱	۰/۶۸	۰/۷۱	۰/۵۹	۰/۶۸	۰/۷۳	۰/۷۸	۰/۸۴	۰/۷۸	۰/۶۵
MAE(mEq/l)	۱/۲۶	۰/۶۳	۱/۱۲	۰/۹۳	۰/۸۸	۱/۳۱	۱/۱۵	۰/۹	۰/۹	۰/۵۲	۰/۸۳	۱/۲۳
RMSE(mEq/l)	۳/۱۵	۰/۱۵	۰/۳۸	۰/۴۷	۰/۳۸	۰/۶	۰/۸۶	۰/۷۹	۰/۷۷	۰/۲۹	۰/۶۵	۰/۹۶

### حلقه‌ی پسماند (Hysteresis)

اطلاع از تغییرات میزان یونهای معلق در مقابل نوسانات ناگهانی و تدریجی دبی رودخانه‌ها از مسائل مهم در مدیریت حوزه‌های آبخیز می‌باشد. در برخی موارد این تغییرات موجب پدیده پسماند می‌شود، به این معنی که تغییرات سولفات وابسته به تغییرات دبی می‌باشد و اگر تغییرات دوره‌ای سولفات با تأخیر در پی تغییرات دبی رخ دهد، احتمال پسماند وجود دارد. (Williams, G., 1989) پنج کلاس معمولی از حلقه‌های پسماند را شامل: ساعتگرد پادساعتگرد، شکل ۸ مانند، خطی و خطی به اضافه یک حلقه معرفی نمود. با بررسی و تحلیل این منحنی‌ها می‌توان از منابع آلاینده و نحوه انتقال آن آگاهی یافت. در شکل ۸ یک رویداد سیل در بازه زمانی آزمون بین ماه‌های بهمن ۸۸ تا تیر ۹۰ برای پارامتر سولفات رسم شده است. در مورد کلراید، با بررسی روابط و نمودارها، این پدیده مشاهده نشد. همانطور که از شکل مشخص است، روش رگرسیون این پدیده را تشخیص نداده و رابطه‌ای خطی را از دو پارامتر دبی و سولفات ارائه کرده است. طبق این رابطه با کاهش دبی، میزان سولفات کاهش می‌یابد. این درحالیست که مدل شبکه عصبی و ترکیبی موجک قادر به مدل کردن شرایطی نظیر پدیده هیستریزس و روابط جریان می‌باشند.

### نتیجه‌گیری

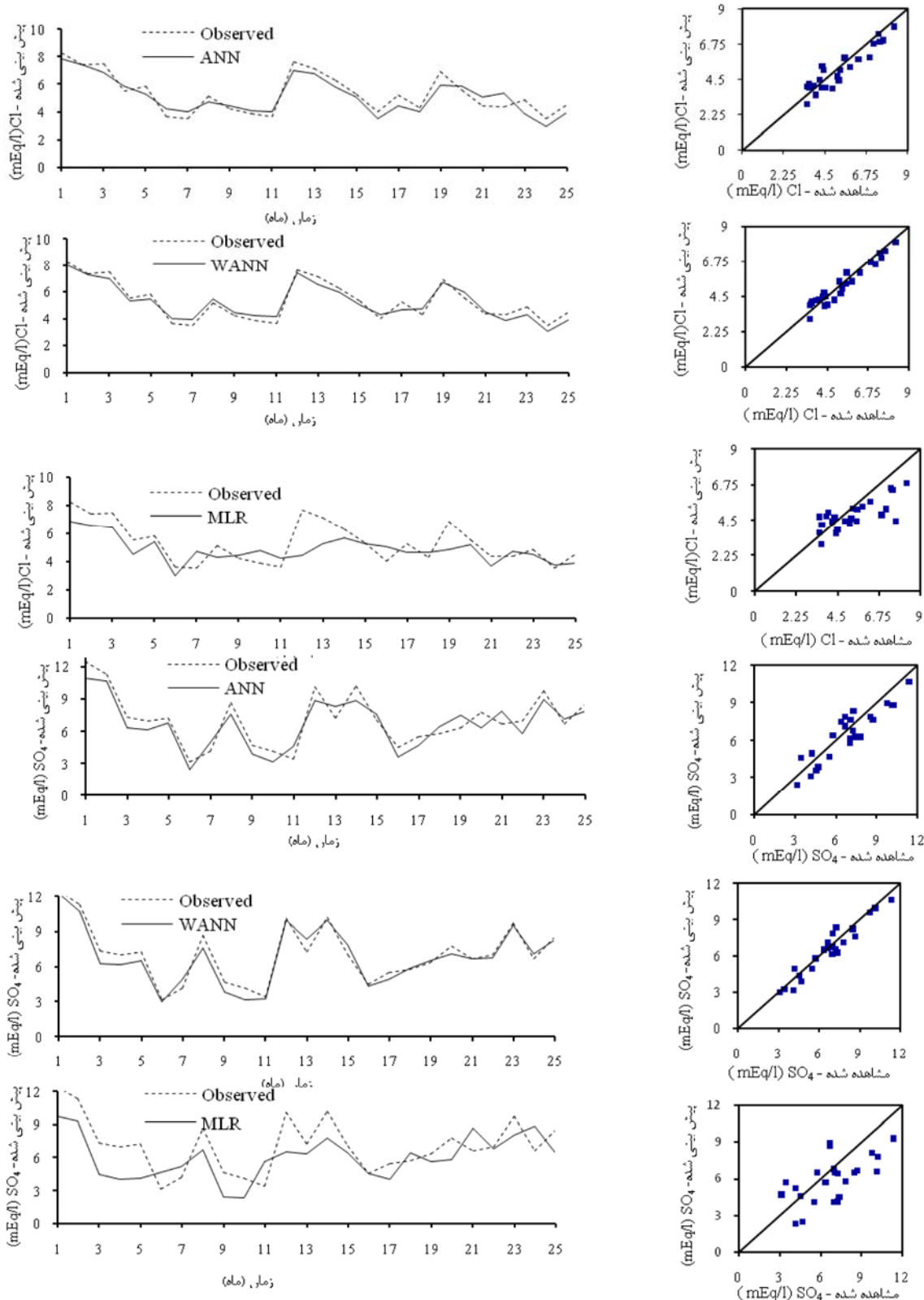
در این تحقیق کاربرد مدل‌های شبکه عصبی، ترکیب شبکه عصبی-موجک و رگرسیون خطی چندمتغیره در پیش‌بینی مقادیر سولفات و کلراید آب رودخانه دز در ایستگاه بامدژ بررسی شد. به دلیل اینکه سری زمانی سولفات و کلراید غیرایستا با ماهیتی نوسانی هستند، فقط لحاظ نمودن تغییرات زمانی آنها راه‌گشا نمی‌باشد.

در این مدل نقاط اطراف خط نیم‌ساز نوسان کرده‌اند که نشان می‌دهد این مدل از پارامترهای بیشتری برای تنظیم وزن‌ها برخوردار بوده است. این حالت در نمودار مربوط به موجک به دقت بیشتری می‌رسد به طوریکه نقاط، بر خط نیم‌ساز انطباق خوبی دارند. شبکه عصبی-موجکی نه تنها قادر به تشخیص تغییرات سری زمانی بوده بلکه نقاط شکستگی، بیشینه و کمینه را به خوبی پیش‌بینی کرده است. با بررسی نقاط شکستگی نمودارهای شبکه عصبی و مدل ترکیبی می‌توان دریافت که عملکرد موجک باعث شده که در قله‌ها با تشخیص داده‌های مؤثر به سمت دقیق‌تر شدن پیش رفته است. ساختار این مدل با تجزیه سیگنال، وزن‌های شبکه عصبی را به گونه‌ای تنظیم می‌کند که روند طبیعی سیستمی را که شبیه‌سازی می‌کند نشان می‌دهد.

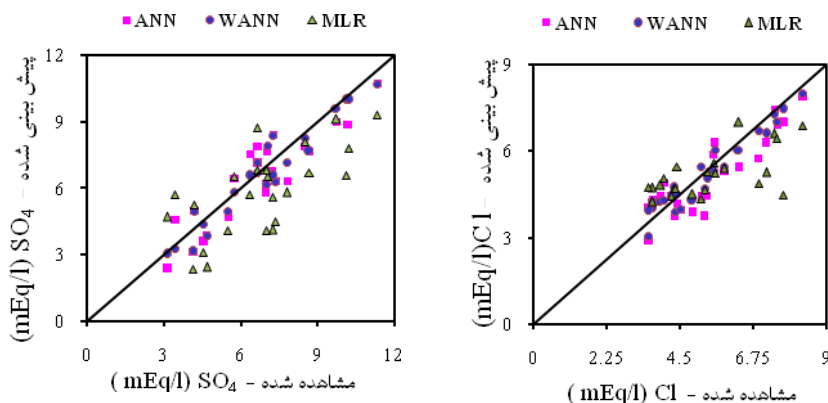
با توجه به جدول ۹، توانایی و ویژگی موجک در تخمین و شبیه‌سازی نقاط پیک (نسبت به دو روش دیگر) دیده می‌شود. جدول ۶ توان پیش‌بینی مقادیر بیشینه و کمینه در ۲۵ ماهه آخر بازه زمانی را توسط سه مدل نشان می‌دهد. استفاده هم‌زمان شبکه‌های عصبی و تبدیل موجک در افزایش ضریب کارایی و کاهش میزان خطای داده‌های محاسباتی به مشاهداتی، موفقیت‌آمیز بوده است. شبکه‌های عصبی و موجکی قادر هستند که یک فضای چندگانه را حتی در صورت ناقص بودن اطلاعات تطبیق دهند. بیش‌ترین مقدار سولفات و کلراید آب در تابستان و پاییز مشاهده می‌شود، لذا آب رودخانه در ماه‌های تابستان و پاییز از نظر سولفات آلوده به نظر می‌رسد و این، گویای روابط این دو پارامتر با میزان دبی و شدت تبخیر است.

جدول ۹ - پیش‌بینی مقادیر بیشینه بازه آزمون توسط سه مدل برای کلراید و سولفات

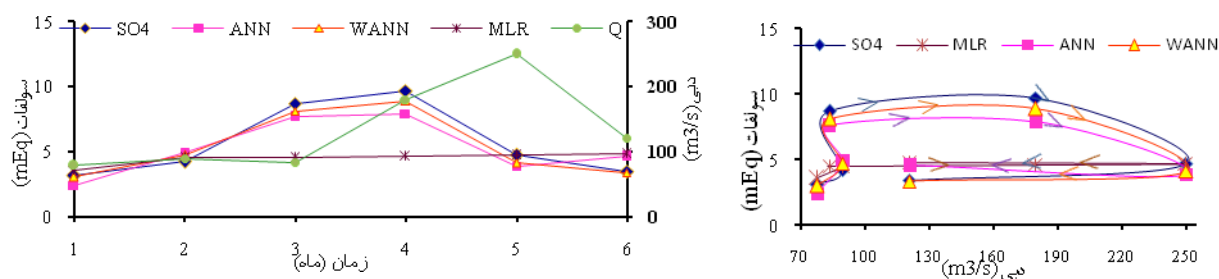
ردیف	(mEq/l) پارامتر سولفات				(mEq/l) پارامتر کلراید				
	تاریخ	MLR	WANN	ANN	واقعی	تاریخ	MLR	WANN	ANN
۱	شهریور ۸۸	۹/۸	۱۲/۲۷	۱۰/۲۳	۱۲/۵	شهریور ۸۸	۵/۸۶	۷/۹۹	۶/۸۷
۲	مهر ۸۸	۹/۳	۱۰/۷	۱۰/۶۹	۱۱/۳	آبان ۸۸	۵/۴۴	۷	۶/۸
۳	مرداد ۸۹	۶/۵۶	۹/۹۶	۸/۸۶	۱/۱	مرداد ۸۹	۴/۷۵	۴/۰۳	۴/۲
۴	مهر ۸۹	۸/۷	۹/۹۶	۸/۸۶	۱۰/۲	شهریور ۸۹	۵/۲۵	۶/۵۸	۶/۳
۵	تیر ۹۰	۹/۱	۹/۵۹	۸/۹۸	۹/۷۶	اسفند ۸۹	۴/۸۷	۶/۷	۵/۷۳
مجموع	-	۴۲/۵۶	۵۲/۴۹	۴۸/۳۵	۵۳/۹	-	۲۷/۹	۳۵/۷۲	۳۳/۸۱
خطا (%)	-	%۲۱/۱	%۴/۸	%۱۰	-	-	%۲۵/۴	%۴/۸۹	%۷/۹
RMSE	-	۲/۴۷	۰/۳۴	۱/۱۷	-	-	۲/۰۴۷	۰/۳۷	۰/۷۷



شکل ۶ - مقایسه مقادیر سولفات و کلراید اندازه گیری شده و پیش بینی برای هر سه مدل



شکل ۷ - مقایسه نتایج پیش‌بینی مدل‌های ANN، WANN، MLR در برابر مقادیر مشاهداتی برای SO4 و Cl



شکل ۸ - نمودار دبی و سولفات مشاهده شده و پیش‌بینی شده، رابطه هیستریزس بین دبی و سولفات در یک بازه از دوره آزمایش

### منابع

گلابی، م.، موسوی جهرمی، س.ج. و کرمی، ب. ۱۳۸۵. شبیه‌سازی EC، SAR، TDS، رودخانه کارون با استفاده از Qnet 2000. هفتمین سمینار بین‌المللی مهندسی رودخانه، دانشگاه شهید چمران اهواز.

نوشادی، م.، سالمی، ح. و احمدزاده، م. ۱۳۸۶. شبیه‌سازی و پیش‌بینی برخی از پارامترهای کیفیت آب رودخانه زاینده رود با استفاده از شبکه عصبی مصنوعی، نشریه آب و فاضلاب، شماره ۶۴، زمستان ۱۳۸۶، صص ۴۹-۶۵.

ودیعتی، م.، نخعی، م.، امرایی، و. و میرعربی، ع. ۱۳۹۲. ارزیابی کیفیت آب رود کارون با استفاده از شبیه‌سازی استنتاج فازی. مجله مهندسی منابع آب، سال ششم، تابستان ۱۳۹۲، صص ۳۹-۴۹.

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.

Haykin, S. 1994. Neural Networks: A comprehensive foundation. MacMillan, New York. 432-444.

Longqin, X. and Shuangyin, L. 2013. Study of short-term water quality prediction model based on wavelet

نتایج تحقیق نشان داد که موجک با درک مکانیسم سری زمانی قادر است تفسیری را از ساختار سری زمانی و اطلاعات مهم و کلیدی را از حافظه آن استخراج کند. مشخص شد که روش رگرسیون دقت لازم برای پیش‌بینی را نداشته و به خصوص در پیش‌بینی نقاط بیشینه و کمینه، بسیار ضعیف عمل می‌کند، در حالیکه دو مدل دیگر پیش‌بینی مناسبی داشتند. مدل ترکیبی موجک عصبی توانست خطا را در مدل‌های شبکه عصبی و رگرسیون کاهش دهد و ضریب کارایی E را بهبود ببخشد. برای تجزیه سری‌های زمانی از موجک مادر Db2 استفاده شد که از نظر شکل هندسی به سیگنال سری زمانی سولفات و کلراید شبیه‌تر بود. در این مقاله با مدل پیشنهادی ارائه شده (که از آنالیز موجک استفاده می‌کند)، نتایج مدل شبکه عصبی بهبود یافت. نوآوری این تحقیق بررسی پدیده هیستریزس در دو پارامتر کیفی کلراید و سولفات بود و نتایج نشان داد که شبکه عصبی و مدل ترکیبی عصبی-موجک به خوبی این پدیده را تشخیص می‌دهند، در صورتیکه رگرسیون قادر به نشان دادن این پدیده نمی‌باشد.

پیشنهاد می‌شود در تحقیقات بعدی به منظور کاهش اغتشاشات سیگنال از سری‌های زمانی روزانه یا هفتگی در بازه زمانی طولانی‌تری استفاده شود. هم‌چنین از پارامتر تبخیر به عنوان سیگنالی تأثیرگذار در روابط هیدرولوژیکی این منطقه در ترکیبات ورودی به مدل شبکه عصبی استفاده شود.

- prediction of daily suspended sediment load in rivers. *Science of The Total Environment*. 409: 2917-2928.
- Rajaei, T., Nourani, V., Zounemat-Kermani, M and Kisi, O. 2011. River Suspended Sediment Load Prediction: Application of ANN and Wavelet Conjunction Model. *Journal of Hydrologic Eng., ASCE*. 8: 613-627.
- Sandhu, N., and Finch, R. 1995. Methodology for flow and salinity estimation in the Sacramento-San Joaquin Delta and Suisun Marsh, Chapter 7. *Artificial neural networks and their applications*, 85: 343-365.
- Singh, R.M. 2012. Wavelet-ANN Model for Flood Events. *Advances in Intelligent and Soft Computing*, Springer. 131: 165-175.
- Williams, G. 1989. Sediment concentration versus water discharge during single hydrologic events in rivers. *Journal of Hydrology*. 111: 89-106.
- Zhou, H.C., Peng, Y., and Liang, G.H. 2008. The research of monthly discharge predictor-corrector model based on wavelet decomposition. *Water Resour Manage, Springer*. 22(2): 217-227.
- neural network. *Elsevier, Journal. Mathematical and Computer Modelling*. 58: 807-813.
- Najah, A., Elshafie, A., Karimi, O and Jaffer, O. 2009. Prediction of Johor river water quality Parameters using artificial neural networks. *Europ J Sci Res*. 28: 422-35.
- Nourani, V., Alami, M.T and Aminfar, M.H.. 2009. A Combined Neural-Wavelet Model for Prediction of Ligvanchai Watershed Precipitation. *Engineerig Application of Artificial Intelligence*. 22: 146-477.
- Nourani, V., Hosseini Baghanam, A., Adamowski, A and Gebremichael, M. 2013. Using self-organizing maps and wavelet transforms for space-time pre-processing of satellite precipitation and runoff data in neural network based rainfall-runoff modeling. *Elsevier, Journal of Hydrology*. 476: 228-243.
- Qian-Jin, L., Zhi-Hue, S., Nu-Fang, F., Hua-De, Z and Lei, A. 2013. Modeling the daily suspended sediment concentration in a hyper concentrated river on the Loess Plateau, China, using the Wavelet-ANN approach. *Elsevier, Journal of Geomorphology*. 186: 181-190.
- Rajaei, T. 2011. Wavelet-ANN combination model for

## Wavelet and ANN Combination Model for Prediction of Sulfate and Chloride Concentrations in Rivers (Case Study: Dez River)

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

Received: Apr. 22, 2014      Accepted: Aug. 27, 2014

### Abstract

One of the most important factors in sustainable development of watersheds is suitable water resources availability in terms of quantity and quality. This study considered artificial neural network (ANN), wavelet analysis and ANN combination (WANN), multi linear regression (MLR) models for monthly Sulfate (SO<sub>4</sub>) and Chloride (Cl) modeling in the Dez River Bamdezh station and investigates the effects of data preprocessing on model performance using discrete wavelet. For this purpose, wavelet analysis and ANN model, observed time series of river discharge and SO<sub>4</sub> and Chloride were decomposed at different scales by wavelet analysis. Then, total effective time series of discharge and SO<sub>4</sub>(Cl) were imposed as inputs to the neural network model for prediction of SO<sub>4</sub>(Cl) in one-month- ahead. Results showed that the WANN model performance was better in prediction rather than the ANN and multi linear regression models. The wavelet analysis model produced reasonable predictions for the extreme values. This model dropped the mean absolute percentage error for the multi linear regression model and the ANN model for the Chloride modeling from 0.84 and 0.64, respectively, to 0.52, and SO<sub>4</sub> modeling from 1.7 and 0.95, respectively, to 0.63. Furthermore, the model could be employed to simulate hysteresis phenomenon for SO<sub>4</sub> modeling, while multi linear regression method is incapable in this event.

**Key words:** Artificial neural network, Wavelet transform, Sulfate(SO<sub>4</sub>), Chloride(Cl), Dez River.

---

1- Assistant Professor, Dept., of Civil Eng., University of Qom. Qom

2- Former MSc Student in Hydraulic Structure, Dept., of Civil Eng., University of Qom, [hjafari@yahoo.com](mailto:hjafari@yahoo.com)

(\* -Corresponding Author Email: [Taher\\_rajaei@yahoo.com](mailto:Taher_rajaei@yahoo.com).)