

مدیریت و تحلیل غیرخطی سیستم گندزدایی شبکه‌های توزیع آب با استفاده از روش‌های داده‌محور

محمد ذونعمت کرمانی^۱، مجید جمالیزاده^۲، عبدالله رضانی چرمهینه^۳

۱- دانشیار، بخش مهندسی آب، دانشگاه شهید باهنر کرمان، کرمان، ایران
(نویسنده مسئول) zounemat@uk.ac.ir

۲- دانش‌آموخته کارشناسی ارشد مهندسی و مدیریت ساخت، گروه عمران، دانشگاه آزاد اسلامی واحد سیرجان، سیرجان، ایران

۳- دانش‌آموخته کارشناسی ارشد مهندسی منابع آب، بخش مهندسی آب، دانشگاه شهید باهنر کرمان، کرمان، ایران

پذیرش ۹۵/۶/۱

(دریافت ۹۴/۹/۳)

چکیده

به منظور تأمین آب آشامیدنی سالم و حذف عوامل بیماری‌زا در شبکه‌های توزیع آب، از واحد کلرزنی استفاده می‌شود. یکی از شیوه‌های مناسب برای تحلیل عملکرد کلر درون شبکه آبرسانی، استفاده از روش‌های داده‌محور است. در پژوهش حاضر شبکه عصبی پرسپترون چندلایه (MLP) با سه الگوریتم آموزش کاهش شیب، شیب مزدوج و BFGS به همراه ماشین بردار پشتیبان (SVM) با تابع کرنل RBF در برآورد غلظت کلر باقیمانده در شبکه‌های آبرسانی احمدآباد دثفه و اهرویه واقع در استان کرمان مورد استفاده قرار گرفت. در این پژوهش داده‌های روزانه شامل دبی آب تولیدی، مصرف کلر و کلر باقیمانده از ابتدای سال ۱۳۹۱ تا پایان ۱۳۹۳ به مدت ۳ سال به کار گرفته شد. به منظور ارزیابی عملکرد مدل‌های مورد بررسی از معیارهای راندمان نش ساکلیف (NS)، ریشه میانگین مربعات خطا (RMSE)، میانگین درصد خطای مطلق (MAPE) و ضریب همبستگی (CORR) استفاده شد. این مقادیر در بهترین حالت مدل‌سازی به ترتیب برابر با 0.9484 ، 0.255 ، 1.081 و 0.974 بودند. با توجه به معیارهای ارزیابی، مدل MLP با الگوریتم BFGS در ۹۰ درصد و با الگوریتم شیب مزدوج در ۱۰ درصد موارد نسبت به نتایج کل مدل‌ها برتری داشتند، در حالی که مدل MLP بر پایه الگوریتم کاهش شیب و مدل SVM در هیچ یک از موارد برتر نبوده‌اند. با توجه به نتایج حاصل از پژوهش حاضر، مدیریت مناسب غلظت کلر را می‌توان با توجه به مقادیر پیش‌بینی شده میزان کلر باقیمانده، در شبکه آبرسانی اعمال نمود به طوری که کاهش عملکرد شبکه پرسپترون و ماشین بردار پشتیبان در شبکه آبرسانی اهرویه نسبت به شبکه احمدآباد دثفه را می‌توان ناشی از مدیریت ناصحیح کلرزنی دانست.

واژه‌های کلیدی: گندزدایی، روش‌های داده‌محور، کلر باقیمانده، شبکه پرسپترون، ماشین بردار پشتیبان

۱- مقدمه

پژوهش‌هایی در داخل و خارج از کشور انجام شده است که در ادامه به برخی از آن‌ها اشاره می‌شود. در سال ۱۳۸۶، عدل و همکاران خطر نشت گاز کلر را در ایستگاه‌های کلرزنی سیستم آب در شهر تهران با استفاده از روش تجزیه و تحلیل درخت خطا مورد ارزیابی قرار دادند. بر اساس یافته‌های حاصل از پژوهش یاد شده، یکی از مهم‌ترین عوامل نامطلوب بودن ایستگاه کلرزنی، ناشی از خطای پرسنل و فشار کاری است (Adl et al. 2008). میرانزاده و همکاران در سال ۱۳۹۰ رابطه بین میزان کدورت، کیفیت میکربی و غلظت کلر باقیمانده در آب آشامیدنی روستاهای شهرستان کاشان را مورد بررسی قرار دادند. نتایج مبین مطلوب نبودن بیش از ۷۵ درصد میزان کلر باقیمانده در آب مصرفی روستاها بوده است (Miranzadeh et al. 2011). صالحی‌آرتمانی و همکاران در سال ۱۳۹۰ مدل‌سازی و ارزیابی ریسک انتشار گاز کلر در تصفیه‌خانه آب جلالیه در شهر تهران را به منظور حفاظت جمعیت و

از بین‌گندزادهای مختلف، کلر و مشتقات آن به دلیل پایین بودن هزینه نسبی، سهولت کاربرد و قابلیت از بین بردن میکروارگانیزم‌های بیماری‌زا در شبکه‌های توزیع آب آشامیدنی، به‌طور گسترده به‌کار می‌روند (Tabesh et al. 2011). یکی از شیوه‌های مناسب برای تحلیل عملکرد کلر درون شبکه آبرسانی، استفاده از روش‌های غیرخطی داده‌محور از جمله شبکه عصبی مصنوعی^۱ و ماشین بردار پشتیبان^۲ است، زیرا افزون بر لزوم مدیریت مناسب هزینه در زمینه مصرف کمتر کلر، رعایت ضوابط استاندارد در مورد میزان کلر باقیمانده در آب نیز الزامی است (The Standard and Industrial Research Organization 1997). در ارتباط با مدل‌سازی و تحلیل کلر در شبکه‌های آبرسانی،

¹ Artificial Neural Network (ANN)

² Support Vector Machine (SVM)

استفاده از یک شیوه نیمه تجربی پرداختند. یافته‌های به دست آمده نشان از توانایی مناسب مدل نیمه تجربی در شبیه‌سازی روند اضمحلال دی اکسید کلر داشته است (Ammer et al. 2014). در سال ۲۰۱۴، نجاری و همکاران اضمحلال کلر را در یک شبکه آبرسانی در شهر بارسلون اسپانیا مدل سازی نمودند. به این منظور از روش الگوریتم ژنتیک برای بهینه‌سازی شبکه استفاده شد (Nejjari et al. 2014). پیرو و ستانسکو در سال ۲۰۱۵ میزان غلظت کلر باقیمانده درون شبکه را با استفاده از مدل عددی بررسی کردند (Perju & Stanescu 2015). در سال ۲۰۱۵، وو و همکاران با استفاده از شبکه‌های عصبی مصنوعی و الگوریتم ژنتیک نسبت به کنترل بهینه کلر در شبکه توزیع آب در استرالیا اقدام کردند (Wu et al. 2015). کارادیرک و همکاران در سال ۲۰۱۶ مدل سازی و پیش‌بینی میزان غلظت کلر را در یک شبکه آبرسانی در ترکیه با به کارگیری روش‌های استوکستیک خودبازگشتی انجام دادند.

نتایج بیانگر قابلیت مناسب روش‌های خودبازگشتی در پیش‌بینی غلظت کلر بوده است (Karadirek et al. 2016). در سال ۲۰۱۶، کارادیرک و همکاران به بررسی میزان کلر در ۱۸ شبکه آبرسانی شهری با استفاده از یک سامانه برخط پرداختند. نتایج کارایی مناسب سامانه در گردآوری اطلاعات و مدیریت کیفی شبکه را نشان داده است (Karadirek et al. 2016).

از یک سوا اهمیت زیاد میزان کلر باقیمانده در سیستم گندزدایی شبکه‌های توزیع آب و از طرف دیگر توانایی بالای شبکه عصبی مصنوعی و ماشین بردار پشتیبان به عنوان روش‌های داده محور در کشف روابط غیرخطی و پیچیده در بین پارامترهای یک سیستم موجب شد که در این پژوهش مدل سازی میزان کلر باقیمانده با روش‌های ANN و SVM انجام شود. این پژوهش با هدف ارزیابی کارایی روش‌های داده محور شامل شبکه‌های عصبی مصنوعی به عنوان یکی از کارآمدترین روش‌های داده محور و ماشین بردار پشتیبان به عنوان یکی از روش‌های نوین داده محور به منظور مدیریت و تحلیل سیستم گندزدایی شبکه آبرسانی با استفاده از تخمین غلظت کلر باقیمانده انجام شد. برای نیل به این هدف آمار و اطلاعات دو شبکه آبرسانی روستایی احمدآباد دثقه و اهرویه واقع در استان کرمان مدل سازی شد.

محیط زیست از ذخیره‌سازی این ماده سمی مورد ارزیابی قرار دادند. با توجه به مضرات ناشی از نشت گاز کلر، پیشنهاد حاصل از این پژوهش بر پایه استفاده از روش‌های نوین و جایگزین برای گندزدایی آب بنا شده است (Salehi Artimani et al. 2012). در سال ۱۳۹۰، تابش و همکاران به بهینه‌سازی میزان تزریق کلر در شبکه‌های آبرسانی توسط الگوریتم ژنتیک پرداختند. در پژوهش صورت گرفته میزان تزریق کلر در محل‌های تزریق با تلفیق یک مدل تحلیل هیدرولیکی و مدل سازی کیفی با وجود یک مدل غیرخطی الگوریتم ژنتیک بهینه شد. نتایج حاصل از این بهینه‌سازی، حاکی از کاهش کلر مصرفی در شبکه بوده است و درصد قرارگیری مقدار کلر باقیمانده در شبکه در محدوده استاندارد افزایش نشان داده است (Tabesh et al. 2011). اصل‌هاشمی در سال ۱۳۹۱ کلر باقیمانده در منابع آب شرب شهری و روستایی استان ایلام را مورد بررسی قرار داد. نتایج نشان داد که کلر باقیمانده آب از کل نمونه‌های گرفته شده، در مراکز شهری ۹۰ درصد مطلوب و ۱۰ درصد نامطلوب و در مناطق روستایی، ۷۵ درصد مطلوب و ۲۵ درصد نامطلوب بوده‌اند (Aslhashemi 2012). در پژوهش رودریگز و سروزد در سال ۱۹۹۹، روش‌های مدل سازی خطی تجربی و غیرخطی به منظور مدل نمودن میزان کلر باقیمانده در شبکه‌های آبرسانی شهری بررسی شد (Rodriguez & Serodes 1999).

بادن و همکاران در سال ۲۰۰۶ با استفاده از شبکه عصبی رگرسیونی، به پیش‌بینی میزان کلر باقیمانده درون شبکه‌های آبرسانی در شهر آدلاید استرالیا پرداختند. یافته‌ها بیانگر مناسب بودن شبکه عصبی مصنوعی رگرسیونی و معرفی آن به عنوان ابزاری کارآمد برای پیش‌بینی وضعیت کلر باقیمانده در شبکه‌های آبرسانی بوده است (Bowden et al. 2006). گیبس و همکاران در سال ۲۰۰۶ روابط میان اضمحلال کلر و پارامترهای توزیع آب را با استفاده از روش‌های داده محور مورد کنکاش قرار دادند. برای این منظور از روش داده محور شبکه عصبی پرسپترون^۱، رگرسیون خطی چندگانه و شبکه عصبی رگرسیونی استفاده شد. نتایج معرف برتری شبکه عصبی پرسپترون نسبت به سایر روش‌های داده محور بوده است (Gibbs et al. 2006). در سال ۲۰۱۴، آمر و همکاران به پیش‌بینی روند اضمحلال دی اکسید کلر در شبکه‌های آبرسانی با

¹ Perceptron Neural Network

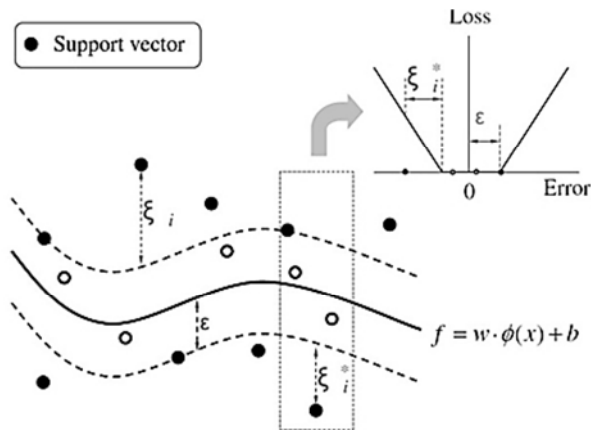


Fig. 1. Insensitive loss function of ε (Yu et al. 2006)
شکل ۱- تابع تلفات ε غیرحساس (Yu et al. 2006)

ساختاری، محدوده خطا به وسیله مسئله بهینه‌سازی در رابطه ۲ بهینه می‌شود

$$\text{Minimize: } \frac{1}{2} (w \cdot w) + C \sum_{i=1}^N (\xi_i + \xi_i^*) \quad (2)$$

$$\text{Constraints: } \begin{cases} y_i - [w \cdot \phi(x_i) + b] \leq \varepsilon + \xi_i \\ [w \cdot \phi(x_i) + b] - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, \quad i = 1, 2, \dots, N \end{cases}$$

که $\phi(x_i)$ پارامتر یا ویژگی ورودی و y_i مقدار هدف است. ε محدوده لوله تقریب را تعیین می‌کند و C ثابت گنجایش است که خطای مربوط به انحراف بیشتر از ε را کنترل می‌کند. برای حل مسئله باید مقادیر C ، ε و پارامترهای مربوط به تابع کرنل^۵ انتخابی مشخص باشند که دقت در تعیین آنها در کاهش خطای مسئله تأثیرگذار است (Dibike et al. 2001). شکل کلی تابع کرنل در رابطه ۳ مشخص شده است که نداشت پارامترها و الگوهای ورودی را با معرفی توابع ویژگی $\phi(x)$ امکان‌پذیر می‌کند. در این حالت که در فضای نگاشت یافته می‌توان به صورت خطی رگرسیون‌گیری کرد

$$K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j) \quad (3)$$

تابع کرنل RBF^۶ رابطه ۴ به‌عنوان یکی از پرکاربردترین توابع در ماشین‌بردار پشتیبان معرفی شده است (Yu et al. 2006)

⁵ Kernel function

⁶ Radial Basis Function (RBF)

۲- مواد و روش‌ها

در پژوهش حاضر از روش‌های داده‌محور با قابلیت‌های فراگیری در مرحله آموزش و صحت‌سنجی در مرحله آزمون با محوریت مدل‌سازی غلظت کلر استفاده شده است. از این رو در ادامه به معرفی روش‌های داده‌محور مورد استفاده شامل شبکه عصبی مصنوعی و ماشین‌بردار پشتیبان پرداخته شده است.

۱-۲- شبکه عصبی مصنوعی

شبکه عصبی پرسپترون چند لایه^۱ از یک لایه ورودی، تعدادی لایه پنهان و یک لایه خروجی تشکیل شده است. مدل عمومی شبکه پرسپترون، شبکه جلورونده با روال آموزش انتشار به عقب است. روال انتشار به عقب به این معناست که پس از مشخص شدن خروجی شبکه، ابتدا وزن‌های لایه پایانی تصحیح شده و سپس به ترتیب، وزن‌های لایه‌های قبلی تصحیح می‌شوند (Abraham 2005).

۲-۲- ماشین‌بردار پشتیبان

در مدل SVM از مجموعه توابع به صورت $f(x) = w \cdot \phi(x) + b$ برای پیش‌بینی استفاده می‌شود که w وزن بردار ϕ و b مقدار بایاس^۲ می‌باشند. برای کاربرد ماشین‌بردار پشتیبان در مسائل رگرسیون، از تابع تلفات^۳ به نام ε غیرحساس^۴ استفاده می‌شود که خطاهایی که در یک فاصله معین از مقادیر واقعی می‌باشند را نادیده می‌گیرد (شکل ۱) (Smola & Scholkopf 2004).

بنابراین امکان دارد میزان خطای کمتر از ε را لحاظ کرد و باید مقداری انحراف از ε را مجاز شمرد. این انحراف به صورت رابطه ۱ بیان می‌شود

$$|\xi|_{\varepsilon} = \begin{cases} 0 & \text{if } |\xi| \leq \varepsilon \\ |\xi| - \varepsilon & \text{otherwise} \end{cases} \quad (1)$$

انحراف مذکور با منظور نمودن متغیرهای کمبود ξ_i و ξ_i^* در رابطه ۲ لحاظ می‌شود و در ادامه بر اساس اصل کمینه‌سازی خطای

¹ Multi-layer Perceptron Neural Network (MLP)

² Bias

³ Loss function

⁴ ε -insensitive

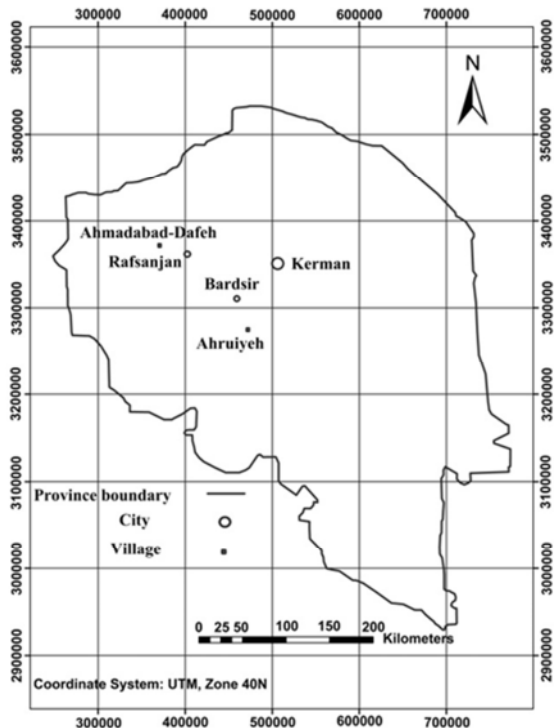


Fig. 2. Geographical location of the studied villages in Kerman Province

شکل ۲- موقعیت جغرافیایی روستاهای مورد استفاده در تحقیق حاضر در استان کرمان

آورده شده است.

برای به دست آوردن ساختار مناسب پارامترهای ورودی به مدل های مورد پژوهش از روش رگرسیون گام به گام^۳ استفاده شد. از آماره MC^4 برای ترتیب اضافه شدن پارامترهای ورودی تأخیر یافته و در نتیجه تشکیل ساختار ترکیب های ورودی به مدل های شبکه عصبی و ماشین بردار پشتیبان در هر شبکه آبرسانی استفاده شد. نحوه اضافه شدن پارامترهای ورودی و در نتیجه ایجاد ترکیب های ورودی به مدل های MLP و SVM بر اساس ترتیب پارامترها در محور افقی نمودار MC است. این انتخاب بر اساس بیشترین رابطه خطی استوار است تا عملکرد هر یک از مدل ها در کشف روابط غیرخطی که قطعاً روابط خطی را نیز شامل می شود، به صورت کامل بررسی شود. به عنوان نمونه، نمودار MC برای شبکه آبرسانی احمدآباد دثفه در شکل ۴ مشخص شده است که نشان می دهد پارامتر RC-1 رابطه بیشینه خطی با خروجی RC داشته

³ Stepwise Regression

⁴ Mallow's C_p

$$K(x_i, x_j) = \exp(-\gamma |x_i - x_j|^2) \quad (4)$$

۳-۲- مطالعه موردی

در پژوهش حاضر از داده های روزانه مربوط به شبکه های آبرسانی روستاهای احمدآباد دثفه رفسنجان و اهرویه بردسیر واقع در استان کرمان از ابتدای سال ۱۳۹۱ تا پایان ۱۳۹۳ (به مدت ۳ سال) استفاده شد. این داده ها شامل پارامترهای دبی آب تولیدی روزانه (Q)، مصرف روزانه کلر (C) و کلر باقیمانده روزانه (RC) و به ترتیب بر حسب متر مکعب در روز، کیلوگرم و میلی گرم در لیتر بودند. در ورودی هر شبکه آبرسانی در هر روز یک مرتبه برداشت انجام شد. در شکل ۲، موقعیت جغرافیایی روستاهای مورد پژوهش در استان کرمان مشخص شده است.

۴-۲- نحوه آماده سازی داده های ورودی به مدل

با توجه به این که داده های مورد پژوهش، سری زمانی (حفظ ترتیب داده ها به صورت روزانه) تشکیل می دهند، ابتدا تأخیر زمانی پارامترهای ورودی به مدل (C, Q, RC) با توجه به خروجی از آن (RC) مشخص شد. در این راستا از توابع خودهمبستگی^۱ و همبستگی متقابل^۲ استفاده شد. برای مثال میزان تأخیر زمانی اعمال شده بر روی پارامترهای ورودی به مدل های MLP و SVM در شبکه آبرسانی احمدآباد دثفه، با توجه به بیشترین مقدار همبستگی میان پارامتر ورودی و خروجی مدل در شکل ۳ با علامت پیکان مشخص شده است. در این نمودارها، محور عمودی میزان همبستگی و محور افقی میزان تأخیر زمانی بر حسب روز را نمایش می دهند.

لازم به ذکر است که پارامترهای Q و C به اندازه ۱ مرتبه و پارامتر RC به علت اهمیت بیشتر تا ۳ مرتبه تأخیر داده شدند. مطابق شکل ۳، در شبکه آبرسانی احمدآباد دثفه به ترتیب پارامترهای ورودی RC به اندازه ۱، ۲ و ۳ روز (RC-1, RC-2, RC-3)، Q به اندازه ۹ روز (Q-9) و C به اندازه ۱ روز (C-1) تأخیر پیدا کرده اند. به همین ترتیب پارامترهای شبکه آبرسانی اهرویه نیز دچار تأخیر شدند که نتیجه کل تأخیرهای صورت گرفته روی ورودی های مدل در جدول ۱

¹ Autocorrelation Function (ACF)

² Cross Correlation Function (CCF)

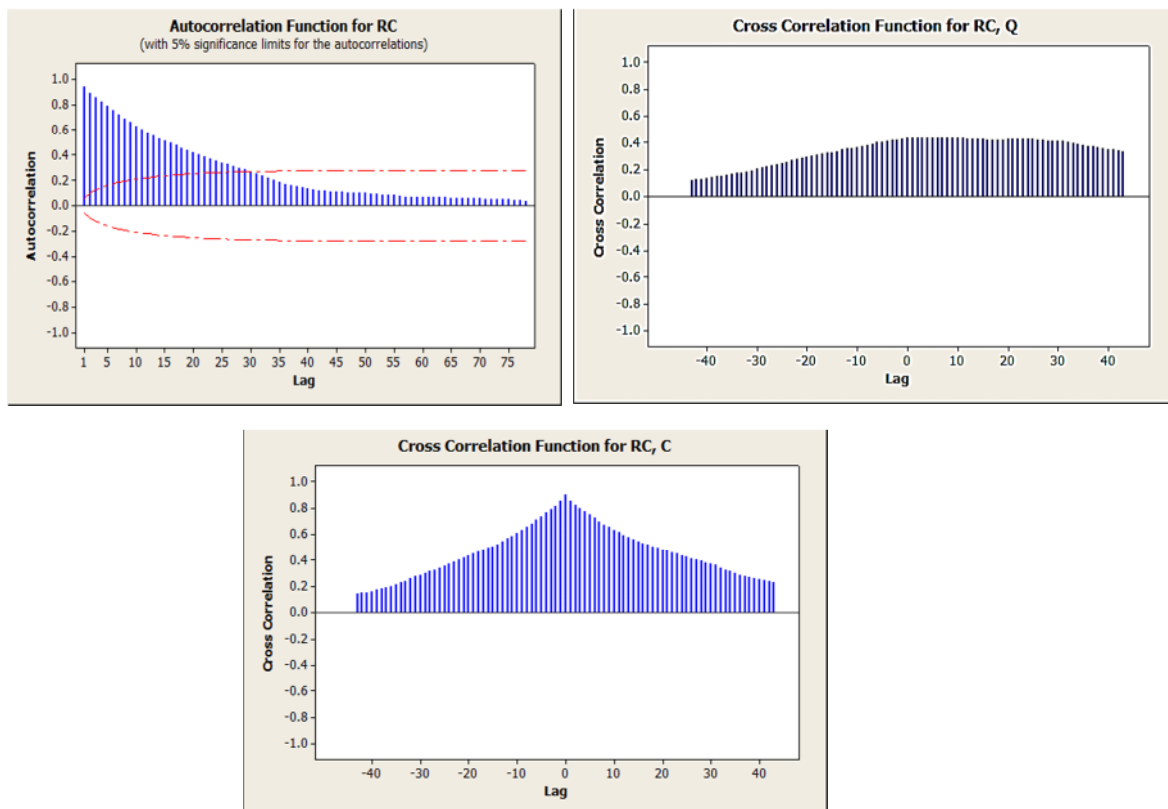


Fig. 3. ACF and CCF plots for Ahmadabad-Dafeh water supply network: (a) ACF daily residual chlorine, (b) CCF of daily water discharge and daily residual chlorine, (c) CCF of daily consumption and daily residual chlorine
شکل ۳- نمودارهای ACF و CCF شبکه آبرسانی احمدآباد دتفه؛ (a) نمودار ACF کلر باقیمانده روزانه، (b) نمودار CCF دیبی روزانه و کلر باقیمانده روزانه، (c) نمودار CCF مصرف روزانه کلر و کلر باقیمانده روزانه

است (مقدار حداکثر MC) که پس از اضافه شدن پارامتر RC-3 این رابطه خطی کاهش شدیدی پیدا کرده است تا با افزایش دقت و کاهش انحراف پیش‌بینی‌کننده‌ها (پارامترهای ورودی) از پاسخ (خروجی مدل)، ایجاد توازن بیشتر در مدل‌سازی فراهم شود. ساختار ترکیب‌های ورودی به مدل‌های MLP و SVM در هر شبکه آبرسانی به صورت جدول ۲ بوده است.

۲-۵- ویژگی مدل‌های به‌کار رفته در تحقیق حاضر

یکی از مدل‌های به‌کار رفته در این پژوهش، شبکه عصبی پرسپترون چند لایه (MLP) شامل یک لایه ورودی، یک لایه پنهان و یک لایه خروجی بوده است که در این شبکه سه لایه از توابع فعال‌سازی سیگموئید لجستیک و تانژانت‌هایپربولیک (Tanh) و الگوریتم‌های آموزش^۱ BFGS، کاهش شیب^۲ و شیب

جدول ۱- میزان تأخیر پارامترهای ورودی به مدل‌های MLP و SVM

Table 1. The amount of lags in input parameters to MLP and SVM models

Water supply network	Q lag (day)	C lag (day)	RC lag (day)
Ahmadabad-Dafeh	9	1	1,2,3
Ahruiyeh	5	1	1,2,3

جدول ۲- ساختار ترکیب ورودی به مدل‌های MLP و SVM

Table 2. Structure of input combinations to MLP and SVM models

Water supply network	Compound No.	Compound structure
Ahmadabad-Dafeh	1	RC-1
	2	RC-1, RC-3
	3	RC-1, RC-3, C-1
	4	RC-1, RC-3, C-1, Q-9
	5	RC-1, RC-3, C-1, Q-9, RC-2
Ahruiyeh	1	RC-1
	2	RC-1, RC-3
	3	RC-1, RC-3, C-1
	4	RC-1, RC-3, C-1, RC-2
	5	RC-1, RC-3, C-1, RC-2, Q-5

¹ Broyden-Fletcher-Goldfarb-Shanno (BFGS)

² Gradient descent

که O_i مقدار اندازه‌گیری شده، P_i مقدار پیش‌بینی شده با مدل، \bar{O} میانگین مقدار اندازه‌گیری شده، \bar{P} میانگین مقدار پیش‌بینی شده با مدل و n تعداد مقادیر اندازه‌گیری شده یا پیش‌بینی شده می‌باشند. هر چه مقادیر NS و CORR به یک نزدیک‌تر باشند کارایی مدل بیشتر و هر چه مقادیر RMSE و MAPE به صفر نزدیک‌تر باشند دقت مدل بیشتر است.

به منظور مدل‌سازی، ۸۰ درصد اول داده‌ها برای آموزش و ۲۰ درصد پایانی برای آزمون به‌کار رفته است. الگوریتم BFGS در اولین گام معادل با روش کاهش شیب است و بر خلاف الگوریتم شیب مزدوج با ماتریس هسین اصلاح می‌شود. بنابراین در ابتدا با به‌کارگیری الگوریتم BFGS، آرایش شبکه و توابع فعال‌سازی لایه‌های پنهان و خروجی در مدل شبکه عصبی پرسپترون تعیین شد و در ادامه الگوریتم‌های کاهش شیب و شیب مزدوج نیز استفاده شدند. مشخصات ترکیب‌های ورودی به مدل MLP در جدول ۳ آورده شده‌اند. همچنین نتایج پارامترهای مدل ماشین‌بردار پشتیبان در جدول ۴ ارائه شده‌اند.

۳- نتایج و بحث

در جدول‌های ۵ و ۶ نتایج معیارهای ارزیابی در ترکیب‌های ورودی به مدل‌های MLP و SVM مشخص شده‌اند تا عملکرد شبکه عصبی پرسپترون و ماشین‌بردار پشتیبان در برآورد غلظت کلر باقیمانده در شبکه‌های آبرسانی مورد پژوهش، بررسی شود. مشخص است در مراحل آموزش و آزمون، شبکه پرسپترون سه لایه مبتنی بر الگوریتم آموزش BFGS دارای بهترین عملکرد بوده است به طوری که برتری آن نسبت به مدل‌های ماشین‌بردار پشتیبان و شبکه عصبی بر مبنای الگوریتم‌های شیب مزدوج و کاهش شیب کاملاً مشهود است. نتایج جدول‌های ۵ تا ۷ علاوه بر مشخص ساختن برتری مدل MLP بر SVM در پژوهش حاضر، حاکی از اهمیت الگوریتم آموزش در عملکرد شبکه عصبی پرسپترون نیز می‌باشد به طوری که شبکه پرسپترون به‌کار رفته برای مدل‌سازی میزان کلر باقیمانده در شبکه‌های آبرسانی مورد بررسی در این پژوهش، در زمان استفاده از الگوریتم BFGS به‌ترتیب در مراحل آموزش و آزمون در ۸۰ و ۹۰ درصد نسبت به نتایج کل مدل‌ها برتری داشته است. علت این موضوع می‌تواند به دلیل استفاده از

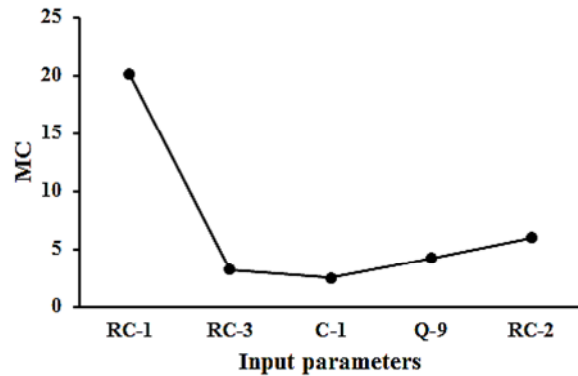


Fig. 4. MC plot for Ahmadabad-Dafeh water supply network
شکل ۴- نمودار MC برای شبکه آبرسانی احمدآباد دنفه

مزدوج^۱ استفاده شده است. ماشین‌بردار پشتیبان (SVM) با تابع کرنل RBF، دیگر مدل مورد بررسی در پژوهش حاضر است. بنابراین در این پژوهش، سه شبکه عصبی پرسپترون تحت نظارت با الگوریتم‌های آموزش متفاوت و ماشین‌بردار پشتیبان مورد بررسی قرار گرفتند. پارامترهای ورودی به مدل‌ها در قالب ترکیب‌های ورودی در جدول ۲ و نحوه استخراج آن‌ها در ارتباط با خروجی مدل در بخش ۲-۴ شرح داده شده‌اند.

به منظور ارزیابی عملکرد مدل‌های مورد بررسی از معیارهای راندمان نش ساکلیف (NS)، ریشه میانگین مربعات خطا (RMSE)، میانگین درصد خطای مطلق (MAPE) و ضریب همبستگی (CORR) استفاده شد (معادلات ۵ تا ۸).

$$NS = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad (5)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2} \quad (6)$$

$$MAPE = 100 \times \frac{1}{n} \left(\sum_{i=1}^n \frac{|O_i - P_i|}{O_i} \right) \quad (7)$$

$$CORR = \frac{\sum_{i=1}^n (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2} \sqrt{\sum_{i=1}^n (P_i - \bar{P})^2}} \quad (8)$$

¹ Conjugate gradient

جدول ۳- مشخصات ترکیب‌های ورودی به مدل MLP
Table 3. Characteristics of input combinations to MLP model

Water supply network	Compound No.	Network structure	Activation function	
			Hidden layer	Output layer
Ahmadabad-Dafeh	1	1-7-1	Tanh	Logistic Sigmoid
	2	2-9-1	Logistic Sigmoid	Logistic Sigmoid
	3	3-7-1	Logistic Sigmoid	Logistic Sigmoid
	4	4-12-1	Logistic Sigmoid	Logistic Sigmoid
	5	5-6-1	Logistic Sigmoid	Tanh
Ahruiyeh	1	1-3-1	Tanh	Tanh
	2	2-8-1	Logistic Sigmoid	Logistic Sigmoid
	3	3-7-1	Logistic Sigmoid	Tanh
	4	4-5-1	Tanh	Logistic Sigmoid
	5	5-8-1	Tanh	Logistic Sigmoid

مدل‌سازی میان متغیرهای موجود در یک سیستم که روابط غیرخطی دارند، بهترین نتیجه را دربرنخواهد داشت بلکه توجه به معماری و ساختار شبکه و انتخاب الگوریتم‌های مناسب بر افزایش دقت و عملکرد مدل مورد استفاده تأثیر به‌سزایی دارد.

وجود کلر باقیمانده با غلظت مطلوب در آب به منظور مقابله با آلودگی میکروبی و حفاظت آب در برابر آلودگی ثانویه ضروری است. بنابراین عمل‌گندزدایی مطلوب آب به غلظت مناسب کلر بستگی دارد. از این رو دقت بالا در تخمین غلظت کلر باقیمانده در سیستم‌های گندزدایی مورد بررسی در این پژوهش، نشان‌دهنده توانایی بالای شبکه عصبی مصنوعی و ماشین بردار پشتیبان در این زمینه است.

ناچیز بودن ریشه میانگین مربعات خطا (RMSE) در جدول‌های ۵ و ۶ نشان‌دهنده دقت بالای روش‌های یاد شده در برآورد غلظت کلر باقیمانده است. مدیریت صحیح گندزدایی، حساسیت بالا به میزان کلر باقیمانده در شبکه توزیع را طلب می‌کند زیرا ورود آلودگی به شبکه آبرسانی می‌تواند به شدت روی سلامت جامعه تأثیر بگذارد. همچنین توجه به این مطلب اهمیت میزان کلر را دوچندان می‌کند که علاوه بر رعایت مقدار حداقل کلر باقیمانده به منظور کنترل کیفیت میکروبی آب، باید به مقدار حداکثر کلر برای کنترل مشکلات مربوط به مزه، بوی آب و جلوگیری از تولید فراورده‌های جانبی سمی نیز توجه کافی شود.

در جدول‌های ۵ و ۶ مشخص است که در مرحله آزمون، مدل شبکه پرسپترون با الگوریتم BFGS در ترکیب شماره ۳ در شبکه آبرسانی احمدآباد دغه و ترکیب شماره ۱ در شبکه آبرسانی اهرویه دارای بهترین عملکرد نسبت به سایر مدل‌ها در کل ترکیب‌های

جدول ۴- نتایج پارامترهای مدل SVM

Table 4. Results of SVM model parameters

Water supply network	Compound No.	Training constants		Parameter of RBF kernel
		C	ϵ	γ
Ahmadabad-Dafeh	1	10	0.05	2
	2	11	0.009	1
	3	10	0.03	0.37
	4	13	0.009	0.25
	5	11	0.009	0.68
Ahruiyeh	1	40	0.009	1.5
	2	11	0.008	0.6
	3	13	0.009	0.333
	4	12	0.008	0.25
	5	11	0.009	0.2

مشتقات دوم (ماتریس هسین) در الگوریتم BFGS بر خلاف الگوریتم‌های شیب مزدوج و کاهش شیب باشد که موجب نرخ مناسب یادگیری در شبکه و کشف بهتر روابط غیرخطی شده است. مدل شبکه عصبی با الگوریتم شیب مزدوج نیز در ۱۰ درصد کل موارد بهترین عملکرد آموزش را از خود برجای گذاشته است و در ۱۰ درصد آزمون داده‌ها، بیشترین دقت را دارا بوده است. در حالی که مدل پرسپترون با الگوریتم کاهش شیب در هیچ یک از موارد برتر نبوده است. دلیل احتمالی این نتیجه ناشی از تفاوت همگرایی الگوریتم‌های کاهش شیب و شیب مزدوج است. شرح این موضوع به این صورت است که در روش کاهش شیب، وزن‌ها و بایاس‌ها در جهت بیشترین شیب کاهش تابع خطا اصلاح می‌شوند، با این حال در کل مسیر سریع‌ترین همگرایی و در نتیجه الزاماً مسیر بهینه حاصل نمی‌شود. از طرفی در روش شیب مزدوج، جستجو بین جهت‌های مزدوج باعث همگرایی سریع‌تری نسبت به روش کاهش شیب می‌شود. بنابراین صرفاً استفاده از روش‌های داده‌محور در

جدول ۵- نتایج معیارهای ارزیابی مدل‌های MLP و SVM در شبکه آبرسانی احمدآباد دقنه

Table 5. Results of evaluation criteria of MLP and SVM models in Ahmadabad-Dafeh water supply network

Compound No.	Model	Training				Testing			
		NS	RMSE (mg/L)	MAPE (%)	CORR	NS	RMSE (mg/L)	MAPE (%)	CORR
1	MLP (BFGS)	0.941	0.029	1.560	0.970	0.9482	0.0256	0.935	0.974
	MLP (gradient descent)	0.888	0.040	3.355	0.946	0.9222	0.031	2.557	0.964
	MLP (conjugate gradient)	0.909	0.036	3.111	0.954	0.935	0.029	1.967	0.968
	SVM (RBF)	0.927	0.032	3.063	0.964	0.933	0.029	2.752	0.966
2	MLP (BFGS)	0.942	0.029	1.498	0.971	0.9481	0.0256	0.974	0.974
	MLP (gradient descent)	0.869	0.043	3.660	0.935	0.909	0.034	3.032	0.955
	MLP (conjugate gradient)	0.867	0.044	3.428	0.932	0.921	0.032	2.445	0.960
	SVM (RBF)	0.925	0.033	1.628	0.962	0.935	0.029	1.294	0.968
3	MLP (BFGS)	0.942	0.029	1.650	0.970	0.9484	0.0255	1.081	0.974
	MLP (gradient descent)	0.850	0.046	4.221	0.929	0.904	0.035	2.980	0.953
	MLP (conjugate gradient)	0.857	0.045	3.783	0.928	0.926	0.031	2.166	0.964
	SVM (RBF)	0.898	0.038	2.329	0.948	0.925	0.031	1.855	0.963
4	MLP (BFGS)	0.943	0.029	1.741	0.971	0.943	0.029	1.170	0.971
	MLP (gradient descent)	0.837	0.048	4.476	0.926	0.883	0.039	3.524	0.946
	MLP (conjugate gradient)	0.914	0.035	3.203	0.958	0.929	0.030	2.447	0.966
	SVM (RBF)	0.890	0.040	1.834	0.944	0.921	0.032	1.274	0.960
5	MLP (BFGS)	0.944	0.028	1.541	0.972	0.947	0.026	1.021	0.973
	MLP (gradient descent)	0.883	0.041	3.937	0.947	0.911	0.034	2.605	0.960
	MLP (conjugate gradient)	0.892	0.039	3.837	0.945	0.926	0.031	2.106	0.963
	SVM (RBF)	0.932	0.031	1.916	0.966	0.935	0.029	1.756	0.967

جدول ۶- نتایج معیارهای ارزیابی مدل‌های MLP و SVM در شبکه آبرسانی اهرویه

Table 6. Results of evaluation criteria of MLP and SVM models in Ahruyeh water supply network

Compound No.	Model	Training				Testing			
		NS	RMSE (mg/L)	MAPE (%)	CORR	NS	RMSE (mg/L)	MAPE (%)	CORR
1	MLP (BFGS)	0.8731	0.0394	4.183	0.938	0.609	0.035	2.088	0.783
	MLP (gradient descent)	0.854	0.042	6.141	0.937	0.496	0.039	4.281	0.763
	MLP (conjugate gradient)	0.844	0.044	5.513	0.944	0.573	0.036	1.934	0.757
	SVM (RBF)	0.8732	0.0393	2.683	0.937	0.479	0.40	2.451	0.741
2	MLP (BFGS)	0.894	0.036	3.057	0.946	0.569	0.036	1.974	0.766
	MLP (gradient descent)	0.875	0.039	4.736	0.943	0.590	0.035	2.217	0.774
	MLP (conjugate gradient)	0.875	0.039	4.736	0.943	0.590	0.035	2.217	0.774
	SVM (RBF)	0.881	0.038	2.282	0.6740	0.476	0.040	2.116	0.740
3	MLP (BFGS)	0.904	0.034	2.993	0.951	0.589	0.035	1.904	0.793
	MLP (gradient descent)	0.881	0.038	4.294	0.941	0.538	0.038	2.647	0.759
	MLP (conjugate gradient)	0.883	0.038	4.165	0.942	0.494	0.039	2.664	0.726
	SVM (RBF)	0.879	0.038	2.136	0.939	0.472	0.040	1.932	0.738
4	MLP (BFGS)	0.881	0.038	4.218	0.940	0.553	0.037	1.828	0.761
	MLP (gradient descent)	0.877	0.039	5.181	0.943	0.423	0.042	4.492	0.723
	MLP (conjugate gradient)	0.882	0.038	3.993	0.944	0.526	0.038	2.063	0.741
	SVM (RBF)	0.880	0.038	2.139	0.939	0.469	0.040	1.929	0.737
5	MLP (BFGS)	0.915	0.032	2.277	0.957	0.594	0.035	1.675	0.786
	MLP (gradient descent)	0.873	0.039	5.534	0.943	0.404	0.043	5.069	0.728
	MLP (conjugate gradient)	0.888	0.037	4.218	0.945	0.529	0.038	2.143	0.746
	SVM (RBF)	0.880	0.038	2.041	0.940	0.469	0.040	1.856	0.738

مؤید این مطلب است که عملکرد شبکه پرسپترون در شبکه آبرسانی اهرویه نسبت به شبکه احمدآباد دئفه کاهش یافته است. این موضوع ممکن است ناشی از مدیریت ناصحیح کلرزنی شبکه اهرویه (بر خلاف شبکه احمدآباد دئفه) باشد. در واقع دقت ناکافی در عملیات گندزدایی می تواند به صورت داده‌هایی نامتناسب در فرایند پیش بینی شبکه عصبی اختلال ایجاد کرده باشد. به بیان دیگر مدیریت صحیح کلرزنی موجب ایجاد یک فرایند قابل درک و کشف توسط شبکه عصبی در شبکه احمدآباد دئفه شده است در حالی که ناتوانی شبکه عصبی در مدل سازی غلظت کلر در روزهای ابتدایی و روزهای میانی در داده‌های مورد آزمون در شبکه اهرویه

جدول ۷- درصد برتری مدل های MLP و SVM نسبت به کل موارد

مدل سازی در پژوهش حاضر

Table 7. Superiority of MLP and SVM models over other modeling cases in this study

Model	Priority percentage (%_	
	Training	Testing
MLP (BFGS)	80	90
MLP (gradient descent)	0	0
MLP (conjugate gradient)	10	10
SVM (RBF)	10	0

موجود در همان شبکه آبرسانی بوده است. نمودارهای مربوط به این ترکیب ها در شکل های ۵ و ۶ آورده شده اند. توجه به این نمودارها

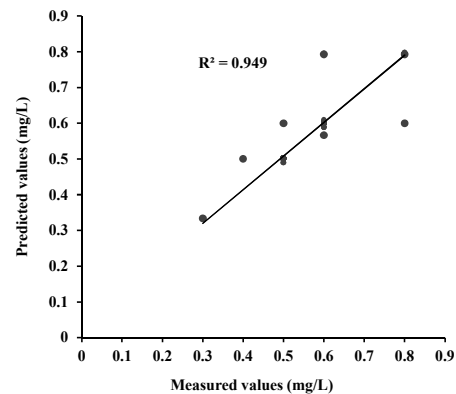
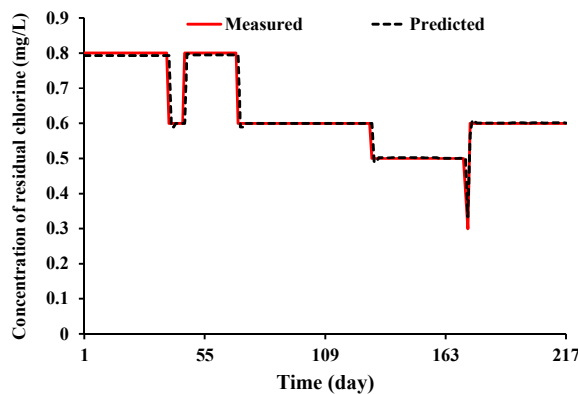


Fig. 5. Measured and predicted values of the best combination in experimental phase for Ahmadabad-Dafeh water supply network (combination No. 3 in MLP-BFGS model): (a) chlorine residual plot, (b) scatter plot

شکل ۵- مقادیر اندازه گیری شده و پیش بینی شده بهترین ترکیب در مرحله آزمون (ترکیب شماره ۳ با مدل MLP-BFGS) مربوط به شبکه آبرسانی احمدآباد دئفه؛ a. نمودار میزان کلر باقیمانده، b. نمودار پراکنش

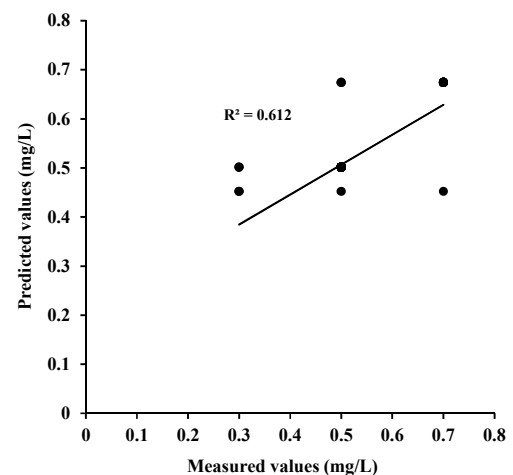
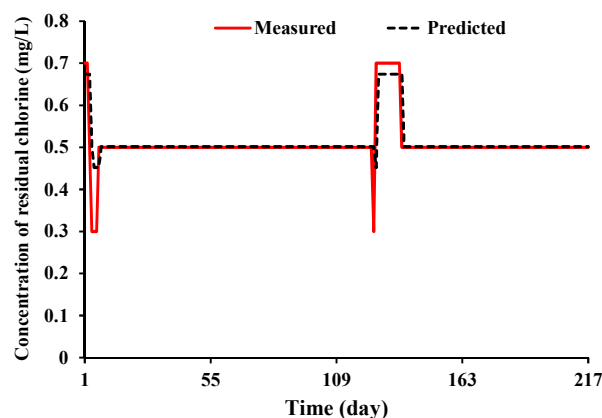


Fig. 6. Measured and predicted values of the best combination in experimental phase of Ahruyeh water supply network (combination No. 1 in MLP-BFGS model): (a) chlorine residual plot, (b) Scatter plot

شکل ۶- مقادیر اندازه گیری شده و پیش بینی شده بهترین ترکیب در مرحله آزمون (ترکیب شماره ۱ با مدل MLP-BFGS) مربوط به شبکه آبرسانی اهرویه؛ a. نمودار میزان کلر باقیمانده، b. نمودار پراکنش

به علاوه نیازی به امکانات و تجهیزات ویژه نیز ندارد. همچنین مزیت منحصر به فرد روش‌های داده‌محور در استفاده از داده‌های اندازه‌گیری شده گندزدایی آب بدون هیچ پیش‌نیاز و ابزار خاص، نه تنها از صرف هزینه نابه‌جا جلوگیری می‌کند، بلکه این روش‌ها (نظیر شبکه عصبی مصنوعی و ماشین بردار پشتیبان) را به‌عنوان ابزاری مناسب و کارآمد در سیستم مدیریت گندزدایی معرفی می‌کند. از جمله نتایج دیگر این پژوهش می‌توان به موارد زیر اشاره کرد:

۱- مدل MLP با الگوریتم‌های آموزش BFGS و شیب مزدوج بر مدل SVM با تابع کرنل RBF در تخمین غلظت کلر باقیمانده برتری دارد.

۲- شبکه‌های آبرسانی احمدآباد دثفه و اهرویه الگوی متفاوتی برای کلر زنی دارند.

۳- مدیریت ناصحیح کلر زنی را می‌توان از دلایل کاهش عملکرد شبکه پرسپترون و ماشین بردار پشتیبان در شبکه آبرسانی اهرویه نسبت به شبکه احمدآباد دثفه برشمرد.

۴- میزان نوسان کلر باقیمانده در شبکه آبرسانی در طول زمان، با شبکه عصبی قابل پیش‌بینی است. این تغییرات غلظت کلر ممکن است ناشی از تغییر شرایط محیطی و حتی اجتماعی باشد.

(شکل ۶) را می‌توان به سوء مدیریت در مصرف کلر نسبت داد. این موضوع نیز منطقی‌تر است که غلظت کلر در زمان‌های مختلف مانند شکل ۵ دارای نوسان نسبی باشد زیرا دلایلی مانند گرم‌تر بودن هوا یا افزایش مهاجرت‌ها در بعضی از فصول و احتمال بیشتر سرایت بیماری، میزان بیشتری از کلر را در شبکه آبرسانی طلب می‌کند. نوسان کلر در طول زمان با شبکه عصبی قابل پیش‌بینی است که ممکن است با توجه به تغییر شرایط محیطی و اجتماعی مورد نیاز باشد.

۴- نتیجه‌گیری

در پژوهش حاضر توانایی دو روش داده‌محور شبکه عصبی پرسپترون و ماشین بردار پشتیبان در برآورد غلظت کلر باقیمانده در شبکه آبرسانی به‌منظور مدیریت و تحلیل سیستم گندزدایی ارزیابی شد. با توجه به نتایج به‌دست آمده در این پژوهش، استفاده از شبکه عصبی پرسپترون سه لایه با الگوریتم BFGS (به‌عنوان روشی داده‌محور با کارایی بالا) در پیش‌بینی میزان کلر باقیمانده در سیستم‌های گندزدایی شبکه‌های توزیع آب پیشنهاد می‌شود. این روش دقت بالایی دارد که از ملزومات سیستم گندزدایی است؛

References

- Adl, J., Mohammadfam, I. & Nezamoddini, Z., 2008, "Evaluation of the chlorine gas leakage risk of drinking water chlorination stations in Tehran using FTA method", *Journal of Jundishapur Scientific Medical*, 6(4), 461-468. (In Persian)
- Abraham, A., 2005, "Artificial neural networks", Sydenham, P.H. & Thorn, R. (Eds.) *Handbook of measuring system design*, John Wiley & Sons, Stillwater, OK, USA.
- Ammar, T.A., Abid, K.Y., El-Bindary, A.A. & El-Sonbati, A.Z., 2014, "Chlorine dioxide bulk decay prediction in desalinated drinking water", *Journal of Desalination*, 352, 45-51.
- Aslhashemi, A., 2012, "Comparison of residual chlorine in drinking water resources in urban and rural areas of Ilam Province", *Journal of Application of Chemistry in the Environment*, 3(12), 17-22. (In Persian)
- Bowden, G.J., Nixon, J.B., Dandy, G.C., Maier, H.R. & Holmes, M., 2006, "Forecasting chlorine residuals in a water distribution system using a general regression neural network", *Journal of Mathematical and Computer Modelling*, 44(5-6), 469-484.
- Dibike, Y.B., Velickov, S., Solomatine, D. & Abbott, M.B., 2001, "Model induction with support vector machines: introduction and applications", *Journal of Computing in Civil Engineering*, 15(3), 208-216.
- Gibbs, M.S., Morgan, N., Maier, H.R., Dandy, G.C., Nixon, J.B. & Holmes, M., 2006, "Investigation into the relationship between chlorine decay and water distribution parameters using data driven methods", *Journal of Mathematical and Computer Modelling*, 44(5-6), 485-498.

- Karadirek, I.E., Kara, S., Muhammetoglu, A., Muhammetoglu, H. & Soyupak, S., 2016, "Management of chlorine dosing rates in urban water distribution networks using online continuous monitoring and modeling", *Urban Water Journal*, 13(4), 345-359.
- Karadirek, I.E., Soyupak, S. & Muhammetoglu, H., 2016, "Chlorine modeling in water distribution networks using ARX and ARMAX model structures", *Desalination and Water Treatment*, 57(25), 11592-11598.
- Miranzadeh, M.B., Hasanzadeh, M., Dehqan, S. & Sabahi-Bidgoli, M., 2011, "The relationship between turbidity, residual chlorine concentration and microbial quality of drinking water in rural areas of Kashan during 2008-9", *Journal of Kashan University of Medical Sciences (Feyz)*, 15(2), 126-131. (In Persian)
- Nejjari, F., Puig, V., Perez, R., Quevedo, M.A., Cuguro, M.A., Sanz, G. & Mirats, J.M., 2014, "Chlorine decay model calibration and comparison: Application to a real water network", *Journal of Procedia Engineering*, 70, 1221-1230.
- Perju, S. & Stanescu, I., 2015, "Numeric modelling of the residual chlorine concentration evolution in a water distribution network", *Revista de Chimie*, 66(6), 886-890.
- Rodriguez, M.J. & Serodes, J.B., 1999, "Assessing empirical linear and non-linear modelling of residual chlorine in urban drinking water systems", *Journal of Environmental Modelling & Software*, 14, 93-102.
- Salehi-Artimani, J., Shamizadeh, H., Alinejad-Shahabi, R. & Arjmand, M., 2012, "Modeling and evaluation of risk of chlorine gas emission in water treatment plants", *Journal of Application of Chemistry in the Environment*, 3(9), 39-50. (In Persian)
- Smola, A.J. & Scholkopf, B., 2004, "A tutorial on support vector regression", *Journal of Statistics and Computing*, 14(3), 199-222.
- Tabesh, M., Azadi, B. & Rouzbahani, A., 2011, "Optimization of chlorine injection dosage in water distribution networks using a genetic algorithm", *Journal of Water & Wastewater*, 22 (1), 2-11. (In Persian)
- The Standard and Industrial Research Organization, 1997, *Physical and chemical characteristics of potable water*, Standard No. 1053, Tehran, Iran. (In Persian)
- Wu, W., Dandy, G.C. & Maier, H.R., 2015, "Optimal control of total chlorine and free ammonia levels in a water transmission pipeline using artificial neural networks and genetic algorithms", *Journal of Water Resources Planning and Management*, 141(7), 123-135.
- Yu, P.S., Chen, S.T. & Chang, I.F., 2006, "Support vector regression for real-time flood stage forecasting." *Journal of Hydrology*, 328, 704-716.