

Journal of Water and Wastewater, Vol. 33, No. 5, pp: 52-66

Investigation of Uncertainty to Artificial Intelligence Models in Tabriz Wastewater Treatment Plant

R. Shahidi Zonouz¹, V. Nourani², M. Dini³

1. PhD. Student in Water and Hydraulic Structures, Dept. of Water and Environment Engineering, Faculty of Civil Engineering, University of Tabriz, Tabriz, Iran
(Corresponding Author) reza800sh@yahoo.com
2. Prof., Dept. of Water and Environment Resources Engineering, Faculty of Civil Engineering, University of Tabriz, Tabriz, Iran
3. Assoc. Prof., Dept. of Civil Engineering, Faculty of Engineering, Azarbaijan Shahid Madani University, Tabriz, Iran

(Received June 10, 2022 Accepted Aug. 12, 2022)

To cite this article:

Shahidi Zonouz, R., Nourani, V., Dini, M. 2023. "Investigation of uncertainty to artificial intelligence models in Tabriz wastewater treatment plant" Journal of Water and Wastewater, 33(5), 52-66.
Doi: 10.22093/wwj.2022.346177.3264. (In Persian)

Abstract

In this paper, the uncertainty of artificial intelligence models for evaluating performance of the activated sludge unit of the Tabriz treatment plant is assessed. In this regard, daily data of pollution parameters, particularly Biochemical Oxygen Demand and Chemical Oxygen Demand, are utilized. All data were collected daily during the years (2015-2020) and the best parameters were selected using the correlation coefficient criterion. The TSSi, TDSi, VSSi, pHi parameters and also, BODe and CODE with a one-day delay were selected as model input and BODe and CODE were selected as model output. The calculations of uncertainties were performed in two models of Feed Forward Neural Network as point prediction and lower upper bound estimation method to provide the Prediction Interval. The LUBE method, unlike the classical methods of calculating PI, estimates PI without the need for data distribution information. In this method, the FFNN was trained with two outputs indicating the upper and lower limits of the prediction. PICP assessment and comparing it with μ values, caused γ values to equal zero that, in the continuation of the calculation process caused CWC extraction with the minimum possible amount and production of PI for computational data and observations with the possibility of controlling random changes in the activated sludge section. So, the convergence of the LUBE method has the ability to effectively control the uncertainty between the parameters of the biological section of activated sludge using PI. The time required to build PI is considerably short. Numerical results show approximately 99% success in calculations and coverage of modeling uncertainties. Providing an oscillating range of uncertainties can be a valuable aid in improving economic conditions as well as reducing activated sludge control time and better treatment plant monitoring. Despite the design criteria for BODe of 20 mg per liter, PI results show a supply of 12% of the design index. However, considering the supply of the remaining 88% in terms of quality standard for the use of effluents and returned water, according to the Deputy of Strategic



Supervision, publication 535, at the rate of 31 mg per liter in the activated sludge sector, the proper performance of the treatment plant is demonstrated. The LUBE method is an efficient method, so by providing an optimized range of fluctuations for computational data, the smallest abnormal changes in the activated sludge section due to controlling the amount of food for the micro-organisms present in this section; also, the pollution indicators with the least computing time are also reported. In addition, due to the high cost of activated sludge in the wastewater treatment sector, from an economic point of view, it also helps reduce costs. According to the non-linear behavior of bacteria during the reduction of food, as well as the control of mortality caused by the reduction of food, it can be considered a very effective tool.

Keywords: Tabriz Wastewater Treatment Plant, Activated Sludge Unit, Uncertainty, Prediction Interval, Upper and Lower Estimate.



مجله آب و فاضلاب، دوره ۳۳، شماره ۵، صفحه: ۵۲-۶۶

بررسی عدم قطعیت مدل‌های هوش مصنوعی در تصفیه‌خانه فاضلاب شهر تبریز

رضا شهیدی زنون^۱، وحید نورانی^۲، مهدی دینی^۳

۱- دانشجوی دکتری آب و سازه‌های هیدرولیکی، گروه عمران آب و محیط‌زیست،

دانشکده مهندسی عمران، دانشگاه تبریز، تبریز، ایران
(نویسنده مسئول) reza800sh@yahoo.com

۲- استاد، گروه عمران آب و محیط‌زیست، دانشکده مهندسی عمران،

دانشگاه تبریز، تبریز، ایران

۳- دانشیار، گروه مهندسی عمران، دانشکده فنی و مهندسی،

دانشگاه شهید مدنی آذربایجان، تبریز، ایران

(دریافت ۱۴۰۱/۳/۲۰ پذیرش ۱۴۰۱/۵/۲۱)

برای ارجاع به این مقاله به صورت زیر اقدام بفرمایید:

شهیدی زنون، ر.، نورانی، و.، دینی، م.، ۱۴۰۱، "بررسی عدم قطعیت مدل‌های هوش مصنوعی

در تصفیه‌خانه فاضلاب شهر تبریز" مجله آب و فاضلاب، ۳۳(۵)، ۵۲-۶۶. Doi: 10.22093/wwj.2022.346177.3264

چکیده

در این پژوهش، عدم قطعیت مرتبط با مدل‌سازی مبتنی بر هوش مصنوعی تصفیه‌خانه فاضلاب تبریز با داده‌های روزانه برای بررسی کارایی آن برای کنترل تغییرات مربوط به شاخص‌های BOD و COD واحد لجن فعال، بررسی شد. داده‌ها به صورت روزانه طی سال‌های ۲۰۱۵ تا ۲۰۲۰ گردآوری و با استفاده از معیار ضریب هم‌بستگی انتخاب شدند. TSS_i ، TDS_i ، VSS_i ، pH_i به همراه BOD_e و COD_e با تأخیر زمانی یک روزه به عنوان ورودی و BOD_e و COD_e برای خروجی مدل‌ها در نظر گرفته شدند. محاسبات در دو مدل شبکه عصبی پیش‌خور به صورت پیش‌بینی نقطه‌ای و روش برآورد حد بالا و پایین برای ارائه فاصله پیش‌بینی با اعمال عدم قطعیت‌ها انجام شد. روش LUBE، برخلاف روش‌های کلاسیک محاسبه PI، بدون نیاز به اطلاعات درباره توزیع داده‌ها به برآورد PI می‌پردازد. در این روش FFNN با دو خروجی که نشان‌دهنده حد بالا و پایین پیش‌بینی است، آموزش داده می‌شود. ارزیابی PICP و مقایسه آن با مقدار μ سبب تساوی مقدار γ با صفر شده که در ادامه روند محاسباتی موجب استخراج عرض هم‌گرایی با حداقل میزان ممکن و تولید PI برای داده‌های محاسباتی و مشاهداتی با امکان کنترل تغییرات تصادفی در بخش لجن فعال را میسر می‌سازد. بنابراین هم‌گرایی روش LUBE به‌طور مؤثر توانایی کنترل عدم قطعیت بین متغیرهای بخش بیولوژیکی لجن فعال با استفاده از PI را دارد. زمان لازم برای ساخت PI، به‌طور قابل ملاحظه‌ای اندک است. استفاده از روش LUBE منجر به ارائه یک بازه پیش‌بینی شامل عدم قطعیت‌های محاسباتی و مشاهداتی می‌شود. نتایج عددی نشان‌دهنده موفقیت تقریبی ۹۹ درصد در محاسبات و پوشش عدم قطعیت‌های مدل‌سازی است. ارائه بازه نوسانی از عدم قطعیت‌ها، کمکی شایان برای بهبود شرایط اقتصادی و همین‌طور کاهش زمان کنترل لجن فعال و رصد بهتر تصفیه‌خانه می‌تواند باشد. با وجود معیار طراحی برای BODE، به میزان ۲۰ میلی‌گرم در لیتر، نتایج PI نشان‌دهنده تأمین ۱۲ درصد شاخص طراحی بوده، ولی با توجه به تأمین ۸۸



درصد مابقی از لحاظ استاندارد کیفی برای کاربری پساب‌ها و آبهای برگشتی طبق نشریه ۵۳۵ معاونت نظارت راهبردی به میزان ۳۱ میلی‌گرم در لیتر، نشان‌دهنده عملکرد مناسب تصفیه‌خانه در بخش لجن فعال است. روش LUBE روشی کارآمد بوده، به طوری که با ارائه بازه‌ای بهینه شده از نوسانات برای داده‌های محاسباتی، کوچکترین تغییرات ناهنجار در بخش لجن فعال برای کنترل میزان غذای میکروارگانیسم‌های موجود در این بخش و همین‌طور شاخص‌های آلاینده‌گی را با کمترین زمان محاسباتی گزارش می‌کند. همچنین با توجه به گرانی خود لجن فعال در بخش تصفیه فاضلاب از منظر اقتصادی نیز کمکی شایسته در کاهش هزینه‌ها کرده و با توجه به رفتار غیرخطی باکتری‌ها در زمان کاهش غذا و همین‌طور کنترل مرگ‌ومیر ناشی از کاهش غذا می‌تواند ابزاری بسیار کارآمد تلقی شود.

واژه‌های کلیدی: تصفیه‌خانه فاضلاب تبریز، واحد لجن فعال، عدم قطعیت، فاصله پیش‌بینی، حد بالا و پایین



۱- مقدمه

با وجود استفاده وسیع از ANN در محاسبات مهندسی، بررسی موارد مشابه نشان‌دهنده استفاده از هوش مصنوعی^۳، برای پیش‌بینی نقطه‌ای خواهد بود که در این راستا اثرات عدم قطعیت‌ها، خطای مدل‌سازی و دقت پیش‌بینی در محاسبات لحاظ نشده است (Khan et al., 2006).

همچنین مدل‌سازی‌ها در AI، به‌طور معمول با روش سعی و خطا، به‌دنبال افزایش دقت برآورد و کاهش پیچیدگی ساختار شبکه‌های عصبی هستند. از سوی دیگر، عملکرد PP در AI با افزایش سطح پیچیدگی و افزایش میزان عدم قطعیت در مدل‌سازی مبتنی بر داده، بوده و به‌طور قابل‌توجهی کاهش می‌یابد. در این صورت اثرات خروجی مدل‌ها به صورت فاصله پیش‌بینی^۴، به جای PP از مدل‌های توسعه‌یافته شبکه عصبی، بسیار کاربردی و منطقی خواهد بود (Khosravi et al., 2010).

چندین روش متداول برای محاسبه PI در پژوهش‌های قبلی ارائه شده است، از جمله روش دلتا، بایاسین، تخمین میانگین و واریانس و روش بوت استرپ. در روش دلتا، فرض همگن بودن خطا و پیروی از تابع دلتا، محدودیتی برای روش دلتا است (Chryssoulouris et al., 1996). روش بایاسین بر اساس آمار بایاسین، برای بیان عدم قطعیت وزن‌های شبکه در ترم توزیع احتمال و ادغام آنها برای به‌دست آوردن توزیع احتمال خروجی ANN بوده و نیازمند حل انتگرال مونت‌کارلو بر روی فضای وزن‌ها است، بنابراین برای کاربرد چند بعدی مسائل دنیای واقعی نامناسب است (Mackay, 1992). در روش تخمین میانگین واریانس، وابستگی واریانس خروجی به مجموعه ورودی‌ها اساسی‌ترین فرض این مدل برای ساختن PI است (Nix and Weigend, 1994). روش بوت استرپ از جمله روش‌های رایج و ساده در محاسبات پیچیده مشتق و ماتریس همبستگی است (Dybowski and Roberts, 2001) بنابراین استفاده گسترده از روش بوت استرپ که در اکثر پژوهش‌ها یک ایراد اساسی برای نادیده گرفتن عدم قطعیت‌ها و فرضیات اولیه این روش با سایر مدل‌سازی‌ها بر آن وارد است.

با توجه به معایب ساخت PI در روش‌های مطرح شده نیاز به یک تابع توزیع احتمالاتی، ضروری به نظر می‌رسد. برای رفع این محدودیت اساسی، در این پژوهش از روش LUBE برای تولید PI

فاضلاب شهری به‌عنوان یکی از اساسی‌ترین منابع آلاینده‌گی در چرخه آب از نظر فیزیکی و شیمیایی است. با هدف افزایش بهداشت محیط‌زیست، هم‌زمان با پیشرفت‌های علمی در زمینه آب و فاضلاب برای ممانعت از آلودگی محیط و کاهش آلاینده‌گی متغیرهای شیمیایی و فیزیکی موجود در فاضلاب‌های شهری، به یکی از راهکارهای مناسب، استفاده از تصفیه‌خانه‌های فاضلاب اشاره کرد. اجرای تکنیک‌های نظارت و کنترل مؤثر برای سیستم‌های فاضلاب، موضوعی شناخته شده برای مهندسی آب و محیط‌زیست است. با این حال، مدل‌سازی یک تصفیه‌خانه فاضلاب^۱ به دلیل پیچیدگی فرایندهای تصفیه یک کار دشوار است (Hanbay et al., 2008).

واحدهای مختلف عملیاتی و فرایندی درگیر در مراحل تصفیه فاضلاب، رفتارهای غیرخطی از خود نشان می‌دهند که توصیف آنها توسط مدل‌های ریاضی خطی کلاسیک دشوار است (Wen and Vassiliadis, 1998). بنا بر این، پایش فرایند قابل‌اعتماد سیستم‌های تصفیه فاضلاب ممکن است با توسعه روش‌های غیرخطی قوی و استفاده از مدل‌های جعبه سیاه که قادر به پیش‌بینی عملکرد تصفیه‌خانه بر اساس مشاهدات گذشته متغیرهای کیفیت هستند، به‌دست آید. در این راستا شبکه‌های عصبی - مصنوعی به‌عنوان زیرشاخه‌ای از مدل‌های جعبه سیاه توانایی تحلیل و پردازش بر روی حجم انبوهی از داده‌ها را برای یافتن رابطه‌ای، برای مدل کردن فرایندهای پیچیده دارند که این کار را توسط آموزش دسته‌ای از داده‌های ورودی در برابر داده‌های خروجی انجام می‌دهند.

از دیگر ویژگی‌های شبکه‌های عصبی نسبت به دیگر روش‌ها، عدم توجه به خواص و روابط فیزیکی در مدل‌های برنامه‌ریزی غیرخطی است. این امر باعث گسترش روزافزون آن در شاخه‌های مختلف علوم مهندسی آب و محیط‌زیست شده است. پژوهش‌های متعددی در زمینه تحلیل متغیرهای آلاینده‌گی موجود در تصفیه‌خانه فاضلاب با استفاده از شبکه‌های عصبی انجام شده، اما تمامی این پژوهش‌های انجام شده به‌صورت پیش‌بینی نقطه‌ای^۲، هستند (Huggi and Mise, 2019, Nourani et al., 2018, Rastegaripour et al., 2019).

¹ Wastewater Treatment Plant (WWTP)

² Point Prediction (PP)

³ Artificial Intelligence (AI)

⁴ Prediction Interval (PI)



متغیرهای ورودی و خروجی انتخاب شدند. ابتدا مدل میانگین متحرک^۲، برای از بین بردن نوسانات شدید لحظه‌ای برای مدل‌سازی مدل‌سازی داده‌های تصفیه‌خانه اعمال شد (Svetunkov and Petropoulos, 2018) و در نهایت آنالیز ANN و LUBE با خروجی BOD و COD، محاسبه شد.

۲-۱- مطالعه موردی

طراحی و اجرای فاز یک تصفیه‌خانه تبریز بر پایه لجن فعال اجرا و با سیستم انتقالی ثقلی، در فاصله چهار کیلومتری غرب شهر تبریز در اراضی روستای قراملک، ضلع جنوبی رودخانه آجی‌چای و در پایین‌ترین نقطه شهر، قرار گرفته است. جانمایی مکانی تصفیه‌خانه شهر تبریز و بخش‌های تشکیل‌دهنده آن به‌طور شماتیک در شکل ۱، نشان داده شده است.

سیستم تصفیه‌خانه فاضلاب تبریز شامل بخش تصفیه فیزیکی، تصفیه بیولوژیکی و گندزدایی است. فرایند تصفیه بیولوژیکی فاضلاب از نوع لجن فعال با هوادهی دیفیوزری و تجهیزات هضم بی‌هوازی برای تصفیه و تثبیت لجن تولیدی است. متوسط دبی سالیانه فاضلاب و طراحی تصفیه‌خانه به ترتیب ۱/۱۵ و ۱/۵ مترمکعب در ثانیه است. حداکثر دبی لحظه‌ای در روزهای بارانی و غیربارانی به ترتیب ۳/۸ و ۲/۵ مترمکعب در ثانیه و غلظت BOD₅ و مواد معلق فاضلاب ورودی به ترتیب ۲۵۰ و ۳۵۰ میلی‌گرم در لیتر و بار آلودگی BOD₅ و مواد معلق فاضلاب ورودی به ترتیب ۳۲۴۰۰ و ۴۵۳۶۰ کیلوگرم در روز است. کیفیت پساب خروجی بر اساس معیارهای طراحی برای غلظت BOD₅ فاضلاب تصفیه شده کمتر از ۲۰ میلی‌گرم در لیتر لحاظ شده است. لجن فعال یکی از متداول‌ترین روش‌های تصفیه بیولوژیکی است. مزایای فرایند لجن فعال ارزان بودن، جذب سریع و کارایی بالا، ظرفیت زیاد جذب، تصفیه مؤثر فاضلاب‌های دارای غلظت مواد آلی، قابلیت احیای جاذب، استفاده مجدد از آن و بازیافت فلزات جذب شده، باعث مطلوب بودن این سیستم برای جذب فلزات سنگین شده است.

عدم دسترسی به داده‌هایی در بازه‌های زمانی کمتر

استفاده شد. پژوهش‌های با استفاده از روش LUBE، در زمینه مهندسی آب مطرح و به‌کار گرفته شده است (Nourani et al., 2019, Nourani et al., 2021, Nourani et al., 2020) نشان داد برای موضوع آب و فاضلاب، مدل‌سازی مبتنی بر PI با متغیرهای خروجی تصفیه‌خانه فاضلاب در بخش لجن فعال تاکنون پژوهشی انجام نشده است. تنها یک مورد در زمینه اندازه‌گیری شاخص‌های BOD و نیتروژن آمونیاکی برای کیفیت آب خروجی تصفیه‌خانه توسط ژو و همکاران که از روش LUBE برای مدل‌سازی PI و برای آنالیز PP با استفاده از الگوریتم LSVM^۱ و بهره‌گیری از آن، پژوهشی را انجام دادند (Zhou et al, 2022).

از جمله نقاط ضعف پژوهش‌های گذشته می‌توان به صرف‌نظر کردن از عدم قطعیت‌های موجود در فیزیک مسئله اشاره کرد، به‌طوری که انجام محاسبات PP صرفاً باعث ارائه گزارش نهایی در مورد کمترین مربعات خطاهای محاسباتی، ضریب تبیین و یا سایر معیارهای ارزیابی مربوطه می‌شود. بنابراین در راستای حذف این روند معمول، ارائه نوسانات در مقدار عدم قطعیت‌ها به‌صورت بازه‌ای از مقدار کمینه و بیشینه می‌تواند کمک زیادی در رصد اطلاعات و نحوه عملکرد هر پدیده‌ای باشد که به‌طور خاص در این پژوهش در مورد تصفیه‌خانه فاضلاب و تغییرات در متغیرهای واحد لجن فعال پرداخته شد.

۲- مواد و روش‌ها

تصفیه‌خانه تبریز از سیستم لجن فعال برای تصفیه فاضلاب شهری استفاده می‌کند. در این پژوهش داده‌های روزانه طی سال‌های ۲۰۱۵ تا ۲۰۲۰ استفاده شد. داده‌های استفاده شده در مراحل محاسباتی، ترکیبی از SS، TSS، COD و BOD₅ که در بخش‌های ورودی، خروجی ته‌نشینی اولیه و ثانویه و خروجی تصفیه‌خانه (بعد از کلرزدنی) که به‌صورت روزانه اندازه‌گیری و ثبت شدند. شاخص‌های TDS، EC، TP، نترات، آمونیاک، pH، دما، DO، TS و VS در بخش‌های لجن خام، لجن غلیظ شده اولیه و لجن برگشتی اندازه‌گیری شدند. ارزیابی بین کل داده‌ها با استفاده از معیار CC، انجام شد و بیشترین ارتباط در مقدار ضریب هم‌بستگی به‌عنوان

² Moving Average (MA)

¹ Least Squares Vector Machine (LSVM)



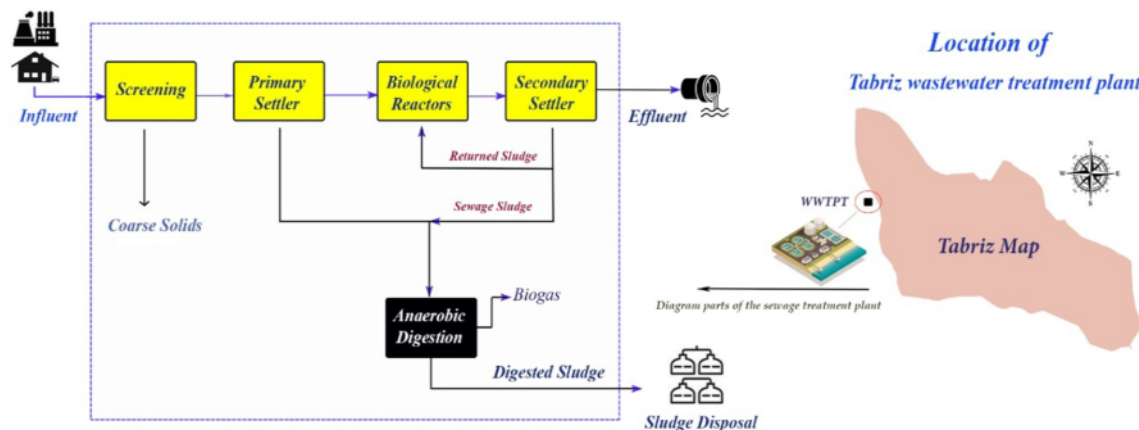


Fig. 1. Location and schematic diagram of Tabriz wastewater treatment plant

شکل ۱- موقعیت مکانی و شماتیکی از دیگرام عملکرد تصفیه‌خانه فاضلاب شهر تبریز

متغیرهای BOD و COD در تصفیه‌خانه تبریز نسبت به سایر متغیرهای دیگر منظم و دقیق‌تر بوده، بنابراین داده‌های تحلیلی ترکیبی از BOD یا COD و ضریب هم‌بستگی بالا سایر شاخص‌های اندازه‌گیری شده با این دو متغیر در مدل‌سازی لحاظ شد. مقدار داده‌ها در چهار گام اساسی ورودی به تصفیه‌خانه، حوضچه ته‌نشینی اولیه، حوضچه ته‌نشینی ثانویه، خروجی از تصفیه‌خانه و همین‌طور خروجی از هر گام نسبت به سایر گام‌ها

(به‌عنوان مثال داده‌های ساعتی یا ۸ ساعته) و بررسی مدل توسعه‌یافته روی عدم قطعیت‌ها، از محدودیت‌های این پژوهش بود. زمان ماند هیدرولیکی با تأخیر زمانی یک روزه در گزارشات لحاظ شد. مقدار کلی داده‌های کمی، در جدول ۱ بیان شده است. BOD و COD مشخصه‌ای برای ارزیابی اثر فاضلاب تخلیه‌شونده بر محیط پذیرنده است، پس مهمترین متغیرها در کنترل کیفیت پساب و توابع هدف آنالیزها محسوب می‌شوند. آمار

جدول ۱- دامنه نوسان متغیرهای آلاینده‌های تصفیه‌خانه تبریز

Table 1. Fluctuation range of pollution variables of Tabriz treatment plant

Variable	Unit	Maximum	Minimum	Average	Std
pH _i	°C	8.61	7.29	8.05	0.19
BOD _i	mg/L	315	220	290.12	12.91
COD _i	mg/L	516.67	362.17	479.16	22.01
SS _i	mg/L	1.5	0	0.61	0.151
TSS _i	mg/L	336.33	234	296.40	14.64
VSS _i	mg/L	235.33	140.5	191.13	15.04
TDS _i	ppm	1366.83	779.83	936.34	64.35
T _i	°C	26.5	12.2	21.04	2.97
EC _i	μSimens/cm	1934	1297	1468.65	91.65
BOD _e	mg/L	26.33	17.33	22.72	1.28

i mean's samples in the inlet section and e mean's effluent data to the activated sludge section in WWTP



جدول ۲- مقایسه ضریب هم‌بستگی بین متغیرهای تصفیه‌خانه با BOD و COD

Table 2. Comparison of the CC between the parameters of the treatment plant with BOD and COD

Input Output	BOD _i	COD _i	TSS _i	BOD (t-1)	COD (t-1)	pH _i	SS _i	VSS _i	TDS _i	T _i
BOD _e	0.429	0.384	0.279	0.945	0.872	0.258	0.039	0.174	0.258	-0.35
COD _e	0.43	0.479	0.332	0.812	0.746	0.279	-0.037	0.226	0.192	-0.27

x_i ورودی مدل‌سازی، t_i داده‌های مشاهده‌ای و $f(x_i, w^*)$ تابع غیرخطی مدل متناظر و w^* متغیرهای واقعی $f(0)$ و e_i خطا با میانگین صفر، $f(x_i, \hat{w})$ میانگین توزیع داده‌های تخمینی است (Heskes, 1996, Khosravi et al., 2010).

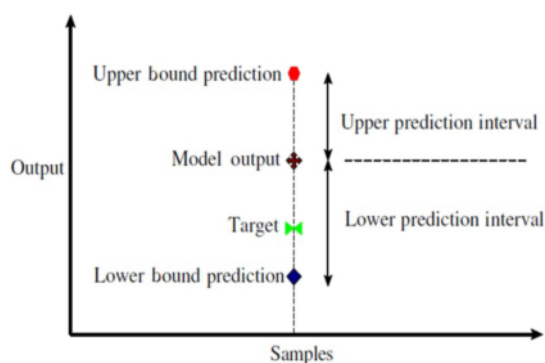


Fig. 2. PI concept

شکل ۲- مفهوم PI

دو معیار ارزیابی کمی PI احتمال هم‌گرایی PICP و MPIW است (Nourani et al., 2021)

$$PICP = \frac{1}{n} \sum_{i=1}^n c_i \quad ; \quad c_i = \begin{cases} 1 & ; t_i \in [L(X_i), U(X_i)] \\ 0 & ; \text{otherwise} \end{cases} \quad (7)$$

که در آن

$L(X_i)$ مرز پایین، $U(X_i)$ مرز بالای حدود پیش‌بینی متناظر برای i امین PI (نمونه)، n تعداد نمونه‌های محاسباتی و MPIW متغیر بهینه‌سازی PICP است

تنظیم مقدار در یک مقیاس، عملیات نرمال‌سازی بر روی داده‌ها محاسبه شد. مقدار CC بین متغیرهای ورودی و خروجی مختلف محاسبه و به‌عنوان بهینه‌ترین حالت در جدول ۲ نتایج ارائه شده است. CC بین متغیرهای ورودی شامل BOD_i، COD_i، TSS_i، هدایت الکتریکی (EC_i)، کل جامدات محلول (TDS_i)، SS_i، pH_i و متغیرهای پساب شامل BOD_i(t-1)، COD_i(t-1)، دما (T_i) و متغیرهای پساب شامل BOD_e، pH_e، VSS_e، TSS_e، TDS_e در جدول ۲، ارائه شده است. بیشترین مقدار CC به‌عنوان ورودی و BOD_e و COD_e به‌عنوان خروجی مدل‌ها انتخاب شدند. بررسی ضریب هم‌بستگی BOD_e بعد از اعمال MA، نشان داد BOD_e(t-1)، هم‌بستگی بالایی با BOD_e(t) و رقمی معادل ۰/۸۷۵۲ است. پس می‌توان فرض کرد که سیستم لجن فعال ماهیت خود جهت دارد. آنالیز شبکه عصبی نشان داد BOD_i(t) و BOD_i(t-1) تأثیر مستقیمی بر کیفیت پساب و داده‌های خروجی دارد. مدل‌سازی شبکه عصبی با استفاده از مدل‌های هوش مصنوعی، برای ۷۵ درصد اولیه داده‌ها به‌عنوان داده‌های آموزشی و ۲۵ درصد بقیه به‌عنوان مجموعه داده‌های صحت‌سنجی استفاده شد.

۲-۳-PI و معیارهای ارزیابی آن

PI دقت پیش‌بینی مقدار محاسباتی، در برابر داده‌های مشاهده‌ای را بیان می‌کند (شکل ۲) که مطابق با معادله ۶ محاسبه می‌شود

$$PI = t_i - f(x_i, \hat{w}) = \left[f(x_i, w^*) - f(x_i, \hat{w}) \right] + e_i \quad (6)$$

که در آن



حالت انجام شده و آخرین مجموعه وزن به عنوان وزن های بهینه شده برای ساخت PI گزارش شده است.

پیش بینی های نقطه ای توسط FFNN هیچ اطلاعاتی در مورد عدم قطعیت ارائه نمی دهند. PI یک شاخص اساسی برای کمی سازی سطح اطمینان در پژوهش متغیرهای آلاینده های پساب مبتنی بر NNFF با استفاده از داده های واقعی WWTP است. روش LUBE روشی جامع برای تأمین کاهش تأثیر عدم قطعیت ها با ساختاری سریع در بخش محاسباتی است. این روش می تواند ابزاری اساسی در رصد عملکرد بخش لجن فعال و کنترل میکروارگانیسم های موجود در این قسمت از تصفیه خانه باشد.

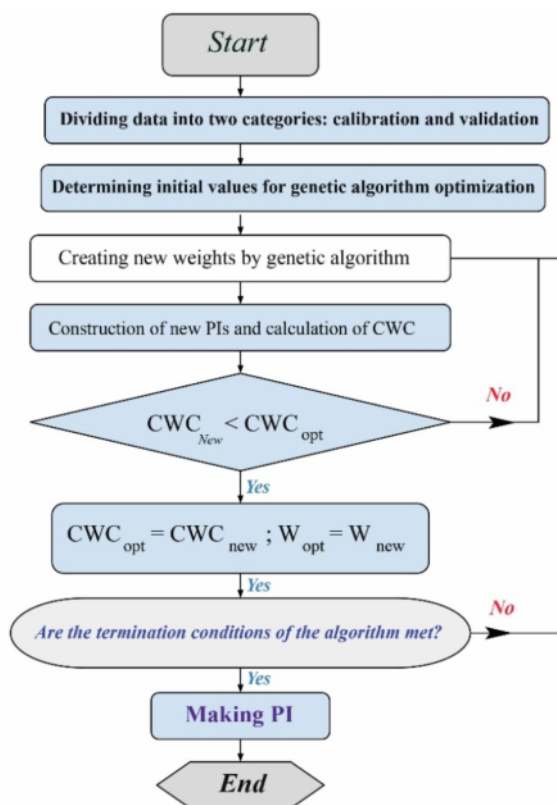


Fig. 3. Flowchart of PI manufacturing steps with LUBE method

شکل ۳- الگوریتم مراحل ساخت PI با روش LUBE

۳- نتایج و بحث

این پژوهش با استفاده از برنامه نویسی در محیط MATLAB تهیه شد. نتایج حاصل از آنالیز داده های مشاهداتی ابتدا از طریق PP، بررسی و برای مقایسه نتایج پیش بینی مدل های FFNN از

$$MPIW = \frac{1}{n} \sum_{i=1}^n [U(x_i) - L(x_i)] \quad (8)$$

میانگین عرض فواصل پیش بینی نرمال شده NMPIW^۱، معیاری بدون بعد برای ارائه عرض متوسط به صورت درصد دامنه توابع هدف برای PI و CWC معیاری برای ارزیابی PICP است

$$NMPIW = MPIW/R \quad (9)$$

$$CWC = NMPIW(1 + \gamma(PICP)e^{-\eta(PICP-\mu)}) \quad (10)$$

$$\gamma = \begin{cases} 0 & ; PICP \geq \mu \\ 1 & ; PICP < \mu \end{cases} \quad (11)$$

که در آن

R به عنوان دامنه توابع اصلی هدف یا به عبارت بهتر دامنه ای برای داده ها، γ و η تعداد تکرارها برای PI بهینه با احتمال پوشش کم و μ متغیر مربوط به سطح اطمینان ظاهری و مشاهداتی مرتبط با PI هستند. برای ایجاد PI می توان از روش LUBE استفاده کرد. روش LUBE، ساختاری بر پایه ANN است و برخلاف PP، دو خروجی برای تخمین حد بالا و پایین PI برای داده های WWTP با استفاده از آن گزارش می کند. اولین خروجی متناظر با حد بالا و دومی نشان دهنده حد پایین برای داده های محاسباتی است. گام بندی و الگوریتم روش LUBE در شکل ۳، نشان داده شده است (Nourani et al., 2021).

در گام اول داده ها در دو دسته کالیبره و صحت سنجی تقسیم بندی می شوند. در گام دوم مقداردهی اولیه بهینه سازی جمعیت اولیه به صورت تصادفی بوده به طوری که مقدار مشاهداتی با اضافه و کم کردن عددی تصادفی به عنوان حد بالا و پایین اولیه در نظر گرفته می شود. برای گام سوم مدل های اولیه FFNN با استفاده از الگوریتم لونیبرگ مارکوارت آموزش داده می شود، سپس در گام چهارم PI خروجی و مقدار PICP، NMPIW و CWC اولیه محاسبه می شود. در گام پنجم، الگوریتم ژنتیک با ایجاد مجموعه ای از وزن های مختلف برای PI و CWC ها را محاسبه کرده که در صورت عدم ارضای معیارهای ارزیابی از گام سوم محاسبات تکرار خواهد شد. در نهایت پس از بررسی شرایط، خاتمه الگوریتم در بهینه ترین

¹ Normal Mean Prediction Interval Width (NMPIW)



متناسب است. BOD و COD از بسیاری جهات مشابهت داشته و میزان اکسیژن خواهی زیستی و شیمیایی را مشخص می‌کنند و در صورت فقدان مواد آلی غیر قابل تجزیه بیولوژیک و همگن بودن فاضلاب مقدارشان یکسان خواهند بود. پس بررسی هر یک به تنهایی در هر مدل مکمل مدل‌های دیگر خواهد بود.

TDS و TSS بیانگر ترکیبی از جامدات محلول و معلق به صورت هم‌زمان از مواد معدنی و آلی بوده، پس بررسی روی این متغیرها به نوعی رصد آلودگی موجود در تصفیه‌خانه خواهد بود. همچنین بررسی TSS و VSS به‌خودی‌خود مشخصات جامدات معدنی معلق^۱ را مشخص می‌کند.

بررسی pH بر میزان حذف فسفر تأثیر مستقیم دارد. کم بودن pH نشان‌دهنده حجم زیادی از مواد معدنی و فلزات سنگین و باعث ایجاد رسوب در لوله‌های انتقال آب می‌شود که اهمیت مسئله از منظر هیدرولیکی، مشخص می‌شود.

نتایج خطا با پژوهش‌های انجام شده توسط (Guo et al., 2015) قابل توجه است. برگردان مقدار نرمال شده در قالب نمودار شکل ۴ مشاهده می‌شود.

نوع آموزش یافته شبکه عصبی برحسب داده‌های مشاهداتی و محاسباتی برای BOD_e و COD_e نشان‌دهنده آموزش بهینه و دقیق شبکه عصبی است.

۳-۲- نتایج حاصل از PI

عملکرد کلی برای ساخت PI مبتنی بر مدل‌سازی ANN بوده و وابسته به ساختار شبکه عصبی و روند آموزش خواهد بود. لازمه نتایج PI، مدل‌سازی با ANN برحسب داده‌های تصفیه‌خانه برای BOD_e و COD_e است. نتایج با سطح اطمینان مرتبط بین ۷۵ تا ۹۰ درصد بررسی شده و بهینه‌ترین حالت برای هر تابع هدف انتخاب شد. ساختار بهینه ANN با بیشترین PICP و کمترین NMPIW برای هر مجموعه از وزن‌های تولید شده (با تغییر وزن‌های اولیه) محاسبه شد. بهینه‌ترین حالت کمینه مقدار غلظت آلاینده‌گی پسماند فاضلاب برای BOD_e و COD_e، در ساخت PI در جدول ۴ گزارش شده است.

نتایج نشان‌دهنده تفاوت در سطح اطمینان مقدار مشاهداتی BOD_e نسبت به COD_e است. مقدار η نیز از این امر پیروی کرده و

معیارهای DC و RMSE استفاده شد. در گام بعدی با استفاده از روش LUBE، محاسبه PI برای متغیرهای خروجی انجام شد.

۳-۱- نتایج حاصل از PP

سری زمانی داده‌های تصفیه‌خانه تشکیل و مدل بهینه برای متغیرهای ورودی و خروجی انتخاب شد. فاز تأخیری اعمالی برای BOD و COD، به لحاظ متجانس بودن ورودی و خروجی در نتایج تأثیر مستقیم داشته اما از لحاظ محاسباتی، به خاطر تابع فعال‌سازی FFNN در بازه (۱-۱) داده‌ها نرمال شدند. به‌منظور کالیبره و صحت‌سنجی مدل داده‌های ۵ ساله روزانه دسته‌بندی شد. ۷۵ درصد نمونه‌ها از ابتدایی‌ترین داده‌ها و داده‌های نهایی در مرحله آموزش گنجانده شد. مقدار بهینه نرون‌ها با استفاده از روش آزمون و خطا تعیین شدند که در محدوده (۱۰-۱) نرون، آزمایش شد. داده‌های تصفیه‌خانه در دو قسمت اساسی بیولوژیکی و شیمیایی قابل تقسیم‌بندی بود، ولی به‌علت بررسی بخش لجن فعال از داده‌های شیمیایی صرف‌نظر کرده و تنها به داده‌های بیولوژیکی اکتفا شده است.

برای متغیرهای خروجی BOD_i و COD_i، با توجه به اختلاف ورودی داده‌ها و ارتباط معنی‌دار با یکدیگر در دو کلاس کاری برای BOD و COD به‌صورت مجزا در جدول ۳ دسته‌بندی و به کار گرفته شد. ساختار بهینه FFNN برای تابع هدف BOD و COD در قسمت خروجی تصفیه‌خانه، به ترتیب از ۴ و ۳ نرون میانی تشکیل شده است. وزن‌های شبکه عصبی برای کمترین خطا تنظیم و در آنالیزها انتخاب شدند.

مقدار کم RMSE و زیاد DC در مرحله صحت‌سنجی نشان‌دهنده دقت بالای محاسبات است. برای جلوگیری از حفظ کردن شبکه، تعداد تکرار برای بهینه‌ترین حالت، محاسبه شد (جدول ۳). به‌علت وجود گیج تثبیت‌کننده جریان، دبی ورودی به تصفیه‌خانه ثابت بوده و به‌طور معمول بررسی شرایط بحرانی و سیلابی ناممکن است. علاوه بر معیار CC، دلایل زیر برای انتخاب متغیرهای آنالیزی را می‌توان نام برد:

BOD_i بین ۱۵۰ تا ۳۰۰ میلی‌گرم در لیتر در شرایط بحرانی، نوسان دارد و میزان آلاینده‌گی فاضلاب شهر تبریز از جمله گروه پساب‌های متوسط شناخته می‌شود. از طرفی میزان اکسیژن خواهی باکتری‌های لجن فعال با میزان مواد آلی موجود در فاضلاب

¹ Fixed Suspended Solids (FSS)



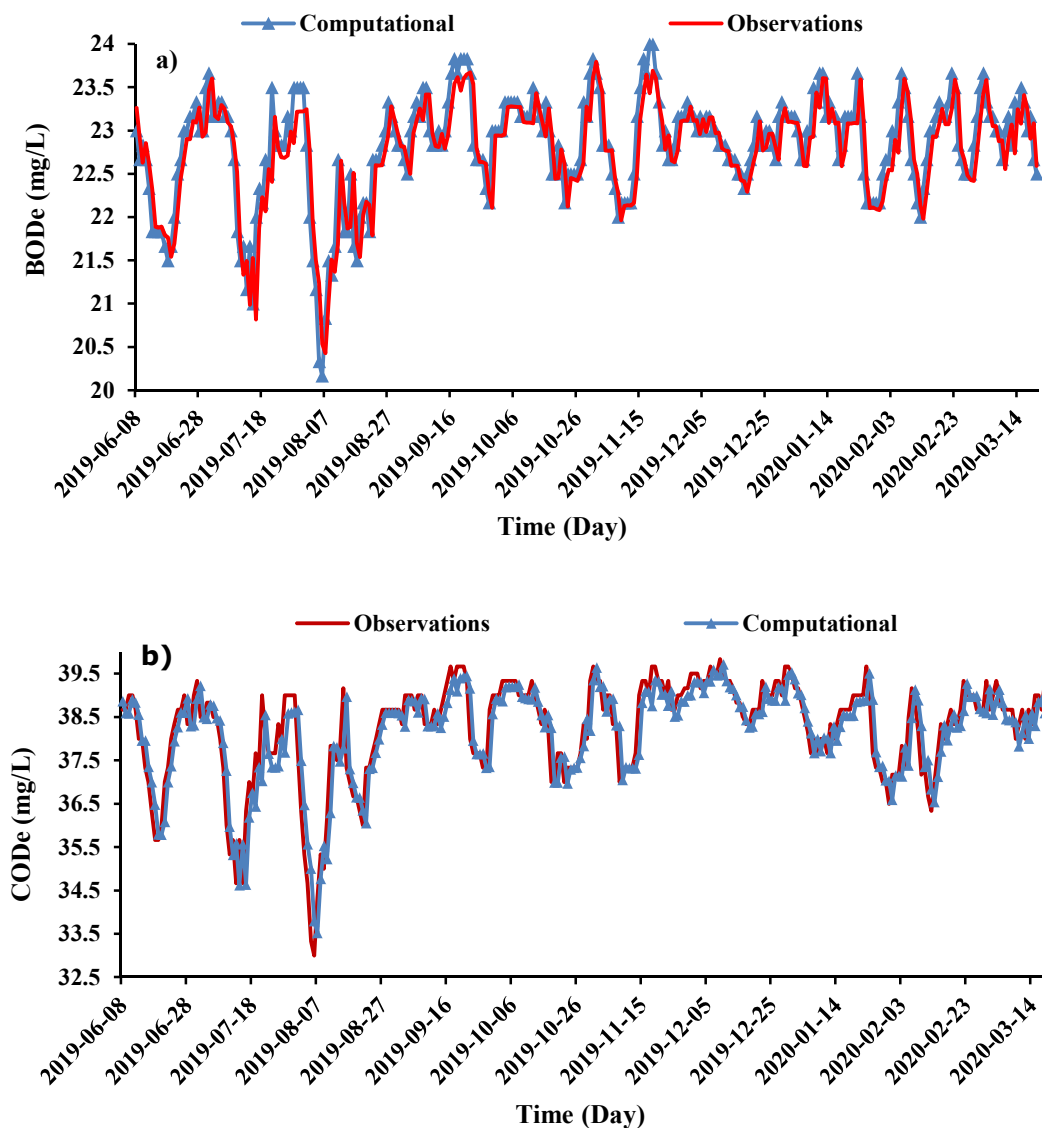


Fig. 4. FFNN computational time series for a) BOD_e and b) COD_e.
 شکل ۴- سری زمانی محاسباتی FFNN برای a) BOD_e و b) COD_e

جدول ۳- متغیرهای ورودی و خروجی و نتایج مدل‌سازی بهینه روش ANN

Table 3. Input and output parameters and best results for ANN method

Model	Output	Inputs				Train		Verify		Epoch	Intermediate neurons	
						RMSE	DC	RMSE	DC			
1	BOD _e (t)	TSS ₁ (t)	TDS ₁ (t)	BOD _e (t-1)	0.091	30.90	0.071	90.72	40	4		
2	COD _e (t)	pH ₁ (t)	TSS ₁ (t)	VSS ₁ (t)	TDS ₁ (t)	COD _e (t-1)	0.122	10.89	0.099	10.78	20	3



جدول ۴- نتایج مدل‌سازی روش LUBE
Table 4. Modeling results of the LUBE method

Model	Output	Epoch	Hidden layer	PICP	μ	η	NMPIW	CWC	Ini Temp ²	Iter SA ¹
1	BOD _e	40	4	0.9095	0.75	50	0.1728	0.1728	10000	300
2	COD _e	40	3	0.8780	0.90	150	0.1945	5.4289	10000	300

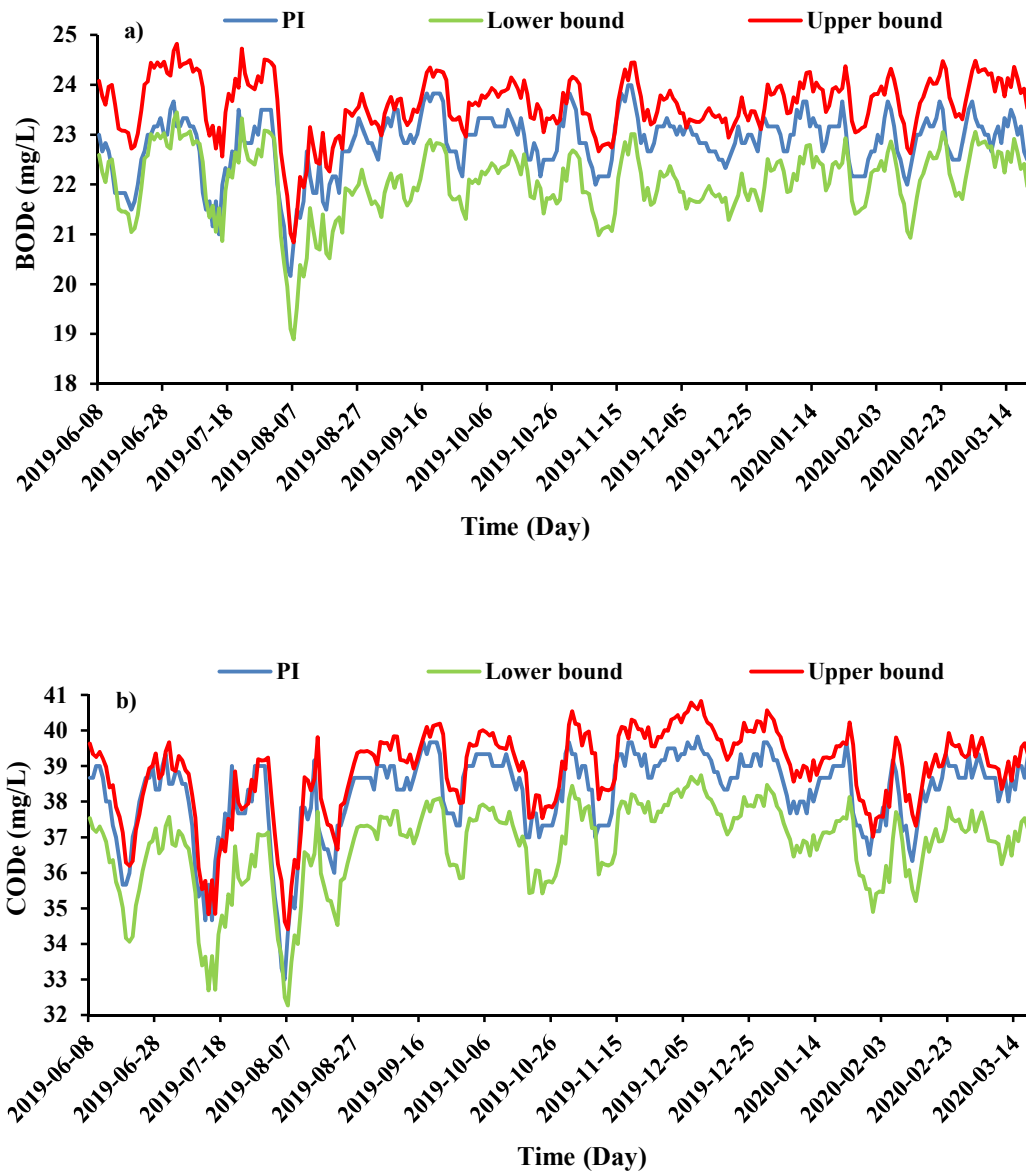


Fig. 5. Modeling results with LUBE method for a) BOD and b) COD_e
 شکل ۵- نتایج مدل‌سازی با روش LUBE برای (a) BOD_e و (b) COD_e

¹ Iteration Simulated Annealing (Iter SA)

² Initial Temperature (Ini Temp)



روش FFNN از روی مقدار RMES و DC برای متغیرهای ورودی و خروجی بر پایه ضریب هم‌بستگی، نشان‌دهنده دقت زیاد محاسباتی است. شبکه عصبی با تابع هدف BOD_e با وجود سه ورودی نسبت به مدل COD_e با پنج ورودی، در هر دو روش محاسباتی عملکرد بهتری را از خود نشان داد. DC و RMSE برای داده‌های آموزش به ترتیب ۰/۹ و ۰/۰۹ و همچنین DC و RMSE صحت‌سنجی به ترتیب ۰/۷ و ۰/۰۷ در بهترین حالت نتیجه گرفته شد. مقایسه نتایج مدل‌های روش LUBE، رقم ۰/۹ برای PICP و ۰/۱ برای CWC نشان‌دهنده برتری مدل ۱ نسبت به مدل ۲ با PICP ۰/۸ و CWC ۵/۴ داشت. مقدار عددی PICP و CWC به‌عنوان متغیرهای ارزیابی عملکرد روش LUBE، نشان‌دهنده هم‌پوشانی ۹۹ درصدی PI نسبت به مقدار مشاهداتی است که موفقیتی چشمگیر برای پیش‌بینی عملکرد پدیده غیرخطی لجن فعال است.

این پژوهش بر توسعه یک الگوریتم بهینه‌سازی تصادفی متمرکز بود که شامل یک روش برای ارائه محدودیت‌ها برای در نظرگرفتن عدم قطعیت‌های مدل کیفیت فاضلاب و پساب شهری بود. بر اساس ترکیبی از الگوریتم بهینه‌سازی و مدل شبیه‌سازی کیفیت پساب، طرح پیشنهادی برای ایجاد یک طرح تخصیص بار پسماند برای مدیریت کارآمد کیفیت فاضلاب قابل‌استفاده است. پیشنهاد می‌شود در راستای پایش بهینه و افزایش بازده تصفیه‌خانه فاضلاب، از این روش به‌عنوان روشی برای کنترل طراحی و مکملی برای محاسبات استفاده شود.

۵- قدردانی

نویسندگان، با توجه به تأمین داده‌های مربوطه و همکاری صمیمانه کارکنان تصفیه‌خانه فاضلاب تبریز و آقای مهندس عباسی معاونت محترم شرکت آب و فاضلاب استان آذربایجان شرقی کمال تشکر را دارند.

تفاوت فاحشی را نسبت به یکدیگر نشان می‌دهد. وجود این تضاد، به علت تاثیر تکرارهای متعدد برای بیشترین هم‌پوشانی با کمترین CWC و NMPIW با هم‌گرایی بالا در مقدار PICP به‌دست آمده است (شکل ۵).

به‌دلیل اینکه PICP و NMPIW رابطه عکس دارند، در نظریه دو معیار ارزیابی نشان داد که نتایج مدل‌سازی برای BOD_e در مقایسه با COD_e عملکرد بهتری را نشان می‌دهد. در مورد ساخت PI با روش LUBE، PI‌ها به‌طور مستقیم از خروجی محاسبات FFNN بدون هیچ فرض اولیه‌ای در مورد توزیع داده‌ها حاصل می‌شوند. PI نشان‌دهنده دقتی است که به‌وسیله آن حد بالا و پایین تولید شده قادر به پوشش مقدار واقعی است. در ضمن با توجه به اینکه روش انتخابی، به‌جای اینکه فقط حاکی از دقت متغیر تخمین رگرسیون صحیح متغیرها باشد، محدوده‌ای را بیان می‌کند که مؤثرتر است. از مزایای چشمگیر روش LUBE زمان کم محاسبات است. برای روش، یک بار آموزش FFNN کافی است.

۴- نتیجه‌گیری

هدف اصلی این پژوهش، بررسی کاربرد مدل FFNN و LUBE در پیش‌بینی BOD_e و COD_e بخش بیولوژیکی تصفیه‌خانه، با لحاظ عدم قطعیت‌ها بود. برای مدل‌سازی عملکرد تصفیه‌خانه فاضلاب، در فاز اول مدل‌سازی، سری زمانی داده‌های اصلی تشکیل، سپس محاسبات پیش‌پردازش (نرمال‌سازی) برای تولید مدل عصبی انجام شد. کیفیت و کمیت داده‌های مشاهده شده به‌شدت بر عملکرد مدل‌های توسعه‌یافته تأثیر گذاشت، بنابراین انتظار می‌رفت با مجموعه داده‌های بزرگتر نتایج بهتری به‌دست آورد. PI‌ها بعد از آموزش با دو نوع خروجی FFNN و متجانس BOD و COD، از روش LUBE با دو خروجی به‌عنوان مرزهای بالایی و پایینی پیش‌بینی ساخته شدند. برای ارزیابی کمی عملکرد PI‌های ساخته شده، از معیارهای NMPIW و PICP استفاده شد. نتایج عددی

References

- Baghanam, A. H., Nourani, V., Sheikhabaei, A. & Seifi, A. J. 2020. Statistical downscaling and projection of future temperature change for Tabriz city, Iran. *IOP Conference Series: Earth and Environmental Science*, IOP Publishing, 012009. Kerala, India.
- Chrysolouris, G., Lee, M. & Ramsey, A. 1996. Confidence interval prediction for neural network models. *IEEE Transactions on Neural Networks*, 7, 229-232.



- Dybowski, R. & Roberts, S. J. 2001. *Confidence Intervals and Prediction Intervals for Feed-forward Neural Networks*. Cambridge University Press. UK.
- Guo, H., Jeong, K., Lim, J., Jo, J., Kim, Y. M., Park, J. P., et al. 2015. Prediction of effluent concentration in a wastewater treatment plant using machine learning models. *Journal of Environmental Sciences*, 32, 90-101.
- Hanbay, D., Turkoglu, I. & Demir, Y. 2008. Prediction of wastewater treatment plant performance based on wavelet packet decomposition and neural networks. *Expert Systems with Applications*, 34, 1038-1043.
- Heskes, T. 1996. Practical confidence and prediction intervals. *Advances in Neural Information Processing Systems*, 9, 176-182.
- Huggi, M. & Mise, S. 2019. Optimized ANN model for ultrasonication wastewater treatment process. *International Journal of Advanced Research in Engineering and Technology*, 10(3), 94-102.
- Khan, M. S., Coulibaly, P. & Dibike, Y. 2006. Uncertainty analysis of statistical downscaling methods. *Journal of Hydrology*, 319, 357-382.
- Khosravi, A., Nahavandi, S. & Creighton, D. 2010. A prediction interval-based approach to determine optimal structures of neural network metamodels. *Expert Systems with Applications*, 37, 2377-2387.
- Mackay, D. J. 1992. A practical Bayesian framework for backpropagation networks. *Neural Computation*, 4, 448-472.
- Nix, D. A. & Weigend, A. S. 1994. Estimating the mean and variance of the target probability distribution. *Proceedings of 1994 IEEE International Conference on Neural Networks (ICNN'94)*, IEEE, 55-60. Orlando, USA.
- Nourani, V., Elkiran, G. & Abba, S. 2018. Wastewater treatment plant performance analysis using artificial intelligence—an ensemble approach. *Water Science and Technology*, 78, 2064-2076.
- Nourani, V., Paknezhad, N. J., Sharghi, E. & Khosravi, A. 2019. Estimation of prediction interval in ANN-based multi-GCMs downscaling of hydro-climatologic parameters. *Journal of Hydrology*, 579, 124226.
- Nourani, V., Paknezhad, N. J. & Tanaka, H. 2021. Prediction interval estimation methods for artificial neural network (ANN)-based modeling of the hydro-climatic processes, a review. *Sustainability*, 13(4), 1633.
- Nourani, V., Sayyah-Fard, M., Alami, M. T. & Sharghi, E. 2020. Data pre-processing effect on ANN-based prediction intervals construction of the evaporation process at different climate regions in Iran. *Journal of Hydrology*, 588, 125078.
- Rastegaripour, F., Saboni, M., Shojaei, S. & Tavassoli, A. 2019. Simultaneous management of water and wastewater using ant and artificial neural network (ANN) algorithms. *International Journal of Environmental Science and Technology*, 16, 5835-5856.
- Svetunkov, I. & Petropoulos, F. 2018. Old dog, new tricks: a modelling view of simple moving averages. *International Journal of Production Research*, 56, 6034-6047.
- Wen, C. H. & Vassiliadis, C. 1998. Applying hybrid artificial intelligence techniques in wastewater treatment. *Engineering Applications of Artificial Intelligence*, 11, 685-705.
- Zhou, M., Zhang, Y., Wang, J., Shi, Y. & Puig, V. 2022. Water quality indicator interval prediction in wastewater treatment process based on the improved BES-LSSVM algorithm. *Sensors*, 22, 422.



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

