طراحی ساختارهای ANFIS و شبکه‌های عصبی GMDH برای پیش‌بینی میزان بهینه مصرف ماده منعقدکننده در فرایند تصفیه آب، مطالعه موردی: تصفیه‌خانه بزرگ آب گیلان

الهیار داغبندان[1] محمد اکبری‌زاده[2]

(دریافت ۹۱/۹/۲۳ پذیرش ۹۲/۳/٤)

چکیده

در این مطالعه با توجه به اهمیت منابع سطحی در تأمین آب شرب و لزوم استفاده از مواد شیمیایی گوناگون در مراحل مختلف تصفیه این گونه آب‌ها، میزان مصرف مواد منعقد کننده در فرایند تصفیه آب مورد بررسی قرار گرفت. یکی از مهم‌ترین قسمت‌های فرایند تصفیه آب، مربوط به میزان مصرف مواد منعقدکننده در واحد انعقاد و لخته‌سازی است. در تصفیه‌خانه، برای تعیین نوع و غلظت منعقدکننده مناسب، از آزمایش جار استفاده می‌شود. این آزمایش وقت‌گیر و همراه با خطا است و نمی‌توان زیاد به نتایج آن استناد کرد. برای رفع این مشکل می‌توان از روش‌های هوشمند استفاده کرد. در این تحقیق داده‌های آزمایشگاهی سال ۹۱-۱۳۹۰ پس از جمع‌آوری و پالایش، مورد مطالعه قرار گرفتند. با بهره‌گیری از سیستم استنتاج فازی-عصبی تطبیقی (ANFIS) و شبکه‌های عصبی نوع GMDH و با استفاده از نتایج تجربی به‌منظور دستیابی به مقدار بهینه مصرفی منعقدکننده پلی‌آلومینیوم کلراید در تصفیه‌خانه رشت، دو مدل غیرخطی ارائه شد. اثر پارامترهای ورودی شامل دما، pH، کدورت، جامدات معلق، هدایت الکتریکی و رنگ بر میزان مصرف منعقدکننده بررسی شد. نتایج نشان داد مدل ANFIS نسبت به مدل GMDH کارایی بهتری برای پیش‌بینی میزان مصرف منعقدکننده پلی آلومینیوم کلراید دارد.

واژه‌های کلیدی: تصفیه آب، انعقاد و لخته‌سازی، پلی آلومینیوم کلراید، ANFIS، GMDH

Design of ANFIS Structures and GMDH Type-Neural Network for Prediction of Optimum Coagulant Dosage in Water Treatment Process Case Study: Great Water Treatment Plant in Guilan Province

A. Daghbandan[1] *M. Akbarizadeh[2]*

(Received Dec. 13, 2012 Accepted May 25, 2013)

Abstract

Given the increasing importance of surface water bodies as supply sources of drinking water and regarding the requirement for using different chemicals at various stages of water treatment processes, it is essential to investigate coagulant consumption in water treatment plants. Determination of the required dosage of coagulants used in the coagulation and flocculation unit is one of the most important decisions in water treatment operations. For this purpose, the jar test is generally used to determine the type and concentration of suitable coagulants in a water treatment plant. However, the test is rather time-consuming and unreliable due to the inaccurate results it yields. Instead, intelligent methods can be employed to overcome this shortcoming of the jar test. In this study, experimental data were collected over the period from 2011 to 2012 and further refined for study. Two non-linear models based on adaptive neuro-fuzzy inference system (ANFIS) and GMDH-type neural networks were then developed and experimental results were used to determine the optimum poly-aluminium chloride dosage for use at Guilan water treatment plant. The effects of input parameters including temperature, pH, turbidity, suspended solids, electrical conductivity, and color were investigated on coagulant dosage. The ANFIS model was found to outperform the GMDH model in predicting the required poly-aluminium chloride dosage.

Keywords: Water Treatment, Coagulation and Flocculation, Poly-Aluminium Chloride, ANFIS, GMDH.

1. Assist. Prof. of Chemical Eng., Dept. of Eng., Guilan University, Guilan (Corresponding Author) (+98 131) 7234501 daghbandan@guilan.ac.ir
2. MSc Student of Chemical Eng., Dept. of Eng., Guilan University, Guilan

۱- استادیار گروه مهندسی شیمی، دانشکده فنی و مهندسی، دانشگاه گیلان (نویسنده مسئول) (۰۱۳۱) ۷۲۳۴۵۰۱ daghbandan@guilan.ac.ir
۲- دانشجوی کارشناسی ارشد گروه مهندسی شیمی، دانشکده فنی و مهندسی، دانشگاه گیلان

۱- مقدمه

برای شناسایی سیستم‌ها و مدل‌سازی فرایندهای پیچیده، استفاده از داده‌های ورودی و خروجی همواره مورد توجه محققان بوده است. برای مدل‌سازی از روشهایی که به محاسبات نرم[1] معروف هستند، استفاده می‌شود. مهم‌ترین اجزای محاسبات نرم، منطق فازی[2]، شبکه‌های عصبی[3] و الگوریتم ژنتیک[4] هستند که در کنترل و شناسایی سیستم‌های پیچیده غیر خطی از قابلیت بالایی برخوردارند [۱].

هدف از طراحی سیستم تصفیه آب شرب، تهیه آب شرب با کیفیت بالا در کوتاه‌ترین زمان ممکن و با حداقل هزینه است. فرایند تصفیه به‌علت وجود فعل و انفعالات فیزیکی و شیمیایی، اغلب با مشکل همراه است [۲]. فرایند تصفیه، وابسته به کیفیت و طبیعت آب خام است. فرایند عمومی تصفیه شامل واحدهای ته‌نشینی اولیه، انعقاد و لخته‌سازی، ته‌نشینی ثانویه، فیلتراسیون و گندزدایی است. واحد انعقاد و لخته‌سازی برای حذف ذرات در فرایند تصفیه ایجاد شده است [۳]. تحقیقات بسیاری در زمینه اثر پارامترهای آب خام و انواع منعقدکننده بر روی واحد انعقاد انجام گرفته است [۴و۵]. دمای آب خام می‌تواند اثر معنی‌داری بر روی میزان مصرف منعقد کننده داشته باشد [۶]. pH آب خام نیز یک فاکتور مهم در فرایند انعقاد است. انعقاد بهینه با استفاده از سولفات آلومینیوم در pH بین ۶ تا ۸ اتفاق می‌افتد [۷]. کدورت[5] نیز بر روی مشخصه‌های فیزیکی و شیمیایی آب مؤثر است [۸]. یکی دیگر از پارامترهای مؤثر مربوط به تصفیه آب، هدایت الکتریکی[6] است. هدف از اندازه‌گیری این پارامتر به‌دست آوردن اطلاعاتی راجع به ماهیت مواد معدنی در آب است [۹].

به‌منظور تخمین میزان بهینه مصرف منعقدکننده در واحد انعقاد از آزمایش جار[7] در آزمایشگاه استفاده می‌شود [۱۰]. میزان مصرف منعقد کننده به پارامترهایی چون دما، pH، کدورت، ذرات حل شده جامد[8]، هدایت الکتریکی و رنگ آب خام بستگی دارد. جار تست عموما وقت‌گیر و همراه با خطا است [۱۱]. این آزمایش برای نشان دادن تأثیر مواد شیمیایی در تصفیه‌خانه‌ها طراحی شده است.

در چند دهه اخیر، تحقیقات بسیاری بر روی میزان مصرف منعقدکننده در فرایند تصفیه آب شرب صورت گرفته است. تاکنون برای پیش‌بینی میزان مصرف، مدل‌های شبکه‌های عصبی و سیستم‌های فازی و معادلات چند جمله‌ای ارائه شده است. بیزر و باچی در سال ۱۹۹۶ دو مدل مجزا از معادلات چندجمله‌ای برای تعیین میزان مصرف منعقدکننده در تصفیه‌خانه کلرفونت[9] فرانسه ارائه کردند [۱۲]. گاگنون و همکاران در سال ۱۹۹۷ مدلی از شبکه‌های عصبی مصنوعی برای تخمین و تعیین میزان بهینه از سولفات آلومینیوم در تصفیه‌خانه استفوی در کبک[10] کانادا ارائه دادند [۱۳]. ون لیون ون در سال ۱۹۹۹ یک مدل از شبکه‌های عصبی مصنوعی برای تصفیه‌خانه آب در جنوب استرالیا توسعه داد که به‌منظور تعیین دز بهینه از میزان مصرف سولفات آلومینیوم ارائه شده بود [۱۴]. یو و همکاران در سال ۲۰۰۰ یک مدل شبکه عصبی مصنوعی برای تصفیه‌خانه چان ژو[11] کره جنوبی ارائه کردند [۱۰]. میر و همکاران در سال ۲۰۰۴ با استفاده از داده‌های ون لیون، میزان بهینه مصرف سولفات آلومینیوم را با استفاده از شبکه‌های عصبی پیش‌بینی کردند [۱۵]. اولین بار در سال ۲۰۰۸ میزان بهینه مصرف منعقدکننده در فرایند تصفیه آب با استفاده از سیستم ANFIS[12] توسط وو و لو ارائه شد. این تحقیق بر روی تصفیه‌خانه شهر کانتی تایپه در تایوان[13] انجام شد. نتایج نشان داد که مدل می‌تواند با دقت خوبی میزان مصرف منعقدکننده را پیش‌بینی کند. [۱۶]. سلیم حدام و همکاران نیز دز سال ۲۰۱۱ با استفاده از سیستم ANFIS، میزان مصرف منعقد کننده در تصفیه‌خانه شهر بودوا الجزایر[14] را پیش‌بینی کردند [۹]. مدل‌های ایجاد شده توسط محققان با توجه به پارامترهای مؤثر بر کیفیت آب خام که به‌عنوان ورودی در نظر گرفته می‌شوند و میزان مصرف منعقدکننده که به‌عنوان خروجی شبکه در نظر گرفته می‌شود، انجام شده است.

در این تحقیق روشهای شبکه‌ای چند هدفی عصبی-فازی (ANFIS) و شبکه‌های عصبی دسته‌بندی گروهی داده‌های عددی[15] برای حل مسئله مورد استفاده قرار گرفتند [۱۷و۱۸].

مدل‌سازی فرایند با استفاده از داده‌های آزمایشگاهی شامل ۳۰۸ داده که مربوط به مصرف پلی آلومینیوم کلراید بود، انجام شد. مدل شامل شش ورودی دما، pH، کدورت، ذرات حل شده جامد، هدایت الکتریکی و رنگ و یک خروجی که میزان مصرف منعقدکننده است، بود. به‌منظور اطمینان از صحت روشهای پیشنهادی، دقت مدل‌های ارائه شده نسبت به رفتار داده‌های تجربی

---

1 Soft Computing
2 Fuzzy Logic
3 Neural Network
4 Genetic Algorithms
5 Turbidity
6 Electrical Conductivity
7 Jar Tests
8 Suspended of Solid
9 Clairfont
10 Stefoy water treatment plant in Quebec
11 Chungju
12 Adaptive Neuro-Fuzzy Inference System (ANFIS)
13 Taipei County, Taiwan
14 Boudouaou, Algeria
15 Group Method of Data Handling (GMDH)

توسط ضریب همبستگی[1] و مجذور میانگین مربعات خطا[2] که بیانگر خطای مدل‌های ارائه شده نسبت به داده‌های تجربی می‌باشند، ارائه شد.

۲- مواد و روشها

۲-۱- مطالعه موردی

تصفیه‌خانه بزرگ آب گیلان وظیفه تأمین آب آشامیدنی شهرها و روستاهای مناطق مرکزی و بخشی از مناطق شرق وغرب را به‌عهده دارد. جمعیت تحت پوشش این طرح، حداکثر دو میلیون و سیصد هزار نفر است. حداکثرآب مورد نیاز جمعیت تحت پوشش به میزان هفت هزار و پانصد لیتر در ثانیه برآورد شده که هشتاد درصد آن، از این تصفیه‌خانه و بقیه از سایر منابع پراکنده منطقه تأمین خواهد شد. آب خام مورد نیاز این تصفیه‌خانه از رودخانه سفیدرود و رودخانه شهربیجار تأمین می‌شود. وضعیت آب این رودخانه‌ها در فصلهای مختلف، بسیار متغیر است. این تصفیه‌خانه در فاصله حدود بیست کیلومتری جنوب شرقی شهر رشت و در مجاورت کانال چپ سنگر قرار گرفته و در دو مرحله احداث شده است که ظرفیت هر مرحله آن سه هزار لیتر در ثانیه است. بهره‌برداری از مرحله اول تصفیه‌خانه از بهمن سال ۱۳۷۸ شروع شده و در حال حاضر با حداکثر ظرفیت در مدار بهره‌برداری است. این تصفیه‌خانه برای حداکثر غلظت مواد معلق ورودی آب خام تا پانزده گرم در لیتر طراحی شده و با افزایش غلظت تا بیست گرم در لیتر نیز قادر به تصفیه آب است.

۲-۲- جمع‌آوری داده‌ها

در واحد آزمایشگاه، میزان مصرف منعقدکننده به پارامترهایی چون دما، pH، کدورت، ذرات حل شده جامد، هدایت الکتریکی و رنگ آب بستگی دارد. از منعقد کننده‌های پلی آلومینیوم کلراید، سولفات آلومینیوم و پلی الکترولیت در فرایند، با توجه به وضعیت آب خام استفاده می‌شود. در تصفیه‌خانه برای نشان دادن تأثیر و نوع مواد شیمیایی و میزان مواد منعقدکننده مصرفی از روش جارتست استفاده می‌کنند. از نتایج حاصل از انجام این آزمایش‌ها برای ارائه یک مدل غیرخطی به‌منظور تعیین میزان مصرف منعقد کننده استفاده می‌شود. مدل‌سازی فرایند با استفاده از داده‌های آزمایشگاهی سال ۹۱-۱۳۹۰ شامل ۳۰۸ داده که مربوط به مصرف پلی آلومینیوم کلراید است، انجام شد. مدل شامل شش پارامتر مؤثر به‌عنوان ورودی و یک خروجی بود. در آزمایشگاه برای کنترل و اندازه‌گیری متغیرهای مؤثر بر میزان مصرف منعقدکننده از دستگاههای اشاره شده در جدول ۱ استفاده شد. میزان مصرف منعقدکننده با استفاده از جارتست تعیین شد. در جدول ۲ به‌عنوان نمونه تعدادی از داده‌های آزمایشگاهی گزارش شده است.

جدول ۱- دستگاههای اندازه‌گیری متغیرهای ورودی مؤثر بر میزان مصرف منعقدکننده

دستگاه	مدل	ساخت
دماسنج	330i/Tetrocon325, WTW	آلمان
pH متر	SenTix 41, WTW	آلمان
کدورت سنج	2100 N, HACH	امریکا
اندازه‌گیری ذرات حل شده جامد	DR/2500, HACH	امریکا
هدایت سنج	330i/Tetrocon325, WTW	آلمان
رنگ (طیف سنج)	DR/2500, HACH	امریکا

جدول ۲- داده‌های آزمایشگاهی حاصل از آزمایش جار

No	ورودی						خروجی
	دما (°C)	pH (-)	کدورت (NTU)	جامدات معلق (mg/L)	هدایت الکتریکی (µs/cm)	رنگ (pt-co)	میزان منعقدکننده (ppm)
۱	۲۳/۴	۸/۲۷	۲۵/۴	۲۹	۱۰۹۳	۲۰۵	۶
۲	۲۳/۹	۸/۰۲	۳۷/۸	۴۶	۱۷۹۰	۳۹۷	۱۱
۳	۲۳/۸	۸/۲۸	۲۶/۶	۹۷	۱۵۰۰	۲۳۷	۴
۴	۱۵/۷	۸/۰۵	۱۵۴	۲۴۹	۴۴۵	۳۵۷۰	۱۳
۵	۲۴/۶	۸/۱۸	۲۳/۶	۸۵	۱۴۴۸	۲۴۴	۵
۶	۱۵/۵	۸/۱۲	۷۴۷	۱۱۲۰	۶۳۷	۳۶۳۴	۲۴
۷	۱۶/۱	۸/۰۹	۸۸	۲۱۰	۶۰۱	۷۴۰	۹
۸	۱۶/۱	۷/۹۹	۱۰۵	۱۵۲	۶۵۶	۹۹۲	۱۰

۲-۳- سیستم استنتاجی عصبی- فازی تطبیقی (ANFIS)

نظریه مجموعه فازی در سال ۱۹۶۵ توسط لطفی عسگرزاده در دانشگاه برکلی امریکا ارائه شد [۱۹]. نظریه مجموعه‌های فازی، نظریه‌ای است برای اقدام در شرایط عدم اطمینان؛ این نظریه قادر است بسیاری از مفاهیم و متغیرها و سیستم‌هایی را که نادقیق و مبهم هستند، صورت‌بندی ریاضی ببخشد و زمینه را برای استدلال، استنتاج، کنترل و تصمیم‌گیری در شرایط عدم اطمینان فراهم آورد. سیستم استنتاج عصبی- فازی تطبیقی از الگوریتم‌های یادگیری شبکه عصبی و منطق فازی به‌منظور طراحی نگاشت غیر خطی بین فضای ورودی و خروجی استفاده می‌کند و همچنین با توجه به توانایی در ترکیب قدرت زبانی یک سیستم فازی با قدرت عددی یک شبکه عصبی در مدل‌سازی فرایند مذکور بسیار قدرتمند است. این سیستم اولین بار در سال ۱۹۹۳ توسط جانگ و همکاران ارائه شد [۲۰].

در مجموع ANFIS از ساختاری پنج لایه با تعدادی متغیر ورودی تشکیل می‌شود که هر ورودی، دو یا چند تابع عضویت

[1] Coefficient of Determination
[2] Root Mean Square Error

دارد. در لایه اول (ورودی) میزان تعلق هر ورودی به بازه‌های مختلف فازی توسط کاربر مشخص می‌شود. با ضرب مقادیر ورودی به هر گره در یکدیگر، وزن قانون‌ها در لایه دوم به‌دست می‌آید. در لایه سوم، عمل محاسبه وزن نسبی قوانین انجام می‌گیرد. لایه چهارم لایه قوانین است که از انجام عملیات بر روی سیگنال‌های ورودی به این لایه حاصل می‌شود. لایه آخر، خروجی شبکه است که هدف آن حداقل نمودن اختلاف خروجی به‌دست آمده از شبکه و خروجی واقعی است [۱۸]. در جدول ۳، پارامترهای ساختاری ANFIS آورده شده است. در این بخش، از مدل سیستم استنتاج فازی- عصبی تطبیقی سوگانو مرتبه اول برای پیش‌بینی میزان بهینه مصرف منعقدکننده استفاده شد [۲۰]. برای متغیرهای ورودی از تابع گاوسین استفاده شد (رابطه ۱). از الگوریتم ژنتیک و روش حل معادلات نرمال[1] (روابط ۲ و۳) برای طراحی و بهینه نمودن پارامترهای غیرخطی در لایه اول (c_i, σ_i) و پارامترهای خطی لایه چهارم (a_0, a_i) (رابطه ۴) استفاده شد [۱۸].

جدول ۳- پارامترهای ساختاری مدل (ANFIS) و (GMDH)

جمعیت اولیه	۱۰۰	مقدار جهش	۰/۱
تعداد تکرار	۱۰۰۰	مقدار تقاطع	۰/۹۸
تعداد تابع عضویت هر ورودی در مدل ANFHS	۲	تعداد قواعد فازی در مدل ANFIS	۶۴
تعداد لایه های مخفی در مدل GMDH	۳	تعداد توابع هدف	۲

$$\mu_A^{(j)} = \exp\left(-\frac{(x_i - c_j)^2}{2\sigma_j^2}\right) \quad (۱)$$

$$A = \begin{bmatrix} 1 & X_{1,1} & X_{2,1} & X_{N,1} \\ 1 & X_{1,2} & X_{2,2} & X_{N,2} \\ 1 & & & \\ 1 & X_{1,M} & X_{2,M} & X_{N,M} \end{bmatrix} \quad (۲)$$

$$a = \left(A_i^T A_i\right)^{-1} A_i^T Y_{i,exp} \quad (۳)$$

$$f_i = a_0 + \sum_{i=1}^{N} a_i X_i \quad (۴)$$

۲-۴- شبکه عصبی نوع GMDH

شبکه‌های عصبی GMDH صورتی از الگوریتم GMDH است که به فرم و ساختار شبکه‌ای بیان شده است. شبکه عصبی GMDH، شبکه‌ای خود سامانده و یک سویه است که از چند لایه و هر لایه نیز از چندین نرون تشکیل شده است. تمامی نرون‌ها از یک ساختار مشابه برخوردارند. وزن‌ها (w) بر اساس روش حل معادلات متعامد (SNE) به‌عنوان مقادیر مشخص و ثابت در داخل هر نرون جایگذاری می‌شود. ویژگی بارزی که در این نوع از شبکه‌ها مشاهده می‌شود، حاکی از آن است که نرون‌های مرحله قبلی و یا لایه قبلی عامل و یا مولد تولید نرون‌های جدید هستند. از میان نرون‌های تولید شده، لزوماً باید تعدادی از آنها حذف شوند تا به این وسیله از واگرایی شبکه جلوگیری به‌عمل آید. اصطلاحاً به این‌گونه نرون‌ها، نرون مرده گفته می‌شود. یکی از مسائل مهمی که در شبکه‌های عصبی مصنوعی چند لایه مطرح می‌شود، طراحی ساختار شبکه است. در این طراحی باید تعداد لایه‌ها و نیز ساختار درونی از قبیل تعداد وزن‌ها و مقادیر اولیه آنها و همچنین تابع تحریک هر نرون به صورت مناسب انتخاب شوند تا یک نگاشت مناسب و ایده‌آل میان داده‌های ورودی و خروجی برقرار شود. یکی از اهداف شبکه‌های عصبی GMDH جلوگیری از رشد واگرایی شبکه و نیز مرتبط کردن شکل و ساختار شبکه به یک یا چند پارامتر عددی است؛ به گونه‌ای که با تغییر این پارامتر، ساختار شبکه نیز تغییر کند [۱۸و۲۱]. این الگوریتم بر اساس تجزیه سری توابع ولترا به چند جمله‌ای‌های دو متغیره درجه دوم پایه‌ریزی شده است (رابطه ۵).

$$G(x_i,x_j)=a_0+a_1x_i+a_2x_j+a_3x_i^2+a_4x_j^2+a_5x_ix_j \quad (۵)$$

در این تجزیه، سری ولترا به مجموعه‌ای از معادلات بازگشتی زنجیره‌ای تبدیل می‌شود، به‌گونه‌ای که مجدداً با جایگذاری جبری هر یک از روابط بازگشتی در یکدیگر، سری ولترا برقرار شود. از روش حل معادلات نرمال (SNE) برای محاسبه ثوابت سری ولترا استفاده می‌شود (روابط ۶ و ۷). الگوریتم این گونه شبکه‌ها به گونه‌ای بسته می‌شود که مدل در زمان کمتری به همگرایی برسد [۱۹].

(۶)

$$A = \begin{bmatrix} 1 & x_{1p} & x_{1q} & x_{1p}^2 & x_{1q}^2 & x_{1p}x_{1q} \\ 1 & x_{2p} & x_{2q} & x_{2p}^2 & x_{2q}^2 & x_{2p}x_{2q} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{Mp} & x_{Mq} & x_{Mp}^2 & x_{Mq}^2 & x_{Mp}x_{Mq} \end{bmatrix}$$

$$a = (A^T A)^{-1} A^T Y_{exp} \quad (۷)$$

شبکه عصبی GMDH متشکل از تعدادی نرون و لایه است که نرون‌های هر لایه به‌صورت غیر خطی با یکدیگر ترکیب شده و نرون‌های جدیدی را تولید می‌نمایند که این نرون‌های جدید، لایه بعدی مدل را تشکیل می‌دهند. هر نرون از دو ورودی ساخته شده است که توسط رابطه خطی سری ولترا با یکدیگر ترکیب می‌شوند و

[1] Solve Normal Equation (SNE)

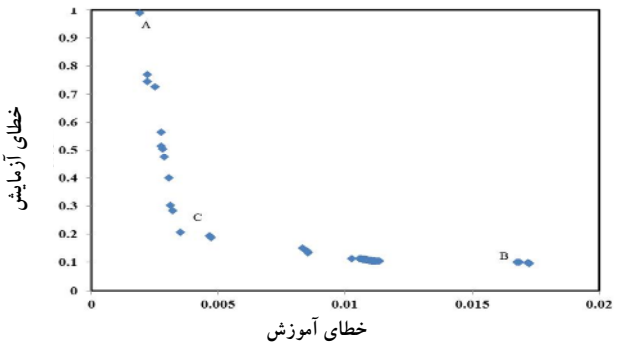
خروجی نرون به‌دست می‌آید. طبق روشی که پیش از این توضیح داده شد، ضرایب a_0 تا a_5 به‌دست می‌آیند. در این روش، نرون‌ها در قالب یک کروموزوم با هم ترکیب می‌شوند. کروموزم که جزئیات ساختاری شبکه GMDH را نشان می‌دهد، از رشته‌هایی که اجزای آن حروف الفبا هستند، ساخته شده است. در این روش کدگذاری، داده‌های مختلف ورودی با حروف الفبا نامگذاری می‌شود. هر کروموزوم، رشته‌ای است که از به هم پیوستن داده‌های ورودی نامگذاری شده، ایجاد شده است.

۳- نتایج و بحث

در واحد انعقاد، به‌منظور حذف فیزیکی ذرات جامد، از منعقدکننده پلی‌آلومینیوم کلراید استفاده می‌شود. آزمایش‌های مورد نظر در واحد آزمایشگاه تصفیه‌خانه آب گیلان انجام و داده‌های تجربی مورد استفاده برای ایجاد مدل‌ها گردآوری شدند. مدل‌سازی با استفاده از داده‌های آزمایشگاهی شامل شش ورودی و یک خروجی انجام شد (جدول ۲). پارامترهای مدل‌های ارائه شده با در نظر گرفتن دو تابع هدف خطای آموزشی (TE) و خطای آزمایشی (PE) (روابط ۸ و ۹) که مربوط به اختلاف نتایج حاصل از مدل‌ها با نتایج تجربی میزان مصرف منعقدکننده است، تعیین می‌شود. این پارامترها باید به گونه‌ای تعیین شوند که مقادیر خطای آموزشی و آزمایشی داده‌ها به‌طور همزمان نسبت به همدیگر حداقل باشد (نقاط مصالحه در نمودار پرتو (شکلهای ۱ و ۲)).

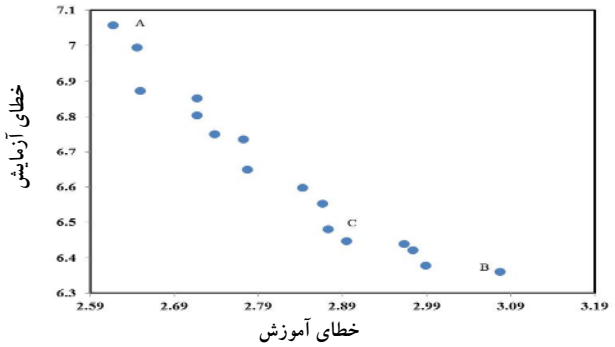
$$TE = \sum_{i=1}^{n} (Y_{i,exp} - Y_{i,train})^2 \qquad (۸)$$

$$PE = \sum_{i=1}^{n} (Y_{j,exp} - Y_{j,prediction})^2 \qquad (۹)$$



شکل ۱- نقاط طراحی برای مدل چندهدفی ANFIS

برای ایجاد مدل‌ها، از هفتاد درصد داده‌ها به‌عنوان داده‌های آموزشی، و مابقی به‌عنوان داده‌های آزمایشی استفاده شد. در نهایت



شکل ۲- نقاط طراحی برای مدل چندهدفی شبکه عصبی نوع GMDH

بعد از مدل‌سازی برای محاسبه خطا و دقت مدل‌ها از روابط مجذور میانگین مربعات خطا و ضریب همبستگی استفاده شد (روابط ۱۰ و ۱۱).

$$RMSE = \left[\sum_{i=1}^{n} (Y_{i,exp} - Y_{i,model})^2 \Big/ n \right]^{0.5} \qquad (۱۰)$$

$$R^2 = 1 - \left[\sum_{i=1}^{n} (Y_{i,exp} - Y_{i,model})^2 \Big/ \sum_{i=1}^{n} (Y_{i,exp})^2 \right] \qquad (۱۱)$$

در شکلهای ۱ و ۲، منحنی مقدار خطای آزمایش بر حسب خطای آموزش که اصطلاحاً منحنی پارتو[1] نامیده می‌شود، نشان داده شده است. کاملاً مشخص است که تمامی نقاط پارتو از دید دو تابع هدف نسبت به هم غیر برترند. همچنین همان‌طور که در شکل مشخص شده است، نقطه A دارای کمترین خطای آموزشی و نقطه B دارای کمترین خطای پیش‌بینی است. با حرکت از نقطه A به سمت نقطه B، خطای مدل‌سازی افزایش و خطای پیش‌بینی کاهش می‌یابد. به‌عبارت دیگر، بهبود در یکی از توابع هدف یک نقطه باعث بدتر شدن تابع هدف دیگر آن شده است. نقطه مصالحه طراحی از بین تمامی نقاط پارتو، نقطه C می‌تواند باشد؛ زیرا خطای مدل‌سازی نسبت به خطای آموزشی در این نقطه، دارای توازن بیشتری نسبت به نقاط دیگر است. پارامترهای مدل‌های ارائه شده به ازای تمامی نقاط پارتو با توجه به دو تابع هدف خطای آموزش و آزمایش به‌طور همزمان محاسبه می‌شوند و مقادیر پارامترهای نقطه مصالحه به‌عنوان مقادیر بهینه در نظر گرفته می‌شوند.

مقادیر خطاهای آموزش و آزمایش نقاط مشخص شده در شکلهای ۱و۲، در جدول ۴ ذکر شده است.

در مدل ANFIS، به ازای هر ورودی در لایه اول، دو تابع

[1] Pareto

جدول ۴- مقدار نقطه بهینه از دید هر تابع هدف به همراه نقطه مصالحه با دو تابع هدف

خروجی		نقطه طراحی	TE	PE
پلی آلومینیوم کلراید	GMDH	A	۲/۶۱۷۴	۷/۰۵۶۳
		B	۳/۰۷۷۵	۶/۳۵۹۷
		C	۲/۸۷۳۲/	۶/۴۷۰۱
	ANFIS	A	۰/۰۰۱۹	۰/۹۸۸۸
		B	۰/۰۱۷۲	۰/۰۹۷۷
		C	۰/۰۰۳۵	۰/۲۰۵۷

عضویت در نظر گرفته شد و مقادیر بهینه ثوابت تابع گاوسین با استفاده از الگوریتم ژنتیک برای نقطه مصالحه محاسبه شد. توابع عضویت گوسی برای نقطه طراحی به‌ازای هر ورودی در شکل ۳، نشان داده شده است.

مقادیر پارامترهای توابع عضویت گوسی به‌کار رفته در مدل ANFIS در جدول ۵ آورده شده است.

جدول ۵- مقادیر ثوابت توابع عضویت گوسی مدل دوهدفی ANFIS

ورودی‌ها	Gaussian MF(1)		Gaussian MF(2)	
	C	σ	C	σ
دما (°C)	۲۳/۳۱۱۱۱	۰/۹۹۳۳۶۵	۱۰/۹۲۶۹۸	۱۲/۸۸۲۱۳
pH (-)	۸/۳۳۶۶۶۷	۰/۱۱۰۶۵۲	۷/۹۰۲۳۸۱	۰/۴۴۲۳۳۸
کدورت(NTU)	۱۷۴/۴۶۷۸	۱۵۶/۵۸۴۹	۲۳۳۴/۲۵۹	۱۵۶/۵۸۴۹
جامدات معلق (mg/L)	۱۰۰۶/۵۸۷	۱۱۷/۸۴	۳۵۶۲/۹۰۵	۱۹۹۳/۳۰۴
هدایت الکتریکی (µs/cm)	۱۵۸۲/۹۵۲	۱۵۸۱/۵۸۵	۱۵۰۵/۳۱	۹۷/۵۷۱۱۹
رنگ(pt-co)	۷۸۲/۱۵۸۷	۱۴۲۰/۳۸۵	۳۲۵۳/۵۸۷	۱۲۹/۷۵

در مدل GMDH نیز با استفاده از الگوریتم ژنتیک، ساختار بهینه به‌دست آمد. در این روش، کروموزومی که کمترین میزان خطای آموزش و آزمایش را دارد، به‌عنوان کروموزوم برتر انتخاب می‌شود. برای نقطه مصالحه طراحی C، خطای آموزش و آزمایش به‌طور همزمان حداقل هستند (شکل ۲). پس از به‌دست آوردن نقطه پاره تو بهینه، کروموزوم برتر به‌دست می‌آید که در اینجا با سه لایه مخفی، کروموزوم مورد نظر به طول 2^4 است که توان ۴، همان تعداد لایه های مخفی + ۱ است. لذا کروموزوم به‌دست آمده ۱۶ رقمی است و هر رقم مربوط به یک ورودی است. کروموزوم برتر برای پیش‌بینی میزان مصرف منعقدکننده در مدل GMDH به صورت جدول ۶ است.

باید توجه داشت که هر عدد متعلق به یک پارامتر ورودی است که ترتیب آنها به شرح زیر است: ۱-دما ۲-pH ۳-کدورت ۴- ذرات جامد معلق ۵- هدایت الکتریکی ۶-رنگ.

در این روش در لایه مخفی اول، نرون‌ها به‌صورت ۲ به ۲ با یکدیگر ترکیب می‌شوند و ترکیب آنها به‌صورت ۱۱۱۴۱۶۵۵۱۲۱۵۲۳۲۳ است. برای سهولت در ارائه جدول، نرون‌ها نامگذاری می‌شوند، لذا برای مصرف منعقدکننده پلی‌آلومینیوم کلراید در لایه اول: Y1=14 Y2=,16 Y3=12 Y4=15 Y5=23 Y6=24. پس از تعیین ضرایب مربوط به نرون‌های لایه اول نوبت به ضرایب لایه دوم می‌رسد که در این لایه طول هر کروموزوم برابر ۴ است و نامگذاری نرون‌ها نیز به این ترتیب Y7=1114 Y8=1655 Y9=1215 Y10=2324 است.

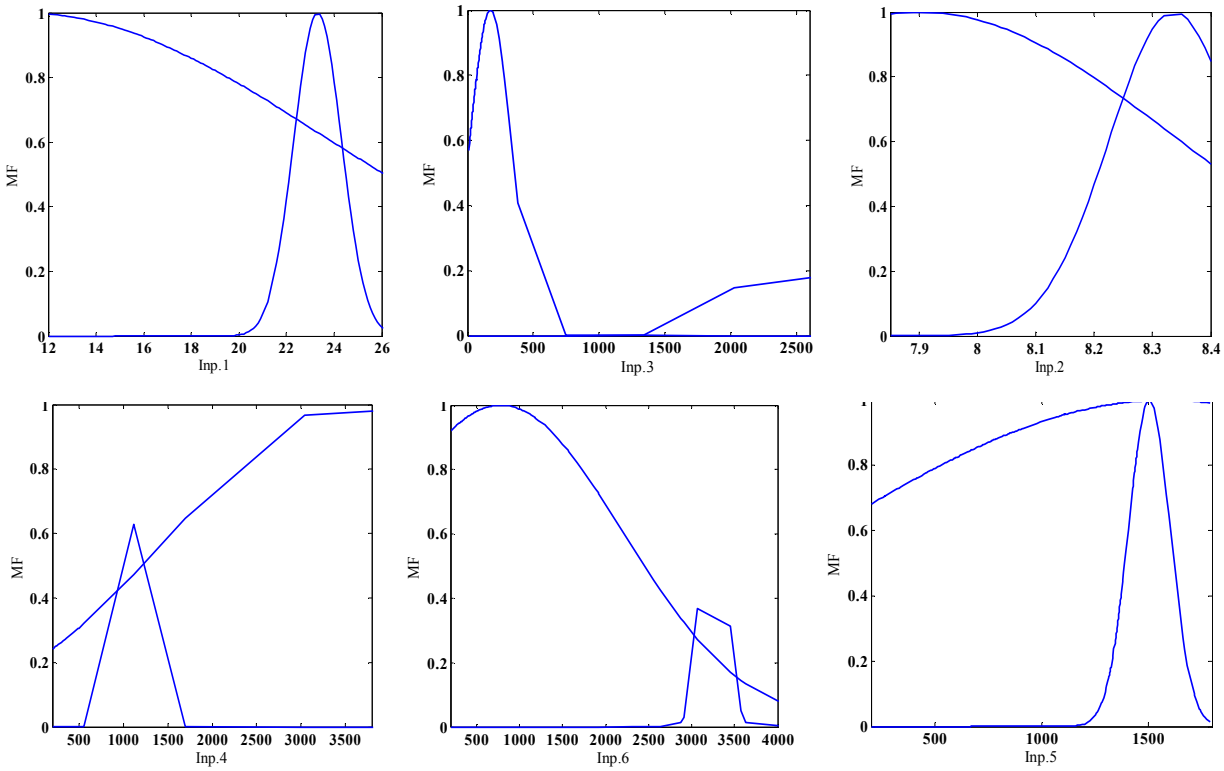
در لایه مخفی سوم طول هر کروموزوم ۸ است و نامگذاری آنها به‌صورت : Y_{11}=11141655 Y_{12}=12152324 است. مدل ریاضی و ثوابت سری ولترا در مدل GMDH برای تمامی نرون‌های مربوط به میزان مصرف منعقدکننده پلی‌آلومینیوم کلراید در جدول ۷ ارائه شده است. در انتهای مدل‌سازی ساختار به‌دست آمده برای پیش‌بینی میزان مصرف منعقدکننده در شکل ۴ نشان داده شده است. Y_i نشان دهنده نرون در لایه‌های مختلف است.

جدول ۶- کروموزوم برتر برای پیش‌بینی میزان مصرف منعقدکننده

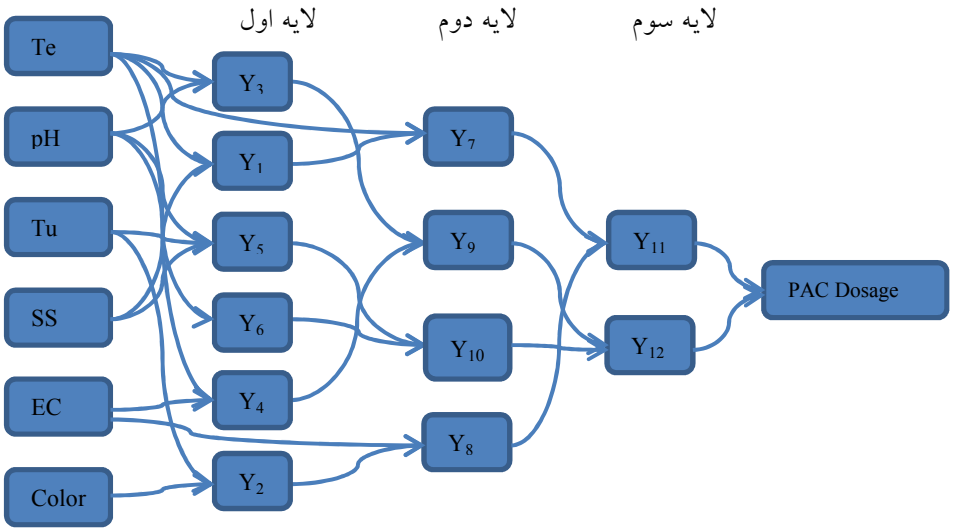
خطای آزمایش	خطای آموزش	کروموزوم بهینه به‌دست آمده توسط مدل شبکه عصبی نوع GMDH															
۶/۴۸۰۱	۲/۸۷۳۲	۴	۲	۳	۲	۵	۱	۲	۱	۵	۵	۶	۱	۴	۱	۱	۱

جدول ۷- مدل ریاضی برای محاسبه میزان مصرف پلی آلومینیوم کلراید با استفاده از شبکه عصبی نوع GMDH

لایه	شماره نرون	مقادیر ثوابت سری ولترا در مدل GMDH
لایه اول	Y_1=	$0.0624 + 0.6127\ Te + 0.0552\ SS - 0.0131\ (Te)^2 - 1.8e\text{-}6\ (SS)^2 - 0.0026\ (Te)\ (SS)$
	Y_2=	$5.4485 + 0.0223\ Tu + 0.0016\ Color - 4e\text{-}6\ (Tu)^2 + 2.3e\text{-}9\ (Color)^2 - 2e\text{-}6\ (Tu)\ (Color)$
	Y_3=	$-1447.1307 + 3.8493Te + 356.2078\ pH + 0.0244\ (Te)^2 - 21.5819\ (pH)^2 - 0.6230\ (Te)\ (pH)$
	Y_4=	$0.1512 + 1.4195\ Te - 0.0033\ EC - 0.0583\ (Te)^2 - 2.8e\text{-}6\ (EC)^2 + 0.0005\ (Te)\ (EC)$
	Y_5=	$2.6847 + 10.8590\ pH - 0.6380\ Tu - 1.2858\ (pH)^2 - 8.3e\text{-}6\ (Tu)^2 + 0.0821\ (pH)\ (Tu)$
	Y_6=	$2.7282 + 11.0558\ pH - 0.2357\ SS - 1.3079\ (pH)^2 - 2.27e\text{-}6\ (SS)^2 + 0.0307\ (pH)\ (SS)$
لایه دوم	Y_7=	$31.7789 - 3.4443\ Te - 1.0684\ Y_1 + 0.0521\ (Te)^2 - 0.0453\ (Y_1)^2 + 0.2280\ (Te)\ (Y_1)$
	Y_8=	$-2.4761 + 1.2766\ Y_2 - 0.0015\ EC - 0.0167\ (Y_2)^2 + 2.5e\text{-}7\ (EC)^2 + 3.5e\text{-}4\ (Y_2)\ (EC)$
	Y_9=	$11.5186 + 1.7401\ Y_3 - 3.9454\ Y_4 - 0.0321\ (Y_3)^2 + 0.3091\ (Y_4)^2 - 0.0617\ (Y_3)\ (Y_4)$
	Y_{10}=	$0.6633 + 3.1189\ Y_5 - 2.2315\ Y_6 - 0.1697\ (Y_5)^2 - 0.0532\ (Y_6)^2 + 0.2279\ (Y_5)\ (Y_6)$
لایه سوم	Y_{11}=	$-2.9489 + 0.7291\ Y_7 + 0.7846\ Y_8 + 0.0677\ (Y_7)^2 + 0.0483\ (Y_8)^2 - 0.133\ (Y_7)\ (Y_8)$
	Y_{12}=	$2.9822 - 0.8669\ Y_9 + 1.0309\ Y_{10} + 0.0585\ (Y_9)^2 - 0.0014\ (Y_{10})^2 - 0.0012\ (Y_9)\ (Y_{10})$
	PAC Dosage	$-1.0513 + 0.4845\ Y_{11} + 0.7213\ Y_{12} - 0.2219\ (Y_{11})^2 - 0.2431\ (Y_{12})^2 + 0.4614\ (Y_{11})\ (Y_{12})$



شکل ۳- توابع عضویت طراحی شده توسط الگوریتم ژنتیک با دو تابع هدف برای مدل ANFIS



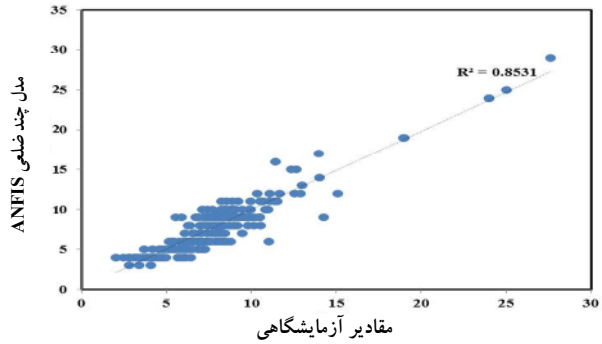
شکل ۴- ساختار شبکه عصبی نوع GMDH برای محاسبه میزان مصرف منعقدکننده پلی آلومینیوم کلراید

خروجی حاصل از مدل‌های ANFIS و GMDH بر حسب خروجی واقعی برای هر دسته داده برای نقطه مصالحه طراحی C در شکلهای ۵ و ۶ نشان داده شده است.

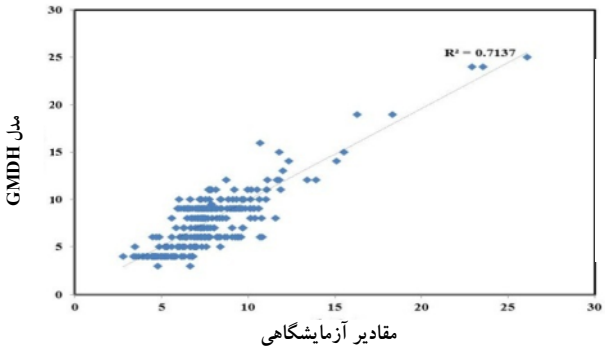
در شکل ۷ مدل‌های ارائه شده برای مدل‌سازی و پیش‌بینی میزان مصرف منعقدکننده در واحد آزمایشگاه با یکدیگر مقایسه شده‌اند. دقت و مقادیر خطای مدل‌ها در جدول ۸ آورده شده است.

همچنین مقادیر حداقل و حداکثر داده‌های تجربی با نتایج حاصل از مدل مقایسه و در جدول ۸ گزارش شده‌اند.

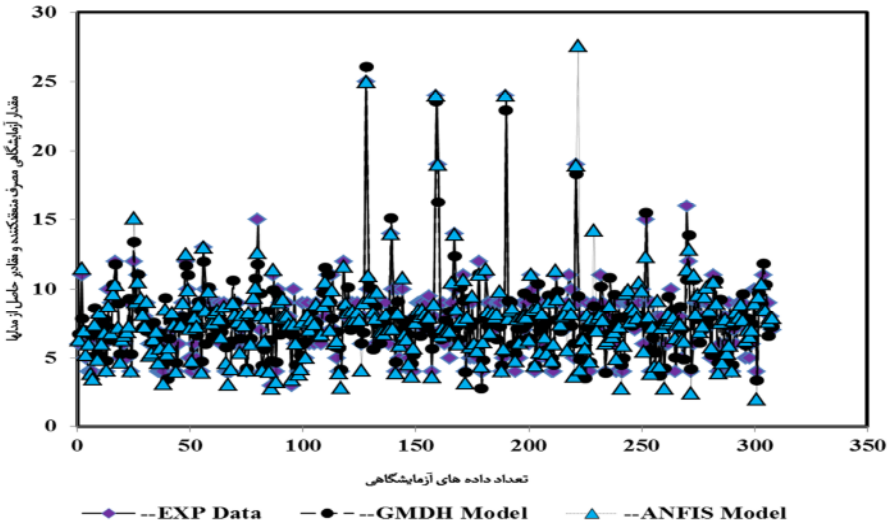
نتایج حاکی از آن است که مدل‌های ارائه شده، کارایی لازم را به‌منظور پیش‌بینی مصرف منعقدکننده دارند و مدل ANFIS عملکرد بهتری نسبت به مدل GMDH از خود نشان داد.



شکل ۵- مقایسه خروجی واقعی و خروجی مدل ANFIS مربوط به مصرف پلی آلومینیوم کلراید



شکل ۶- مقایسه خروجی واقعی و خروجی مدل GMDH مربوط به مصرف پلی آلومینیوم کلراید



شکل ۷- مقایسه مدل های ارائه شده برای مدل سازی و پیش بینی میزان مصرف منعقدکننده پلی آلومینیوم کلراید در واحد آزمایشگاه

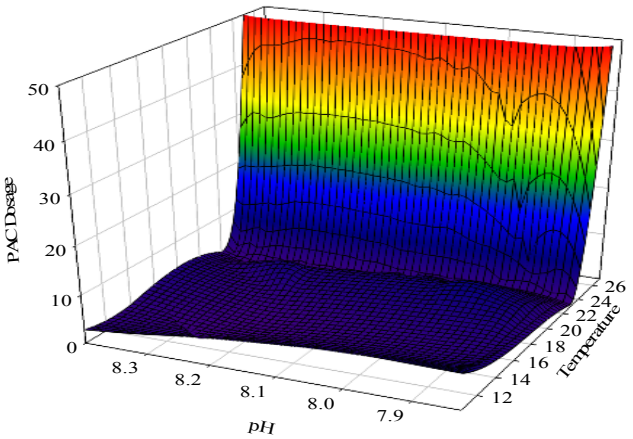
جدول ۸- خطای حاصل از مدل سازی مربوط به میزان مصرف منعقدکننده در واحد آزمایشگاه

مدل	دز حداقل [ppm]		دز حداکثر [ppm]		RMSE	R^2
	EXP.	Model	EXP.	Model		
ANFIS	۳	۲/۷۸۵۵	۲۵	۲۵	۱/۲۶۹۷	۰/۹۷۶۲
GMDH	۳	۴/۷۹۹۸	۲۵	۲۶/۱۰۱۲	۱/۶۴۵۰	۰/۹۶/۱

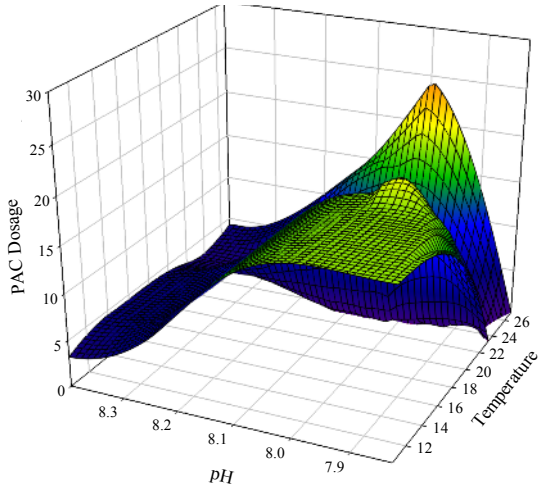
جدول ۹- مقادیر بهینه متغیرهای مستقل مؤثر بر میزان مصرف منعقدکننده شامل ورودی‌ها و خروجی

نوع مدل ---	ورودی‌های مدل						خروجی مدل
	دما (°C)	pH (-)	کدورت (NTU)	جامدات معلق (mg/L)	هدایت الکتریکی (µs/cm)	رنگ (pt-co)	میزان مصرف منعقدکننده (ppm)
GMDH	۱۸/۰۹۰	۸/۰۸۸	۴۰۶/۶۰۱	۱۴۲۳/۶۳۵	۱۲۵۶/۲۰۰	۵۷۱/۷۱۴	۰
ANFIS	۲۳/۰۵۵	۸/۰۵۲	۱۲۷۸/۵۰۰	۲۵۴۱/۷۰۰	۱۲۵۸/۱۰۰	۲۰۳۳/۰۰۰	۰

در نهایت برای بهینه‌سازی پارامترهای مؤثر بر میزان مصرف منعقدکننده پلی‌آلومینیوم، از مدل‌های چند هدفی شبکه‌های عصبی نوع GMDH و ساختارهای فازی ANFIS که برای مدل‌سازی فرایند مورد استفاده قرار گرفتند، استفاده شد. مدل‌های ایجاد شده، تابع متغیرهای مستقل دمای آب خام، pH آب خام، ذرات حل شده جامد، کدورت، هدایت الکتریکی و رنگ آب خام هستند. برای



شکل ۸- اثر متغیرهای مستقل دما و pH آب خام بر روی میزان مصرف منعقدکننده پلی آلومینیوم کلراید در شرایط بهینه برای مدل GMDH



شکل ۹- اثر متغیرهای مستقل دما و pH آب خام بر روی میزان مصرف منعقدکننده پلی آلومینیوم کلراید در شرایط بهینه برای مدل ANFIS

بهینه‌سازی میزان مصرف از الگوریتم ژنتیک استفاده شد. مقادیر بهینه پارامترهای مؤثر بر میزان مصرف منعقدکننده برای هر دو مدل در جدول ۹ آورده شده است. در این شرایط میزان مصرف منعقدکننده حداقل است. اثر پارامترهای دما و pH آب خام بر روی میزان مصرف در شرایط بهینه در شکل‌های ۸ و ۹ نشان داده شده است. در شکل‌های ۸ و ۹ بقیه پارامترها، مقدار بهینه خود را خواهند داشت. نتایج نشان می‌دهد که با افزایش دمای آب خام تا ۲۳ درجه سلسیوس، میزان مصرف پلی‌آلومینیوم کلراید کاهش و بعد از آن افزایش می‌یابد و در pH بین ۸/۱ تا ۸/۲، میزان مصرف، حداقل است.

۴- نتیجه‌گیری

نتایج نشان می‌دهد که سیستم استنتاج فازی-عصبی تطبیقی (ANFIS) و مدل شبکه عصبی نوع GMDH می‌توانند با دقت قابل قبولی میزان مصرف منعقدکننده را پیش‌بینی نمایند.

بیشترین تأثیر بر روی میزان مصرف، مربوط به دما و pH آب خام است. بهترین عملکرد منعقدکننده در pH بین ۸/۱ تا ۸/۲ است. میزان مصرف منعقدکننده در کدورت‌های بسیار پایین افزایش می‌یابد. با افزایش کدورت، ذرات حل شده جامد، هدایت الکتریکی و رنگ، میزان مصرف منعقدکننده افزایش و با افزایش دمای آب خام تا ۲۳ درجه سلسیوس، میزان مصرف کاهش و بعد از آن افزایش می‌یابد.

استفاده از بهینه‌سازی چند هدفی در طراحی سیستم ANFIS و GMDH برای سیستم‌ها و فرایندهای پیچیده، مجموعه نقاط طراحی بهینه غیر برتر (پارتو) را به‌دست می‌دهد و طراحان با در دست داشتن تمامی نقاط طراحی بهینه، می‌توانند نقطه مصالحه طراحی را انتخاب کنند.

مدل‌سازی رفتار سیستم‌ها با استفاده ترکیبی از ANFIS و الگوریتم ژنتیک و روش حل معادلات نرمال (SNE) به‌صورت چندهدفی، منجر به نتایج خوبی در داده‌های آموزش دیده و همچنین در پیش‌بینی رفتار داده‌های آموزش ندیده، می‌شود.

در مجموع، مدل ANFIS کارایی بهتری برای پیش‌بینی میزان مصرف منعقدکننده پلی آلومینیوم کلراید دارد.

۵- قدردانی

از مسئولان محترم شرکت آب و فاضلاب گیلان و تصفیه‌خانه بزرگ آب شهر رشت که صمیمانه در انجام هر چه بهتر این پروژه یاری رساندند، تشکر و قدردانی به‌عمل می‌آید.

۶- مراجع

1. Maniezzo, V. (1994). “Genetic evolution of topology and weight distribution of neural networks.” *IEEE. Trans. On Neural Networks*, 5(1), 39-53.
2. Bratby, J. (2006). *Coagulation and flocculation in water and wastewater treatment*, IWA Pub., USA.

3. Hu, C., Lu, H., Qu, J., Wang, D., and Rut, J. (2006). "Coagulation behavior of aluminum salts in eutrophic water: Significance of Al13 species and pH control." *Environmental Science and Technology*, 40(1), 325-331.
4. Dentel, S.K. (1991). "Coagulant control in water treatment." *Critical Reviews in Environmental Science and Technology*, 21(1), 41-135.
5. Edzwald, J. (1993). "Coagulation in drinking water treatment: Particles, organics and coagulants." *Water Science and Technology*, 27(11), 21-35.
6. Hanson, A.T., and Cleasby, J.L. (1990). "The effects of temperature on turbulent flocculation: Fluid dynamics and chemistry." *J. American Water Works Association*, 82(11), 56-73.
7. Gregor, J., Nokes, C., and Fenton, E. (1997). "Optimising natural organic matter removal from low turbidity waters by controlled pH adjustment of aluminium coagulation." *Water Research*, 31(12), 2949-2958.
8. Juntunen, P., Lu, K., Konen, M., Pelo, M., Lehtola, M.J., and Hilton, Y. (2012). "Modeling of water quality: An application to a water treatment process." *Applied Computational Intelligence and Soft Computing*, Article ID: 846321.
9. Heddam, S., Bermad, A., and Dechemi, N. (2012). "ANFIS-based modelling for coagulant dosage in drinking water treatment plant: A case study." *Environmental Monitoring and Assessment*, 184 (4), 1953-1971.
10. Yu, R.F., Kang, S.F., Liaw, S.L., and Chen, M.C. (2000). "Application of artificial neural network to control the coagulant dosing in water treatment plant." *Water Science and Technology*, 42 (3-4), 403-408.
11. Shariff, R. (2007). *Real-time artificial intelligence control and optimization of a full scale WTP*, American Water Works Research Foundation.
12. Bazer-Bachi, A., Puech-coste, R., and Probst, J.L. (1990). "Mathematical modelling of optimal coagulant dose in water treatment plant." *Rev. Sci Edu*., 3 (4), 377-397. (In Persian)
13. Gagnon, C., Grandjean, B., and Thibault, J. (1997). "Modelling of coagulant dosage in a water treatment plant." *Artificial Intelligence in Engineering*, 11(4), 401-404.
14. Leeuwen, V. (1999). "Empirical mathematical models and artificial neural networks for the determination of alum doses for treatment of southern Australian surface waters." *Aqua*, 48(3), 115-127.
15. Maier, H.R., Morgan, N., and Chow, C.W.K. (2004). "Use of artificial neural networks for predicting optimal alum doses and treated water quality parameters." *Environmental Modelling and Software*, 19(5), 485-494.
16. Wu, G.-D., and Lo, S.-L. (2008). "Predicting real-time coagulant dosage in water treatment by artificial neural networks and adaptive network-based fuzzy inference system." *Engineering Applications of Artificial Intelligence*, 21(8), 1189-1195.
17. Jamali, A., Narmian-zadeh, N., Ashraf, H., and Jamali, Z. (2008). "Robust Pareto design of ANFIS networks for nonlinear systems with probabilistic uncertainties." *IEEE*, *International Sysmposium, Istanbul*, 300-304.
18. Nariman-Zadeh, N., and Jamali, A. (2007). "Pareto genetic design of GMDH-type neural networks for nonlinear systems." *Proceeding of IWIM*, University of Guilan, Iran
19. Zadeh, L.A. (1965). "Fuzzy sets." *Information and Control*, 8(3), 338-353.
20. Jang, J.S.R. (1993). "ANFIS: Adaptive-network-based fuzzy inference system." *Systems, Man and Cybernetics, IEEE Transactions,* 23(3), 665-685.
21. Ivakhnenko, A., and Ivakhnenko, G. (1995). "The review of problems solvable by algorithms of the group method of data handling (GMDH)." *Pattern Recognition and Image Analysis*, 5, 527-535.