

# یک سیستم استنتاج تطبیقی عصبی-فازی به منظور کنترل آماری فرآیند داده‌های خود همبسته

محمد رضا وکیلی\*

دانشگاه آزاد اسلامی، واحد اردبیل، گروه مهندسی صنایع، اردبیل، ایران

عباس سقایی

دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران، گروه مهندسی صنایع، تهران، ایران

امین محمودی

مرکز آموزش عالی فنی و مهندسی بوئین زهرا، گروه مهندسی صنایع، قزوین، ایران

**چکیده** نمودارهای کنترل سنتی مبتنی بر این فرض اساسی هستند که داده‌های فرآیند بطور متوالی مستقل از یکدیگر بوده و دارای توزیع نرمال می‌باشند. این در حالی است که در بسیاری از موارد در دنیای واقعی از جمله در فرآیندهای شیمیایی و پیوسته این فرض اساسی وجود ندارد و بین داده‌های جمع‌آوری شده از فرآیند نوعی خودهمبستگی وجود دارد. استفاده از نمودارهای کنترل سنتی در فرآیندهای خود همبسته غیرقابل اطمینان بوده و باعث افزایش اخطارهای اشتباه می‌شود. یکی از روش‌های توسعه داده شده به منظور کنترل فرآیندهای خودهمبسته، شناسایی ساختار سری‌های زمانی فرآیند و استفاده از مقادیر باقیمانده‌ها به منظور کنترل فرآیند است. در این مقاله از یک مدل مبتنی بر سیستم‌های تطبیقی عصبی-فازی به منظور شناسایی ساختار سری‌های زمانی و پیش‌بینی استفاده شده است. همچنین نمودارهای کنترل باقیمانده مبتنی بر این سیستم برای داده‌های خود برگشتی درجه ۲  $AR(2)$  طراحی می‌شود. در نهایت با استفاده از داده‌های شبیه‌سازی شده، کارایی روش پیشنهادی در نمودار میانگین متحرک موزون نمایی و برای درجه‌های مختلفی از همبستگی مورد ارزیابی قرار گرفته و نشان داده می‌شود که روش پیشنهادی برای داده‌های با همبستگی زیاد دارای کارایی بسیار مناسب است.

**کلمات کلیدی** کنترل فرآیند آماری، نمودارهای کنترل، سیستم استنتاج عصبی-فازی تطبیقی، داده‌های خود همبسته

## ۱- مقدمه

می‌گیرند. به هنگام استفاده از نمودارهای کنترل استاندارد دو فرض اساسی در نظر گرفته می‌شود بطوریکه داده‌ها تولید شده به وسیله یک فرآیند تحت کنترل بطور مستقل از توزیع نرمال با میانگین  $\mu$  و انحراف معیار  $\sigma$  پیروی می‌کنند. در این حالت می‌توان گفت مشخصه کیفی در زمان  $t$  بصورت  $x_t = \mu + \varepsilon_t$  مدل می‌شود. در رابطه بیان شده  $\varepsilon_t$  از توزیع نرمال مستقل با میانگین صفر و انحراف معیار  $\sigma$  پیروی می‌کند. در این حالت می‌توان از نمودارهای کنترل معمول برای ارزیابی شرایط فرآیند استفاده کرد. یکی از مفروضات مهم، مستقل بودن مشاهدات

تکنیک‌های کنترل کیفیت آماری<sup>۱</sup> (SPC) بطور گسترده در مراکز تولیدی به منظور کنترل میانگین و واریانس مشخصه کیفی محصولات مورد استفاده قرار می‌گیرد. در علم SPC از سیگنال‌های آماری به منظور کشف نوسانات در فرآیند تولید، شناسایی منبع و منشاء تغییرات، بهبود در کارایی و حفظ فرآیند در سطح کیفی مورد نیاز استفاده می‌شود. یکی از ابزارهای پر کاربرد در SPC نمودارهای کنترل می‌باشند که توسط شوارت در دهه ۱۹۲۰ مطرح شدند. نمودارهای کنترل به منظور ابزاری برای بررسی تحت کنترل بودن فرآیند تولید مورد استفاده قرار

\* (Corresponding author) mreza vakili@iauardabil.ac.ir

تاریخ پذیرش: ۱۳۹۳/۱۱/۲۴

تاریخ دریافت: ۱۳۹۳/۹/۲

از تکنیک‌های کلاسیک آماری همانند کمترین مربعات خطا و یا حداکثر درست‌نمایی انجام می‌شد. اگر چه این روش‌ها بسیار در ادبیات مورد استفاده قرار گرفته است ولی باید توجه داشت که معمولاً فرضیات آنها ارضاء نمی‌شود. این موضوع در بسیاری از موقعیت‌های واقعی باعث ضعف کارایی نمودارهای کنترل می‌شود. همانگونه که قبلاً اشاره شده نمودارهای کنترل باقیمانده‌ها بر مبنای این حقیقت استوار شده‌اند که، زمانیکه فرآیند تحت کنترل است داده‌ها دارای ساختار مشخصی هستند و این در حالی است که در موقعیت خارج از کنترل، بدون اصلاح ساختار اولیه مقدار خطای پیش‌بینی افزایش می‌یابد. بنابراین، به منظور اجرای موثرتر نمودارهای باقیمانده‌ها، آموزش داده‌ها بدون محدودیت تبعیت از فرم معینی همچون روش‌های سری زمانی خطی بسیار مهم است. به همین دلیل استفاده از روش‌های متفاوت که فرضیات محدودکننده کمتر و انعطاف‌پذیری بیشتری داشته و در موقعیت‌های دنیای واقعی توانایی بیشتری داشته باشند، در سال‌های اخیر بیشتر رواج پیدا کرده‌اند.

برخی از این تکنیک‌ها، روش‌های آموزش ماشین<sup>۱۱</sup> همانند شبکه‌های عصبی مصنوعی (ANN<sup>۱۲</sup>) و ماشین‌های بردار پشتیبان (SVM<sup>۱۳</sup>) می‌باشند. این تکنیک‌ها آموزش مستقیم داده‌ها را بدون در نظرگیری فرض محدودکننده‌ای برای ساختار داده‌ها امکان پذیر می‌سازند. دولی و گو<sup>۱۴</sup> [۱۱] مدل‌های شبکه عصبی را به منظور کشف تغییرات مثبت در میانگین و واریانس داده‌ها اجرا کردند. هوآرنگ و هوبل<sup>۱۵</sup> [۱۲] چندین مدل شبکه‌های عصبی را به منظور ساخت نمودار کنترل شوهارت  $\bar{x}$  و به منظور شناسایی موقعیت‌های خارج از کنترل یکپارچه کردند. چنگ و او<sup>۱۶</sup> [۱۳] یک شبکه تطبیقی فازی-عصبی را منظور شناسایی و طبقه‌بندی تغییرات میانگین در حالیکه که داده‌ها را مستقل از هم در نظر گرفتند، پیشنهاد دادند.

ایسام و موحامد<sup>۱۷</sup> [۱۴] از روش رگرسیون بردار پشتیبان<sup>۱۸</sup> برای رسم نمودار کنترل باقیمانده‌های چندمتغیره و به منظور کنترل یک فرآیند خودهمبسته استفاده کرد. همچنین ایسام و همکاران [۱۵] از روش رگرسیون بردار پشتیبان به منظور کنترل فرآیند خودهمبسته چندمتغیره غیرخطی مورد استفاده قرار دادند. ارکات<sup>۱۹</sup> و همکاران [۱۶] از شبکه‌های عصبی به منظور کنترل یک فرآیند چندمتغیره که دارای همبستگی بود استفاده کردند.

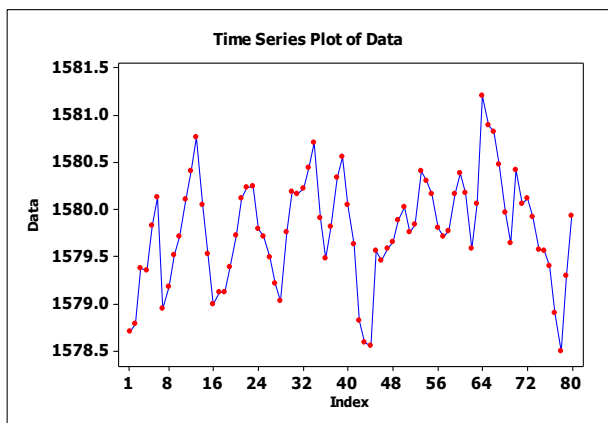
آنها از این ابزار به منظور پیش‌بینی و سپس با استفاده از باقیمانده‌ها برای رسم نمودار  $MCUSUM_{20}$  بهره بردند. سلیمانی و نورالسنا<sup>۱۷</sup> [۱۷] به پایش پروفایل خطی با در نظر

است. در صورت وجود همبستگی بین مشاهدات در طول زمان نمودارهای کنترل معمولی کارایی خود را از دست می‌دهند و منجر به انعکاس تعداد زیادی زنگ خطر اشتباه شده و باعث گمراهی مهندسان کیفی می‌شوند. در برخی از فرآیندهای تولیدی، مخصوصاً فریندهای تولید پیوسته مانند فرآیندهای شیمیایی فرض استقلال بین داده‌های جمع‌آوری شده وجود ندارد و عمدتاً بین نمونه‌های گرفته شده از خط تولید نوعی همبستگی وجود دارد که نباید آن را نادیده گرفت.

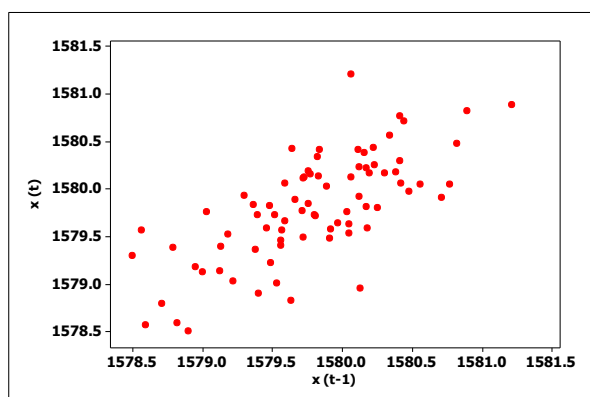
از جمله اولین کسانی که کنترل کیفیت آماری را با فرض وجود همبستگی بین داده‌ها در واحد زمانی در نظر گرفتند می‌توان به تحقیقات الوان و رابرتس<sup>۲</sup> [۱] و همچنین واردل<sup>۳</sup> و همکاران [۲] اشاره کرد. روش‌های کنترل داده‌های خود همبسته شامل دو رویکرد مبتنی بر مدل و مدل آزاد است. رانگر و ویلیمین<sup>۴</sup> [۳] در یک فرآیند (۱)  $AR^5$  روش‌های نمودارهای کنترل باقیمانده‌ها، میانگین‌های دسته‌ای موزون و ناموزون را مقایسه قرار دادند. از جمله رایج‌ترین روش‌ها برای کنترل داده‌های خودهمبسته روش هانتر<sup>۶</sup> [۴] است. در روش او، پارامترهای ساختار سری زمانی تخمین زده شده و سپس با استفاده از باقیمانده که اختلاف بین مقادیر واقعی و پیش‌بینی شده هستند اقدام به کنترل فرآیند می‌شود. در حقیقت از آنجائیکه باقیمانده‌ها بطور متوالی از یکدیگر مستقل هستند، بنابراین، می‌توان از نمودارهای کنترل استاندارد برای باقیمانده‌ها استفاده کرد. نورالسنا و واقفی<sup>۷</sup> [۵] نشان دادند که عدم وجود شرط استقلال، بر روی متوسط طول دنباله (ARL) نمودارهای کنترل تاثیر می‌گذارد و آنها را غیر واقعی می‌کند. همچنین فرآیند تولید با داده‌های که بطور متوالی همبسته هستند موجب سیگنال‌های ناصحیح شده و کارایی یک نمودار کنترل از بین می‌رود [۷و۸]. اورلاندو<sup>۸</sup> و همکاران [۸] یک نمودار جمع تجمعی تک‌متغیره بر مبنای باقیمانده‌ها را برای داده‌های خودهمبسته پیشنهاد دادند. لوردو<sup>۹</sup> و همکاران [۹] یک رویه تنظیم رگرسیون را به منظور کنترل فرآیند خودهمبسته ارائه دادند و نشان دادند که نمودارهای کنترل باقیمانده‌ها در کشف انحرافات میانگین بهتر عمل می‌کنند. همچنین سنوسی<sup>۱۰</sup> و همکاران [۱۰] نشان دادند که نمودارهای کنترل باقیمانده‌ها توانایی بهتری در کشف انحرافات میانگین دارند.

در اکثر تحقیقات اولیه در زمینه کنترل فرآیند خود همبسته، تخمین پارامترهای سری زمانی و محاسبه باقیمانده‌ها با استفاده

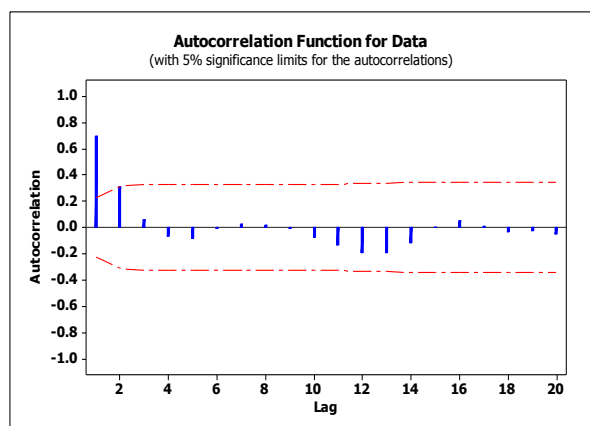
به شدت همبسته است. به منظور نمایش چگونگی اثر همبستگی بین داده‌ها از نمودار مشاهدات انفرادی برای ارزیابی فرآیندی که داده‌های آن در شکل ۱ نشان داده شد، استفاده می‌شود. نمودار کنترل مشاهدات انفرادی بصورت شکل ۴ است.



شکل ۱: یک فرآیند با داده‌های خود همبسته



شکل ۲: نمودار پراکندگی  $x_t$  نسبت به  $x_{t-1}$



شکل ۳: تابع خودهمبستگی مشاهدات برای نمونه‌های شکل ۱

گیری خودهمبستگی بین داده‌ها پرداختند. فرآنکو و همکاران [۱۸] یک روش نمونه‌گیری جدید را به منظور کاهش تاثیر خودهمبستگی داده‌ها روی نمودارهای کنترل ارائه کردند. دو و لو<sup>۲۱</sup> [۱۹] نمودار حداقل فاصله اقلیدسی<sup>۲۲</sup> (MED) را بر مبنای رگرسیون بردار پشتبان<sup>۲۳</sup> به منظور کنترل یک فرآیند خود همبسته ارائه کردند. فرآنکو و همکاران [۲۰] یک طراحی اقتصادی نمودارهای کنترل شوهارت را به منظور کنترل یک فرآیند خودهمبسته با در نظرگیری نمونه گیری skip را ارائه کردند. با توجه به مرور ادبیات بیان شده، مشاهده می‌شود که یکی از قدرتمندترین ابزارهای پیش بینی و مدل سازی سری‌های زمانی بنام سیستم‌های تطبیقی استنتاج عصبی-فازی<sup>۲۴</sup> (ANFIS) مورد استفاده قرار نگرفته است. بنابراین، در این تحقیق نحوه استفاده از ابزار پیش بینی ANFIS به منظور رسم نمودارهای کنترل باقیمانده‌ها در فرآیندهای دارای خودهمبستگی مورد بررسی قرار خواهد گرفت. همچنین کارایی روش پیشنهادی با شبیه‌سازی ارزیابی شده است.

## ۲- بیان مساله

همانگونه که قبلاً اشاره شده در دنیای واقعی و در بسیاری از فرآیندهای تولیدی، مخصوصاً فرایندهای تولیدی پیوسته مانند فرآیندهای شیمیایی عمدتاً بین نمونه‌های گرفته شده از خط تولید نوعی همبستگی وجود دارد و فرض استقلال بین داده‌های جمع‌آوری شده وجود ندارد. حال در ادامه می‌خواهیم تاثیر وجود همبستگی بین داده‌ها را با یک مثال مورد بررسی قرار دهیم.

شکل ۱ نشان دهنده نمونه‌های گرفته شده از دمای حرارت یک کوره سرمیک‌سازی است که باید مورد کنترل قرار گیرد. بر اساس بیسگارد و کولاچی<sup>۲۵</sup> [۲۱] نمونه‌های گرفته شده از یک فرآیند تولیدی با داده‌های خودهمبسته و در موقعیت تحت کنترل است که از طرف گروه مهندسی به عنوان نمونه‌های مینا در حالت کارایی خوب فرآیند مورد استفاده قرار می‌گیرند. به منظور به تصویر کشاندن وجود همبستگی بین داده‌ها در شکل ۲ نمودار پراکندگی  $x_t$  (مشاهدات در زمان  $t$ ) نسبت به  $x_{t-1}$  (مشاهدات در زمان  $t-1$ ) را نشان می‌دهد. همانطور که ملاحظه می‌شود هر مشاهده نسبت به مشاهده قبلی خود دارای یک همبستگی مثبت است. به منظور محاسبه میزان ضریب همبستگی هر مشاهده نسبت به مشاهدات قبلی خود، نرم افزار Minitab مورد استفاده قرار گرفته و نمودار مربوط به ضرایب همبستگی در تاخیرهای مختلف به صورت شکل ۳ است. همانطوریکه در شکل ۳ مشاهده می‌شود هر مشاهده نسبت به دو مشاهده ما قبل خود

است. در حالت کلی داده‌های مربوط به یک مدل AR(P) به صورت رابطه ۱ مدل می‌شوند.

$$x_t = a + \rho_1 x_{t-1} + \rho_2 x_{t-2} + \dots + \rho_p x_{t-p} + \varepsilon_t \quad (1)$$

که مشابه یک مدل رگرسیون چندمتغیره است. در مدل فوق  $a$  و  $\rho$  ها مقادیر ثابت و نامعلوم هستند که باید به منظور پیش‌بینی داده‌ها تخمین زده شوند. همچنین  $\varepsilon_t$  یک عبارت تصادفی با توزیع نرمال با میانگین صفر و انحراف معیار ثابت  $\sigma$  است. حال می‌توان با هدف کمینه کردن میانگین مربعات خطا، تخمین پارامترهای  $a$  و  $\rho_i$  را بدست آورد. پس از تخمین پارامترهای رگرسیونی می‌توان مقادیر پیش‌بینی داده‌ها را ( $\hat{x}_t$ ) را بدست آورد. بنابراین، مقادیر باقیمانده‌ها بصورت رابطه ۲ محاسبه می‌شود.

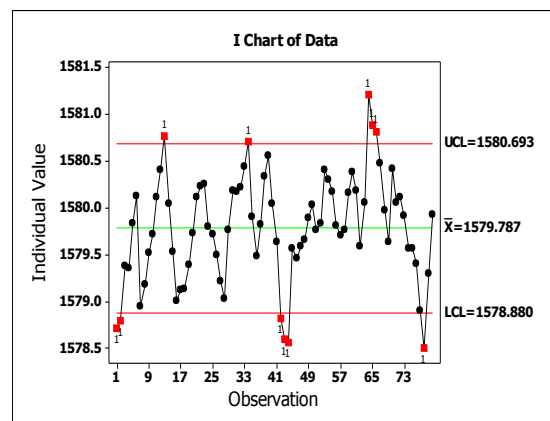
$$e_t = x_t - \hat{x}_t \quad (2)$$

در موقعیت‌هایی که فرآیند تحت کنترل است، رفتار میانگین باقیمانده‌ها باید بطور تصادفی و در اطراف صفر باشد. درحالی‌که برای موقعیت‌هایی که فرآیند خارج از کنترل باشد، تغییر در میانگین فرآیند باعث یک تغییر در میانگین باقیمانده‌ها می‌شود. بدین ترتیب می‌توان با استفاده از نمودارهای کنترل مبتنی بر باقیمانده‌ها، یک فرآیند تولید خودهمبسته را مورد کنترل و پایش قرار داد. ما در این تحقیق به دنبال این هستیم تا با استفاده از یکی از قدرتمندترین ابزارهای پیش‌بینی یعنی سیستم‌های استنتاج تطبیقی عصبی-فازی (ANFIS) و با رسم نمودارهای کنترل باقیمانده‌ها یک فرآیند خودهمبسته را مورد کنترل و پایش قرار دهیم.

## ۲-۲) رویکرد مدل آزاد

در این رویکرد، به جای استفاده از اصل داده‌های خودهمبسته برای کنترل فرآیند، از میانگین دسته‌ای داده‌ها استفاده می‌شود. علت این کار در این است که با میانگین‌گیری از یک سری داده خودهمبسته، همبستگی بین آنها از بین رفته تقریباً مستقل از یکدیگر می‌شوند [۲۲]. یک از پارامترهای مهم و تاثیرگذار در این روش‌ها تعیین اندازه هر دسته است.

معمولاً روش‌هایی که در این رویکرد وجود دارند دو دسته‌اند (الف) میانگین‌گیری بصورت ساده<sup>۲۰</sup> و (ب) میانگین‌گیری موزون<sup>۲۱</sup>.



شکل ۴: نمودار کنترل مشاهدات انفرادی برای داده‌های شکل ۱

همانطور که ملاحظه می‌شود در چندین نقطه فرآیند خارج از کنترل اعلام می‌شود و این در حالی است که نمونه‌ها تحت کنترل بوده و در بهترین حالت خط تولید جمع‌آوری شده‌اند. این هشدارهای اشتباه به دلیل نادیده گرفتن وجود همبستگی در داده‌ها و استفاده از نمودارهای کنترل معمول برای داده‌های دارای خودهمبستگی است. دو رویکرد عمده برای مقابله با همبستگی بین داده‌ها و کنترل فرآیندهای خود همبسته وجود دارد. (۱) رویکرد مبتنی بر مدل<sup>۲۶</sup> و (۲) رویکرد مدل آزاد<sup>۲۷</sup>. در ادامه مختصراً در رابطه با این دو رویکرد توضیح داده می‌شود.

## ۱-۲) رویکرد مبتنی بر مدل

یکی از روش‌هایی که برای ارزیابی داده‌های خود همبسته معمولاً مورد استفاده قرار می‌گیرد، استفاده از نمودارهای کنترل برای پایش باقیمانده‌ها است. این دسته از روش‌ها بدین گونه عمل می‌کنند که ابتدا با استفاده از یک مدل سری زمانی مقدار هر یک از مشاهدات با استفاده از مشاهدات قبلی پیش‌بینی می‌شوند و سپس از اختلاف بین مقادیر واقعی و پیش‌بینی مقادیر باقیمانده‌ها<sup>۲۸</sup> بدست می‌آیند [۷و۸]. در نهایت این مقادیر باقیمانده‌ها برای رسم نمودار کنترل و ارزیابی فرآیند استفاده می‌شود. استفاده از یک مدل سری زمانی مناسب با خطایی کمتر به منظور پیش‌بینی در استفاده از این روش‌ها بسیار پراهمیت است. مدلی بهتر است که کمترین میزان خطا را داشته باشد و همچنین کمترین مقدار همبستگی بین باقیمانده‌ها را داشته باشد. هرگاه هر داده برای پیش‌بینی به  $p$  مشاهد قبلی وابسته باشد به آن مدل خودبرگشتی<sup>۲۹</sup> درجه  $p$  گفته می‌شود (AR(P)). همانگونه که در شکل ۳ مشاهده می‌شود، بیشترین همبستگی بین مشاهدات با انحراف زمانی یک و دو است و یا به عبارت دیگر هر مشاهده به دو مشاهد قبلی خود وابسته است یعنی مدل مورد نظر AR(۲)

مختلف است که با یکدیگر ارتباط دارند. خروجی این شبکه وابسته به پارامترهای قابل تنظیم این گره‌ها است. قواعد یادگیری شبکه تعیین می‌کند که چطور این پارامترها به منظور کمینه کردن خطا به روز شوند. یک سیستم استنتاج فازی یک چارچوب بر مبنای تئوری فازی و قواعد اگر-آنگاه فازی است. ساختار یک سیستم استنتاج فازی شامل سه جزء اصلی است:

(۱) پایگاه قواعد، (۲) پایگاه داده و (۳) یک مکانیزم استدلال. یک پایگاه قواعد فازی شامل قواعد اگر-آنگاه فازی است. به عنوان مثال، یک قاعده می‌تواند "اگر قیمت کم باشد آنگاه تقاضا زیاد می‌شود" باشد بطوریکه کم و زیاد در این قاعده متغیرهای کلامی هستند. پایگاه داده، توابع عضویت بکاربرده شده در قواعد فازی و همچنین مکانیزم استدلال، رویه استنتاج خروجی از متغیرهای ورودی را اجرا می‌کند [۲۴]. فرض کنید سیستم فازی دارای دو متغیر  $x$  و  $y$  ورودی و یک متغیر خروجی  $z$  باشد. به علاوه پایگاه قواعد شامل دو قاعده اگر آنگاه فازی مشابه قواعد شرح داده شده توسط تاکاجی و سوچنو<sup>۳۵</sup> [۲۵] باشد (رابطه ۵):

$$\text{Rule 1: If } X \text{ is } A_1 \text{ and } Y \text{ is } B_1 \Rightarrow f_1 = p_1x + q_1y + r_1 \quad (5)$$

$$\text{Rule 2: If } X \text{ is } A_2 \text{ and } Y \text{ is } B_2 \Rightarrow f_2 = p_2x + q_2y + r_2$$

بطوریکه اگر  $f(x, y)$  یک چند جمله‌ای درجه یک همانند بالا باشد آنگاه مدل به عنوان مدل فازی درجه یک سوچنو<sup>۳۶</sup> نامیده می‌شود. معماری ANFIS در شکل ۵ نشان داده شده است بطوریکه نوع توابع بکار برده شده برای هر گره در یک لایه مشابه می‌باشند.  $O_{1,i}$  خروجی آامین گره از آامین لایه است. خروجی هر لایه می‌تواند بصورت زیر نشان داده شود.

لایه ۱: گره‌ها در این لایه، گره‌های تطبیقی هستند. خروجی هر گره می‌تواند همانند رابطه ۶ نشان داده شود.

$$O_{1,i} = \mu_{A_i}(x) \quad (6)$$

بطوریکه  $x$  مقدار ورودی به گره  $i$ ،  $A_i$  برچسب متغیر کلامی مربوطه و  $O_{1,i}$  تابع عضویت متغیر کلامی  $A_i$  می‌باشند. تابع عضویت می‌تواند یکی از توابع عضویت معروف که در جدول ۱ آمده است باشد: بطوریکه  $\sigma_1, c_1, b_1, a_1$  مجموعه پارامترها هستند که شکل تابع با تغییر آنها تغییر خواهد کرد و همچنین این پارامترها در طول الگوریتم مورد آموزش قرار می‌گیرند.

(الف) در میانگین‌گیری ساده ابتدا داده‌های خود همبسته به دسته‌های با اندازه  $b$  تقسیم می‌شوند سپس به هنگام میانگین‌گیری در هر دسته به هر داده وزن یکسان تعلق می‌گیرد. در این صورت میانگین دسته  $j$  بصورت رابطه ۳ می‌شود:

$$\bar{x}_j = \frac{1}{b} \sum_{i=1}^b x_{(j-1)b+i}, j = 1, 2, \dots \quad (3)$$

میانگین‌های دسته‌ای را می‌توان بر روی یک نمودار کنترل مشاهدات انفرادی استاندارد رسم و تحلیل کرد. تعیین مقدار پارامتر  $b$  با دو هدف کاهش همبستگی به میانگین‌ها و کسب یک مقدار قابل قبول از  $ARL^{33}$  انجام می‌شود.

(ب) در میانگین‌گیری موزون نیز داده‌ها به دسته‌های با اندازه  $b$  تقسیم می‌شوند با این تفاوت که به هنگام میانگین‌گیری در هر دسته وزن همه داده‌ها یکسان در نظر گرفته نمی‌شوند. یک از روش‌های پیشنهادی در این زمینه بصورت رابطه ۴ است که در آن  $b$  اندازه هر دسته و  $\phi$  اندازه همبستگی بین داده‌ها است [۳].

$$\bar{x}_j = \sum_{i=1}^b w_i x_{(j-1)b+i}, j = 1, 2, \dots \quad (4)$$

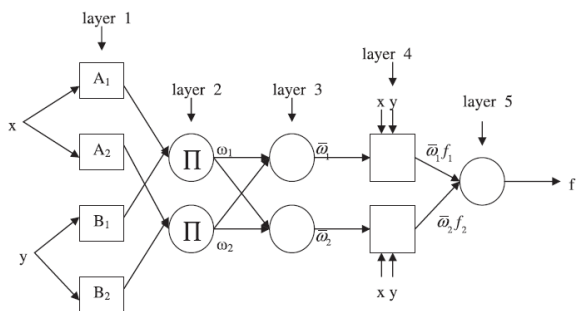
$$w_i = \frac{-\phi}{(b-1)(1-\phi)}, w_i = \frac{1}{b-1}, i = 2, \dots, w_b = \frac{1}{(b-1)(1-\phi)}$$

روش میانگین‌های دسته‌ای زمانی می‌تواند مفید باشد که داده‌ها به طور مرتب جمع‌آوری شوند. در اغلب کارخانه‌های شیمیایی و فرآیندهای پیوسته برخی از داده‌های فرآیند هر چند ثانیه مشاهده و ثبت می‌شوند. بدیهی است تحت این شرایط، میانگین‌های دسته‌ای می‌تواند کاربرد بالقوه زیادی داشته باشند.

### ۳- روش پیشنهادی: سیستم‌های استنتاج عصبی-فازی تطبیقی

#### ۳-۱- ساختار سیستم‌های استنتاج عصبی-فازی تطبیقی

سیستم‌های استنتاج تطبیقی عصبی-فازی (ANFIS) برای اولین بار توسط جنگ<sup>۳۳</sup> [۲۳] با آموزش یک سیستم استنتاج فازی (FIS<sup>۳۴</sup>) در چارچوب یک شبکه عصبی تطبیقی معرفی شد. ساختار یک شبکه عصبی تطبیقی شامل تعدادی گره در لایه‌های



شکل ۵: معماری شبکه تطبیقی ANFIS

$$o_{5,i} = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (10)$$

الگوریتم ترکیبی یادگیری ANFIS:

الگوریتم یادگیری به منظور تنظیم همه پارامترهای قابل تنظیم (پارامترهای تابع هدف و پارامترهای نتیجه‌گیری) و با هدف بدست آوردن مقادیر پارامتر خروجی از سیستم ANFIS که با داده‌های آموزش مطابقت داده می‌شوند مورد استفاده قرار می‌گیرند. فرض کنید پارامترهای تابع هدف ثابت باشند [۲۳]. خروجی ANFIS می‌تواند به صورت رابطه ۱۱ نمایش داده شود.

$$f = \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2 \quad (11)$$

$$= \bar{w}_1(p_1x + q_1y + r_1) + \bar{w}_2(p_2x + q_2y + r_2)$$

$$= (\bar{w}_1x)p_1 + (\bar{w}_1y)q_1 + (\bar{w}_1)r_1 + (\bar{w}_2x)p_2 + (\bar{w}_2y)q_2 + (\bar{w}_2)r_2$$

بطوریکه یک ترکیب خطی از پارامترهای نتیکه‌گیری  $p_i, q_i, r_i$  با آنجائیکه خروجی سیستم نسبت به پارامترها خطی است، روش کمترین مربعات با هدف مینیمم کردن خطای پیش‌بینی مورد استفاده قرار می‌گیرد. الگوریتم ترکیبی یادگیری در هر تکرار شامل دو مسیر است، مسیر رفت و مسیر برگشت [۲۳، ۲۶].

(۱) در مسیر رفت بردار متغیرهای ورودی ارائه شده و مقدار گره خروجی هر لایه تا لایه ۴ محاسبه می‌شود. پارامترهای نتیجه‌گیری با استفاده از روش کمترین مربعات تعیین می‌شوند. بعد از تعیین پارامترهای قسمت نتیجه‌گیری با روش کمترین مربعات خطا تعیین شده و سپس مقدار خطا برای همه داده‌های آموزشی محاسبه می‌شوند.

جدول ۱: انواع توابع عضویت مورد استفاده در ANFIS

نام تابع عضویت	تابع عضویت
تابع زنگوله ای شکل	$\mu_{A_i}(x) = \frac{1}{1 + \left  \frac{x - c_i}{a_i} \right ^{2b_i}}$
تابع گوسی	$\mu_{A_i}(x) = e^{-\frac{1}{2} \left( \frac{x - c_i}{\sigma_i} \right)^2}$
تابع دوزنقه‌ای	$\mu_{A_i}(x) = \begin{cases} 0 & , x \leq a_i \\ \frac{x - a_i}{b_i - a_i} & , a_i \leq x \leq b_i \\ 1 & , b_i \leq x \leq c_i \\ \frac{d_i - x}{d_i - c_i} & , c_i \leq x \leq d_i \\ 0 & , d_i \leq x \end{cases}$
تابع مثلثی	$\mu_{A_i}(x) = \begin{cases} 0 & , x \leq a_i \\ \frac{x - a_i}{b_i - a_i} & , a_i \leq x \leq b_i \\ \frac{c_i - x}{c_i - b_i} & , b_i \leq x \leq c_i \\ 0 & , c_i \leq x \end{cases}$

لایه ۲: مقدار هر گره در این لایه ثابت بوده و نشان دهنده شدت خروجی قاعده نام است. خروجی این لایه حاصل ضرب سیگنال‌های همه ورودی‌ها بوده و بصورت رابطه ۷ می‌باشند.

$$o_{\tau,i} = w_i = \mu_{A_i}(x) \cdot \mu_{B_i}(y), \quad i = 1, 2, \dots \quad (7)$$

لایه ۳: هر گره در این لایه ثابت بوده و نسبت شدت نامین قاعده را محاسبه می‌کند و بصورت رابطه ۸ است.

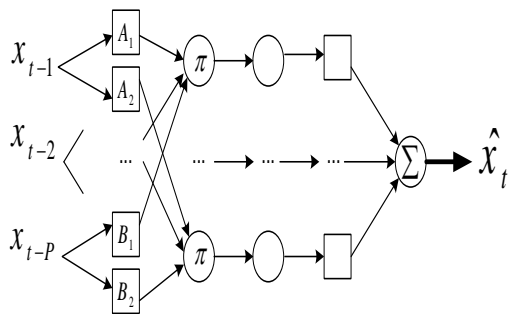
$$o_{\tau,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2, \dots \quad (8)$$

لایه ۴: گره‌ها در این لایه اجرا کننده خروجی هر قاعده هستند و بصورت رابطه ۹ می‌باشند.

$$o_{4,i} = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad (9)$$

پارامترهای  $p_i, q_i, r_i$  به عنوان پارامترهای بخش نتیجه هستند و در طی مرحله آموزش به روز رسانی می‌شوند.

لایه ۵: تک گره موجود در این لایه یک گره ثابت است که مقدار نهایی پارامتر خروجی را به عنوان مجموعی از سیگنال‌های ورودی نشان می‌دهد و همانند رابطه ۱۰ محاسبه می‌شوند.



شکل ۶: معماری شبکه روش پیشنهادی

از فرآیند مورد کنترل باشند. در ادامه به منظور تشریح بیشتر روش پیشنهادی مثال کنترل دمای حرارت کوره سرامیک که در حالت ایده‌آل و تحت کنترل فرآیند تولید جمع‌آوری شده بودند (داده‌های شکل ۱) در نظر گرفته می‌شود. همانگونه که قبلاً مورد بررسی قرار گرفت از آنجائیکه این داده‌ها دارای خود همبستگی هستند بنابراین، استفاده از نمودارهای کنترل معمول فرآیند را خارج از کنترل نشان داد با اینکه فرآیند تحت کنترل است و داده‌ها از بهترین وضعیت خط تولید جمع‌آوری شده‌اند.

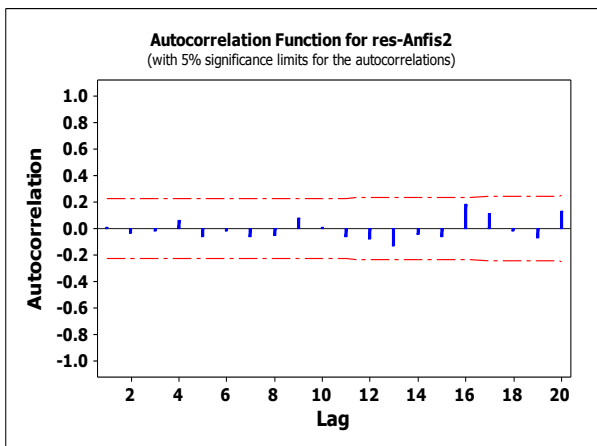
در مثال مطرح شده از مشاهدات دارای خود همبستگی، هر مشاهده به دو مشاهده قبلی خود وابسته است. بنابراین، به منظور استفاده از ANFIS به عنوان یک مدل سری زمانی به منظور پیش‌بینی، متغیرهای ورودی  $x_{t-1}$  و  $x_{t-2}$  است. و بنابراین،  $\hat{x}_t$  به عنوان پارامتر خروجی ANFIS و به عنوان مقدار پیش‌بینی مشاهده  $x_t$  در نظر گرفته می‌شود. شکل ۷ ساختار شبکه مربوط به مثال مورد بررسی در نرم افزار متلب نشان داده شده است. پس از پیش‌بینی هر مشاهده، مقادیر باقیمانده برای هر مشاهده محاسبه می‌شود. نمودار پراکندگی باقیمانده‌ها در شکل ۹ نشان داده شده است. پس از محاسبه باقیمانده‌ها به منظور بررسی کارایی مدل ANFIS، مقادیر همبستگی باقیمانده‌ها محاسبه می‌شود. همانگونه که در شکل ۱۰ نشان داده شده است، مقدار همبستگی نسبت به سه مشاهده قبلی بسیار کاهش یافته و حتی نزدیک به صفر است. پس از ذخیره کردن سیستم فازی آموزش دیده، مقادیر هر یک از مشاهدات را با استفاده از دو مشاهده قبلی خود مورد پیش‌بینی قرار می‌دهیم. در شکل ۸ مقایسه بین مقادیر واقعی و پیش‌بینی شده برای مشاهدات شکل ۱ نشان داده شده است. حال با توجه به اینکه استقلال بین داده‌ها برقرار شده است بنابراین، می‌توان از نمودارهای کنترل برای ارزیابی باقیمانده‌ها

(۲) در مسیر برگشت، خطاهای خروجی پخش شده و پارامترهای توابع عضویت با استفاده از الگوریتم گرادینان نزولی<sup>۳۷</sup> به روز می‌شوند. لازم به ذکر است که، پارامترهای نتیجه‌گیری تحت شرط ثابت بودن پارامترهای توابع عضویت بهینه می‌شوند. از اینرو، همگرایی نگرش ترکیبی یادگیری بسیار سریعتر از روش اصلی پس انتشار<sup>۳۸</sup> است.

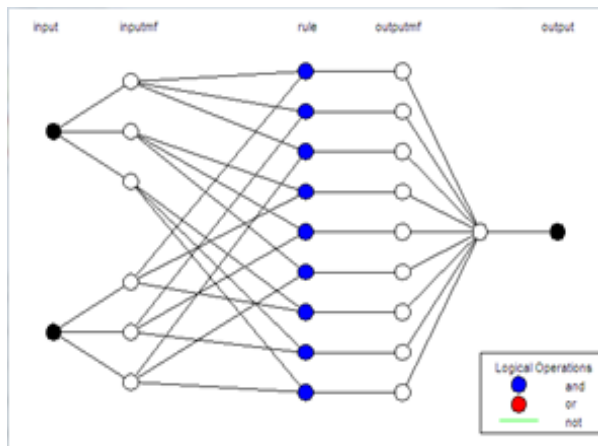
### ۳-۲- ساختار روش پیشنهادی

همانگونه که قبلاً اشاره شد یکی از روش‌هایی که برای کنترل آمارهای یک فرآیند خود همبسته مورد استفاده قرار می‌گیرد، استفاده از نمودارهای کنترل برای پایش باقیمانده‌ها است. در این دسته از روش‌ها ابتدا نیاز به یک مدل پیش‌بینی مناسب است تا از طریق آن اختلاف بین مقادیر واقعی و پیش‌بینی به عنوان مقادیر باقیمانده‌ها بدست آیند. استفاده از یک مدل سری زمانی مناسب به منظور پیش‌بینی در استفاده از این روش‌ها بسیار پراهمیت است. مدلی بهتر است که کمترین میزان خطا را داشته باشد و همچنین کمترین مقدار همبستگی بین باقیمانده‌ها را ایجاد نماید. روش ارائه شده در این تحقیق به دنبال استفاده از یکی از قدرتمندترین ابزارهای پیش‌بینی یعنی تکنیک ANFIS است. اگر مشخصه کیفی یک فرآیند خودهمبسته از مدل  $AR(P)$  تبعیت کنند، میانگین فرآیند در هر دوره زمانی تابعی از میانگین‌های فرآیند در  $P$  دوره قبل است. بنابراین، بردار ورودی‌های سیستم ANFIS به منظور پیش‌بینی هر داده، داده‌های  $P$  دوره قبل بوده و خروجی سیستم مقدار پیش‌بینی مشخصه کیفی در دوره مورد پیش‌بینی است.

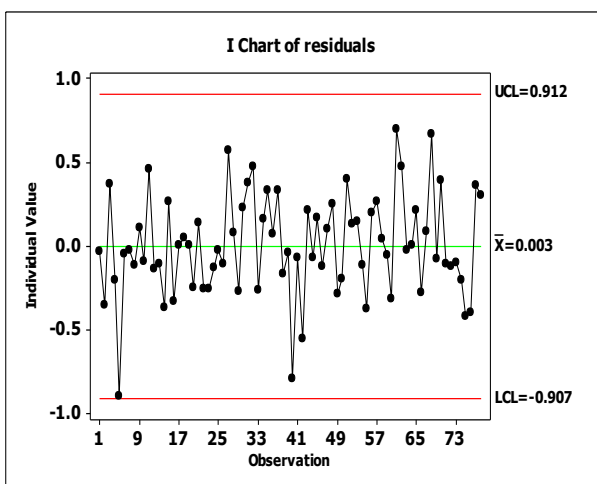
شکل ۶ نشان دهنده ساختار شبکه مربوط به روش پیشنهادی در حالت کلی است. به عنوان مثال، اگر ما بخواهیم یک نمودار کنترل باقیمانده‌ها برای یک مدل  $AR(2)$  ایجاد کنیم در اینصورت سیستم مورد نظر دارای دو متغیر ورودی خواهد بود. برای تعیین تعداد و نوع توابع عضویت برای متغیرها ورودی روش سعی و خطا مورد استفاده قرار می‌گیرد. به عبارت دیگر تعداد و انواع مختلف توابع عضویت برای هر متغیر ورودی اجرا شده سپس ترکیبی که کمترین خطا را داشته باشد انتخاب می‌شود. بعد از فاز طراحی شبکه، آموزش و اعتبارسنجی شبکه با استفاده از داده‌های آموزشی انجام می‌شود. داده‌های مورد استفاده برای آموزش شبکه باید از بهترین وضعیت فرآیند تولید جمع‌آوری شده باشند و یا دارای ویژگی‌های مشابه داده‌های جمع‌آوری شده



شکل ۱۰: تابع همبستگی باقیمانده‌ها

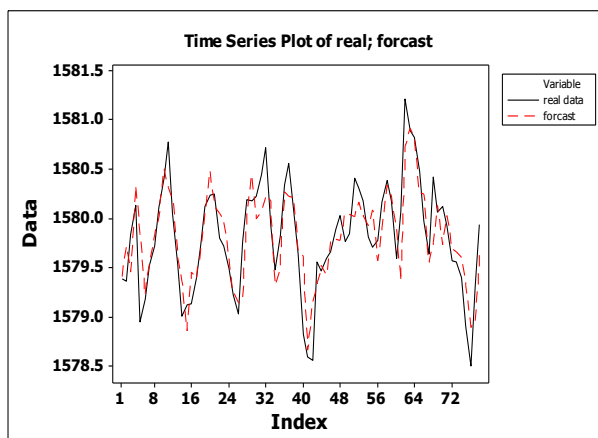


شکل ۷: ساختار شبکه مثال مورد بررسی در نرم افزار متلب

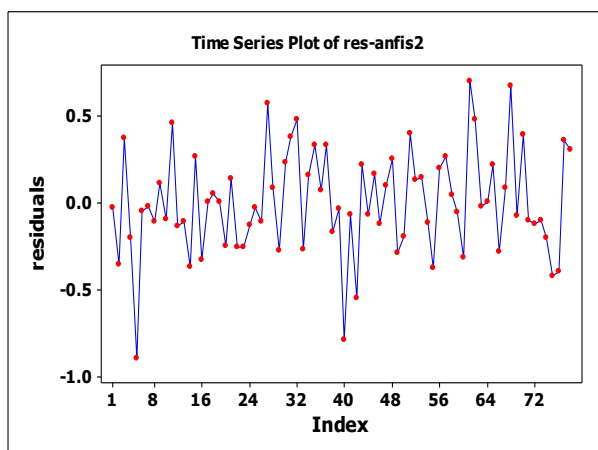


شکل ۱۱: نمودار کنترل مشاهدات انفرادی برای باقیمانده‌ها

نقاطی نزدیک حدود کنترل قرار گرفته‌اند. همانگونه که مشاهده شد، استفاده از سیستم ANFIS برای رسم نمودارهای کنترل باقیمانده‌ها به منظور کنترل داده‌های خودهمبسته دارای کارایی خوبی بود. حال در ادامه با استفاده از شبیه‌سازی، قدرت نمودارهای کنترل باقیمانده مبتنی بر ANFIS را در کشف انحرافات میانگین فرآیند در وضعیت مبنا مورد سنجش قرار دهیم. برای اینکار ابتدا با تولید داده‌های آموزشی برای وضعیت مبنا، سیستم فازی مورد آموزش قرار خواهد گرفت و سپس با تولید داده‌هایی که دارای انحراف در میانگین هستند قدرت روش پیشنهادی مورد بررسی قرار می‌گیرد. به منظور ارزیابی قدرت روش پیشنهادی مقادیر  $ARL^{3\sigma}$  برای حالت‌های تحت کنترل و خارج از کنترل محاسبه شده و برای نمودارهای مختلف مقایسه می‌شود تا بهترین نمودار برای این روش شناسایی شود.



شکل ۸: مقایسه مقادیر واقعی و پیش‌بینی حاصل از ANFIS

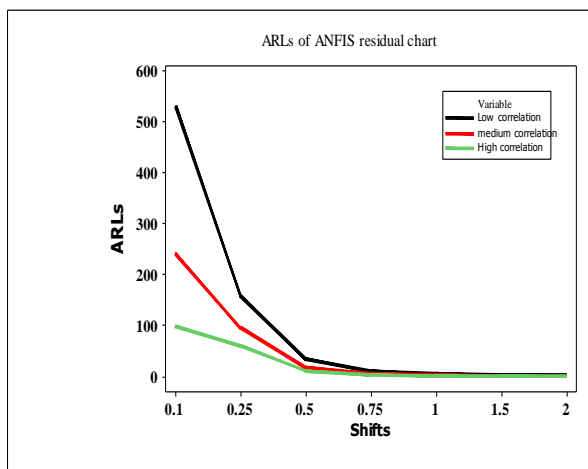


شکل ۹: نمودار پراکندگی باقیمانده‌ها

استفاده کرد. نمودار کنترل مشاهدات انفرادی برای باقیمانده‌ها همانند شکل ۱۱ نشان داده شده همانگونه که انتظار داشتیم فرآیند تحت کنترل اعلام شده است و فقط در مشاهدات ۵ و ۴۱

## ۴- نتایج شبیه‌سازی

از انحراف در میانگین فرآیند ابتدا در ۱۰ تکرار مجزا ۱۰۰۰ داده خارج از کنترل تولید شده و سپس مقدار متوسط طول دنباله محاسبه می‌شود. لازم به ذکر است که پارامترهای ARL<sub>1</sub> و ARL<sub>2</sub> دو معیار بسیار مهم در ارزیابی کارایی نمودارهای کنترل می‌باشند. بطوریکه هر چه مقدار ARL<sub>1</sub> برای نمودار بیشتر باشد بهتر است یعنی تعداد اعلام خطرهای اشتباه کمتر خواهد بود. همچنین هر چه مقدار ARL<sub>2</sub> کمتر باشد بهتر است یعنی سرعت کشف انحرافات با دلیل در میانگین فرآیند بیشتر خواهد بود. جدول ۲ نشان دهنده نتایج شبیه سازی برای سه درجه مختلف از مقادیر همبستگی بین داده‌ها است. از نمودار EWMA به منظور رسم نمودار کنترل باقیمانده‌ها استفاده شده و نتایج با یکدیگر مقایسه شده است. اعداد داخل جدول نشان دهنده ARL<sub>1</sub> به ازاء مقادیر مختلف از انحرافات می‌باشند. براساس جدول ۲ همواره عملکرد نمودار EWMA مبتنی بر باقیمانده‌های ANFIS از باقیمانده‌های مبتنی بر رگرسیون کلاسیک بهتر است. همچنین با مقایسه مقادیر ARL<sub>1</sub> متوجه می‌شویم که عملکرد روش پیشنهادی زمانیکه همبستگی بین داده‌ها زیاد است بهتر است. به منظور ارزیابی عملکرد روش پیشنهادی نسبت به ARL<sub>1</sub> نمودار رفتار ARL<sub>1</sub> را نسبت به انحرافات مختلف و برای مقادیر مختلف همبستگی در شکل ۱۲ نشان داده شده است. همانگونه که در شکل ۱۲ نشان داده شده همواره عملکرد روش پیشنهادی برای فرآیندهای با مقدار همبستگی زیاد ( $\phi_1 = 0.9$ ,  $\phi_2 = 0.6$ ) بهتر و جذابتر است. هم چنین از شکل می‌توان به قدرت روش پیشنهادی در کشف انحرافات کوچک برای فرآیندهایی با همبستگی زیاد پی برد. برای انحرافات بزرگتر از  $1\sigma$  تقریباً برای مقادیر مختلف همبستگی عملکردی مشابه و بسیار خوبی نشان داده شده است.



شکل ۱۲: نمودار ARL نسبت به مقادیر مختلف انحراف در میانگین

ما در این تحقیق به منظور اعتبارسنجی صحت روش پیشنهادی خود (مبتنی بر ANFIS) نیاز به یک سری داده معتبر داشتیم که بدین منظور از داده‌های مقاله بیسگارد و کولاچی<sup>۴۰</sup> [۲۱] استفاده کردیم که از مدل AR(۲) تبعیت می‌کردند. بنابراین در این بخش کارایی روش پیشنهادی نسبت به روش‌های رگرسیون کلاسیک (مقاله [۲۱]) برای رسم نمودارهای کنترل مبتنی بر باقیمانده‌ها و با رسم نمودارهای کنترل EWMA<sup>۴۱</sup> برای یک فرآیند با مدل AR(۲) مورد ارزیابی قرار می‌گیرد و نتایج روش پیشنهادی با روش رگرسیون کلاسیک مقایسه خواهد شد. بنابراین، به منظور شبیه‌سازی می‌توان داده‌ها را مطابق رابطه ۱۲ تولید کرد.

$$x_k = \mu_0 + \phi_1(x_{k-1} - \mu_0) + \phi_2(x_{k-2} - \mu_0) + \varepsilon_k \quad (12)$$

که در آن  $\phi_1$  ضریب همبستگی هر داده نسبت به داده قبلی خود و  $\phi_2$  ضریب همبستگی هر داده نسبت به دو داده قبلی خود است.  $\varepsilon_k$  عبارت تصادفی است که دارای توزیع نرمال با میانگین ۰ و واریانس  $\sigma^2$  می‌باشند. و همچنین  $\mu_0$  میانگین برآوردی فرآیند است. در این بخش فرآیند شبیه‌سازی به طور مجزا برای سه دسته همبستگی، زیاد، متوسط و کم انجام می‌شود. روش انجام کار بدین گونه است که ابتدا با استفاده از معادله فوق ۱۰۰۰۰ داده تحت کنترل تولید می‌شود. سپس با استفاد از نرم افزار Matlab اقدام به طراحی و بهینه‌سازی سیستم ANFIS می‌شود. برای این منظور دو داده قبلی برای هر داده به عنوان مقادیر ورودی سیستم ANFIS مورد استفاده قرار می‌گیرند و مقدار واقعی هر داده به عنوان خروجی به سیستم ANFIS داده می‌شود. پس از فاز آموزش شبکه ANFIS مقادیر باقیمانده‌ها که اختلاف بین مقادیر واقعی و مقادیر پیش‌بینی حاصل از سیستم ANFIS می‌باشند محاسبه می‌شوند. سپس با استفاده از نرم افزار Minitab نمودار کنترل EWMA برای مقادیر باقیمانده‌ها ایجاد شده و در نهایت با تعیین حدود کنترل مبنا برای نمودار مورد نظر مرحله طراحی به پایان می‌رسد. سپس در مرحله ارزیابی، به منظور محاسبه متوسط طول دنباله در حالت خارج از کنترل (ARL<sub>1</sub>)، پارامترهای مربوط به نمودارهای کنترل طوری طراحی می‌شوند که مقدار متوسط طول دنباله در حالت تحت کنترل (ARL<sub>2</sub>) برای همه نمودارها حدود ۴۵۰ شود تا از این طریق بتوان تحلیل صحیحی نسبت به ARL<sub>1</sub> داشت. همچنین به منظور محاسبه متوسط طول دنباله در حالت خارج از کنترل (ARL<sub>1</sub>) نیز برای یک مقدار مشخص

جدول ۲: نتایج شبیه‌سازی برای مقادیر مختلفی از  $\phi$ 

ANFIS						روش پیش‌بینی
رگرسیون						
$\phi_x = 0.1$	$\phi_x = 0.5$	$\phi_x = 0.1$	$\phi_x = 0.9$	$\phi_x = 0.5$	$\phi_x = 0.1$	مقدار انحراف (شیفت)
$\phi_y = 0.6$	$\phi_y = 0.2$	$\phi_y = 0.1$	$\phi_y = 0.6$	$\phi_y = 0.2$	$\phi_y = 0.1$	
۴۵۱	۴۵۱	۴۵۱	۴۵۰	۴۵۰	۴۵۰	۰
۶۲۶/۳۹	۴۳۵/۵۲	۹۳۹/۵۸	۹۸/۴۷	۲۴۱/۴۳	۵۳۱/۴۸	۰/۱
۱۷۳/۳۰	۱۰۳/۵۶	۲۱۰/۴۷	۶۰/۰۷	۹۴/۶۴	۱۵۷/۹۳	۰/۲۵
۳۳/۴۱	۲۰/۳۲	۴۶/۹۸	۱۰/۸۳	۱۶/۱۳	۳۳/۶۲	۰/۵
۱۰/۴۴	۵/۶۱	۱۳/۵۳	۲/۱۷	۴/۷۳	۱۰/۱۷	۰/۷۵
۵/۴۳	۲/۳۲	۶/۳۹	۱/۲۳	۱/۷۶	۴/۱۵	۱
۴/۱۸	۲/۲۳	۵/۲۶	۰/۹۹	۱/۲۶	۲/۳۵	۱/۵
۳/۵۵	۱/۷۴	۳/۵۷	۰/۹۸	۱/۱۶	۱/۷۵	۲

## ۵- نتیجه‌گیری

از آنجا که همبستگی‌های بزرگتر عملکرد بهتری دارد (شکل ۱۲). به عنوان تحقیقات آتی می‌توان به استفاده از روش پیشنهادی به منظور کنترل فرایندهای خودهمبسته چندمتغیره و همچنین همبستگی‌های غیرخطی اشاره کرد.

### مراجع

- [1] Alwan. LC, Roberts. HV. (1988), Time-series modeling for statistical process control. J Bus Econ Stat, 6, pp.87-95.
- [2] Wardell. DG, Moskowitz H, Plante. RD. (1994), Run-length distributions of special-cause control charts for correlated processes. Technometrics 36(1), pp.3-17.
- [3] Runger . GC, Willemain. TR. (1995), Model-based and model-free control of autocorrelated processes. J Qual Technol 27(4), pp.283-292.
- [4] Harris. T.J, Ross. W.H. (1991), Statistical process control procedures for autocorrelated observations, The Canadian Journal of Chemical Engineering, 69, pp.48-57.

در این مقاله به بررسی اثر حضور همبستگی بین مشاهدات یک فرآیند تولیدی در استفاده از نمودارهای کنترل کلاسیک پرداخته شد و نشان داده شد که استفاده از نمودارهای کنترل کلاسیک برای داده‌های خود همبسته کارایی مناسبی نداشته و باعث اختراهای اشتباه تحلیل‌گران کیفی خواهد شد. در تحقیق ارائه شده، نمودارهای کنترل باقیمانده‌ها مبتنی بر روش پیش‌بینی ANFIS به منظور کنترل و کشف انحرافات در یک فرآیند خود همبسته مورد بررسی قرار گرفت. یکی از مهم‌ترین مزیت‌های روش پیشنهادی نسبت به روش‌های کلاسیک سری‌های زمانی در قدرت مدل کردن روابط غیرخطی بین متغیرهای کنترل و نداشتن فرضیات محدودکننده است. در شبیه‌سازی انجام شده نشان داده شد که استفاده از نمودار کنترل EWMA مبتنی بر باقیمانده‌های سیستم ANFIS برای انحرافات بزرگتر از  $1\sigma$  به ازاء همبستگی‌های مختلف عملکرد بسیار خوبی را دارد. ولی برای انحرافات کوچک در میانگین به