

پایش ناپارامتری پروفایل‌ها و تصمیم‌گیری درمورد کیفیت فرایند با استفاده از مدل اثرات آمیخته

فاطمه حاجی احمدی^۱، رسول نورالسنا^{۲*}

۱. کارشناس ارشد دانشکده مهندسی صنایع، دانشگاه آزاد اسلامی واحد تهران جنوب

۲. استاد دانشکده مهندسی صنایع، دانشگاه علم و صنعت ایران

(تاریخ دریافت ۹۴/۱/۲۹ - تاریخ دریافت اصلاح‌شده ۹۴/۱۲/۸ - تاریخ تصویب ۹۴/۱۲/۲۶)

چکیده

در بسیاری از کاربردهای کنترل کیفیت آماری فرایند، کیفیت یک فرایند از طریق یک یا چند تابع یا یک یا چند متغیر، تحت عنوان پروفایل بیان می‌شود. پایش پروفایل‌ها ثبات رابطه تابع پروفایل را در طول یک بازه بررسی می‌کند. نمودارهای کنترل پروفایل‌های ناپارامتری در شرایطی که توصیف این رابطه به صورت پارامتری پیچیده باشد اهمیت دارد. بیشتر نمودارهای کنترل کیفیت آماری موجود در ادبیات به منظور پایش پروفایل‌های پارامتری طراحی شده‌اند. ساختار همبستگی درونی پروفایل‌ها که در عمل کاربرد زیادی دارد در اغلب مدل‌سازی‌های پیشین در نظر گرفته نشده است. این پژوهش بر پایش ناپارامتری پروفایل‌ها در شرایطی تمرکز دارد که پروفایل‌ها همبستگی دارند و فرض استقلال برقرار نیست. نمودار کنترل فاز دوم پیشنهادی که مدل اثرات آمیخته را با به کارگیری چارچوب کلی تحلیل واریانس هموارسازی اسپلاین با طرح کنترل هتلینگ تلفیق کرده است، دقت کافی و انعطاف‌پذیری بسیار بیشتری به‌ویژه در مورد داده‌های رسته‌ای دارد. نتایج شبیه‌سازی نشان می‌دهد روش پیشنهادی به شکل مؤثری تغییرات در پروفایل‌ها را شناسایی می‌کند.

واژه‌های کلیدی: فاز دوم کنترل کیفیت آماری، مدل‌های ناپارامتری اثرات آمیخته، همبستگی درون پروفایل‌ها، هموارسازی.

مقدمه

سانگ [۱۲] درباره کاربردهای پروفایل بحث کردند. در ادبیات تحقیق، درکل طرح‌های پایش در دو پایش فاز اول و دوم مطالعه شده‌اند. هدف پایش فاز اول بررسی برقراری وضعیت تحت کنترل مشاهدات و فرایند و نیز تشخیص تفاوت میان شرایط تحت کنترل و حضور علل اکتسابی است. در فاز دوم کنترل کیفیت آماری، هدف کشف تغییرات فرایند نسبت به پارامترهای برآورده‌شده فرایند تحت کنترل در سریع‌ترین زمان ممکن است. عملکرد رویکرد فاز دوم به‌طور عمومی از طریق شاخص متوسط طول گردش که معرف متوسط تعداد مشاهدات تا صدور علامت تغییر است ارزیابی می‌شود. زمانی عملکرد رویکرد پایش، مطلوب ارزیابی می‌شود که در کشف تغییرات پروفایل‌ها ARL^۱ (متوسط طول گردش) کمتری داشته باشد. در مطالعات متعددی کیفیت محصول یا فرایند با استفاده از رابطه غیرخطی بین دو یا چند متغیر مشخص شده است. در میان تحقیقات انجام‌گرفته در این زمینه

در سال‌های اخیر کنترل کیفیت آماری در شرایطی که کیفیت فرایند یا محصول توسط تابعی قابل تعریف باشد، از جنبه نظری و کاربردی توسعه چشمگیری یافته است. پایش پروفایل‌ها در زمینه‌هایی مانند صنعت خودروسازی [۱]، تولید قطعه [۳ و ۲]، کالیبراسیون [۴]، گردشگری [۵]، اقتصادی [۶] فرآیند بارگذاری [۷] کاربردی کارا داشته‌اند. پروفایل‌ها انواع مختلفی دارند که انتخاب مدل مناسب برای آن‌ها اهمیت ویژه‌ای دارد [۸]. هر مدلی که برای کاربردی خاص استفاده می‌شود باید ضمن سادگی، به‌گونه مناسبی ویژگی‌های مشاهدات را بیان کند. همچنین، برای هر مدل باید روش‌های پایشی طراحی کرد که به‌گونه‌ای مؤثر تغییرات را شناسایی کند و همچنین برای تفسیر هشدارهای خارج از کنترل نیز مناسب باشد. تعدادی از محققان مانند استور و بریل [۹]، کنک و آلبین [۱۰]، محمود و وودال [۱۱]، وودال و همکاران و ونگ و تی

در سال‌های اخیر، استفاده از تحلیل آماری غیرپارامتری به‌ویژه در تحلیل داده‌هایی که دارای اغتشاش هستند گسترش بیشتری یافته است. استفاده از روش‌های غیرپارامتری در کاربردهای صنعتی واقعی که فرض‌های نرمال و تصادفی بودن به‌ندرت برقرار است، بی‌گمان مطلوب و کاربردی است. تحلیل رگرسیون غیرپارامتری، محدودیت نیاز به تعداد پارامترهای کم رگرسیون پارامتری را رفع می‌کند و توانایی مدل‌سازی با کمترین خطا را در شرایط ابعاد زیاد^۲ دارد. وظیفه اصلی رگرسیون غیرپارامتری ارائه نوعی از اطلاعات دقیق با توجه به برآورد به‌دست‌آمده است. در حوزه پایش غیرپارامتری پروفایل‌ها در فاز دوم که همبستگی درون پروفایل‌ها را لحاظ می‌کنند تحقیقات بسیار اندکی صورت گرفته است. در این میان، روش‌های مدل‌سازی آماری پیشرفته‌ای توسعه یافته‌اند که می‌توانند تغییرات درون پروفایل‌ها را از طریق مدل‌های اثرات آمیخته در نظر بگیرند. جنسن، برچ و وودال [۲۳ و ۲۲] پایش پروفایل‌های خطی و غیرخطی را با به‌کارگیری مدل‌سازی اثرات آمیخته خطی و غیرخطی ارائه‌شده توسط لبرد و ویر [۲۴] تحلیل کردند.

در این پژوهش، به‌منظور برآزش مناسب و نیکوی پروفایل‌های غیرخطی که ویژگی همبستگی درونی دارند، از مدل آنالیز واریانس هموارساز ارائه‌شده از سوی گو^۳ استفاده شده است و در طرح نمودار کنترل پیشنهادشده مجذوراتی هتلینگ امکان پایش آثار تصادفی پروفایل‌های دارای همبستگی درونی فراهم شده است.

در ادامه پژوهش، تدوین مدل آمیخته ناپارامتری به‌طور خلاصه بررسی می‌شود. در بخش بعد، نمودار کنترل پیشنهادی مبتنی بر طرح مجذوراتی هتلینگ با استفاده از برآوردهای مدل آمیخته رگرسیون ناپارامتری ارائه می‌شود. سپس روش معرفی‌شده از طریق شبیه‌سازی عددی تشریح می‌شود. در نهایت، نتیجه‌گیری براساس معیار متوسط طول دنباله نمودار کنترل با استفاده از شبیه‌سازی مونت کارلو ارائه می‌شود.

رویکرد پایش ناپارامتری مدل اثرات آمیخته

مدل‌های آمیخته در مدل‌سازی و توصیف داده‌هایی استفاده می‌شود که رابطه بین متغیر پاسخ و چند متغیر کمکی

می‌توان به مطالعات موردی بریل [۱۳]، والکر و رایت [۱۴]، جیانگ و همکاران [۱۵] و ویلیامز و همکاران [۱۶] اشاره کرد. براساس اینکه پروفایل‌های غیرخطی پیچیدگی تقریباً بیشتری دارند، تحقیقات کمتری در راستای توسعه رویکردهایی برای پایش این‌گونه پروفایل‌ها انجام گرفته است. در زمینه پایش پروفایل‌های غیرخطی، جین و شی [۱۷] طرحی انطباق‌پذیر را ارائه دادند که به اطلاعات تحت کنترل نیاز نداشت یا میزان بهره‌گیری از آن اندک بود. آن‌ها با استفاده از روشی تکرارشونده دو نمودار مجذوراتی هتلینگ را برای هر یک از مشاهدات جدید مقایسه کردند. واکر و رایت [۱۴] از مدل‌های جمعی تعمیم‌یافته برای مقایسات مجموعه‌ای از منحنی‌ها استفاده کردند. آنان جزء تابعی داده‌ها را از طریق رویکرد هموارسازی و تغییرات تبیین‌نشده را از طریق پارامترهای رگرسیونی مدل‌سازی کردند. در روش پیشنهادی آنان منحنی‌های نامتعارف از طریق آزمون فیشر متمایز می‌شدند. ویلیام و همکاران [۱۸] طرح آنان را با معرفی آماره مجذوراتی هتلینگ گسترش دادند. دینگ و همکاران [۱۹] از انواع مختلفی از روش‌های کاهش داده‌ها از جمله آنالیز اجزای اصلی و آنالیز اجزای مستقل، به‌منظور پایش پروفایل‌های غیرخطی استفاده کردند. جنسن [۲۰] به‌منظور در نظر گرفتن ساختار همبستگی از یک مدل ترکیبی غیرخطی برای پایش پروفایل‌های غیرخطی استفاده کرد. در بسیاری از موارد کاربردی تغییرات میان پروفایل‌های تحت کنترل بسیار زیاد است و از طریق نوفه‌های توزیع‌های مستقل نرمال، تبیین نمی‌شوند. منحنی‌های رشد نمونه‌های متداول چنین وضعیت‌هایی هستند. در این موارد بیشتر تغییرات ناشی از تغییرات هر مشاهده با مشاهده دیگر است و نمی‌توان تغییرات پروفایل‌ها را از طریق تغییرات نوفه‌های تصادفی درون هر منحنی تبیین کرد. پی‌نبر [۲۱] نمونه کاربردی بسیار مناسبی در زمینه تولید بیان کرد. مطالعات قبلی پایش پروفایل‌ها بر فرض استقلال پروفایل‌ها و خطاهای تصادفی انجام گرفته بود، اما در بسیاری کاربردها برقراری این فرض‌ها امکان‌پذیر نیست؛ برای نمونه، امکان دارد همبستگی فضایی در مورد اندازه‌گیری ابعاد یا همبستگی سریالی در مورد مشاهدات جمع‌آوری‌شده طی زمان وجود داشته باشند.

طرح مشاهدات اختیاری مطلوب باشد. آنان نشان دادند رهیافت مذکور دارای مزیت سرعت محاسبات و سهولت به‌کارگیری است. زو و همکاران [۳۲] از آماره GLR برای کشف جابه‌جایی‌های بالقوه استفاده کردند. البته زمانی که هریک از پروفایل‌ها از طرح اختیاری تبعیت کنند، این آماره شرایط کاربرد ندارد. به‌منظور در نظر گرفتن پایش در شرایط وجود همبستگی و ناپارامتری، کیو و همکاران [۳۱] نمودار کنترل ابداعی را در پایش فاز دوم پیشنهاد دادند که از تلفیق هموارسازی موضعی خطی و میانگین متحرک موزون نمایی به‌دست آمده است. در این روش، همبستگی درون پروفایل‌ها از طریق مدل‌های تأثیرات آمیخته‌شی و همکاران [۳۳] و رایس و زو [۳۴] توضیح داده شده است. این مدل ساختار واریانس کواریانس انعطاف‌پذیری دارد. از مزیت‌های بهره‌گیری از این روش، سرعت محاسبات و نیز راحتی استفاده از سوی ارائه‌کنندگان این روش است. در نمودار پیشنهادی کیو و همکاران [۳۱] در فاز دوم به معلوم‌بودن پارامترهای فرایند تحت کنترل نیازی نیست، بلکه پارامترهای مربوط به فرایند تحت کنترل را از اطلاعات آن برآورد می‌کند. در مدل مزبور از تلفیق هموارسازی اطلاعات پروفایل با هموارسازی خطی کرنل و طرح موزون هموارسازی نمایی حاصل شده است. کیو و همکاران [۳۱] استفاده از این رویکرد را هنگامی که فرایند تحت کنترل دارای همبستگی درون پروفایلی است، برای تخمین پارامترها در فاز دوم بررسی کردند. ابدل سلام و برچ و جنسن [۳۴] فاز اول نمودارهای کنترلی پروفایل‌ها را از طریق مدل‌های آمیخته پارامتری و شبه‌پارامتری پایش کردند. آنان تأثیرات آمیخته را با تلفیق دو روش جدید بررسی کردند. روش اول شامل رویکرد ناپارامتری از طریق رگرسیون آمیخته پی اسپلین و روش دوم رویکرد شبه‌پارامتری از طریق پایش مقاوم آمیخته پروفایل بود. بسیاری از تحقیقات پیشین پایش پروفایل بر مدل‌سازی پارامتری خطی و غیرخطی در هریک از آثار ثابت و تصادفی تحت فرض تعیین صحیح مدل صورت گرفته بود، اما تحقیق ابدل سلام و همکاران [۳۴] شرایطی را مطالعه کرد که در آن حداقل مدل پارامتری خانواده پروفایل‌ها دارای ابهام است. در تحقیق مذکور، با استفاده از دو رویکرد ناپارامتری و شبه‌پارامتری و با تلفیق رویکرد مدل آمیخته از

براساس یک یا چند عامل گروه‌بندی می‌شود و همچنین در زمینه داده‌های زمانی کاربرد دارد که به‌طورعمومی شامل اندازه‌گیری‌های مکرر یا اندازه‌های مختلف از یک آزمودنی می‌شود. این مدل‌ها اغلب در زمینه‌های مطالعات کلینیکی، بیوپزشکی، علوم زمین، هیدرولوژی و پردازش تصویر ابزاری قدرتمند به‌شمار می‌روند. مدل‌های آمیخته از جمله رویکردهایی است که در کنار رهیافت‌های مدل‌های حاشیه‌ای و مدل‌های ساختار مارکوف برای تحلیل این نوع داده‌ها استفاده می‌شود. مطالعات اخیر در زمینه تحلیل داده‌های زمانی در دیگل [۲۵] دمیدنکو علاوه‌بر سایر تحقیقات شایان توجه است. در مطالعات زمانی در بسیاری کاربردها ممکن است مدل‌های پارامتری محدودیت زیادی داشته باشند و دست کم در برخی موارد در تحلیل‌های داده‌های اولیه ضعیف باشند. نکته شایان توجه در زمینه تفاوت اساسی داده‌های زمانی و متقاطع این است که داده‌های زمانی و متقاطع مستقل‌اند. چالش اصلی تحلیل داده‌های زمانی، احتساب همبستگی درون آزمودنی‌هاست که مدل‌های آمیخته در راستای رفع این مشکل ابزاری ارزشمند و پرتوان به‌شمار می‌روند. آبرامویچ، آنجلان [۲۶]، گو [۲۷]، آنتونیادونیس و ساپاتیناس [۲۸]، موریس و کارول [۲۹] و موریس و همکاران [۳۰] به برازش تأثیرات تصادفی در داده‌های تابعی^۴ پرداختند.

نمودارهای کنترل مطرح‌شده درمورد پایش پروفایل‌ها نیازمند برقراری فرض‌های بنیادینی چون معین‌بودن یا غیرتصادفی‌بودن مشاهدات درون پروفایل و نیز نامتغیربودن از یک پروفایل به پروفایل دیگر است. در بسیاری کاربردهای صنعتی این مفروضات اعتبار لازم را ندارند؛ بنابراین، هنگامی که ساختار مشاهدات درون پروفایل‌ها دارای توزیع یکسان بوده و متغیر تصادفی از توزیع مفروض پیروی کند و یا در موارد متداول دیگری که رخداد مشاهدات، در زمان‌های مختلف از مفقودشدگی برخوردار است. همچنین، مواردی مانند پروفایل‌های عمود چگال که واکر و رایست [۱۴] بررسی کردند از جمله موضوعات چالش‌برانگیز مبحث پایش پروفایل‌ها محسوب می‌شوند. کیو و زو [۳۱] نمودار کنترلی را بر مبنای روش هموارسازی خطی کرنل و میانگین متحرک موزون نمایی پیشنهاد کردند تا برای مدیریت مشکلات پایش ناپارامتری در شرایط

همکاران [۳۹] و گو و واهبا [۴۰] درباره روش‌های عددی برای برازش هموارسازی مدل اسپلاین بحث کرده‌اند. آرونزاجن [۴۱] و واهبا [۳۷] تفسیر و شرح کاملی از اطلاعات درباره فضا و شرایط توابع مورد استفاده در این مدل را ارائه کرده‌اند. فرم کلی هموارسازی اسپلاین-SS-ANOVA با داده‌های نرمال، در رابطه ۱ مشاهده می‌شود.

(۱)

$$y_i = f(x_1(i), \dots, x_d(i)) + \varepsilon_i$$

که در آن $i=1, \dots, n$ و $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)'$ دارای توزیع

$$N(0, \sigma^2 I_{n \times n}) \quad \text{و} \quad t_\alpha \in T^{(\alpha)} \quad \text{که} \quad T^{(\alpha)} \quad \text{فضای شمارا و}$$

$$\sigma^2 \quad \text{و} \quad (x_1, \dots, x_d) = X_i \in \mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \dots \otimes \mathcal{X}_d$$

ممکن است نامعلوم باشد. فرض می‌کنیم $f \in W_p^m$

به عبارت دیگر:

(۲)

$$W_p^m = \left\{ f : f^{(v)}, m-1, \int_0^1 (f^{(m)}(t))^2 dt < \infty \right\}$$

برای آنکه f شرایط اندازه‌پذیری را داشته باشد، تجزیه یکتای تحلیل واریانس تابع f به صورت رابطه ۳ تعریف می‌شود:

$$f(X) = \mu + \sum_{j \in J_1} f_j(x_j) + \sum_{(j_1, j_2) \in J} f_{j_1, j_2}(x_{j_1}, x_{j_2}) + \dots + \sum_{(j_1, \dots, j_d) \in J} f_{j_1, \dots, j_d}(x_{j_1}, \dots, x_{j_d}) \quad (۳)$$

به عبارتی، فضای مدل SS-ANOVA دارای عناصر تجزیه فوق است. در این مدل، هر بخش در اولین مجموع آثار اصلی و هر بخش در دومین مجموع آثار متقابل دوسویه را بیان می‌کنند. تجزیه SS-ANOVA هنگامی که تمام x_k ها متغیرهایی از نوع گسسته باشند به تحلیل کلاسیک مدل D-طرفه تحلیل واریانس منجر می‌شود؛ بنابراین، تجزیه ANOVA SS توسعه‌ای طبیعی از تجزیه تحلیل واریانس کلاسیک از دامنه‌های گسسته به دامنه‌های عمومی و از فاصله فضاهای متناهی بعد به بی‌نهایت بعد است. این روش ضرب تانسور RKHS را به زیرفضاهای معنی‌دار تجزیه می‌کند. تجزیه SS-ANOVA مانند تجزیه

طریق در نظر گرفتن مجموعه پروفایل‌ها به عنوان نمونه تصادفی از جامعه کلی خود، همبستگی درون پروفایل‌ها بررسی شد. در هریک از رویکردها، دو آماره T^2 هتلینگ به منظور شناسایی پروفایل‌های ناهمسان به کار گرفته شد. روش ناپارامتری آنان استفاده از روش رگرسیون اسپلاین تاوانیده در مدل‌های آمیخته بود.

۱. مدل اثرات آمیخته آنالیز واریانس اسپلاین

هموارسازی SS-ANOVA

۱-۱. مدل آنالیز واریانس هموارسازی

قابلیت کاربردی اصلی مدل تحلیل واریانس اسپلاین هموارسازی به طور ویژه در مورد داده‌های زمانی و طولی-زمانی است. این مدل محدودیت الزام ساختار خاص خطاها و نیز داده‌ها را ندارد. میانگین توابع این نوع داده‌ها با به کارگیری چارچوب متحدی از مدل‌های اسپلاین عمومی برآورد می‌شوند و این نوع مدل قابلیت مدل‌سازی در دامنه دلخواه را دارد [۳۶]. تحقیقات گذشته در مورد مدل‌های عمومی اسپلاین بر استقلال مشاهدات تأکید داشتند، اما در مطالعات بعدی این مدل در مدل‌سازی داده‌های زمانی و زمان-طول از روش‌های متداول به طور شایان توجهی پیشی گرفت. هموارسازی اسپلاین یکی از فنون آماری پرتوان، پیشرفته و انعطاف‌پذیر است. این رویکرد شامل انواع اسپلاین مانند چندجمله‌ای، دوره‌ای، دایره‌ای، تین پلیت،^۵ ال-اسپلاین، بردار اسپلاین و آنالیز واریانس هموارسازی اسپلاین است.

مدل آنالیز واریانس اسپلاین همواری (SS-ANOVA) چارچوبی بسیار عمومی را برای تحلیل داده‌ها به منظور مدل‌سازی و یادگیری در زمینه‌های گوناگون فراهم می‌سازد. در این مدل مشاهدات گسسته، اغتشاشی مستقیم و غیرمستقیم با ورودی‌های چندگانه و همبستگی‌های متغیر پاسخ و انواع مختلفی از ساختارها ممکن است با هم همراه شوند. جزئیات بیشتر درباره مفاهیم بنیادین این مدل در تحقیقات واهبا [۳۶] و گو [۲۷] یافت می‌شود.

یک چارچوب ریاضی برای برازش این مدل در دستنمایی تاوانیده و در روش‌های خاص اسپلاین از سوی چن [۳۷]، گو و واهبا [۳۸] و واهبا [۳۶] توسعه داده شده است. گو و

غیرپارامتری کیو و زوو فرض می‌شود زمانی که فرایند تحت کنترل آماری است، رابطه بین متغیر پاسخ و متغیر پیشگو و خطاها به صورت مدل آمیخته بیان می‌شود.

$$y_{ij} = g(x_{ij}) + f_i(x_{ij}) + \varepsilon_{ij} \quad (۴)$$

$g(x_{ij})$ تابع عمومی یا تأثیرات ثابت است که برای تمام پروفایل‌ها یکسان است و $f_i(x_{ij})$ تأثیرات تصادفی مربوط به نمونه i ام است، به طوری که $i = 1, 2, \dots, m$ و $j = 1, 2, \dots, n_i$ و ε_{ij} ها خطاهای خودهمبسته است؛ به عبارت دیگر، $\varepsilon_i(x)$ خطای اندازه‌گیری است که ناشی از هر دو اثر ثابت و تصادفی است و $g(x)$ تابع میانگین جامعه و $f_i(x)$ انحراف i امین تابع از تابع میانگین جامعه یا $g(x)$ است که تابع اثر تصادفی i امین نمونه نامیده می‌شود. فرض می‌شود مشاهدات نمونه i ام به صورت (x_i, y_{ij}) در نظر گرفته می‌شود، به طوری که y_{ij} نشان‌دهنده مشاهده i ام در نمونه i ام است. از یک سو، مقادیر x ها نیز ثابت است و برای تمام پروفایل‌ها یکسان فرض می‌شود (حالت داده‌های متعادل) A و همچنین فرض می‌شود مشاهدات متناظر در پروفایل‌های متوالی مستقل از یکدیگرند. از سوی دیگر، تأثیرات تصادفی مستقل از یکدیگر و مستقل از خطاها در نظر گرفته می‌شود.

۳-۱. ارتباط تجزیه ANOVA SS با مدل‌های

آمیخته

با فرض $\phi_v(t) = t^{v-1} / (v-1)!$ ، $v = 1, 2, \dots, m$ و با تعریف $R^1(s, t) = \int_0^1 (s-u)_+^{m-1} (t-u)_+^{m-1} du / ((m-1)!)^2$ که در آن $(x)_+ = x$ به‌ازای $x \geq 0$ و در سایر شرایط $(x)_+ = 0$. همچنین، تعریف می‌شود

$$\sum_{n \times n} \{R^1(t_i, t_j)\}_{i,j=1}^n \quad \text{و} \quad T_{n \times m} \{\phi_v(s, t_i)\}_{i=1, v=1}^n$$

به صورت رابطه γ بیان می‌شود:

$$y = Td + u + \varepsilon \quad (۷)$$

در شرایطی که بیش از یک اثر تصادفی وجود دارد و برآورد آن‌ها مطلوب است، ارتباط مدل آنالیز واریانس هموارسازی اسپلاین و مدل آمیخته به صورت رابطه δ بیان می‌شود:

تحلیل واریانس کلاسیک، به ساختارهای سلسله‌مراتبی منجر می‌شود که برای انتخاب مدل و تفسیر مفیدند. می‌توان تجزیه‌های متفاوت SS-ANOVA را براساس عملگرهای مختلف میانگین (یا به عبارتی تجزیه‌های مختلف فضای حاشیه‌ای) به دست آورد؛ بنابراین، تجزیه SS-ANOVA باید جایگزین برخی از مدل‌های ثابت به عنوان یک مدل عمومی برای ساخت مدل‌های چندمتغیره غیرپارامتری در نظر گرفته شود. همچنین، می‌توان از آن‌ها برای ساخت زیرمدل‌های فرعی برای مؤلفه‌های مدل پیچیده‌تر استفاده کرد. تصمیم مستقیم کمترین مربعات وزنی توانیده به صورت رابطه ϵ تعریف می‌شود:

(۴)

$$\min_{f \in W, \tau} \left\{ \frac{1}{n} (Y - f)^T C^{-1} (Y - f) + \lambda \int_0^1 (f^{(m)}(t))^2 dt \right\}$$

که در آن λ برقرارکننده توازن بین نکویی برآزش و همواری برآورد مدل است و پارامتر همواری نام دارد. اثبات می‌شود $\hat{f} = (\hat{f}(t_1), \dots, \hat{f}(t_n))^T = Ay$ برآورد ناپارامتری است که در آن A ماتریس هت (هشین) است.

(۵)

$$M(\lambda, \tau) = \frac{y'W(I-A)y}{[\det^+(W(I-A))]^{\frac{1}{n-m}}}$$

روش بیشینه درستنمایی تصمیم‌یافته GML^۶ یکی از روش‌هایی است که برآورد پارامتر مدل را ارائه می‌دهد. برآورد پارامتر با کمینه کردن تابع M حاصل می‌شود. در این پژوهش، از برآورد GML برای تخمین پارامترها استفاده می‌شود که نسبت به سایر روش‌های رقیب مانند GCV^۷ ثبات بیشتری دارد.

۲-۱. مدل اثرات آمیخته

مدل اثرات آمیخته به صورت $y = X\beta + Zu + \varepsilon$ در نظر گرفته می‌شود که در آن X اثر ثابت یا تابع عمومی و u اثر تصادفی است و دارای توزیع $MN(0, \Sigma)$ است که در آن Σ ماتریس واریانس کواریانس و مثبت معین است. همچنین، فرض می‌شود $\text{cov}(\varepsilon_i, u_i) = 0$. براساس مدل

تأثیرات تصادفی کافی کارا دارد. همچنین، جنسن و همکاران [۲۰] نشان دادند وقتی $X_i = Z_i$ و همچنین وقتی ماتریس Z_i درون ماتریس X_i قرار می‌گیرد، میانگین تخمین‌های تأثیرات تصادفی برابر با صفر است. می‌توان از آماره T^2 به صورت رابطه ۱۰ در فاز ۲ استفاده کرد:

$$T_i^2 = (\hat{u}_i)' \Sigma^{-1} ((\hat{u}_i)) \quad (10)$$

که Σ از فاز ۱ به دست می‌آید. برای تخمین u_β می‌توان از برآورد حاصل از توزیع رابطه ۱۱ استفاده کرد:

$$u_\beta \sim N(0, \sigma^2 \theta_\beta \Sigma_\beta / n\lambda) \quad (11)$$

در فاز ۲، برخلاف روش حداقل مربعات برای تخمین تأثیرات تصادفی هر نمونه جدید از مقادیر تأثیرات ثابت به دست آمده از فاز ۱ استفاده می‌شود. در نتیجه، در محاسبات شبیه‌سازی برای تخمین ماتریس واریانس کوواریانس تأثیرات تصادفی باید واریانس مقادیر ثابت نیز معلوم باشد. در نتیجه، با فرض m مشخص، در محاسبات شبیه‌سازی می‌توان تخمین واریانس تأثیرات ثابت را حساب کرد و آماره T_i^2 را به دست آورد. حد کنترل بالا هم به گونه‌ای انتخاب می‌شود که ARL تحت کنترل مشخصی به دست آید.

$$UCL_{T^2} = \chi_{\alpha, (n)}^2 \quad (12)$$

مطالعه شبیه‌سازی

اعتبارسنجی مدل پیشنهادی به صورت گام‌های زیر تشریح می‌شود:

۱. تولید پروفایل‌های مناسب تحلیل

مجموعه داده‌های این تحقیق مشابه مطالعه کیو و همکاران [۳۱] تولید می‌شوند. فرض می‌شود $x_{ij} \sim U(0,1) i=1, \dots, n$ و بدون آنکه از عمومیت مسئله کاسته شود، فرض می‌شود x ها مقادیر ثابتی اختیار می‌کنند. نرمال بودن خطاها فرض اساسی دیگری است که در این مدل فرض می‌شود. در حالت کلی، متغیر مستقل باید هنگام وارد کردن در مدل، مقیاس شود؛ به عبارت دیگر، مقادیر مقیاس شده بین صفر تا یک باشند. همچنین، مانند تحقیق هاوکینز و اوول ول [۴۳] به رفتار متوسط

$$Y = Td + \sum_{\beta=1}^p u_\beta + \varepsilon = Td + Zu + \varepsilon \quad (8)$$

که در آن T ماتریس $n \times 2$ بعدی که اولین ستون دارای درایه‌های واحد و دومین ستون $t_i - 0/5$ است و $\varepsilon \sim N(0, \sigma^2 C)$ و $u_\beta \sim N(0, \sigma^2 \theta_\beta \Sigma_\beta / n\lambda)$ و $Z_{n \times np} = (I_n, \dots, I_n)$ و $D = \text{diag}(\theta_1 \Sigma_1, \dots, \theta_p \Sigma_p)$ و $\hat{u} = D\phi$ و C ماتریس همبستگی و در غیر این صورت ماتریس همبستگی C تعریف می‌شود. λ پارامتر هموارسازی اسپلاین است که در این مطالعه از روش GML به دست می‌آید. از حل معادله هارویل [۴۳] مقدار پارامترهای ϕ , d به شکل رابطه ۹ است:

$$\begin{pmatrix} T^T C^{-1} T & T^T C^{-1} Z D \\ D Z^T C^{-1} T & n\lambda D + D Z^T C^{-1} Z D \end{pmatrix} \begin{pmatrix} d \\ \phi \end{pmatrix} = \begin{pmatrix} T^T C^{-1} Y \\ D Z^T C^{-1} Y \end{pmatrix} \quad (9)$$

ماتریس Σ معکوس پذیر است؛ بنابراین، برآورد مدل آنالیز واریانس هموارسازی اسپلاین در زیر فضای H در صورتی که $\phi = Z^T c$ به صورت $\hat{u}_\beta = \theta_\beta \Sigma_\beta c$ برآورد می‌شود که برآوردگری ناریب خطی (BLUP) با بهترین پیش‌بینی است.

نمودار کنترل پیشنهادی

در این بخش، طرح پایش غیرپارامتری پروفایل در شرایط وجود همبستگی درون پروفایل‌ها ارائه می‌شود. فرض بر این است که تابع رگرسیونی تحت کنترل (g) معلوم در نظر گرفته می‌شود. در کاربرد، این تابع از مجموعه داده‌های تحت کنترل برآورد می‌شود. با توجه به اینکه مطالعه حاضر در فاز ۲ صورت می‌گیرد، پارامترهای تابع عمومی و میانگین و واریانس تابع هر پروفایل معلوم فرض می‌شود. فرض کنید z نمونه در فاز دوم جمع‌آوری شده است. مجموعه‌ای از مشاهدات به صورت $(x_i, y_{ij}), i=1, \dots, m$ وجود دارد که بردار x_i شامل n متغیر پیشگو است، $x_i = (x_{i1}, x_{i2}, \dots, x_{in})^T$ جنسن و همکاران [۲۳] نشان دادند برای بررسی تحت کنترل بودن فرایند کنترل

به‌منظور مقایسه کارایی مدل پیشنهادی، متوسط طول گردش تحت کنترل نمودار کنترل SS-ANOVA با دو نمودار MENPC و FENPC مربوط به تحقیقات کیو و همکاران [۳۱] و زو و همکاران [۴۵] ارزیابی می‌شود. بررسی متوسط طول گردش در ۱۰۰۰۰ بار تکرار مونت کارلو نتایج جدول ۱ را تأمین کرده است. همان‌گونه که ملاحظه می‌شود، در شرایطی که همبستگی درونی وجود دارد، مدل‌های آمیخته کارایی مناسب‌تری نسبت به مدل رقیب دارند.

۳. ارزیابی متوسط طول گردش در صورت اعمال شیفت

در مرحله بعد، عملکرد خارج از کنترل نمودار مطرح‌شده ارزیابی می‌شود. مشابه مطالعه کیو و همکاران [۳۱] مدل‌های خارج از کنترل به‌صورت تغییر تابع مدل عمومی پروفایل‌ها در نظر گرفته می‌شود. اگر تغییر در مؤلفه میانگین پروفایل‌ها یا به‌عبارتی تابع عمومی پروفایل‌ها در نظر گرفته شود، رابطه تحت کنترل و خارج از کنترل از طریق روابط ۱۳ و ۱۴ بیان می‌شود:

$$y_{ij} = \begin{cases} g_1(x_{ij}) + f_i(x_{ij}) + \varepsilon_{ij} \\ g_0(x_{ij}) + f_i(x_{ij}) + \varepsilon_{ij} \end{cases} \quad (13)$$

بر اساس مطالعه متوسط طول گردش خارج از کنترل، در صورتی که تغییر در مؤلفه میانگین کلی به‌صورت خطی (رابطه ۱۴) در نظر گرفته شود:

$$g_1(x) = 2\theta(x - 0.5) \quad (14)$$

در این حالت $\delta(x) = g_1(x) - g_0(x)$ یک خط مستقیم است. پارامتر θ بزرگی تغییر را کنترل می‌کند، به‌طوری‌که تغییرات خطی در شیفت‌های مختلف می‌کند. در این پژوهش $\theta = 0.2, 0.6, 1.2, 2.4$ بررسی می‌شود. فرض می‌شود مدل تحت کنترل به‌طور دقیق شناخته شده است و حد کنترل نمودار پیش در دستیابی به ARL تحت کنترل ۲۰۰ تنظیم شده است. در ادامه، رویکرد معرفی شده که مبتنی بر مدل‌سازی غیرپارامتری اثرات آمیخته است، در مقایسه با نمودارهای توسعه‌یافته کیو و همکاران [۳۱] و جنسن و برچ [22] ارزیابی می‌شود.

طول گردش در وضعیت پایدار پرداخته می‌شود. از $m = 100$ پروفایل فاز اول که مؤلفه اثر تصادفی آن دارای توزیع $b.MN_n(\cdot, \Sigma) \sim [f_i(x_{i1}), \dots, f_i(x_{im})]^T$ ، که در آن $\Sigma = (\rho_{jk})$ به‌طوری‌که $\rho_{jk} = 0.2^{|x_{ij} - x_{ik}|}$ و $j, k = 1, 2, \dots, n$ تعریف ساختار ماتریس واریانس-کواریانس به‌صورت بالا نوعی از ساختار را تولید می‌کند که ماتریس واریانس-کواریانس تابعی غیرخطی از متغیرهای مستقل است. همچنین، برای سادگی فرض می‌شود تابع عمومی پروفایل‌ها و به‌عبارت دیگر میانگین کلی صفر است ($g = 0$). نمونه تحت کنترل شبیه‌سازی شده حجم داده تقریباً بزرگی به‌صورت $n = 200, m = 500$ دارد. به‌این ترتیب، پروفایل‌هایی با ۲۰ اثر تصادفی تولید شده است.

۲. ارزیابی متوسط طول گردش تحت کنترل

متوسط طول گردش فرایند تحت کنترل حدود ۲۰۰ تعیین شده است. پروفایل‌ها با توجه به بردار اثر تصادفی به‌صورت $[f_i(x_{i1}), \dots, f_i(x_{im})]^T \sim b.MN_n(0, \Sigma)$ مقادیر b در سه سطح ۰/۲۵ و ۰/۵ و ۱ بررسی شده‌اند.

جدول ۱. نتایج مقایسه‌ای ARL خارج از کنترل در سه نمودار کنترل MENPC و FENPC و SSANOVAC

	ARL	SDRL
MENPC	۰/۲۵	۲۰۳
	۰/۵	۲۰۳
	۱	۲۰۳
FENPC	۰/۲۵	۳۴
	۰/۵	۲۰
	۱	۱۴,۲
SSANOVAC	۰/۲۵	۲۰۹
	۰/۵	۲۰۹
	۱	۲۰۹

تصادفی، مسئله همبستگی درون پروفایل‌ها و طرح اختیاری را به صورتی کارا در نظر می‌گیرد. از آماره مناسب مجزوراتی هتلینگ، برای پایش اثرهای تصادفی استفاده شد. مطالعات محاسباتی شبیه‌سازی مونت کارلو نمودار پیشنهادی و شاخص متوسط طول گردش تحت کنترل و خارج از کنترل بیان می‌کند نمودار ارائه شده عملکرد مطلوب و اثربخشی در شناسایی تغییرات پروفایل‌ها داشته است. همچنین، نمودار معرفی شده نسبت به نمودارهای رقیب برتری داشته است. نمودار پایش پروفایل پیشنهادی در مقایسه با روش‌های پارامتری معمول اثربخشی و انعطاف‌پذیری بیشتری داشته است. از آنجاکه رویکردهای شناسایی پروفایل‌های خارج از کنترل در ادبیات تحقیق مورد توجه بوده است استفاده از سایر متریک‌های جدید مانند فاصله متریک‌های کوک و ماهالانوبیس قابلیت طرح شدن در موضوعات تحقیقات آینده را دارد.

جدول ۲. نتایج مقایسه‌ای متوسط طول گردش خارج از کنترل در MENPC و FENPC و SSANOVAC

θ	MENPC	FENPC	SSANOVAC
۰/۲	۱۳۰	۱۳۹	۹۵
۰/۶	۲۰/۷	۲۸/۴	۱۳
۱/۲	۶/۶۴	۸/۴۳	۱
۲/۴	۲/۸۸	۳/۶۸	۱

نتیجه‌گیری

در این پژوهش، رویکردی به منظور پایش پروفایل‌های دارای ساختاری پیچیده و نوفه‌ای همراه با همبستگی درونی ارائه شد. نمودار کنترل پیشنهادی براساس مدل سازی غیرپارامتری اثرات آمیخته مطرح شد. مدل رگرسیونی اثرات آمیخته آنالیز واریانس هموارسازی اسپلاین، با توجه به قابلیت آن در تخمین آثار اصلی و اثر

مراجع

- Amiri, A., Jensen, W.A., Kazemzadeh, R.B., (2010) "A Case Study on Monitoring Polynomial Profiles in the Automotive Industry." *Quality and Reliability Engineering International*, 26, 2010, 509-520.
- Yeh, A.B., Huwang, L., Li, Y.M., (2009) "Profile Monitoring for a Binary Response." *IIE Transactions*, 41, pp. 931 – 941.
- Noorossana, R., Izadbakhsh, H., "Profile Monitoring for Multinomial Responses" (In Farsi). *International Journal of Industrial Engineering and Production Management*, in press.
- Noorossana, R., Eyvazian, M., Amiri, A., Mahmoud, M.A., (2010) "Statistical Monitoring of Multivariate Multiple Linear Regression Profiles in Phase I with Calibration Application." *Quality and Reliability Engineering International*, 26, pp. 291-303.
- Izadbakhsh, H., Noorossana, R., Saghaei, A., (2010) "Phase II Profile Monitoring for Ordinal Responses." *Advanced Manufacturing Technology*, (Under Review), 2012.
- Ho, L.L., El, Said M., Kim, R.W., "Monitoring the Parameters of the Market Model by Linear Profile Procedures." *Economic Quality Control*, 25, 2010, pp. 81-96.
- Zou, C., Ning, X., Tsung, F., (2010) "LASSO-Based Multivariate Linear Profile Monitoring." *Annals of Operations Research*, 192, pp. 1-17.
- Woodall, W.H., Spitzner, D.J., Montgomery, D.C., Gupta, S., (2004) "Using Control Charts to Monitor Process and Product Quality Profiles." *Journal of Quality Technology*, 36, 2004, pp. 309-320
- Stover, F. S., and Brill, R. V. (1998), "Statistical Quality Control Applied to Ion" Chromatography Calibrations, *Journal of Chromatography*, A, 804, 37-43.
- Kang, L., and Albin, S. L. (2000), "On-Line Monitoring When the Process Yields a Linear Profile." *Journal of Quality Technology*, 32, 418-426.
- Mahmoud, M. A., and Woodall, W. H. (2004). "Phase I Monitoring of Linear Profiles with Calibration Application." *Technometrics* 46. 380-391.
- Woodall, W.H. (2007), "Current Research on Profile Monitoring." *Revista Produção*, 17, 420-425.

13. Brill, R.V.,(2001) "A Case Study for Control Charting a Product Quality Measure that is a Continuous Function Over Time." Presentation at the 45th Annual Fall Technical Conference, Toronto, Ontario,
14. Walker, E., Wright, S.P., "Comparing Curves Using Additive Models." *J. Qual. Technol*, Vol. 34, 2002, pp. 118–129.
15. Jeong, M. K. , Lu, J. C. , Wang, N. (2006). "Wavelet-based SPC procedure for complicated functional data." *International Journal of Production Research*, Volume 44, Issue 4 , 729 – 744.
16. Hotelling, H.H. (1947). "Multivariate Quality Control Illustrated by the Air Testing of Sample Bombsights," *Techniques of Statistical Analysis*, 111-184.
17. Williams, J.D., Woodall, W.H., Birch, J.B., (2007) "Statistical Monitoring of Nonlinear Product and Process Quality Profiles." *Quality and Reliability Engineering International*, Vol. 23, pp. 925–941.
18. Jin, J. and Shi, J. (2001) "Automatic feature extraction of waveform signals for in-process diagnostic performance improvement." *Journal of Intelligent Manufacturing*, 12, 140–145.
19. Williams, J. D., Woodall, W. H., and Birch, J. B. (2003) "Phase I Monitoring of Nonlinear Profiles." paper presented at the 2003 Quality and Productivity Research Conference, Yorktown Heights, New York.
20. Ding, Y., Zeng, L., Zhou, S. (2006) "Phase I Analysis for Monitoring Nonlinear Profiles in Manufacturing Processes." *Journal of Quality Technology*, 38, pp. 199-216.
21. Jensen, W. A.; Birch, J. B.; and Woodall, W. H. (2006). "Profile Monitoring via Linear Mixed Models." Technical Report 06-2, Department of Statistics, Virginia Polytechnic Institute & State University.
22. Peynabar, K., and Jin, J. (2011). "Characterization of non-linear profiles variations using mixed-effect models and wavelet's." *IIE Transactions*, Volume 43, Issue 4, 275 – 290.
23. Jensen, W.A. and Birch, J.B. (2009) " Profile monitoring via nonlinear mixed models." *Journal of Quality Technology*, 41, 18–34.
24. Jensen, W.A., Birch, J.B. and Woodall, W.H. (2008) "Monitoring correlation within linear profiles using mixed models." *Journal of Quality Technology*, 40, 167–183.
25. Laird, N. M., and Ware, J. H. (1982), "Random Effects Models for Longitudinal Data." *Biometrics*, 38, 963–974. [266]
26. Diggle, P. J., Heagerty, P., Liang, K. Y. and Zeger, S. L. (2002). "Analysis of Longitudinal Data." (2nd ed). Oxford University Press, Oxford, U.K.
27. Abramovich, F. and Angelini, C. (2006), "Testing in Mixed-Effects FANOVA Models." *Journal of Statistical Planning and Inference*, 136, 4326-4348.
28. Gu, C. (2013) "Smoothing Spline ANOVA Models." New York: Springer.
29. Antoniadis, A. and Sapatinas, T. (2004). "Estimation and Inference in Functional Mixed-effects Models." Technical Report, Department of Mathematics and Statistics, University of Cyprus, Nicosia
30. Morris, J. S. and Carroll, R. (2006). "Wavelet-based Functional Mixed Models." *Journal of the Royal Statistical Society: Series B*, 68(2), 179-199.
31. Morris, J. S., Arroyo, C., Coull, B. A., Ryan, L. M. and Gortmaker, S. L. (2006). "Using Wavelet-Based Functional Mixed Models to characterize Population Heterogeneity in Accelerometer Profiles: A Case Study." *Journal of the American Statistical Association*, 101(476), 1352-1364.
32. Qiu, P., Zou, C., & Wang, Z. (2010). "Nonparametric profile monitoring by mixed effects modeling." (with discussions). *Technometrics*, 52, 265–277.
33. Qiu, P., & Zou, C. (2010). "Control chart for monitoring nonparametric profiles with arbitrary design." *Statistica Sinica*, 1655–1682
34. Shi, M., Weiss R. E., and Taylor, J. M. G. (1996), "An Analysis of Paediatric CD4 Counts for Acquired Immune Deficiency Syndrome Using Flexible Random Curves." *Applied Statistics*, 45, 151-163.
35. Rice, J. A. and Wu, C. O., (March 2001), "Nonparametric Mixed Effects Models for Unequally Sampled

- Noisy Curves.” *Biometrics*, 57, 253-259.
36. Abdel-Salam, A.-S., and Birch, J. B. (2009), “Profile-Monitoring Analysis With Fixed and Random Effects Using Nonparametric and Semiparametric Methods.” contributed paper presented at the Joint Statistical Meetings, August 2009, Washington, DC. [286,287]
 37. Wahba, G. (1990) “Spline Models for Observational Data.” Philadelphia: Society for Industrial and Applied Mathematics.
 38. Chen, Z. (1993) Fitting multivariate regression Journal of the Royal Statistical Society. Series B (Methodological), Vol. 55, No. 2, pp. 473-491
 39. Demidenko, E. (2004), “Mixed Models: Theory and Applications”. New York: Wiley.
 40. Gu, C., Bates, D. M., Chen, Z. and Wahba, G. (1989) “The computation of GCV functions through Householdertridiagonalization with application to the fitting of interaction spline models.” *SIAM J. Matrix Anal. Applic.*, 10, 457-480
 41. Wahba, G. (1978) “Improper priors, spline smoothing and the problem of guarding against model errors in regression.” *J. R. Statist. Soc. B*, 40, 364-372.
 42. Aronszajn, N. (1950) “Theory of reproducing kernels.” *Trans. Am. Math. Soc.*, 68, 337-404.
 43. Kimeldorf, G. S. and Wahba, G. (1971). “Some results on Tchebycheffian spline functions.” *J. Math. Anal. Appl.* 33 82-94.
 44. Harville, D. (1976). “Extension of the Gauss-Markov theorem to include the estimation of random effects.” *Ann. Statist.* 4 384-395.
 45. Hawkins, D. M., and Olwell, D. H. (1998), “Cumulative Sum Charts and Charting for Quality Improvement.” New York: Springer-Verlag.
 46. Zou, C., Tsung, F. and Wang, Z. (2008) “Monitoring profiles based on non-parametric regression methods.” *Technometrics*, 50, 512-526

واژگان انگلیسی به ترتیب استفاده در متن

1. Average Run Length
2. High Dimensionality
3. Gu
4. Functional data
5. Tin plate
6. General Maximum Likelihood
7. General Cross Validation
8. Balanced data
9. Best Linear Unbiased Prediction