



ارائه شده توسط:

سایت ترجمه فا

مرجع جدیدترین مقالات ترجمه شده

از نشریات معتبر



ارائه شده توسط:

سایت ترجمه فا

مرجع جدیدترین مقالات ترجمه شده

از نشریات معتبر

صف بندی داده‌های دسته‌ای ناجور با کاربردی برای کنترل کمیت محصول

انتهایی دسته

خلاصه

فرایندهای دسته‌ای بطور عمومی توسط مسیرهای غیر زوج توصیف می‌شود که بدلیل وجود متغیرهای دسته به دسته‌ای می‌باشد. مقدار محصول نهایی دسته‌ای معمولاً در پایان این مسیر غیر زوج (uneven) اندازه‌گیری می‌شود. معمولاً به صف کردن تفاوت‌های زمانی لازم است برای مسیرهای اندازه‌گیری شده و مقدار محصول انتهایی دسته به منظور ردیابی اجراء فرایند آماری و طرح‌های کنترل می‌باشد. صرف نظر از مسیرهای همزمان‌سازی با طول‌های متغیر با استفاده از یک متغیر شاخص یا زمان دینامیکی منحرف کننده، این مقاله، یک روش تازه برای صف‌بندی داده‌های دسته‌ای غیر زوج، توسط شناختن پنجره‌ی کوتاه مدل‌های PCA&PLC در اولین و سپس اعمال این مدل‌های شناخته شده برای مسیرهای کوتاه‌تر و پیش‌بینی مقدار محصول انتهایی دسته‌ای را پیشنهاد می‌کند. بعلاوه داده‌های دسته‌ای غیر زوج همچنین می‌توانند برای طول دسته‌ی تعیین شده با استفاده از حرکت تخمین پنجره، صف‌بندی شوند. روش پیشنهادی و کاربرد آن برای کنترل مقدار محصول انتهایی دسته‌ای با یک مثال شبیه‌سازی از تخمیر دسته‌ی غذا برای محصولات پنی‌سیلین را نشان می‌دهد

مقدمه

فرایند دسته‌ای/نیمه دسته‌ای نقش قابل توجهی را در تولید حجم کم، محصولات مقدار اضافه‌شده‌ی زیاد مانند پلیمرهای اختصاصی، داروسازی‌ها و مواد شیمیایی خوب دارد. این مورد بطور اساسی بدلیل انعطاف‌پذیری آن و مقیاس بالای فرایندهای آزمایشگاهی می‌باشد.

به منظور بدست آوردن شیوه‌های دسته‌ای مطلوب و سازگار، ردیابی موثر و کنترل این فرایندها در طول هر مرحله از عملکرد، حیاتی می‌باشند. برای مثال، بازیابی قبلی از شرایط غیر عادی از یک فرایند نه تنها انرژی و مواد نارس را ذخیره می‌کند بلکه همچنین یافتن علت و جبران آن را با یک استراتژی کنترل مناسب، ممکن می‌سازد.

فرایند ردیابی و تکنیک‌های کنترلی بطور وسیعی در نوشتیجات، نظارت شده است. بدلیل پیچیدگی فرایندهای دسته‌ای، معمولاً گسترش مدل‌های مکانیکی بر پایه‌ی اصول شیمی فیزیک مشکل می‌باشد به هر حال، یک مقدار کثیری از متغیرهای فرایند مانند دما، فشار و سرعت جریان، اغلب بطور روزمره در ساخت ماشین‌الات‌های مدرن، اندازه‌گیری می‌شود. بنابراین، روش‌های کنترل فرایند آماری چند متغیره (MSPC) که بر پایه‌ی اطلاعات تاریخی فرایند می‌باشند، مطلوب و یا روش‌های استاندارد زوج برای فرایند ردیابی و کنترل، می‌شوند. در میان آنها، روش چندگانه‌ی آنالیز اجزاء اصلی (PCA) و روش چندگانه‌ی ساختارهای پوشیده‌ی طرحی (PLS) که توسعه یافته‌های PCA و PLS برای بکار بردن ماتریس‌های سه بعدی می‌باشند که بطور وسیعی استفاده می‌شوند.

فرض پایه‌ای متضمن، روش چندگانه‌ی PCA و روش چندگانه‌ی PLS برای کنترل فرایندها و ردیابی می‌باشد که تمام مدت‌های زمانی دسته‌ای مشابه می‌باشند و مسیرهای دسته‌ای مناسب، همزمان می‌شوند برای اطمینان از اینکه پیشامدهای مشابه در دوره‌های زمانی مشابه، رخ داده است. به هر حال، عملکرد صنعتی، طول دسته‌ی بسیاری از فرایندها به ندرت ثابت می‌شود بعنوان معیاری برای انتهای یک دسته که معمولاً برای روبرو شدن با بعضی از نیازمندی‌های کیفیتی نسبت به زمان می‌باشد. اختلات و تغییرات در شرایط عملیاتی و مواد نارس می‌توانند به آسانی منجر به طول‌های دسته‌ای غیر زوج شود. در یک رفتار، برای روبرو شدن با احتیاجات کیفیتی در پایان اجرای دسته‌ای شوند. بنابراین آن برای صف‌بندی این داده‌های دسته‌ای غیر زوج برای استفاده آنها در فرایند کنترل و ردیابی، لازم است. الگوهای پتانسیلی مسیرهای دسته‌ای غیر همزمان و روش‌های ممکن برای همزمان شدن آنها در [10] توضیح داده می‌شود. ساده‌ترین حالت زمانی پیش می‌آید که مسیرهایی با طول‌های متفاوت در قسمت‌های زمانی مشترک، همپوشانی می‌کنند، مسیرهای دسته‌ای هم می‌توانند برای طول حداقل از تمام مسیرها یا برای طولانی‌ترین طول از همه‌ی مسیرها، توسط برآورد قسمت‌های مفقوده از مسیرهای کوچکتر با استفاده از الگوریتم‌های داده‌های مفقوده،

بریده شوند. خصیصه‌ی آخری، نیازمند دسته‌هایی با اندازه‌ی کافی طویل برای شناخت مدل در تخمین داده‌ی مفقودی می‌باشد. روش دیگر یافت یک متغیر شاخص مناسب می‌باشد که می‌تواند برای جایگزینی بعد زمانی برای مسیرهای دسته‌ایی هم زمان با طول متغیر می‌باشد.

مسیرهای دسته‌ایی با طول‌های متغیر همچنین می‌توانند از طریق تفسیر قطعات موضعی شده‌ی متراکم و گسترده‌ی داخل آنها، هم زمان شوند. روش‌های معمول برای چنین همزمانی شامل ارتباط بهینه شده‌ی منحرف کننده (COW) و زمان دینامیکی منحرف کننده (DTW) می‌باشد. توانایی آنها در صف‌بندی پروفایل‌های اسپکتروسکوپی و کروماتوگرافی در [17,18] بیشتر مقایسه شده است. روش COW بطور اساسی، برای جابجایی پیک صحیح در پروفایل‌های کروماتوگرافی، طراحی شده است و آن مقداری محدودیت برای اصلاح خط پایه از طریق ضریب اصلاح دارد که یک اندازه‌گیری مشابه را پیشنهاد می‌کند. کاربرد COW در کنترل فرایند دسته‌ایی on-line می‌تواند در [20] یافت شود. روش DTW که بطور وسیعی، گسترش یافته است برای جور بودن با رخداد‌های مشابه بین سیگنال-ها، در همزمانی مسیرهای دسته‌ایی، کاربرد وسیعتری دارد. DTW، اصولی از برنامه‌ی دینامیکی برای حداقل کردن اندازه‌گیری غیر مشابه با یک فاصله بین دو مسیر را استفاده می‌کند و آن ممکن است بعضی از خصوصیات بردارها در زمان، بعضی فشردگی‌ها و گستردگی‌ها را جابه‌جا کند برای اینکه یک فاصله‌ی حداقل بدست آورده شود. DTW اولین بار در کروماتوگرافی، بعنوان یک ابزار نظارت و آشکارسازی نادرستی با مرجع ویژه برای کاربردهای فرایند زیستی، ارائه شد. سپس Kassidas و همکاران، الگوریتم DTW را برای کنترل فرایند آماری دسته‌ایی، بعنوان یک ابزار برای همزمانی دسته‌ها با طول‌های متفاوت و یا زمان‌های محلی متفاوت، پذیرفتند. پس از آن نسخه‌های اصلاح شده از DTW در نوشتیجات برای همزمانی مسیر دسته‌ایی، پیشنهاد شدند. برای مثال، یک الگوریتم DTW تن‌مند در [23] از طریق ترکیب یک روش کمترین مربعات پنجره‌ی حرکتی با مشتق DTW برای پرهیز از نقاط انفرادی و کاهش تمایل هم ترازى نسبت به مسیر مرجع، پیشنهاد شد. به تازگی، یک انطباق از روش Kassidas و همکاران در [9] با دستاورد همزمان on-line مسیرهای دسته‌ایی، ارائه شده است.

در هر دو مسیر با طول‌های متفاوت، DTW می‌تواند برای همزمانی آنها بکار برده شود چنانچه آنجا معمولاً یم راه‌حل بهینه برای حداقل فاصله بین این دو مسیر بر طبق معیارهای خاص وجود دارد. چنین مکانیسم همزمانی، دینامیکی فرایند را کوچک در نظر می‌گیرد. بعنوان نتیجه DTW ممکن است دینامیک‌های فرایند ذاتی را در طول تنظیم قطعات متمرکز شده، منحرف کند و بنابراین توانایی برای آشکارسازی نادرستی و تشخیص با استفاده از داده‌های همزمانی می‌تواند در بعضی مواقع کاهش یابد. مسیرهای متغیر فرایند، اندازه‌گیری کیفیت فرایند پایانی دسته در پایان اجراء دسته، برای صف‌بندی، بعلاوه بدلیل تغییر مدت زمان دسته برای دسته‌های کوتاهتر، نیاز می‌باشد. بعد از همزمانی برای روش متغیر شاخص یا روش DTW، کیفیت محصول انتهایی دسته‌ی بعدی در مدت زمان دسته‌ی همزمان شده برای دسته‌های کوتاهتر می‌تواند به سادگی برای مقادیر مشابه در نقاط پایانی اصلی‌شان، فرض شود. به هر حال، این مورد بدلیل تغییر مدت زمان‌ها برای آن دسته‌های کوتاهتر، بخصوص در حالتی از یک تغییر بزرگ در مدت زمان‌های دسته، ممکن نیست درست باشد. یک روش پی‌درپی، برای پیش‌بینی مقادیر کیفیت محصول انتهایی دسته‌ی بعدی در طول زمان دسته‌ی همزمان شده برای دسته‌های کوتاهتر با استفاده از یک مدل از پیش تعیین شده، می‌باشد. چنین مدلی می‌تواند با استفاده از مدل پیشنهاد شده در [26] شناخته شود جایی که یک سری از دسته‌های شبیه خلق شده با نقاط پایانی دسته‌شان، همزمان می‌شوند و مدل PLS از داده‌های دسته‌ی شبیه همزمان شده شناخته می‌شود. چنین داده‌های دسته‌ی شبیه از پنجره‌های مختلفی از اجزاء دسته‌ی نی‌آیند، مدل PLS شناخته شده، بطور ضروری یک مدل PLS پنجره‌ی حرکتی، می‌باشد. آن می‌تواند برای پیش‌بینی کیفیت محصول انتهایی دسته‌ی بعدی، استفاده شود تا زمانی که دینامیک فرایند برای دوره‌ی زمانی آن با پوشش پنجره‌های حرکتی، تغییر نکند.

با استفاده از روش شناسایی در [26]، این مقاله، یک روش دو مرحله‌ای را برای بصف کردن داده‌های دسته‌ی غیرزوج را پیشنهاد می‌کند. اولین مرحله شناخت مدل PCA&PLS پنجره‌ی کوتاه توسط صف‌بندی دسته‌هایی با طول‌های غیر زوج برای نقاط انتهایی آنها می‌باشد. و دومین مرحله اعمال مدل PCA شناخته شده برای تخمین مسیرهای مفقود می‌باشد. اعمال مدل PLS شناخته شده، برای پیش‌بینی کیفیت محصول انتهایی دسته‌ی بعدی برای دسته‌های کوتاهتر می‌باشد. مقاله سازمان داده می‌شود چنانچه که در ادامه می‌آید. مدل‌های PCA&PLS در این مقاله، بطور

مختصر در بخش 2 ارائه می‌شود. روش‌های پیشنهادی سپس در بخش 3 ارائه و تفسیر می‌شوند. مطالعه‌ی حالتی از اعمال روش پیشنهادی برای صف‌بندی داده‌های دسته‌ای غیر زوج و کنترل کیفیت محصول انتهایی با انتخاباتی از طول‌های دسته‌ی متغیر برای تخمیر دسته‌ی غذا از پنی‌سیلین در بخش 4 داده می‌شود و بعضی از نتایج در بخش 5 آورده می‌شود.

2. مقدمات

برای روش‌های چندگانه PLS متغیرهای فرایند به دو گروه تقسیم می‌شوند: یکی مقادیر پیش‌بینی کننده مانند مسیرهای متغیر فرایند اندازه‌گیری شده را عهده‌دار می‌شود و دیگری مقادیر پاسخی مانند متغیرهای کیفیتی محصول انتهایی دسته‌ی اندازه‌گیری شده را به عهده می‌گیرد. هر گروه از داده‌ها بطور اصلی، یک ماتریس سه بعدی با اندازه‌ی $I \times J \times K$ می‌باشد، جایی که I تعداد دسته‌هایی با داده‌ی در دسترس می‌باشد J تعداد متغیرهایی است که اندازه‌گیری می‌شود و K تعداد نمونه‌های جمع‌آوری شده در طول اجراء دسته می‌باشد. اگرچه در اینجا احتمالات مختلفی برای آشکار شدن داده‌ها وجود دارد، روش آشکار شدن طریقه‌ی دسته‌ای، روش بسیار منطقی برای مدل متفاوت در میان دسته‌ها می‌باشد. داده‌های آشکار شده بیشتر مرکز میانه می‌باشند و برای واحد شدن واریانس مدرج می‌شوند و PLS انجام دهنده بر روی داده‌ها بدست آمده در یک مدل پنهانی از شکل را نتیجه می‌دهد:

$$\mathbf{X} = \mathbf{T}_1 \mathbf{P}_1^T + \mathbf{E}_1, \quad (1)$$

$$\mathbf{Y} = \mathbf{U}_1 \mathbf{Q}_1^T + \mathbf{F}_1, \quad (2)$$

جایی که \mathbf{X} ماتریس $I \times J \times K_x$ برای متغیرهای پیش‌بینی کننده می‌باشد، \mathbf{Y} یک ماتریس $I \times J_y \times K_y$ برای متغیرهای پاسخی می‌باشد و \mathbf{P}_1 و \mathbf{Q}_1 ماتریس‌های بارگذاری کننده می‌باشند. در اینجا A تعداد متغیرهای پنهانی می‌باشد. مشابه‌ی \mathbf{T}_1 و \mathbf{U}_1 با یک ماتریس قطری \mathbf{B}_1 با $\mathbf{T}_1 = \mathbf{XW}_1(\mathbf{P}_1^T \mathbf{W}_1)^{-1}$ و $\mathbf{U}_1 = \mathbf{T}_1 \mathbf{B}_1$ مربوط می‌شوند جایی که \mathbf{W}_1 ماتریس وزنی می‌باشد. سرانجام \mathbf{E}_1 و \mathbf{F}_1 ماتریس‌های باقیمانده می‌باشند. روش چندگانه‌ی شناخته‌شده‌ی مدل PLS می‌تواند برای کنترل کیفیت محصول انتهایی دسته از طریق بکاربری مسیر، اعمال شود. بخصوص

اگر آنجا هیچ اندازه‌گیری برای متغیرهای کیفیتی محصول انتهایی نباشد یا متغیرهای پاسخی در نظر گرفته نشده‌اند، روش‌های چندگانه‌ی PCA می‌تواند بجای یک مدل ساختار همبستگی از تمام متغیرهای فرایند اندازه‌گیری شده، اعمال شود:

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E}, \quad (3)$$

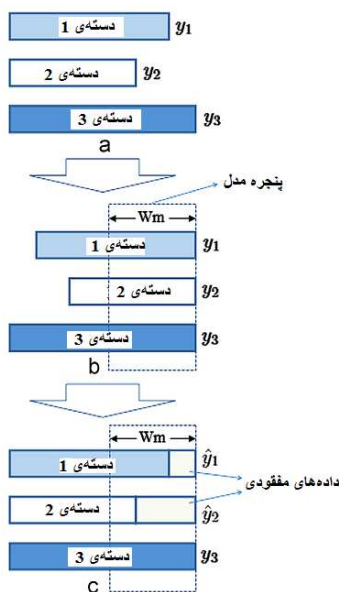
جایی که \mathbf{P} ، \mathbf{T} و \mathbf{E} ماتریس‌های متناظر برای مدل PCA می‌باشند.

3. روش هم‌ترازی داده‌های دسته‌ی غیرزوج (uneven)

فرض بحرانی زمانی ایجاد می‌شود که آشکار شدن طریقه‌ی دسته‌ی انجام‌دهنده، می‌باشد که همه‌ی دسته، طول مشابه دارند. این فرض اطمینان می‌دهد از اینکه تعداد ستون‌ها در ماتریس داده‌های آشکار شده، برای هر ردیف مشابه می‌باشد که اجازه می‌دهد تکنیک‌های PCA و PLS به آسانی اعمال شوند. برای مسیرهای دسته‌ی غیرزوج، در ابتدا به صف کردن داده‌های دسته‌ی غیرزوج و سپس شناخت مدل‌های PCA&PLS برای فرایند ردیابی و کنترل لازم می‌باشد.

اصول روش هم‌ترازی داده‌های دسته‌ی غیرزوج پیشنهاد شده، با استفاده از یک مثال نشان داده شده در شکل 1 (a) توضیح داده می‌شود که دسته با طول متفاوت را نشان می‌دهد. متغیرهای برای کیفیت محصول انتهایی دسته، تنها در پایان اجراء هر دسته یعنی y_1, y_2 و y_3 برای هر سه دسته در مثال، اندازه‌گیری می‌شود. مشابه دسته‌های شبیه در [26]، ابتدا به سمت نقاط پایانی، صف‌بندی می‌شود و سپس مدل پنجره برای شناخت مدل‌های PCA و PLS روش چندگانه برای پنجره‌ی انتخاب شده، انتخاب می‌شود، چنانچه که در شکل 1 (b) نشان داده شده است. طول پنجره‌ی انتخاب شده بعنوان W_m مشخص می‌شود و مدل‌های PCA&PLS پنجره‌ی کوتاه برای به ترتیب W_m -PCA و W_m -PLS اعمال می‌شود. با ارزش است توجه به اینکه ساختار داده‌های مدل‌های W_m -PCA و W_m -PLS برای متغیرهای پیش‌بینی کننده، می‌تواند بعنوان یک حالت خاصی از روش متغیر با شیوه‌ی تاخیر زمان، توجه شود جایی که تمها ردیف آخر داده‌های آشکار شده برای هر دسته‌ای، برای شناخت مدل‌های PAC&PLS انتخاب می‌شود و بنابراین شناخت مدل‌های PAC&PLS برای ارتباط نزدیک فرایند با دینامیک‌های فرایند در پایان مرحله‌ی

پایانی اجراء دسته، پیش‌بینی می‌شوند. مدل‌های PCA&PLS شناخته شده، بصورت ضروری مدل‌های پنجره‌ی حرکتی می‌باشند. چنین روش شناسایی مدل معتبر است تا زمانی که دینامیک‌های فرایند در دوره‌های زمانی آن پوشش پنجره‌های حرکتی، تغییر نکند. این حالتی برای بیشتر فرایندهای فاز انفرادی می‌باشد که می‌تواند توسط یک مدل PCA/PLS مدل شود.



شکل 1 هم‌ترازی داده‌های دسته‌ی غیرزوج (uneven)

سابقاً مدل‌های W_m -PCA و W_m -PLS برای پنجره‌ی مدل انتخاب شده، شناخته می‌شوند، آنها برای طویل شدن مسیرهای کوتاه‌تر با استفاده از الگوریتم‌های داده‌های مفقود و برای پیش‌بینی کیفیت محصول نهایی دسته، برای آن دسته‌ای کوتاه‌تر، استفاده می‌شوند. چنین آرایه‌ای استفاده‌ی کامل از مدل W_m -PCA شناخته شده را برای تخمین و استفاده‌ی کامل از PLS شناخته شده را برای پیش‌بینی، هدف‌گیری می‌کند. همانطور که در شکل (a) نشان داده شده است دسته‌ی 1 و 2، دسته‌های کوتاه‌تری نسبت به دسته‌ی 3 می‌باشند و بنابراین داده‌ی مفقودی در دسته‌ی 1 و 2 می‌تواند توسط جای دادن مدل W_m -PCA شناخته شده برای یک موقعیتی که تمام قسمت مفقوده را پوشش دهد که در شکل 1 (c) مشاهده می‌شود. چندین روش اسنادی داده‌ی مفقودی در نوشتیجات، پیشنهاد شده است. ایده‌ی رایج آنها، استفاده از الگوی داده‌ی اصولی برای استنباط قسمت مفقودی از قسمت شناخته شده می‌باشد. پنداشتن الگوریتم مفقودی، تجسم با نقشه نامیده می‌شود. بعنوان یک مثال، اصول آن می‌تواند بطور

اختصار در ادامه توضیح داده شود. برای دسته‌ی آن مقادیر پیش‌بینی کننده‌ی X_i می‌تواند جزءبندی شود چنانکه $X_i = [X_i^* \hat{X}_i]$ جایی که X_i^* داده‌ی اندازه‌گیری شده را نشان می‌دهد در حالی که \hat{X}_i داده‌ی مفقودی را نمایش می‌دهد. ماتریس بارگذاری P از مدل Wm-PCA می‌تواند جزءبندی شود، بنابراین چنانکه $P = [P^* \hat{P}]$ جایی که P^* و \hat{P} به ترتیب با X_i^* و \hat{X}_i مطابقت می‌کنند. در ابتدا، برداری نشان $\tau = X_i^* P$ توسط X_i^* و \hat{X}_i تعیین می‌شود. سپس بردار علامت τ ، داده‌های اندازه‌گیری شده و داده‌های مفقودی را مربوط می‌کند. آن برای این است که داده‌های اندازه‌گیری شده، بهینه شده و بهتر جور شوند و بردار علامت بهینه شده $\hat{\tau}$ همچنین باید کاندید قابل دسترس بهتری برای جور شدن داده‌های مفقودی شود بنابراین داده‌های مفقودی می‌توانند از بردار نشان‌ی بهینه $\hat{\tau}$ نتیجه شوند که از حداقل توابع هدف در زیر بدست می‌آیند.

$$J = \frac{1}{2} (X_i^* - \tau P^{*T}) (X_i^* - \tau P^{*T})^T. \quad (4)$$

آن می‌تواند مشاهده شود که تابع هدف، برای جور کردن داده‌های شناخته شده با استفاده از بردار نشان‌ی τ برای بهینه شدن، هدف‌گیری می‌کند. بردار نشان‌ی بهینه می‌تواند بزرگ‌ترین و تحلیل‌ی توسط مشتق‌گیری نسبت به τ برای تابع هدف و قرار دادن آن برابر صفر، می‌تواند بدست آورده شود.

$$\frac{dJ}{d\tau} = -P^{*T} (X_i^* - \tau P^{*T})^T = 0, \quad (5)$$

$$\hat{\tau} = X_i^* P^* (P^{*T} P^*)^{-1}. \quad (6)$$

وقتی ماتریس $P^{*T} P^*$ بد حالت است، محاسبه‌ی $\hat{\tau}$ فرایند بهینه شده، توصیه می‌شود. سابقاً بردار علامت بهینه، بدست آورده می‌شود، \hat{X}_i می‌تواند محاسبه شود چنانکه در زیر آمده است:

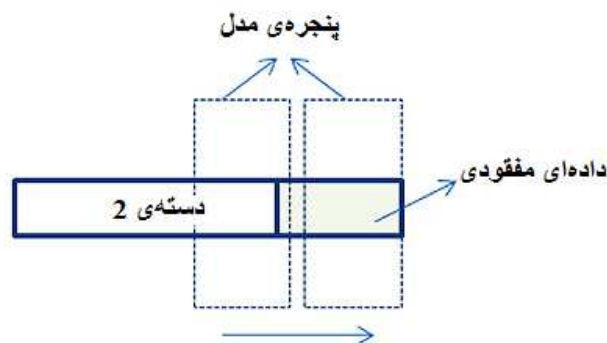
$$\hat{X}_i = \hat{\tau} \hat{P}^T. \quad (7)$$

کیفیت محصول انتهایی دسته‌ی بعدی y برای دسته‌های کوتاه‌تر می‌تواند با استفاده از مدل Wm-PLS شناخته شده، پیش‌بینی شود. مدل Wm-PLS در موقعیت مشابه مدل Wm-PCA قرار می‌گیرد و آن، داده‌های مفقودی تخمین زده شده‌ی \hat{X}_i از مدل Wm-PCA را برای پیش‌بینی کیفیت محصول انتهایی دسته‌ی بعدی، استفاده می‌کند:

$$\hat{y}_i = [X_i^* \hat{X}_i] W_1 (P_1^T W_1)^{-1} B_1 Q_1^T \quad (8)$$

در این روش کیفیت محصول انتهایی دسته، می‌تواند برای دسته‌های کوتاه‌تر، مشابه کیفیت محصول انتهایی دسته‌ی به روز شده‌ی y_1 و y_2 ، بروز شود، چنانکه در شکل 1 (C) مشاهده می‌شود. طول پنجره‌ی مدل برای مدل‌های Wm-PCA&WmPLS می‌تواند برای انعکاس دینامیک‌های موقعیتی فرایند، نسبتاً کوچک باشد، سپس تخمین داده‌های مفقودی می‌تواند در یک روش برگشتی، انجام شود چنانکه در شکل 2 مشاهده می‌شود. اگرچه پنجره‌ی مدل ممکن نیست در طول تمام قسمت‌های داده‌ی مفقودی پوشیده شود، آن می‌تواند بطور آغازی برای پوشش تنها قسمتی از داده‌های مفقودی، قرار گیرد و سپس برای تخمین تمام داده‌های مفقودی بطور برگشتی، بسمت جلو حرکت کند. چنین روشی می‌تواند بعنوان یک برآورد کننده‌ی پنجره‌ی حرکتی برای داده‌های مفقودی، نسبت داده شود. یک مزیت چنین روشی امکان استفاده‌ی بیشتر از داده‌های شناخته شده برای تخمین داده‌های مفقودی می‌باشد. در عمل، طول پنجره‌ی مدل می‌تواند بطور ابتکاری با توجه به دینامیک‌های موقعیتی برای تخمین دقت و پیش‌بینی با استفاده از مدل‌های Wm-PCA&Wm-PLS، انتخاب شود.

یک فرض کلی برای استفاده از مدل Wm-PCA شناخته شده برای تخمین داده‌های مفقودی و استفاده از مدل WM-PLS شناخته شده برای پیش‌بینی کیفیت محصول انتهایی دسته‌ی بعدی، می‌باشد که مدل‌های Wm-PCA&Wm-PLS معتبر می‌باشند برای تمام دوره‌ی زمانی که پنجره‌ی مدل پوشیده می‌شود. تحت چنین فرضی، داده‌های دسته می‌توانند عملاً برای طول دسته‌ی ویژه‌ای بین کوتاه‌ترین و بلندترین دسته‌ها با استفاده از تخمین پنجره‌ی حرکتی، صف‌بندی شوند.



شکل 2. تخمین داده‌های مفقودی بطور برگشتی

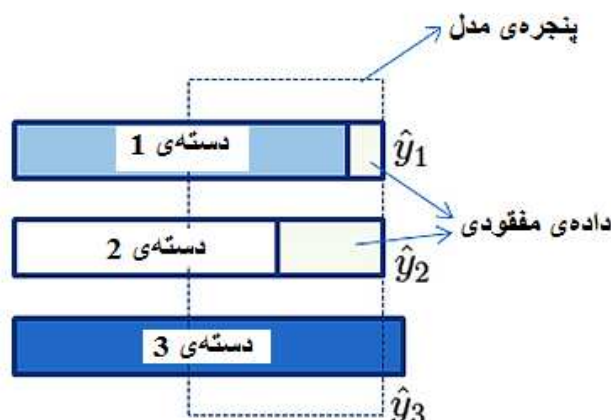
چنانچه که در شکل 3 نشان داده شده است پنجره‌ی مدل در یک موقعیتی بین کوتاهترین و بلندترین دسته‌ها برای صف‌بندی داده‌های دسته، حرکت داده می‌شود تا یک طول خاصی بین کوتاهترین و بلندترین دسته‌ها، ایجاد شود. سپس داده‌های مفقودی برای دسته‌ی 1 و 2 می‌توانند با استفاده از مدل Wm-PCA تخمین زده شود و کیفیت محصول انتهایی دسته برای تمام این سه دسته می‌تواند با استفاده از مدل Wm-PLS شناخته شده، پیش‌بینی شود. برای دسته‌ی بدون داده‌ی مفقودی مانند دسته‌ی 3 در شکل 3، کیفیت محصول انتهایی دسته در زمان انتهایی دسته‌ی فرض شده، می‌تواند به سادگی با استفاده از تمام داده‌های شناخته شده از متغیرهای پیش‌بینی کننده، پیش‌بینی شود.

$$\hat{y}_i = X_i^* W_1 (P_1^T W_1)^{-1} B_1 Q_1^T \quad (9)$$

بنابراین تمام داده‌های دسته می‌توانند برای داشتن طول مشابه با کیفیت محصول انتهایی دسته‌ی بروز شده در نقاط انتهایی دسته‌ی فرضی، کاملاً به صف شوند. برای مثال پنانکه در شکل 3 نشان داده شده است مقادیر کیفیت محصول انتهایی دسته‌ی به روز شده \hat{y}_1, \hat{y}_2 و \hat{y}_3 می‌باشند. با استفاده از داده‌های دسته‌ی ردیف شده با طول دسته‌ی ویژه، مدل‌های PCA&PLS جدید می‌توانند نتایجاً برای کنترل و ردیابی فرایند، شناخته شوند. توانایی طول دسته‌ی خاص برای داده‌های دسته‌ی غیر زوج ردیف، می‌تواند برای کنترل کیفیت محصول انتهایی دسته مفید باشد چنانکه انتخاب‌های قابل انعطاف برای طول‌های دسته موجود باشد و یک استراتژی کنترل بهینه از این انتخاب‌ها می‌تواند با توافق طول‌های ادامه‌یافته‌ی دسته و تلاش‌های کنترلی قابل دسترس، انتخاب شود. با این حال، طول‌های دسته نمی‌توانند دلخواهانه برای بعضی از اجرای کوتاهتر، طویل شود چنانکه عملکرد ممکن است به آسانی به لحاظات بحرانی مانند Yield حداکثر محصول، نائل شود. در چنین حالتی، اجرای ادامه یافته، ممکن است برای محصول زیان‌آور باشد و مدل‌های PCA&PLS پنجره‌ی کوتاه شناخته شده برای دوره‌ی زمانی طویل شده بدلیل تغییر دینامیک‌های فرایند، غیر معتبر می‌شوند. بنابراین بسط چنین دسته‌های کوتاهتری باید از روش پیشنهاد شده محروم شود.

4. مطالعه‌ی حالت

به منظور شناسایی و معتبر ساختن داده‌های دسته‌ی غیر زوج بالا، یک شبیه‌ساز ارزیابی مقایسه‌ایی برای فرایند تخمیر دسته‌ی خوراکی پنی‌سیلین، استفاده می‌شود. شبیه‌ساز، Pensim نامیده شده، بر پایه‌ی یک سری مدل‌های مکانیکی مفصل می‌باشد که فرایند تخمیر را توصیف می‌کند. متغیرهای فرایند به شرح زیر در طول ساعت‌ها فرایند تخمیر، جمع‌آوری می‌شود: سرعت هوا دهی، قدرت هم‌زنی، دمای خوراک سابستریت، شرایط سابستریت، غلظت اکسیژن حل شده، حجم culture (کشت میکروب در آزمایشگاه)، غلظت مونوکسید کربن، pH، دمای ظرف تخمیر و سرعت خوراک سابستریت. سرعت خوراک سابستریت، متغیر فرایند دستکاری شده می‌باشد. در حالی که کیفیت محصول انتهایی دسته، اندازه‌گیری غلظت جرم زیستی در پایان اجرای دسته می‌باشد. به استثنای پروفایل‌های سرعت خوراک سابستریت مختلف برای اجرای دسته، فرایند تخمیر دسته‌ی خوراک، همچنین تحت معرض اختلالات سرعت هوا دهی، قدرت هم‌زنی، سرعت خوراک سابستریت و دمای خوراک سابستریت می‌باشد. بعلاوه غلظت محلول سابستریت خوراکی در حال نوسان اطراف حجم ثابتی از 600g/l بعنوان یک نتیجه از تغییرپذیری در خواص مواد نارس می‌باشد. تمام این اختلالات یا نویز در شرایط عملی و مواد نارس برای طول‌های دسته‌ی مختلف در عملکرد، سهم دارند. فرض می‌شود که غلظت جرم زیستی هدف در پایان اجرای دسته 12g/l می‌باشد و نمونه‌ها برای سنجش آزمایشگاهی در 160 مین و 180 مین ساعت برای دیدن بدست آورده می‌شوند اگر هدف یافت شود. اگر هدف یافت شود اجرای دسته فوراً متوقف می‌شود. در غیر این صورت، دسته به صورت مداوم برای تمام مدت 200 ساعت اجراء می‌شود. بنابر این بطور محتمل دسته‌ها به ترتیب طول 160، 180 و 200 ساعت را دارند. سایر معیارها برای انتهای اجرای یک دسته می‌تواند همچنین برای تولید اجرای دسته با طول‌های متغیر، اجرا شود.



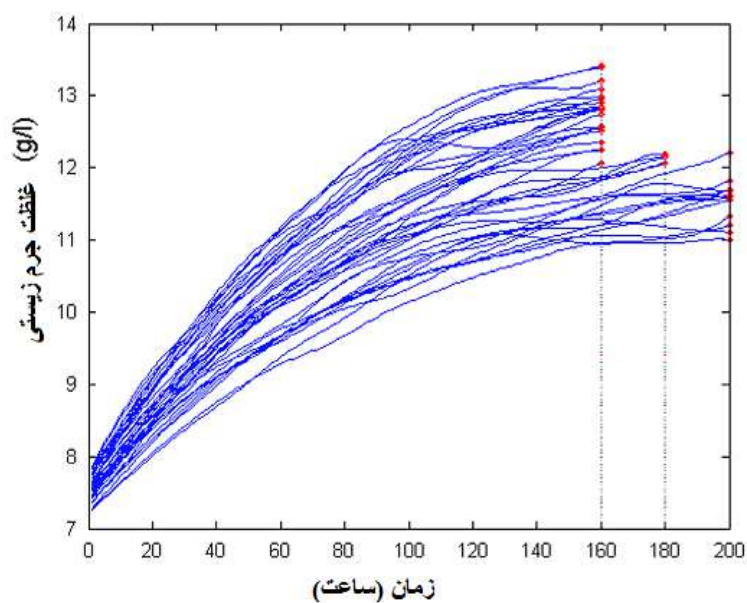
شکل 3. صف‌بندی داده‌های دسته‌ی غیر زوج (uneven) برای داشتن طول خاص بین کوتاهترین و بلندترین

دسته‌ها

40 دسته با طول‌های متغیر بعنوان یک مثال برای توضیح روش هم‌ترازی داده‌های غیرزوج، مسیرهای غلظت جرم زیستی، برای 40 دسته بدست آمده از شبیه‌سازی در شکل 4 رسم می‌شوند. می‌تواند دیده شود که این دسته‌ها، طول‌های دسته‌ای متغیر می‌باشند و بلندترین دسته 200 ساعت طول می‌کشد. توجه کنید مه این غلظت جرم زیستی فرض می‌شود اندازه‌گیری نشده اما در اینجا برای هدفی گویا نشان داده شده‌اند.

بر طبق روش هم‌ترازی داده‌های غیر زوج نشان داده شده در شکل 1 این 40 دسته ابتدا برای نقاط پایانی، صف‌بندی می‌شوند و یک پنجره‌ی مدل برای شناخت مدل‌های Wm-PCA&Wm-PLS انتخاب می‌شود. مدل Wm-PCA شناخته شده بیشتر برای تخمین مسیرهای مفقودی در دسته‌های کوتاهتر بکار می‌رود، بنابراین، تمام دسته‌های حاصل، طول مشابه 200 ساعت، بعد از خوراک داده‌ی مفقودی می‌باشند. سرعت خوراک سابس‌تریت بعدی برای دانستن در پیشرفت، پنداشته می‌شود از این رو آن، فرایند ورودی می‌باشد. بنابراین آن توسط الگوریتم‌های داده‌های مفقودی، تخمین زده می‌شود در حالی که سایر متغیرهای فرایند برای دسته‌های کوتاهتر، تخمین زده می‌شوند. بخصوص الگوریتم داده‌های مفقودی، تجسم نقشه نامیده شده، که در اینجا اعمال می‌شود. سایر الگوریتم‌های داده‌خ‌ی مفقودی مانند روش نشانه‌ی مرتب شده، همچنین می‌تواند اعمال شود.

متغیر فرایند غلظت اکسیژن حل شده بعنوان یک مثال، مسیر غلظت اکسیژن حل شده‌ی بعدی در مقابل مسیر واقعی آن، برای یک دسته با طول اجرای 160 ساعت، در شکل 5 نشان داده شده است جایی که سه حالت $W_m=80h$ ، $W_m=100h$ و $W_m=120h$ مقایسه می‌شوند. می‌توان مشاهده کرد که طول‌های پنجره‌ی مدل انتخاب شده، صحت تخمین داده‌های مفقودی را تحت تاثیر قرار می‌دهد. شکل 5 نشان می‌دهد که در این کاربردهای ویژه، بخش $W_m=100h$ در تمام شبیه‌سازی‌های بعدی، انتخاب می‌شود به هر حال، انتخاب طول پنجره‌ی مدل کاملاً ابتکاری و وابسته به فرایند می‌باشد. بطور کلی، طول‌های پنجره‌ی بزرگتر برای الگوریتم‌های داده‌ی مفقودی مناسب می‌باشد.

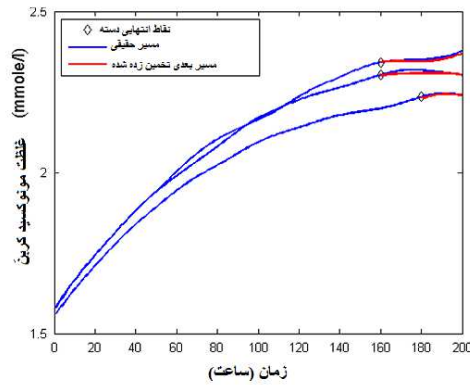


شکل 4. دسته‌هایی با طول‌های متغیر

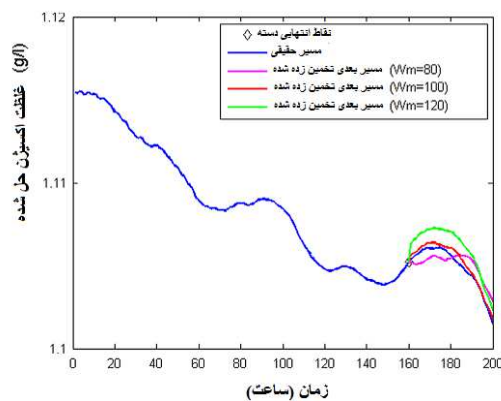
بنابراین W_m نباید بیش از اندازه بزرگ باشد برای اطمینان ساختن از اینکه دینامیک‌های فرایند در طول پنجره‌ی مدل خیلی زیاد تغییر نکند.

هم‌ترازی دسته‌ی کوتاهتر با استفاده از $W_m=100h$ با روش DTW در شکل 6 مقایسه می‌شود جایی که مسیر مرجع برای روش DTW انتخاب می‌شود برای اینکه مسیر برای تمام دسته‌های طول کامل با پروفایل‌های سرعت خوراک سبستریت، مشابه متوسط شود. می‌توان مشاهده کرد که اجرای هم‌زمان با استفاده از روش DWT، ارتباط نزدیکی به مسیر مرجع دارد چنانکه مسیر کوتاهتر بطور اساسی برای نزدیک بودن به مسیر مرجع بدون در نظر گرفتن

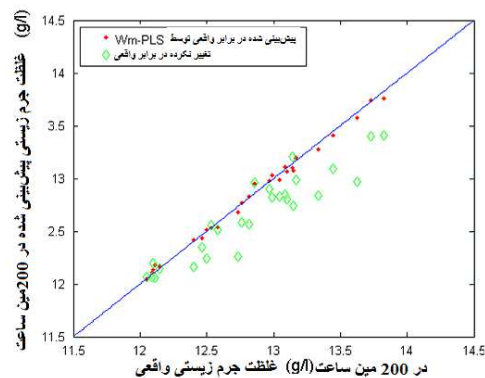
دینامیک‌های فرایند، پیچیده می‌شود. بنابراین روش‌های همزمان شده با روش DTW می‌توانند راه را از مسیر واقعی منحرف کنند در حالی که روش هم ترازوی پیشنهاد شده تنها مسیر مفقودی بعدی را تخمین می‌زند و مسیرهای شناخته شده‌ی قبلی دست نخورده نگه داشته می‌شوند.



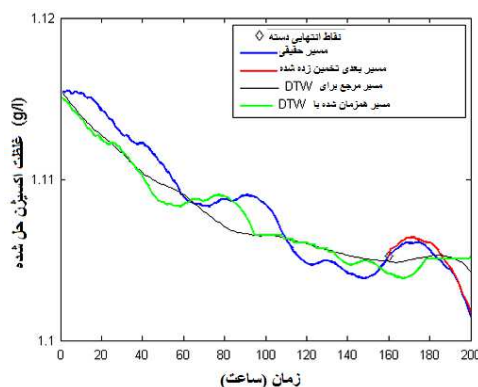
شکل 7. تخمین داده‌های مفقودی برای سه دسته‌ی کوتاهتر



شکل 5. تاثیر پنجره‌ی مدل روی تخمین داده‌های مفقودی



شکل 8. پیش‌بینی کمیت انتهایی دسته برای دسته‌ی کوتاهتر



شکل 6. مقایسه‌ی روش‌های هم‌ترازی داده‌های دسته‌ی غیر زوج

بدین منظور شرح بیشتر ظرفیت مدل $Wm-PCA$ شناخته شده برای طویل کردن مسیرهای کوتاهتر در شکل 7 نشان داده می‌شود تخمین مسیرهای غلظت مونوکسید کربن بعدی، برای سه دسته‌ی کوچکتر و مقایسه‌ی آنها با مسیرهای واقعی که بدست می‌آیند اگر این دسته‌ها در 200 ساعت امتداد یافته بودند. شکل، بطور واضح نشان می‌دهد که $Wm-PCA$ شناخته شده با استفاده از $Wm=100h$ بطور موفقیت‌آمیزی، مسیر مونوکسید بعدی را برای این سه دسته تخمین می‌زند. در اینجا تعدادی اصول برای $Wm-PCA$ برای بدست آوردن 16 اصول حسابداری برای تغییرپذیری 90% داده‌ها، انتخاب می‌شود.

غلظت‌های جرم زیستی بعدی برای دسته‌های کوتاهتر همچنین می‌تواند با استفاده از مدل $Wm-PCA$ شناخته شده، پیش‌بینی شود، یکبار دیگر تمام مسیرهای مفقودی قابل دسترس می‌باشند. در اینجا تعدادی از متغیرهای نهفته برای $Wm-PCA$ برای 8 شدن، جاسازی بیش از 70% تغییرپذیری در فضای x و بیش از 98% تغییرپذیری در فضایی y انتخاب می‌شود. غلظت‌های جرم زیستی، تخمین زده شده و مقادیر واقعی آنها برای دسته‌های کوتاهتر آنها در شکل 8 نشان داده شده است جایی که غلظت‌های جرم زیستی در 200h برای دسته‌های کوتاهتر آن همچنین توسط امتداد یافتن انطباق شبیه‌سازی‌ها، بدست آورده می‌شود در شکل 8 کیفیت محصول انتهایی دسته‌ی غیر صف شده، را نشان می‌دهد برای همچنین تمام دسته‌ها، بعضی از کیفیت‌های محصول انتهایی دسته از نقاط انتهایی اصلی

برای دسته‌های کوتاه‌تر آن ثابت نگه داشته می‌شود. آن می‌تواند مشاهده شود که مقادیر کیفیت محصول انتهایی دسته‌های غیر صف‌بندی شده، بطور مداوم پایین‌تر از مقادیر واقعی شان می‌باشد.

این بدلیل این حقیقت است که دسته‌ی کوتاه‌تر برای طول 200h صف‌بندی می‌شود و نتیجتاً مقادیر کیفیت محصول انتهایی دسته می‌تواند در طول مدت زمانی دسته‌ی امتداد یافته، رشد کند. داده‌ی دسته‌ی غیر زوج، کاملاً صف‌بندی می‌شوند، یک بار دیگر تمام مسیرهای مفقودی به 200h امتداد می‌یابد و کیفیت محصول انتهایی دسته‌ی 200 مین ساعات برای این دسته‌های کوتاه‌تر، به روز می‌شود. داده‌های دسته‌ی به صف شده سپس می‌تواند برای شناخت مدل PLS جدید استفاده شود که اشاره به مدل PLS با دسته‌ی کامل (FBL-PLS) دارد. صحت FBL-PLS می‌تواند بیشتر توسط آزمایش دسته‌هایی با طول دسته‌ی واحد 200h معتبر شود. با استفاده از مسیرهای متغیر فرایند از این دسته‌های آزمایش، غلظت جرم زیستی در 200h مین ساعات می‌تواند با استفاده از FBL-PLS پیش‌بینی شود و مقادیر پیش‌بینی شده با مقادیر واقعی شان، مقایسه می‌شود.



این مقاله، از سری مقالات ترجمه شده رایگان سایت ترجمه فا میباشد که با فرمت PDF در اختیار شما عزیزان قرار گرفته است. در صورت تمایل میتوانید با کلیک بر روی دکمه های زیر از سایر مقالات نیز استفاده نمایید:

لیست مقالات ترجمه شده ✓

لیست مقالات ترجمه شده رایگان ✓

لیست جدیدترین مقالات انگلیسی ISI ✓

سایت ترجمه فا ؛ مرجع جدیدترین مقالات ترجمه شده از نشریات معتبر خارجی



این مقاله، از سری مقالات ترجمه شده رایگان سایت ترجمه فا میباشد که با فرمت PDF در اختیار شما عزیزان قرار گرفته است. در صورت تمایل میتوانید با کلیک بر روی دکمه های زیر از سایر مقالات نیز استفاده نمایید:

لیست مقالات ترجمه شده ✓

لیست مقالات ترجمه شده رایگان ✓

لیست جدیدترین مقالات انگلیسی ISI ✓

سایت ترجمه فا ؛ مرجع جدیدترین مقالات ترجمه شده از نشریات معتبر خارجی