

Direction R&I

**Article Interne**

SINBAD

Récupération de chaleur fatale : Comment définir des intervalles de temps de séries temporelles optimaux pour le modèle thermique Time Slice Model

# Manoa RAFILIPOJAONA(1), Karine VOLTZ(1), Alan Jean-Marie(2), Sara EL MEDLAOUI(3) Ufuk HALISDEMIR(1)

1. Altran Research EST, Département R&I, Parc d’Innovation – 950 Boulevard Sébastien Brant 67404 – Illkirch-Graffenstaden France
2. Altran Lab, Département R&I, 2 rue Paul Dautier 78457 Velizy Villacoublay France
3. Altran Tech., Département R&I, 92B boulevard Solidarité 57070 METZ

## IDD.SINBAD.01 (Thématique de recherche) : Combinaison technologique récupérant la totalité des énergies fatales d’un système.

## EDC.SINBAD.01.01 (Question de recherche): Comment définir des intervalles de temps de séries temporelles optimaux pour le modèle thermique Time Slice Model ?

Résumé

Dans cet EDC, nous proposons d’améliorer la méthode du pincement étendue aux procédés industriels dont les grandeurs physiques des flux varient dans le temps, aussi appelé le *time slice model* (TSM). Le TSM consiste à segmenter le procédé en intervalles de temps dans lesquels les grandeurs physiques des flux peuvent être considérées constantes. L’objectif est de trouver un modèle de la série temporelle qui optimise sa précision en fonction de sa complexité. Pour ce faire, plusieurs variantes de la méthode de segmentation Bottom-up ont été implémentées. Les méthodes se basent sur le changement de corrélation entre les variables, sur la distance à la valeur moyenne et sur la distance à la médiane. Les résultats suggèrent que les méthodes des moindres carrés et de l’erreur absolue sont plus probantes car elles détectent les points de ruptures de manière plus systématique. Bien que les résultats obtenus soient prometteurs, nous proposons une piste d’amélioration en adaptant le pas de segmentation en fonction de la vitesse d’évolution des grandeurs physiques.

#### Keywords : Data science; Data mining; Intégration énergétique; Efficacité énergétique dans l’industrie.

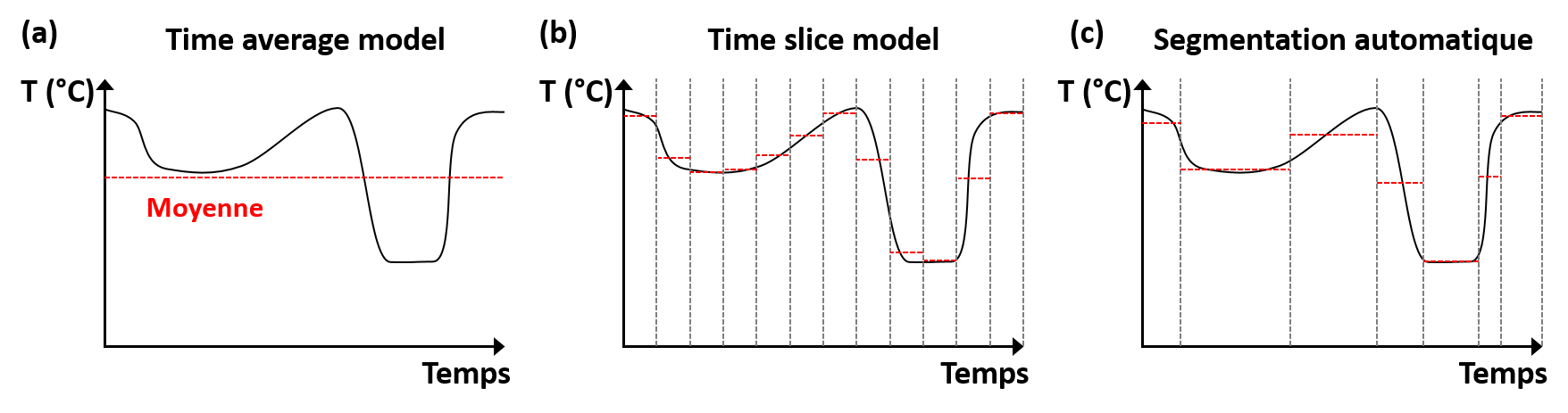
### Introduction

L’aspect énergétique des procédés industriels a longtemps été négligé à cause du faible coût des énergies fossiles et d’un développement rapide pour satisfaire une demande en croissance exponentielle. Mais l’augmentation du prix de l’énergie et la prise de conscience de l’impact de l’activité humaine sur l’environnement a remis le thème de l’efficacité énergétique au centre des considérations sociétales1.

Améliorer la performance énergétique des procédés industriels nécessite des méthodes d’optimisation systématiques et rigoureuses. Les premières méthodes datent des années 70 et se sont concentrées sur la récupération de la chaleur fatale dans l’industrie avec la méthode du pincement développé par Linnhoff *et al.*2.

La méthode du pincement est une méthodologie qui se base sur les lois de la thermodynamique afin d’identifier des échanges possibles entre des sources et des puits de chaleurs. L’objectif est de réduire les besoins en utilités externe chaudes et froides pour réaliser des économies mais aussi réduire les émissions de gaz polluant. Cette méthode fournit une procédure systématique pour la création de réseaux d’échangeurs qui permet de récupérer la chaleur fatale3.

Malgré un développement rapide des techniques d’intégrations énergétique ces dernières décennies, un rapport récent de l’ADEME4 souligne qu’une grande partie des besoins de chaleur dans l’industrie française est toujours perdue sous forme de chaleur fatale.

Historiquement, l’intégration énergétique et plus spécifiquement la récupération de chaleur fatale, a été principalement appliquée au processus continu1. Les processus discontinus, ou ceux dont les grandeurs physiques varient dans le temps, ont été ignorés parce que considérés comme trop complexes mais aussi parce que les besoins en énergie de ces procédés sont plus faibles que dans les processus continus. Néanmoins il a été démontré qu’il existait de nombreuses opportunités d’économies. Étant donné que les procédés batch aient été moins considérés que leurs homologues continus, il est d’autant plus important de s’y intéresser5–9.

**Figure 1**: **Exploitation des données capteurs issus de site industriels**. (a) La méthode du pincement standard, les grandeurs physiques des flux (ici la température est représentée), (b) le Time Slice Model (TSM) dans lequel le procédé est divisé en intervalle de temps (constant) et (c) l’approche proposé par ANAGREEN, l’objectif est de développer un algorithme qui automatise et optimise la segmentation.

L’objectif de cette étude est d’explorer la possibilité d’améliorer un modèle thermodynamique basé sur une version modifiée de la méthode du pincement qui permet de prendre en compte les variations dans le temps des grandeurs physiques des flux d’un procédé. Ce modèle, appelé le *Time Slice Model* (dénommée ci-après TSM), consiste à segmenter le procédé en intervalles de temps dans lesquels les grandeurs physiques peuvent être considérées constantes ([Fig. 1 (b)](#Fig1)). La méthode du pincement standard ([Fig. 1 (a)](#Fig1)) est ensuite appliquée sur chaque intervalle de temps. Cette méthode, bien que simpliste, permet d’identifier de nombreuses piste d’améliorations suivant les spécificités du process considéré10:

* Identifier les échanges directs entre les flux lorsqu’ils existent conjointement. Un réseau d’échangeur est créé sur chaque intervalle de temps. Le réseau final est la somme des réseaux optimisés sur chaque intervalle de temps.
* Reprogrammer les flux (*rescheduling*) afin de maximiser les possibilités d’échange direct. Cette approche est limitée par les contraintes liées au bon fonctionnement des procédés mais peut permettre de réaliser des économies par une simple reprogrammation du procédé, sans investissement à amortir.
* Intégrer le stockage de chaleur pour maximiser les opportunités de récupération de chaleur. Cette solution est souvent économiquement défavorable puisqu’elle nécessite l’intégration d’éléments complexes et couteux11. Néanmoins, de nombreuses études récentes suggèrent que cette approche peut être économiquement viable dans certains cas spécifiques.

Une estimation précise du coût est primordiale pour la mise en place d’un projet de rénovation ou de création de réseau d’échangeurs. Puisque le coût est directement lié à la quantité de chaleur récupérable, il est donc nécessaire d’estimer ce potentiel de récupération avec précision. Dans le cas des procédés dont les grandeurs physiques des flux varient dans le temps et dans le cadre du TSM, il faudrait idéalement segmenter en intervalle infiniment petit ([Fig. 1 (b)](#Fig1)). Cependant, bien qu’un découpage en intervalle de temps fin permette de quantifier avec précision la quantité de chaleur récupérable, cette approche nécessite un temps de calcul long. De plus, le réseau final est la somme des réseaux optimisés obtenus sur chaque intervalle de temps. Augmenter le nombre d’intervalle rend la solution finale plus complexe et trop couteuse.

Une approche plus adaptée est celle de la segmentation automatique des séries temporelles ([Fig. 1 (c)](#Fig1)). L’objectif de cet EDC est donc d’explorer les techniques de segmentation automatique des séries temporelles afin d’optimiser le TSM.

Dans la prochaine section, nous introduisons rapidement la méthode du pincement qui permet de définir les objectifs énergétiques et de créer un réseau d’échangeur optimisé sur chaque intervalle de temps. Nous détaillons également les différents algorithmes utilisé pour segmenter les séries temporelles. Nous présentons et discutons ensuite les résultats obtenus avec l’approche proposée dans cette étude. Nous finissons par une conclusion et des perspectives à ces travaux.

### Principe

*II.1 La méthode du pincement*

Il y a deux problèmes principaux lors du design d’un procédé industriel. Le premier est celui du design et du fonctionnement des unités individuelles, le deuxième est celui du design du système global. La méthode du pincement adresse le deuxième point.

La méthode considère l’ensemble des flux d’un procédé pour déceler des potentiels de récupération de chaleur. Tout flux ayant une température d’entrée supérieure à la température de sortie est défini comme un flux chaud, et inversement tout flux ayant une température d’entrée inférieure à la température de sortie est défini comme un flux froid. Dans le cas le plus simple, des utilités froides pourraient être utilisé pour refroidir les flux chauds et des utilités chaude pour réchauffer des flux froids. La méthode du pincement fournit une méthodologie pour réduire la consommation énergétique, elle se décompose en deux étapes principales. La première consiste à découper le procédé en intervalle de température dans lesquels la quantité de chaleur disponible ou requise est calculée. Ce découpage, ainsi que des considérations thermodynamiques simple, permettent de déterminer le minimum d’énergie externe (en utilités chaude et froide) nécessaire au bon fonctionnement du procédé. La deuxième étape de la méthode fournit des règles dérivées de concepts thermodynamiques qui permet la synthèse d’un réseau d’échangeurs. Ces échangeurs permettent le transfert de chaleur entre flux chaud et flux froid et leur placement selon les règles de la méthode du pincement permet d’atteindre l’objectif énergétique calculé dans la première étape1.

**Figure 2: Illustration du principe de fonctionnement de l'algorithme Bottom-up**. L’algorithme commence par diviser le signal original en de nombreux petits segment (Step 0) puis fusionne séquentiellement les segments voisins selon un critère de similarité (Steps 1, 2, …) jusqu’à obtenir le nombre de segment désiré (Result)12.

Bien que la méthode ne soit pas basée sur des concepts mathématiques rigoureux, elle a été appliquée à de nombreux cas industriels avec succès permettant des économies d’énergie de 10% à 40%. D’autres méthodes ont été développées mais sont hors du cadre de la présente étude3.

*II.2* *La segmentation des séries temporelles multivariées avec l’algorithme Bottom-up*

La segmentation des séries temporelles peut être utilisée pour extraire des segments dans lesquels le signal est considéré constant, pour identifier des points de rupture ou simplement pour compresser le signal dans une représentation plus compacte. L’objectif de cette étude est d’étudier la segmentation multivariée des séries temporelles. L’approche multivariée de la segmentation consiste à combiner les informations de plusieurs capteurs d’un système de tel sorte que cette fusion permette d’obtenir des informations supplémentaires comparé au cas où ces données étaient considérées séparément.

Une série temporelle multivariée est définie par :

Ou est un ensemble fini de échantillons de dimension n (n variables) indexés par des points dans le temps . Un segment de T est un ensemble de points consécutifs dans le temps :

La segmentation d’une série temporelle en une partition de segment disjointes deux à deux peut être formulée de la façon suivante :

Où , et

L’objectif de la présente étude est de déterminer des segments sur lesquels les séries temporelles peuvent être considérées homogène. Cette notion est formalisée par une fonction de coût qui décrit l’homogénéité de ces segments internes. Cette fonction de coût peut être définie comme la distance entre la valeur actuelle de la série temporelle et la valeur d’une fonction simple définie sur cet intervalle (comme par exemple la variance).

L’algorithme *Bottom-up* est une approche séquentielle de détection de point de rupture. L’algorithme commence par diviser le signal original en de nombreux petits segments puis fusionne séquentiellement les segments voisins selon un critère d'homogénéité (défini par une fonction de coût que nous discuterons plus tard) jusqu’à obtenir le nombre de segment désiré (ou un autre critère d’arrêt). La méthode est illustrée sur la [Fig. 2](#Fig2). Les avantages de cette approche sont sa complexité de calcul linéaire et son concept relativement simple. Cependant, la procédure de fusion peut être instable car elle est réalisée sur des petits segments pour lesquels la signification statistique est moindre12.

Le coût de la fusion de deux segments voisins peut être déterminé en comparant l’homogénéité des deux segments séparés à l’homogénéité du segment fusionné. Dans la présente étude nous avons exploré quarte approches pour déterminer le coût d’un segment, elles sont détaillées dans les paragraphes suivants.

**Décomposition en valeurs singulières**

La première approche se base sur la décomposition en valeurs singulières (ou single value décomposition en anglais, dénoté SVD ci-après), elle se base sur les travaux de Spiegel *et al.*13. La SVD permet de décomposer la matrice du segment , défini sur l’intervalle , de la manière suivante :

Où les éléments diagonaux de contiennent les valeurs singulières de en ordre décroissant, et et sont les vecteurs singuliers gauches et droits correspondants.

En sélectionnant les premières valeurs singulières (les plus importantes) et leurs vecteurs propres correspondant, le modèle SVD projette les données sur un sous espace de taille . Pour extraire des indicateurs d’anomalies ou d’instabilités dans la progression des variables étudiées, il est utile de calculer l’erreur de reconstruction qui est définie comme l’erreur quadratique moyenne entre la projection , et le segment original. L’erreur de reconstruction est formellement déterminée par la mesure ( en anglais) qui est définie par :

Le but de la méthode est d’utiliser la comme indicateur de l’homogénéité des segments indépendants et fusionnés. Finalement, la fonction de coût est définie par :

Pour résumer, la méthode se base sur le fait que la SVD permet d’approximer la série temporelle de manière satisfaisante lorsque la corrélation entre les variables reste linéaire. En d’autres termes, le model va générer une erreur de reconstruction importante lorsqu’il y a un fort changement de corrélation entre les variables. On suppose ensuite que deux segments adjacents peuvent être fusionnés si l’erreur de reconstruction est faible, c’est-à-dire si la corrélation entre les variables dans ces intervalles de temps contigus ne change pas considérablement.

Le choix de la taille du sous espace est une question fondamentale de la SVD. Il existe de nombreuses méthodes heuristiques pour fixer et la taille du sous espace est souvent choisie pour que le vecteur projeté représente 95% de la variance des données originales14. Il s'agit en fait de trouver une valeur de qui optimise la précision du modèle en fonction de sa complexité. Une étude récente de Gavish et Donoho15 a étudié la question en détail et propose une méthode pour déterminer un optimal en fonction des données étudiées. Dans notre cas le problème est simple puisque notre espace original est composé de trois variables, la taille du sous espace est donc ou , pour cette étude nous avons choisi .

**Analyses en composantes principales**

L’analyse en composante principales (dénoté PCA ci-après) consiste à calculer les valeurs et vecteurs propres de la matrice de covariance de définie par16 :

Où les sont les valeurs moyennes des variables sur l’intervalle . De manière similaire à la SVD, la PCA se base sur la décomposition de la matrice de covariance sous la forme :

Où est une matrice qui contient les valeurs propres de en ordre décroissant et une matrice contenant les vecteurs propres correspondants. On tronque la décomposition sur un sous espace de taille . L’erreur de reconstruction est définie par :

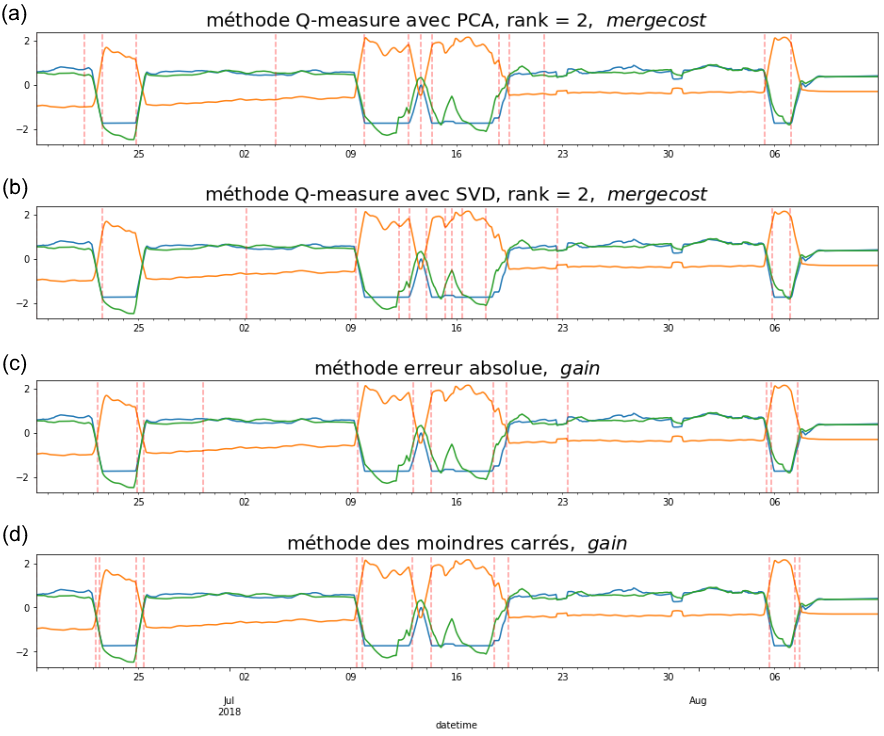
et le coût de la fusion entre deux segments par l’équation .

**Les méthodes des moindres carrés et de l’erreur absolue**

Nous avons également exploré la possibilité de définir des fonctions de coût basées sur les méthodes des moindres carrés et de l’erreur absolue12. Les fonctions de coûts associées sont :

Et

pour la méthode des moindres carrés et la méthode de l’erreur absolue respectivement.

Pour résumer, la Q-measure segmente le signal original en fonction de la stabilité de la corrélation entre les variables alors que la méthode des moindres carrés détecte des changements dans la moyenne et que la méthode de l’erreur absolue est sensible au déplacement du point central.

**Figure 3 : Résultat de segmentation** **sur un flux spécifique du procédé Midrex**. Résultat de la segmentation avec la méthode Q-measure issue de la PCA (a) et de la SVD (b). Résultat de la segmentation avec les méthodes de l’erreur absolue (c) et de la méthode des moindres carrés (d). Ces graphiques sont expliqués plus en détail dans le texte.

**Homogénéité de deux segments voisins**

Nous avons utilisé deux méthodes pour déterminer l’homogénéité de deux segments voisins. Le « *mergecost* » représente le coût de fusion de segment et est défini par :

Le *mergecost* est utilisé pour la Q-measure13,16. La deuxième méthode est utilisée avec les méthodes du moindre carré et de l’erreur absolue12 :

Le *mergecost* calcule simplement le coût du segment fusionné alors que le *gain* soustrait le coût des segments indépendants au coût du segment fusionné. Pour rappel, l’algorithme Bottom-up fusionne les deux segments contigus pour lesquels le *mergegost* ou le *gain* est le plus faible jusqu’à un critère d’arrêt. Nous avons étudié deux critères d’arrêt possible. Le premier est de fixer le nombre de segment désiré, dans cas l’algorithme fusionnera des segments jusqu’à atteindre le nombre de segment prédéfinis. Le deuxième critère d’arrêt se base sur l’évolution de l’homogénéité, en effet une valeur faible d’homogénéité indique la fusion de segments voisins similaires alors qu’une valeur élevée indique la fusion de segments hétérogènes. Nous avons défini un critère basé sur l’augmentation abrupte du minimum de coût de fusion.

### Résultats

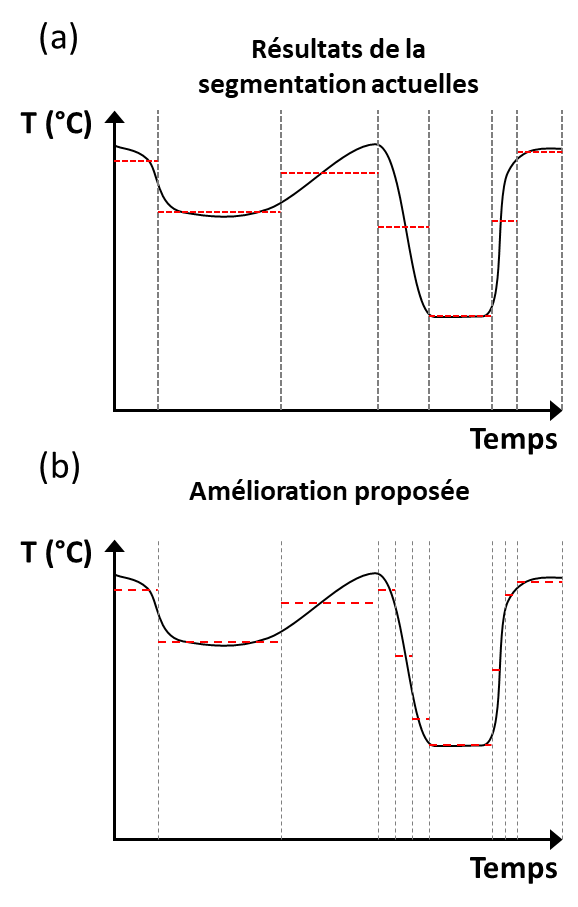
*III.1 Cas d’étude : Le procédé Midrex*

Le procédé Midrex est un procédé de réduction du minerai de fer et utilise généralement le gaz naturel comme source de gaz réducteur et source d’énergie. Il est composé de deux éléments principaux, le four dans lequel la réaction entre le gaz réducteur et les oxydes de fer présent dans le minerai prend place et le reformeur dans lequel le gaz réducteur est préparé. Ces deux éléments sont accompagnés d’un système de production et de récupération de chaleur.

*III.2 Discussion*

La [Fig. 3](#Fig3) illustre les résultats des algorithmes présenté dans la [section II.2](#SectionII2) appliqués à un flux spécifique du procédé. Le flux est représenté par trois variables : la température d’entrée, la température de sortie et le débit calorifique massique. La série temporelle est composée de 80 000 points enregistrées entre le 18 Juin et le 12 Août. Notons que pour la , les variables sont centrées et réduites. Nous précisons également qu’un lissage sur les données a été appliqué, comme suggéré dans l’article de Spiegel et al13. Un filtre passe-bas Savitzky-Golay d’ordre 1 et avec une fenêtre de 501 points a été appliqué pour réduire le bruit. Finalement, nous avons utilisé les mêmes paramètres pour faciliter la comparaison des résultats obtenus : un nombre de segment initial fixé à 160 000 (5 minutes par segment) et le nombre de segment final fixé à 14.

Nous remarquons que pour la ([Fig. 3](#Fig3) (a)&(b)), la série temporelle est principalement segmentée lorsqu’il y a des changements importants dans la corrélation entre les variables. Cependant on remarque aussi quelques situations pour lesquelles l’algorithme semble ne pas tenir compte de changements importants dans la corrélation. Pour la PCA par exemple, un changement est identifié autour du 25 Juin. Par contre, un jour plus tard, un autre segment est attendu à la fin de la période de variation importante des variables, mais n’est pas identifié par l’algorithme. Pour la SVD, le changement de corrélation du 25 Juin n’est tout simplement pas identifié. On note également la présence de segment là où il n’y a pas de changement en apparence (autour du 04 et du 22 Août pour la PCA et autour du 02 et du 23 Août pour la SVD).

D’après la [Fig. 3](#Fig3) (c) et (d), il semble que les méthodes des moindres carrés et de la valeur absolue soient plus efficaces pour la segmentation des séries temporelles étudiées. En effet, ces méthodes de segmentation sont sensibles aux changements de la valeur moyenne ou de la médiane respectivement et des segments sont créés lorsque les dérivées de ces grandeurs accélèrent ou décélèrent rapidement. On peut observer une illustration de ces propriétés autour du 25 Juin, l’algorithme segmente la série temporelle à la fois lorsque le débit calorifique massique (les température d’entrées et de sorties) diminue fortement (augmentent) et lorsqu’il se stabilise. Puisque l’objectif final est d’estimer la quantité de chaleur récupérable avec plus de précision, il est souhaitable que l’algorithme isole les intervalles ou les grandeurs physiques varient fortement dans des segments séparés. On note toutefois que la méthode de l’erreur absolue souffre des mêmes limitations que les méthodes , des segments sont créés autour du 29 Juin et du 23 Août alors que la série temporelle semble stable.

### Conclusions et perspectives

Dans cet EDC nous proposons d’améliorer la méthode du pincement étendue aux procédés industriels dont les grandeurs physiques des flux varient dans le temps, aussi appelé le time slice model. Le but est d’estimer la quantité de chaleur récupérable avec plus de précision, c’est une étape cruciale pour la réalisation d’un projet industriel puisqu’elle permet d’estimer les coûts d’investissement, d’opération, les économies réalisées ainsi que le temps de retour sur investissement. L’approche proposée est de trouver automatiquement des intervalles de temps sur lesquels les séries temporelles issues de sites industriels peuvent être considérés constantes.

**Figure 4: Piste d'amélioration de la segmentation automatique**. (a) Illustration des résultats des segmentations obtenus dans cet article et (b) Piste d’amélioration proposée. En plus de détecter les points de rupture nous proposons d’adapter le pas de segmentation de manière adaptative en fonction de l’évolution des grandeurs physiques.

Pour la segmentation nous avons implémenté la méthode Bottom-up dont le principe est de diviser le signal original en de nombreux petits segments puis de fusionner séquentiellement les segments voisins selon un critère d'homogénéité, jusqu’à obtenir le nombre de segment désiré (ou selon un autre critère d’arrêt). Nous avons exploré quatre méthodes pour évaluer l’homogénéité de deux segments voisins ainsi que deux fonctions de coûts de fusion. Ces variantes de la méthode bottom-up ont été appliquée à une série temporelle issue d’une usine Midrex. La série temporelle est composée de trois variables nécessaires relatives à un flux et nécessaire à l’application du modèle thermique.

Les résultats montrent que les méthodes qui se basent sur les changements de corrélation entre les variables () se comportent de manière instable. En effet, une analyse détaillée de la [Fig. 3](#Fig3) (a) et (b) montre que cette approche ne détecte pas certains points de rupture mais aussi qu’elle segmente la série temporelle à des endroits où les grandeurs physiques semblent constantes. Les résultats obtenus avec les méthodes des moindres carrés et de l’erreur absolue semblent plus probantes car elles prennent en compte les points de ruptures de manière plus systématique. Néanmoins la méthode de l’erreur absolue souffre des mêmes limitations que les méthodes , des segments sont créés à des endroits ou la série temporelle semble constante. Notons que ces comportements instables peuvent être dû au fait que le nombre de segment final est fixé. Pour évaluer ces approches de manière plus précise, la suite logique de ce travail est de comparer les résultats obtenus avec le modèle thermique. Ces travaux seront présentés dans un EDC indépendant de celui-ci.

Finalement, bien que les résultats obtenus soient prometteurs, il existe des pistes d’amélioration. La [Fig. 4 (a)](#Fig4) illustre les résultats obtenus avec les méthodes de segmentation présentées dans cet article. La [Fig. 4 (b)](#Fig4) est une proposition d’amélioration de la segmentation automatique. En plus de détecter les points de rupture nous proposons d’adapter le pas de segmentation de manière adaptative en fonction de l’évolution des grandeurs physiques. Le critère d’évolution des grandeurs physique doit être optimisé pour le modèle thermique. Encore une fois, il s'agit de trouver un modèle de la série temporelle qui optimise sa précision en fonction de sa complexité.

### Références

1. Kemp, I. C. *Pinch analysis and process integration: a user guide on process integration for the efficient use of energy*. (Butterworth-Heinemann, 2007).

2. Linnhoff, B. & Hindmarsh, E. The pinch design method for heat exchanger networks. *Chem. Eng. Sci.* **38**, 745–763 (1983).

3. Klemeš, J. J. & Kravanja, Z. Forty years of Heat Integration: Pinch Analysis (PA) and Mathematical Programming (MP). *Curr. Opin. Chem. Eng.* **2**, 461–474 (2013).

4. ADEME. La chaleur fatale. https://www.ademe.fr/sites/default/files/assets/documents/chaleur\_fatale-8821-2018-06\_pdf.pdf (2017).

5. Kemp, I. C. *Process integration: Process change and batch processes*. (ESDU International PLC, 1990).

6. Fernández, I., Renedo, C. J., Pérez, S. F., Ortiz, A. & Mañana, M. A review: Energy recovery in batch processes. *Renew. Sustain. Energy Rev.* **16**, 2260–2277 (2012).

7. Krummenacher, P. Contribution to the heat integration of batch processes (with or without heat storage). (École Polytechnique Fédérale de Lausanne, 2001).

8. Pourali, O., Amidpour, M. & Rashtchian, D. Time decomposition in batch process integration. *Chem. Eng. Process. Process Intensif.* **45**, 14–21 (2006).

9. Zhao, X. G., O’neill, B. K., Roach, J. R. & Wood, R. M. Heat Integration For Batch Processes. *Chem. Eng. Res. Des.* **76**, 685–699 (1998).

10. Brunner, F. & Krummenacher, P. *Introduction à l’intégration énergétique de procédés par l’Analyse Pinch*. (2017).

11. Abdelouadoud, Y., Lucas, E., Krummenacher, P., Olsen, D. & Wellig, B. Batch process heat storage integration: A simple and effective graphical approach. *Energy* **185**, 804–818 (2019).

12. Truong, C., Oudre, L. & Vayatis, N. Selective review of offline change point detection methods. *Signal Process.* **167**, 107299 (2020).

13. Spiegel, S., Gaebler, J., Lommatzsch, A., De Luca, E. & Albayrak, S. Pattern recognition and classification for multivariate time series. in *Proceedings of the Fifth International Workshop on Knowledge Discovery from Sensor Data - SensorKDD ’11* 34–42 (ACM Press, 2011). doi:10.1145/2003653.2003657.

14. Singhal, A. & Seborg, D. E. Clustering multivariate time-series data. *J. Chemom.* **19**, 427–438 (2005).

15. Gavish, M. & Donoho, D. L. The Optimal Hard Threshold for Singular Values is 4/sqrt(3). *ArXiv13055870 Stat* (2014).

16. Abonyi, J., Feil, B., Nemeth, S. & Arva, P. Principal Component Analysis based Time Series Segmentation – A New Sensor Fusion Algorithm. 23 (2004).