

DATA SCIENCE PROJECT
LBIRE2234

Développement d'un modèle opérationnel de détection de collages en temps réel lors de coulées continues en aciérie

Étudiant.e.s :

Abdelali DSSAM
Cyril JOLIET
Nour-El Houda TEBER
Lison VAN ASBROUCK

Supervision :

Pr. Patrick BOGAERT
Pr. Emmanuel HANERT

Client :

EBDS Engineering

Table des matières

Résumé	2
1 Introduction	3
2 Contexte : Détection d'un collage en coulée continue	4
3 Matériels et méthodes	6
3.1 Données utilisées	6
3.2 Software et packages utilisés	7
3.3 Modèle	7
3.3.1 Calculs	8
3.3.2 Vérifications	9
3.3.3 Fonctions	10
3.3.4 Paramètres	11
3.4 Indice de certitude de détection (DCI)	12
4 Résultats et discussion	13
4.1 Caractérisation du bruit	13
4.2 Cas détectés	14
4.3 Sorties du modèle	14
4.4 Analyse de sensibilité	16
4.5 Améliorations	17
5 Conclusions et perspectives	19
Bibliographie	20
A Fonctions utilisées	21
A.1 <i>process_data</i>	21
A.2 <i>plot</i>	22
A.3 <i>linear</i>	24
A.4 <i>temp</i>	25

Résumé

Dans ce rapport, nous présentons un outil de détection visant à identifier le phénomène de collage dans le processus de coulée continue de l'acier, où l'acier solidifié adhère à la lingotière, entraînant à la fois des défauts et des dommages aux équipements ainsi que des interruptions de production. Notre algorithme s'appuie sur une analyse par régression linéaire, permettant d'identifier la forme en "V" dans les variations de température, qui est caractéristique des collages. Notre analyse repose sur des données enregistrées par des capteurs à fibre optique, offrant une grande qualité de mesure. Des vérifications sont effectuées sur les pentes des droites, sur les R^2 des régressions ainsi que sur les variations de température et une analyse de sensibilité nous a permis de fixer des conditions pour déterminer si un collage a lieu. Si ces conditions sont respectées une préalarme est déclenchée avant de sonner l'alarme si la préalarme est maintenue dans le temps. Un indice allant de 0 à 1 est aussi calculé pour indiquer la certitude avec laquelle un collage est détecté. Notre modèle a permis de détecter les collages sans déclencher de fausses alarmes pour plusieurs mois de données analysées provenant des aciéries *AMDG* (Pologne) et *Industeel* (Belgique). Cependant certaines améliorations pourraient encore être apportées notamment grâce aux mesures enregistrées par la deuxième fibre optique de la lingotière qui permettraient d'assurer une certaine fiabilité en cas de défaillance de la première fibre optique.

1 Introduction

La coulée continue est une étape essentielle dans le processus de production sidérurgique, permettant de transformer l’acier liquide en produits semi-finis de haute qualité (Fig. 1). Cette technique implique un contrôle pointu des conditions thermiques au sein de la lingotière, où l’acier commence sa solidification, pour garantir l’efficacité et la sécurité du processus (VIJAYARAM, 2019). En effet, le phénomène appelé *collage* qui peut arriver dans ce type d’opération, constitue l’un des défis majeurs. Une adhérence inappropriée de l’acier solidifié à la lingotière peut entraîner des arrêts de production coûteux, des dommages aux équipements et une dégradation de la qualité des produits finis (MAZUMDAR et RAY, 2001).

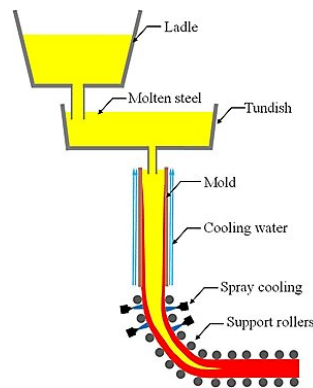


FIGURE 1 – Schéma illustrant le principe de la coulée continue (VIJAYARAM, 2019).

Pour répondre à ces enjeux, les technologies de surveillance utilisant des fibres optiques se sont imposées comme des solutions de pointe. Contrairement aux méthodes traditionnelles de mesure utilisant des thermocouples, elles offrent une analyse thermique fine et distribuée tout le long de la lingotière, permettant de détecter des petites variations de température qui précèdent souvent l’apparition de défauts comme le collage. Cette capacité à détecter précocement les anomalies offre la possibilité d’intervenir rapidement, réduisant ainsi le risque de défauts structuraux ou de fissures dans l’acier en cours de solidification (LI et al., 2024).

Dans ce rapport, nous présentons un modèle opérationnel de détection des collages basé sur l’analyse des données de température fournies par ces capteurs à fibre optique. Ce modèle a pour objectif d’identifier, en temps réel, les signes avant-coureurs d’un collage en analysant les fluctuations thermiques mesurées dans la lingotière. En optimisant ce suivi, nous pouvons non seulement améliorer la sécurité des opérations de coulée continue, mais aussi accroître la productivité en minimisant les interruptions de production et les risques de détérioration des équipements. Ainsi, la mise en place d’un système de surveillance fiable, capable de prévenir ces incidents, représente un atout majeur pour la sidérurgie actuelle. Ce rapport détaillera donc la conception de ce modèle ainsi que ses performances pour répondre à ces défis critiques.

2 Contexte : Détection d'un collage en coulée continue

La technique de coulée continue consiste à verser l'acier liquide dans une lingotière maintenue à une température spécifique grâce à un système de refroidissement. Les métaux en fusion en contact avec la lingotière commencent rapidement à se solidifier. Des rouleaux d'extraction tirent la brame et la font avancer dans la machine de coulée, tandis que l'eau continue de la refroidir. La brame peut être coupée lorsque la section est complètement solidifiée (EBDS ENGINEERING, 2015).

Quand une fracture se produit en raison du collage près du ménisque, la propagation de la fissure s'étend dans l'espace et le temps (TRAN et al., 2020). Lorsque cette fissure se forme, l'écart est alors rempli par de l'acier liquide qui se solidifie, augmentant ainsi la zone déjà collée sur la lingotière. Ensuite, la fine couche formée à l'endroit de la fissure sera de nouveau déchirée à un emplacement légèrement décalé dans la direction de coulée (Fig. 2).

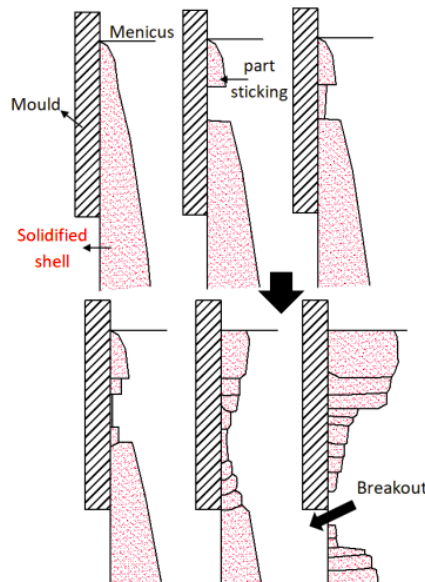


FIGURE 2 – Propagation d'une fissure due au phénomène de rupture par collage (TRAN et al., 2020).

Cependant, dans certains cas, l'écart causé par le collage peut se refermer spontanément sans causer de dommages. Ceci est dû à un décollage de l'acier. La fissure formée par le collage se referme alors naturellement, rétablissant ainsi la continuité de la brame. Ce phénomène est identifiable par une remontée des températures après le refroidissement distinctif du collage.

Le phénomène de collage affecte non seulement la production fluide de la coulée continue, mais endommage également les équipements. Pour éviter cela, les aciéries utilisent généralement des systèmes de surveillance thermique pour la détection en temps réel des conditions internes de la lingotière. Ces systèmes, qui reposent généralement sur des thermocouples (TC) intégrés dans les plaques de cuivre du moule, mesurent les variations de température qui surviennent pendant la solidification initiale de l'acier en fusion dans la

lingotière (DUAN et al., 2019).

En effet, un collage suit une propagation horizontale et verticale avec des caractéristiques claires et repérables grâce à un suivi des températures (Fig. 3). Ce type de propagation en deux dimensions crée une forme de V dans les relevés de température montrant un pic de chaleur suivie d’une baisse de température (LI et al., 2024).

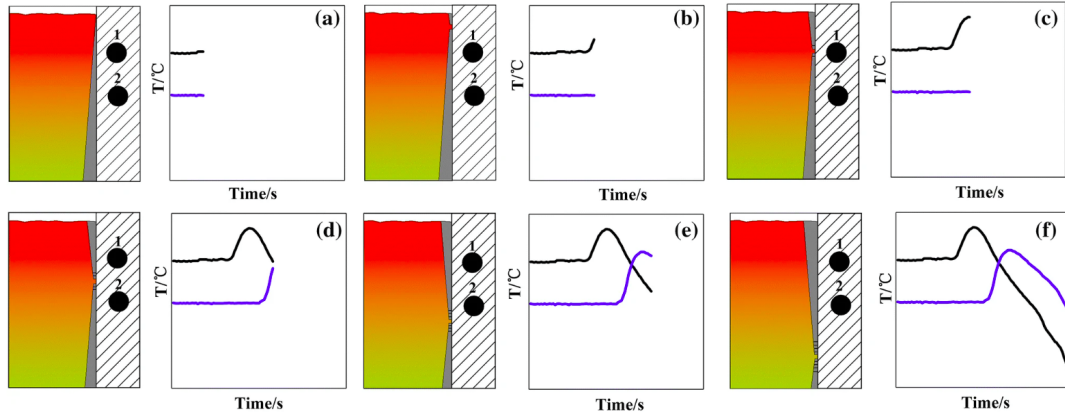


FIGURE 3 – Diagramme des caractéristiques de la température et du processus de formation du point de rupture : (a) Température dans des conditions normales sans rupture, (b) le point de rupture émerge près du ménisque, (c) le point de rupture descend progressivement et se rapproche du thermocouple (TC) de la première rangée, (d) le point de rupture traverse le TC de la première rangée, (e) le point de rupture traverse le TC de la deuxième rangée, (f) le point de rupture descend jusqu’à la sortie du moule (DUAN et al., 2019).

La prédiction des collages et fissures qui s’ensuivent est donc possible en alertant de la formation d’un collage lorsque les signes de sa propagation thermique spécifique sont détectés. Cependant, des mauvais diagnostics sont possibles en raison notamment de thermocouples défectueux qui ne mesurent pas correctement les températures dans la lingotière (LI et al., 2024).

Une solution émergente dans le domaine et qui est notamment adoptée par l’entreprise *EBDS Engineering* (EBDS ENGINEERING, 2024), est l’utilisation de fibres optiques pour la mesure des conditions thermiques dans la lingotière, en remplaçant ainsi les thermocouples. L’un des principaux avantages des fibres optiques réside dans leur sensibilité de mesure bien plus grande que celle des thermocouples (LI et al., 2024). Les mesures avec ces fibres optiques sont également mieux réparties tout au long de la lingotière ce qui permet des mesures plus précises et localisées des variations de température sur la surface de la plaque de cuivre (DUAN et al., 2019).

Le modèle que nous avons construit dans ce projet est basé sur ces mesures de température obtenues par fibres optiques dans la lingotière pour évaluer l’occurrence de collages de l’acier liquide.

3 Matériels et méthodes

3.1 Données utilisées

Dans le cadre de ce projet, nous avons travaillé avec des données collectées à l'aide de capteurs de température intégrés dans des fibres optiques au sein de la lingotière. Ces données qui nous ont été transmises par *EBDS Engineering*. Elles concernaient l'évolution de coulées continues dans les aciéries d'*Industeel* (Belgique) et d'*AMDG* (Pologne) qui font partie du groupe *ArcelorMittal*. Les données collectées présentaient des cas de collage ainsi que des variations température aléatoires, sans collages, considérées comme du bruit.

La lingotière, qui joue un rôle essentiel dans le refroidissement de l'acier, est composée de quatre faces : deux grandes faces équipées de fibres optiques et deux petites faces équipées de thermocouples. Afin de détecter les collages, nous avons concentré notre analyse sur les grandes faces de la lingotière, en utilisant une seule fibre optique par face. Ces fibres permettent de détecter des anomalies thermiques caractéristiques des collages, tels que des augmentations soudaines de température suivies d'une diminution, souvent associées à une forme en "V" sur les courbes de température.

Les tables de données exploitées contenaient un large éventail de variables liées au processus de production. Cependant, notre analyse s'est principalement orientée sur les colonnes concernant :

- la date et l'heure (**DATE** et **TIME**),
- la vitesse de coulée (**SPEED**),
- les températures mesurées par les capteurs positionnés sur les fibres optiques.

La colonne **TIME** représente l'heure exacte d'enregistrement des données au cours du processus de production, exprimée au format **hh:mm:ss**. Elle permet de suivre les variations des différents paramètres mesurés (comme la vitesse ou les températures) avec un intervalle d'une seconde entre chaque enregistrement. La colonne **DATE** contient la date à laquelle les mesures ont été enregistrées. Pour faciliter l'analyse, nous avons fusionné ces deux colonnes en une seule colonne appelée **DATETIME**, combinant la date et l'heure en un format unique. Les colonnes dédiées aux capteurs représentent les mesures de température collectées en **degrés Celsius**. Chaque colonne correspond à un capteur spécifique, et les valeurs qu'elle contient indiquent les températures successives mesurées par ce capteur à chaque seconde, comme indiqué dans la colonne **TIME**.

3.2 Software et packages utilisés

L'ensemble du développement du modèle a été réalisé en Python. Ce qui nous a permis de tester, analyser et visualiser nos résultats aisément grâce aux nombreuses bibliothèques existantes décrites brièvement ci-dessous.

1) Chargement et traitement des données

- **Pandas** a été utilisé pour gérer les données brutes issues des capteurs, en les chargeant depuis un fichier CSV.
- **Numpy** a servi à générer les vecteurs représentant les positions des capteurs (coordonnées X et Y) et à effectuer des calculs numériques rapides
- **Datetime** a été utilisé pour combiner les colonnes DATE et TIME, générer des incréments temporels d'une seconde, et définir des intervalles de 60 secondes pour l'analyse.
- **math** a été utilisé pour effectuer des arrondis lors du positionnement des capteurs dans une grille, ce qui garantit une représentation géométrique précise des données captées.

2) Analyse géométrique et statistique

- **sklearn.linear.model** a permis de calculer des régressions linéaires sur les variations de température pour chaque côté du motif en "V".

3) Visualisation et communication des résultats

- **matplotlib.colors** a été utilisé pour attribuer une échelle de couleurs aux variations de température dans les graphiques.
- **matplotlib.pyplot** a été utilisé pour générer des graphiques mettant en évidence les anomalies.

3.3 Modèle

Pour détecter la présence d'un collage dans la lingotière, nous sommes partis de la géométrie du phénomène lorsque l'on observe les variations thermiques au sein de la lingotière. Lorsqu'un collage a lieu, le profil de variation de température est très caractéristique et constitue donc une base solide pour identifier ce phénomène. En effet, lorsque l'acier colle à la paroi de la lingotière, un schéma en "V" apparaît sur le profil des variations de température, correspondant à une soudaine montée de température suivie d'un refroidissement rapide. Nous avons donc élaboré une méthode basée sur des régressions linéaires afin d'identifier les deux droites qui constituent ce "V".

La détection d'un collage se fait en plusieurs étapes et selon différents paramètres. Dans cette section, nous présenterons d'abord en section 3.3.1 les calculs réalisés pour obtenir les valeurs de ces paramètres. Nous détaillerons ensuite en section 3.3.2 les étapes suivies par l'algorithme pour la vérification de ces paramètres qui mènera à une alarme si un collage a lieu. Ces étapes sont reprises dans un schéma conceptuel illustrant l'algorithme (Fig. 5). Les fonctions utilisées par l'algorithme ainsi que la description des paramètres et

leurs valeurs se retrouvent dans les section 3.3.3 et 3.3.4 respectivement. Les valeurs de ces paramètres ont été choisies arbitrairement et feront l'objet d'une analyse de sensibilité dans la section 4.4 de ce rapport.

3.3.1 Calculs

L'analyse des données se fait chaque seconde sur l'historique de mesure des 60 dernières secondes. La première étape du modèle consiste en l'identification de "points critiques" définis comme tous les points pour lesquels la variation de température est supérieure au *Seuil* fixé à 1°C. Tant que le nombre de points critiques, sur une plage de temps de 60 secondes, est inférieur à 30, aucun calcul n'est effectué et la plage de temps suivante est analysée. Une fois le nombre de 30 points critiques atteint, seuls les maximums de l'historique de mesure sur 60 secondes pour chaque capteurs sont utilisés. L'algorithme détermine ensuite un point de mesure comme étant la base du "V" où les droites s'intersectent. Le point sélectionné est celui qui, parmi les maximums, a la position en Y la plus basse, correspondant à la mesure maximale la plus ancienne dans l'intervalle de 60 secondes. Le point sélectionné doit également posséder au moins 5 points critiques dans son voisinage proche afin d'éviter de prendre un point isolé.

Après avoir défini la base du "V", les données des maximums sont séparées en deux groupes à droite et à gauche du point et une régression linéaire est calculée sur chacun des groupes. La Fig. 4 représente l'historique des variations de température pour chaque capteur réaligné par rapport à la droite de régression. La moyenne des variations de température après la régression est calculée afin de vérifier si un refroidissement a lieu, ce qui confirmerait la présence d'un collage. Cette valeur est calculée pour les deux groupes dans les paramètres *variation* à droite et à gauche. La dernière valeur calculée par notre modèle est la moyenne des variations de température sur les deux dernières secondes de la plage de temps. Ce paramètre (T) permet de détecter si les températures remontent. Ce qui signifie qu'un décollage a lieu.

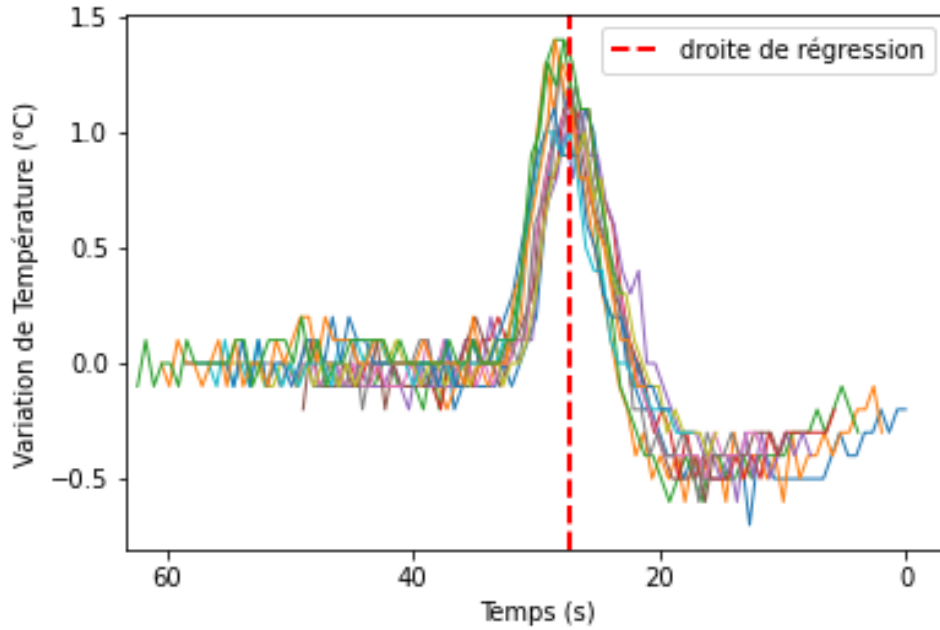


FIGURE 4 – Historique des variations pour chaque capteur réaligné par rapport à la droite de régression

3.3.2 Vérifications

Le modèle effectue des vérifications sur toutes les valeurs calculées préalablement afin de détecter la présence d'un collage et sonner l'alarme. Le résultat de l'analyse est facilement visualisable par la sortie du modèle sous forme d'un tableau de bord qui devient orange lorsqu'un collage est suspecté et rouge lorsqu'il est confirmé, le tableau de bord restant vert lorsqu'il n'y a rien à signaler. Une préalarme (couleur orange dans le tableau de bord) est déclenchée si :

- Les pentes des deux régressions sont de signes opposés et dans une plage de valeurs allant de -15 à 15.
- Les coefficients de détermination des droites (R^2) sont supérieurs à 0,8.
- La moyenne des variations de température après les droites est négative.

Si ces conditions sont respectées pendant plus de 5 secondes, une alarme est alors déclenchée (couleur rouge dans le tableau de bord). Si, après avoir sonné l'alarme, la moyenne des variations de température sur les deux dernières secondes est positive, un décollage est suspecté et le système repasse en préalarme (orange). Le diagramme présenté à la Fig. 5 reprend les différentes étapes effectuées par le modèle pour détecter les collages.

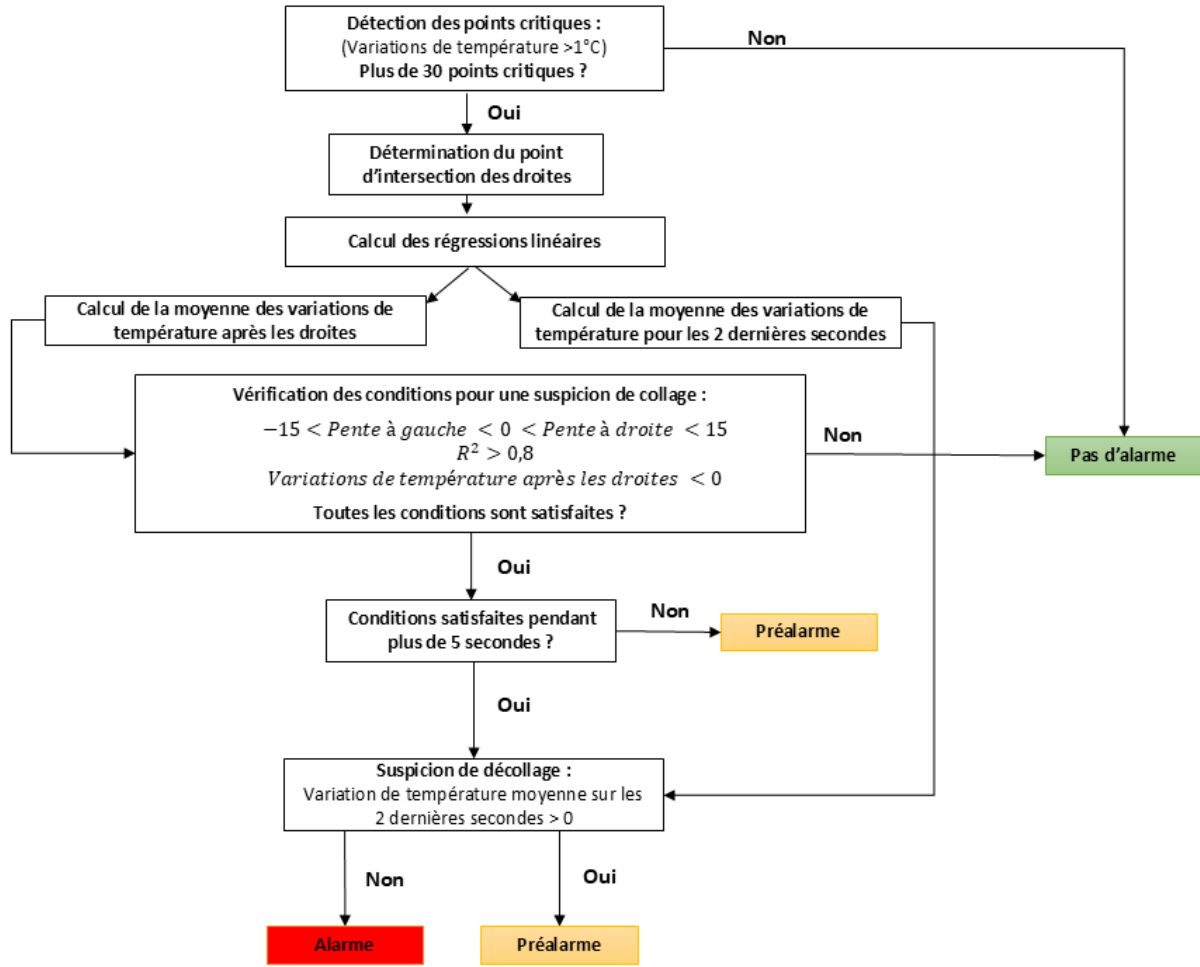


FIGURE 5 – Description générale du modèle permettant de détecter un collage.

3.3.3 Fonctions

Nous utilisons plusieurs fonctions dans notre modèle, toutes regroupées dans la librairie "*fun.py*" dont une brève description de chacune d'entre-elles est donnée ici. Pour plus de détails concernant ces fonctions, voir annexe 1.

- ***process_data*** : Traite les données de température des capteurs de fibre optique pour calculer les variations de température sur une période de temps donnée de 60 secondes. Des coordonnées X-Y sont ensuite assignées aux données en fonction de la vitesse de coulée et de la position des capteurs dans la lingotière. Pour finir, les données des capteurs en dehors de la largeur de coulée sont exclues.
- ***plot*** : Crée une représentation graphique intuitive avec un fond coloré selon le niveau d'alarme, une échelle de couleurs pour interpréter les variations de température, des lignes de régression (s'il y a lieu) et la valeur de l'indice de détection (voir section 3.4 pour l'explication de cet indice).
- ***linear*** : Calcule les régressions linéaires et renvoie les valeurs de pentes et de R^2 .
- ***temp*** : Calcule la moyenne des variations de température après les régressions.

3.3.4 Paramètres

TABLE 1 – Paramètres utilisés et valeurs attribuées

Paramètre	Description	Valeur Fixée
Seuil	Valeur à partir de laquelle un point de mesure est considéré comme critique.	1 °C
High	Nombre minimal de points critiques pour calculer les régressions.	> 30 points
slope_L et slope_R	Pentes des droites de régressions (gauche et droite)	$-15 < slope_L < 0 < slope_R < 15$
R^2	Coefficient de détermination minimal pour valider une régression.	> 0.8
variation_L et variation_R	Variation moyenne de température après la régression (gauche et droite).	< 0 °C
T	Moyenne des variations de température sur les deux dernières secondes.	> 0 °C

Les principaux paramètres utilisés dans notre modèle sont repris dans la table 1. Le nombre de points critiques minimum à partir duquel les calculs sont effectués (*High*) à été fixé arbitrairement à 30 de façon à limiter le nombre de faux positifs (détection d'un collage lorsqu'il n'y en a pas) sans trop retarder le moment de détection des collages. Les bornes de pentes acceptées pour la détection des collages ont été fixées de manière à capter toute la gamme de pentes observée lors de collages, avec une pente négative à gauche et positive à droite. Les valeurs de variations de température après les droites (*variation_L* et *variation_R*) doivent être négatives pour bien confirmer la signature d'un collage avec une diminution des températures après leur soudaine augmentation (Fig. 4). Si la moyenne des variations de température des deux dernières secondes (*T*) est positive, cela signifie que la température remonte et que l'acier s'est probablement décollé. Le niveau d'alerte peut alors redescendre. Les valeurs des paramètres le plus critiques (*Seuil* et R^2) seront discutées plus en détails dans l'analyse de sensibilité (section 4.4).

3.4 Indice de certitude de détection (DCI)

L'Index de Certitude de Détection (DCI) est conçu pour mesurer le niveau de certitude quant à la présence d'un collage. Il varie entre 0 et 1. Un DCI de 0 indiquant qu'il n'y a pas de collage, et de 1 qu'un collage est certain. Il est recalculé pour chaque plage de temps analysée en suivant la progression illustrée dans la figure 6.

1. **Initialisation** : Pour chaque plage de temps de 60 secondes, l'indice est initialisé à 0.
2. **Partie 1** : L'indice augmente proportionnellement à la valeur de *High*, le nombre de points critiques. Dans cette première partie, l'indice prend la valeur minimale entre $(High/100)$ et 0.35, ce qui représente le nombre de points critiques détectés. Lorsque la valeur de l'indice atteint 0.35, cela signifie qu'il y a 35 points critiques détectés, donc plus de 30, et que l'algorithme va enclencher les calculs pour vérifier la présence d'un collage ou non (voir section 3.3.1).
3. **Partie 2** : Après avoir atteint la valeur de 0.35, le DCI est incrémenté de 0.1 si des droites de régressions ont pu être calculées. Cela signifie que les points critiques détectés ne sont pas des maximums de variation de température isolés mais qu'une géométrie de V apparaît sur le profil des variations de température. Ensuite, une incrémentation de 0.1 est ajoutée à la valeur du DCI pour chaque seconde où une préalarme sort du tableau de bord (couleur orange). Enfin, lorsqu'une alarme est enclenchée, l'indice passe directement à 1 puisqu'un collage est confirmé (couleur rouge du tableau de bord).
4. **Décollage** : Lorsqu'un décollage survient, l'algorithme sortira à nouveau une préalarme (couleur orange sur la tableau de bord). Le DCI est alors redescendu à la valeur de 0.6, correspondant à l'état de préalarme.
5. **Retour à l'initialisation** : Finalement, lorsqu'il n'y a plus d'alarme ou de préalarme, l'indice se remet à 0 et recommence les étapes de la partie 1.

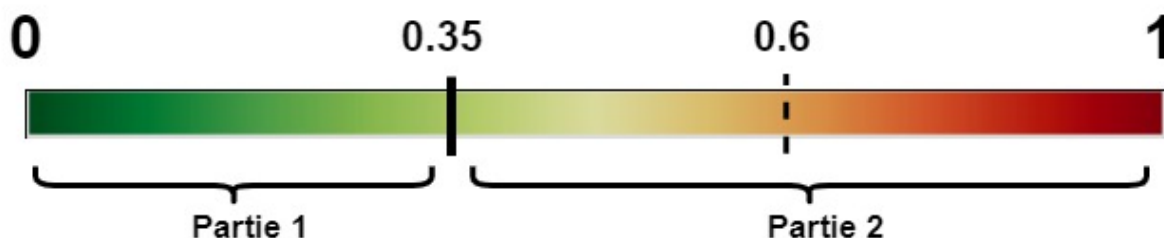


FIGURE 6 – Représentation des différentes parties du calcul de l'indice.

4 Résultats et discussion

4.1 Caractérisation du bruit

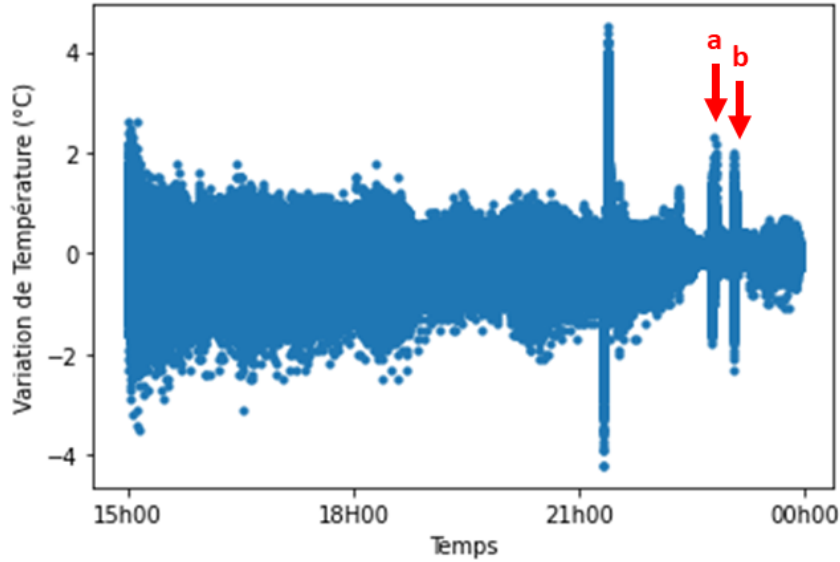


FIGURE 7 – Variations de température sur un jeu de données provenant de l’entreprise *AMDG* contenant deux collages a et b (indiqués par des flèches rouges).

L’évaluation de la performance du modèle a été faite sur base de données de température mesurées dans la lingotière au cours du temps dans deux entreprises : *AMDG* et *Industeel*. Une première analyse des variations de température sur un jeu de données a permis de comparer l’amplitude du bruit, correspondant aux variations de températures ne provenant pas d’un phénomène de collage, avec les variations enregistrées dans le cas de deux collages (Fig. 7). Nous pouvons observer que les variations de température correspondant à du bruit sont du même ordre de grandeur que lors de collage. Face à ces premières observations, une technique de détection consistant à repérer les outliers s’écartant trop du bruit de fond pour identifier la présence d’un collage s’est révélée inutilisable.

4.2 Cas détectés

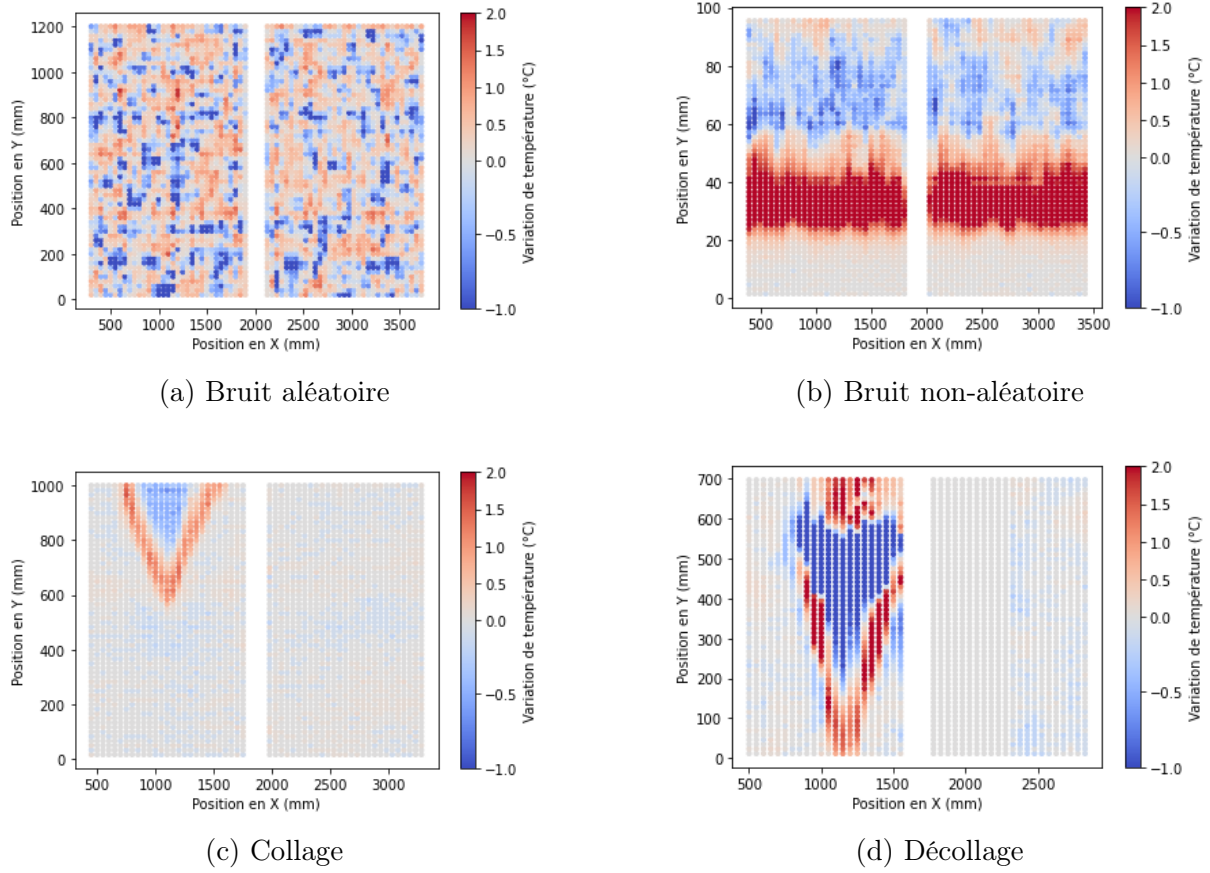
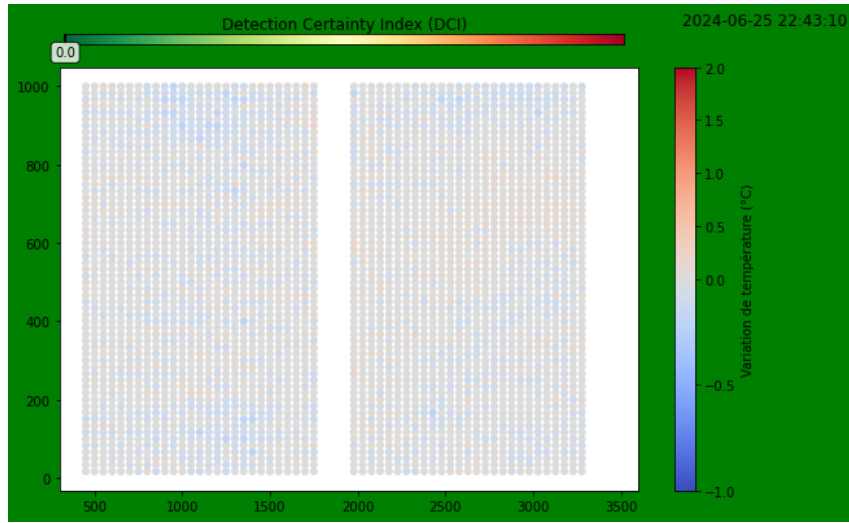


FIGURE 8 – Exemples illustrant l’historique des variations de température sur 60 secondes pour différents cas. Les données proviennent de l’entreprise *AMDG* pour les cas a, b et c et de l’entreprise *Industeel* pour le cas d.

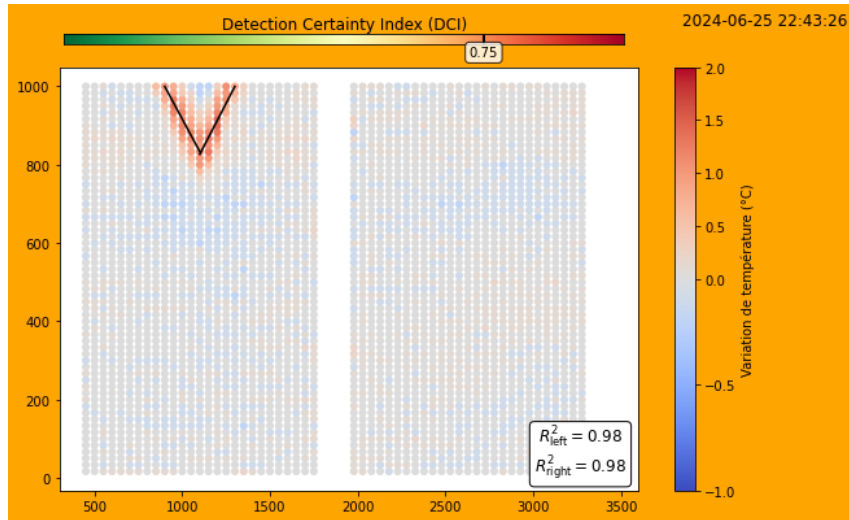
L’ensemble des données analysées contenait majoritairement des cas de variations de températures correspondant à du bruit sans présence de collage (Figs.8a et 8b) mais également quelques exemples de collage (Fig. 8c) ainsi que de décollage (Fig. 8d). En utilisant les valeurs de paramètres décrites dans la table 1, notre modèle a détecté avec succès tous les événements de collage. Par ailleurs, il n’a signalé aucun collage dans les données de bruit, démontrant ainsi sa fiabilité et son aptitude à éviter les fausses détections susceptibles de générer des alertes erronées dans la gestion de la coulée continue.

4.3 Sorties du modèle

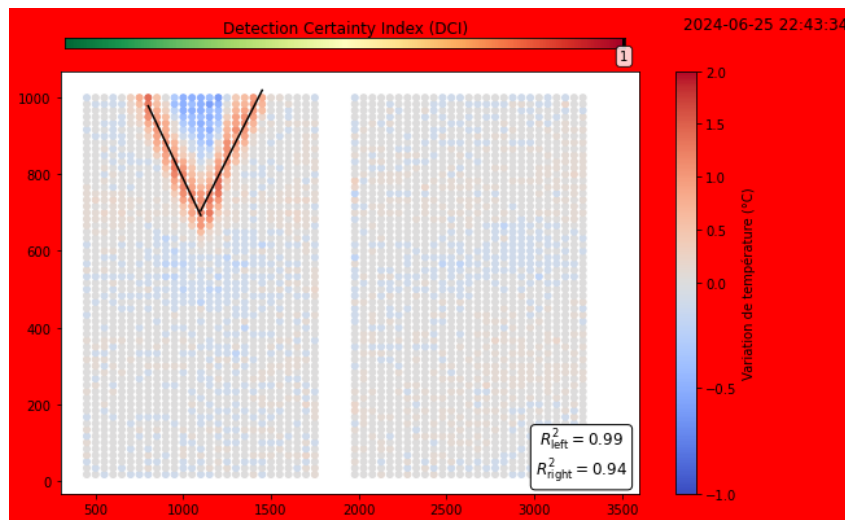
Le modèle permet de visualiser l’évolution de la température dans la lingotière, ce qui permet une détection aisée des collages (Fig. 9). Grâce à ce tableau de bord, l’utilisateur est en mesure de suivre l’évolution d’éventuels processus de collage, en ayant à sa disposition l’indice de détection ainsi que la couleur du tableau de bord comme aide à la décision.



(a) Pas d'alarme, tableau de bord en vert, DCI à 0.



(b) Alarme orange, suspicion de collage, DCI en augmentation.



(c) Alarme rouge, collage confirmé, DCI à 1.

FIGURE 9 – Exemples de sorties du modèle sous forme de tableaux de bords pour les 3 types de situation présentant l'historique des variations de température sur 60 secondes.

4.4 Analyse de sensibilité

TABLE 2 – Nombre de faux positifs selon les paramètres de *Seuil* et de R^2

Seuil/R^2	0.5	0.6	0.7	0.8	0.9
0.6	30	28	28	28	0
0.8	3	3	0	0	0
1	1	1	0	0	0
1.2	0	0	0	0	0
1.4	0	0	0	0	0

TABLE 3 – Nombre de secondes durant lesquelles une alarme de détection d'un collage (orange ou rouge) est maintenue selon les paramètres de *Seuil* et de R^2

Seuil/R^2	0.5	0.6	0.7	0.8	0.9
0.6	66	66	66	66	66
0.8	66	66	66	66	66
1	56	56	56	56	56
1.2	33	33	33	33	33
1.4	0	0	0	0	0

L'impact des paramètres de *Seuil* et de R^2 a été analysé sur le nombre de fausses alarmes pour un jeu de données de bruit sans collage dans un premier temps ainsi que sur la rapidité de détection des collages pour un jeu de données présentant un collage dans un deuxième temps. Ces deux paramètres sont les plus influents sur les sorties du modèle. En effet, le *Seuil* est le point de départ de la détection, il définit à partir de quelle valeur un point de mesure est considéré comme point critique. Le R^2 quant à lui est une mesure importante de la détermination d'une droite de régression.

La table 2 reprend le nombre de fausses alarmes obtenues sur un échantillon de bruit (sans collage) pour différentes valeurs des paramètres de *Seuil* et de R^2 . La table 3 quant à elle reprend le nombre de secondes durant lesquelles une alarme (orange ou rouge) est maintenue sur des données contenant des collages en fonction des différentes valeurs de *Seuil* et de R^2 . En effet, à chaque pas de temps d'une seconde, le modèle définit à nouveau la situation dans la lingotière sur base des conditions décrites dans la section 3.3.1. Si un collage est détecté, il y aura donc une alarme pour chaque seconde jusqu'à ce que les conditions ne soient plus respectées, ce qui définit la fin du collage. Plus le nombre de secondes est élevé, plus l'alarme a été maintenue longtemps et donc plus le collage a été détecté tôt. Un nombre plus faible signifie donc une détection plus tardive du collage.

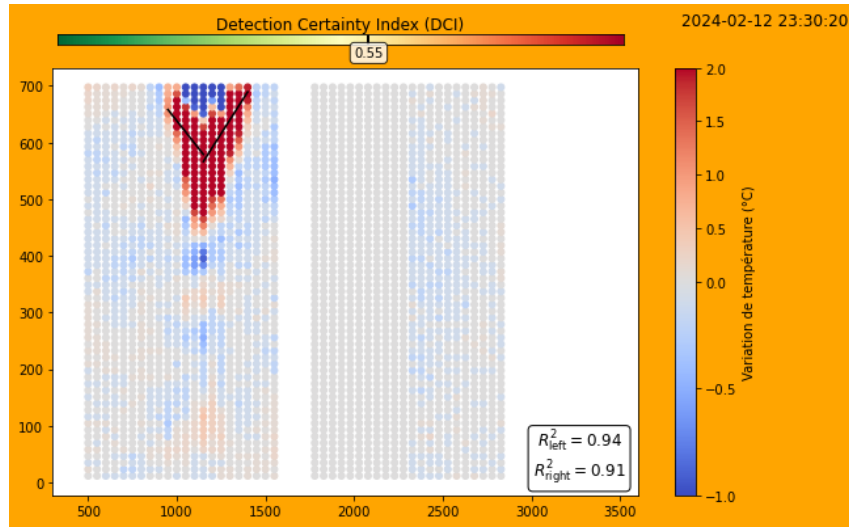
La table 2 montre que plus les valeurs de *Seuil* et de R^2 sont élevées, plus le nombre de fausses alertes diminue. Cependant, la table 3 nous montre que si le *Seuil* est trop haut, le collage est détecté plus tardivement (près de trente secondes plus tard lorsque le *Seuil* est à 1.2) ou peut même ne pas être détecté du tout, ce qui est le cas lorsque le *Seuil* est élevé à 1.4. On peut donc déduire de cette analyse que la valeur minimale de R^2 imposée n'affecte pas la détection des collages mais limite le nombre de faux positifs

tandis que la valeur de *Seuil* affecte à la fois le nombre de faux positifs mais aussi la détection des collages. Ces 2 tables nous permettent donc de fixer les valeurs de *Seuil* et de R^2 respectivement à 1°C et 0.8.

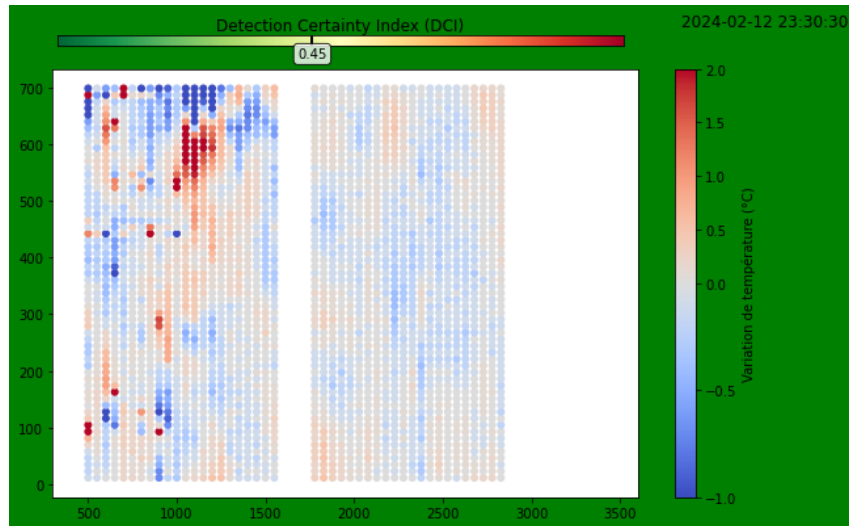
4.5 Améliorations

Les résultats obtenus avec ce modèle montrent que les techniques de régression linéaire permettent de modéliser de manière satisfaisante la géométrie associée au phénomène de collage. En effet, la forme en "V" étant très caractéristique, elle constitue un élément facilement exploitable pour la détection des collages.

Néanmoins, des pistes d'amélioration de ce modèle existent, notamment par l'intégration des mesures issues de la deuxième fibre optique, située plus bas dans la lingotière. Actuellement, le modèle repose uniquement sur les données fournies par une seule fibre optique. L'ajout des mesures de la deuxième fibre pourrait non seulement renforcer la robustesse du modèle, mais aussi permettre de vérifier le bon fonctionnement de la première fibre en cas de défaillance ou de signal ambigu. Cependant, la qualité des mesures variant fortement d'une fibre à l'autre, l'application du même modèle sur la deuxième fibre pour confirmer les résultats de la première n'est pas concluante. En effet, certains collages détectés sur la première fibre ne l'étaient pas sur la deuxième en raison des différences de mesures (Fig. 10). Néanmoins, le modèle pourrait reposer sur les mêmes principes en adaptant les valeurs des paramètres pour mieux intégrer et exploiter les données issues de la deuxième fibre.



(a) Mesures de la première fibre



(b) Mesures de la deuxième fibre

FIGURE 10 – Différences de mesures entre les deux fibres pour un même phénomène, décalé de 10 secondes en raison de la distance entre les fibres.

5 Conclusions et perspectives

L'objectif de ce projet était d'élaborer un modèle informatique capable de détecter la présence de collage dans une lingotière sur base des variations de température mesurées par une fibre optique insérée dans celle-ci, tout en évitant les fausses alarmes. Nous avons donc construit un outil de détection qui utilise la géométrie du processus de collage en se basant sur 4 règles d'identification faisant appel à la variation de température, les pentes des droites détectées sur les maximums des variations par régression linéaire, le coefficient de détermination R^2 de ces droites ainsi que la présence d'une diminution des températures suivant les maximums. Nous pouvons conclure que l'algorithme présenté dans ce rapport répond à l'objectif fixé même si quelques améliorations sont encore possibles. Le modèle actuel est déjà capable de détecter les phénomènes de collage et permet une visualisation aisée de leur avancée. Celui-ci ne génère pas non plus de fausses alarmes qui auraient ralenti la production et sont donc à éviter.

Références

- DUAN, H., WANG, X., BAI, Y., YAO, M., & GUO, Q. (2019). Integrated Approach to Density-Based Spatial Clustering of Applications with Noise and Dynamic Time Warping for Breakout Prediction in Slab Continuous Casting. *Metallurgical and Materials Transactions B*, 50(5), 2343-2353. <https://doi.org/10.1007/s11663-019-01633-w>
- EBDS ENGINEERING. (2015, mai). EBDS Engineering - Sticker Breakout Detection Process. Récupérée novembre 18, 2024, à partir de <https://www.youtube.com/watch?v=owpDEw4lJEo>
- EBDS ENGINEERING. (2024, juillet). EBDS ENGINEERING EMERALD Comparison TC vs 1 Fiber. Récupérée novembre 18, 2024, à partir de <https://www.youtube.com/watch?v=a0AL28pBWSs>
- LI, Q., WEN, G., CHEN, F., TANG, P., HOU, Z., & MO, X. (2024). Irregular initial solidification by mold thermal monitoring in the continuous casting of steels : A review. *International Journal of Minerals, Metallurgy and Materials*, 31(5), 1003-1015. <https://doi.org/10.1007/s12613-023-2798-z>
- MAZUMDAR, S., & RAY, S. K. (2001). Solidification control in continuous casting of steel. *Sadhana*, 26(1), 179-198. <https://doi.org/10.1007/BF02728485>
- TRAN, H.-S., CASTIAUX, E., & HABRAKEN, A.-M. (2020). 2D thermal finite element analysis of sticker breakout in continuous casting. *Procedia Manufacturing*, 50, 376-383. <https://doi.org/10.1016/j.promfg.2020.08.069>
- VIJAYARAM, D. T. (2019). Production of Continuous Castings in Foundries. https://www.researchgate.net/publication/355915987_Production_of_Continuous_Castings_in_Foundries

A Fonctions utilisées

A.1 *process_data*

Description générale

Cette fonction prend en entrée des données temporelles provenant de capteurs de température et de vitesse, et les traite pour créer un tableau de variations de température en fonction de la position et du temps. Elle est utilisée pour analyser les changements de température au fil du temps dans un cadre spatial (coordonnées X et Y) sur une période donnée. Cette fonction est utile pour visualiser et analyser la distribution des variations de température dans un contexte spatial en utilisant un système de coordonnées X-Y basé sur la position des capteurs et la progression temporelle.

Entrées

- *df* : DataFrame contenant les données des capteurs avec des colonnes pour le temps, la largeur, la vitesse et les températures mesurées par les capteurs.
- *start_time* : Temps de début de la période d'analyse.
- *end_time* : Temps de fin de la période d'analyse.
- *captors* : Nom du groupe de capteurs (ex. 'A', 'B', 'C', etc.).
- *n_capt* : Nombre de capteurs par face.

Étapes principales

1. **Filtrage temporel** : Sélection des données entre *start_time* et *end_time*.
2. **Calcul des variations de température** : Pour chaque capteur, la variation de température est calculée comme la différence entre la mesure actuelle et la mesure précédente.
3. **Création d'une grille de positions X et Y** : Les positions X des capteurs sont fixes, tandis que les positions Y sont calculées en fonction de la vitesse moyenne sur la période étudiée.
4. **Construction d'un DataFrame final (*final_df*)** : Contient les coordonnées (X, Y) et les variations de température pour chaque capteur, à chaque instant.
5. **Filtrage des capteurs hors de la largeur maximale** : Les capteurs situés en dehors de la largeur maximale sont exclus.
6. **Ajustement des positions X des capteurs** : Les positions X des capteurs actifs sont modifiées pour correspondre à la configuration finale.

Sorties

- *final_df* : Un DataFrame contenant trois colonnes : X qui est la position horizontale du capteur ; Y, la position verticale calculée en fonction du temps et de la vitesse moyenne et Variation_t qui est la variation de la température par rapport à la mesure précédente.
- *Speed* : Vitesse moyenne sur la période étudiée, ajustée par un facteur de division.

A.2 *plot*

Description générale

La fonction `plot` génère un tableau de bord visuel permettant d’analyser et de visualiser les variations thermiques détectées par les capteurs. Elle prend en entrée des données sur les coordonnées des capteurs (X , Y), leurs variations de température (`Variation_t`), ainsi que des paramètres optionnels tels que des prédictions de régression linéaire pour chaque côté et des coefficients de corrélation (R^2). La fonction crée une représentation graphique intuitive avec un fond coloré, une échelle de couleurs pour interpréter les variations, des lignes de régression (s’il y a lieu) et un indicateur visuel, le *Detection Certainty Index (DCI)*, qui varie en fonction des conditions vérifiées.

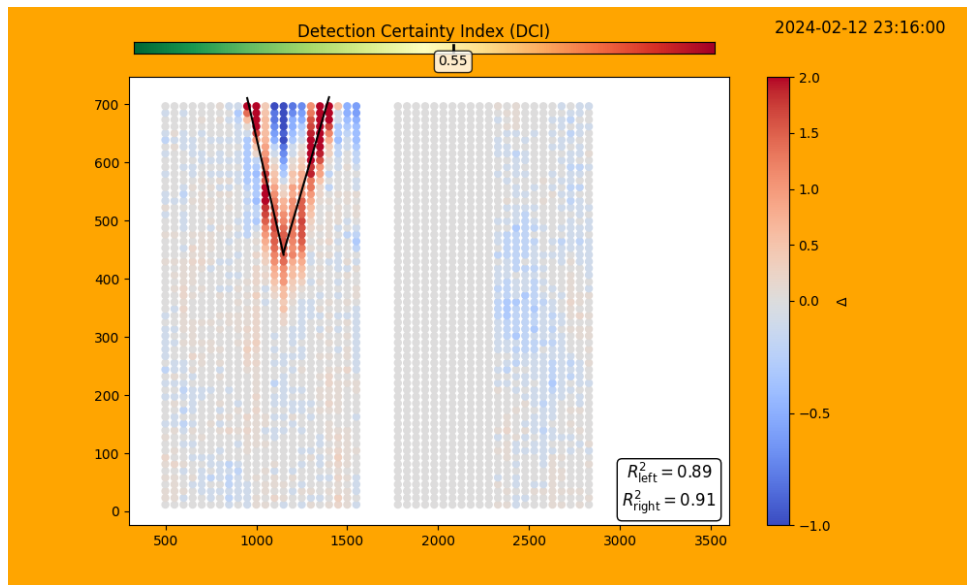


FIGURE 11 – Exemple de tableau de bord affiché par la fonction `plot`.

Entrées

La fonction `plot` accepte les paramètres suivants :

- `df` : Le jeu de données principal contenant les colonnes X , Y (coordonnées des capteurs) et `Variation_t` (variations de température).
- `end_time` : Une chaîne de caractères qui indique le temps et la date représenté sur le tableau de bord.
- `background_color` : La couleur d’arrière-plan utilisée pour indiquer soit le vert (pas d’alarme), l’orange (pré-alarme) ou le rouge (alarme).
- `X_left`, `y_left_pred` : Les coordonnées X et les prédictions Y pour la régression linéaire du côté gauche, s’ils sont calculées.
- `X_right`, `y_right_pred` : Les coordonnées X et les prédictions Y pour la régression linéaire du côté droit, si disponibles.
- `r2l` : Le coefficient de corrélation R^2 de la régression pour le côté gauche.
- `r2r` : Le coefficient de corrélation R^2 de la régression pour le côté droit.
- `coefficient` : Un indice normalisé (entre 0 et 1) indiquant la certitude de la détection, appelé : *Detection Certainty Index (DCI)*.

Étapes principales

La fonction `plot` suit les étapes suivantes pour générer le tableau de bord :

1. **Création de la figure principale** : La fonction initialise une figure et un axe principal avec des coordonnées (X , Y) et une échelle de couleurs représentant les variations de température (`Variation_t`). Cela génère une carte colorée permettant de visualiser ces variations de manière intuitive avec une échelle de couleurs (`colormap`).
2. **Vérification des paramètres de régression** : Si les données de régression linéaire (`X_left`, `y_left_pred`, `X_right`, `y_right_pred`) sont fournies, la fonction trace les lignes correspondantes sur la figure. Les coefficients de corrélation R^2 pour les côtés gauche (`r2l`) et droit (`r2r`) sont également affichés dans un encadré.
3. **Affichage de l'indice de certitude** : Un rectangle en dégradé de vert à rouge est ajouté en haut de la figure pour représenter visuellement le *indice de certitude*, et cette valeur est donc affichée.
4. **Affichage du tableau de bord** : Enfin, le tableau de bord complet est affiché à l'écran.

A.3 *linear*

Description générale

Cette fonction applique une régression linéaire entre deux ensembles de données, **X** et **Y**. Elle utilise `LinearRegression` de `sklearn` pour calculer la droite de régression. Les résultats incluent les valeurs de **X**, les prédictions associées de **Y** (**y_pred**), la pente de la droite (**slope**) et le coefficient de détermination (R^2) qui mesure la qualité de l'ajustement.

Paramètres

- **x** : Série de données correspondant aux coordonnées X .
- **y** : Série de données correspondant aux valeurs Y .

Sorties

- **X** : Les valeurs transformées de **x** sous forme matricielle.
- **y_pred** : Les valeurs prédites de Y basées sur la régression linéaire.
- **slope** : La pente de la droite de régression.
- **r2** : Le coefficient de détermination (R^2), qui évalue la précision de la régression.

Utilisation

Cette fonction est utilisée pour effectuer les régressions linéaires séparées sur les deux côtés (gauche et droit), et donc obtenir les paramètres .

A.4 *temp*

Description générale :

La fonction `temp` calcule la variation moyenne de température après un pic pour chaque capteur dans une liste fournie. Elle prend en entrée la liste de capteurs a droite et a gauche du base de "V" et le jeu de données, puis identifie les pics de température et calcule la variation moyenne après chaque pic.

Paramètres :

- `L` : Liste des capteurs (`X`) à analyser, ce sont les capteurs avec les valeurs de variations maximales en dessus de la base de "V".
- `df` : Le jeu de données principal avec les variations de température.

Étapes effectuées :

- **Initialisation** : on commence par créer une liste vide `variations_after_peak` est créée pour stocker les variations après le pic de température.
- **Itération sur les capteurs** : La fonction parcourt chaque capteur dans la liste `L`, filtre les données pour ce capteur et localise le pic de température avec `idxmax()` sur `Variation_t`.
- **Sélection des données après le pic** : on selectionne apres les variations apres le pic.
- **Calcul de la variation moyenne** : Les variations de température après le pic sont ajoutées à la liste `variations_after_peak`, et la moyenne de ces variations est calculée.
- **Retour du résultat** : La fonction renvoie la variation moyenne après le pic (`vap`).

Valeur retournée :

- `vap` : La variation moyenne de température après le pic pour les capteurs spécifiés, pour pouvoir détecter si on a une chute après le "V".