|  |  |  |  |
| --- | --- | --- | --- |
|  |  | | |
| To | Bernard Beauzamy | | |
| Company | Société de Calcul Mathématique SA | | |
| From | G. Fricout, A. Gay | | |
| Date | Date | | |
| Pages | 1 of 3 | | |
| Copy |  | | |
|  |  | | |
| Subject | **Description des données fournies pour le projet d’usure des Rouleaux de Laminage** | | |
| Ref |  | Programme :  Axis: | Project: |

# Rapide rappel du contexte

Explicité plus précisément dans les annexes au contrat

L'étude consiste à prédire l'usure de rouleaux de laminage à chaud. Ces rouleaux sont utilisés pour réduire l'épaisseur de brames d'acier, pour les faire passer de 20cm à 1-3mm d'épaisseur en une douzaines de 'cages', chacune constituées de deux rouleaux de travail (un en haut 'top', un en bas 'bottom'). On se penche plus précisément ici sur les deux dernières cages (6e et 7e cages 'finisseuses', notées F6 et F7). Quatre rouleaux sont donc étudiés ici, nommés F6t, F6b, F7t et F7b.

Le laminage des brames se composent de campagnes successives, chaque campagne étant constitué de 100 à 300 brames environ, passées à la suite. Les rouleaux sont changés à la fin des campagnes. **L'usure des rouleaux n'est mesurée qu'une fois qu'ils sont changés, donc à la fin de chaque campagne.**

Le dataset comprend donc les détails des brames constituant chaque campagne et les usures mesurées sur les rouleaux à la fin de chacune de ces campagnes. Le dataset est composé de 395 campagnes pour un total de 51 710 brames.

## Format de l'usure

L'usure de chaque rouleau est mesurée après la campagne sur toute la longueur du rouleau, obtenant ainsi un profile d'usure, tel que visible en exemple sur la Figure 1.

L'estimation du profile complet étant très complexe voire impossible en l'absence de donnée localisée pour les brames, le travail pourra se concentrer sur une étude de l'usure moyenne au centre du cylindre, appelée par la suite *usure au centre*. Cette usure au centre est aussi représentée sur la Figure 1.

Une image contenant carte, texte

Description générée avec un niveau de confiance très élevéUne image contenant carte, texte

Description générée avec un niveau de confiance très élevé

Figure 1 – Deux exemples de profils de campagnes.

# Organisation des fichiers

Les données liées aux campagnes de laminage et d'usures des rouleaux sont stockées dans deux fichiers principaux et un dossier :

* WearDataForDatamining.xlsx
* WearCentres.xlsx
* Wear\_profiles\ (dossier)

## WearDataForDatamining.xlsx

Ce fichier Excel contient les données et caractéristiques de l'ensemble des brames et campagnes.

### Onglet Strips\_data

Le premier onglet 'Strips\_data' contient la liste de toutes les brames, chacune identifiée par son numéro de campagne (de 1 à 395) et son numéro au sein de la campagne (commençant à 1 jusqu'à maximum 306). Pour chaque brame, de nombreuses caractéristiques physiques sont présentes, en cas de non-compréhension de la nature physique d'une des variables, ne pas hésiter à revenir vers les équipes d'ArcelorMittal. Quelques colonnes remarquables sont :

* Strip Grade Family : Il s'agit d'une catégorisation réalisée 'à la main' par les équipes du site concerné.
* Strip lengths : Il ne faut pas oublier que lors du laminage, l'épaisseur de la brame diminuant, sa longueur augmente. Ainsi, pour chaque brame, sa longueur en sortie de la cage 6 n'est pas la même que celle en sortie de la cage 7. La longueur en sortie de la cage 5 correspond donc à celle en entrée de la cage 6.
* Shifting : Cette valeur correspond à un décalage de la brame par rapport au centre du rouleau, dans un objectif d'une meilleure répartition des forces le long du rouleau. Dans le cas d'une étude du profil d'usure, cette valeur peut être importante, auquel cas une discussion avec les équipes d'ArcelorMittal est recommandée pour clarifier les sens de décalage par rapport au rouleau.
* Colonnes AS à BN : Ces dernières valeurs correspondent à d'autres mesures ou à d'autres modèles (propres à l'usine) pour certaines valeurs, d'où leur mise à part.

### Onglet Campaign\_data

Le second onglet 'Campaign\_data' contient les données propres à chaque campagne, et peuvent être croisées avec celles du premier onglet, grâce au numéro de campagne (de 1 à 395). Quelques colonnes remarquables :

* Line Up : Comme pour les familles de brames, il s'agit d'une catégorisation des campagnes.
* Colonnes J à M : Nom du fournisseur et modèle du cylindre. Cette information peut être intéressante en tant que catégorisation (clustering) des rouleaux, et est utilisée dans les modèles internes d'ArcelorMittal.
* Colonnes Y à AH : Décompte pour chaque campagne du nombre de brames appartenant à chaque famille, spécifiées dans l'onglet précédent.

## WearCentres.xlsx

Ce fichier Excel contient tout simplement la mesure de l'usure au centre pour les quatre rouleaux de chaque campagne, toujours identifiées par leur numéro de 1 à 395.

Beaucoup de valeur sont manquantes, et (d'après quelques observations personnelles), certaines pourraient être aberrantes (notamment quelques valeurs négatives). Mais le tri et traitement est laissé au soin du Data Analyst et les données sont livrées brutes.

## Dossier Wear\_profiles\

Ce dossier contient un fichier Excel par campagnes (quelques campagnes sont manquantes néanmoins), chaque fichier contenant les mesures des profils d'usure pour les quatre rouleaux.

Les profils d'usures sont mesurés une fois le rouleau enlevé de la ligne et refroidi, avant qu'il ne soit corrigé pour revenir à une forme 'neuve' et réutilisé pour une prochaine campagne. Ils sont mesurés sur 64 points le long du rouleau, et dans les fichiers, une abscisse est liée à chaque point de mesure.

Chaque fichier contient deux onglets f6 et f7 (les premiers onglets) contenant les mesures pour les rouleaux *top* et *bottom*. Pour chaque rouleau (F6t, F6b, etc.) sont présents le profil avant la campagne (FIRST FRESH GROUND), après la campagne (Adjusted Profile) et après correction (avant d'être réutilisé pour une prochaine campagne, Real Profile - 2ND FRESH GROUND). La valeur qui nous intéresse étant l'usure, c'est à dire la quantité de matière enlevée, elle est obtenue en soustrayant le profile après campagne au profil avant campagne. **Cette mesure de l'usure est présente dans la colonne appelée *Top Wear, mm* ou *Bottom Wear, mm***