DAS Departamento de Automação e Sistemas
CTC Centro Tecnológico
UFSC Universidade Federal de Santa Catarina

On-line diagnostic tool for hot strip mill

Relatório submetido à Universidade Federal de Santa Catarina

como requisito para a aprovação da disciplina:

DAS 5511: Projeto de Fim de Curso

Renato Kendy Murata

Florianópolis, Novembro de 2019

On-line diagnostic tool for hot strip mill

Renato Kendy Murata

Esta monografia foi julgada no contexto da disciplina

DAS 5511: Projeto de Fim de Curso

e aprovada na sua forma final pelo

Curso de Engenharia de Controle e Automação

Prof. Ubirajara Franco Moreno

____

Banca Examinadora:

Van Thang Pham
Orientador na Empresa

Prof. Ubirajara Franco Moreno
Orientador no Curso

Prof. Fabio Luis Baldissera
Responsável pela disciplina

Bernardo Schwedersky Barancelli, Avaliador

Gustavo Ardigó de Souza, Debatedor

Eduardo Marschall, Debatedor

Dedico este trabalho ao meu professor de francês Michel Abes (in memorian).
Sua vocação em ensinar e transmitir experiências me fez me encantar com um outro país.
Merci beaucoup!

Acknowledgements

This work represents the end of one cycle that started in September of 2017 with the double degree program in France. I would like to thank everyone that make that experience as splendid as it was.

First, I would like to thank my family, specially my father Nelson, my mother Marineuza and my sister Cintia, to the unconditional support they gave me in my academic life. Without their incentive, I would never have left my hometown at age 14 to dream bigger.

I would like to thank the amazing friends that I met during this experience. They made me feel like home even being in another country. Xiaowen, Gao, Alejandra, Rodrigo, Sibel, Humberto, Iryna, Caio, Lucas, Carmen, Teresa, Vitor, Mônica, Isaac, Josué, Débora, Diego, Laura and all the others. I was lucky to experienced so many great moments with you. A special thank to Corentin, for all the support and partnership and for being the person that I can always count with.

To my brazilian university UFSC, the french *école* CentraleSupelec and all the inspiring teachers that I had the honor to learn from them, my sincere gratitude. A special thanks to my supervisor during this program Prof. Ubirajara Franco Moreno for all the support and disponibility.

Finally, I would like to thank all the COSMO team from ArcelorMittal Maizières Research and specially my supervisor Van Thang Pham. You were all very kind, patient encouraging. Thanks to you, my first professional experience was enriching and pleasant.

Resumo

A indústria siderúrgica está em constante evolução. O processo de laminação a quente, introduzido no início do século XX, revolucionou a indústria do aço, tornando o custo de produção de laminas de aço significativamente menores. O processo consiste em transformar grandes barras de aço em chapas extremamente finas através de uma série de rolos. Com o advento dos sistemas de controle, a automatização do processo de laminação a quente deixou o processo ainda mais eficiente e rápido. Novas tecnologias trazem novos desafios, nos processos de laminação a quente atuais, os rolos compressores chegam a uma velocidade de 120 metros por minuto. Nessas condições, qualquer defeito ou falha no sistema deve ser detectado o mais rápido possível, para evitar danos ou produtos defeituosos. O presente trabalho apresenta diferentes métodos para implementar um sistema de detecção de falhas. Primeiramente é desenvolvido um sistema para detecção de falha nos *"loopers"*. Esse sistema consiste em analisar os sinais adquiridos nas fábricas com diferentes métodos de processamento de sinal, notadamente estatística descritiva, STFT e EMD, e usando técnicas de aprendizado de máquina, classificar as informações extraídas dos sinais em dois grupos: nominal (quando o sistema está em seu funcionamento normal) e falho (quando há alguma falha no sistema). Esse método provou-se eficaz na detecção de falhas. Em seguida, foi proposto um método baseado em modelo para identificação de falhas no subsistema de *"strip steering"*. No entanto, a implementação desse método pode ser terminada devido a ausência de dados experimentais, necessários para a validação do modelo matemático do sistema.

Palavras-chave: Laminação a quente. Sistema de detecção de falhas. Processamento de sinal. Aprendizado de máquina.

Abstract

The steel industry is constantly evolving. The hot strip mill process, introduced in the early twentieth century, reshaped the steel industry, making the cost of producing steel sheets significantly lower. The process consists in transform thick steel slabs into thin coils using a series of compressing rolls. With the advent of control systems, the automation of the hot rolling process has made it even more efficient and faster. Nowadays, the rolling speed can reach 120 meters per minute. In this condition, any default or failure must be detected as soon as possible to avoid damages and non-quality products. This document present different method to implement a fault detection system. First, it is developed one fault detection system on loopers. This system analyses the signals recorded by the plant's data acquisition system with data processing methods, notably descriptive statistics, STFT and EMD. By using machine learning techniques, the features extracted from the signals are separated into two groups: the nominal (when the system has no default) and the fault (when the system has a fault). This method proved to be efficient for fault detection on loopers. Then, it was proposed a model-based method for fault detection on the strip steering subsystem. Although, the method implementation wasn't possible due to the lack of experimental data. Those data are necessary to validate the mathematical model of the system.

Keywords: Hot strip mill. Fault detection system. Signal processing. Machine Learning

List of Figures

Figure 1 – Finishing mill plant (From [1]) 22
Figure 2 – Measurement & Control department (Source: personal archive) 24
Figure 3 – R&D Maizières campus organization (From: ArcelorMittal) 24
Figure 4 – Hot strip mill (From [2]) 27
Figure 5 – Arcelor's finishing mill schema (From [3]) 27
Figure 6 – Variables to be controlled (From [3]) 28
Figure 7 – Florange's finishing mill looper (Source: ArcelorMittal) 29
Figure 8 – Strip steering system (From [4]) 30
Figure 9 – Cobble fault (ThyssenKrupp Steel Europe AG) 31
Figure 10 – Shearing-tail fault (ThyssenKrupp Steel Europe AG) 32
Figure 11 – Plausibility test scheme (Source: personal archive) 33
Figure 12 – Hardware redundancy scheme (Source: personal archive) 34
Figure 13 – Software redundancy scheme (Source: personal archive) 34
Figure 14 – General description of the model-based fault diagnosis scheme (Source: personal archive) 35
Figure 15 – Signal processing structure (From [5]) 40
Figure 16 – Empirical mode decomposition (Source: Towards Data Science) 43
Figure 17 – Nominal and fault signals (Source: personal archive) 50
Figure 18 – Nominal and fault signals, another example (Source: personal archive) . 51
Figure 19 – Fault detection system architecture (Source: personal archive) 52
Figure 20 – Relation between the maximum value and the median of the modeling signals (Source: personal archive) 52
Figure 21 – Relation between the standard deviation and the mean value of the signals (Source: personal archive) 53
Figure 22 – STFT results (Source: personal archive) 54
Figure 23 – Application of EMD to nominal case (Source: personal archive) 55
Figure 24 – Application of EMD to fault case (Source: personal archive) 55
Figure 25 – Model validation results (Source: personal archive) 57
Figure 26 – Strip steering model (From [4]) 61
Figure 27 – Simplified Simulink model (Source: personal archive) 62

List of Tables

Table 1 – Numeric values of model validation 58
Table 2 – The most efficient techniques for fault detection on loopers 58
Table 3 – Strip steering parameters 60

List of abbreviations and acronyms

SSU: Shared Services Unit

MC: Measurement & Control department

COSMO: Control , Simulation and Models

HAGC: Hydraulic automation gauge control

HSM: Hot strip mill

FD: Fault detection

FDI: Fault detection and isolation

FDIA: Fault detection and isolation and Analysis

FDF: Fault detection filter

DO: Diagnostic observer

PRRG: Parity based residual generator

LCF: Left coprime factor

SDFA: Set of disturbances that causes false alarms

FDR: Fault detection rate

FFT: Fast Fourier transform

STFT: Short time Fourier transform

EMD: Empirical mode decomposition

SVM: Support vector machine

Contents

1	**INTRODUCTION**	**21**
1.1	**Motivation of the project**	**21**
1.1.1	ArcelorMittal	22
1.1.1.1	ArcelorMittal Global R&D	23
1.1.1.2	ArcelorMittal Maizières Research SA	23
1.1.1.3	Measurement & Control department	23
1.1.1.4	COSMO team	24
1.1.2	Project objectives	25
1.2	**Structure of the document**	**25**
2	**HOT STRIP MILL**	**27**
2.0.1	Finishing Mill Control Systems	28
2.0.1.1	Screw position	28
2.0.1.2	Torque on the looper	29
2.0.1.3	Work roll speed	29
2.0.1.4	Strip lubrication	29
2.0.1.5	Strip steering	29
2.1	**Possible faults on Hot Strip Mill plant**	**30**
2.1.1	Faults on actuators	30
2.1.2	Cobbles	31
2.1.3	Shearing-tails	31
3	**FAULT DETECTION THEORY**	**33**
3.1	**Plausibility test**	**33**
3.2	**Hardware redundancy**	**33**
3.3	**Software redundancy**	**34**
3.3.1	Analytical model-based fault detection	34
3.3.2	Residual generation	35
3.3.2.1	General residual generator	35
3.3.2.2	Optimal fault detection filter	36
3.3.2.3	Parity space method	37
3.3.3	Fault detectability, isolability and identifiability	37
3.3.4	Structural fault detectability	37
3.3.5	Structural fault isolability	38
3.3.6	Structural fault identifiability	39
3.4	**Signal processing**	**39**

3.4.1 Symptom generation . . . 40
3.4.1.1 Descriptive statistics . . . 40
3.4.1.2 Fast Fourier transform (FFT) . . . 41
3.4.1.3 Short time Fourier transform (STFT) . . . 42
3.4.1.4 Empirical mode decomposition (EMD) . . . 42

4 DEVELOPMENT OF AN APPLICATION FOR FAULT DETECTION 45
4.1 The proposed signal processing method for fault detection . . . 45
4.1.1 Data reading . . . 45
4.1.1.1 Pre-processing . . . 45
4.1.2 Data processing . . . 46
4.1.3 Data classification . . . 46
4.1.3.1 Support Vector Machine (SVM) . . . 47
4.1.3.2 Classification Tree . . . 48
4.1.3.3 Cross-Correlation . . . 48
4.2 Implementing the signal processing method for fault detection on loopers . . . 49
4.3 Data selection . . . 50
4.3.1 Pre-processing . . . 51
4.4 Data processing . . . 51
4.4.1 Descriptive statistics . . . 52
4.4.2 Short-time Fourier Transform (STFT) . . . 53
4.4.3 Empirical Mode Decomposition (EMD) . . . 54
4.5 Data Classification . . . 56
4.5.1 Classification Tree . . . 56
4.5.2 Support vector machine . . . 56
4.5.3 Cross Correlation . . . 56
4.6 Implementation results . . . 56
4.7 Experimental results and validation . . . 57

5 MODEL-BASED FAULT DETECTION ON STRIP STEERING . . . 59
5.1 The proposed model-based method for fault detection . . . 59
5.2 Implementing the model-based method for fault detection on strip steering . . . 59
5.2.1 Mathematical equations . . . 60
5.2.2 Model implementation . . . 61
5.2.2.1 Model Validation . . . 62

6 SUMMARY AND FUTURE WORK . . . 63
6.1 Summary . . . 63

6.2 Future work 63

BIBLIOGRAPHY 65

1 Introduction

The 20th century lead the world to a new era with a chain of events that had a great impact in our society. In the 1960s a big quantity of theories was developed and had a great contribution in controls design. Nowadays, with the control theory consolidated in the industry and the automation degree of process continuously growing, there is an increasing demand for higher system perform and product quality. To increase the perform of the system and the product quality, it is required more system safety and reliability.

One way to have a safer and more reliable process is by building a fault diagnosis system. Those diagnosis systems are built to detect faults that a simply process monitoring system would not detect, like sensor faults and actuator faults. A fault diagnosis system can be developed in many ways, the main methods are the hardware redundancy schemes, plausibility test, software redundancy schemes and by signal processing. During this project, it was developed a fault detection system using the software-redundancy test and the signal processing.

The fault diagnosis system aims to detect fault in a hot strip mill plant. The plant transforms thick stabs of steel into thin coils. It is an expensive process because of the wear of actuators, that must be often changed, and due to the size of the system, any fault can be dangerous for the operators that are working next to it. The rolling process, Fig. 1, is an important method in the metal industry. The product must have very precise dimension, in the micro meter range.

1.1 Motivation of the project

The hot strip mill process consists hundred of sensors and actuators to ensure the process operation. The rolling speed can reach up to 120m/min. At that speed, any fault can have serious damages on the plant, and also have an influence on the product quality. Even with the most advanced sensors and actuators, and making a regular maintenance of the components, the process is not immune to faults.

According to a study made by SCHAFFLER on the ThyssenKrupp Steel Europe AG's facilities in Germany shown that one repair on the work rolls of the hot strip mil plant costs around €21000 euros. In one year, this problem happened 5 times, totalizing an extra charge of €105000 euros per year due to unplanned stoppage. Also, 5 unplanned roller replacement of 7 minutes each costs €35000 euros. So, in this case, the unpredictable faults in one hot strip mill during year costs around €140000 euros.

Figure 1 – Finishing mill plant (From [1])

One hot strip mill plant can produce from 360 tonnes to 3300 tonnes of steel plates per year [6]. In 2018, ArcelorMittal produced 118 million tonnes of brute steel. When we consider that just one hot strip mill costs €140000 euros per year in maintenance caused by unplanned stoppage, anything that can be done to improve the time to identify defaults and optimize the maintenance would have a big financial impact.

1.1.1 ArcelorMittal

ArcelorMittal is the world's leading steel and mining company. With more than 200.000 employees in 60 countries. ArcelorMittal is present in the world as no other steelmaker, producing 91,9 million of tons of steel and producing one revenue of US$96,03 billions in 2018, being placed as 123rd world's biggest company in 2017 by Fortune Global 500.

The company mine iron ore and coal, most of then are used to supply their own steel-making operations. ArcelorMittal is the leading supplier of all major markets including automotive, construction, household appliances and packing with a high quality steep products. The steel is produced all over the world. Europe is the biggest producer corresponding to 47% of the steel production, the Americas produces 38% and the other regions 18%. ArcelorMittal's products are organised according to region, each region is specialized in a set of products.

1.1.1.1 ArcelorMittal Global R&D

The research and development have a big importance in the ArcelorMittal's strategies. The main goal of the R&D sector is to realise ArcelorMittal's ambitions in technological innovation, and ensuring future growth with focus on sustainability. The company has 10 research centers distributed in North America, Europe and South America, with more than 1400 researchers. The R&D is highly business oriented, it is focused on seven key areas: automotive, packaging, construction, general industry, energy market, long products and special plates. Those centers ensures a shorter time to market and improved competitiveness.

1.1.1.2 ArcelorMittal Maizières Research SA

The project took place in Maizières's campus. The Maizière campus is the biggest global R&D campus. With 24 hectares and 45.000 m2 of laboratories, the Maizière center has 570 permanent employees and receives around 100 interns per year. The campus hosts four different research centres: Maizière Process, Maizière Products, Maizières Mining and Shared Services Unit (SSU). Those centres are specialized in automotive products, packaging, process development and mining processing.

1.1.1.3 Measurement & Control department

The internship took place in the Measurement & Control Department (MC). The department is part of the Maizière Process center. The MC department is composed of three teams: Advanced Process Instrumentation and product non destructive evaluation; Surface Properties, Image & Data processing; and COntrol, Simulation and MOdels (COSMO). The main mission of those teams is to bring relevant solutions for the different process involved in the productive chain of the steel products. Those solutions are searched in the fields of process control, instrumentation and product inspection and evaluation.

Figure 2 – Measurement & Control department (Source: personal archive)

1.1.1.4 COSMO team

The COSMO team was composed by four research-engineers and four interns. It searches for advanced control solutions for different process of the steel industry, it is also responsible for modeling and simulation of systems present in the industry. The team also works on the on-line integration for process monitoring and predictive maintenance. The project 'On-line diagnosis toll for Hot Strip Mill' is one of the team's project.

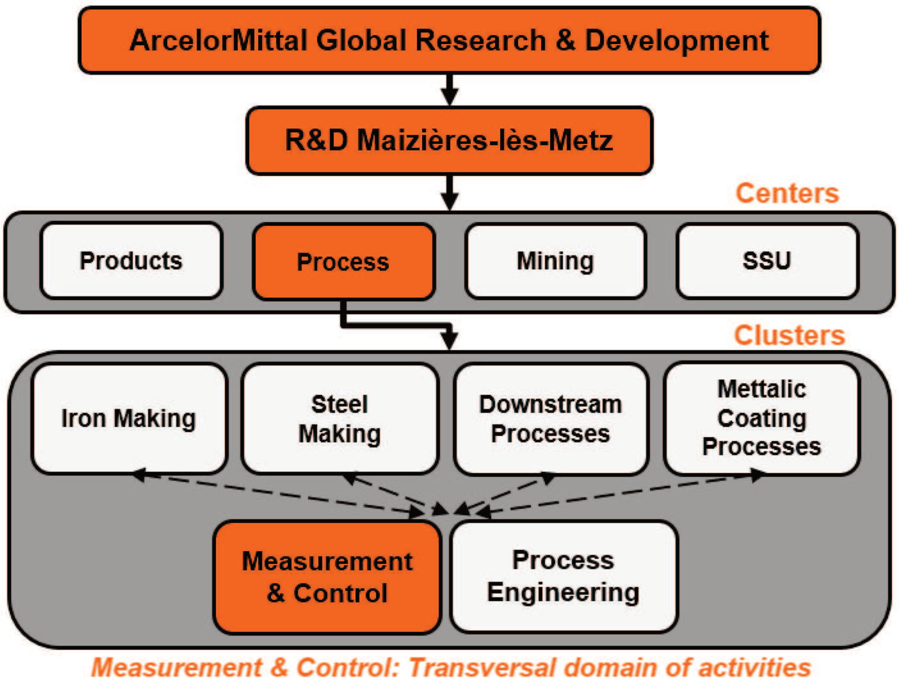


Figure 3 – R&D Maizières campus organization (From: ArcelorMittal)

1.1.2 Project objectives

The objective of the internship is to carry out a feasibility of different fault detection methods for hot strip mill process. The idea is to take advantage of the huge amount of data that is recorded with the hot strip mill plant data acquisition system to perform an on-line fault detection of the system. The methods can be model-based methods, data-oriented method as well as knowledge-based methods. To develop such a system, the following tasks are expected:

- Bibliography revision of preview works with the same subject
- Try different methods on real industrial historic data
- Develop a toolbox for facilitating further usage
- Establish a final report to capitalize the study

1.2 Structure of the document

This document is divided into six chapters. In chapter 1, the motivation of the project and its main objectives is given. The enterprise organization and the details about the place where the internship took place is also presented in chapter 1. Chapter 2 presents the hot strip mill system and its components. It also contains the possible faults on hot strip mill plant that were found in the bibliography revisions. In chapter 3 the fault detection theory is presented, this chapters gives an overview on two different methods: signal processing and model-based. Chapter 4 proposed a signal processing method for fault detection, the method is applied for the looper subsystem and its results are analyzed. In chapter 5 a model based method for fault detection on strip steering subsystem is presented. Chapter 6 summarizes the main results of this project and presents and outlook for future work.

2 Hot strip mill

Hot strip mill is one process that transforms steel slabs (10m long, 1.5m wide and 25cm thick) into thin sheets (1km long, 1m wide and 3.5mm thick). A hot strip mill is a complex process that has 8 units in the following order: Reheat furnace, roughing mill, transfer table, coilbox, crop shear, finishing mill, runout table and coiler. A scheme of this structure is presented in Fig. 4.

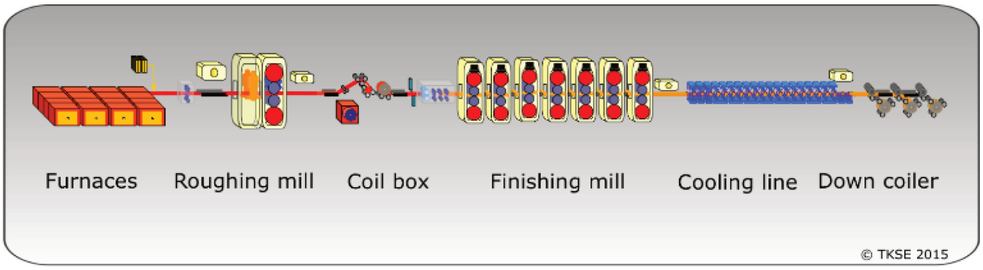


Figure 4 – Hot strip mill (From [2])

The reheat furnace increases the slab temperature to approximately 1200°C to reach the desired properties and facilitate the milling process. In the roughing mill, the slab thickness is reduced from 25cm to 3cm. The slab's length increases proportionally with the thickness reduction, considering that its width does not have major changes during the whole process. The roughing mill stand is reversing, it means that the slab is rolled back and forth until it reaches the desired thickness. Then, the slab is transported by the transfer table to the finishing mill area. The coilbox is used to reduce the difference of the temperature along the slab. Reducing the temperature drop between the head and tail ends of the strip from 100°C to around 10°C. The even temperature on the bar avoid great tension peaks in the finishing mill [7] known as bite impact.

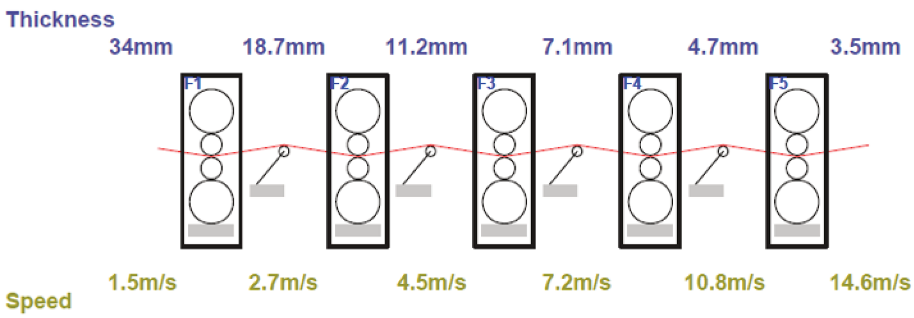


Figure 5 – Arcelor's finishing mill schema (From [3])

The finishing mill system in ArcelorMittal can have from 5 to 7 continuous rolling stands. Each stand consists in a pair of work rolls and backup rolls. The working rolls have direct contact with the slab and must be regularly exchanged because of the roll wear.

The backup rolls usually have twice the diameter of a working roll, around 1.5m, and they are used to give more mechanical stability and reduce the deformation of the working rolls. Between each stand there is one looper. The looper is used to control the tension between stands, the tension control is necessary because of the speed difference between two stands. That speed mismatch causes a tension change, upon which the looper compensate. To do it, the looper stretched the slab, dealing with the extra amount of strip. The looper function is limited, if there's a big amount of extra strip beyond the looper's reach, the strip will be stored into two stands, creating the dangerous phenomena called cobble [2]. Cobble problem is dangerous because the tension values on each stand are high. It goes from 5 MPa after the first stand to 10 MPa before the last [8].

2.0.1 Finishing Mill Control Systems

The main objectives of the controls systems are to deliver a uniform strip and guarantee a safe and continuous operation. The main measures to control the product quality it's the strip profile, width and shape. After the rolling mill process the strip profile is thicker in the middle due to roll deflection but it should be kept within 1% of the strip thickness [4]. To reach those characteristics, the finishing mill plant can have many control loops. Some variables that are controlled on the finishing mill system is shown if Fig. 6.

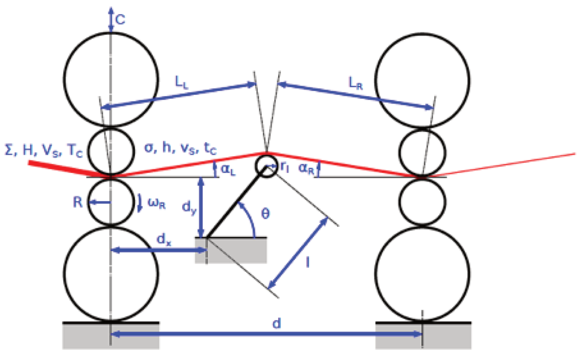


Figure 6 – Variables to be controlled (From [3])

2.0.1.1 Screw position

The rolls on the stand can move up and down using a screw. The screw position controls the thickness reduction after each stand. Its reference is calculated using the initial strip thickness and the desired final value. Then the screw position for each stand is calculated in such a way that the thickness reduction is balanced between the stands.

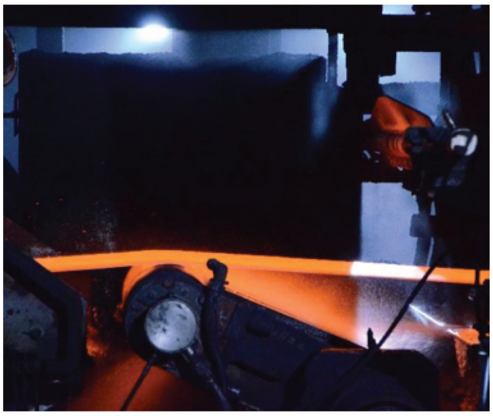


Figure 7 – Florange's finishing mill looper (Source: ArcelorMittal)

2.0.1.2 Torque on the looper

The looper deals with the extra amount of strip generated due to velocity mismatch between two stands. The torque reference is related to an inter-stand tension reference. To control the inter-stand tension, the looper controls its angle from the horizontal plane.

2.0.1.3 Work roll speed

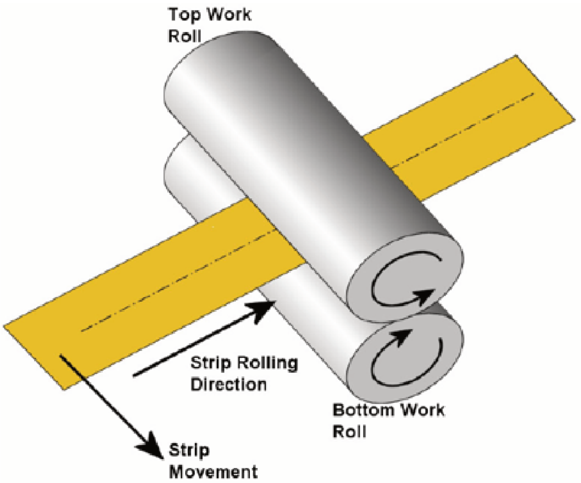
Work roll speed depends on the thickness reduction of the strip, due to the conservative mass flow. The difference between the entry rolls speed and the exit roll speed has a direct influence on the inter-stand torque on the strip. Due to this direct influence, the work roll speed is determined in function of the desired looper angle.

2.0.1.4 Strip lubrication

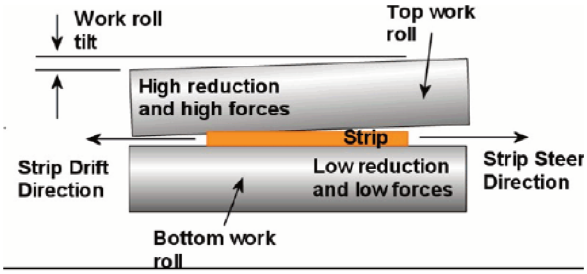
Automatic control of strip lubrication is not always applied on finishing mill plants. In theory, the lubrication reduces the surface defect, increase mill capability and save energy consumption [9]. The lubrication control is made applying a mixture with water and oil in the working rolls. This mixture influences the friction coefficient between the working roll and the strip.

2.0.1.5 Strip steering

When the strip is being rolled by the work rolls, there's a strip movement perpendicular to the direction of rolling occurs [10]. To compensate this movement generated by the rolling, it is changed the screw position of just one side of the rolls, the same side of the direction of the perpendicular movement, as shown in Fig. 8a. The difference of the screw position on each side applies more tension in just one side of the strip, the strip moves to the side where the tension is lower, correcting the strip trajectory, Fig. 8b. Applying more tension in one side of the strip to correct its trajectory is called strip steering. This control can be made manually by a highly trained operator or by an automatic control.



(a) Strip movement



(b) Strip steering

Figure 8 – Strip steering system (From [4])

2.1 Possible faults on Hot Strip Mill plant

From the bibliography research that was made, we found the most commons faults that may occurs in the hot strip mill process. The fault diagnosis system on hot strip mill is a field that is still developing, so it was not available a big quantity of articles and thesis about this theme. The most relevant results of this research are discussed below.

2.1.1 Faults on actuators

On rolling force control systems, a fault diagnosis observer for a hydraulic automation gauge control (HAGC) is presented on [7]. The HAGC is responsible for the rolling force control in most of the HSM from ArcelorMittal. Some of the most common faults in hydraulic automation gauge control systems is the change of valve core response time, the leakage of hydraulic cylinder, dump coefficient of cylinder and dump coefficient of roller system of mill. On [7], the whole rolling force control system is modeled. Then, each fault dynamic is expressed in function of the system's variables. One fault diagnosis observer is made for each fault, in a way that each observer is sensible for just one fault. A monitoring tool to quantify valve stiction presents in control loops is proposed on [8]. The monitoring tool uses the plausibility test to identify is the malfunction of the valve is caused by stiction. When the valve stiction reaches defined thresholds the monitoring tool notifies the operator about maintenance. Work roll eccentricity is another fault that may occur [9]. The main effects of work roll eccentricity are oscillations with twice the rotational frequency.

2.1.2 Cobbles

A cobble is a deviation in the strip travel between two stands [2]. It happens when the tolling process is run as usual, but suddenly the strip bows up between the stands. The looper can't control the tension anymore and the strip doesn't pass through the rolls, pilling up the strip between two steps and creating a dangerous situation to the operator next to it. A cobble is a severe fault in the rolling process. When it happens, the rolling process must stop to repair the process manually. It takes from 30 minutes up to a whole day [2].

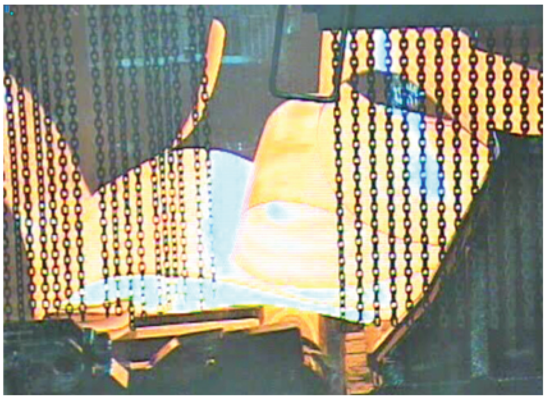


Figure 9 – Cobble fault (ThyssenKrupp Steel Europe AG)

2.1.3 Shearing-tails

This fault is described on [2]. It happens when the end of the strip breaks. When the damaged strip is rolled, it leaves marks in the working rolls. The damaged working rolls must be exchanged, otherwise it will leave marks on the surface of the consecutive strips.

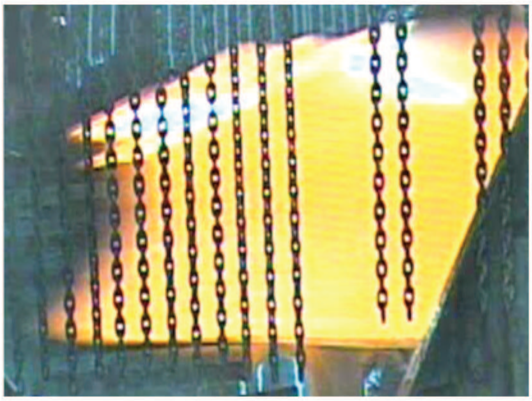


Figure 10 – Shearing-tail fault (ThyssenKrupp Steel Europe AG)

3 Fault detection theory

There are three different types of fault diagnosis systems. The FD (fault detection), FDI (fault detection and isolation) and FDIA (fault detection and isolation and analysis). The first one is the most basic, it just detects the occurrence of faults that could be occurred in any part of the process and lead to undesired behaviour. The FDI detects the faults and isolate it, so the system can localize which fault occurred, analysing the abnormal behaviour of the system. The FDIA isolate the faults and determine its type, magnitude and cause. The main techniques for fault detection are explained below.

3.1 Plausibility test

This is the simplest model. It is based on the check on physical laws to verify if the component working property. It requires knowledge of the component and its efficiency is limited to faults that generates a loss of plausibility. The plausibility test is most recommended for fault detection only, the fault identification and isolation are limited to some special cases.

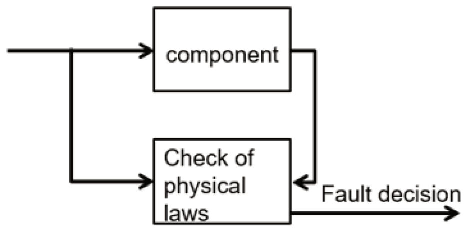


Figure 11 – Plausibility test scheme (Source: personal archive)

3.2 Hardware redundancy

This scheme uses one copy of the component to check its state. A fault is present in the system if the difference between the output of both components is important. This scheme has a high reliability and the fault isolation is automatic . But for some process, like the hot strip mill, the components are huge and expensive. Creating one physical copy of it will require too many resources. This method is more indicate for critical systems where fault can have severe results, as in the aerospace industry.

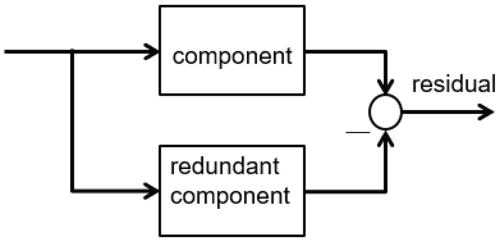


Figure 12 – Hardware redundancy scheme (Source: personal archive)

3.3 Software redundancy

The software redundancy aims to copy the hardware redundancy but spending less resources. Instead of physically duplicate the component, it is made one mathematical model of it. The mathematical model can be made by using different modelling methods. The software redundancy method gives much more information about the fault than the plausibility test method, and have the same advantages of the hardware redundancy but using less resources. There's many ways to build a mathematical model of the process, the analytical model-based scheme is discussed below.

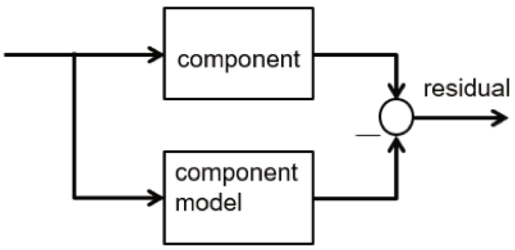


Figure 13 – Software redundancy scheme (Source: personal archive)

3.3.1 Analytical model-based fault detection

Analytical model-based fault diagnoses was originated in the early 70's, with a high influence of the observer theory established at that time. Since then, its efficiency has been demonstrated by different applications in the aeronautical and automotive industry [11].

In this case, the main goal is to build a model that describes the expected behaviour of the system, without any faults. The same input applied to the real system is applied to the process model and the difference of both outputs,called residual, is compared. The residual is analysed and then the system will decide about the state of the system, as shown in Fig. 14.

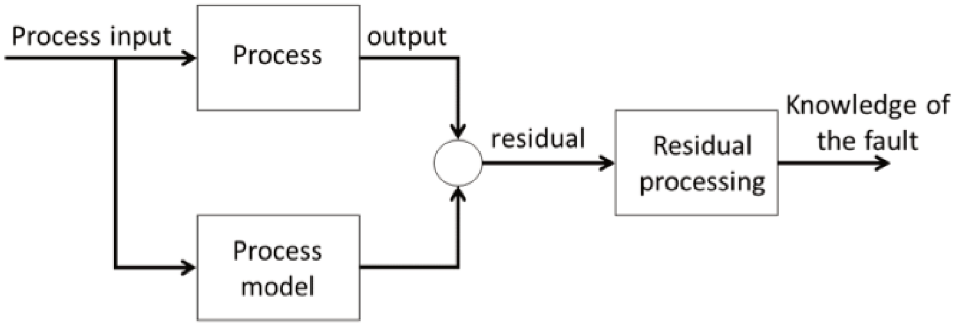


Figure 14 – General description of the model-based fault diagnosis scheme (Source: personal archive)

3.3.2 Residual generation

The residual generation is the first step of the fault diagnosis system. The residual is the difference between the output of the real system and the output of the analytical model, and it carries information about possible faults that has occurred in the system.

It has numerous methods to generate residuals like the fault detection filter (FDF), diagnostic observer (DO), parity based residual generator (PRRG) and general residual generation. All those methods share the same objectives.

Firstly, the residual signal should must be decoupled from the input. It means that when there are no faults, disturbances or model imprecision the residual must converge to zero. The influence of unknown inputs and model imprecision should be minimized. The facility of the on-line implementation and the design facility should be considered. Three methods for residual generation are presented below, all those methods were given by [5].

3.3.2.1 General residual generator

This approach follows the intuitive solution that the residual is calculated only doing the difference between the measured output and the expected output of the nominal plant.

$$r(s) = y(s) - \hat{y}(s) = y(s) - G_{yu}(s)u(s) \tag{3.1}$$

Where $r(s)$ is the residual, $y(s)$ is the measured output and $\hat{y}(s)$ is the expected output. This implementation is intuitive but it does not consider the influences of the system initial states and system model uncertainty. The residual model can be extended to the following expression.

$$r(s) = y(s) - \hat{y}(s) + C(sI - A)^{-1}x(0) + \Delta y(s) \tag{3.2}$$

The term $C(sI - A)^{-1}x(0)$ represents the initial state of the system and $\Delta y(s)$ the model uncertainty. To suppress the influence of the model uncertainty, one solution is to make one closed loop structure as one output observer with the following structure.

$$\widehat{y}(s) = G_{yu}(s)u(s) + L(s)(y(s) - \widehat{y}(s)) \tag{3.3}$$

Rewriting the previous expression in the state space model, we must define the L matrix ensuring $A - LC$ stable and with the desired dynamic. It was proven on [5] that the residual of the closed loop structure presented can be expressed by the formula

$$r(s) = \widehat{M}_u(s)y(s) - \widehat{N}_u(s)u(s) \tag{3.4}$$

where the transfer matrices $\widehat{M}_u$ and $\widehat{N}_u$ are the left coprime factor (LCF) of the $G_{yu}(s)$ function and can be obtained with the following expression.

$$\widehat{M}_u = (A - LC, -L, C, I), \widehat{N}_u = (A - LC, B - LD, C, D) \tag{3.5}$$

This method is easy to design but it requires complex computation, which can difficult the on-line realization. Choosing L is also a problem, and it was the subject of various residual generation approached in the last three decades.

3.3.2.2 Optimal fault detection filter

In this section the optimal tuning of the fault detection filter is presented. This regulation minimizes the set of disturbances that causes false alarms (SDFA) in function of a given fault detection rate (FDR). Having the following state space system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + E_d d(t) + E_f f(t) \\ y(t) &= Cx(t) + Du(t) + F_d d(t) + F_f f(t) \end{aligned} \tag{3.6}$$

we must generate residuals using one FDF of the form

$$\begin{aligned} \dot{x}_a &= Ax_a(t) + Bu(t) + L(y(t) - y_a(t)) \\ y_a(t) &= Cx_a(t) + Du(t) \\ r(t) &= V(y(t) - y_a(t)) \end{aligned} \tag{3.7}$$

The optimal L and V gains can be obtained using the following equations

$$\begin{aligned} L_{opt} &= (E_f F_f' + Y_f C')(F_f F_f')^{-1} \\ V_{opt} &= (F_f F_f')^{\frac{1}{2}} \end{aligned} \tag{3.8}$$

with Y_f being the solution of the Riccati equation

$$AY_f + Y_f A' + E_f E_f' - (E_f F_f' + Y_f C')(F_f F_f')^{-1}(F_f E_f' + CY_f) = 0 \tag{3.9}$$

This function can be numerically solved using the function care from MATLAB®.

$$\begin{aligned} A_e' X E_e + E_e' X A_e - (E_e' X B_e + S_e) R_e^{-1} (B_e' X E_e + S_e') + Q_e = 0 \\ A_e = A' \quad B_e = C' \quad Q = E_f E_f' \\ R = F_f F_f' \quad S = E_f F_f' \quad E = I_{size(A)} \end{aligned}$$

3.3.2.3 Parity space method

As shown on [11] parity equations are a straightforward method for fault detection. Developed on the 80's, this method is simple do construct and implement, it's a method indicated to detect additive faults and it has similar results of more complicated methods presented in previous sections.

First we need a model $G_m(s)$ of the process $G_p(s)$. The residual can be calculated by the following formula

$$r(s) = [G_p(s) - G_m(s)]u(s) + G_p(s)f_u(s) + f_y(s) \tag{3.10}$$

If the model is accurate, $G_m(s) = G_p(s)$, so the first part of the equation is equal to zero. The residual value is only different to zero if the additive input fault $f_u(s) \neq 0$ and the additive output fault $f_y(s) \neq 0$

3.3.3 Fault detectability, isolability and identifiability

A model-based fault detection cannot detect all the faults of the system, there are some conditions have to be validated to detect, isolate and identify a fault, in a FDI point of view. The theorems used in this section were obtained on [5].

3.3.4 Structural fault detectability

A fault is detectable if its occurrence, independent of its size and type, would cause a change in the nominal behaviour of the system output. To define in which conditions a fault can be detectable, we recall the following theorem

Theorem 1. *Given the system*

$$\dot{x} = (A + \Delta A_f)x + (B + \Delta B_f)u + E_f f$$

$$y = (C + \Delta C_f)x + (D + \Delta D_f)u + F_f f$$

Where f is the additive fault vector and $\Delta A_f, \Delta C_f, \Delta B_f$ and ΔD_f are the multiplicative faults given by

$$\Delta A_f = \sum_{i=1}^{l_A} A_i \Theta_{A_i}, \Delta B_f = \sum_{i=1}^{l_B} B_i \Theta_{B_i}$$

$$\Delta C_f = \sum_{i=1}^{I_{C_f}} C_i \Theta_{C_i}, \Delta D_f = \sum_{i=1}^{l_D} D_i \Theta_{D_i}$$

An additive fault f_i is detectable if and only if

$$C(sI - A)^{-1}E_{f_i} + F_{f_i} \neq 0$$

with E_{f_i}, F_{f_i} denoting the i-th column of matrices $E_f F_f$ respectively, a multiplicative fault Ω_{A_i} is detectable if and only if

$$C(sI - A)^{-1}A_i(sI - A)^{-1}B \neq 0$$

a multiplicative fault Θ_{B_i} is detectable if and only if

$$C(sI - A)^{-1}B_i \neq 0$$

A multiplicative fault Θ_{C_i} is detectable if and only if

$$C_i(sI - A)^{-1}B \neq 0$$

A multiplicative fault Θ_{C_i} is detectable if and only if

$$D_i \neq 0$$

From Theorem 1 , we can conclude that an additive fault is detectable if the transfer function between the fault and the system output is not zero. A multiplicative fault Θ_{D_i} is always detectable and the detectability of fault Θ_{B_i} can be interpreted as input controllability and the Θ_{C_i} the observability. The multiplicative fault Θ_{A_i} can change the system dynamics.

It is evident that detecting additive faults can be realized independent of the system input, and multiplicative faults is possible do be detected if the input signal is not zero. This conclusion explain why sensor faults are usually modelled as additive faults and actuators faults as multiplicative faults.

3.3.5 Structural fault isolability

In this context, it is considered a system with the influence of two faults, unknown inputs are not taken into account. We take two faults, ξ_i ,ξ_j ,$i \neq j$ are possible to isolate if the changes in the system output caused by these two faults are distinguishable. To check the conditions of fault isolability, we now use a the following corollary.

Corollary 1. *Given the system used in theorem 1, then ξ with fault transfer matrix*

$$G_\xi(s) = [G_{\xi 1}(s)...G_{\xi l}(s)]$$

Is structurally isolable if and only if

$$rank(G_\xi(s)) = \sum_{i=1}^{l} rank(G_{\xi i}(s))$$

For additive faults for example, this result means that the faults are isolable only if the number of faults is not larger than the number of sensors. With multiplicative fault, the fault transfer matrix is different so it may demand more sensors. It is possible to archive fault isolation even if this corollary is not satisfied by removing some assumption as having previous knowledge of faults, or assuming that a simultaneous occurrence of faults is impossible.

3.3.6 Structural fault identifiability

Fault identifiability characterizes the mapping from the system output to the faults under consideration. If this mapping is unique, then the faults are identifiable. The formal definition is presented below.

Definition 1. *Given the system used in theorem 1, then ξ with fault transfer matrix*

$$G_{\xi}(s) = [G_{\xi 1}(s) ... G_{\xi l}(s)]$$

the fault vector is called structurally identifiable if $G_{\xi}(s)$ is invertible and its inverse is stable and casual.

It is a necessary condition the fault transfer function be invertible and stable, otherwise the fault identifiability structure will be equivalent to the fault isolability scheme.

3.4 Signal processing

The basic architecture for fault detection using signal processing is shown is Fig. 15. In this case we consider that the system's outputs carry information about faults. The symptom generation will catch all relevant information about the possible faults present in the signal. Typical faults usually have influence on time domain functions, like mean values, limit values, standard deviation or frequency domain functions like spectral power densities, etc. The main drawback for signal processing methods is its inefficiency with dynamic systems, where the input signals changes very often. In other hand, this method is easy to implement and it does not requires much information about the system.

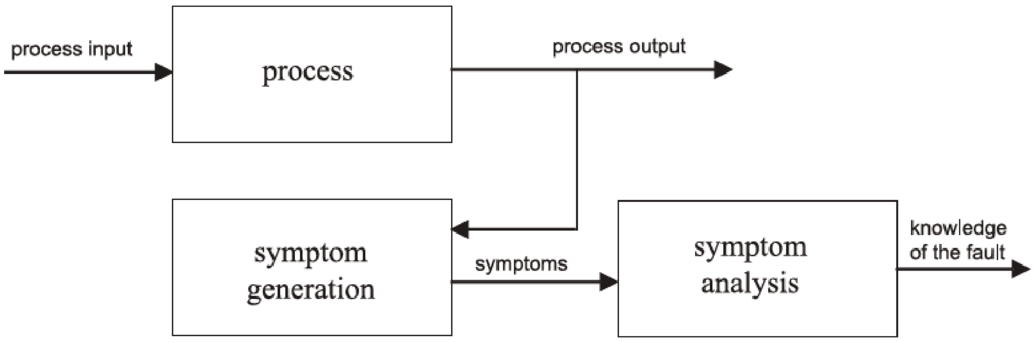


Figure 15 – Signal processing structure (From [5])

3.4.1 Symptom generation

The main methods for symptom generation are the time domain functions, frequency domain functions and time-frequency domain functions. The most common functions used for features extraction of a signal are presented below.

3.4.1.1 Descriptive statistics

Some values that can have information about faults can be easily extracted using time domain functions. Examples of functions are expressed below.

- Arithmetic mean

  If we have a vector with the values $a_1, a_2, \ldots, a_n$, the arithmetic mean A is defined by the formula:

$$A = \frac{1}{n}\sum_{i=1}^{n} a_i = \frac{a_1 + a_2 + \cdots + a_n}{n}$$

- Median

  The median is the middle value between the higher and the lower value of a list of values.

- Mode

  The mode is the common value of a lift of values.

- Standard deviation

The standard deviation calculates the amount of dispersion of a list of values and can be calculated by the following formula

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

where $\bar{x}$ is the arithmetic mean and N is the amount of values of the list.

- Variance

  The variance is the expectation of the squared deviation of a random variable from its mean [12].

$$Var(X) = \mathrm{E}\left[(X-\mu)^2\right]$$

- Range

  The range of a set of data is the difference between the largest and smallest values [12].

- Limit values

  The maximum and the minimum value of a list of values.

3.4.1.2 Fast Fourier transform (FFT)

The fast Fourier transform is the discrete version of the Fourier transform. The Fourier transform convert a signal in time domain into the frequency domain. The Fourier transform plays a very important roll in signal processing because it represents a signal as sum of complex sinusoids. Having a signal described as a sum of complex sinusoids, it is possible to make a spectral analysis of the signal.

The spectral analysis is a important component of signal processing. For example, it can give information on the wear of mechanical components of one system by monitoring its vibration, and it can also identifies cyclic behaviours hidden in a signal.

The fast Fourier transform (FFT) is one algorithm to perform a Fourier transform of a dataset. The FFT can be calculated by the following formula [13] with $y(t)$ being the signal.

$$Y(k) = \sum_{t=1}^{N} y(t)e^{-i2\pi kt/N} \qquad k = 0, \ldots, N-1,$$

The fast Fourier transform is recommended for stationary signals because it applies the formula over the entire signal.

3.4.1.3 Short time Fourier transform (STFT)

The short time Fourier transform is another method based on the Fourier transform. This method instead of applying the Fourier transform in the whole signal, as the FFT does, it splits, over the time, the entire signal into smaller parts. So the FFT is applied for each part of the signal.

This is a solution for spectrum analysis when it comes to non-stationary signals. When we take the whole signal and divide it into smaller parts, each part is considered to be stationary, so it makes sense to apply the FFT over the smaller part. As a result, we can extract information about the evolution of the most important frequencies of the signal over the time. The STFT is considered to be a time-frequency method because it carries information about both variables.

The discrete STFT can be calculated with the following formula [2]

$$X(m,\omega) = \sum_{n=-\infty}^{\infty} x[n]w[n-m]e^{-j\omega n}$$

where $X(m,\omega)$ are the Fourier coefficients depending on time index m and frequency ω, x is the signal and w is the window function.

3.4.1.4 Empirical mode decomposition (EMD)

The empirical mode decomposition is another time-frequency method for signal processing of non-stationary signals. It's an empirical method, different from STFT and FFT, and it decomposes the signal into intrinsic mode functions (IMF).

IMFs represents specific oscillations of the original signal [2]. This method was developed by NASA's engineer Norden E. Huang on 1988. According to its definition, one IMF should satisfy two conditions "... the number of extrema and the number of zero crossings must either equal or differ at most by one" and "the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero" [14] The EMD method can be perform by the following algorithm [15]

1. Initialize: r_1 = x (t) , and i = 1

2. Extract the i^{th} IMF

   a) Initialize: $h_{i(k-1)} = r_i$, $k = 1$

   b) Extract the local extrema and the minima of $h_{i(k-1)}$

   c) Cubic spline interpolation of local extrema from upper and lower envelopes of $h_{i(k-1)}$

d) Calculate the mean $m_{i(k-1)}$ of the upper and lower envelopes of $h_{i(k-1)}$

e) Let $h_{ik} = h_{i(k-1)} - m_{i(k-1)}$

f) If h_{ik} is an IMF then set $IMF_i = h_{ik}$, else go to step (b) with $k = k + 1$

3. Define $r_{i+1} = r_i - IMF_i$

4. If r_{i+1} still has least 2 extrema then go to step (2) else decomposition process is finished and r_{i+1} is the residue of the signal

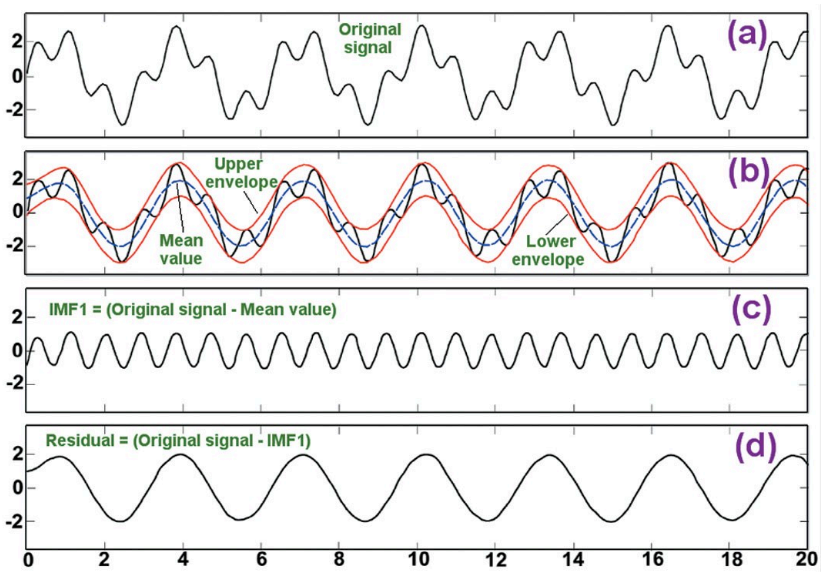


Figure 16 – Empirical mode decomposition (Source: Towards Data Science)

4 Development of an application for fault detection

In this chapter, it is proposed a signal processing method for fault detection. The method combine signal processing techniques with machine learning algorithms. Its implementation for fault detection on loopers is shown in section 4.2. Then the results of the validation are discussed.

4.1 The proposed signal processing method for fault detection

The hot strip mill plant, as presented on section 2, has a lot of sensor and actuators, the signals transmitted and emitted by those components contains information about the system state. On ArcelorMittal's steel mills, those signals are stored using the IBA® measurement system. The main idea is to take advantage of the huge amount of signals that were record to build a signal processing fault detection system.

The proposal is to build a fault detection system that can be divided into three parts:

1. Data reading
2. Data processing
3. Data classification

4.1.1 Data reading

This is the most simple task of the proposed method. The system will be developed on MATLAB® platform, and the signals are recorded using the IBA® measurement system. As the acquisition system is already complete and working, it's just necessary to convert the IBA® files into MATLAB® ".dat" format.

4.1.1.1 Pre-processing

The hot strip mill plant record hundreds of signals not all of them are useful or contains information about the state of the system. In this part, the useful signals for fault detection are selected and the useless part of those signals are removed. Some signals like the input of the motor that controls the looper for example, when the looper is not used, the signal is zero and it doesn't have any useful information.

The pre-processing take care of the useless signals. It is an important task because the signal processing requires a considerable computational effort, like the Fourier or Wavelet transforms. When the data is clean, less computational effort is required so the system is more efficient and faster.

4.1.2 Data processing

The data processing is the most important part of the method. It's necessary to use the right methods for features extraction, so the data classification will have more success and the fault detection system will be more precise and accurate.

For features extraction, the proposed technique is to try different methods studied on section 3.4.1:

- Descriptive statistics
- Short time Fourier transform
- Empirical mode decomposition

Those three methods were chosen because they have completely different characteristics. Descriptive statistics like the mean, variance and standard deviation are very easy to calculate and can give a lot of information about the signal. The STFT is a theoretical tool based on the Fourier transform and it's possible to analyse the evolution of the frequencies over the time of one signal. The EMD is different from both of them, it's a empirical method and can be useful when the signal is too complex or non-stationary.

4.1.3 Data classification

The features extracted by the methods presented in the previous section need to be classified into fault or nominal data. As there's an important amount of signals recorded, it was chosen two machine-learning methods: support vector machine and classification tree. The cross-correlation method is used to classify the EMD results, following the signal processing method presented on [5]. In this section, a basic theory about those classification method is given.

4.1.3.1 Support Vector Machine (SVM)

The support vector machine is a learning model that can be used for data classification. In our case, we have to give the data from the feature extraction labeled into two groups: the nominal and the fault data. The SVM algorithm will create an optimal hyperplane to separate the both data. The optimal hyperplane created to separate the training data can be used to classify new data. So, by using a big amount of labeled and recorded data, it is possible to train a SVM model to classify the new data.

The mathematical basis to develop a system with such characteristics is given by [16].
The data needed for training is given in the following form

$$(y_1, x_1)...(y_n, x_n) \quad x \in \mathbb{R}^n, y \in \{-1, +1\} \tag{4.1}$$

The hyper plane is defined as

$$H(\omega, b) = \{\forall x | \omega^T x + b = 0\} \tag{4.2}$$

where b is the parameter needed to calculate the normal distance of the hyperplane to the origin and ω is the normal vector to the hyperplane. The prameters in Equation 4.2 are scalable, meaning that

$$H(\omega, b) = \{\forall x | c\omega^T x + cb = 0\} \tag{4.3}$$

with $c \in \mathbb{R}$ and $c \neq 0$ lead to the same hyperplane as Equation 4.2 The optimal hyperplane has to fulfil the condition below

$$min_{x_i} |\omega^T x_i| \stackrel{!}{=} 1 \tag{4.4}$$

The euclidean distance of a point x_i to the hyper plane is

$$d(H; x) = \frac{|\omega^T x_i + b|}{||\omega||} \tag{4.5}$$

The support vectors are the the data points x_i closest to the hyperplane. The distance between the data points and the hyperplane is called margin, must reach a maximum. The distance can be calculated by the following formula

$$\zeta(H) = min_{x_i} d(H; x_i) = \frac{1}{||\omega||}[min_{x_i} |\omega^T x_i + b|] = \frac{1}{||\omega||} \tag{4.6}$$

The maximum ζ is reached when $||\omega||$ is minimized. The optimization problem below must be solved

$$\Theta(\omega) = argmin_{\omega, b}[\frac{1}{2}||\omega||^2] \tag{4.7}$$

The following constraint must be assumed to not violate the condition 4.4

$$y_i((\omega^T x_i) + b) \geq 1, i = 1, ...n \tag{4.8}$$

The optimization problem with constraints can be solved using the Lagrange's method of multipliers

$$\overline{\omega} = \sum_{i=1}^{n} \overline{\alpha}_i y_i x_i \tag{4.9}$$

for $\overline{\omega}$, with $\overline{\alpha}_i$ being the Lagrange multipliers

$$\bar{b} = -\frac{1}{2}\overline{\omega}[x_r + x_s] \tag{4.10}$$

for $\bar{b}$, where the indiced r and s indicate the support vectors

$$\overline{\alpha}_r, \overline{\alpha_s} > 0, y_r = 1, y_s = -1 \tag{4.11}$$

finally, the optimal hyperplane is

$$f(x) = sign(\overline{\omega}^T x + b) \tag{4.12}$$

4.1.3.2 Classification Tree

The classification tree is a very simple and efficient tool that can be used in data classification. The classification tree can be seen as a binary tree. Each root node of the tree represents one input and the leaf nodes represents the output of the tree. The C4.5 algorithm [17] is one solution to implement a classification tree.

- Check for base cases
    1. For each attribute a
        a) Find the feature that best divides the training data such as information gain from splitting on a
    2. Let a best be the attribute with the highest normalized information gain
        a) Create a decision node node that splits on a_best
- Recurse on the sub-lists obtained by splitting on a best and add those nodes as children of node

4.1.3.3 Cross-Correlation

The cross-correlation function on signal processing tells the similarity of two given signals. The function is defined by the following expression

$$C_{x_1 x_2}(\sigma) = \lim_{T_f \to \infty} \frac{1}{T_f} \int_{-Tf/2}^{T_f/2} x_1(t) x_2(t+\sigma) dt \tag{4.13}$$

The method proposed by [5] for fault detection using cross-correlation is to calculate the auto-correlation between the IMF's obtained with the EMD method, discussed on section

3.4.1.4, in the feature extraction part. The auto-correlation is the cross-correlation function applied to the same signal divided into two parts. Having the results of the auto-correlation, the magnitude and the position where the correlation was more important is saved. Those values gives information about the symmetry of the signal. If the highest magnitude of the correlation is when the position is close to zero, it means that the IMF is highly symmetric. The method proposes to use those symmetric information to define thresholds to classify the data into two states.

In our case, instead of define thresholds to classify the data, we propose to use the machine learning methods, SVM or classification tree. So, as training data, instead of the features extracted by the data processing part, the machine learning methods used the information about similarity of the IMFs obtained with the cross-correlation function.

4.2 Implementing the signal processing method for fault detection on loopers

The signal processing method for fault detection on loopers was implemented using the recorded data from 2016 to 2018 from Florange factory. The hot strip mill plant has an iba® measurement system. This system acquires hundreds of signals during the whole milling process. The signal-based fault detection model was made to identify fault on loopers. In the looper subsystem, 18 signals are recorded by the iba® measurement system. Some of the most important signals recorded are:

- Angular position measure/reference
- Speed measure/reference
- Torque measure/reference
- Servomotor reference
- Pression on hydraulic cylinder

There are 2297 files recorded in total. It's known that 1827 files represent the nominal behaviour of the system. The other 470 files were recorded when the system was on its faulty state. It's not defined the exact fault presented in the system, or how many faults can be found on the faulty files. It is just known that those files don't represents the nominal behaviour of the system.

4.3 Data selection

To apply the fault-detection techniques, it needs to be defined the signal that will be used. In coordination with specialists and analysing separately each of the 18 signals, the servomotor reference was chosen. The servomotor is part of the looper's control system. When the strip must be tensioned, the angular reference of the looper increases. To move up the looper, the servomotor opens one valve that increases the pressure on the hydraulic cylinder, which will move up the looper. When the looper reaches it desired position, the servomotor closes the valve, so the pressure on the hydraulic cylinder is maintained constant and the looper stays fixed in the same position. When this behaviour is not observed, it's considered that a system fault has occurred. We can compare both fault and nominal servomotor reference signals below.

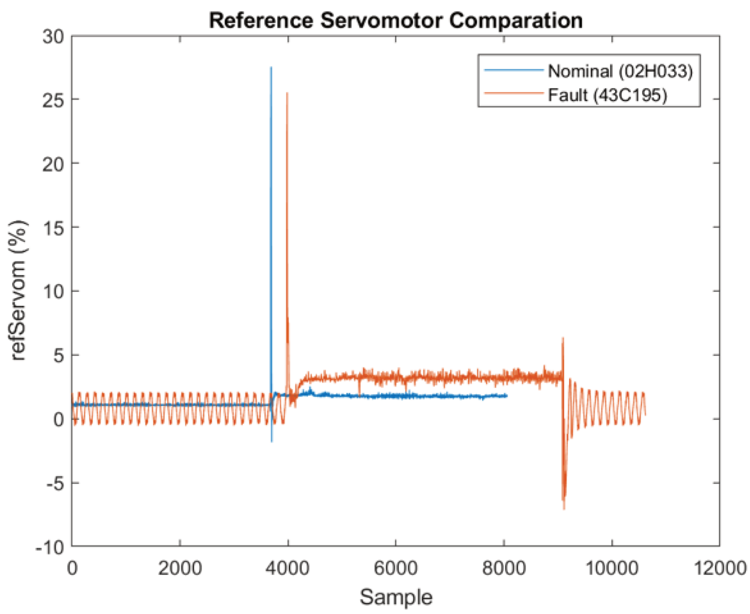


Figure 17 – Nominal and fault signals (Source: personal archive)

On this first example, we can see a clear difference between the nominal and the fault signal. The nominal signal doesn't come back to the same point as before, considering the ideal behaviour, but it keeps a constant value not far from it, which is acceptable considering that this is real measurements recorded from the factory. In Figure 18, it's possible to compare another fault and nominal signals extracted from the iba® files. In this case, the difference between nominal and fault signals are not as evident as before, which makes more difficult the state identification.

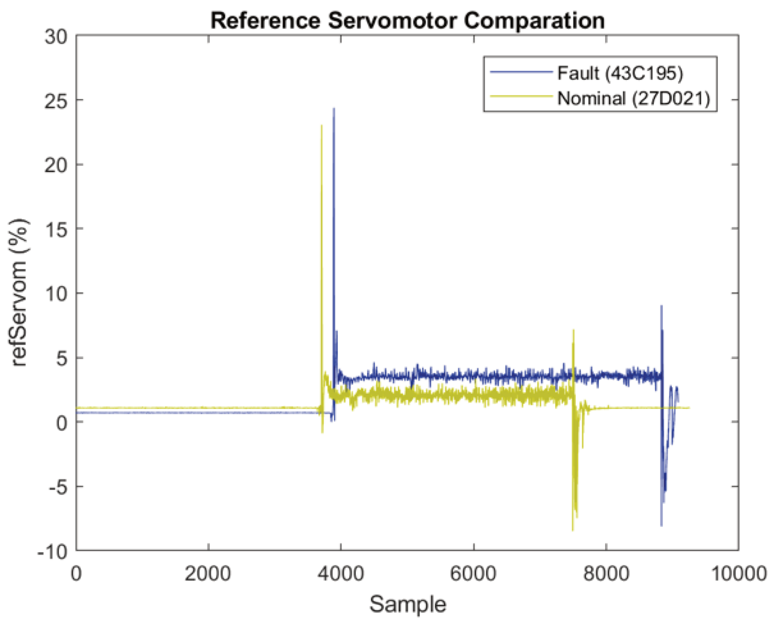


Figure 18 – Nominal and fault signals, another example (Source: personal archive)

4.3.1 Pre-processing

From the 2297 MATLAB® files, 60% were used for modeling and 40% for validation. As the data were separated by year (2016, 2017 and 2018), the chosen files were selected randomly, with the intention to minimize the influence of the period of the year that it was recorded. For each file, it was selected just the period when the angular reference for the looper changes. This is the part that contains the most information about the faults.

4.4 Data processing

After a visual analysis of the signals from the servomotor, we implemented the methods proposed in previous section of this chapter for feature extraction and classification: descriptive statistics, STFT and EMD. As we don't have further information about faults, it's considered only 2 states systems: nominal and faulty.

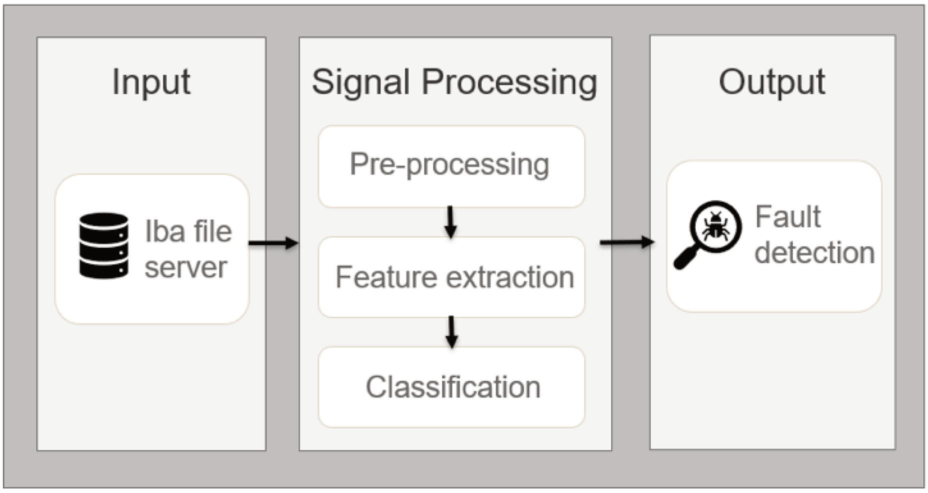


Figure 19 – Fault detection system architecture (Source: personal archive)

4.4.1 Descriptive statistics

In this case, from each signal, the mean, median, minimum value, maximum value, standard deviation and the variation were calculated using MATLAB®. Those values were separated into two groups, fault and nominal, the figures below show the relation between 4 of those 6 variables for each state.

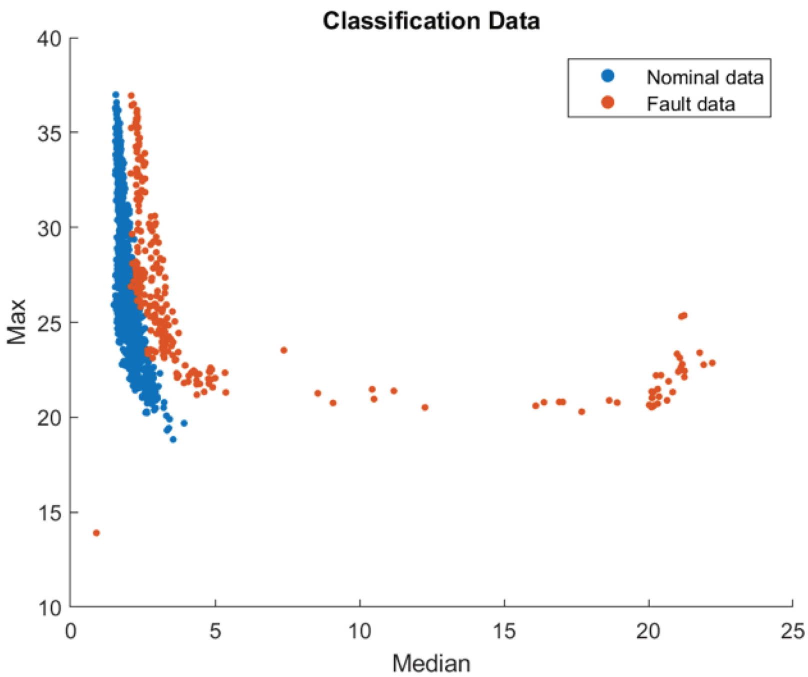


Figure 20 – Relation between the maximum value and the median of the modeling signals (Source: personal archive)

Each point of this graph represents one MATLAB® file, it's possible to see a clear difference between the fault and nominal data for the relation between the median and the max value of each file. It means that those features are useful for the classification purposes. From Figure 21 it's possible to see that the relation between the mean and the standard deviation doesn't give enough information to classify the data between fault or nominal.

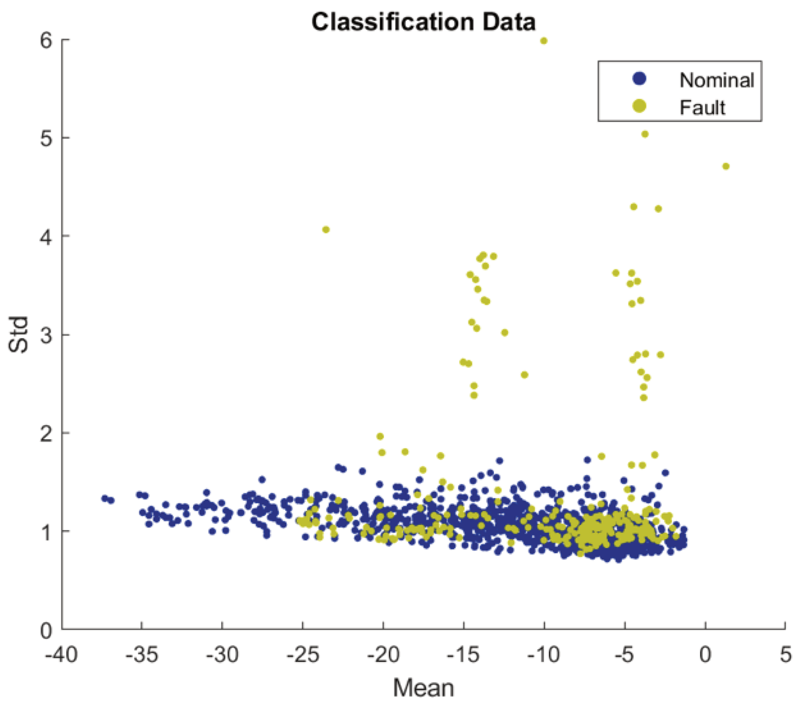
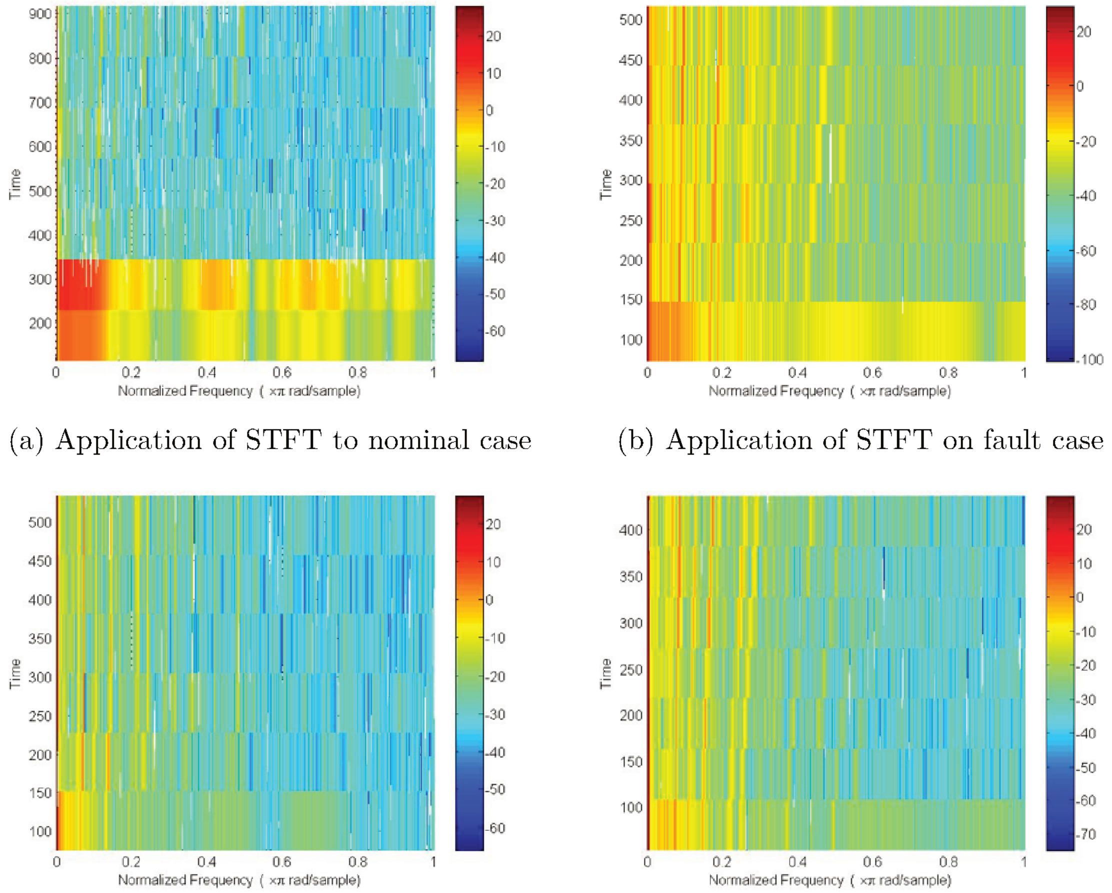


Figure 21 – Relation between the standard deviation and the mean value of the signals (Source: personal archive)

4.4.2 Short-time Fourier Transform (STFT)

In the figures below, we can compare the result of the application of STFT on four different files. In Fig. 22a for a time between 0 and 350, the amplitude of the frequencies is higher than in other periods of the time. It's possibly caused by the change of the angle's reference shown in Fig. 17 and Fig. 18. In the fault case (Fig.22b and Fig.22d), both spectrograms don't show a big difference on the amplitude in the time-frequency domain which is reasonable considering that the fault signals are much noisier. Fig.22c shows a spectrogram of a signal when the system is on its nominal state, in this case, the difference between Fig.22d is not very clear, this flue difference can cause trouble in the fault identification system.



(a) Application of STFT to nominal case

(b) Application of STFT on fault case

(c) Application of STFT to nominal case

(d) Application of STFT on fault case

Figure 22 – STFT results (Source: personal archive)

4.4.3 Empirical Mode Decomposition (EMD)

The IMF in the first line of the graphic shows the detailed changes in the original time signal. Each following line gives the remaining time-dependent behaviour of the signal until the last line shows the residual [2]. In this case, as it will be applied the cross-correlation on each signal, in both fault and fault-free cases the symmetry is not evident. The last results (EMD and STFT) indicates that a time-frequency analysis of the signal may not give satisfactory results.

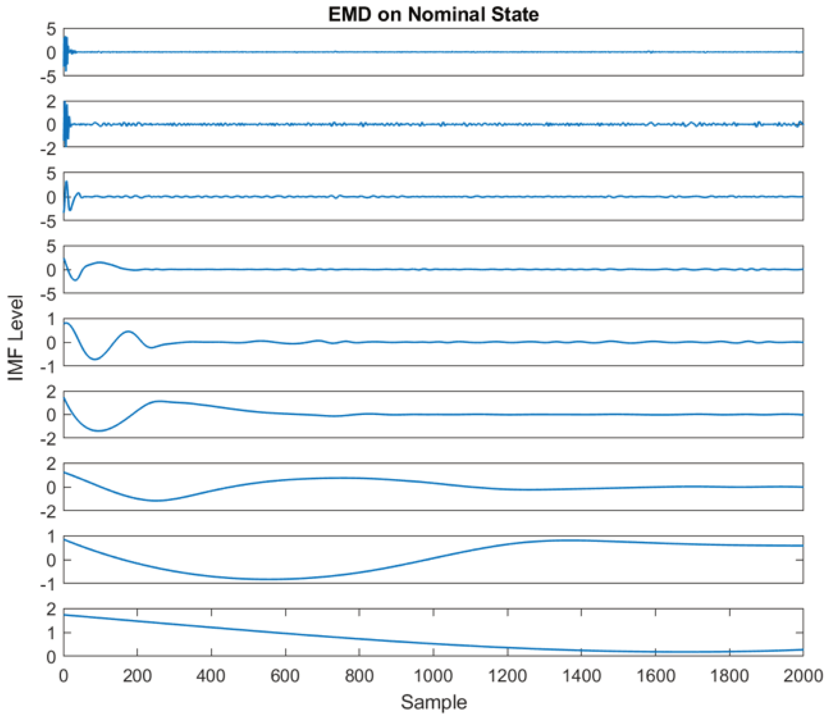


Figure 23 – Application of EMD to nominal case (Source: personal archive)

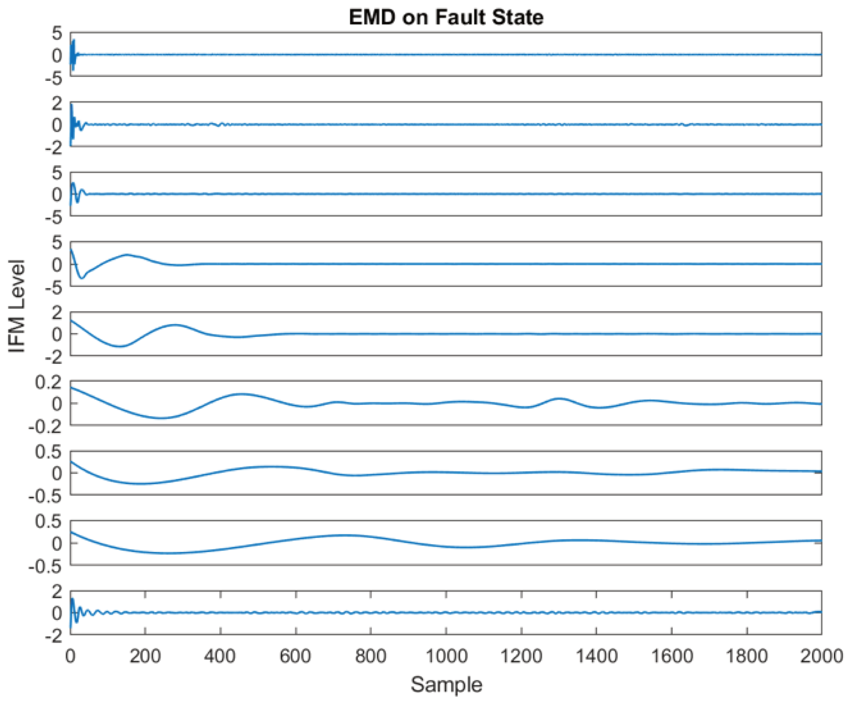


Figure 24 – Application of EMD to fault case (Source: personal archive)

4.5 Data Classification

After the feature extractions, the next step is the classification. To implement different classification methods, it was used the MATLAB® Classification Learner app. The basic statistical features were classified using SVM and classification tree. The STFT results were sorted using classification three. The EMD results were first classified using cross-correlation and then with SVM.

4.5.1 Classification Tree

The classification tree was applied to the basic statistics features and to the STFT data. For the basic statistics features, each feature (mean, median, maximum value, minimum value, standard deviation and variation) is an input to the classification tree, as result, it indicates the state of the system. The STFT data were also used as input in the classification tree. The results are presented in the next section.

4.5.2 Support vector machine

The SVM was used to classify the features extracted by the data processing methods into the two system states. The classification tree has the same input as the SVM. So our objective here is to compare which classification method is more efficient to identify faults in the signal that is being used.

4.5.3 Cross Correlation

The features extracted using the EMD method were auto-correlated. From each result of the autocorrelation it was extracted the maximum point of correlation. If the signal is symmetric for example, the highest autocorrelation would be in 0. That information was used as input for the support vector machine training.

4.6 Implementation results

In total, we had 2297 files recorded. Around 80% represented the system on its nominal behaviour and the other 20% the system on its fault state. For modeling, 60% of each type of file were used. So the other 40% were reserved to do the validation. The files were chosen randomly, to minimize the influence of the period when the file was recorded.

To validate our fault detection models we build a 'receiver operating characteristic' (ROC) space using only the validation files. The ROC space is defined by calculating the true positive rate (TPR) and false positive rate (FPR) of each system and defining the FPR as the x axis and the TPR as the y axis.

The TPR indicates how sensible the model is to identify faults. One system that identifies all the faults would have a TPR of 100%. The FPR can be interpreted as probability of false alarm. The false alarm is when the signal is on its normal behaviour, but the fault identification system identifies a fault. The perfect system would have 100% of TPR and 0% of FPR. Analyzing the ROC space of each fault detection model that was developed will indicate how precise and accurate our system is. The formula to calculate the TPR and FPR are presented below.

$$TPR = \frac{TP}{P}$$

$$FPR = \frac{FP}{N}$$

Where TP is the number of positive cases correctly guessed by the fault detection system, and P is the total number of positive cases in the data. FP is the number of negative incorrectly guessed by the system and N is the total number of negative cases in the data.

4.7 Experimental results and validation

The ROC space is shown in Fig.25 and the numeric values are presented in table 1.

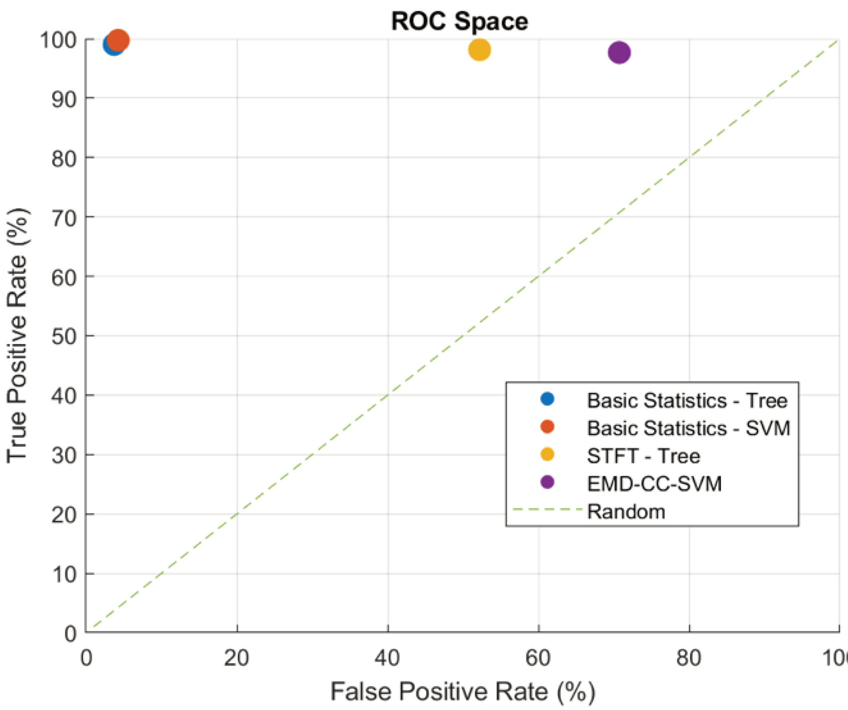


Figure 25 – Model validation results (Source: personal archive)

Table 1 – Numeric values of model validation

Model	False positive rate	True positive rate
Descriptive statistics - Tree	3.7%	99.0%
Descriptive statistics - SVM	4.2%	99.7%
STFT - Tree	52.2%	98.2%
EMD-CC-SVM	70.7%	97.8%

It's clear that for fault detection, the best results are the combination between basic statistics and a given classification method. The classification tree has a singly better fault positive rate, and a worst true positive rate comparing to the SVM. Considering that this is a fault detection system, it's better to have a false alarm instead of doesn't identify one fault. So, the basic statistics features combined with SVM classifier is the best choice.

The other two model has shown worst results. It shows that for fault detection on loopers, time-frequency analysis aren't good methods for features extraction. However, those methods should be considered when it comes to fault identification, but to perform a fault identification it is required more information about the fault.

Another reason that should be considered to understand why the more complex methods didn't give good results is that is not known how many type of faults were present on the files. Both features extraction systems, STFT and EMD, had a big false alarm rate. It means that with the system is too sensible to identify faults.

So, the best signal processing method for fault identification on loopers are the descriptive statistics combined with support vector machine or classification tree.

Table 2 – The most efficient techniques for fault detection on loopers

Model	False positive rate	True positive rate
Descriptive statistics - Tree	3.7%	99.0%
Descriptive statistics - SVM	4.2%	99.7%

5 Model-based fault detection on strip steering

After speaks with specialists from the factories. It was noticed that the loopers aren't a big problem on the hot strip mill plant. Those faults are not the most frequent and it's not a reason for a big concern from the operators.

The most problematic part of the hot rolling process is the strip tracking. Usually the strip gets out of its desired path, this deviation can cause serious faults like the 'cobble' which also very dangerous for the operators.

This chapter proposes a model-based fault identification system on strip steering.

5.1 The proposed model-based method for fault detection

Model-based methods are much more complex to develop than signal based ones. It requires an accurate model of the system and the fault detection, isolation and identification is much more complex. In other hand, those models are usually better on fault identification and much more flexible to changes on the plant.

As seen in previous sections, there's many ways to build a model-based fault detection system. The proposed solution in presented below.

1. Build a dynamic model
    - Validate the model
2. Study the faults on the system and build model for the faults
3. Linearize the system
4. Apply the methods presented on section 3.3.2
5. Validate the fault identification system with real data

5.2 Implementing the model-based method for fault detection on strip steering

The first step to implement the proposed solution is to build a dynamic model for the system. When the strip steering control, see 2.0.1.5, was implemented, one dynamic model was developed. The idea was to use the same model for fault detection on the strip path.

5.2.1 Mathematical equations

Those equations were developed by ArcelorMittal's engineers when the strip steering control was developed [18]. The non-linear equations shows the relation between the parameters presented on table 3.

Table 3 – Strip steering parameters

Parameter	Description
ΔP_j	Differential rolling force
Δh_j	Strip thickness profile
ΔS_j	Stand tilting
ΔK_j	Differential stand stretch
ΔT_j^{am}	Upstream differential of strip tension
ΔT_j^{av}	Downstream differential of strip tension
Z_j	Strip off-centre
E_j	Young's module
s_j	Work roll speed
b_j	Work roll length
l_j^0	Inter-stand length
l_j^v	The screw inter-axis length
T_j^{av}	The front strip tension
T_v^{am}	The back strip tension
h_j	Strip thickness
w_j	Strip width
$c_j^{fh}, c_j^{fT_{am}}, c_j^{fT_{av}} c_j^{gh}$	Constants to represent the gradient of the strip parameters
$c_j^{gT_{am}}, c_j^{gT_{av}}, K_j^h, K_j^f, K_j^l, P_j, g_j$	Constants to represent the gradient of the strip parameters

The main equations governing the system are presented below.

The differential rolling force equation

$$\Delta P_j = c_{j-1}^{fh} \Delta h_{j-1} + c_j^{Tam} \Delta T_j^{am} + c_j^{fTav} \Delta T_j^{av} \tag{5.1}$$

The exit stand wedge equation

$$\Delta h_j = \left(\frac{w_j}{(l^v)_j^2 K_j^h} + \frac{6 w_j}{b_j^2 K_j^f} \right) (\Delta P_j - 2 P_j) Z_j + \frac{\Delta P_j}{K_j^l} + \frac{w_j}{l_j^v} \Delta S_j - \frac{w_j}{l_j^v (K_j^h)^2} P_j \tag{5.2}$$

The angle between the strip and the mill axis

$$\dot{\alpha}_j = \frac{s_j}{w_j}\left(\frac{c_j^{gh}}{1+g_j}+\frac{1}{h_j}\right)\Delta h_j + \frac{s_j}{w_j}\left(\frac{c_{j-1}^{gh}}{1+g_j}-\frac{1}{h_{j-1}}\right)\Delta h_{j-1} + \frac{s_j c_j^g Tav}{w_j(1+gj)}\Delta T_j^{av} + \frac{s_j c_j^{gTam}}{w_j(1+g_j)}\Delta T_j^a m \tag{5.3}$$

The upstream differential of strip tension equation

$$\Delta T_j^{am} = 3\left(\frac{w_j E_j}{(l_f^0)^2}+\frac{T_j^{am}}{w_j}\right)(Z_j - Z_{j-1}) + \frac{w_j E_j}{l_j^0}(2\alpha - \alpha_{j-1}) + 3\frac{l_j^0 T_j^{am}}{w_j}\alpha_j \tag{5.4}$$

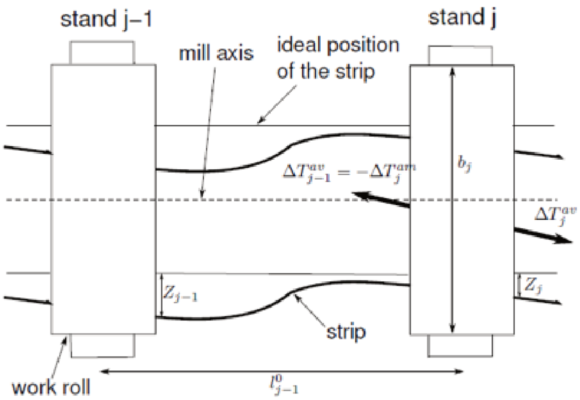
The strip off-centre

$$\dot{Z}_j = s_j \alpha_j \tag{5.5}$$

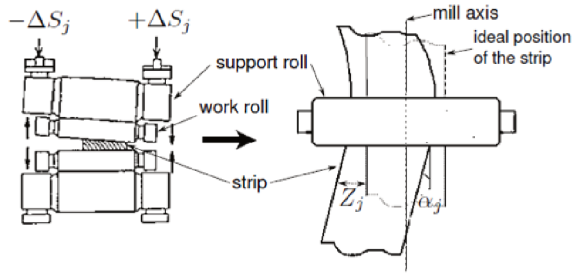
The coupling between two successive stands equation

$$\Delta T_{j-1}^{av} = -\Delta T_j^{am} \tag{5.6}$$

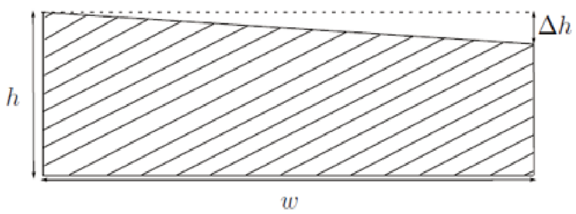
A graphical representation of the variables is presented below.



(a) Strip behaviour between two stands



(b) Front and top view of strip rolling process



(c) Strip profile

Figure 26 – Strip steering model (From [4])

5.2.2 Model implementation

The equations were rewritten to make the system in function of two inputs, stand stretch (ΔK) and stand tilting (ΔS), and its previous states. The equations were implemented on Simulink® and the simplified architecture with only two stands is shown in Fig. 27.

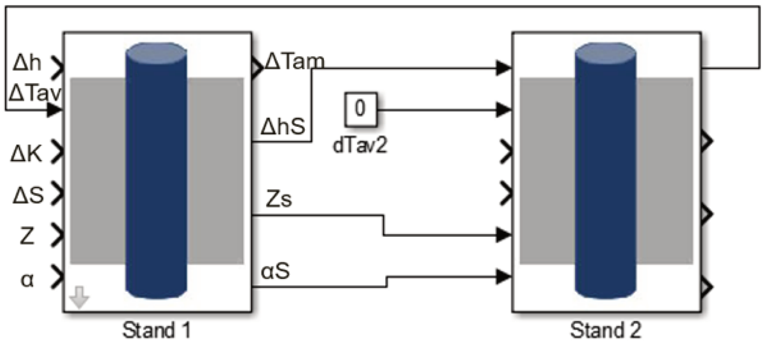


Figure 27 – Simplified Simulink model (Source: personal archive)

5.2.2.1 Model Validation

To continuous this project, the next step is to validate the model. To validate it, it's required real data recorded from the factory. Those data are not available yet, so this project is stopped till the data acquisition.

6 Summary and future work

6.1 Summary

The project presents two different methods for fault detection on hot strip mill plants: signal processing and model-based. The signal processing method was used for fault detection on looper by analysing the servo-motor signal. The signals were firstly analysed through short time Fourier transform, empirical mode decomposition and basic descriptive statistics. Then the data was classified using machine-learning methods. The method that combined descriptive statistics and support vector machine gave great results, with more than 99% of true positive rate and less than 5% of false positive rate for fault detection. The model-based method proposed for fault detection on strip steering system couldn't be finished due to lack of real data to validate the model of the process.

Most part of the project objectives were accomplished. An extensive bibliography revision were made to study the techniques of fault detection in the steel industry. Two different methods for fault detection were studied and one was successfully implemented. A toolbox for facilitating further usage wasn't developed because the fault detection on loopers turned out to be an infrequent and not relevant fault. So it was made a decision to don't spend more resources on this subject.

6.2 Future work

The signal processing method for fault detection proved its efficiency. The same method could be used for fault detection on other parts of the hot strip mill system. The recorded data available was separated into two groups: fault and nominal, those circumstances made impossible the development of a complete fault diagnostic system. Organizing the fault data into subgroups, each one with one different fault, would allow the implementation of a fault isolation and identification system.

The model-based method, the next step is to validate the existing model with real data from industry. Having a valid model of the strip steering system, it's possible to do a study of the failures on the strip steering system and model it. With both models it's possible to build model-based fault diagnosis system with the techniques studied in this project.

Bibliography

1 M., L. C. D. Design of fault diagnosis observer for hagc system on strip rolling mill. *Journal of Iron and Steel Research*, v. 13, p. 27 – 31, 2006. Citado 2 vezes nas páginas 11 and 22.

2 A., R. *Approach for Improved Signal-Based Fault Diagnosis of Hot Rolling Mills*. Tese (PhD Thesis) — Universität Duisburg-Essen, 2016. Citado 5 vezes nas páginas 11, 27, 28, 42, and 54.

3 LOIC, B. *Hot Rolling Friction Control Through Lubrification*. Tese (PhD Thesis) — Université de Lorraine, 2017. Citado 3 vezes nas páginas 11, 27, and 28.

4 B., C.-W. *Strip Tracking in Hot Strip Mills*. Tese (Marter's thesis) — Cardiff University, 2011. Citado 4 vezes nas páginas 11, 28, 30, and 61.

5 DING, S. X. *Model-based Fault Diagnosis Techniques*. United States: Springer, 2008. Citado 7 vezes nas páginas 11, 35, 36, 37, 40, 46, and 48.

6 LENARD, J. G. *Primer on Flat Rolling - History of Hot Strip Mills*. Netherlands: Elsevier, 2014. Citado na página 22.

7 J., M. Maintenance and quality related monitoring of rolling mill main drives. AISE Annual Convention, p. 1 – 14, 2002. Citado na página 27.

8 S., Y. K. *Finishing Mill Modelling and Looper Control*. Tese (Marter's thesis) — University of Alberta, 2005. Citado na página 28.

9 G., B. J. Lubrification of hot-strip-mill rolls. *Wear*, v. 23, p. 203 – 208, 1973. Citado na página 29.

10 B., F. A. *Adaptive Fuzzy Logic Steering Controller for a Steckel Mill*. Tese (Marter's thesis) — University of Johannesburg, 2005. Citado na página 29.

11 R., I. Model-based fault detection and diagnosis - status and applications-. *Control Engineering Practice*, v. 5, p. 709 – 719, 1997. Citado 2 vezes nas páginas 34 and 37.

12 WOODBURY, G. *An Introduction to Statistics*. United States: Cengage Learning, 2001. Citado na página 41.

13 STOICA, P.; MOSES, R. *SPECTRAL ANALYSIS OF SIGNALS*. United States: PRENTICE HALL, 2005. Citado na página 41.

14 HUANG, S. S. N. Hilbert–huang transform and its applications. *World Scientific*, 2005. Citado na página 42.

15 PENG P. W. TSE, F. C. Z. A comparison study of improved hilbert–huang transform and wavelet transform: Application to fault diagnosis for rolling bearing. *Mechanical Systems and Signal Processing*, 2005. Citado na página 42.

16 ABE, S. *Support Vector Machines for Pattern Classification*. United Kingdom: Springer Verlag London, 2010. Citado na página 47.

17 KOTSIANTIS, S. Supervised machine learning: A review of classification techniques. *Informatica*, 2007. Citado na página 48.

18 MALLOCI J. DAAFOUZ, C. I. R. B. P. S. I. Robust steering control of hot strip mill. *10th European Control Conference,*, 2009. Citado na página 60.