MYRIAM COUSINEAU-PELLETIER

COMBUSTION MONITORING FOR BIOMASS BOILERS USING MULTIVARIATE IMAGE ANALYSIS

Mémoire présenté
à la Faculté des études supérieures de l'Université Laval
dans le cadre du programme de maîtrise en Génie chimique
pour l'obtention du grade de Maître ès sciences (M. Sc.)

DÉPARTEMENT DE GÉNIE CHIMIQUE
FACULTÉ DES SCIENCES ET DE GÉNIE
UNIVERSITÉ LAVAL
QUÉBEC

2009

RÉSUMÉ

Les procédés de combustion sont utilisés dans la plupart des industries chimiques, métallurgiques et manufacturières, pour produire de la vapeur (chaudières), pour sécher des solides ou les transformer dans des fours rotatifs (ou autres). Or, les combustibles fossiles qui les alimentent (ex. : gaz naturel) sont de plus en plus dispendieux, ce qui incite plusieurs compagnies à utiliser d'autres sources de combustibles tels que de la biomasse, des rejets inflammables produits par le procédé lui-même ou des combustibles fossiles de moindre qualité. Ces alternatives sont moins coûteuses, mais de composition, et donc de pouvoir calorifique, plus variable. De telles variations dans la chaleur dégagée par la combustion perturbent l'opération des procédés et la qualité des produits qui dépendent de ces installations. De nouvelles stratégies de contrôle de la combustion doivent donc être élaborées afin de tenir compte de cette nouvelle réalité. Il a été récemment démontré que l'énergie dégagée par la combustion est corrélée à l'aspect visuel de la flamme, principalement sa couleur, ce qui permet d'en quantifier les variations par imagerie numérique. L'objectif de ce projet industriel consiste à faire la démonstration que l'analyse d'images multivariées peut servir à l'identification du comportement d'une chaudière à biomasse. La chaudière à biomasse opérée par Irving Pulp & Paper Ltd (Saint-John, Nouveau-Brunswick) fera office d'exemple. Les résultats montrent qu'un modèle bâtit à partir des informations fournies par les images ainsi que les données de procédé donne de bonnes prédictions de la quantité de vapeur produite ($R^2_{modèle}$=93.6%, $R^2_{validation}$=70.1%) et ce, 2,5 minutes à l'avance. Ce projet est la première étape du développement d'une nouvelle stratégie de contrôle automatique de la combustion de biomasse, capable de stabiliser l'énergie dégagée, malgré les variations imprévisibles dans le pouvoir calorifique et les proportions des combustibles utilisés provenant de différentes sources.

ABSTRACT

Biomass is increasingly used in the process industry, particularly in utility boilers, as a low cost source of renewable, carbon neutral energy. It is, however, a solid fuel with some degree of moisture which feed rate and heat of combustion is often highly variable and difficult to control. Indeed, the variable bark properties such as its carbon content or its moisture content have an influence on heat released. Moreover, the uncertain and unsteady bark flow rate increases the level of difficulty for predicting heat released. The traditional 3-element boiler control strategy normally used needs to be improved to make sure the resulting heat released remains as steady as possible, thus leading to a more widespread use biomass as a combustible. It has been shown in the past that the flame digital images can be used to estimate the heat released by combustion processes. Therefore, this work investigates the use of Multivariate Image Analysis (MIA) of biomass combustion images for early detection of combustion disturbances. Applied to a bark boiler operated by Irving Pulp & Paper Ltd, it was shown to provide good predictions, 2.5 minutes in advance, of variations in steam flow rate (R^2_{fit}=93.6%, R^2_{val}=70.1%) when information extracted from images were combined with relevant process data. This project is the first step in the development of a new automatic control scheme for biomass boilers, which would have the ability to take proactive control actions before such disturbances in the manipulated variable (i.e. bark flow and bark properties) could affect steam production and steam header pressure.

AVANT-PROPOS

La vie nous mène parfois où l'on ne croirait jamais aller. Sans mon directeur de recherche, Carl Duchesne, à qui je dois cette idée, je n'aurais probablement pas pensé faire des études graduées. Je l'en remercie donc grandement puisque ce fut une expérience de vie enrichissante d'où je suis sortie grandie et meilleure. Non seulement m'a-t-il proposé un projet des plus intéressants dans un domaine d'actualité dont l'avenir est prometteur, mais il a aussi contribué à sa réalisation par sa grande disponibilité et ses nombreuses idées. J'ai eu la chance d'avoir un directeur aidant et compréhensif et j'en suis très contente.

Si mon directeur de recherche a vu en moi le potentiel de faire des études graduées, c'est surtout grâce à mes parents qui ont su me stimuler intellectuellement tout au long de ma vie en me faisant connaître différents champs d'activité pour que j'acquière les qualités qui m'ont permis de réaliser les projets de ma vie. Merci pour votre soutien et votre amour.

Mes remerciements seraient incomplets sans mentionner les gens qui font partie de ma vie. Merci donc à mes amies, Sarah et Julie, qui m'écoutent et me divertissent. Merci à la complice de ma vie, ma sœur, Paule qui me comprend sans même que j'explique. Finalement, merci à mon amour, Pascal, qui m'aime et me supporte quoiqu'il arrive. Vous rendez la vie tellement plus intéressante!

Je me dois absolument de souligner la contribution de mes collègues, Ryan, Jayson, Pierre-Marc et Julien, qui ont toujours été à mes côtés. Durant les deux dernières années, ils ont dû subir mes nombreuses péripéties et répondre à mes questionnements. Merci Ryan qui récemment s'est transformé en dictionnaire anglais-français!

Ce projet a été un projet d'équipe et je veux remercier tous les gens qui, tout comme Carl, m'ont aidé et qui ont participé à son accomplissement. Premièrement, Don McCabe et toute l'équipe de chez Irving Pulp and Paper qui nous ont accueillis. Votre disponibilité, votre collaboration et votre chaleureux accueil ont été très appréciés. Marlene, ta compagnie a été des plus agréables. J'ai découvert une fille extraordinaire et une bonne amie. Merci pour ton aide.

Merci

À la meilleure de toutes les sœurs et à
notre Teddy qui a su illuminer nos vies

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

1.1 ENERGY PRODUCTION IN THE PROCESS INDUSTRY

Combustion is used in most process industries, either as part of their main process or as a utility (i.e. steam generation). Indeed, the metallurgical, the food, as well as the chemical industry commonly use utility boilers, rotary kilns, and various types of furnaces. The food industry needs combustion processes to dry a wide variety of products, such as fish or soy meal. The pulp and paper industry, for example, uses kilns to perform a chemical reaction called calcining which requires heat to regenerate spent lime. Metallurgical industries also use kilns to carry out physical and chemical reactions within the ore to modify the material properties (i.e. drying, desulphurizing, phase change, etc). Most industries also use utility boilers for producing steam, mainly for supplying heat to the process. In chemical pulping large amounts of steam is required in the cooking stage where high temperatures are needed for delignification. Steam production can also be part of a power cogeneration process where a turbine is fed with steam and electricity is produced for equipment.

Combustibles mainly used in combustion processes are fossil fuels such as petroleum products, coal or natural gas. The cost of those products is steadily increasing over the years and, in turn, yields significant fluctuations in production costs. Indeed, the price of natural gas, for example, tripled between 1995 and 2005 (Giroux, 2008). Fossil fuels are non-renewable resources and the sustainability of these raw materials is more and more compromised. Increasing demand in conjunction with depleting fossil fuels reserves are among the main causes for increasing costs. Another argument against fossil fuels is related with their environmental impact. Indeed, extracting and burning fossil fuels are major sources of greenhouse gas (i.e. CO_2) which raises environmental concerns. As it is shown in Table 1, the amount of pollution, in mg/MJ of energy produced, is considerably higher when a boiler is fed with fossil fuels compared to bark especially with the CO_2 emissions. Indeed, according to the Institute of Bioenergy, since the amount of CO_2

released by the combustion of bark is equivalent to the one absorbed by the respiration process, the CO_2 balance can be considered nil.

Table 1 : Net emissions created by the combustion of various fuels (mg/MJ of heat created)

	SO_2	NOx	CO	CO_2	Dust
Oil boiler	140	40	50	78 000	5
Natural gas boiler	0	40	50	52 000	0
Coal boiler	340	70	4 500	104 000	60
Bark boiler (teared wood)	10	45	16	0	4

Therefore, the rising costs of fossil fuels as well as associated sustainability and environmental issues are compelling arguments for process industries to replace fossil fuels by alternative low-cost, renewable and carbon neutral fuels as their energy supply in their combustion processes.

Combustible process by-products and/or wastes, either gaseous or liquid, have traditionally been used to reduce fossil fuel consumption and increase energy efficiency. Nowadays, other kinds of combustibles are used. Biomass (e.g. sludge, bark, plant, vegetable or animal derived materials) is indeed increasingly used in the process industry as an alternative, renewable and in most cases carbon neutral fuel. Indeed, according to the Forest Product Association of Canada, biomass and small hydroelectric station are representing 60% of the energy source in the pulp and paper industry. Irving Pulp and Paper, located in St-John (New Brunswick, Canada), is one of those industries using biomass (i.e. bark) to reduce operational cost and environmental impact. A bark boiler is used to produce the steam they need in their pulp making process. However, not only process industries make use of biofuels to decrease their energy costs. Many public institutions such as schools, hospitals or even residential complexes can also be heated

using biomass boilers. Amqui's hospital located in Gaspésie plans to save over 100,000$/year using their new biomass boiler as their main heating system (Gagné, 2009).

1.2 PROBLEMS WITH THE COMBUSTION OF BIOMASS

Biomass boilers are very promising alternatives to traditional fossil fuel boilers, but cost reduction and improved environmental performance is made at the expense of using a much less stable fuel. Indeed, biomass composition (i.e. mainly carbon content), moisture content, and flow properties are typically highly variable which leads to an unsteady heat of combustion. Carbon content mainly depends on the biomass sources and carbon loss during storage (e.g. fermentation within bark piles reduces the amount of carbon available for combustion). Moreover, it is difficult to achieve a steady biomass feed flow to the boiler since biomass consists of wet solid particles which degree of humidity depend on biomass source and weather conditions (when stored outside). When transported using screw feeders, the wet solids have a tendency to stick to the screw shaft which modifies the effective volume of materials discharged at each screw revolution (Bell et al. 2003). Achieving consistent screw fullness is also difficult, creating further flow disturbances (i.e. flow tends to be lumpy). Furthermore, obtaining flow measurements of such materials is not straightforward and these are rather imprecise. The resulting variations in heat released by biomass combined with their unsteady flow properties are two factors that bring important pressure fluctuations within the biomass boiler steam drum, hence in steam production rate. Variations in steam production can lead to serious operation problem with equipments using steam for heating such as, in the chemical pulping industry, the pulp digesters which require a certain steam pressure (or temperature) in order to achieve the desired reaction temperature. Another example is when the steam from biomass boilers is also used in a cogeneration plant, as is the case with Irving Pulp & Paper. When it is anticipated that weather conditions could result in a power shortage, then steam production is increased in order to produce more electricity and even become independent of the local network. Since bark is too variable, the operation is often switched to fossil fuels until

weather conditions have improved. This results in higher operating costs, especially when such a situation happens regularly.

To compensate for these variations, operators can either cut biomass feed rate when heat released is too high or burn fossil fuels through auxiliary fossil fuel burners when biomass combustion is not sufficient to meet steam demand. These control decisions are, however, generally taken after combustion disturbances have already affected steam production (i.e. feedback control). In certain critical situations when steam production is required to be steady, the use of biomass alone may be too risky. In these cases, when the boiler is equipped with auxiliary fossil fuel burners, steam is mainly produced through fossil fuel combustion until the process returns to normal operation.

Improved equipment design and instrumentation may lead to some reduction in the variability of biomass flow properties, but carbon content would still remain untraceable. Indeed, a variety of agitation devices are commercially available to ensure that solids entering the screw feeders are in a highly flowable state, leading to a more uniform distribution of the product and a higher and more consistent degree of filling of the screw (Bell et al. 2003). On-line moisture sensors mainly based of near-infrared (NIR) probes or arrays have also been developed for estimating solids moisture content on conveyor belts. The system developed and commercialised by FPInnovation Paprican (www.paprican.ca) for wood chips is one example. Feedforward control actions could be taken to reduce potential upsets to the boiler before wet biomass is fed into the combustion chamber. Nevertheless, carbon content is typically measured using bomb calorimetry, which is mainly an off-line test performed in laboratory, and requiring some time to be performed (i.e. cannot be easily adapted for on-line use). Therefore, even with implementation of the improvements discussed above, biomass flow properties would not be perfectly steady, and variations in carbon content will still cause important disturbances in steam production.

1.3 OBJECTIVES OF THIS THESIS

The long term objective of this research program is to develop a new combustion control strategy in order to stabilize the heat released by biomass combustion, hence leading to a more widespread use of such an alternative to fossil fuels. The 3-elements control strategy typically used in boiler control (Smith and Corripio, 2005) works well with fossil fuels which have much more stable combustion energy and are delivered steadily to the boilers (i.e. flow rates of gas or liquid fuels are easier to control). This strategy involves feedforward plus feedback control of the water level in the steam drum and feedback steam pressure control using the fuel flow rate. Whenever the steam pressure changes due to a change in steam demand, an action is taken to adjust the fuels flow. Since fossil fuels are generally very stable, consistent closed-loop dynamic responses are achieved. However, when biomass is used, the lumpy flow behaviour combined with fluctuations in heat of combustion introduces additional disturbances in the combustion chamber. That is, important disturbances enter the process through the manipulated variable (i.e. fuel flow) which is not the case with fossil fuels, except for fuel line pressure variations which are easily removed using flow controllers in cascade control loops. To help achieve steadier steam production, a more efficient use of biomass and a lower use of auxiliary fuels, the 3-elements boiler control scheme needs improvements.

The short term objective of this thesis therefore is to investigate early detection of heat of combustion disturbances introduced by changes in biomass flow properties, composition and moisture content, directly within the combustion chamber (i.e. the earliest time at which variations in heat released can be observed). The output of such an observer could later be incorporated into a new biomass boiler control strategy. More specifically, the forecast ability of such an observer will be studied. Indeed, forecasting steam production disturbances up to 2-3 minutes in the future would give a sufficient amount of time for the operators (or a controller) to react and attenuate these disturbances.

1.4 OVERVIEW OF THE PROPOSED APPROACH

One way to measure fluctuations in heat released by the combustion of biomass is to rely on a set of temperature sensors (i.e. thermocouples or pyrometers). However, the combustion of biomass generally occurs on a large grate and is therefore not as spatially confined as when combustion occurs at a burner outlet (gas or liquid fuels). This would require a great number of temperature sensors (i.e. an array) carefully distributed within the combustion chamber to make sure that biomass combustion is monitored effectively on the 2-dimensional grate area. Initial investment, shut down for the installation of those temperature sensors, as well as their maintenance over time may represent important operational cost.

An alternative approach consists of using ***area*** sensors, such as cameras, to monitor biomass combustion. Nowadays, high temperature RGB cameras are often already installed on boilers and kilns for process safety purpose; operators look at flame images provided by these video cameras to make sure that the flame is well behaved and that combustion is maintained within a safe region (i.e. qualitative use of the combustion images). Furthermore, it has been recently established that the heat released by a combustion process can be indirectly quantified using RGB flame images, analysed using the Multivariate Image Analysis family of approaches (Yu and MacGregor, 2004; Szatvanyi and *al*., 2006). Indeed, it was shown that heat released is highly correlated to visual properties of the flame and that of combustion chamber (colors mainly), which allows its variations to be quantified using combustion images. Moreover, disturbances in biomass feed rate or properties should be first observed in the combustion chamber through the visual appearance of the flame and surroundings before having an impact on steam pressure.

This work will therefore investigate a new machine vision approach based on Multivariate Imaging techniques for detecting disturbances related to the biomass unsteady flow and/or changing properties using images inside the combustion chamber. This approach should apply to biomass boilers in general, with some adaptation. However, the proof of concept will be illustrated using Irving Pulp and Paper's bark boiler, which was already equipped with two high temperature video cameras, and sustained variations in bark properties affect steam production.

This thesis is organized as follows. Chapter 2 covers the review of literature. The first part is about the general operation of a boiler as well as the 3-elements combustion control strategy typically used when burning fossil fuels. Few improvements that have been made and some other that could be made when using bark will also be discussed at the end of chapter 2. The end of chapter 2 covers the literature review. The design of industrial experiment as well as the information acquisition will be covered in chapter 3. Chapter 4 is about multivariate imaging technique and covers the techniques used such as multivariate image analysis (MIA) and multivariate image regression (MIR). Each steps from the image acquisition to the modelling using images will be explained. The results will be discussed in chapter 5 after which conclusion will be drawn in chapter 6. Finally, in the last chapter 7, some recommendations will be made to increase the stability and make the process easier to operate.

CHAPTER 2 BACKGROUND

2.1 BOILER OPERATION AND CONTROL

The understanding of the principles surrounding boilers is important. As shown in Figure 1, a boiler is an equipment that is composed of a combustion chamber where fuels are burned and combustion takes place. Walls around that combustion chamber are covered with small tubes in which water circulates. The small diameter of the tubes increases contact area between water inside the tubes and the hot flue gas produced inside the combustion chamber. This high contact area increases heat transfer rate between the hot flue gas and the water inside the tubes. Fresh water is first fed in the steam drum located near the top of the boiler and, by natural convection, circulates in the tubes to be heated up and evaporated to produce steam.

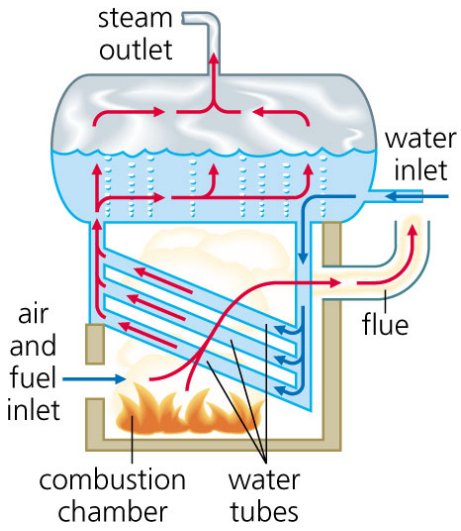

Figure 1: Schematic representation of a steam boiler (http://www.lenntech.com)

This chapter will cover boiler operation in some details. Further information on boiler operation or boiler control strategy can be found in the *Boiler Operator's Handbook*

(Heselton, 2005) and in the *Principles and Practice of Automatic Process Control* (Smith and Corripio, 2005). First, the various water treatments required before it can be safely fed into the boiler will be presented. The control strategy around the water level in the steam drum of the boiler (Smith and Corripio, 2005) will also be explained. Then, in order to have a better understanding of the fuel-air control strategy, the combustion reactions will be discussed and explanation will be given on the different kinds of fuels that can be used to produce steam, as well as how these fuels are supplied to the boiler. Finally, the 3-elements control strategy (Smith and Corripio, 2005) used in industry will be discussed and a few suggestions about possible improvements will be made.

2.1.1 WATER

The control strategy surrounding the water used in a boiler is critical. The amount of water contained in the steam drum located at the top of the boiler has to be steady. Indeed, a high level of water can result in impurities in the steam system due to water dragging. On the other hand, tubes could be severely damage by overheating if they happen to be empty due to the insufficient amount of water in the steam drum. The control strategy required to maintain steam drum water level will be discussed but, first, the water pre-treatments needed before its use is explained.

2.1.1.1 WATER PRE-TREATMENTS

Fresh water needs to be treated before it is fed to the boiler to make sure the evaporation does not lead to a fouling layer inside the tubes due to accumulation of solid impurities that cannot evaporate. Fouling would decrease the heat transfer efficiency or, in a more serious manner, plug the tubes and perturb the normal operation of the boiler. To prevent this, water is first softened (Heselton, 2005) which consists in removing impurities such as calcium and magnesium thus preventing the fouling layer which acts as an undesired insulation.

Another problem that can happen when using fresh water is corrosion. Corrosion is due to the presence of oxygen inside the fresh water. The water pre-treatments then need a de-aeration stage (Heselton, 2005) to make sure the tubes will not be prematurely damaged by corrosion.

2.1.1.2 WATER LEVEL CONTROL STRATEGY

The level of water inside the steam drum is typically controlled using a feedforward/feedback control strategy (Smith and Corripio, 2005) as shown in Figure 2. LT1 is the level transmitter allowing for feedback control actions on the feed water flow. A lower and an upper limit are also selected such that process operators have enough time to react when these limits are reached.

The main disturbance that affects the drum level is changes in steam demand. For this reason and because of the importance of maintaining a steady level in the steam drum a feedforward loop anticipating changes in water need is added to the feedback loop. Both control strategies are shown in Figure 2, where steam flow is measured and transmitted by FT5 allowing feedforward action on water valve opening. The idea behind this control loop is that each unit of steam needed requires the same amount of water to be fed in the drum. Before adding a feedforward loop, the effect of steam demand on level has to be well understood. Note that some issues related to the behaviour of bubbles are worth explaining. The bubbles have a lower specific volume compared to that of water, which allows them to displace the liquid and flow upward. When the steam demand goes up, the steam pressure drops which causes two effects; first, the existing bubbles collapse to form bigger bubbles and, second, a certain amount of hot water flashes into steam bubbles. This results in a momentary higher apparent volume of water in the steam drum due to the increased volume of bubbles going upward in the tubes. Without the feedforward loop, the feedback loop would decrease the flow of feed water. This phenomenon is called swell. The opposite situation is a shrink. It happens when the steam demand suddenly goes down causing the steam pressure to increase. Bubble size gets smaller while some bubbles are condensed into water causing the apparent volume of liquid in the drum to decrease. The steam flow

transmitter provides the feedforward signal while the level controller compensates for unmeasured flow or inaccurate measurements.

The last component shown in Figure 2, is the flow transmitter FT2. This transmitter is in a cascade loop added on the feed water valve to compensate for changes in water line pressure in the system. Those changes in water feed pressure can happen at any time and introduce undesired flow disturbances. Indeed, the water line pressure affects water flow rate even though the valve has not changed position.

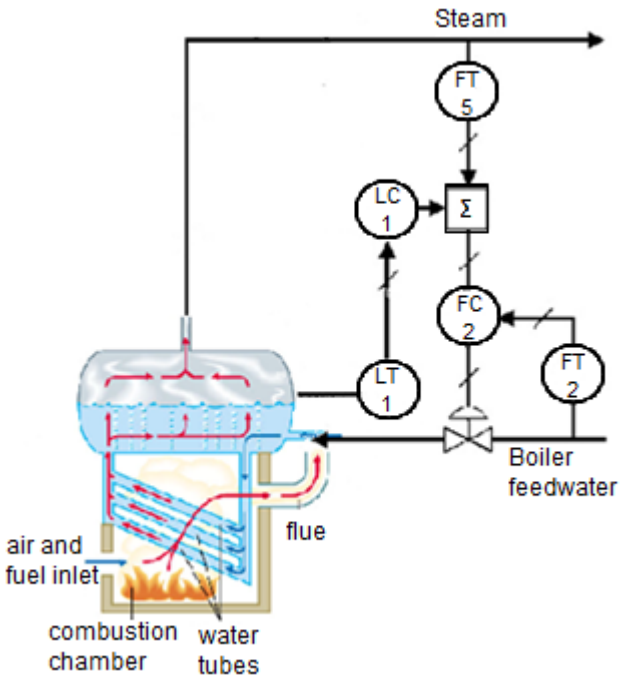

Figure 2 : Steam Drum Water Level Control Strategy (two of the 3-elements boiler control scheme) (adapted from Smith and Corripio, 2005)

2.1.2 FUELS

To produce enough heat to evaporate the water inside the tubes, combustion has to take place. Combustion is a chemical reaction between a fuel (i.e. a source of carbon) and an oxidant. The overall combustion reaction (1) happens in two steps. Carbon goes through

a first oxidation process (2) and form carbon monoxide which can be further oxidized to produce carbon dioxide (3). Only the first reaction occurs when the amount of air is insufficient. It is called incomplete combustion.

Overall reaction:

$$2C + 2O_2 \rightarrow 2CO_2 \qquad \Delta H \approx -786 \text{ kJ/mol} \qquad (1)$$

Intermediate reactions:

$$2C + O_2 \rightarrow 2CO \qquad \Delta H \approx -221 \text{ kJ/mol} \qquad (2)$$

$$2CO + O_2 \rightarrow 2CO_2 \qquad \Delta H \approx -565 \text{ kJ/mol} \qquad (3)$$

The enthalpy of reaction of the first oxidation (2) is -221 kJ/mol while the second oxidation reaction, from carbon monoxide to carbon dioxide (3), generates slightly more than twice as much heat as that of the first oxidation (i.e. 565 kJ/mol). Incomplete combustion thus leads to the lost of two third of the heat that one carbon can released. To make sure the fuel is used in an efficient way, and that complete combustion is obtained, air is introduced in excess of the stoichiometric ratio. Besides, another argument in favour of making sure complete combustion occurs is that carbon monoxide is toxic, and its emissions are ruled by environmental laws. The excess of air also prevents the production of smoke and pollutants due to the presence of carbon monoxide in the combustion chamber exhaust (i.e. flue gas). On the other hand, the greater the excess of air is, the lower the combustion efficiency. Indeed, even if pre-heated air is introduced in the combustion chamber, its temperature is lower than that inside the combustion chamber. The heat required to warm up the excess air and make sure the combustion chamber temperature is high enough to guarantee water evaporation is lost through the stack gases. Therefore, a compromise has to be made in the control of excess air and fuel flow to ensure the process is safe and does not release an important amount of pollutant while making sure

economical objectives are reached by optimizing efficiency. This makes the control strategy surrounding air and fuel flows important.

In this section, the combustion control strategy will be discussed but first, the various kinds of combustibles that can be used in combustion processes will be reviewed. A fuel consists of any material that can be oxidized in presence of oxygen and energy. Fuel is burned in order to obtain energy to heat up another gas, liquid or solid. Even if gas and liquids are the fuels that are mostly used in combustion processes, solid fuels are increasingly used in combustion processes. Since the boiler investigated in this study uses mainly bark, but also natural gas and oil, both fossil fuel and biomass combustible will be discussed.

2.1.2.1 FOSSIL FUELS

Fossil fuels are an energy source considered non-renewable due to both the fact that it takes a long period to form and that reserves are being depleted much faster than new ones are being formed or discovered. The issue of using alternative fuels is to be seriously considered, but fossil fuels are still the primary source of energy in the world.

Fossil fuels are found in the earth's crust where dead plants and animals have been caught between layers of sediments for over hundreds of millions of years. Under the exposure of high levels of heat and pressure caused by the accumulation of sediments, the organic matter has fossilized, leading to the formation of hydrocarbons, made of carbon and hydrogen. The fossil fuels sources are mainly petroleum (crude oil), coal and natural gas. The biomass boiler operated by Irving Pulp and Paper uses both oil and natural gas.

Oil is a product obtained from the distillation of petroleum. It has been classified into 6 groups, labelled 1 to 6, according to their properties such as boiling point, carbon chain length or viscosity. Carbon chain length increases with class number and so does the boiling point and viscosity which are two characteristics related to the number of carbons within the molecule. Oil number 6, also called Bunker C, is often used in furnaces and

boilers. It also has the longest carbon chain. Its viscosity is so high that it has to be pre-heated before being sent to the burners. Another disadvantage of bunker C is the high amount of pollutants it contains, particularly sulphur compounds which lead to the formation of sulphur dioxide upon combustion.

Natural gas can be found in earth in its associated form which means as a pocket of gas over crude oil or in a non-associated form, pure, in natural gas fields where it is trapped in porous rock formations. It contains mostly methane (CH_4) but also significant amounts of ethane (C_2H_6), propane (C_3H_8), butane (C_4H_{10}), pentane (C_5H_{12}) and other impurities that first have to be removed prior to be used as a fuel. Since combustion of natural gas releases less carbon dioxide per unit of energy produced, it is considered a cleaner fossil fuel. Though, natural gas still generates more pollutant than renewable energy source such as biomass, and is definitely more expansive to purchase.

2.1.2.2 BIOFUELS

In contrast to fossil fuels, biofuels are renewable combustibles because they are produced from recently dead biological material that can be rapidly replenished such as plants. The most commonly used source of biofuels are photosynthetic plants, but animal residues can also be used. Biofuels are available in either its solid, liquid or gaseous state. Liquid and gaseous biofuels can be produced using two different strategies. One of them consists of growing crops of high sugar or starch content, which are then used to produce ethanol by fermentation. The other alternative is based on the fact that few plants contain high amounts of vegetable oil that can be heated and then burned. Biofuels in solid state, called biomass, can either come from plants, such as wood chips, bark or even tree trimming residues, typically left in the forest after tree trimming, or from animals like, for example, pig or chicken waste, manure or even bacteria removed from waste water treatment which are most of the time buried. Irving Pulp and Paper's boiler uses, as mentioned, natural gas and/or oil but only if the main combustible, bark, cannot release enough heat to meet steam demand.

More specifically, bark is delivered by trucks every day and stored in the yard in natural conic shape stock piles. The angle between the surface on which the bark is poured and the surface of the pile is known as the angle of repose (Friedman and Robinson, 2002). This angle is different according to the density, the surface area and the friction coefficient of the material. Even though bark is stored outside under any weather, rain does not lead to significant additional accumulation of moisture due to the fact that the angle of repose of snow and rain is smaller than that of bark. Thus, the rain and snow do not accumulate on the bark piles. The ideal storage time is around 3 to 4 months. During this period, fermentation occurs. Carbons inside the bark components are oxidized and this reaction releases significant amount of energy. That energy helps evaporating excess bark moisture content. It was found that between 3 to 4 months of storage time (Simard, 2008), the carbon lost in the fermentation worth the amount of water evaporated from the pile.

The bark taken from the yard needs to go through a few sorting stages. The first step is a screening stage that removes large pieces. The largest passing is 2”×2”×2”. Those rejected pieces of bark are taken to a hog that grinds the large particles before returning to the screening stage. The passing bark goes under a magnet whose function is to remove every metallic piece. The bark is then ready to be fed in the boiler.

2.1.3 FUEL AND AIR FLOWS CONTROL STRATEGIES

The fuel and air flow control strategy with natural gas and oil is the same. Using biomass makes the strategy different. The general strategy using either one of the fossil fuels will be explained and, then, the differences that need to be made when using biomass will also be clarified.

2.1.3.1 CONTROL STRATEGY USING FOSSIL FUEL

The way the combustibles are fed inside the boiler influences the design and size of the boiler equipment. In the case of Irving's bark boiler, both fossil fuels are feed inside the boiler using a burner equipped with a pulveriser. The atomisation created by the pulveriser decreases the length of the flame allowing a larger width. This helps reducing the size, and thus the cost, of the boiler. The contact between the two reactants is also better due to the smaller size of combustible particles. For safety reason, the flame shall not touch the tubes that could be damaged by excessive heat. Unlike feeding bark, using fossil fuels allow the heat released by the flame to be steady since their characteristics are stable.

Fuel flow is often set according to the amount of steam needed. As mentioned earlier, air is introduced in the boiler in excess of the stoichiometric ratio. Since the air stream is cooler than the materials inside the combustion chamber, it reduces the combustion temperature and thus the efficiency is significantly reduced. This is why the amount of excess of air has to be limited. Preliminary tests covering the whole operation range are made to establish a relationship between fuel flow and the amount of air required to obtain complete combustion at this particular fuel flow rate. The amount of excess of air to use is determined by achieving a safe concentration of oxygen in the exhaust gas (i.e. flue or stack gas) within certain limits defined *a priori*. The excess of air needed typically decreases with increasing fuel flow. Indeed, the higher the fuel flow, the better is the mix between reactants due to turbulence. It is therefore easier to achieve complete combustion at higher fuel flow rates. As shown in Figure 3, the flow of combustibles is set in order to control steam drum pressure. The air-fuel relationship discussed previously (i.e. the F_A/F_F block in Figure 3) is then used to compute the appropriate air flow rate required to make sure complete combustion is achieved.

As for fresh water flow rate, both fuel and air flow rates are under feedback control in order to reject flow disturbances such as fluctuations in line pressure. These flow controllers are integrated within cascade loops as shown in Figure 3.

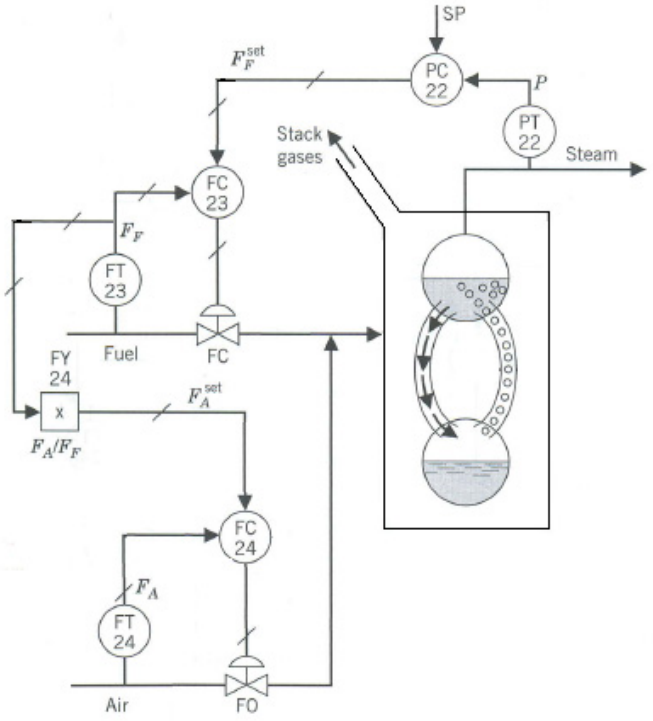

Figure 3 : Air and fuel control strategy including flow cascade loop (adapted from Smith and Corripio, 2005)

A deeper analysis of the control scheme showed in Figure 3, a problem still persists requiring the control strategy to be improved. If the steam demand suddenly decreases, the pressure inside the drum then increases due to the greater accumulation of steam. The value being higher than the steam drum set point, the drum pressure controller reduces the fuel flow by closing the control valve. Since the set point on the combustible flow has been lowered down, the air flow is also reduced. During a short period of time, the air flow rate to the burner is larger than the flow needed to maintain complete combustion, which does not cause any safety issues. If, on the other hand, the fuel flow is increased due to a higher steam demand, the pressure controller opens the fuel valve first, which suddenly results in a poorer air/fuel ratio and potentially incomplete combustion and safety issues. The control scheme is thus not appropriate for this situation because it does not guarantee that air flow rate is always sufficient for every situation. Such a lower air/fuel ratio can be very dangerous since an insufficient amount of air in the combustion chamber results in

unburned fuel in the stack gases. This does not only mean economic losses and potential environmental hazards, but this situation can lead to an explosive mixture inside the combustion chamber due to accumulation of either carbon monoxide created by the incomplete combustion or unburned fuel. The control scheme shown in Figure 4 was modified to account for this safety issue.

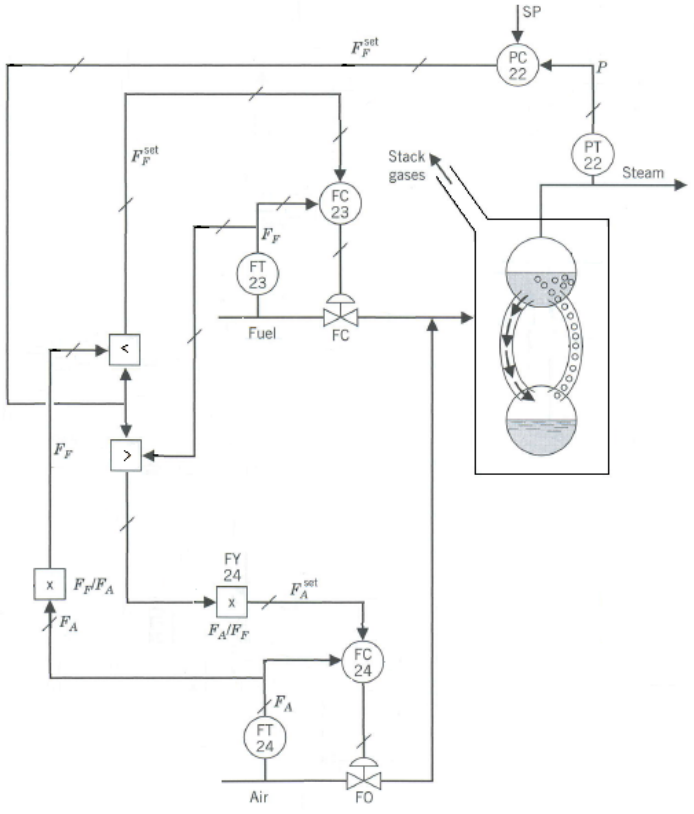

Figure 4: Modified air and fuel flow implemented control strategy to ensure safe combustion conditions (adapted from Smith and Corripio, 2005)

This modified control scheme resolves the safety issue by making sure that air flow rate always changes first and that if, for some reason, air flow does not change (i.e. sticky valve, or else), then fuel flow is maintained at its current value. At any time, the fuel flow set point (F_F^{set}) taken into account is the lowest value between the one requested by the pressure controller and the fuel flow required for complete combustion based on the current

air flow rate (computed using the fuel/air ratio). On the other hand, the air flow rate set-point (F_A^{set}) is computed based on the pre-determined air/fuel ratio using the greatest value between the actual fuel flow rate and the fuel flow set point requested by the pressure controller. This ensures that fuel flow is always lower or equal to what is necessary for complete combustion. This will be particularly useful when sudden changes in steam demand occur such as when start-up or shut-down of some process units are performed. The situation where a lower steam demand occurs will be explained as well as the opposite situation where the steam demand increases. Going through both situations will help understand how the modified control loop works.

When steam pressure sharply increases due to a sudden reduction in steam demand, then less combustibles is needed, and the pressure control reacts by lowering the fuel flow rate set point. At that very moment, the controller compares the value of the new fuel flow rate set point with the calculated fuel flow based on the actual air flow and identifies the smallest value. In this particular case, the value of the new set point on the combustible is the smallest. At the same moment, the controller also needs to set a value for the new air flow. To do that, the actual fuel flow is compared to the fuel flow set-point and the highest value is taken in the computation of the air flow. Since the fuel flow had not enough time to decrease yet, the actual fuel flow is still larger than its new set point. The air flow is then fixed considering the actual fuel flow value. The more the actual fuel flow decreases, the more the air flow is also decreased. The transient period is thus richer in air than usual.

The opposite situation occurs when the steam pressure is lower than its set point which requires the combustible flow to increase. The controller compares the value of the fuel set point with the calculated fuel flow based on actual air flow rate. Since the requested new set point is higher than its calculated value, the calculated value is selected. On the other hand, the actual value of the fuel flow is compared to the fuel flow set point and since the highest value is the one of the new set point, the air flow is calculated based on the new set point which results in increasing air flow. The more the real air flow increases and the more the calculated fuel flow based on the air flow increases. This increases the fuel flow set point since it is determined by the lowest value which will be the calculated flow. This will prevail, as long as the calculated flow does not reach the value fixed by the steam

pressure set point. This pattern ensures that the air flow is always increased first or decreased last which makes the combustible and air mixture always richer than its normal value during transient periods.

As mentioned, a compromise needs to be made on the air flow to meet both economical and safety objectives. Making sure the air flow is always in excess helps avoid producing an explosive carbon monoxide mixture due to incomplete combustion reaction. Economical objectives can be reached by minimising the excess air flow rate which requires very tight control of air/fuel ratio. This will help maintaining the excess of air introduced in the boiler at an optimum value. The goal is to maintain the amount of oxygen and/or of carbon monoxide in the stack gases between an upper and a lower limit. A CO and/or O_2 analyzer helps finding and adjusting the air/fuel ratio to an appropriate value. Oxygen level in the exhaust is maintained close to the lower limit. This helps the combustion efficiency to be as high as possible by preventing loss of energy in the exhaust due to an important excess of air fed.

2.1.3.2 CONTROL STRATEGY USING BIOMASS

The control strategy described here could be used with most of the bark boilers. Once bark has gone through preliminary sorting stages, it is fed to the boiler. Since bark is a solid combustible, it is carried from the bin to the boiler using conveyors. Four screw feeders take the bark on the conveyor and leave it inside a discharge. At the end of this discharge, an air burst throws the bark at a certain location onto the boiler grid. As seen on Figure 5, primary and secondary air flow for bark combustion, unlike fossil fuels, are not fed at a burner tip, but are rather fed in the boiler using separate ducts. As bark is thrown on the grid located at the bottom of the boiler, air is fed both under and over the gapped grid. Undergrate air prevents bark and ash from falling through the grid holes. The inclined grid is equipped with a device named a stocker that shakes the grid. The vibrations thus created allow the bark to go from one end of the grid to the other at a speed determined by the frequency of the vibrations. Vibrations also help distributing the bark load evenly on the

grid. A good distribution improves the air-bark contact and allows reducing the air/bark ratio. It is also important to understand that the wetter is the bark, the thinner the bed needs to be. The air feeding strategy is therefore very different and more complex than the one used for combustion of fossil fuels. However, the impact of manipulating the grate vibration frequency is still to be quantified since this parameter has never been changed in this project.

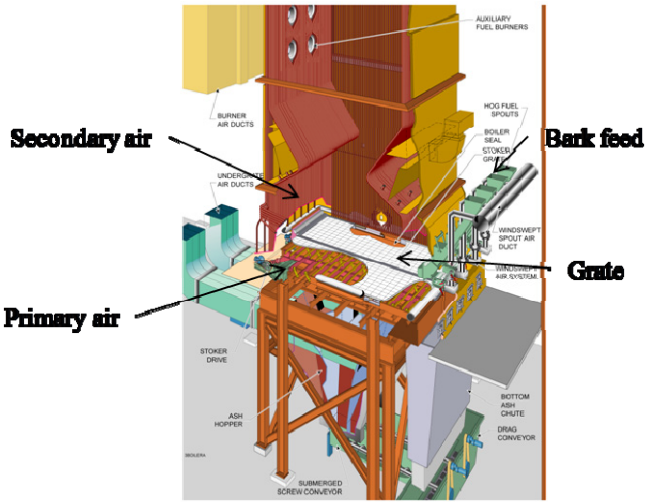

Figure 5: Irving Pulp and Paper's bark boiler

As mentioned previously, the total amount of air required for bark combustion is partly fed under and partly fed over the grid. For safety reasons, the grid is at the bottom of the boiler to make sure the fire does not come in contact with the tubes where water is evaporated. The undergrate air is mainly used as primary air (combustion reaction 2) which produces an incomplete combustion of carbon into carbon monoxide that goes up in the boiler. The overgrate air is the secondary air. It makes sure that combustion is complete by further oxidizing carbon monoxide into carbon dioxide (combustion reaction 3), which releases two thirds of the combustion heat as close as possible to the tubes. Heat transfer is thus more efficient. The secondary air is fed uniformly over the grid because the gas produced by the combustion tends to be distributed equally in the combustion chamber. Primary air needs to be fed differently.

After discussing with an expert involved in several start-ups of biomass boilers (Simard, 2008), it was determined that most biomass boiler should be operated differently depending on bark moisture content. The undergrate air blown towards the end, in the middle and at the beginning of the grid will be different. Bark is thrown at the far end of the grid by an air spout and from there it goes all the way down the grate and fall in the ash bin at beginning of the grid. When bark is very wet, the first step is to evaporate the excess of humidity before it can burn. Most of the primary air will then be fed in the middle of the grid where the bark will be dry enough to burn. A small amount will be feed at the end of the grid to allow the moisture to be evaporated and remaining of the total air flow will be fed at the beginning to burn the residual carbon left inside the ashes. On the other hand, if the bark entering the boiler is already dry enough to burn, the major part of the primary air will be feed at the end on the grid where the bark is thrown. The more the bark will progress down the grid and the less residual unburned carbon will remain and, for this reason, the air fed under those sections will be smaller.

This air control strategy can be implemented as long as bark moisture content is measured (or known from the stockpiles feeding policy) before feeding in the boiler. To design such an air distribution control scheme, preliminary tests need to be performed to determine the optimal undergrate air flow distribution for each situation. One way to make sure the amount of secondary air is not too high is to measure the O_2 concentration in the stack gases. If the concentration is too high, both the secondary air flow and the primary air flow are too high because both reactions occur on the grid and the secondary air does not contributes to oxidise the carbon monoxide which have been oxidised already by the primary air. To verify whether such a situation occurs, one simple test is to cut the secondary air flow. If steam production increases, secondary air flow rate is too high and cools down the combustion chamber. Therefore, reducing both primary and secondary air flows in this particular case improves heat transfer efficiency.

2.1.4 OTHER POTENTIAL IMPROVEMENTS WHEN BURNING BARK

As it was mentioned before, some variations in the control strategy used with fossil fuels needs to be made. Indeed, since the combustible is burned on a grate instead of inside a burner, the combustion process is very different. Therefore, some modifications were made to ensure a proper air flow distribution. However, most of the biomass boilers have a control strategy that has not been improved more. The main reason for that is the lack of research on biomass combustion. Indeed, this process being still at its beginning, the researches on the other possible improvements are limited as well as the documentation on the subject. Therefore, in order to better understand the biomass boiler issues and their solutions, an experienced engineer, Gaétan Simard, who participated to start-ups and operated a few biomass boilers, have been consulted. The next section will cover some of the improvements mentioned during this consultation, solution that could be brought to any biomass boiler.

2.1.4.1 FIXING THE AIR FLOW PROPERLY

As explained earlier, the combustible flow is used as the manipulated variable to maintain the steam pressure at a desired set-point. The air flow is then fixed accordingly. The flow of natural gas and oil (liquid) can be easily measured using conventional flow meters allowing the computation of how much air is required and how much energy will be released. When it comes to bark, a solid, the flow is a calculated value rather than a measured value. The speed of the screw feeders only provides an estimate of the bark flow rate. This value does not take into account the fact that screws could be only partially full. Since the air flow is set according to an uncertain value, it brings even more uncertainties into the boiler operation. If the air flow is fixed with the bark flow assuming the screws are full, the chances of having too much air are to be considered. Indeed, if the screws are not full and the amount of air required is set according the screw speed assuming their fullness, the amount of air is then too high. This lowers the combustion efficiency because the excess of air cools down the combustion chamber and a part of the heat produced is not used to evaporate the water but is used to heat up the air inside the combustion chamber.

For this reason, the air flow should be set according to a more accurate value. Steam pressure could be a good value to fix the amount of air needed. Knowing that to produce more steam more energy needs to be released and to release more energy more carbon is required, it becomes easy to know in advance the amount of air necessary. The screw speed could then be adjusted if the amount of steam required is not reached because if all the steam needed is still not produced with a certain screw speed it means that the amount of carbon feed inside the boiler is insufficient. Increasing the screw speed would then increase the amount of carbon fed in the boiler no matter how full the screws are. The amount of carbon fed then becomes more important than the bark flow itself. This strategy is better because the amount of carbon changes throughout the year. Indeed, the compactness of the bark changes the amount of carbon per unit of bark. So, for the same flow, the carbon content will change and for this reason, the same flow will not always produce the same amount of steam.

2.1.4.2 ASH RECIRCULATION

Combustion of solid fuels produces ash that goes in the stack or is thrown at the end of the grid. Before being released in the environment, the stack gases have to go through different cleaning stages, involving electric precipitators and/or cyclones that remove the ash. The ash thus produced by bark combustion still contains some carbon residuals when some of the bark did not have a sufficiently long residence time on the grid. Since it still contains a certain amount of carbon, getting rid of it by burying it is a loss of energy and money. This ash could, if the boiler configuration allows it, be reintroduced inside the boiler to be burned again. The amount of bark fed inside the boiler could then be significantly reduced and the combustion efficiency increased.

However, biomass combustion is still hard. Indeed, even if everything possible is done to ensure the stability of those processes, the disturbances of the feeding will always be high. Even if the instrumentation of the boiler is highly updated, the properties (i.e. moisture content, carbon content, heat of combustion) of the biomass will always change

and will always be hard to measure. Therefore, another control strategy, involving new sensors, needs to be use. For this reason, images will be used for the monitoring and the control of biomass combustion.

2.2 REVIEW OF PAST COMBUSTION IMAGING LITERATURE

A large body of literature already exists on the use of various combustion imaging techniques in the fields of burner and internal combustion engine design, where the primary concern is flow visualization, and measuring the concentration of chemical species involved in combustion reactions. These require sophisticated spectroscopic imaging techniques. For example, Laser Induced Fluorescence (LIF) and Raman imaging were used to measure the concentration of CH_4, CH radicals, CO, and OH species within methane flames (Dyer and Crosley, 1982; Namasian et al., 1989; Karpentis and Barlow, 2005). In another example, Waterfall et al. (1997) used electrical capacitance tomography (ECT) to monitor flame position and size, and the effect of air/fuel ratio within an internal combustion engine.

Combustion imaging in process applications have mainly used grey scale images to extract flame features related to combustion performance and safety considerations. A method for detecting the presence of a flame in fossil fuel fired steam boilers was proposed by Bae et al. (2006). Various geometrical and luminous flame properties were extracted from grey level images, and classified using Artificial Neural Networks into arbitrarily defined states related to combustion performance (Victor et al., 1991; Bertucco et al., 2000; Zhang et al., 2008). Flame classification was also proposed for monitoring fuel and air flow rates (Tao et al. 1995), and for on-line tracking of pulverized fuels from different sources (Xu et al, 2005). Flame morphological and luminous features were also used for predicting a large number of combustion properties such as flicker rate (Huang et al., 1999, Lu et al. 2004), unburnt carbon, CO_2 and NO*x* emissions, ignition point, spreading angle, and temperature (Shimoda et al., 1990; Lu et al, 1999, 2000, 2004; Yan et al., 2002). An

improved flame segmentation algorithm for rotary kilns was also developed by Sun et al. (2008) using the Gabor wavelets.

A few research studies used RGB color images for extracting flame characteristics. Wang et al. (2002) took advantage of the three color channels to estimate flame temperature (i.e. bi-color method) and to predict NOx concentration in the exhaust of a power boiler. Some unpublished color features were also extracted in the work reported by Keyvan (2005) after segmentation of RGB flame images. These features were correlated to air/fuel ratio obtained within a multiple burners glass furnace.

More complex spectrometric devices were also investigated to obtain additional information about chemical species within the combustion chamber. Allen et al. (1983) used infrared spectrometers and neural networks to estimate the flame temperature inside a utility boiler. An infra-red spectrometer has also been used by Yamaguchi et al. (1997) as a sensor to detect the air/fuel ratio of a premixed flame in a gas boiler. Finally, Romero et al. (2005) have reviewed the application of VIS-IR spectrometry to provide real-time flame stoichiometry and temperature measurements at the burner level in natural gas-fired glass furnace applications.

An important issue with most of the grey scale and color imaging techniques reviewed above is that the flame geometrical and luminous properties are extracted directly in the image space. Hence, every incoming image needs to be segmented in order to identify the pixels belonging to the region of interest (i.e. the flame) from which combustion features are calculated. In industrial combustion systems, the air and the fuel are non-premixed and the flames are highly turbulent. The flames bounce around continuously, change location, size, and shape even if they release the same amount of energy. Thus, using traditional image space segmentation approaches requires spending important computational efforts in tracking flame properties (i.e. location in the image) that are not fully relevant for estimating heat released.

A more efficient method was first proposed by Yu and MacGregor (2004), investigating a liquid waste boiler, and later applied by Szatvanyi et al. (2006) to a rotary kiln fired using two fuels where one was a process by-product. They have shown that using

Multivariate Image Analysis (MIA) for extracting information from RGB images of turbulent flames was a very effective method, since segmentation is rather performed in a feature space (i.e. transformation of the original image) where image pixels are classified according to their colors (i.e. spectral signature) regardless of their location within the image. These color features were demonstrated to be closely related to the heat released by the combustion system. Hence, two flame images collected under similar combustion conditions (i.e. releasing similar amounts of heat) will have very similar spectral characteristics within the feature space, even if their shape, size and location within the combustion chamber is different. Segmentation in the feature space would be performed at only a few occasions; initially and every time the model is updated and not for every incoming image as is the case for the traditional image space segmentation approaches. In this research, Multivariate Image Analysis (MIA) will be used to extract relevant color features from RGB flame images that are the most highly correlated with steam production in biomass boilers. These concepts will be explored further in the remainder of this thesis.

However, this work deals with different issues. As it has been mentioned, solid flow creates more disturbances in the flame images. As seen on Figure 6, two images taken one second apart will have a very different appearance without releasing a different amount of energy. To our knowledge, this problem has never been investigated in the past with biomass boilers even though it is increasingly being used in industries.

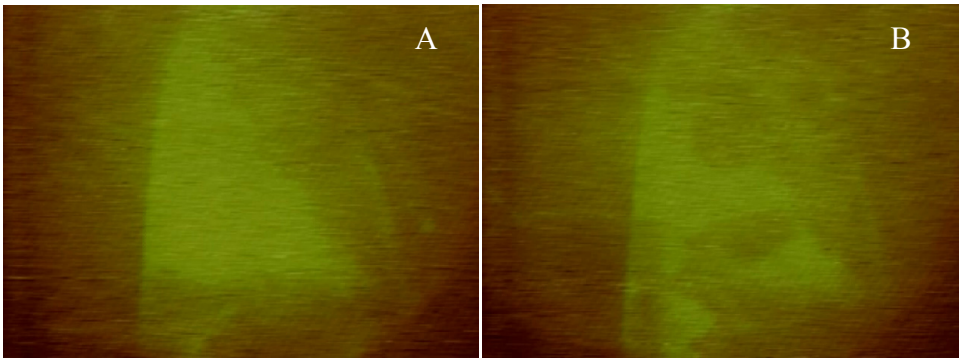

Figure 6: Flame images from inside Irving's bark boiler taken A) at time t B) at time t + 1s

CHAPTER 3 DESIGN OF INDUSTRIAL EXPERIMENTS AND DATA COLLECTION

3.1 IRVING PULP AND PAPER'S BARK BOILER IMAGES

The methodology presented here can be applied to any biomass boiler but the proof of concept will be made on Irving Pulp and Paper's bark boiler. Therefore, their bark boiler will be presented in this section.

Every combustion process needs to be equipped with a system that allows the operator to know whether the combustible is burning or not. This security requirement ensures the safety of the process by preventing accumulation of unburned combustible inside the combustion chamber which could lead to an explosive situation. Some systems have scanners that give analog signal to the operator but most of the combustion processes, including Irving's bark boiler, are equipped with high temperature cameras looking inside the boiler giving thus a visual feedback on the flame to the operator. As shown on Figure 7, Irving Pulp and Paper's bark boiler has two cameras. The first one is located on the third floor and looks across the grate where the bark is thrown. The second one is on the fifth floor located under the fossil fuels burners looking down at the grate.

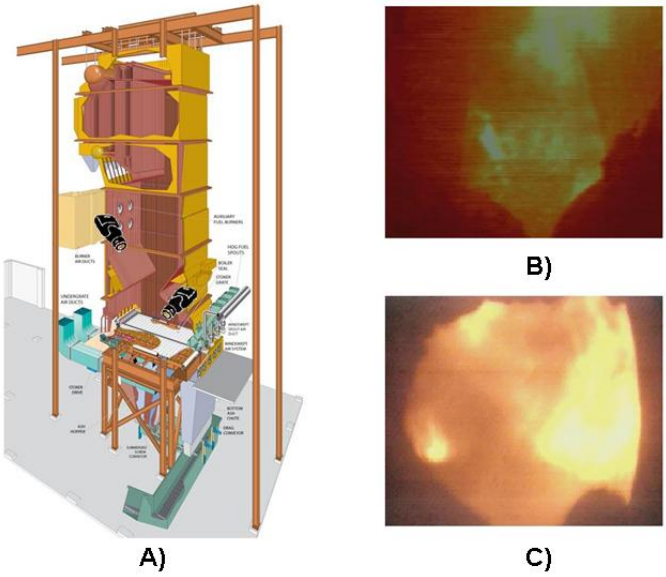

Figure 7: Location of the two high temperature cameras on Irving's bark boiler (A) with the view provided by both camera on fifth floor (B) and on third floor (C)

The image shown in Figure 7B is taken from the 5th floor camera and shows the combustion of bark on the grate, viewed from the top. The second image (Figure 7C) was taken by the 3rd floor camera, and provides a side view of bark combustion on the grate. The 3rd floor camera has a limited field of view while images from camera on fifth floor shows a better field of view even if located over the arches of the boiler. Both high temperature video cameras (Quadtek M530) provide 8 bit RGB images with a spatial resolution of 576×768 pixels.

3.2 DESIGN OF INDUSTRIAL EXPERIMENTS

A series of experiments were performed on bark flow rate and properties as well as on the fossil fuel flow rates (i.e. oil and natural gas) in order to generate a rich database for building dynamic models for steam production rate. These tests were conducted in June 2008 at Irving's Pulp and Paper mill.

3.2.1 FOSSIL FUELS EXPERIMENTS

The objective was to vary the flow rate of fossil fuels to quantify their influence on steam production. However, bark is the less expensive combustible and was burnt at all times except under very special circumstances. It was burnt during the experiments made on fossil fuels. Bark properties and feed rate were maintained as constant as possible during the fossil fuel experiments. Bark flow and properties are hard to control, and these were obviously changing during these experiments as they usually do during normal operation. The changes made on fossil fuels were therefore made large to distinguish their contribution to steam production from that of bark. Tests were made on both fossil fuels (i.e. natural gas and oil) and data were acquired at a sampling rate of 10 seconds. This was the fastest sampling rate available from Irving Pulp and Paper's data archiving system.

3.2.1.1 NATURAL GAS TESTS

Tests on natural gas were performed on the afternoon of June, 17th 2008. No natural gas was used until the beginning of the tests. Only then, natural gas was burned using one of the four gas burners. Its flow was increased so the contribution of the natural gas on the steam production was around 40 kilo pounds per hour. At the beginning of the test (between 1000-2000 seconds) shown in Figure 8, bark flow was reduced and natural gas flow rate was increased to make the effect of natural gas clearer, but changes in both were made such that steam production remained approximately constant (i.e. both fuels were varied according to their respective heat of combustion, but total heat released is about the same). These conditions were kept steady for about half an hour, long enough to let the process stabilize. Then, step changes were made on natural gas. Without changing any other variables, the natural gas flow was doubled. Because the pressure of this new flow would have been too high in one burner, two burners were used. Once again, everything else was kept steady until the process had reached a new steady-state and then, the natural gas flow was returned to its initial value until the end of the test.

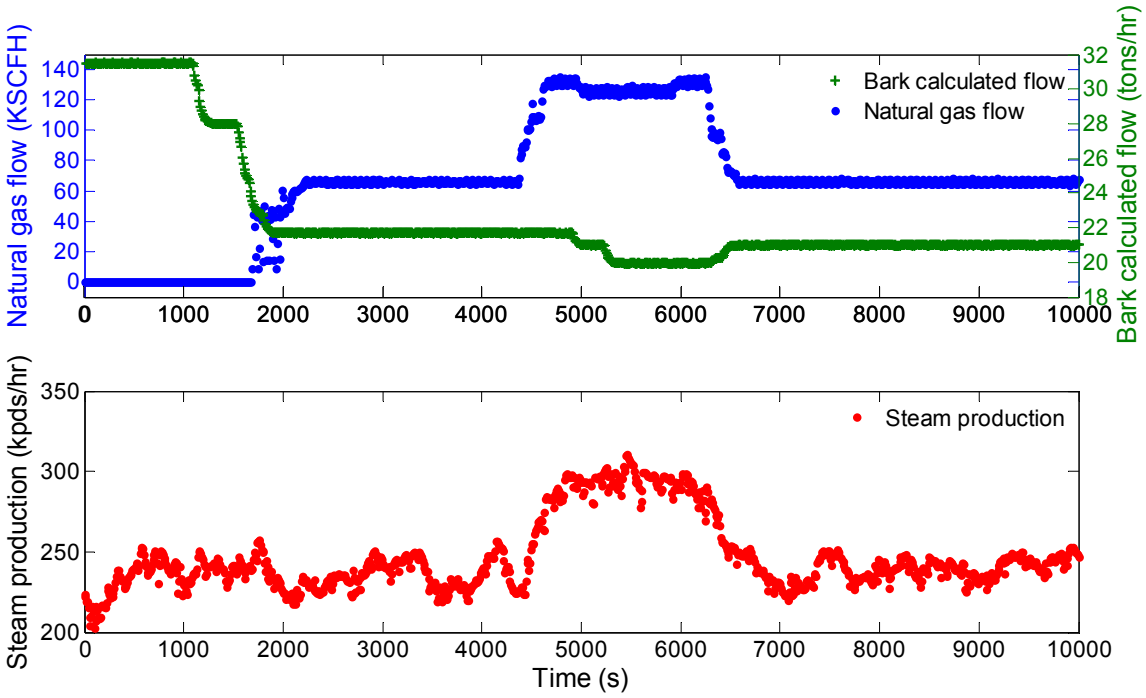

Figure 8: Natural gas experiments

3.2.1.2 OIL TESTS

The next day (June, 18th 2008), the tests on oil, the other fossil fuel normally used at Irving P&P, were made as shown in Figure 9. Once again, bark was also used throughout the test duration, which was implemented similarly as for natural gas. Note that the oil flow meter seems inaccurate since the oil flow value is 1.2 before the test started even if no oil was burnt. The test started by lowering bark flow rate and increasing oil flow rate (burnt using one of the four burners). As for the natural gas test, steam production was maintained at about the same level through an appropriate balance between the two fuels. Both combustibles were kept constant for about half an hour to let the process reach steady-state. Step tests on oil flow rate were then implemented first by increasing oil flow rate by a twofold factor and kept at that level until the new steady-state was achieved. Then oil flow was returned to its initial and the test was stopped after stabilisation. The step change on oil flow rate in percentage was smaller than the one on natural gas since oil is more difficult to burn. Therefore, the variance of steam production introduced by bark is more important compared to what it was in the natural gas test.

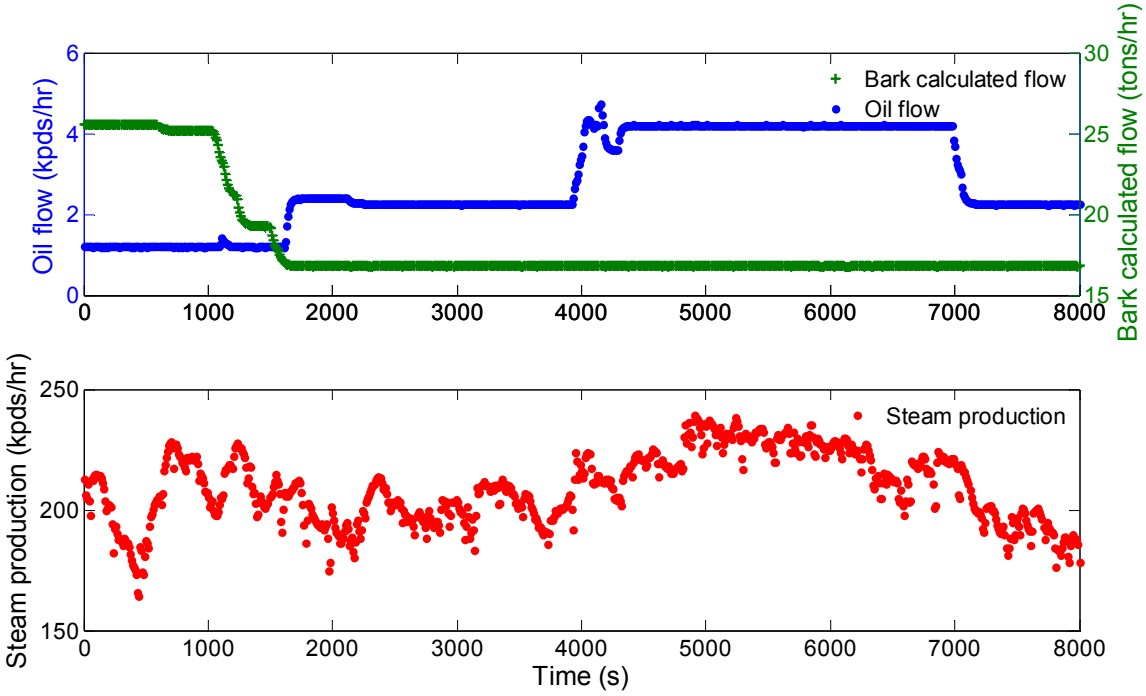

Figure 9: Oil experiments

3.2.2 BARK EXPERIMENT

Bark is different from fossil fuels. This solid combustible can change in two ways. First of all, as it was the case with natural gas and oil, its flow can vary. The main difference between biomass such as bark and fossil fuels lies in the fact that biomass properties can also be different from time to time. Unlike any fossil fuel whose composition is very stable and well known, bark properties such as carbon content and humidity that are constantly changing are most of the time unknown. Therefore, since two factors can have an impact on steam production, two kinds of test were made. Indeed, the moisture content as well as the carbon content needs to be changed along with the bark flow and that, for each combination of properties. This section will describe the way the properties of the bark are changed. The results of the tests made on bark will after that be explained.

3.2.2.1 CHANGING BARK PROPERTIES

The series of tests performed on bark did not focus on flow rate only, but also on its properties to test whether changes in its combustion energy could be detected using the proposed imaging approach. Bark moisture and, therefore, its carbon content were changed simultaneously so any situations would be represented. The operation during summer, winter as well as other seasons was simulated by adjusting moisture content in the bark fed to the boiler.

Different types of bark stock piles are maintained at Irving P&P plant, each having a different composition (bark, sawmill dust, wood chip rejects, pieces of fiberboard) and moisture content, allowing the operators to introduced designed disturbances in bark feed properties. The two mostly used bark piles are the covered bark which is dry, and piled bark which is subject to weather conditions, and therefore is wet. Sawmill dust and wood chip rejects are continuously mixed with the bark feed to the boiler. Therefore, changing the amount of moisture in the feed is relatively easy. The ratio of dry/wet bark fed to the boiler is changed by adjusting the relative amount of dry and wet bark the loader driver pours into the bin located at the beginning of the process. For example, a ratio dry:wet of 3:9 means that for each bucket of dry bark loaded into the bin, 3 similar buckets of wet bark are loaded into the bin as well.

Changing bark moisture content seems straightforward, but following the bark path from the yard to the boiler will help understand the difficulties associated with changing bark properties. First, the bark is taken in the yard by a loader and the desired recipe is charged into the reclaim container. The screening stage is the following step. The accepted bark then goes in another bin. This last bin feeds the conveyor that brings the bark to the boiler's screw feeders.

According to Irving's operators, bark takes about 5 minutes to go from the reclaim container through the screening stage, to the bin. From the bin, the different conveyors then take an additional 2 minutes to bring this bark to the boiler. Considering the residence time within both bins, changing bark recipe (i.e. moisture) takes about half an hour to reach the bark boiler. Therefore, the bark mixtures are not changed very often. If the operator observes that the bark is too humid, the short term solution does not rely on changing the

bark mixture, but temporarily burning fossil fuels. This will have to be taken into account when a control strategy will be put online.

3.2.2.2 BARK TESTS RESULTS

Those tests were performed using no other combustibles but bark. The amount of moisture in the bark feed was progressively changed from 12:0 (dry bark) to 3:9 (very wet bark). Bark flow rate was also changed when moisture content allowed it. Indeed, when the bark feed is too wet, increasing flow rate may kill the fire on the grate since most of the available energy within the combustion chamber would be used for drying the wet bark, lowering the temperature to a point where combustion is no longer possible. A description of each test will be given.

On the morning of June 20th, 2008, the two wetter bark mixtures were fed in the boiler. These tests are shown in Figure 10 C) and D) for dry:wet mixtures of 4:8 and 3:9, respectively. Step changes were made around 4000 seconds during the 4:8 mixture to investigate the effect of bark flow rate steam production when bark is wet. However, bark flow was quickly resumed to its initial value since the bark was pilling up on the grid instead of burning. The fire was being killed by the high moisture content. The wettest bark mixture (3:9) is shown in Figure 10 D). Bark flow was maintained constant during the test since any variation would affect steam production.

Two drier mixtures were also studied. The driest one, the 12:0 mixture, was fed during the morning of June 18th, 2008 and the results are shown in Figure 10 A). The positive step change is followed by a decrease in steam production. According to the operator, this phenomenon is due to the sudden lower amount of bark fed. The temporarily insufficient amount of bark on the feeding conveyor did not allow the increase of the amount of bark fed inside the boiler. The last mixture, using equal amount of dry and wet bark (6:6), is the one normally fed in the boiler. All the tests when bark was fed were based on this mixture, except when the mixture was varied as discussed above. A representative period of time (afternoon of June 18th 2008) was selected for illustration purpose and is shown in Figure 10 B). A few steps changes on bark flow rate were also implemented when

feeding 6:6 mixture. The increase of the screws speed results into a higher steam production. When looking at the bark flow-steam production dynamics, it is clear it has a more complex, apparently non-linear, behaviour compared to that of fossil fuels since bark dynamics depends on its physical properties. It mostly depends on bark moisture content, which varies with time, as well as the degree of filling of the screws.

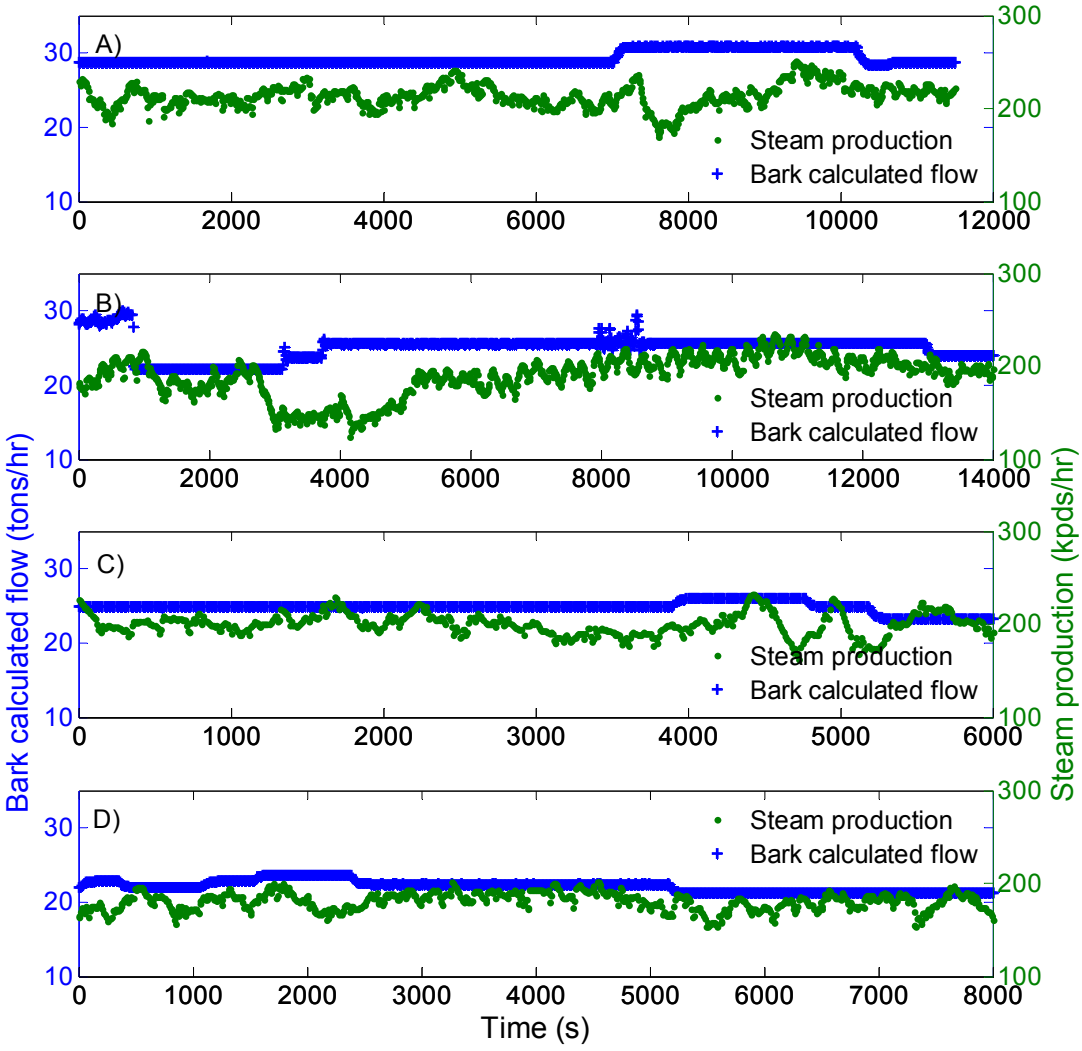

Figure 10: Bark tests from drier to wetter mixtures (A) 12:0 (B) 6:6 (C) 4:8 (D) 3:9

3.3 PROCESS DATA AND IMAGE ACQUISITION

During each dynamic experiment, images and process data were acquired. Digital RGB images were extracted from the video captured using both cameras. The effective capture rate was 1 frame every 1 or 2 seconds which was the fastest rate the camera could record images. All measurements routinely available through the boiler instrumentation were also collected (10 seconds averages), including temperatures, air and fuel flow rates, water temperature and steam pressure and flow rate. In total, 154 300 images and 34 564 observations on 320 variables were collected. The number of images and data available for each test is shown in the following table.

Table 2: Information available for each test

		Observations available for each variable	Images available from each camera
Fossil fuels tests	**Natural gas**	1000	8789
	Oil	800	3749
Bark tests (from dryer to wetter)	**12:0**	1150	10107
	6:6	1400	12705
	4:8	600	5086
	3:9	800	7243

Preliminary test were made to identify the important process measurements (variables). A PLS model (Erikson et al., 2006) showed that 28 variables were the most important for the prediction of steam production. The criterion used for this selection was a statistics known as the variable importance on projection (VIP). The reader is referred to Erikson et al. (2006) for details on the VIP statistic. The variables that had a VIP higher than 1 were considered the most important in the model. This rule of thumb is typically used in the chemometrics literature (Erikson et al., 2006). The following table shows the most important process variables as well as the tag number and a short description. Although it should have an important impact on combustion and steam production, grate vibration frequency could not change during the tests at Irving.

Table 3: Important TAG number and description

TAG description	TAG number	TAG description	TAG number
FEEDWATER FLOW CONTROL	53FC102.PV	#1 BARK FEEDER AMPERES	53II3306.PV
NORTH UNDERGRATE AIR FLOW CONTROL	53FC107.PV	#2 BARK FEEDER AMPERES	53II3307.PV
SOUTH UNGERGRATE AIR FLOW	53FC108.PV	#3 BARK FEEDER AMPERES	53II3308.PV
BARK BOILER OIL FLOW	53FC143.PV	#4 BARK FEEDER AMPERES	53II3309.PV
BARK FEEDERS FLOW	53FC300.PV	CONDENSATE FROM STRIPPING COLUMN FLOW	50FC2236.PV
#4 BARK FEEDER FLOW	53FC311.PV	RED AVERAGE (cam3)	-
NATURAL GAS FLOW	53FC567.PV	COLOR AVERAGE (cam5)	-
SOUTH REAR UNDERGRATE AIR FLOW	53FC620.PV	RED AVERAGE (cam5)	-
NORTH REAR UNDERGRATE AIR FLOW	53FC623.PV	GREEN AVERAGE (cam5)	-
DNCG FAN DISCHARGE FLOW	53FI857.PV	BLUE AVERAGE (cam5)	-
TOTAL BARK AIR FLOW	53FX087.PV	FUEL OIL TEMP	53TI090.PV
TOTAL UNDERGRATE AIR FLOW	53FX088.PV	STEAM DRUM TEMP B	53TI161.PV
TOTAL NORTH UNDERGRATE AIR FLOW	53FX107A.PV	STEAM DRUM TEMP C	53TI162.PV
TOTAL SOUTH UNDERGRATE AIR FLOW	53FX108A.PV	STEAM DRUM TEMP D	53TI163.PV

CHAPTER 4 MULTIVARIATE IMAGING TECHNIQUES

4.1 APPROACH OVERVIEW

This section will give an overview of the proposed approach. As shown in Figure 11, the design of experiments allowed gathering images and data. Images were first extracted from the captured videos and then synchronized with the process database. Useful features are extracted from the score density histogram obtained using principal component analysis (PCA). Images features are combined with process data to build input and output data matrices for regression ($\mathbf{X}_{\mathbf{MIR}}$ and $\mathbf{Y}$). The model built is then analyzed and is modified iteratively if the predictive performance is not satisfactory until the final model is obtained.

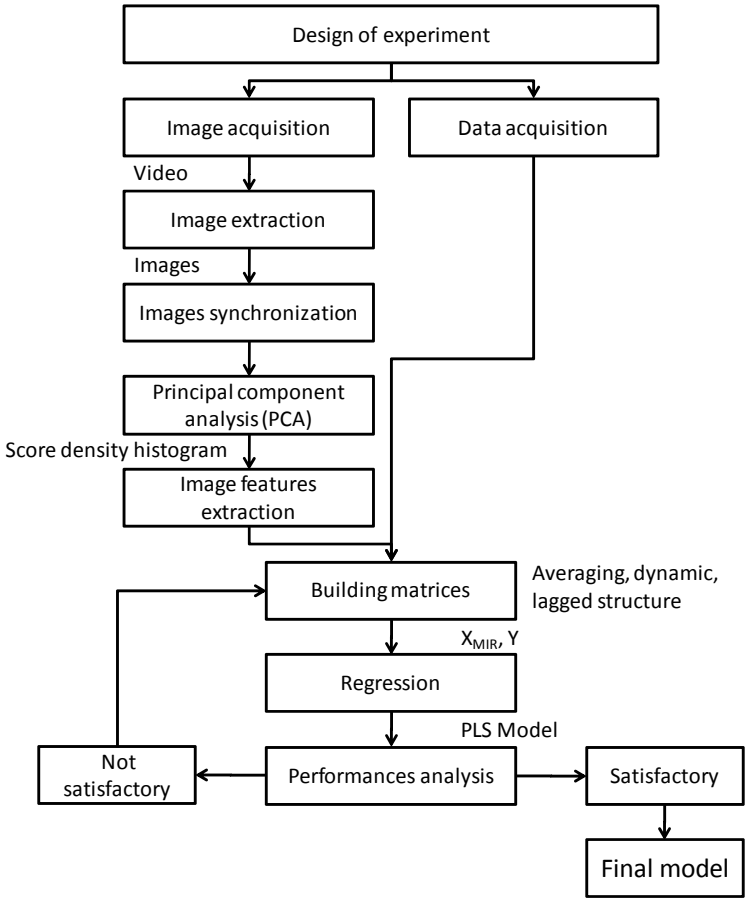

Figure 11: Overview of the proposed Multivariate Imaging Approach for predicting steam production

The procedure used for extracting image features, involving PCA and score histograms (called Multivariate Image Analysis, MIA), for generating input and output data matrices reflecting process dynamics, and for building the regression model will be explained in more details in this chapter.

4.2 MULTIVARIATE IMAGES ANALYSIS (MIA)

Multivariate Image Analysis (MIA) was originally proposed by Esbensen and Geladi (1989) as a very effective method for extracting color, or more generally, spectral

features from multivariate digital images (i.e. images having several color or spectral channels). Since color features are the most interesting characteristics for predicting steam production, this method is presented in this section.

4.2.1 DIGITAL IMAGES

Images captured by both cameras are RGB images. An RGB image, also called a true color image, is captured by a camera equipped with a charge couple device (CCD) sensor made from an array of photosensitive sensors (semiconductor material) of a given rectangular size, each defining one single pixel within the resulting image. To capture the three-color images, the CCD has three types of sensors, some sensitive to the wavelengths corresponding to the red color, the green, and the blue colors. The sensors are distributed on the CCD according to the *Bayer pattern*. When the incoming light from the scene reaches the CCD sensor, a photo-electric effect converts the energy of the light (intensities within a wavelength range) into a proportional accumulation of electrical charges. When the camera exposure time is completed (user-defined), each sensor of the CCD is discharged (in parallel) and an electric current results. An analog to digital converter takes the analog signal (i.e. discharge electric current), and converts it into a digital signal, which are integers varying from 0 to an upper value defined by the converter resolution (the number of bits of the camera). For example, an 8-bit camera will describe light intensity into 256 grey levels (i.e. 0-255). All these numbers are organized according to their original position on the CCD sensor which allows creating an image of the scene. Once the image is captured, a signal is sent to the computer for recording.

Once digitized, these images become a 3-way array of data $\underline{\mathbf{X}}$ ($x \times y \times \lambda$), with two spatial dimensions (i.e. x and y) defining the pixels of the image (i.e. discretization of the scene), and a third (λ) containing the light intensities at each wavelength as captured by the camera CCD as shown in Figure 12. The resolution of the RGB images (3 channels) captured by both cameras used at Irving P&P is 576×768 (height×width). Therefore, the i^{th} RGB combustion image is $\underline{\mathbf{X}}_i$ (576×768×3).

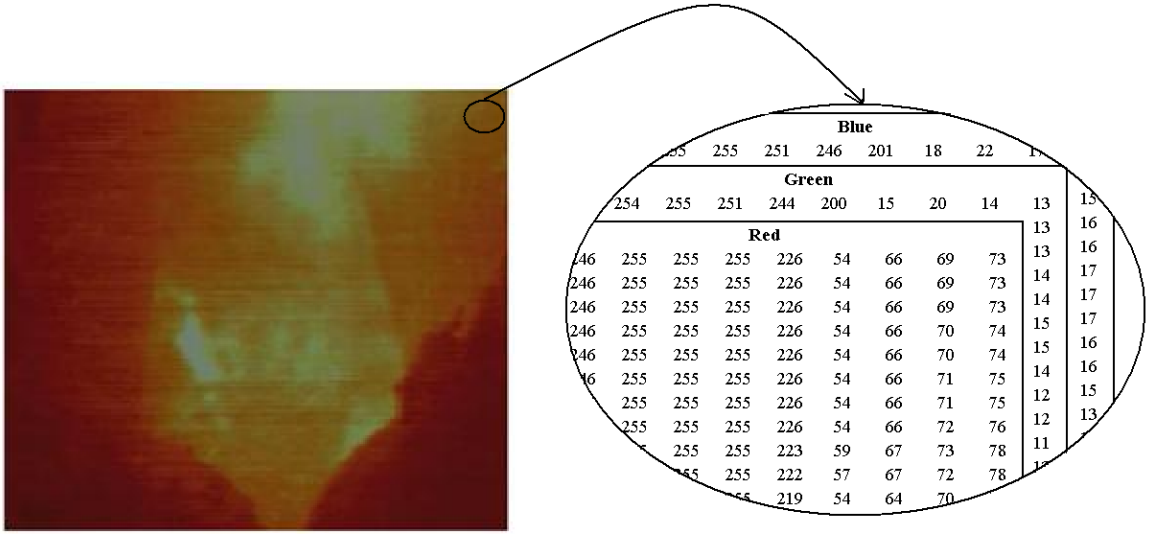

Figure 12: Image digitization

4.2.2 PRINCIPAL COMPONENT ANALYSIS (PCA)

Each image is then reorganized into a matrix $\mathbf{X_i}$ $(x \times y) \times 3$ by collecting the RGB color intensities of each pixel row-wise as shown on Figure 13. Each reorganized image $\mathbf{X_i}$ has now 576×768 rows (i.e. as many as the number of pixels of the original image), and 3 columns for the red (R), green (G), and blue (B) intensities of each pixel.

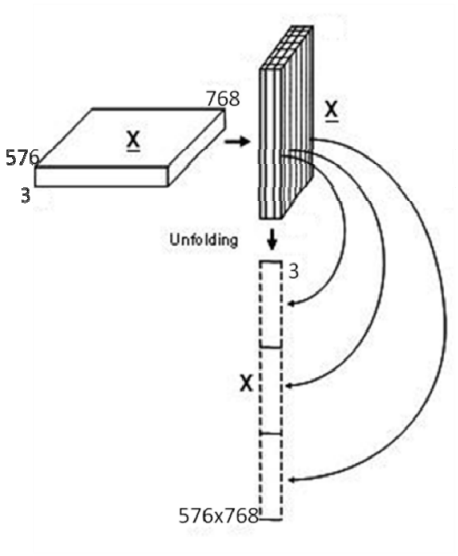

Figure 13: Unfolding operation of an array $\underline{\mathbf{X}}$ into a matrix $\mathbf{X}$ before applying PCA

When applied to a single image, the next step in MIA consists of applying Principal Component Analysis (PCA) on the image matrix $\mathbf{X_i}$. This particular step finds the orthogonal direction where the most variation into the data is observed. PCA is based on the same principles as performing singular value decomposition (SVD). Since the matrix $\mathbf{X}$ has as much as 442,368 rows (i.e. 576×738), SVD is rather performed on the $\mathbf{X^TX}$ matrix instead. The $\mathbf{X^TX}$ matrix is of much lower dimensional matrix which, therefore, is advantageous for computation efforts and results in the same principal directions as if PCA was directly applied to $\mathbf{X}$. This is called the kernel algorithm (Geladi and Grahn, 1996). To illustrate the method, a 3-D plot can be created using the values of red, green and blue intensities for one image as the three axes.

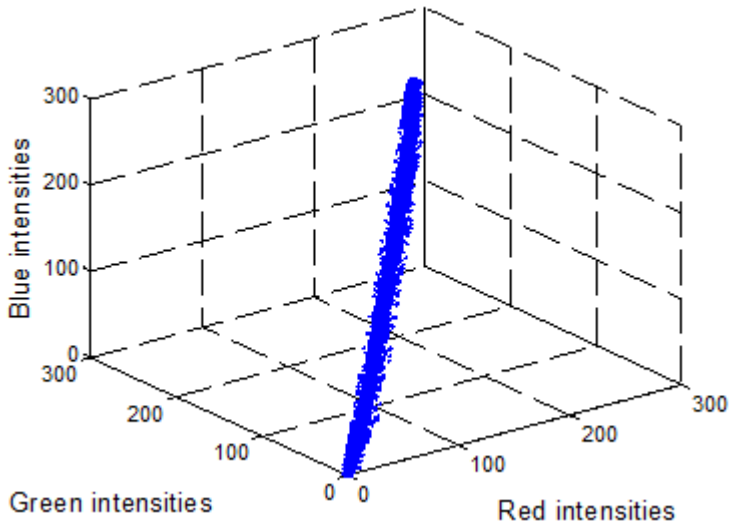

Figure 14: 3-D plot representing the distribution of red, green and blue intensities for each pixel of a combustion image

As it can be seen in Figure 14, where each point represents one single pixel of the original image, the pixels of an image are oriented in a certain direction allowing the reduction of the number of dimensions of the 3-D plot. The direction of the most variation into the 3-D plot is called a loading vector $\mathbf{p_1}$ while the second direction of the largest source of variation into the data that is also orthogonal with $\mathbf{p_1}$ is called the loading vector $\mathbf{p_2}$. The third direction, also perpendicular to the first two, which would represent the third direction, would be the last loading vector $\mathbf{p_3}$. Adding a third dimension to the first two would be the same as performing a rotation of the original axes. PCA or SVD of $\mathbf{X}$ (or $\mathbf{X^TX}$) allows computing the loading vectors $\mathbf{p_1}$, $\mathbf{p_2}$ as well as $\mathbf{p_3}$. The projection of the data ($\mathbf{X}$) onto these principal directions are called the score vectors and are obtained by projection (i.e. $\mathbf{t_i}=\mathbf{Xp_i}$). Mathematically, the PCA decomposition is represented as followed.

$$\mathbf{X_i} = \sum_{\mathbf{a}=1}^{\mathbf{A}} \mathbf{t_a}\, \mathbf{p_a^T} + \mathbf{E_i} \qquad (4)$$

where A is the number of principal components, the $\mathbf{t_a}$ vectors are the score vectors and the $\mathbf{p_a}$ vectors are the loading vectors. As explained earlier, the loadings are orthonormal vectors defining linear combinations of the RGB color intensities explaining most color variations in an image; $\mathbf{p_1}$ is oriented in the direction of the greatest amount of variance in the RGB color space whereas $\mathbf{p_2}$ is the linear combination explaining the second greatest source of variance, and so on. The fact that the pixels are already oriented in a certain direction allows, once again, a space reduction from a 3-dimensional space onto a 2-dimensional space (i.e. the plane defined by the first two components). Since most of the information contained within RGB images is generally captured using 2 components (A=2), this leaves the 3rd dimension in the PCA decomposition residuals **E**. Dimension reduction for RGB images it typically less important than for multi- or hyper-spectral images. Since any color is some linear combination of RGB intensities, the loading vectors are defining a new two dimensions color space. Those two vectors (not of unit length on the image for visualisation purpose) are shown in Figure 15.

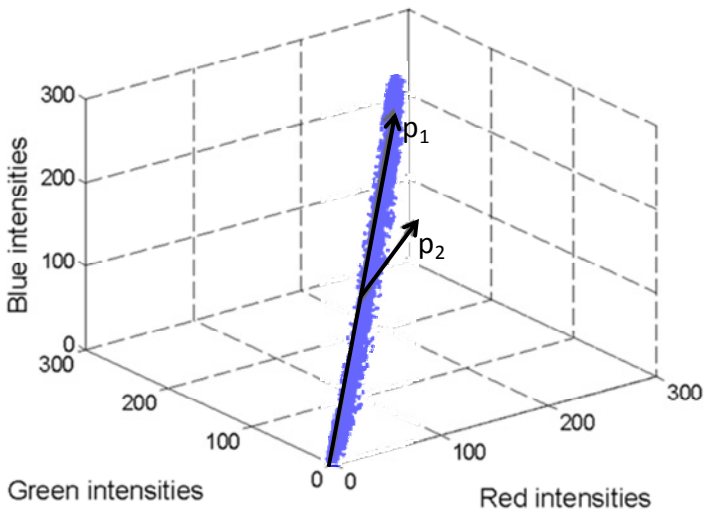

Figure 15: Loadings vectors $\mathbf{p_1}$ and $\mathbf{p_2}$ together define a plane in $\Re^3$ space.

Note that, in general, when using PCA on a data matrix, the latter is mean-centered and scaled in some way, typically to unit variance, in order to explain variations about the

mean of the variables, and to attribute equal importance to each variable which usually have different engineering units. In the case of images, mean centering and scaling is not necessary since the mean light intensities are meaningful (i.e. changes in mean color from image to image is important to capture) and the various color channels are measured in the same units and span a similar range (i.e. 0-255 for a 8 bit image).

4.2.3 SCORE DENSITY HISTOGRAMS

As discussed above, each pixel of the 3-D plot can be projected onto the plane defined by the first two new directions (principal components). The scores are, in fact, the new coordinates of each point on this new plane whereas the residuals **E** correspond to the distance of each observation off that plane (i.e. orthogonal projection error). The score vectors (i.e. $\mathbf{t_a = X\,p_a}$) can be interpreted as color intensities in this new color space.

Since score plots contain as many points as number of pixels of the image (i.e. 576×768 = 442,368), it is a common practice to enhance the visual appearance of these plots by displaying them as density histograms discretized into 256×256 bins (Esbensen and Geladi, 1989; Geladi and Grahn, 1996; Yu and MacGregor, 2004), as shown in Figure 16. To do so, the scores are first scaled so their values become integers between 1 and 256 using the following equation.

$$T_{k,i} = Round\left(\frac{t_{k,i} - t_{k,\min}}{t_{k,\max} - t_{k,\min}} \times 255\right) + 1 \qquad (5)$$

Where k represents the principal component number and i represents the observation number (i = 1..442,368).

To obtain the score density histogram for one image, the range spanned by T_1 and T_2 is first discretized using a 256×256 rectangular grid. Each bin of the grid corresponds to a particular combination of T_1 and T_2 values. The count of the number of pixels having a given T_1-T_2 combination is stored in the appropriate position within a 256×256 square matrix **H**. The $h_{i,j}$ element of matrix **H** contains the total number of pixels having a T_1-T_2 combination corresponding to position (i,j) in the grid. This matrix can be represented as a density histogram, which is simply the 2-D joint density distribution of the first two scores. A hot color map is used to indicate pixel density (represented by element $h_{i,j}$) for each T_1-T_2 combination; black means zero pixels and white means highest pixel density.

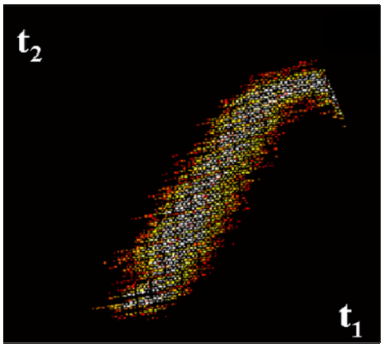

Figure 16: Score plot density histogram for a particular bark combustion image.

When MIA is to be applied to a set of J images, obtained, for example, during a design of experiments, as it is the case with this project, the orientation of the plane used to project the pixels (i.e. the loading vectors) as well as the scaling range of the score vectors t_1 and t_2 need to be common for all images of the set in order to make a comparison possible between images (i.e. score density histograms for all images are congruent). Using the scaling range on t_1 and t_2 allows the reference to be the same which let the different combinations of t_1-t_2 to represent the same color in the original image and this, despite the image.

A common PCA model for J images is obtained by applying PCA on the kernel matrix $\Sigma(\mathbf{X_i}^T\mathbf{X_i})$, i=1,2..,J (Geladi and Grahn, 1996; Yu and MacGregor, 2004). This yields a global model of all J images with a single common set of loading vectors. Score vectors are again computed, for each single image, as $\mathbf{t_{a,i}} = \mathbf{X_i}\ \mathbf{p_a}$.

4.2.4 IMAGE FEATURE EXTRACTION USING MIA

The MIA approach classifies the pixels according to their color features instead of their position in the original image. However, the location of the pixel in the original image is stored in memory which allows going back and forth from the score plot to the original image and so on to see what color is represented by different regions of the score plot, and therefore, to explore the information contained in the images. The $\mathbf{t_1}$-$\mathbf{t_2}$ score plot is used to explore the image color features using MIA in a very similar way as when PCA is used to analyze process data. Score scatter plots reveal the presence of clusters of observations that are similar in a multivariate space. In image analysis, such a plot shows clusters of pixels (observations) sharing similar colors or spectral features.

Score density histograms therefore provide an unsupervised classification of pixels according to their color (or spectral signature) regardless of their location in the original image. For example, two regions of the histogram shown in Figure 17 have been selected (see purple and green masks in Figure 17B) and all pixels having $\mathbf{t_1}$-$\mathbf{t_2}$ values falling within these selected regions are shown in the flame image of Figure 17C using a color overlay. Clearly, all pixels falling under the purple and green masks belong to the luminous and the non-luminous regions of the flame image, respectively.

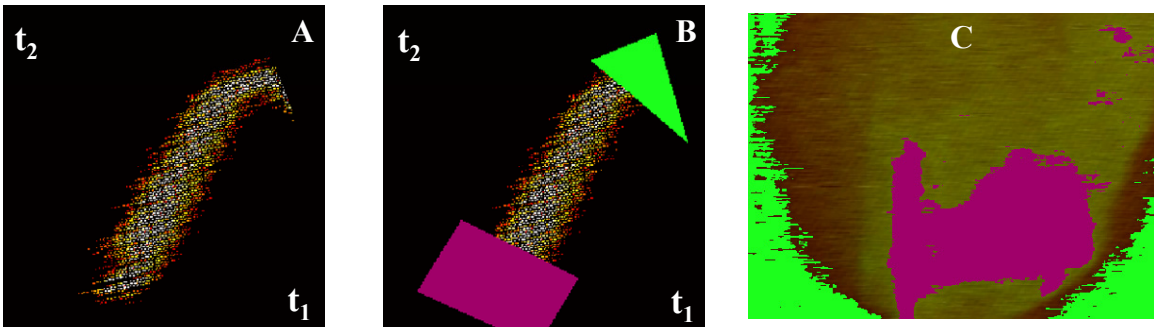

Figure 17: Multivariate Image Analysis: A) $\mathbf{t_1}$-$\mathbf{t_2}$ score density histogram of a combustion image, B) masks capturing pixels of similar colors, C) overlay of pixels falling under each mask and capturing the luminous (purple) and the non-luminous (green) regions of the bark combustion flames

Moreover, since pixels were classified individually according to their color only (i.e. the spatial inter-correlation of the pixels was not taken into account), another flame image having a similar color distribution but different flame shape, location or size, would have a very similar score density histogram. This is important in this case since we are mostly interested in heat released by the combustion (i.e. color distribution) rather than flame geometry. This shows a clear advantage for MIA over traditional imaging techniques, which would require segmenting the flame region in the image space every time and more computational efforts would result from this first step. Indeed, since the important information is in the flame color and not in the flame location in the original image, MIA provides a fast way to extract color features. Moreover, the common set of loadings and scaling range for the scores imply that the various regions of score plot represent the same color characteristics for every image. Therefore, the flame luminous region will always be at the same place in the score plot resulting in less computational efforts since this spot has to be found only once (or once every time the model is updated with new data). When the flame has been isolated for one image, it has also been isolated for all the other images. Furthermore, some aspects of flame geometry are also captured by MIA. For example, flame size is related to the number of pixels belonging to the luminous region.

4.3 MULTIVARIATE IMAGE REGRESSION

Once the images have been collected, a model between images and steam production need to be built. Some problems need to be solved first. This section will cover the problems that need to be overcome as well as their solutions. The structure of the various dynamic models tested in this project will then be described. The way to use those models to get a forecast on the steam production a few steps ahead (i.e. minutes) will also be covered.

4.3.1 REGRESSION DIFFICULTIES USING BARK COMBUSTION IMAGES

The main issues before building a regression model arise from 1) the very different dimensionality of the images versus steam and process data, 2) the noisy nature of bark images requiring some averaging of the image features, and 3) synchronization of the images and process data. These issues will be discussed in turn.

4.3.1.1 DIMENSIONALITY ISSUE

For each sampling time, a multivariate combustion image (i.e. array or matrix of data) and a scalar measurement of steam production are available. The dimensionality difference between those data is something to be resolved before using those matrices. Indeed, the image array has 3 wavelengths (corresponding to the red, green and blue intensities for each pixels) that represents each pixel of the spatial resolution of the original image which is 576×768. The dimensions of each matrix containing the information on each image are thus 576×768×3. Although the multivariate image is first transformed into a score plot density histogram (which is divided in 256×256 (65 536) bins), reducing the dimensions of the array from a 3-way array to a matrix, a difference with the steam production dimension (a scalar) still exists. Therefore, formulating a regression problem between multivariate images and some response variables requires solving this

dimensionality issue first. It mainly involves computing a row vector of relevant image features from the score density histograms. Various alternative methods for doing this were explored in details by Esbensen *et al*. (1992) and Yu and MacGregor (2003, 2004).

The relevant characteristics likely to be correlated with the steam production are extracted from the score plot density histogram. A vector of image features are typically computed from the score density histograms. This vector can be combined with some process data altogether in a regressor matrix $\mathbf{X_{MIR}}$ (n×p) and used in a regression model to predict a set of response variables, here steam production **Y** (n×1). The image features used in this research are similar to those proposed by Yu and MacGregor (2004). First, a mask is constructed for the segmentation of the luminous and the non-luminous regions of the score plot of each combustion image. While the pixels that fall inside the mask are considered to be part of the luminous region, the pixels of the score plot that do not fall inside the mask are considered as being part of the non-luminous region. The two masks shown in Figure 17 were drawn for illustration purpose only, but these are very close to those used for extracting the image features. The final mask is created in the score space and the related pixels falling into this mask are highlighted in the original image. This is done a few times until almost all the pixels of the luminous area are highlighted by the mask which is now considered precise enough to be selected as the one used for the analysis. This mask **M** is a squared matrix having the same dimensions than the score density histogram (i.e. 256×256). The selected regions of **M** have a value of 1 while the non selected regions have a value of 0. The 9 following features are then computed using the score density histogram of each combustion image, according to Yu and MacGregor (2004).

1. **<u>Luminous features</u>**

   1.1. <u>Luminous region area</u> (A) defined as the total amount of pixel having the same color then the bins under the defined mask where the numbers $h_{i,j}$ are the numbers of pixels falling at a particular position (i,j).

$$A = \sum_{i,j} h_{i,j} \qquad \forall (i,j), m_{i,j} = 1 \qquad (6)$$

1.2. Flame brightness (B) is obtained by first converting the score space into a grey-scale plane (**L**) instead of a color plane. The conversion coefficient vector helps convert the RGB signal into a corresponding intensity (Matlab Image Processing Toolbox, MathWorks, Natick, MA)

$$l_{i,j} = \left[\mathrm{R\ G\ B}\right]_{i,j} \begin{bmatrix} 0.299 \\ 0.587 \\ 0.114 \end{bmatrix} \qquad (7)$$

where $[\mathrm{R\ G\ B}]_{i,j}$ is computed using the following equation.

$$\left[\mathrm{R\ G\ B}\right]_{i,j} = t_1(i)\,p_1^T + t_2(j)\,p_2^T$$
$$\mathrm{t_1(i)} = \mathrm{i}\frac{(t_{1,\max} - t_{1,\min})}{255} + t_{1,\min} \quad \mathrm{t_2(j)} = \mathrm{j}\frac{(t_{2,\max} - t_{2,\min})}{255} + t_{2,\min} \qquad (8)$$

The sum of all the luminous intensity levels of each pixels falling inside the luminous area is then computed as the flame brightness.

$$B = \sum_{i,j} h_{i,j} l_{i,j} \qquad \forall (i,j), m_{i,j} = 1 \qquad (9)$$

1.3. Uniformity of flame brightness (U) defined as the standard deviation of the flame brightness throughout the luminous region.

$$U = \sqrt{\frac{\sum_{i,j} h_{i,j} l_{i,j}^2 - \sum_{i,j} h_{i,j} l_{i,j}}{\sum_{i,j} h_{i,j}}} = \sqrt{\frac{\sum_{i,j} h_{i,j} l_{i,j}^2 - B}{A}} \qquad \forall (i,j), m_{i,j} = 1 \quad (10)$$

1.4. Average brightness of the non luminous area (W) calculated as the brightness of the non-luminous area divided by the total amount of pixels of this region.

$$A = \frac{\sum_{i,j} h_{i,j} l_{i,j}}{\sum_{i,j} h_{i,j}} \qquad \forall (i,j), m_{i,j} = 0 \qquad (11)$$

2. Color features

2.1. Average color of the whole flame image along the first (s_{1m}) and the second principal components (s_{2m})

$$s_{1m} = \frac{\sum_{i,j} h_{i,j} i}{N} \;,\; s_{2m} = \frac{\sum_{i,j} h_{i,j} j}{N} \qquad i, j = 1, ..., 256 \quad (12)$$

where N is the total amount of pixels within the images.

2.2. Average color of the flame luminous region (s_{1f}, s_{2f}) which is the same features as the one described in 2.1 but specific to the luminous region.

$$s_{1f} = \frac{\sum_{i,j} h_{i,j} i}{A} \;,\; s_{2f} = \frac{\sum_{i,j} h_{i,j} j}{A} \qquad \forall (i,j), m_{i,j} = 1 \qquad (13)$$

2.3. Number of color of the luminous region (N_c) computed using the following equation. All the regions under the mask where at least one pixel is falling are considered as one color inside the luminous region.

$$N_c = \sum_{i,j} 1 \qquad \forall (i,j), h_{i,j} m_{i,j} \neq 0 \qquad (14)$$

Each one of those 9 features are computed for each combustion image and stored in a matrix.

4.3.1.2 UNSTEADY BARK FEED ISSUE

As it has been already mentioned, bark flow is highly variable compared to liquids or gases. Indeed, a flow delivered by a pump is continuous and could be precisely measured if required using flow meters. Unlike liquids and gases, solids are delivered using conveyors and screw feeders. Moreover, solid material particles do not disperse themselves making sure all the space available is used evenly. Solid particles stay where they were put which results in an unequal distribution of the particles in the available space. This brings an unsteady combustible flow to the boiler. Indeed, conveyors, where a non homogeneous distribution of bark lays, bring the bark to the screw feeders that throw it in a discharge from where it is then blown on a grid, resulting in a very non homogeneous "lumpy" bark feed flow rate to the boiler.

This creates rapid movements of the flame appearance as was shown in Figure 6. The temperature inside the combustion chamber, however, does not change as quickly as the flame, and part of the high frequency variations in the combustion does not affect steam production due to the large thermal inertia. To filter out some of these high frequency variations, the flame characteristics extracted from small number of consecutive images were averaged. The number of images taken for the averaging will be discussed in the results chapter. The corresponding process conditions were also averaged over the same time periods to maintain information consistency with respect to the images.

4.3.1.3 TIME SYNCHRONIZATION

Since the link needs to be found between $\mathbf{X}_{\mathbf{MIR}}$ (containing information on images or on both images and process data) and **Y** which is the steam production, the time of the rows in the **X** matrix have to be synchronized with those of **Y**. Therefore, images, process data and steam production were all synchronized prior to constructing these matrices. The process data and steam production were all acquired using the same system and are already synchronized. Images, however, were collected using a different system, had to be synchronized with data. Therefore, the image taken has the one synchronized with the data

as well as with the steam production was chosen to be the one which acquisition time was either the same or lower than the data's acquisition time. Since the acquisition time were most of the time slightly different, the image acquisition time was more often chosen has the time that gave the smallest difference between the process data and the image acquisition time as long as that difference was not higher than 4 seconds. Therefore, the image was acquired first which is consistent with the fact that if a disturbance is introduced in the boiler, the flame appearance should change before the steam production. Indeed, considering the observation of the flame appearance gives an idea on the following steam production, images have to come first.

4.3.1.4 REGRESSION MATRICES CONSTRUCTION SUMMARY

Many problems have to be solved before having matrices ready for modelling. Those issues have been presented separately and this section will give a summary of all the different steps in chronological order as presented in Figure 18. First, images and process data were synchronized. To resolve the dimensionality issue, image features were computed from each image and a matrix $\mathbf{X_{MIR}}$ was then built using image features and process data. Since the bark flow is lumpy which created too much variability in the image features, the matrix **X(t)** was averaged using m images which yields $\bar{\mathbf{X}}(\mathbf{t})$. This averaged matrix is now ready for the modelling.

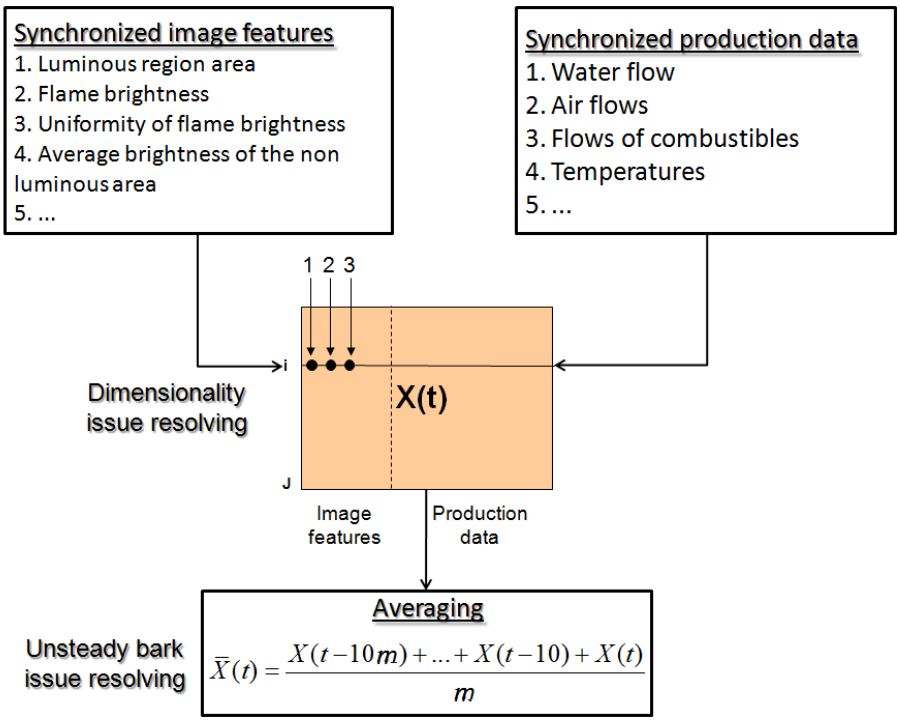

Figure 18: Matrix construction summary

4.3.2 MULTIVARIATE IMAGE REGRESSION AND DYNAMIC MODELLING

The fundamental underlying relationship between two variables can be found using the appropriate balance equations. Taking into account every factor that can have an impact on the system formed by those two variables will result into a series of equations leading to the exact relation between the two considered variables. However, this method is easier to apply in theory where the amount of variables involved are more limited than in practice. Using balance equations to find the dynamic of a real process is harder since a large number of variables often needs to be considered and some phenomena may be unknown or difficult to quantify, such as energy lost in the process, or simply because it is not measured. For those reasons, empirical dynamic model identification approaches are often used in the process industries.

The identification of a process model is usually done using iterative method. An initial model structure is first chosen, either the simplest one or based on *a priori*

knowledge of process behaviour. Then, some process data are used to find the model parameters and the other part of the process data available is used to validate the model. If the results are good on both sets of data, then the iterative process ends. In this research, the predictive ability in validation, measured by the multiple correlation coefficient (R^2), was taken as the criteria to maximise. On the other hand, if the model performance is not satisfactory, the model structure is modified and the parameters have to be re-estimated and validated until an acceptable model is obtained. Therefore, model validation is very important since it confirms the capacity of the model to perform on new sets of data. A model could have a very good fit on data used to find the model coefficients because those coefficients are too numerous and overfitting of those particular data occurs. However, the model fitness in validation may be poor since those new data have not been used to build the model. It is therefore important to select a model structure that allows a good fit of the data, but general enough to maintain fitness on data sets not used to build the model.

In this research, two non-parsimonious dynamic model structures were preliminary studied for relating images and process data to steam flow rate. Finite impulse response coefficients were obtained using the following two structures: the output error (OE) model and the exogeneous variable auto-regressive (ARX) model (Ljung, 1999). The model parameters were estimated using PLS regression in both cases (Duchesne and MacGregor, 2001). Both structures provided good identification results based on the available data. Those two models will be presented in this section for a single-input single-output first, and then for the multi-input case. The results will be discussed in the results chapter.

4.3.2.1 OUTPUT ERROR (OE) MODEL

The first model tested in this research is a very simple one. The model used is based on the discrete structure called output error. The modelled output y is a function of the current and past values of the input variables x. The number of past values is chosen according to the process settling time. The general structure is shown using equation (15).

$$y(k) = \frac{B(z^{-1})z^{-d}}{F(z^{-1})}x(k) + e(k) \qquad (15)$$

where the y(k), x(k) and e(k) correspond respectively to the steam production, the inputs containing whatever information required and the noise taken at each sampling time noted here k. The $B(z^{-1}) = b_1(z^{-1})+b_2(z^{-2})+...+b_{nb}(z^{-nb})$ and $F(z^{-1}) = f_1(z^{-1})+f_2(z^{-2})+...+f_{nf}(z^{-nf})$ symbols are polynomials in the backwards shift operator (z^{-1}) whereas z^{-d} is the process dead-time. Note that in this work, the coefficients resulting from the long division of the numerator and denominator polynomials are estimated rather than the coefficient of the polynomial themselves (i.e. $y(k) = B^*(z^{-1})z^{-d}x(k) + e(k)$). The number of past input variables to use (also called lags) need to be sufficient to cover the process settling time. Moreover, when a prediction is needed, the error term is set to 0. Therefore, if we consider no dead time, the prediction of the y is shown in the equation (16).

$$\hat{y}(k+1 \mid k) = b_1 x(k) + b_2 x(k-1) + ... + b_n x(k-n+1) \qquad (16)$$

where all the b's are the coefficient of the regression estimated using PLS regression. The latter is an alternative to Multi-Linear Least-Squares Regression (MLR) improving the estimation of impulse response coefficients (i.e. the b's) in presence of collinearity (Duchesne and MacGregor, 2001). Since the columns of the regressor matrix are past lags of the input variables, collinearity is generally present even with an orthogonal design of experiments. The amount of past x's used will be determined in the results chapter. A prediction for additional steps ahead can be obtained using this structure. This prediction is computed using the equation (17) where n is the number of past input lags used in the model and where d, the dead time, is considered to be 0.

$$\hat{y}(k+p \mid k) = b_1 x(k+p) + b_2 x(k+p-1) + \ldots + b_n x(k-n+1) \qquad (17)$$

Since the information is available until sampling time k, all the inputs beyond this time need to be estimated using either some assumptions about future variations or using mathematical equations describing these if available. The results using this structure will be discussed in the next section.

4.3.2.2 EXOGENEOUS VARIABLE AUTO-REGRESSIVE (ARX) MODEL

The second model tried is based on the controlled auto-regressive structure. Once again, this structure is discrete and can be used with the sampled data available at Irving P&P. This model is similar to the OE structure, but allows for a colored noise model to be identified, that is the output y is computed as a function of past values of y and x. The number of past input and output lags to use in the model can be different from one another, but both of them are selected through the same iterative procedure. The general parsimonious ARX model structure is shown in equation (18).

$$A(z^{-1})y(k) = B(z^{-1})z^{-d}x(k) + e(k) \qquad (18)$$

The difference with the OE structure is that polynomial $A(z^{-1}) = a_1(z^{-1})+a_2(z^{-2})+\ldots+a_{na}(z^{-na})$ becomes a common auto-regressive component for both the process transfer function as well as the noise model. Non-parsimonious identification of impulse response coefficients for the ARX structure consists of building regression matrices for which any row expresses the current output y(k) as a function of past values of x and y. Once again, when a prediction is needed, the error term e(k) is set to 0. Therefore, the prediction, considering no dead time, of the y is shown on the equation (19).

$$\hat{y}(k+1\,|\,k) = -a_1 y(k) - \ldots - a_{ny} y(k-ny+1) + b_1 x(k) + \ldots + b_{nx} x(k-nx+1) \qquad (19)$$

where nx and ny correspond to the number of past input and output lags, respectively. The a and b coefficients are again estimated using PLS regression. The numbers nx and ny used will be determined in the results chapter. This structure also allows to compute prediction for more steps ahead. This prediction ŷ will be a function of the past x's and y's as well as the estimated x's until x(k+p-1).

4.3.2.3 STRUCTURE FOR MULTIPLE INPUTS

The OE and ARX structures were presented for a single input variable only. Some modifications need to be made when using multiple inputs as it is the case here. Since the matrix **X** has more than one column (i.e. 9 image features and 28 process data), the equation becomes the following:

$$\begin{aligned}\hat{y}(k+1\,|\,k) = {} & -a_1 y(k) - \ldots - a_{ny} y(k-ny+1) + b_{1,1} X_1(k) + b_{1,2} X_2(k) + \ldots \\ & + b_{1,37} X_{37}(k) + b_{nx,1} X_1(k-nx+1) + \ldots + b_{nx,37} X_{37}(k-nx+1) \\ = {} & \sum_{i=1}^{ny} -a_i y(k-i+1) + \sum_{i=1}^{nx} \sum_{j=1}^{37} b_{i,j} X_j(k-i+1)\end{aligned} \qquad (20)$$

where $a_1,\ldots,a_{ny}$ are equal to 0 when using the OE structure.

Note that the dead-time (i.e. z^{-d}) was considered to be the same for all inputs since it was smaller than the sampling rate used to build the dynamic models.

4.3.2.4 BUILDING MATRICES SUMMARY

The PLS regression technique is used in this project. This linear regression technique involves a regressor matrix $\mathbf{X_{MIR}}$, containing the past input/output lags of each variable, and a response matrix Y containing the variable to be predicted. How these matrices were obtained is shown in Figure 19.

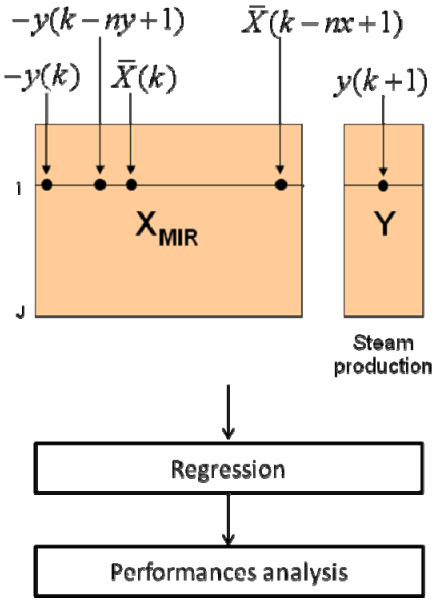

Figure 19: Building the $\mathbf{X_{MIR}}$ and **Y** matrix summary

Using those matrices, the regression coefficients are estimated using PLS and the performances of the resulting model are analyzed.

4.3.3 STEAM PRODUCTION PREDICTION

The main objective of this research is to be able to predict steam production so an action can be taken to stabilize the steam production before it is affected by the different disturbances. Different approaches can be used to get predictions a few steps ahead. One of them is to develop a model as described in the previous sections and use the structure of this model as well as the coefficients found to make a prediction of the output a few steps

ahead. This method requires the development of mathematical equations based on the original model found.

In this work, it was decided to construct a lagged structure matrix $\mathbf{X_{MIR}}$ with respect to $\mathbf{Y}$ and the coefficients of this new model were then computed with PLS. The lags were selected by trial and error, to maximise predictive ability.

CHAPTER 5 RESULTS AND DISCUSSION

This section presents and discusses all the results obtained in this work. First, the dynamics between each combustible and steam production will be shown. This will serve as preliminary tests to help understand the behavior of the process and determine the structure of the final model to be found. The reader will also see that simple ARX models can be used to model steam production by both fossil fuels. However, steam produced using bark is harder to model using only a simple ARX structure and images could lead to a better model. Therefore, models based on each camera will be shown and the performance of each one will be analyzed and compared to see whether both cameras or only one of them help predict steam production by bark. It will be shown that process data give enough information to model the natural gas and oil behavior but some important pieces of information are missing when only bark is used. It will be shown that images complement process data by providing supplementary information. A strategy using simple ARX models for fossil fuels contributions on steam production and a more complex model using images and data for the contribution of bark on steam production will be explained. Finally, the final model will be shown and its performance will be discussed.

5.1 STEAM PRODUCTION DYNAMICS FOR DIFFERENT FUELS

Identifying the relationship between the manipulated variable and the variable to be controlled involves trying different models. Among all the data available, a certain amount should be kept aside from the identification procedure and used as testing data to see how well the model performs on data that were not used to identify it, a procedure called validation. The data chosen as being part of the validation set needs to have values inside the range of the values used to build the model. The model's parameters can also be

compared to the information already available on the process and it is important that the model agrees with measurements. Any model can be tried but some models could be better for some process. To know what model to try first, all the information already available on the process to be identified needs to be used.

In order to identify the relationship between a manipulated variable and a variable to be controlled, significant changes are required on the manipulated variable according to some designed experiments minimizing the correlation among the manipulated input variables and unforeseen events (i.e. disturbances). A change is considered significant when a valuable change is observed in the variable to be controlled due to the variation in the manipulated variable and not only due to the presence of noise in the process. To identify the relationship between the flow rate of each fossil fuel and steam production, two steps changes have been made on the natural gas flow and the oil flow (manipulated variables) and even if bark was also burnt during this period of time, its flow was as constant as possible. The values of the steps were chosen according to normal values used at the mill. However, some disturbances such as changing bark properties were unavoidable and this explains some of the noise observed in the steam production flow rate (variable to be controlled). As shown in Figure 20, the amount of noise introduced by the use of bark seems more important during the oil test (B) than during the natural gas test (A), since the oil step change released less energy. However, the results are still reasonable.

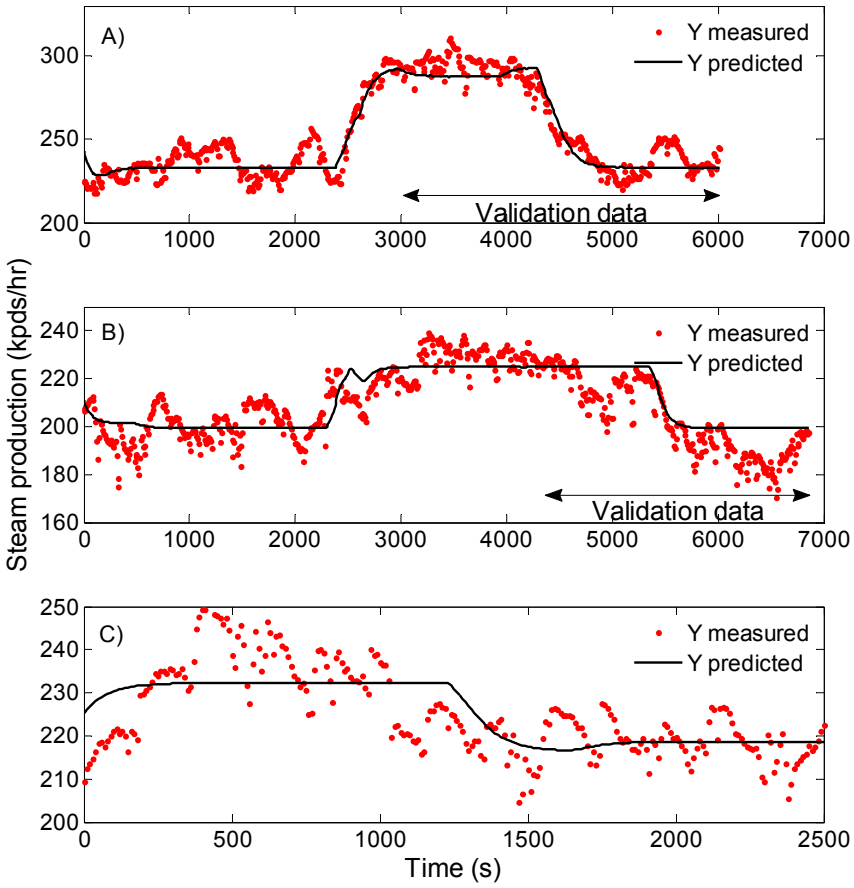

Figure 20: Single-input single-output dynamic identification results between the flow rates of natural gas (A), oil (B), and bark (C) and steam production using ARX models

The dynamics between bark flow rate and steam production is harder to capture using bark flow rate as the single input. When using only bark, as it was the case during the period of time shown in Figure 20 C, the fluctuations in steam production are greater compared to when fossil fuels were used. No validation data was kept for the bark dynamic model since there was an insufficient amount of data where the effect of bark flow rate only could be isolated from other changes and disturbances. Fossil fuels were used at a few occasions during this period and changes in bark properties such as heat of combustion were observed. However, the main objective of this section of the thesis is to get a rough idea of the dynamics of the three fuels and, therefore, validation was not so critical at this

point. Note that the initial variations in model predictions are due to initial conditions, which were determined based on a much larger dataset. Again, the models identified in this section were built with the only goal of obtaining a rough idea about the dynamics of the fuels.

The initial models used to build a dynamic relationship between steam production and the various fuels were based on a first order exogeneous auto-regressive (ARX) model structure. These models were identified in their parsimonious form using the Matlab Systems Identification Toolbox. A few other model structures were tried but steam production depends on the past operating conditions which means that if the boiler is cold, the steam produced by a certain flow will be different from the steam that could be produced if the boiler has been on for a long time, the ARX model was found to be the better one.

Using an ARX model to identify the dynamic between both fossil fuels flow rate and the steam production gives good results. Indeed, looking at Figure 20, it can be seen that the predicted steam production calculated using the ARX model is representing the real steam production properly. The performance of both models is shown in Table 4. Considering the higher variability of the steam produced when using oil due to the fact that a higher percent of steam was produced using bark, poorer results were expected for oil identification.

The time constants (τ) and gains for the first order transfer functions identified for each test are provided in Table 4. Considering the uncertainty range, these were found to vary between 1.5 and 2.5 minutes. The percent explained variance (or multiple correlation coefficients R^2 traditionally used in statistics) in fit and on validation data are also provided in this table. Time delays were very small for each fuel tested and were therefore neglected. As seen in Table 4, the time constant of each test is about the same.

Table 4: Natural gas, oil and bark identifications results

Fuel tested	Fit R^2 (%)	Validation R^2 (%)	τ (min)	Gain
Natural gas	82.9	91.2	1.80±0.45	0.9±0.2 kpds/KSCF
Oil	71.2	78.0	1.35±0.25	13.1±3.0 kpds/ kpds
Bark	45.8	-	1.30±0.35	6.5±2.2 kpds/tons

Those dynamics between each combustible will help having an idea on the best way to build the final model. For example, knowing that the time constants were found to be between 1 minute and 2 minutes and 30 seconds and that to cover the dynamic at least 5 data samples are needed during one time constant, one data every 12 to 30 seconds is needed. Those sampling times were tried and 30 seconds was selected. Shorter sampling times did not increase the predictive ability of the models. Moreover, a sampling time of 30 seconds simplifies the synchronization of images with the data sampled at a rate of 10 seconds, which is a factor of 30 seconds.

The results shown in this section indicate that simple models (i.e using flow rates only) could be used to describe fossil fuel dynamics contribution to steam production. However, when bark is used, alone or in combination with fossil fuels, a more complex model is required since the bark flow rate alone is insufficient to determine the amount of steam produced by the bark. Indeed, the value of the flow is inaccurate and does not give information on the changing properties such as the carbon content. Therefore, another sensor will have to be used to extract this information. In this research, images are proposed as a new ***area*** sensor to have indirectly this information.

5.2 DIFFERENCE BETWEEN BOTH CAMERAS

As mentioned, Irving P&P's boiler is equipped with two high temperature cameras each providing a different field of view of the combustion taking place within the boiler. As shown in Figure 7, the camera located on the third floor of the boiler gives a perpendicular view of the grate while the camera located on fifth floor looks down at the grate where the bark is burning. Some tests were made to investigate the predictive ability obtained with combustion images gathered using both cameras. Periods of time where **only bark** was burning were taken to build models of steam production by bark. This includes the bark moisture and flow rate tests discussed previously. In this section, only images were considered in the $\mathbf{X}_{\mathbf{MIR}}$ matrix (no process data was used nor past outputs). The image features used in $\mathbf{X}_{\mathbf{MIR}}$ were averaged using the current and the past two images. Ten past lags of those averaged images taken at 30 seconds intervals were used to build the model. That is, a history of 300 seconds (5 minutes) was used in the models for predicting current steam production rate. The modeling results using both cameras are shown in Figure 21.

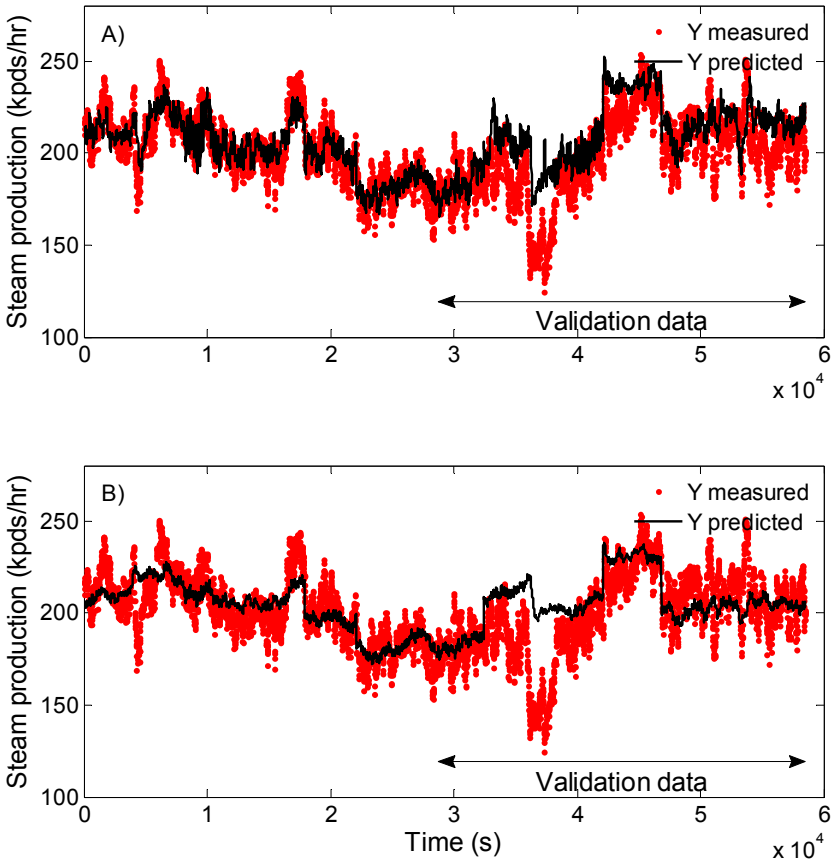

Figure 21: Dynamic model identification results for bark tests based on each camera separately (one step ahead predictions). A) 5th floor camera; B) 3rd floor camera. Dots correspond to measured steam production and solid lines to model predictions.

The percent explained variance of steam production by the models (i.e. R^2), both in fit and validation and for each camera, are presented in Table 5. The number of latent variable (LV) chosen for each model is also presented. The number of LVs was selected using a leave-group-out cross-validation procedure (Eriksson et al., 2006). The 5th floor camera provided the best results, with about 70% of variations in steam production (i.e. training set) explained by the images only, and slightly above 50% on the validation set. Some discrepancies between the measured and predicted steam flow are observed in Figure 21A, particularly on the validation data. Some efforts to improve these results, by adding information into the model, will be shown in the next sections. However, when compared to the results obtained with the 3rd floor camera, it is clear that the images provided by the 5th floor camera are more informative, most likely due to a better field of view of the bark

combustion zone. Nevertheless, the 3rd floor camera captures most of the trends and supports the assumption that combustion images will be very useful in capturing disturbances in bark combustion. The 3rd floor camera was not investigated further and only the 5th floor camera images will be used in the remainder of this thesis. The camera on the 3rd floor is less information rich, and provides redundant information compared to the 5th floor camera. A model using both cameras was also tried without any improvement of the predictive ability and the decision was taken to use the 5th floor camera only.

Table 5: Comparison between camera 5 and camera 3 performances

Camera	Number of LV	Fit R^2 (%)	Validation R^2 (%)
5th floor	7	69.9	53.5
3rd floor	7	59.4	24.9

Some improvements have been brought further to see if the model built using the 5th floor camera could be improved. The results presented in this section covered only those periods when bark was burned. It will be shown that, although images are essential for capturing the contribution of bark for steam production, these images provide little information when fossils are burned. Adding process data to the models will prove to be useful for situations when fossil fuels only, or in combination with bark are used for producing steam.

5.3 DIFFERENCE IN THE INFORMATION PROVIDED BY IMAGES AND PROCESS DATA

Most of the time, models are built using information provided by traditional sensors, such as temperature, flow, pressure and concentration measurements. In this research,

predicting steam production will require using both process data and images. This section will establish the difference between the information given by both sources.

To illustrate the difference between the information given by process data and images, a model has been build using only process data. This will show that some of the crucial information, such as variations in carbon content and other biomass properties, are not carried by process data. Another example is the inaccurate biomass flow rate measurements. A model was built using only process data in the $\mathbf{X}_{\mathbf{MIR}}$ matrix. The variables used in this model are those described in Chapter 3. The lag structure used for data was similar to the one used for images, that is ten past lags for each variable, taken at intervals of 30 seconds, were used to build the model. The information on each process data at time t-30, t-60 and so on until t-300 seconds were thus collected into the $\mathbf{X}_{\mathbf{MIR}}$ matrix to build the model. Periods of time where the three main combustibles were burned are used to build the model. The results are shown in Figure 22.

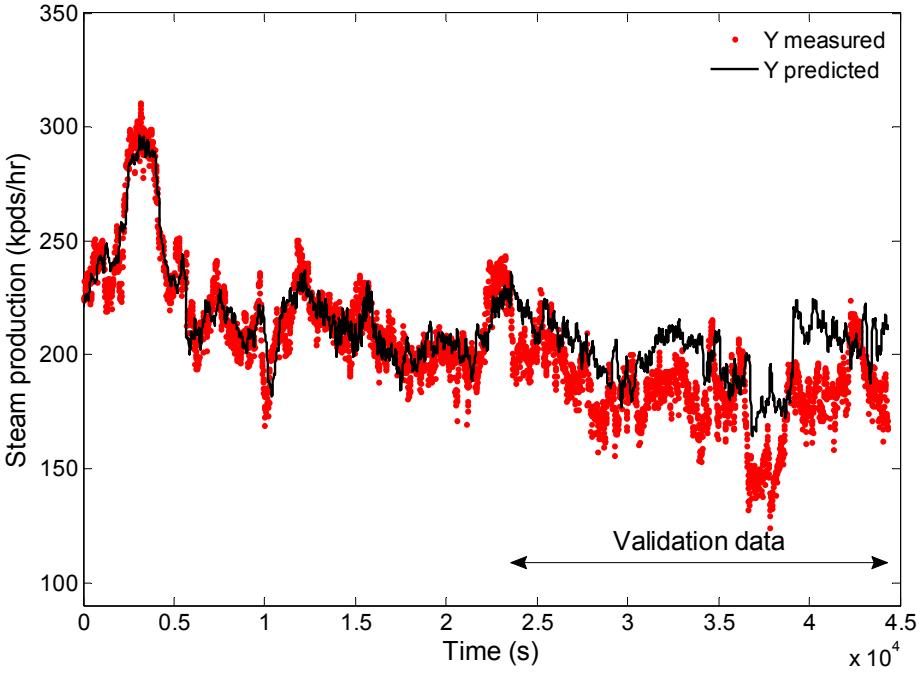

Figure 22: One-step ahead prediction model built using process data only (steam produced using natural gas or oil and/or bark)

As observed in this figure, the results of the model built using 7 latent variables are fairly good on the training data set, where a R^2 in fit of 83.2% was achieved. The model is then applied to a validation set and a much lower R^2 of 28.0% was obtained. Some on the variation were not captured by this model. Indeed, the validation data are not represented by the model properly. The disturbances brought by the bark flow or the changing moisture content are not captured by the model because those values are either inaccurate or completely missing in the datasheet, and therefore, this model seems to lack robustness.

On the other hand, some information was also missing when the model was built using only images, as was observed in Figure 21 (on a different data set including only bark, no fossil fuels). Information not provided by images but available from process data includes the effect of variables such as temperature of the feed water, air flows, etc. Better results could be obtained combining images and process data.

Therefore, the information provided by images and process data is different but complementary. Process data give information on the temperature of the different flows, the values of those flows as well as any other useful value that can be quantified such as information on temperatures or water feeding rate. On the other hand, those data do not provide precise information on bark flow, and very little or no information about bark moisture and other properties such as combustion energy. As mentioned, bark flow is calculated using the screw feeders' speed which is an approximation of the real flow rate. To illustrate the inaccurate calculated bark flow, a situation where the calculated bark flow rate is steady but where the steam produced varies widely is presented in Figure 23A. A comparison is made on Figure 23B with a critical situation where the steam production had to be steady. In critical situations, operators stabilize the steam production by lowering the amount of bark fed inside the boiler. The steam is therefore produced using fossil fuels. In this particular situation, natural gas was used.

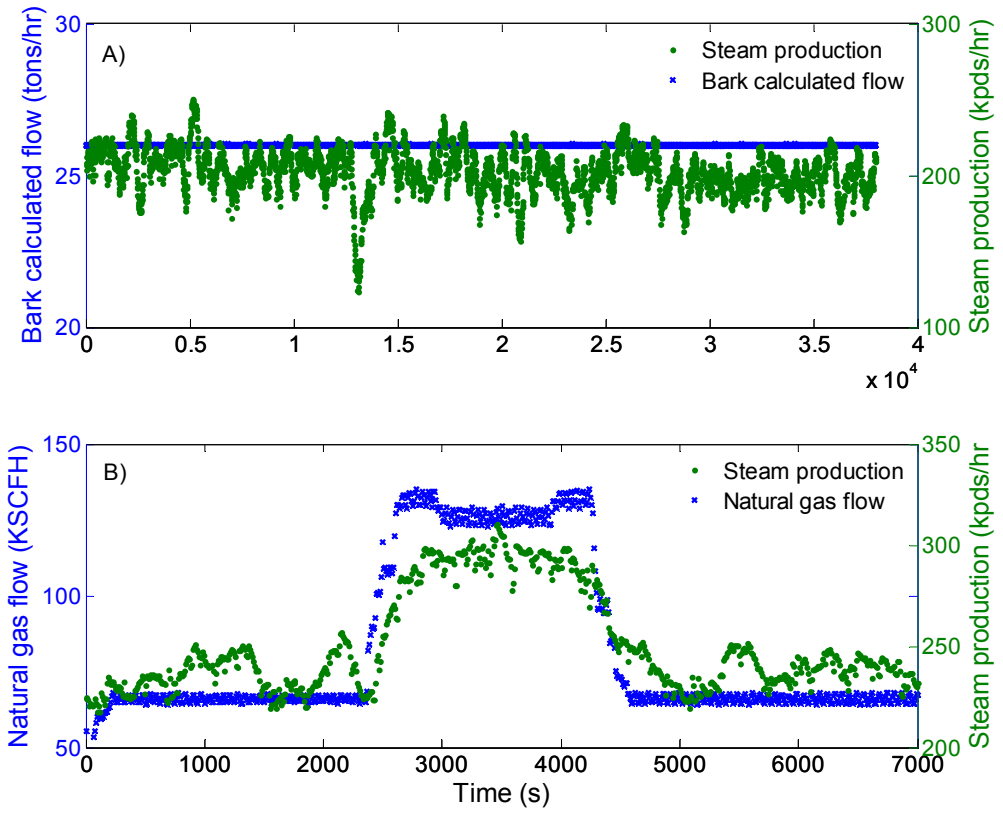

Figure 23: Comparison of steam production variability when using bark only (A) and a combination of bark and natural gas (B)

As it is shown, the variability of steam production is lower when fossils are used compared with the situation where only bark is fed. The more the percent of steam is produced using fossil fuel and the lower the steam variability is. The variability can be quantified using the standard deviation. In the case shown on Figure 23A), the standard deviation is 20.43 which is a value higher than in the case where natural gas was used since the standard deviation of the steam production shown on Figure 23B) is 5.99 which is 4 times smaller. Therefore, the bark calculated flow is inaccurate since the flow varies much more leading to important variations in steam production.

In addition to have inaccurate values, process data also have missing information. Indeed, the heat of combustion of the bark is an unknown value that changes according to the amount of carbon inside. The bark flow and the amount of energy it contains are both important data. This information is given within the images. Indeed, it was shown that the flame appearance is highly correlated with the heat it releases (Yu and MacGregor, 2004; Szatvanyi and *al*., 2006). Thus, using images allow this important information to be available. Whenever an unmeasured disturbance is introduced by bark, the image will provide information on those variations. Therefore, process data gives information on fossil fuel and images and process data are needed for bark contribution on steam production.

5.4 IMAGES FOR BARK CONTRIBUTION ONLY

Irving P&P boiler is fed using three different combustibles. Therefore, the total steam production is the sum of the contribution of each of them, thus requiring incorporating more information into an overall dynamic model for predicting total steam production. However, increasing the number of input and output variables generally leads to a more complex model and a greater level of difficulty to identify. The easiest way to build a model is always to try simplest structure first. Then add complexity as required to obtain a satisfactory model. To build a dynamic model for steam production when more than one fuel is used, it was decided to identify each of the three models separately after decoupling the contribution of each fuel.

It was shown that single-input single output first order ARX models were sufficient for describing the dynamics between the fossil fuel flow rates (inputs) and steam production. Hence, these two models (one for oil and the other for natural gas) could allow computing the contribution of fossil fuels to steam production. Unlike bark, the stability of fossil fuels makes them easily described by simple first order models since those two combustibles are well understood. Indeed, their properties such as the flow or the heat of combustion are known accurately allowing the estimation of the heat released by each of them. However, a very different situation occurs when using bark since its flow rate and

heat of combustion are either inaccurate or unavailable and need to be found in a different way. For this reason, first order ARX models could be used to compute the contributions of the fossil fuels on steam production and a separate, more complex model could be build using images and process data to get the contribution of bark. The use of these three models in parallel would naturally lead to predict the total amount of steam produced.

Moreover, another argument in favor of this approach is the fact that the 5th floor camera is located underneath the fossil fuel burners and therefore, should only provide information about bark contribution. Hence, it was decided to use the ARX models for the fossil fuels to predict their contribution to total steam production. These contributions were then subtracted from total steam produced which should, in theory, provide a good estimate of bark contribution. The more complex model for bark (using data and images as discussed in previous sections) could be developed using this estimate, therefore decoupling the identification of the models for each fuel, even if they were used simultaneously. This strategy was tested with little success. The combustion images shown in Figure 24 will help understand why the results were not conclusive.

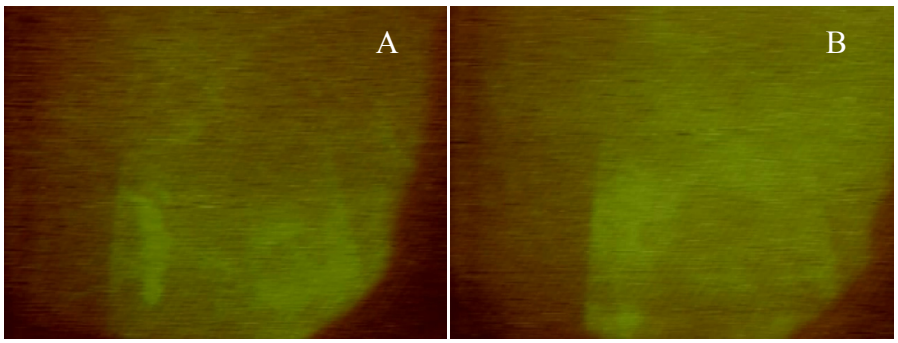

Figure 24: Comparison of bark combustion images without (A) and with (B) natural gas burned simultaneously (images were taken by the 5th floor camera)

Shown in Figure 24 are two situations where the bark flow feed rate and recipe (i.e. ratios of stockpiles) is similar. The image on the left shows combustion of bark alone. The image on the right represents a situation where the same amount of bark is burned but

where natural gas is simultaneously burned using two of the four auxiliary burners. As it can be seen, the resulting images are different. When using both fuels, the flame (Figure 24B) is a lot brighter especially on the top right corner. This is due to the natural gas flame. Even if the 5th floor camera is located underneath the auxiliary burners (used for oil and natural gas), the distance between the burners and the camera is not sufficient to allow bark combustion alone to be visualized in the images. Therefore, the light created by the combustion of the natural gas illuminates the boiler's walls. As those walls can be seen within the field of view of the 5th floor camera, the resulting images contain some information on fossil fuels as well. Since this information cannot be removed from the image color intensities, the images contain redundant information about the effect of fossil fuels and this leads to overestimate the contribution of bark to total steam production. The information provided by the image about fossil fuel combustion are not sufficient for building a single model based solely on images for predicting steam production under any combinations of the three main fuels. However, it is important enough to justify building a single, more complex model, in which the contribution of oil, natural gas and bark are identified simultaneously. Therefore, the final model is a complete model including the effect of all combinations of bark, natural gas and oil used in practice on steam production. Such a model relies on both process data and images of the combustion chamber.

5.5 FINAL MODEL

The final model was built using a matrix **X** containing the 9 images features as well as 28 process variables (i.e. before including lags). Those variables were found in a preliminary analysis to be important for predicting steam production. Data were taken for each situation when combustion was made using either natural gas and bark, oil and bark or only bark but with different amount of moisture in it. Therefore, a single model is built for all three fuel combinations used in practice. Building a separate model for each situation only led to marginal improvements.

The one step ahead prediction results for total steam production are shown in Figure 25, for both the output error (OE) structure (Figure 25A) and the ARX structure (Figure 25B). As it was the case for the previous dynamic models, ten lags of matrix **X** were used to build the $\mathbf{X_{MIR}}$ matrix. The lag intervals where again selected to be every 30 seconds and each data inside the matrix **X** was an average of the current data and the 2 previous data acquired at time t-10 and t-20 seconds. Models with increasing number of output lags were built for the ARX structure, but the one shown in Figure 25B has only one past lag of the output y. The choice of one past lag for the output will be discussed later in this section.

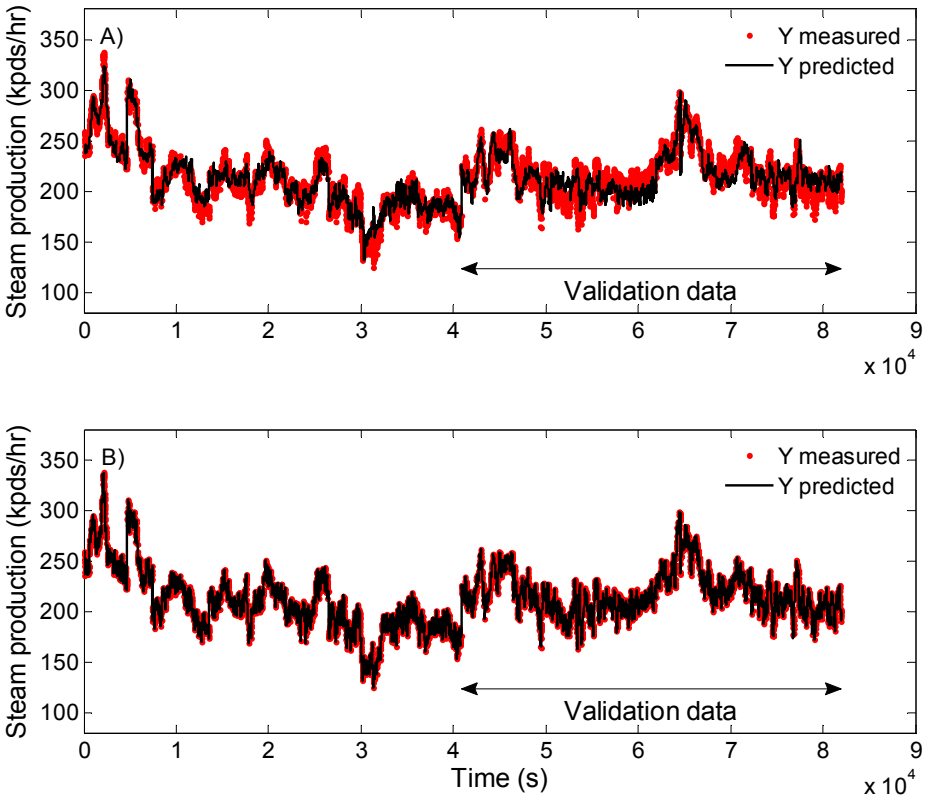

Figure 25: Results of simultaneous identification of dynamic models for all three fuels using the OE structure (A) and the ARX structure (B). One step ahead predictions are shown. The ARX model uses one past output lag.

The results of those models are promising as shown in Table 6. The results for the ARX model are a lot better since the autoregressive noise model embedded within its structure captures some of the structure information unexplained by images and process data. This model heavily relies on the output lag and it becomes easy to predict the steam production at time t when the previous steam production is known.

Table 6: Comparison of the performance of the final models

Structure	Number of LV	Fit R^2 (%)	Validation R^2 (%)
OE	9	89.0	55.6
ARX	30	99.9	99.8

The choice of having only one output lag is arbitrary. However, adding more output lags has been tried. The ARX model using ten past lags for the inputs but an increasing number of output lags were built and their predictive ability in fit and validation are shown in Figure 26.

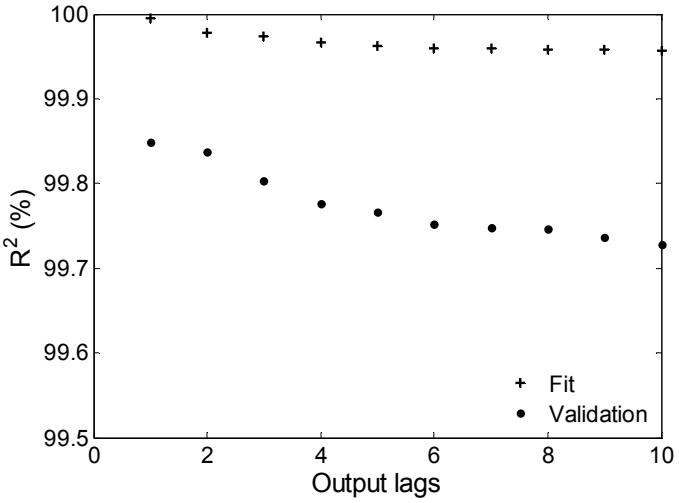

Figure 26: Prediction ability (R^2) of the various ARX models with increasing number of output lags

Based on Figure 26, the ARX model predictions remain fairly constant after adding one past output lag. Therefore adding more lags only increases the model complexity but only results in marginal improvements in the results. Only one output lag will be used for the rest of this thesis.

To assess the forecast ability of those models, the OE and ARX model structures were applied for predicting steam production up to p future time steps (p=1,2,…,10) using the images and data available up to current time. Each step ahead corresponds to a prediction 30 seconds in the future (i.e. 5 steps ahead = 2.5 minutes in the future). To do this, matrices were built using the same structure than the one shown previously. The matrix $\mathbf{X_{MIR}}$ was built based on the recursive structures. Since the predicted steam production is made for p steps ahead in the future, as shown by Equation 17 for the OE structure (an equivalent expression for the ARX model is also available), assumptions on how future values of the input data (i.e. images and process data) would behave in the future is required. In this work it was assumed that future data and images would remain constant and equal to last measurement collected on them (i.e. at current time). Therefore, if poor bark is fed to the boiler, the operator will have an early warning of what could happen in p steps if the bark quality does not improve. This approach requires re-estimating model parameters using, once again, PLS regression. The predictive ability of both model structures (OE and ARX) in fit and validation are presented in Figure 27 for p=1,2,…10 steps ahead.

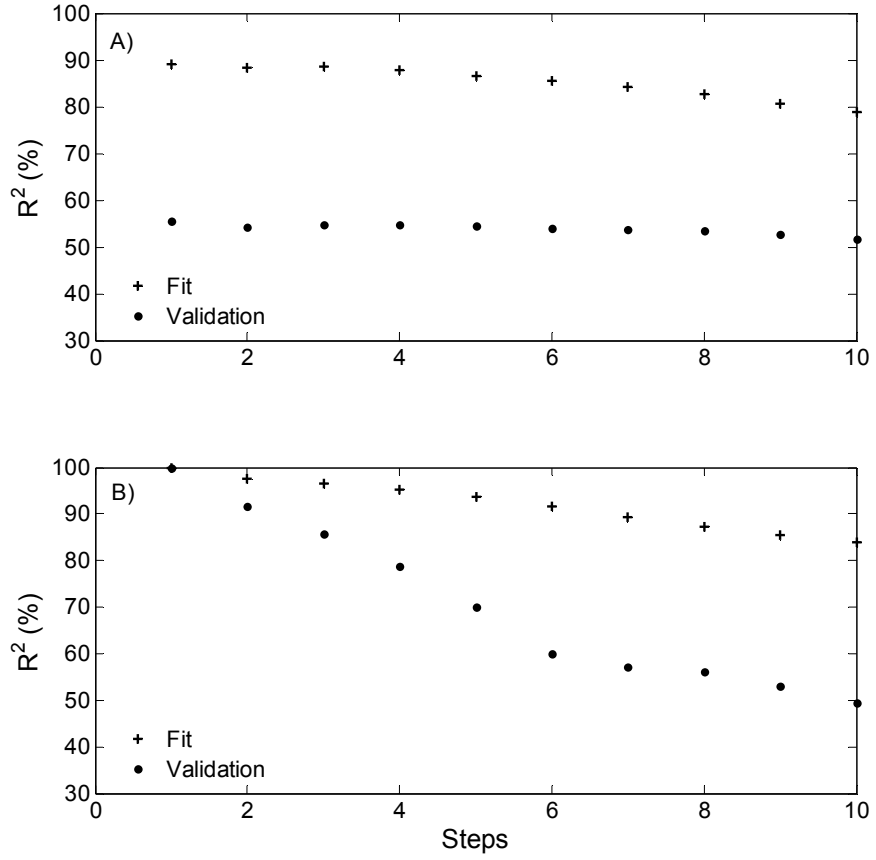

Figure 27: Model predictive ability for several steps ahead forecast of steam production using the OE structure (A) or the ARX structure (B)

It can be observed that prediction performance reduces progressively as prediction horizon increases. This situation was expected since if new set-point changes or disturbances are introduced in the process in between current time t and time t+p (p steps ahead forecast), this information is not considered in the model. Variations occurring during this period may not be fully explained by the images and data collected up to current time. After discussing with Irving P&P's engineers, it was found that, in practice, if their operators were warned at least 2 minutes in advance that steam production upsets will occur due to changes in bark properties, this would leave a sufficient amount of time to react proactively. However, the performance improvement obtained by using these forecasts into an advanced control scheme needs to be quantified. The 5 steps ahead

forecast corresponds to this situation in which steam production is predicted 2.5 minutes in the future. The fit and validation results for both OE and ARX structures are acceptable up to 5 steps ahead. Model predictions for p=5 are provided in Figure 28. Their R^2 statistics are also given in Table 7.

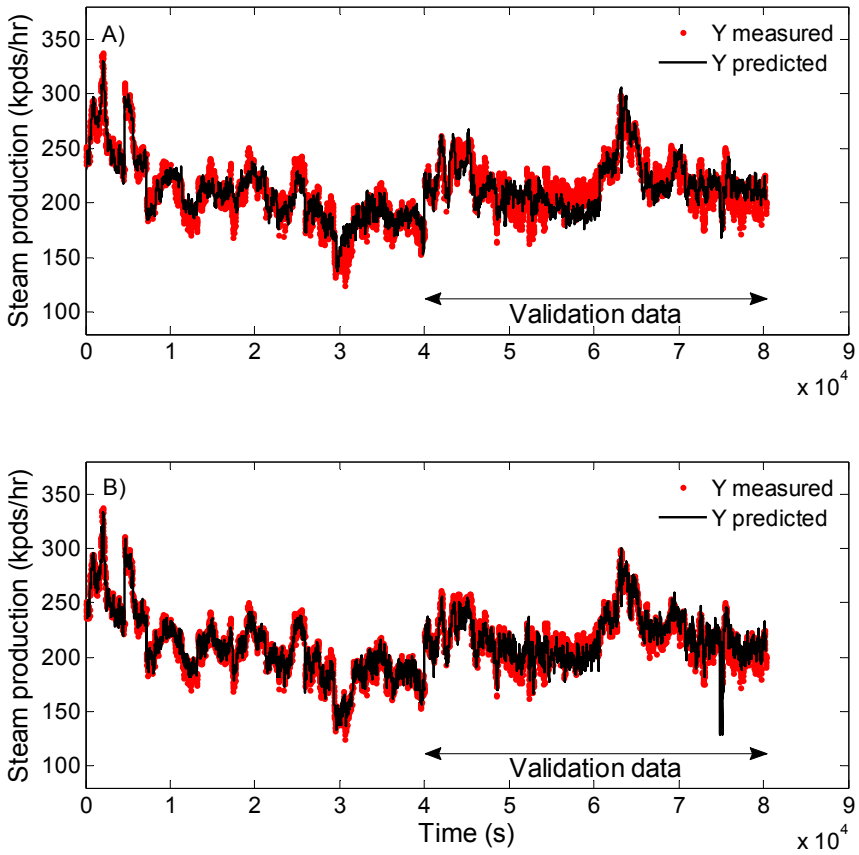

Figure 28 : Steam production 5 steps ahead forecast using the OE structure (A) and the ARX structure (B) (steam produced using natural gas or oil and/or bark)

Table 7: Comparison between both model performances in prediction (5 steps ahead)

Structure	Number of LV	Fit R^2 (%)	Validation R^2 (%)
OE	13	86.7	54.5
ARX	24	93.6	70.1

The results of the ARX model are generally better, especially on validation data, since it captures more variations in steam production. These results were expected because current steam production is correlated with previous values of steam production due to the boiler thermal inertia. If the boiler is cold and the combustibles are fed at their maximum values to increase the steam production, the value will not be as high as if the boiler would have been running for a long time. Therefore, the process seems likely to be described by an auto-regressive structure. However, this model seems to react more strongly to certain situations. For example, at a time of approximately $7.5x10^4$ seconds, an increase in the flow of the wettest bark mixture was tested. This almost killed the fire on the grate, and led to a sudden drop in steam production. The ARX model clearly underestimates steam production in this situation compared to the OE model which prediction was much closer to the measured steam production. Further long-term validation of the models is required to confirm which dynamic structure performs best in terms of predictive ability and robustness. The autoregressive disturbance of the ARX model seems to exaggerate certain situations such as the one discussed previously.

CHAPTER 6 CONCLUSION

Combustion is a process used in most industries such as in chemical industries where boilers are used to produce steam for the process or in metallurgical industries where rotary kiln are used to dry or to change a property in the material. The rising cost of the combustibles mainly used in combustion processes, which are fossil fuels, as well as environmental issues surrounding the use of those products are undeniable reasons for industries to change the combustible used. Biomass is nowadays a combustible frequently used instead of fossil fuels.

The biomass (bark) boiler operated by Irving Pulp and Paper's is one example which was used in this study to illustrate the opportunities offered by combustion imaging techniques. Their boiler is now mainly fed with bark. Bark is a solid fuel with some degree of moisture which feed rate and heat of combustion are highly variable. The high sources of variations brought by the use of a solid product such as bark forces the improvement of the control strategy normally used with fossil fuels combustibles. This research program aimed at developing combustion monitoring tools for biomass boilers, to help react to combustion disturbances in a very proactive and efficient manner through innovative automatic control schemes.

This was accomplished by performing an extensive review of boiler operation and control in chapter 2. This allowed the understanding of the process including the general operation of a fossil fuel boiler as well as the differences when using bark as a major combustible. Once the difficulties and issues surrounding the operation of biomass were identified (chapter 2), some basic possible improvements on the process were raised. In order to overcome the difficulties when using any kind of solid biomass as the main combustible in a boiler, a new control strategy was also suggested based on previous work also presented in chapter 2. Indeed, the changing properties of the combustible that cannot be easily quantified force the operation of a biomass boiler to be significantly different

from the operation of a boiler fed with fossil fuels. The main purpose of this work was to identify a new control strategy based on machine vision approach that could detect disturbances related to the biomass unsteady flow and/or changing properties using images already available on the site. Irving's bark boiler is already equipped with two high temperature video cameras. The images that are now only used to help the operators to know whether bark is burning well or not based on a view of the combustion chamber. The basic technique used to analyze images which is called MIA for multivariate image analysis is explained in details in chapter 4 as well as the different regression techniques used to related images and process data with the steam production which are developed in chapter 5. The different results of this research are shown is chapter 6.

In this research, we have shown that combining multivariate RGB images of the combustion chamber, relevant process data and chemometrics methods such as principal component analysis (PCA) and partial least square (PLS), allows achieving good predictions of steam flow rate when applied to a bark boiler (i.e. a typical biomass boiler). Using a controlled auto regressive (ARX) model type which requires having information on the past gives very good results (R^2_{fit}=93.6% and R^2_{val}=70.1%). Moreover, it was shown that this model could be used to forecast variations in steam production up to 2-3 minutes in the future with very reasonable accuracy. This should give some time for the operators to react proactively to these incoming combustion disturbances. Such a model could be incorporated within a new biomass boiler control strategy for stabilizing steam production.

A proof of concept that steam production can be predicted using bark combustion images has been made in this research. This project is the starting point for the development of a new control strategy for biomass boilers. Indeed, since steam production can be predicted using an image sensor, some work on the design and implementation of such a control strategy based on this new machine vision sensor can be pursued. Future work could include an analysis on the advantage for the operators to have a prediction of the steam production. Indeed, these techniques using multivariate image analysis have not yet been tested on real industrial time-varying processes. Some tests could therefore be made during a short period of time to see if the use of an on-line steam production

prediction based on the model built in this research would improve the efficiency of the boiler. Indeed, the prediction obtained by this model could help reduce the use of fossil fuels and contribute to stabilizing steam production. A comparison of the boiler performance could then be made between the situation where the operator has this helpful tool which is a predicted steam production that would occur if he doesn't change anything in the process and the previous situation where the operators did not had access to this information. Moreover, since no industrial application of control scheme has been installed yet using information gathered using MIA, this research would not be complete without considering the implementation of an eventual automatic feedback control scheme on this bark boiler. Indeed, a complete system of feedback process control based on the model developed in this research could be installed so the correction of the process could be made automatically, thus helping to reduce the impact of the combustion disturbances brought by the time varying bark properties. As soon as the tolerance limits of the fire intensity would be reached, the monitoring system based on the MIA model would adjust the speed of the screw feeders or other manipulated variables to make sure that heat released (indirectly measured by image color intensities) would remain steady. Then and only then, the power and the economical advantages of this approach would be quantified and highlighted.

CHAPTER 7 RECOMMENDATIONS

Some improvements can be made to ensure a better stabilisation of the process. This section will give a description about the possible improvements that would get the boiler operation better such as improving the images quality, having a steadier bark flow and moisture or making useful changes into the operation.

7.1 IMAGES QUALITY AND LOCATION

The importance of image quality is crucial and could be improved. Two things could help having better images. First, more frequent cleaning of the camera lens and mirror (used by the 5th floor camera) could help enhance image quality. Since the model is built using information on images relatives on color properties, the sharper the colors are and the more the relevant information taken from those images are useful. The resulting final model would probably be better when it comes to predicting changes due to a variation in heat released by the bark due to a change in its quality or its feed flow rate. Even if the best results were obtained with the 5th floor camera, it was found difficult to obtain a good contrast with this camera compared to the 3rd floor camera, which do no make use of a mirror.

Finding the best location of the high temperature cameras could also be investigated further. As mentioned, the information given by the camera located on third floor provides less information due to the fact that the field of view is limited to a small portion of the grate where the fire occurs. For this reason, the images captured are not as representative of the whole combustion process as the 5th floor camera is and, therefore, does not provide as much information for modelling purposes. The same argument can be applied to the camera located on fifth floor. Even if images provided by this camera are given a field of view big

enough to be useful, some improvement would lead to a greater image. Indeed, its location could be changed from its actual position which is over the arches. Those arches are limiting the field of view provided by the camera. Locating the camera under these arches could provide a complete view of the bark combustion area. The resulting images would then show the whole combustion zone providing an even more representative field of view of the fire. The information taken from those images would then be more precise leading to a better model.

Positioning the camera below the arches could also help simplifying the models. Indeed, as mentioned in the results, since images contain information on bark as well as on fossil fuels combustion, the model built is a general model including bark and fossil fuels. However, the heat released by both fossil fuels is easier to compute since their properties are well known and are stable. The relative steam production is then easy to obtain based on the relevant flow rates. For this reason, the contribution of fossil fuels to steam production could be removed from the model built using images. The steam produced by fossil fuels could be subtracted from the total steam produced and images could be used to predict the steam produced by the bark only. The resulting model would be simpler and may be more precise. Installing the camera below the arches would focus even more on bark combustion and remove any information about fossil fuels. Indeed, the camera being under the arches, the burners where both fossil fuels are fed would have less effect on the lighting of the images taken by the camera. The images collected closer to the grate might help give a better account of bark combustion on the grate.

7.2 BARK FLOW

An important source of disturbance in this process is due to unsteady bark flow. The screw feeders that feed the bark to the boiler are not always full which brings uncertainty about the bark flow and variations in the steam production. Bark is already a combustible that brings many disturbances. Indeed, its properties such as its moisture or its carbon content are continuously changing. Reducing disturbances at the source may contribute

significantly in achieving a steady steam production. The unsteady flow is one of the disturbances that come with the use of bark as a combustible. Indeed, since this solid is feed using a screw feeder, the flow is a calculated value based on the screw feeder speed assuming they are completely full. In addition to flow measurement uncertainties, some design modifications may be implemented to help delivery steadier biomass flow such as incorporation of laser detector to measure the fullness of the screw feeders.

7.3 BARK MOISTURE CONTENT

Bark moisture is an important parameter since it has a direct impact on the boiler operation. The air feeding strategy should be different according to the bark moisture. To increase efficiency, the amount of air fed in each section under the grate should be changed depending upon whether the bark is wet or dry. On-line moisture sensors could be used to track these variations before bark is fed in the boiler so the proper air feeding strategy could be adapted by the operators in a timely fashion.

7.4 AIR FLOW CONTROL STRATEGY

The combustible flow is used as the manipulated variable to control steam drum pressure. The appropriate air flow is determined based on fuel flow. Considering that bark flow is calculated according to the screw feeder speed assumed completely full, the amount of air may often be too high. This lowers the combustion efficiency because the excess of air cools down the combustion chamber and a part of the heat produced is not used to evaporate the water but is used to heat up the air inside the combustion chamber. For this reason, the air flow should be set according to a more accurate measurement such as the steam pressure set point. The screw speed could then be adjusted to meet the steam production required regardless on the carbon content of the bark flow.

BIBLIOGRAPHY

Allen, M. G., Butler, C.T., Johnson, S.A., Lo, E.Y. and Russo F. (1993), An Imaging Neural Network Combustion Control System for Utility Boiler Applications, Combust. Flame, 94, 205-214.

Bae, H., Kim, S., Wang, B.-H., Lee, M.H. and Harashima, F. (2006), Flame Detection for the Steam Boiler Using Neural Networks and Image Information in the Ulsan Steam Power Generation Plant, IEEE Transactions on Industrial Electronics, 53(1), 338-348.

Bell, T.A., Couch, S.W., Kreiger, T.L. and Feise, H.J. (2003), Screw feeders: A guide to selection use, Chemical Engineering Progress, 99, 2, ABI/INFORM Trade & Industry pg. 44.

Bertucco, L., A. Fichera, G. Nunnari, and A. Pagano (2000), A Cellular Neural Networks Approach to Flame Image Analysis for Combustion Monitoring, In: Proc. 6th IEEE Int. Workshop Cell. Neural Networks Appl., 455-459.

Burnham, A.J., Viveros, R. and MacGregor, J.F. (1996), Frameworks for Latent Variable Multivariate Regression, J Chemom, 10:31-45.

Duchesne, C. and MacGregor, J.F. (2001), Jackknife and bootstrap methods in the identification of dynamic models, J. Process Contr., 11(5), 553-564.

Dyer, M.J. and Crosley, D.R. (1982), Two-dimensional imaging of OH laser-induced fluorescence in a flame, Optics Letters, 7(8), 382-384.

Erikson, L., Johanson, E., Kettaneh-Wold, N., Trygg, J., Wikstrom, C. and Wold, S. (2006), Multivariate and Megavariate Data Analysis: Part 1 – Basic Principles and Applications, Umetrics Academy.

Esbensen, K.H. and Geladi, P. (1989), Strategy of Multivariate Image Analysis (MIA), Chemom. Intell. Lab. Syst., 7 , 67-86.

Esbensen, K. H., Geladi, P. and Grahn, H. (1992), Strategies of Multivariate Image Regression (MIR), Chemom. Intell. Lab. Syst., 14, 357-374.

Friedman, S. P. and Robinson, D.A. (2002), Particle shape characterization using angle of repose measurements for predicting the effective permittivity and electrical conductivity of saturated granular media, Water Resources Research, 38(11), 1236.

Gagné, G. (2009), La biomasse servira à chauffer l'hôpital d'Amqui, Le Soleil, 21 mai 2009.

Geladi, P. and Grahn, H. (1996), Multivariate Image Analysis, Wiley: Chichester, UK.

Geladi, P. and Kowalski, B.R. (1986), Partial Least-Squares Regression: A Tutorial, Anal Chim Acta, 185:1-17.

Giroux, A. (2008), Industrial Consumption of Energy (ICE) Survey – Summary Report of Energy Use in the Canadian Manufacturing Sector, 1995-2005.

Heselton, K. E. (2005), Boiler Operator's Handbook, Fairmont Press inc., Lilburn, Georgia, USA.

Höskuldsson, A. (1988), PLS Regression Methods, J Chemom, 2:211-228.

Huang, Y., Yan, Y., Lu, G. and Reed, A. (1999), On-line Flicker Measurement of Gaseous Flames by Image Processing and Spectral Analysis, Meas. Sci. Technol., 10, 726-733.

Karpetis, A.N. and Barlow, R.S. (2005), Measurements of flame orientation and scalar dissipation in turbulent partially premixed methane flames, In: Proceedings of the Combustion Institute, 30, 665–672.

Keyvan, S. (2005), Diagnostics and Control of Natural Gas-Fired furnaces via Flame Image Analysis using Machine Vision & Artificial Intelligence Techniques, Report DOE/CH/11032, US Department of Energy (DOI 10.2172/862201).

Ljung, L. (1999), System Identification: Theory for the user, 2nd edition, Prentice Hall Press, Englewood Cliffs, N.J.

Lu, G., Yan, Y., Huang, Y. and Reed, A (1999), An Intelligent Vision System for Monitoring and Control of Combustion Flames. Meas. Control, 32, 164-168.

Lu, G., Yan, Y., Ward, D.D. (2000), Advanced Monitoring, Characterisation and Evaluation of Gas Fired Flames in a Utility Boiler. Journal of the Institute of Energy, 73, 43-49.

Lu, G., Yan, Y. and Colechin, M. (2004), A Digital Imaging Based Multifunctional Flame Monitoring System, IEEE Transactions on Intrumentation and Measurement, 53(4), 1152-1158.

Namazian, N., Kelly, J., Schefer, R.W., Johnston, S.C. and Long, M.B. (1989), Nonpremixed bluff-body burner flow and flame imaging study, Experiments in Fluids, 8, 216-228.

Romero, C., Li, X., Keyvan, S. and Rossow, R. (2005), Spectrometer-based combustion monitoring for flame stoichiometry and temperature control, Applied Thermal Engineering, 25, 659–676.

Shimoda, M., Sugano, A., Kimura, T., Watanabe, Y. and Ishiyama, K. (1990), Prediction Methods of Unburnt Carbon for Coal Fired Utility Boiler Using Image Processing Technique of Combustion Flame, IEEE Transactions on Energy Conversion, 5, 640-645.

Simard, G., Personal communication, December 2008.

Smith C. A. and Corripio, A. (2005), Principles and Practice of Automatic Process Control, Third Edition, Wiley.

Sun, P., Chai, T. and Zhou X.J. (2008), Rotary Kiln Flame Image Segmentation Based on FCM and Gabor Wavelet Based Texture Coarseness, In: Proceedings of the 7th World Congress on Intelligent Control and Automation, June 25-27, 2008, Chongqing, China, 7615-7620.

Szatvanyi, G., Duchesne, C. and Bartolacci, G. (2006), Multivariate Image Analysis of Flames for Product Quality and Combustion Control in Rotary Kilns, Ind. Eng. Chem. Res., 45, 4706-4715.

Tao, W., H. Burkhardt (1995), Vision-Guided Flame Control Using Fuzzy Logic and Neural Networks, Part. Part. Syst. Charact., 12, 87-94.

Victor, J., Costeira, J., Tomé, J., Sentieiro, J. (1991), A Computer Vision System for the Characterization and Classification of Flames in Glass Furnaces, Conference Record of the IEEE Industry Applications Society Annual Meeting, 1109-1117.

Wang, F., Wang, X.J., Ma, Z.Y., Yan, J.H., Chi, Y., Wei, C.Y.,Ni, M.J. and Cen, K.F. (2002), The Research on the Estimation for the NOx Emissive Concentration of the Pulverized Coal by the Flame Image Processing Technique, Fuel, 81, 2113-2120.

Waterfall, R.C., R. He, C.M. Beck, (1997), Visualizing combustion using electrical impedance tomography, Chemical Engineering Science, 52(13), 2129-2138.

Xu. L., Yan, Y., Cornwell, S., Riley, G. (2005), Online Fuel Tracking by Combining Principal Component Analysis and Neural Network Techniques, IEEE Transactions on Intrumentation and Measurement, 54(4), 1640-1645.

Yamaguchi, T., Grattan, K.T.V., Uchiyama, H., Yamada, T. (1997), A Practical Fiber Optic Air-Ratio Sensor Operating by Flame Color Detection, Rev. Sci. Instrum., 68, 197-202.

Yan, Y., Lu, G. and M. Colechin (2002), Monitoring and characterisation of pulverized coal flames using digital imaging techniques, Fuel, 81, 647-656.

Yu, H. and MacGregor, J.F. (2003), Multivariate Image Analysis and Regression for Prediction of Coating Content and Distribution in the Production of Snack Foods, Chemometr. Intell. Lab., 67, 125-144.

Yu, H., and MacGregor, J.F. (2004), Monitoring Flames in an Industrial Boiler Using Multivariate Image Analysis, AIChE J., 50(7), 1474-1483.

Zhang, H.-L., Zou, Z., Li, J., Chen, X.-T. (2008), Flame image recognition of alumina rotary kiln by artificial neural network and support vector machine methods, J. Cent. South Univ. Technol., 15, 39−43.

ANNEXE 1: MATLAB SCRIPTS

SCRIPTS SUMMARY

Some of the MATLAB scripts used in this work are presented. A list of all the scripts presented as well as a short description is given.

- **Image extraction**: Takes each video recorded from both camera and extracts each frame
- **Synchronization**
  - Image time: Stores the time acquisition of each image
  - Data time: Stores the time acquisition of each data
  - Synchronizing images and data: Compared the image time and the data time acquisition and finds the tinniest difference as long as this difference is higher than 4 seconds. This is the synchronized image.
- **MIA**
  - Finding score density histogram scaling: Finds the minimum and the maximum values of t_1 and t_2 relatives to each synchronized images
  - Building score density histogram: Takes each synchronized images and build the score density histogram (**H**) based on the same scaling
  - Luminous region finding: This script allows the user to display a score density histogram, choose a region of interest and highlight the corresponding pixels in the original image back and forth until the luminous region is found.
  - Computing image features: Takes each synchronized image, computes the 9 image features and stores the values for each image in one row.
- **Building matrices**
  - Averaging: Takes each row of the matrix **X** and replace it with the average value of the row k, k-1, ..., k-n+1.
  - Matrices $\mathbf{X_{MIR}}$ and **Y**: Takes the averaged matrix **X** and build matrices $\mathbf{X_{MIR}}$ and **Y** according to the dynamic needed (numbers of past inputs and outputs lags, prediction)
- **Regression**: PLS regression between matrix $\mathbf{X_{MIR}}$ and **Y**

IMAGE EXTRACTION

```
% Prend un film et le numérise en enregistrant les valeurs des pixels de
% chaque image dans une structure.
[nom, chemin] = uigetfile('*.wmv','Veuillez sélectionner le vidéo à numériser');
video_name = [chemin, nom];
cd(chemin)

video=mmread(video_name,1:1,[],[],'disableAudio');%Pour estimer le nombre de frame dans le
film.

if video.nrFramesTotal < 0 % Lorsqu'on a une estimation elle est négative
    frame_tot = -1*video.nrFramesTotal;
else
```

```
    frame_tot = video.nrFramesTotal;
end

warning off
sstep=400;
compteur1 = fix(frame_tot/sstep);%On définit le nombre de structures qui vont contenir
exactement sstep images.

% Structures pleines (comportant sstep images)
for i=1:compteur1

    video=[];
    ii=1+sstep*(i-1);
    video=mmread(video_name,ii:i*sstep,[],[],'disableAudio');
    nommatricevideo = sprintf('cam520080204%d',(i+1000));% On met (i+1000) pour que
lorsqu'on voudra faire un film avec ces figures, la première figure du film soit la figure 1
et non la 10...
    save(nommatricevideo,'video')
    i

end


%Dernière structure...
if rem(frame_tot,sstep)~=0 %On s'assure qu'il y ait vraiment un histogramme incomplet.

    video = mmread(video_name,((sstep*compteur1)+1):frame_tot,[],[],'disableAudio');
    nommatricevideo = sprintf('cam520080204%d',(compteur1+1+1000));
    save(nommatricevideo,'video')

end
```

SYNCHRONIZATION

IMAGE TIME

```
% Sortir les heures de toutes les images du film

% Ouvre une fenêtre qui permet à l'utilisateur de choisir le fichier
% dans lequel se trouve les images qui serviront à faire la régression.
clear all

fichier = uigetdir('C:\','Veuillez sélectionner le dossier contenant les images
numérisées');

% Définir l'heure réelle de début du film
display('Entrez date et heure du début du film')
an = input('Année : ');
mois = input('Mois : ');
jour = input('Jour : ');
heure = input('Heure : ');
minute = input('Minute : ');
seconde = input('Seconde : ');

temps = [];
contenu = (dir([fichier,'\','*','.mat']));
mat_contenu = struct2cell(contenu)';
[l,c] = size(mat_contenu);
b=0; %Compteur pour changer de ligne dans imagesdufichier

for i = 1:l
    if mat_contenu{i,4} == 0
        b = b+1;%À la fin de la boucle, b représente le nombre de matrice dans le fichier.
        imagesdufichier(b,1) = mat_contenu(i,1);
    end
end
```

```
% Définir le cd pour la fonction imread.

cd(fichier)
for i=1:b

    nomimage = char(imagesdufichier(i));
    z = load(nomimage);

    if size(z.video.frames,2) ~= 0

        a = ones(size(z.video.times,2),1);
        debut = [an*a mois*a jour*a heure*a minute*a seconde*a];
        offset = [0*a 0*a 0*a 0*a ((0*a)-4*a) 0*a-32*a];
        images = [0*a 0*a 0*a 0*a 0*a z.video.times'];

        real_time = debut + offset + images;

        % Le temps des images est égal à celui enregistré dans la matrice conçue par mmread.
À ce temps, on ajoute l'heure du début du vidéo et on soustrait l'OFFSET par rapport au PhD
qui est de 4 minutes et 32 secondes. On obtient alors le temps synchronizé
        format long
        temps_images_syncro_PhD = datenum(real_time);

        temps = [temps; temps_images_syncro_PhD];
    end
end
```

DATA TIME

```
% Sortir les heures de toutes les données de production

% Ouvre une fenêtre qui permet à l'utilisateur de choisir le fichier
% dans lequel se trouve les images qui serviront à faire la régression.
[nom, chemin] = uigetfile('*.*','Veuillez sélectionner le fichier excel');
fichier_name = [chemin, nom];
cd(chemin)

[data, txt] = xlsread(fichier_name);

a = size(txt,1);

format long
temps_data = datenum(txt(7:a,3));
```

SYNCHRONIZING IMAGES AND DATA

```
% Trouver l'image qui est la plus proche de la donnée i tout en étant plus petite que 0
(l'image doit venir avant la production de vapeur...) mais plus grande que -4secondes (le
délai entre l'image et la production de vapeur doit être de moins de 4 secondes...)
clear all

format long
% Images
[nom, chemin] = uigetfile('*.mat','Veuillez sélectionner le fichier pour le temps des
images','C:\');
nomcomplet = [chemin, nom];
cd(chemin)
temps_images = load(nomcomplet);
temps_images = temps_images.temps;
temps_images = temps_images(:,2);

% Données
[nom, chemin] = uigetfile('*.mat','Veuillez sélectionner le fichier pour le temps des
données','C:\');
nomcomplet = [chemin, nom];
```

```
cd(chemin)
temps_data = load(nomcomplet);
temps_data = temps_data.temps_data;
temps_data = temps_data(:,2);

% Syncroniser les données

delaimax = 4/(3600*24);% Un intervalle d'une journée représente 1. Pour avoir 9 secondes, il
suffit de diviser 4 par (3600*24)sec
syncro_results = [];

for i = 1:size(temps_data,1)

    % On trouve la différence entre le temps de la data i et tous les temps des images... On
trouve ensuite la différence la plus petite tout en respectant les contraintes.
    diff_temps = temps_data(i,1)*ones(size(temps_images,1),1)-temps_images;

    for j = 1:size(diff_temps,1)
        if diff_temps(j,1) < 0
            % On s'assure que l'image vient avant la donnée de production
            % de vapeur. Si la différence entre les 2 temps moins
            % que 0 alors l'image est après et on l'élimine
            diff_temps(j,1) = 1e10; % On met un gros nombre pour l'éliminer.
        else diff_temps(j,1)> delaimax
            % On s'assure que le délai entre l'image et la donnée est, au
            % plus, 9 secondes...
            diff_temps(j,1) = 1e10;
        end
    end

    if min(diff_temps) ~= 1e10
        % On s'assure qu'il y ait au moins une image qui réponde aux
        % critères.
        x = find(diff_temps == min(diff_temps)); % indice de l'image la plus près
        syncro_results = [syncro_results; i temps_data(i,1) x temps_images(x,1)];
    end
end
```

MIA

FINDING SCORE DENSITY HISTOGRAM SCALING

```
clear all
X = [];
Z = zeros(3);

t1min = realmax;% Valeurs assez élevées pour qu'elles changent au premier score calculé.
t1max = -realmax;
t2min = realmax;
t2max = -realmax;

question1 = 1;

while question1 == 1

fichier = uigetdir('C:\','Veuillez sélectionner le dossier contenant les images à
comparer');

contenu = (dir([fichier,'\','*','.jpg']));

% Convertit la structure en une cellule de p*n*m (nombre de fields*nombre
% d'images*1.

mat_contenu = struct2cell(contenu)';

% On veut s'assurer de conserver seulement les images et non les fichiers
% contenus dans fichier.
```

```
[l,c] = size(mat_contenu);
a=0; %Compteur pour changer de ligne dans imagesdufichier

for i = 1:l
    if mat_contenu{i,4} == 0
        a = a+1;%À la fin de la boucle, a représente le nombre d'images dans le fichier.
        imagesdufichier(a,1) = mat_contenu(i,1);
    end
end

cd(fichier)

% On ouvre chaque image du fichier et on la transforme en une matrice de
% trois colonnes (RGB) pour ensuite calculer la matrice de Kernel relative à
% cette image (matrice de variance-covariance X'X). On peut alors trouver la
% matrice de Kernel globale pour toutes les images.

for i = 1:a
    nomimage = char(imagesdufichier(i));
    RGB = imread(nomimage);
    [largeur,hauteur,profondeur]=size(RGB);
    X = double(reshape(RGB, largeur*hauteur,profondeur));
    Z = Z + (X'*X);
end

question1 = input('Y a-t-il dautres images à analyser? (oui=1)');
end
% On applique une décomposition SVD. La matrice P contient les eigenvectors
% de Z, PT est la matrice transposée de P et E est une matrice dont la
% diagonale est composée des eigenvalues de Z.
[P,E,PT] = svd(Z);

% On peut maintenant trouver les scores (t) relatifs à chaque image et ainsi
% trouver les extrémums de l'ensemble des images.
question1 = 1;
while question1 == 1

    fichier = uigetdir('C:\','Veuillez sélectionner le dossier contenant les images à
comparer');

    contenu = (dir([fichier,'\','*','.jpg']));

    % Convertit la structure en une cellule de p*n*m (nombre de fields*nombre
    % d'images*1.

    mat_contenu = struct2cell(contenu)';

    [l,c] = size(mat_contenu);
    a=0; %Compteur pour changer de ligne dans imagesdufichier

    for i = 1:l
        if mat_contenu{i,4} == 0
            a = a+1;
            imagesdufichier(a,1) = mat_contenu(i,1);
        end
    end

    % Définir le cd pour la fonction imread.
    cd(fichier)

    for i = 1:a
        nomimage = char(imagesdufichier(i));
        RGB = imread(nomimage);
        [largeur,hauteur,profondeur]=size(RGB);
        X = double(reshape(RGB, largeur*hauteur,profondeur));
        T = X*P;
        t1 = T(:,1);
        t2 = T(:,2);
    % On vérifie que les valeurs de t1 et t2 sont dans les limites des extrémités.
        if min(t1) < t1min
            t1min = min(t1);
```

```
        end
        if max(t1) > t1max
            t1max = max(t1);
        end
        if min(t2) < t2min
            t2min = min(t2);
        end
        if max(t2) > t2max
            t2max = max(t2);
        end

    end

    question1 = input('Y a-t-il dautres images à analyser? (oui=1)');
end
```

BUILDING SCORE DENSITY HISTOGRAM

```
% On doit d'abord définir t1min, t1max, t2min, t2max et Z
[nom, chemin] = uigetfile('*.mat','Veuillez sélectionner le fichier des extrémités','C:\');
nomcomplet = [chemin, nom];
cd(chemin)
ext = load(nomcomplet);
t1min = ext.t1min; t1max = ext.t1max; t2min = ext.t2min; t2max = ext.t2max; Z = ext.Z;
[P,E,PT]=svd(Z);


% ------------------------------------------------------------------------
% Lorsqu'on veut faire un histogramme à partir d'images dont on doit
% enlever (ou conserver seulement) une partie qui est nuisible dans l'image.
% On doit aussi loader la matrice B dans la fenêtre de commandes. La
% matrice B est l'image dont les 0 représente les pixels sous le masque.
% BW = ext.BW;
% [imask,jmask] = find(BW==0);
% ------------------------------------------------------------------------


% Définir les numéros d'images syncronisées avec les données.
[nom, chemin] = uigetfile('*.*','Veuillez sélectionner le fichier excel de la
syncronisation','C:\');
nomcomplet = [chemin, nom];
cd(chemin)
no_images = xlsread(nomcomplet);
no_images = no_images(:,3); % Vecteur de nX1

nom_fichier_initial = input('Numéro cam, date(ex. : 320080617)');
nom_fichier = 'vide';

% Définir le répertoire avec les matrices des images numérisées d'une
% journée.
fichier = uigetdir('C:\','Veuillez sélectionner le dossier contenant les images numérisées
de cette journée');
cd(fichier) %Pour loader les différentes matrices, on doit être dans ce répertoire.

H = ones(256,256,size(no_images,1));
for i = 1:size(no_images,1)
    nom_fichier_comparaison = nom_fichier;
    if rem(no_images(i,1),400) ~= 0
        no_fichier = fix(no_images(i,1)/400)+1;

    else
        no_fichier = fix(no_images(i,1)/400);

    end

    nom_fichier = sprintf('cam%d%d',nom_fichier_initial,(1000+no_fichier));
    tf = isequal(nom_fichier, nom_fichier_comparaison);
```

```
    if tf == 0
        % On empêche de loader le même fichier pour les images consécutives
        % qui sont dans le même fichier.
        Images = load(nom_fichier);

    end

    % Le numéro de l'image à aller chercher dans Images correspond à
    % no_images - 400*(no_fichier-1) (ex : 404 représente l'image 4 du
    % fichier 2 donc 404 - 400(2-1) = 4
    RGB = Images.video.frames(1,(no_images(i,1)-400*(no_fichier-1))).cdata;
%     % -----------------------------------------------------------------
    % On redéfinit les pixels sous le mask pour les faire
    % ''disparaitre''.
    for m = 1:length(imask),
        RGB(imask(m),jmask(m),1) = 0;
        RGB(imask(m),jmask(m),2) = 0;
        RGB(imask(m),jmask(m),3) = 0;
    end;
%     % -----------------------------------------------------------------
    [largeur,hauteur,profondeur]=size(RGB);
    X = double(reshape(RGB, largeur*hauteur,profondeur));
    T = X*P;%Chaque colonne de la matrice T est un vecteur qui représente un score de X.
    t1 = T(:,1);
    t2 = T(:,2);


    s1 = ones(size(t1,1),1)+round(((t1 - (t1min*ones(size(t1,1),1))))./((t1max-
t1min))).*255);
    s2 = ones(size(t2,1),1)+round(((t2 - (t2min*ones(size(t2,1),1))))./((t2max-
t2min))).*255);


    nbrgroupes = 256;%Ces trois valeurs demeurent fixes pour toutes les
%     images ce qui nous permet de nous assurer que les mêmes caractéristiques de chaque
image se retrouve dans la même région puisque le scaling est le même.

    % On doit maintenant faire une matrice H. Initialement, cette matrice contient des 0. On
passe les valeurs de t1new et t2new et la valeur contenue dans H à la position (t1new,
t2new) augmente de 1.
    Amat = zeros(nbrgroupes, nbrgroupes);

    for k = 1:length(s1)
        % A(#rangée, #colonne)... comme t1 représente les colonnes et t2 les rangées ont met
A(t2new, t1new)
        Amat(s2(k), s1(k)) = Amat(s2(k), s1(k)) + 1;
    end
    %-----------------------------------------------------------------

    A = flipud(Amat);%On change la matrice pour qu'elle devienne la représentation exacte de
plot(t1,t2) et non son image miroir.

% On change les valeurs de l'histogramme pour en faire l'affichage. Les nouvelles valeurs %
sont comme un code de couleur. Utile pour faire des masques puisqu'il faut l'afficher.
% [i,j] = find(151<=A);
% for m = 1:length(i),
%     A(i(m),j(m)) = 256; % Those bins in A having 151 or greater hits are reassigned to  %
a pixel intensity value of 256
% end;
%
% [i,j] = find(76<=A & A<=150);
% for m = 1:length(i),
%     A(i(m),j(m)) = 226;
% end;
%
% [i,j] = find(34<=A & A<=75);
% for m = 1:length(i),
%     A(i(m),j(m)) = 201;
% end;
%
% [i,j] = find(16<=A & A<=33);
```

```
% for m = 1:length(i),
%     A(i(m),j(m)) = 176;
% end;
%
% [i,j] = find(7<=A & A<=15);
% for m = 1:length(i),
%     A(i(m),j(m)) = 151;
% end;
%
% [i,j] = find(3<=A & A<=6);
% for m = 1:length(i),
%     A(i(m),j(m)) = 126;
% end;
%
% [i,j] = find(A==2);
% for m = 1:length(i),
%     A(i(m),j(m)) = 101;
% end;
%
% [i,j] = find(A==1);
% for m = 1:length(i),
%     A(i(m),j(m)) = 76;
% end;
%
% [i,j] = find(A==0);
% for m = 1:length(i),
%    A(i(m),j(m)) = 1; % Those bins in A having NO hits are reassigned to a pixel
% intensity value of 1 (smallest value representing black)
% end;

    H(:,:,i) = A;
end
```

LUMINOUS REGION FINDING

```
clear all

% Pour faire des mask, on doit d'abord faire afficher une image. On peut aussi appliquer des
masques déjà définis.
% Loader la matrice qui contient tous les masques

% % -------------------------------Masques prédéfinis------------------% [nom, chemin] =
uigetfile('*.mat','Veuillez sélectionner le fichier des masques','C:\');
% nomcomplet = [chemin, nom];
% cd(chemin)
% masques = load(nomcomplet);
% masques = masques.mask;
%
% scoreplotpouruneimageFINAL
%
% for i = 1:size(masques,3)
%     BW = masques(:,:,i);
% ------------------------------------------------------------------------
% -------------------------Masques fait manuellement-------------------
scoreplotpouruneimageFINAL

% On utilise la commande roipoly qui nous permet de tracer un mask
% en forme de polygone comme on veut. La matrice BW est une image
% dont les pixels sous le mask sont blanche (valeur 1) et les autres
% sont noirs (valeur 0).
%
Indic = 1;

while Indic == 1

    BW = roipoly;

% % ------------------------------------------------------------------
```

```
% % Faire afficher le mask dans le score plot
    B = ~BW; % B is the mirror of BW (i.e. B has 0 for selected pixels and 1 for background
pixels)

    % A représente l'histogramme de l'image.
    % Element by element multiplication of B with A will result in C (with 0 for the
selected pixels)
    C = double(B).*A;

    [i,j] = find(C==0);

    colormap(hot(256)); % HOT is a Matlab built-in colormap
    map = colormap;

    CRGB = uint8(round(ind2rgb(C,map)*255));

    % On fait en sorte que les pixels sous le mask soient roses.
    for m = 1:length(i),
        CRGB(i(m),j(m),1) = 160;
        CRGB(i(m),j(m),2) = 0;
        CRGB(i(m),j(m),3) = 105;
    end;

    imshow(CRGB,'InitialMagnification','fit');
    xlabel('t1');
    ylabel('t2');


    % Identifie les pixels de l'image initiale qui se trouve sous le masque
    Anouveau = A;
    [i,j] = find(Anouveau==1);
    for m = 1:length(i),
        Anouveau(i(m),j(m)) = 0; % Those bins in A having NO hits are reassigned to a pixel
intensity value of 0
    end;


    % Les pixels sous le mask deviennent avec une valeur de -1 et les autres pixels
conservent leur valeur de 0. On a redonné des valeurs nulles aux endroits qui n'avaient
aucun point dedans (qui était quand même 1).

    BWneg = -1*double(BW);

    Amaskneg = Anouveau.*BWneg;

    Amiroir = flipud(Amaskneg);% On remet la matrice A dans sa forme originale pour que les
positions correspondent à celles de t2new.

    [i,j] = find(Amiroir < 0);% On trouve les coordonnées des carreaux sous le mask.

    % On veut créer un vecteur qui nous indiquera les positions où la valeur dans t2new et
celle de t1new correspondent aux positions des carreaux sous le mask. On commence donc par
faire un vecteur de 0 de la longueur de t1new pour ensuite assigner une valeur différente
aux positions correspondantes.
    pixelmaskt1 = zeros(size(t1new));


    for a = 1:length(i)
        for b = 1:length(t1new)
            if i(a)==t2new(b) && j(a)==t1new(b)

                pixelmaskt1(b) = 1;% On met des valeurs

            end
        end
    end

    % On peut maintenant redonner la forme de l'image initiale à pixelmaskimage pour faire
en sorte que les pixels du mask se retrouvent à l'endroit original sur l'image.
    pixelmaskimage = reshape(pixelmaskt1, largeur, hauteur);
```

```
    [i,j]=find(pixelmaskimage == 1);% les pixels sous le mask ont des valeurs de 1

    image = RGB;

    for a = 1:length(i)
        image(i(a),j(a),1) = 28;
        image(i(a),j(a),2) = 255;
        image(i(a),j(a),3) = 28;
    end

    figure(2);
    imshow(image,'InitialMagnification','fit');

  Indic = input('Voulez-vous refaire le mask?: (0 = non, 1 = oui): ');
    if Indic == 1
        close(2)
    end
 end
```

COMPUTING IMAGE FEATURES

```
% Loader les histogrammes
[nom, chemin] = uigetfile('*.mat','Veuillez sélectionner le fichier des
histogrammes','C:\');
nomcomplet = [chemin, nom];
cd(chemin)
histogramme = load(nomcomplet);
histogramme = histogramme.H;
%%
% Loader le masque qui représente la flamme.
[nom, chemin] = uigetfile('*.mat','Veuillez sélectionner le fichier du mask de la
flamme','C:\');
nomcomplet = [chemin, nom];
cd(chemin)
masques = load(nomcomplet);
masques = masques.mask;

%% --------------------- Partie pour trouver RGB final -----------------
% Loader les extrémités.
[nom, chemin] = uigetfile('*.mat','Veuillez sélectionner le fichier des extrémités','C:\');
nomcomplet = [chemin, nom];
cd(chemin)
ext = load(nomcomplet);
t1min = ext.t1min; t1max = ext.t1max; t2min = ext.t2min; t2max = ext.t2max; Z = ext.Z;
[P,E,PT]=svd(Z);

p1 = P(:,1);
p2 = P(:,2);

RGB_final = zeros(256,256,3);
for i = 0:1:255
    t1 = (i*(t1max-t1min)/255)+t1min;
    for j = 0:1:255
        t2 = (j*(t2max-t2min)/255)+t2min;
        RGB = t1*p1'+ t2*p2'+[1 1 1];
        for k = 1:3
            if RGB(1,k) < 0
                RGB(1,k)=0;
            end
        end

        RGB_final((i+1),(j+1),:) = reshape(RGB,1,1,3);
    end
end
ccc = RGB_final;
RGB_final(:,:,1) = flipud(ccc(:,:,1));RGB_final(:,:,2) = flipud(ccc(:,:,2));RGB_final(:,:,3)
= flipud(ccc(:,:,3));

%% ---------------------- Propriétés de la flamme ---------------------
```

```
L = 0.299*RGB_final(:,:,1)+0.587*RGB_final(:,:,2)+0.114*RGB_final(:,:,3);

% On prédéfinit la matrice contenant les nouvelles images binnées.
features = zeros(size(histogramme,3),9);

pixels_total = sum(sum(histogramme(:,:,1)));

mask = masques;
inv_mask = ~mask;

for j = 1:size(histogramme,3)

    image = histogramme(:,:,j);
    c = image.*mask;

%     ---------------Luminous Features-------------------------
%     Luminous Region Area A (nombre de pixels qui se trouvent dans la région lumineuse.
    A = sum(sum(c));
    features(j,1) = A;

    % Flame Brightness B
    b = c.*L;
    B = sum(sum(b));
    features(j,2) = B;

    % Uniformity of Flame Brightness U
    if A ~= 0
        u = c.*L.*L;
        u = sum(sum(u));
        U = ((u-B)/A)^(0.5);
        features(j,3) = U;
    else
        features(j,3) = 0;
    end

    % Average Brightness of the nonluminous area W
    inv_c = image.*inv_mask;
    inv_b = inv_c.*L;
    if sum(sum(inv_c)) ~= 0
        W = sum(sum(inv_b))/sum(sum(inv_c));
        features(j,4) = W;
    else
        features(j,4) = 0;
    end

%     --------------Color Features----------------------
%     Average Color of the whole Image
    SS1 = sum(image,1);
    SS2 = sum(image,2);
    S1 = 0;
    S2 = 0;
    for i = 1:length(SS1)
        S1 = S1 + SS1(1,i)*i;
        S2 = S2 + SS2(i,1)*i;
    end
    features(j,5) = S1/pixels_total;
    features(j,6) = S2/pixels_total;

%     Average Color of the Luminous Region
    SS1_flame = sum(c,1);
    SS2_flame = sum(c,2);
    S1_flame = 0;
    S2_flame = 0;
    for i = 1:length(SS1_flame)
        S1_flame = S1_flame + SS1_flame(1,i)*i;
        S2_flame = S2_flame + SS2_flame(i,1)*i;
    end
    if A ~= 0
        features(j,7) = S1_flame/A;
        features(j,8) = S2_flame/A;
    else
```

```
        features(j,7) = 0;
        features(j,8) = 0;
    end

%     Total Number of Colors in the Flame Region
    [row,colomn] = find(c~=0);
    N = length(row);
    features(j,9) = N;
end
```

BUILDING MATRICES

AVERAGING

```
function [x_moy, y] = moyennage(a, x_ini,y_ini)

% Avec cette fonction, on prend un x comprenant des données
% de production et on moyenne a rangées à la fois
% Lorsqu'on veut donner un poids différent des valeurs à moyennées
% g = 0.7;
% div = (g^0*(1-g)+g^1*(1-g)+g^2*(1-g)+g^3*(1-g)+g^4*(1-g)+g^5*(1-g));
% poids = [g^5*(1-g) g^4*(1-g) g^3*(1-g) g^2*(1-g) g^1*(1-g) g^0*(1-g)]./div;

if a>1
    x_moy = zeros((size(x_ini,1)-a+1),size(x_ini,2));
    for i = a:size(x_ini,1)
        x_moy(i-a+1,:) = mean(x_ini(i-a+1:i,:));%poids*(x_ini(i-a+1:i,:));
    end

    y=y_ini(a:length(y_ini),:);

else
    x_moy = x_ini;
    y = y_ini;
end
```

MATRICES X_{MIR} AND Y

```
clear all

function [x_new,y_new] = prediction(pas,b,a,YY,x,y)
% YY doit toujours être plus petit ou égale à b

[N,sx] = size(x);
sy = size(y,2);

% Construire le modèle sans prédiction revient au même que tomber en
% prédiction à 1 pas et ainsi considérer l'erreur nulle.
if pas == 0
    pas = 1;
end

y_new = y((b*a+1+(pas-1)*a):N);
% pas réfère au nombre de pas à l'avance qu'on veut prédire. pas = 3, alors
% notre prédiction sera y(k+3|k).

% a réfère au nombre de lack entre les images. On prend une image toutes
% les minutes donc toutes les 6 images.

% b réfère au nombre d'images dans la dynamique. On veut avoir une
% dynamique de 15 images donc pendant 15 minutes.

% YY réfère au nombre de y qu'on veut dans la régression. Si on prend 3 y,
% on aura y(k), y(k-1) et y(k-2)
```

```
compteurx = 1:a:a*b;% On prend donc les images 1,7,13,19,25,...
matricex = zeros(size(y_new,1),(b+(pas-1))*sx);
for i = 1:b

    matricex(:,(((i-1)*sx)+1):i*sx) = x(compteurx(1,i):(compteurx(1,i)+(N-(b*a+1+(pas-
1)*a))),:);

end

% On ajoute x(k+1), x(k+2),... en supposant que ce sont les mêmes que x(k).
% On dit donc voici ce qui va arriver si jamais les choses ne changent pas.
if pas > 1
    for i = (b+1):(b+pas-1)
        matricex(:,(((i-1)*sx)+1):i*sx) = matricex(:,((b-1)*sx+1):b*sx);
    end
end

compteury = compteurx(1,(b-YY+1):b)+a;
matricey = zeros(size(y_new,1),YY*size(y,2));
if YY ~= 0
    for i = 1:YY

        matricey(:,(((i-1)*sy)+1):i*sy) = y(compteury(1,i):(compteury(1,i)+(N-(b*a+1+(pas-
1)*a))),:);

    end
end

x_new = [matricey matricex];
```

REGRESSION

```
function [ymeasured,ypred,n,sizemod,rmod,rval,rcum] = modeliser(xmod,ymod,xval,yval)

%-------------------------------------------------------------------------
% Les lettres majuscules sont pour les variables preprocesser alors que les lettres
minuscules sont pour les variables sans preprocessing.
%-------------------------------------------------------------------------

options.display    = 'off';%Permet de faire apparaître le % de variation capturé par chaque
PC dans la command window.
options.plots      = 'none';%Fait en sorte qu'aucun graphique s'affiche à l'appel de la
fonction pls.

% Pour enlever le preprocessing et retrouver les valeurs initiales
% data = preprocess('undo',sp,datap)
s = preprocess('default','Autoscale');
[Xmod,spx] = preprocess('calibrate',s,xmod);
[Ymod,spy] = preprocess('calibrate',s,ymod);

Xmod = double(Xmod);
Ymod = double(Ymod);

% En utilisant les structures de preprocessing on peut faire le
% preprocessing approprié pour les vecteurs de validation.
Xval = preprocess('apply',spx,xval);
Yval = preprocess('apply',spy,yval);

Xval = double(Xval);
Yval = double(Yval);

a = 30; % Nombre de composantes à investiguer
Q2 = zeros(1,a);
R2 = zeros(1,a);
Rcum = zeros(1,a);

for i = 1:a
```

```
    model = pls(Xmod, Ymod, i, options);
    Ypred_mod = model.pred{1,2};
    ypred_mod = preprocess('undo',spy,Ypred_mod);
    ypred_mod = double(ypred_mod);
    valid = pls(Xval,Yval,model,options);
    Ypred_val = valid.pred{1,2};
    ypred_val = preprocess('undo',spy,Ypred_val);
    ypred_val = double(ypred_val);

    coeffq = corrcoef([ypred_val yval]);
    Q2(1,i) = 100*coeffq(2,1)*coeffq(2,1);
% Le coefficient R2 est calculé selon la vrai formule. Par contre, on
% aurait très bien pu prendre le coefficient de corrélation au carré.
%     rsquare(1,i) = ((Ypred_mod - ybar)'*(Ypred_mod - ybar))/((Ymod -

    R2(1,i) = model.detail.ssq(i,5);

    ymeasured = [ymod;yval];
    ypred = [ypred_mod;ypred_val];

    coeffrcum = corrcoef([ymeasured ypred]);
    Rcum(1,i) = 100*coeffrcum(2,1)*coeffrcum(2,1);

end

[i,n] = find(Rcum == max(Rcum));

model = pls(Xmod, Ymod, n, options);
Ypred_mod = model.pred{1,2};
ypred_mod = preprocess('undo',spy,Ypred_mod);
ypred_mod = double(ypred_mod);
valid = pls(Xval,Yval,model,options);
Ypred_val = valid.pred{1,2};
ypred_val = preprocess('undo',spy,Ypred_val);
ypred_val = double(ypred_val);

ymeasured = [ymod;yval];
sizemod = length(ymod);
ypred = [ypred_mod;ypred_val];

rmod = model.detail.ssq(n,5);
coeffq = corrcoef([ypred_val yval]);
rval = 100*coeffq(2,1)*coeffq(2,1);
coeffcum = corrcoef([ymeasured ypred]);
rcum = 100*coeffcum(2,1)*coeffcum(2,1);
```