Università degli Studi di Padova

Dipartimento di Ingegneria Industriale

Corso di Laurea Magistrale in Ingegneria Chimica e dei Processi Industriali

Tesi di Laurea Magistrale in Ingegneria Chimica e dei Processi Industriali

Real-time determination of optimal switching times for a new H_2 production process with CO_2 capture: a supervised learning approach

Relatore:

Prof. Massimiliano Barolo

Correlatori:

Prof. Martin van Sint Annaland

Dr. Marcella Porru

Laureando:

Luca Zanella

1147626

Anno accademico 2017 - 2018

To my parents

Abstract

THE CALCIUM-COPPER LOOPING PROCESS is a new concept that combines H_2 production via reforming of natural gas with inherent CO_2 capture, using CaO as CO_2 sorbent and a Cu/CuO chemical looping combustion for sorbent regeneration. The process comprises three dynamic stages, conducted sequentially in a packed bed reactor.

This thesis addresses the determination of the optimal duration of each stage; this is an important requirement to match satisfactory operating targets. For the purpose, economic and quality criteria (e.g. maximization of revenues from syngas) are proposed. Switching from one stage to the other is triggered by the fulfilment of these criteria. Due to disturbances entering the system, an offline determination is unsuitable to guarantee optimal stage duration. The estimation accuracy can be substantially improved by employing a real-time approach. As switching criteria rely on time-profiles of some key variables whose measurements are often delayed, especially outlet compositions, an inferential model is needed. With this in mind, Gaussian Process Regression (GPR) models are employed as soft sensors. They are trained on historical data and employed for predictions during the future process cycles.

GPR is a powerful, relatively recent method of supervised learning. Its application is simpler compared to popular machine learning methods such as artificial neural networks, since no network architecture, number of units and regularization scheme need to be specified. GPR is found to provide accurate estimations also during the first stage of the process, which is characterized by quite a large variability among the experimental runs. Once a switching criterion is chosen, the actual stage duration is compared to the average of the stage durations computed offline and the real-time determination using GPR. The proposed method is capable to infer stop criteria with an accuracy of $\pm$ 5%. The obtained results confirm the potential of GPR-based soft sensors as a mean for reliable real-time process monitoring.

Riassunto

L'accordo di Parigi, ratificato il 12 Dicembre 2015, prevede di limitare l'aumento della temperatura media globale al di sotto dei 2 °C rispetto ai livelli preindustriali. Si ritiene che le conseguenze dei cambiamenti climatici avranno ripercussioni non solo sul piano ambientale, ma anche sul piano economico, causando perdite tra il 5 e il 20% del prodotto interno lordo (PIL) globale annuo. Al contrario, è previsto che porre in atto misure mitigative comporterebbe costi pari solamente all'1% dello stesso (Stern, 2007). Tra le varie cause del surriscaldamento globale, spiccano le emissioni di CO_2, le quali costituiscono la prima causa dell'effetto serra. Ne consegue l'importanza di sviluppare processi sostenibili, contraddistinti da una ridotta impronta di carbonio (*carbon footprint*). La sostituzione di combustibili tradizionali con risorse rinnovabili appare un processo troppo lungo per contribuire al raggiungimento della soglia dei 2 °C nei tempi richiesti. Risulta pertanto necessario rivolgere l'attenzione ad alternative efficaci nel breve periodo. In questa tesi, l'attenzione è rivolta ad un processo che integra la produzione di H_2 con la cattura di CO_2 (Abanades *et al.*, 2010). Questa tecnologia si basa sul *reforming* di gas naturale in presenza di CaO che funge da sorbente della CO_2. La rigenerazione di CaO è ottenuta sfruttando il *loop* chimico Cu/CuO, allo scopo di ottenere l'energia necessaria. Il reattore utilizzato è un letto fisso, il cui solido è costituito da CaO, Cu e un catalizzatore a base di Ni, supportato su di un carrier di Al_2O_3. Il processo, condotto adiabaticamente, si articola in tre stadi, eseguiti sequenzialmente: (A) *sorption enhanced reforming*, (B) ossidazione del Cu e (C) rigenerazione di CaO contemporaneamente alla riduzione del CuO (Martini, Berg, *et al.*, 2016).

Ogni stadio di processo è intrinsecamente dinamico. Di conseguenza, risulta di interesse identificare la durata ottimale di ciascuno di essi, al fine di raggiungere i target operativi ottimali, in primis la redditività d'impianto. Questo è di interesse specie nell'ottica di un futuro *scale-up* del processo.

In questa tesi, si propone una metodologia che mira a determinare l'intervallo di tempo dopo il quale effettuare lo *switch* da uno stadio al successivo. L'approccio seguito è totalmente *data-derived*. Come primo passo, vengono proposti alcuni criteri di identificazione del tempo ottimale di *switch*. Tali criteri sono applicati offline a diversi set di dati, ottenuti attraverso una precedente campagna sperimentale impiegando un *set up* su scala di laboratorio. Per quanto riguarda lo stadio A, i valori ottimali di durata risultano particolarmente dispersi. Ciò rende necessario un loro impiego in tempo reale (*real-time*). Per questo è richiesta la disponibilità dei profili temporali di alcune variabili,

ad esempio composizioni, tipicamente soggette a ritardi nella misura. Per ovviare al problema, è richiesto un modello inferenziale che consenta una stima in tempo reale di tali *variabili primarie*, sulla base del valore assunto da *variabili secondarie*, facilmente misurabili online.

In questo lavoro, si propone l'uso della regressione Gaussiana, allo scopo di sviluppare sensori virtuali per la stima dei profili temporali delle variabili d'interesse, in primis le composizioni in uscita dal reattore. I processi Gaussiani possono essere impiegati per apprendimento supervisionato: essi consentono di condurre una regressione senza bisogno di ricorrere a parametrizzazioni esplicite delle funzioni cercate. Il metodo si fonda sui principi dell'inferenza Bayesiana: si pone una distribuzione a priori di carattere Gaussiano sulle possibili funzioni e la si converte in una distribuzione a posteriori, una volta acquisiti i dati. La media della distribuzione a posteriori corrisponde al valore predetto, mentre la varianza rappresenta l'intervallo di confidenza dello stesso. Il metodo rende possibile lo sviluppo di modelli flessibili, adatti alla rappresentazione di sistemi complessi. Tradizionalmente, altri metodi di *machine learning*, quali le reti neurali artificiali, sono stati utilizzati per la realizzazione di sensori virtuali. L'impiego dei processi Gaussiani a sistemi complessi, appare tuttavia più agevole, in quanto questi non richiedono di specificare l'architettura della rete o il numero delle unità, ma soltanto il tipo di funzione covarianza.

Una volta concluso l'addestramento (*training*) del modello di regressione, la qualità delle predizioni ottenute è stata testata su di un sottoinsieme dei dati acquisiti, detto sottoinsieme di *test*. Nella maggior parte dei casi, i modelli ottenuti forniscono predizioni soddisfacenti, specie per quanto riguarda le variabili degli stadi A e C. Inoltre, l'impiego dei modelli in tempo reale con finalità predittive fornisce performance soddisfacenti. Tuttavia, in una reale applicazione impiantistica del processo, è possibile che si verifichino variazioni nell'attività catalitica o della morfologia del solido, dovute a fenomeni di invecchiamento o avvelenamento al decorrere dei cicli del processo. É lecito attendersi che tali fenomeni determinino variazioni nei profili temporali delle variabili impiegate nella definizione dei criteri adottati, con conseguente disallineamento tra il processo e il modello. In questi casi può rendersi necessario eseguire l'addestramento di nuovi modelli. Questo problema è stato affrontato solo parzialmente nel presente lavoro, a causa della difficoltà nel trovare un indicatore quantitativo affidabile per stabilire univocamente quando il riaddestramento sia necessario o no.

La riscontrata efficacia dei modelli di regressione Gaussiana per la realizzazione di sensori virtuali, fa sì che questo metodo risulti appetibile per essere esteso a diversi processi, con finalità di monitoraggio e controllo.

Contents

List of Figures

List of Tables

Introduction

The Paris agreement of December 2015 established a target of 2 °C as the maximum increase in Earth's temperature by 2100 (*Paris Agreement* 2015). Global temperature will keep on rising by the end of the century (See Fig. 1 on the following page). Besides environmental reasons, the consequences of climate change are foreseen to cause economic losses between 5% and 20% of the annual gross domestic product (GDP). Conversely, mitigating actions are expected to cost 1% of global GDP each year (Stern, 2007). Since greenhouse effects are mainly due to anthropogenic CO_2 emissions (Archer, 2005) a great effort should be put in the development of processes with a low carbon footprint. Limiting CO_2 concentration in atmosphere below 450 ppm by 2100 is foreseen to maintain global warming below 2 °C. In principle, this goal could be achieved by replacing fossil fuels with renewable resources, but due to their limited deployment, this solution seems promising only in the medium-to-long term. In this scenario, the development of short-to-midterm alternatives still relying on (a better use of) fossil fuels is necessary (Medrano, 2017). With this in mind, process intensification, that is the design approach leading to cleaner, safer and energy-efficient technology (Reay, Ramshaw, and Harvey, 2013), will play a major role.

The Calcium-Copper (Ca-Cu) looping process is a novel concept for hydrogen production (Abanades *et al.*, 2010). It combines the capture of CO_2 on a solid material with the reforming of natural gas, the benchmark technology from which about 48% of the hydrogen produced worldwide is obtained (Martini, Berg, *et al.*, 2016). The process is carried out under adiabatic conditions in a packed bed reactor, whose solid phase consists in CaO, employed as a CO_2 sorbent, Cu and a Ni-based catalyst usually supported on a Al_2O_3 carrier. The process comprises three stages:

Stage A hydrogen production via steam reforming and simultaneous carbonation of the CaO on solid bed to $CaCO_3$ with the by-product CO_2;

Stage B oxidation of Cu with air;

Stage C regeneration of CaO and simultaneous reduction of CuO to Cu with a fuel gas. CO_2 released from this stage is captured and stored.

As reactions proceed, in each step the solid phase undergoes some modifications along time, both in temperature and composition. Due to the intimate correlation between gas and solid phase, changes in the solid bed also determine variations in the outlet gas composition. Therefore, each stage is inherently dynamic.

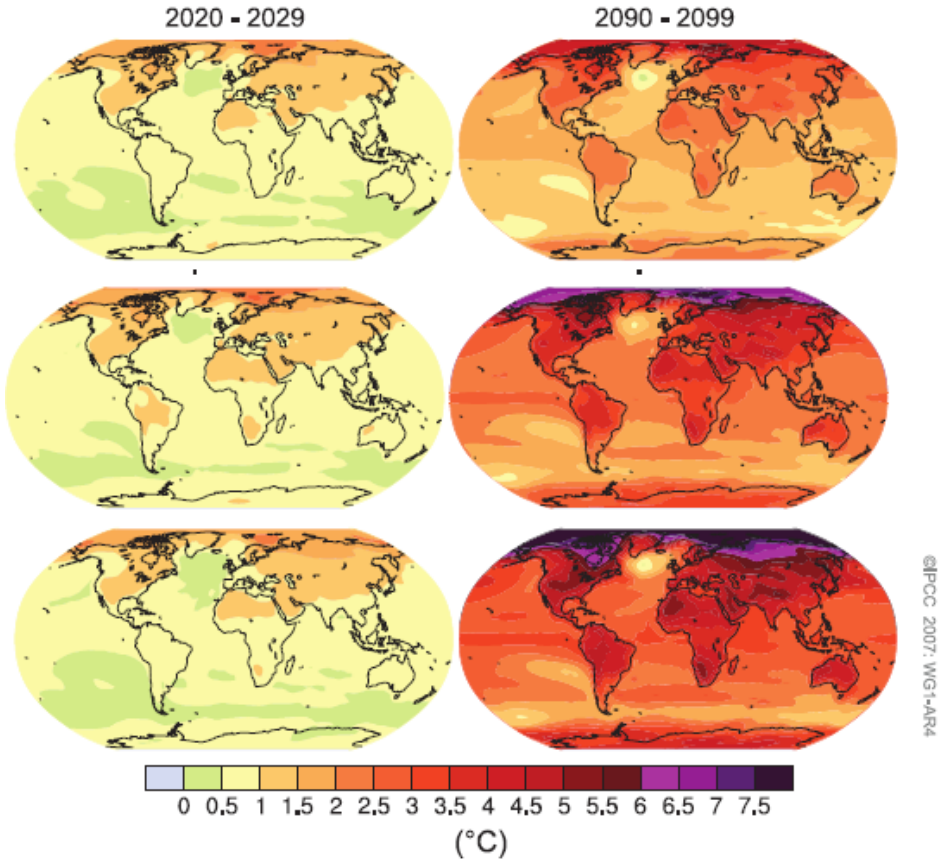


Figure 1: Three scenarios showing the projection of surface Earth's temperatures in the periods: (left) 2020-2029, (right) 2090-2099, with respect to Earth's temperature in 1980-1999. Image from ICCP, 2007.

The main purpose of this thesis work is to propose a methodology to assess the duration of each stage. When in operation, the switching from one reactor mode to the following should be performed at a moment such that the process is optimally profitable. For example, consider to prolong the duration of stage C after the solid bed has already been regenerated for stage A: this would make the operation economically inefficient, since the reactor would not be available to begin stage A, hence to produce hydrogen (the valuable product). Analogous considerations can be done for steps B and A, for which a stage prolongation might even lead to a off-spec grade hydrogen. Conversely, a premature stage termination would cause the solid bed to be in an inadequate state for the beginning of the following step. Notice that this task differs from classical real-time optimization, where the set of operating conditions that satisfy an economic objective is found. In this work, operating conditions are assigned *a priori*, while the optimal stage duration to maintain the operation economical is determined.

As a first, simplistic solution, one may consider to set a predetermined duration for each stage. This approach, based on past experience is acceptable for lab-scale investigations, but unfeasible in the industrial practice. In this case defining a proper

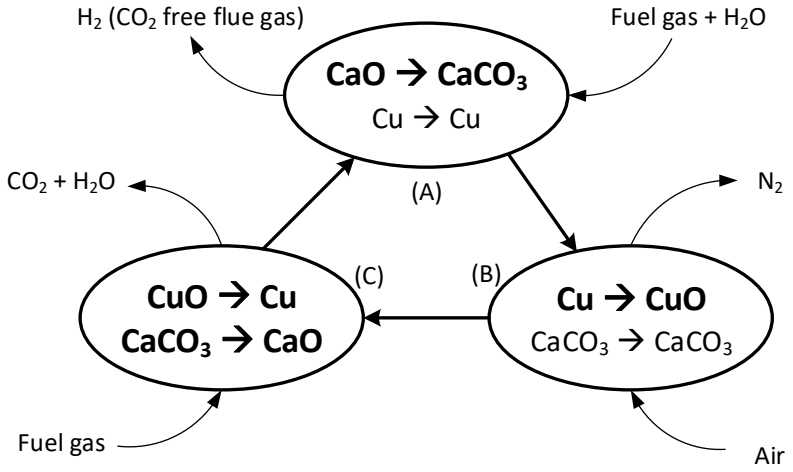


Figure 2: General scheme of the Ca-Cu three steps chemical loop. Image adapted from Abanades *et al.*, 2010.

criterion, whatever it may be (e.g. switching when some key process variables assume certain values or when the optimum of an economic objective function is found), is mandatory to match optimal operating targets. Once defined, a switching criterion can be used in two modes, either offline or in real-time. The first mode is carried out through the following steps:

- a suitable switching criterion is identified;
- process simulations or experiments at the operating conditions of interest are conducted to assess the time-profile of key variables. This way, the stage duration for the given set of operating conditions is computed.
- when the real process is in operation, the duration of each stage is scheduled according to the pre-determined value for the same set of operating conditions.

Experiments or detailed models can be used in this approach. However, an offline scheduling of stage duration does not allow to deal with disturbances entering the system, process uncertainties and failures. This is limiting in the view of a future process up-scale. For this reason, in general an online determination is preferred. The real-time approach is performed according to the following steps:

- a proper switching criterion is defined, as in the offline case;
- when the process is in operation, the value of key variables employed by the switching criterion is acquired at any time instant. This allows to assess the course of the stage and to terminate it when the criterion is satisfied.

Unfortunately, in the industrial practice not all the key variables are available in real-time. The conventional strategy to approach monitoring problems relies upon field instrumentation, hardware-sensors and off-line laboratory analysis (Corona *et al.*, 2012). Lack of measurements is usually attributed to the cost of installation of some instrumentations or, more often, to the insufficient automation of sensitive analysis (e.g. instruments with slow dynamics). In the specific case, composition measurements are often affected by substantial time delays. In such situations, the development of a *soft sensor* is required. Soft sensors are softwares that estimate the value of hard-to-measure primary variables through online measurements of easy-to-measure, secondary variables.

Different soft-sensing techniques are reported in literature. The following classification usually holds, based on the type of model employed:

1. model based;

2. data derived.

Model based soft sensors use *first-principle* models to make inferences. These models are based on conservation equations and constitutive relations (Seborg *et al.*, 2010 and Porru, Alvarez, and Baratti, 2013). Examples include mechanistic models and filtering techniques, such as the extended Kalman filter routinely employed in the process industry. They are often referred to as *knowledge-driven* models since they provide a systematic way of organizing the physical understanding about the process (Bonvin *et al.*, 2016). Main advantages of first-principle models are that it is relatively easy to interpret their results and that they can be used for extrapolation. However, the development of reliable first principle models often requires a deep understanding of the phenomena involved. Though this is not a problem for simple systems, it becomes an issue in case of complex processes, for which acquisition of adequate knowledge is sometimes impractical. In this case, lack of knowledge even in some minor aspects of the underlying physics might result in poor model performances. Conversely, even assuming that the set of modelling equations could be written for such processes, model complexity might make their solution too computationally demanding for their use in real-time. This is the problem faced with the Ca-Cu looping process.

Data-derived soft sensors employ data-driven models to make inferences. This model class is very broad. Examples include principal component analysis (PCA), partial least-squares regression (PLS) and machine learning (ML) models. As opposed to the first-principle ones, data-derived are not knowledge-driven. They are obtained via system identification employing input and output data, without any knowledge about the physical characteristics of the process. As a result, the information these models provide is the only one encoded in the data used to realize it. Therefore, empirical models do not generate new information. The main drawback in the use of these

models relies in their limited extrapolation capabilities. Furthermore, usually their parameters do not have a clear physical meaning, making them appear as *black boxes*. Last, they require the physical process under investigation to be available, so that experiments can be carried out, data collected and the model identified (Bonvin *et al.*, 2016). Quite a remarkable amount of data is usually needed to build them, but this does not seem a limit to their application, as vast quantities available in the industry nowadays. Resorting to this class of model is necessary when the underlying physics of the process is not fully understood.

In the present work, *Gaussian Process Regression* (GPR) models were employed as soft sensors to serve as estimators of the key variables employed to compute the optimal stage duration. GPR is a relatively new method of *supervised learning* (Rasmussen and Williams, 2006). The commonly agreed definition for supervised learning is:

> The machine learning task of learning a function that maps an input to an output based on example input-output pairs (Russell and Norvig, 2010).

Several machine learning algorithms have been developed in the recent years, including artificial neural networks (ANN), support vector machines (SVM), fuzzy logic and genetic algorithms. Due to the recursive nature of the Ca-Cu looping process, employing a machine learning algorithm to model it seemed reasonable: as new data are continuously collected during operation, they can be employed to train models with increasingly reinforced predictive capabilities. This might prove useful if variation in catalyst activity or solid morphology occur, resulting from aging or poisoning, after several runs. These might cause drifts in the time profiles of key variables, leading to increased process-model mismatches. Within the multitude of machine learning methods, the choice of the most appropriate one is dictated by the problem. GPR is based on probabilistic kernel functions as an alternative to the typical multilayer perceptrons used in artificial neural networks (Gupta, 2018). In GPR, a process is modelled without giving explicit parametrization to the underlying function. Conversely, a distribution over the possible functions that are consistent with the observed data is found. As this is a Bayesian method, a prior Gaussian distribution is placed over the functions and is then converted into a posterior once data are acquired. Though Gaussian Processes have been raising increasingly attention in the last years, very few applications in soft sensing have been reported (Mei *et al.*, 2016).

Gaussian Processes have very useful features. First, they act as *smoothing devices* (Rasmussen and Williams, 2006) and allow to reduce overfitting, by incorporating an automatic trade-off between model fitting and complexity (Rasmussen and Williams, 2006). Very good fits can be obtained without even the need for cross-validation. Unlike linear regression, there is no need to postulate a fitting function upfront. Compared to other machine learning methods such as ANN, whose application requires the selection of

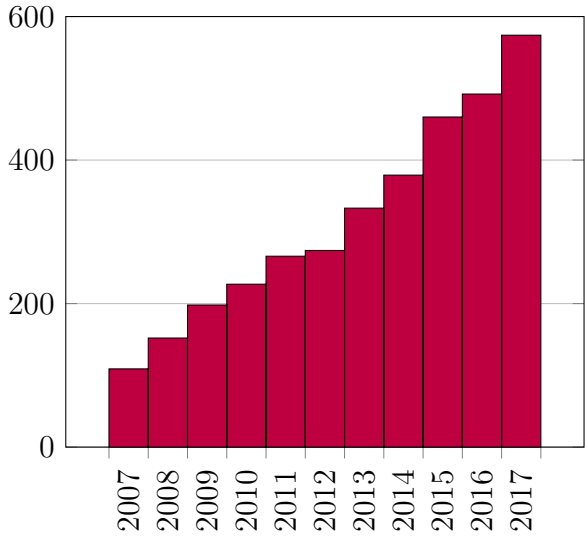


Figure 3: Number of publications on Scopus with "Gaussian Process Regression" in the title during the past ten years (2007-2017).

several factors, including network architecture, number of units, regularization scheme, GP relies on the sole assumption of a covariance function. In addition, a big advantage over both ANN and SVM is that GPR gives confidence intervals for the predictions. As all these characteristics are desirable, when it comes to design a reliable soft-sensor, GPR seemed an appropriate modelling tool for the goal of this work.

The present work follows an approach which is mainly based on data. All the employed set of data were collected by the Chemical Process Intensification group at Eindhoven University of Technology TU/e, where an experimental setup for the Ca-Cu looping process is available. The entire thesis work was carried out within the group of Control Systems at the same university. As a first step, detailed data analysis was performed. Besides providing a useful insight into the process, investigation of the time profiles of several variables (outlet mole fraction, axial and outlet temperatures) is useful to identify suitable switching criteria. Several criteria are proposed for the aforementioned process stage durations and tested offline employing the available datasets. The choice of the most suitable one is up to the decision-maker. Finally, the real-time determination of each stage duration is performed employing some of the criteria proposed. At the purpose, GPR models trained on a subset of the available datasets, are employed as soft-sensors and their performances are evaluated on test datasets. Outlet composition of gas phase is inferred by means of the GPR model trained on operating variables (i.e. time, pressure, inlet composition).

The thesis is organized as follows:

Chapter 1 – The Calcium - Copper Looping Process introduces the features of the Ca-Cu looping process. Each stage is described with some details. For sake of clarity, some key concepts are displayed resorting to the results of

some simulations by Martini, Berg, *et al.*, (2016), Martini, Martínez, *et al.*, (2017) and Berg, (2015). Finally, a possible industrial process implementation is briefly described, to give the reader an example of what a future up-scale could consist of.

Chapter 2 – Data analysis includes a description of the time profiles obtained from experiments. This investigation is used as a basis to gain further insight into the dynamic behaviour of process and to expand the discussion of Chapter 1.

Chapter 3 – Offline Determination of the Switching Time proposes several criteria to assess the optimal stages duration. The switching time is computed offline for all the experimental runs available. Processed data are employed for the purpose.

Chapter 4 – Bayesian Inference and Gaussian Processes introduces key concepts of Bayesian inference and Gaussian Processes used for regression. Key concepts used in model selections are described. The procedure employed to predict the outputs is also presented.

Chapter 5 – Online Determination of the Switching Time concerns the use of GPR models as soft sensors. The procedure employed to perform model training is described. The proposed algorithms to assess the duration of each stage are shown. Finally, the advantages of employing a real-time determination of the switching time is displayed.

Conclusions and Outlooks are summarized in this last chapter. The main outcomes of the present work are displayed, assessing advantages and drawbacks of the methods developed. At the end, some outlooks on the use of GPR models as soft sensors are presented, together with future challenges and perspectives.

Chapter 1

The Calcium - Copper Looping Process

THIS CHAPTER provides the reader with a brief overview on the role of hydrogen in today's economy. Later, the Ca-Cu looping process for H_2 production is also introduced. Section 1.1 shows hydrogen market segmentation, while Section 1.2 describes the main features of the process stages. The final part of this chapter shows an advanced process scheme with potential industrial applications.

1.1 Background

The world hydrogen demand in 2010 has been estimated in 43 Mtons. A growth up to 50 Mtons is expected by 2025 (Fraile *et al.*, 2015). Main fields of application include industry and mobility; the so-called *Power to Gas* technology is also raising attention. *Industry* covers 90% of the market share and a 3.5% annual growth in the H_2 consumption is foreseen in the next years. The industrial subsegments of major interest are (Fraile *et al.*, 2015):

- chemical industry (63% of the total demand). Most of H_2 is employed for the production of ammonia (84%). Methanol (12%), nylon (2%) and polyurethane (2%) follow;
- refineries (30% of the total consumption). H_2 is used mostly for hydrocracking and hydrotreating processes;
- metal processing (6%);
- others (1%), especially semiconductors and glass production, for which an annual growth of 2.3% and 3.4% is foreseen in the next years.

Besides its use as a chemical, H_2 is becoming also as a promising energy carrier, as its final use does not imply CO_2 release to the atmosphere. Worldwide energy demand will increase by 48% in the period 2012-2040, as a result of the growth in global population (EIA, 2016). As a large part of the energy demand is required to satisfy the need from large scale industrial processes, H_2 is receiving increasing attention also in this regard.

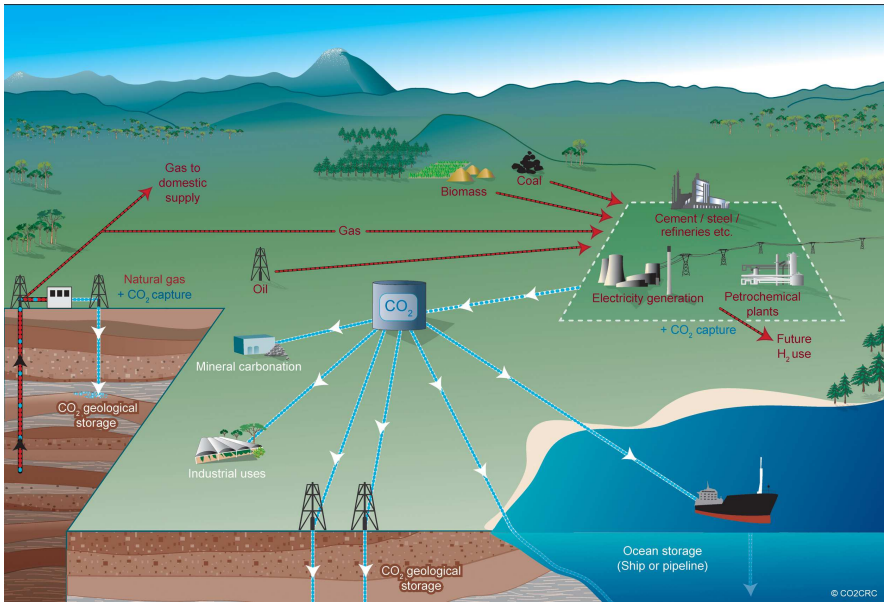


Figure 1.1: Schematic overview of a Carbon Capture and Storage (CCS) process (from ICCP, 2007).

Mobility is expected to be the fastest H_2 growing market in the period 2025-2030. Transport emissions are currently responsible for 32% of annual greenhouse gases (GHG) emissions in Europe. The target established by EU Commission is to attain a 80% reduction in CO_2 emissions by 2050, corresponding to a decarbonization of 95% [1]. This situation is pushing several EU governments, like those of Germany, France and Scandinavian countries to speed up the development of hydrogen-based mobility.

Power To Gas consists in the injection of H_2 into the natural gas pipeline network, with a blending ratio up to 15%. Given that: (i) this solution might contribute to a reduction of GHG emissions from the energy system; (ii) natural gas demand is increasing annually at a pace of 1% per year and (iii) injecting H_2 into pipelines requires only minor modifications of the existing network, a growth in this sector is also expected.

Despite the potential of H_2 as a substitute from fuels in energy production and mobility, nowadays about 80% of H_2 employed in industry is produced from fossil fuels via steam reforming (Helmi, 2017). Therefore, conventional technologies lead to relevant CO_2 emissions. The novel concept of Ca-Cu looping technology allows inherent CO_2 capture within an enhanced steam reforming framework (See Section 1.2 for the process details). Once captured, CO_2 can be concentrated and stored in sites, such as geological formations, oceans, saline formations, depleted oil or gas reservoirs (Fig. 1.1). This way, CO_2 release into the atmosphere is avoided. This technology is referred to as Carbon Capture and Storage (CCS) . As explained in Section 1.2, since Ca-Cu technology allows highly concentrated streams of CO_2, carbon storage fits perfectly with this process. Actually, there exist several applications of CO_2, besides storage. These include its

[1] McKinsey, Roadmap 2050

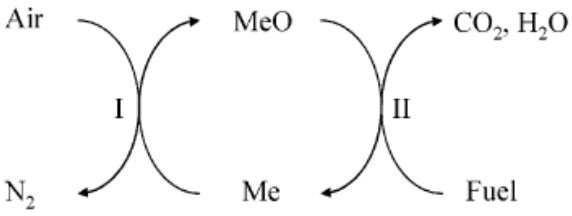


Figure 1.2: Schematic of the chemical looping process: (I) oxidation, (II) reduction. Image from Noorman, Sint Annaland, and Kuipers, 2007.

use in greenhouses, (as it is done by the Shell refinery in Pernis), or as a feedstock for growing algae. Other CO_2 applications as a starting molecule for chemical synthesis have been envisaged (Tuinier, 2011), but CO_2 chemical stability and scarce reactivity require severe and uneconomical process conditions. For this reason, storage will be the simplest solution to reduce CO_2 emissions.

1.2 Process concept

The Ca-Cu chemical looping reforming is a novel process for high purity H_2 production from natural gas with integrated CO_2 capture (Abanades *et al.*, 2010), envisaged as a future solution for both power generation and hydrogen production (Martini, Martínez, *et al.*, 2017). This technology relies on the traditional reforming of natural gas, the most economical method to produce H_2 on a large scale (J. R. Rostrup-Nielsen and T. Rostrup-Nielsen, 2002). This is a well-established technology, but it still presents some drawbacks: high temperatures are needed in the reforming section and H_2 purity of the reformer gas is not high, making additional purification steps necessary. In addition, traditional reforming results in quite high CO_2 emissions. One may consider separating CO_2 from flue gases, but this would involve high purification costs as well as a significant decrease in the overall energy efficiency of the process, due to the dilution of the CO_2 stream. Although worldwide CO_2 emissions associated to large scale H_2 production only contribute to the 3% of global emissions (Meerman *et al.*, 2012), the fast rise in H_2 demand will make the percentage increase in the incoming years. Therefore, the development of fossil fuels-based hydrogen production processes with integrated CO_2 capture remains a promising alternative.

The Ca-Cu looping reforming consists in three stages, cyclically carried out in a packed bed reactor. Switching across stages is performed by changing operating conditions in the fixed bed, mainly temperature and pressure, and alternating the gaseous reactants to be fed. Ideally, operation should be adiabatic. The bed is made up by three different solids: a Ni-based catalyst for SMR, CaO as CO_2-sorbent material and Cu which is involved in CaO regeneration. A detailed description of the three stages is further presented.

1.2.1 Stage A - Sorption Enhanced Reforming

In stage A sorption enhanced reforming (SER) takes place. The goal of this stage is to maximize H_2 purity and CH_4 conversion. The reactor is fed with CH_4 and H_2O in a given steam-to-carbon ratio (S/C) . Clearly, in its industrial implementation the feed would not consist in pure CH_4 but natural gas. The main reactions yielding to H_2 production are steam-methane reforming (SMR) and water-gas shift (WGS), displayed in (R1.1) and (R1.2), respectively. Both reactions are heterogeneously catalysed over Ni:

$$CH_4 + H_2O \rightleftharpoons 3\,H_2 + CO, \quad \Delta_r H^{\ominus} = 206\,\mathrm{kJ\,mol^{-1}} \tag{R1.1}$$

$$CO + H_2O \rightleftharpoons H_2 + CO_2, \quad \Delta_r H^{\ominus} = -41\,\mathrm{kJ\,mol^{-1}} \tag{R1.2}$$

Typical temperatures of the solid bed (T_s^{out}) and inlet gas (T_G^0) result from a trade-off between the need to push the equilibrium of (R1.1) to the right hand side, and to not undermine (R1.2). A reasonable operating process window is $T_s^{\mathrm{out}} = 650\,^{\circ}\mathrm{C}$, $T_G^0 = 575\,^{\circ}\mathrm{C}$, respectively.

The global process is endothermic, resulting from the strong endothermicity of (R1.1) and the slight exothermicity of (R1.2). Therefore, high heat duties are required to boost H_2 production. Both reactions are equilibrium limited, hence reaching high CH_4 conversion and H_2 purity is challenging. To shift the equilibrium towards products, CO_2 is captured by CaO:

$$CaO + CO_2 \rightleftharpoons CaCO_3\cdot \quad \Delta_r H^{\ominus} = -178\,\mathrm{kJ\,mol^{-1}} \tag{R1.3}$$

Combination of (R1.1), (R1.2) and (R1.3) gives the name SER to this stage. As CO_2 removal pushes equilibrium of (R1.2) and, consequently, the one of (R1.1) to the right-hand side, both CH_4 conversion and H_2 throughput increase. (Besides CO_2 adsorption, competitive approaches exist to shift equilibrium towards products, such as separation of H_2 from product gases using selective membranes (Kikuchi, 1995), while the second is attained by including a CO_2 sorbent into the solid bed (Shimizu *et al.*, 1999). In addition, (R1.3) makes the overall stage slightly exothermic, thus providing the heat required to trigger (R1.1):

$$CaO + CH_4 + 2\,H_2O \rightleftharpoons 4\,H_2 + CaCO_3 \quad \Delta_r H^{\ominus} = -14\,\mathrm{kJ\,mol^{-1}} \tag{R1.4}$$

This makes the operation much less intensive compared to standard SMR. Moreover, energy costs associated to adsorption are expected to be lower than typical post-combustion CO_2 capture technologies, such as scrubbing with amine-based solvents Tuinier, 2011.

As stated in the Introduction, each stage is inherently dynamic. At the purpose, Fig. 1.3 shows simulated time-profiles of $x_{H_2}^{out}$ from literature (Berg, 2015). These results were obtained by simulating a system of unsteady partial differential equations describing conservation laws along time and axial coordinate. The model is pseudo-homogeneous: reaction, axial mass and heat dispersion are simulated, while radial and mass dispersion as well mass and heat transfer resistances between solid and gas phase are neglected. Consequently, temperature is homogeneous between the two phases. Detailed description can be found in Berg, (2015) and Martini, Berg, *et al.*, (2016). In Fig. 1.3, three different H_2 purity plateaus are distinguishable

- pre-breakthrough: $x_{H_2}^{out}$ is at the highest purity reached, usually above 95%, when most of CaO on the solid bed is still not carbonated, being available for CO_2 capture;
- first breakthrough: $x_{H_2}^{out}$ sets to a lower value. During this time interval SER is still effectively occurring because of the leftover CaO on solid bed.
- second breakthrough: $x_{H_2}^{out}$ strongly decreases, because no SER is quantitatively taking place since almost no CaO is left.

Notice that Fig. 1.3 also shows that sizes of first and second plateaus augment with the initial fractions of CaO, while no changes in the third plateau are observed. This is because the third plateau refers to H_2 purity once the bed has been fully converted.

Stage A is considered over when CaO has been fully carbonated. Then, CaO has to be regenerated. Typical techniques suggested in literature are either pressure swing or flow of high temperature gases through the solid bed, to release CO_2 captured. However, these are characterized by low thermal efficiencies and high equipment costs (Berg, 2015). In this work, the regeneration of CaO is achieved by means of the Ca-Cu chemical looping (Abanades *et al.*, 2010). The chemical looping is carried out in stages B and C, which are described in the following subsections.

1.2.2 Stage B - Oxidation

Stage B is preparatory for the subsequent regeneration of CaO. During this stage, the first part of the loop takes place. The main reaction involved is Cu oxidation:

$$\mathrm{Cu} + \frac{1}{2}\,\mathrm{O_2} \longrightarrow \mathrm{CuO} \quad \Delta_r H^{\ominus} = -156\,\mathrm{kJ\,mol^{-1}} \tag{R1.5}$$

The goal of this stage is to maximize the amount of Cu oxidised, while minimizing $CaCO_3$ calcination, since the formed CO_2 could not be captured owing to its dilution with N_2. The reactor is fed with pressurized diluted air. Typical T_G^0 is about 350 °C. Both oxidant

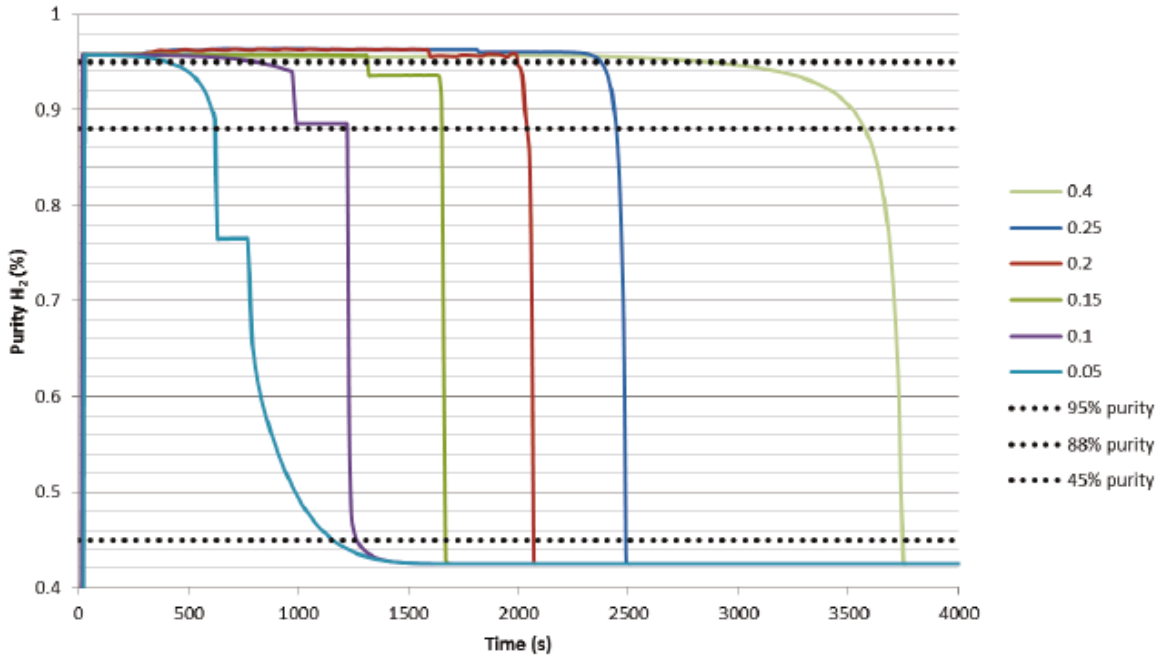


Figure 1.3: Comparison of H_2 purity in the product gas in stage A (SER) for different initial bed compositions obtained via simulation through a 1-D pseudo-homogeneous model. In all cases, the three regions of H_2 purity are clearly noticeable (from Berg, 2015).

dilution, limited T_G^0 and operation under pressure are desired to decrease the entity of CO_2 release. Indeed, flowing considerable amounts of inert gases limits temperature increase, thus diminishing the risk of hotspots formation that may drive to undesired calcination, solid agglomeration or other undesired side-reactions. Analogously, high pressure reduce CO_2 releases. O_2 dilution can be achieved by recycling gases from this stage and mixing them with air. Since effluents from the reactor contain some CO_2, recycling part of this stream has the double benefit of saturating the amount of CO_2 at the equilibrium and reducing the actual CO_2 loss to the sole fraction which is not recycled back.

1.2.3 Stage C - regeneration

In stage C, regeneration of the CaO is carried out:

$$CaCO_3 \rightleftharpoons CaO + CO_2 \quad \Delta_r H^\circ = 178\,\mathrm{kJ\,mol^{-1}} \tag{R1.6}$$

The goals of this stage are the regeneration of the CO_2 sorbent, keeping CO_2 released as pure as possible, and the guarantee of operation being energy neutral. This is where the chemical looping comes to the aid. Heat needed to trigger calcination is provided by reduction of the CuO formed in the previous stage. One or more reducing species

(usually a mixture of all the ones available in the process) are fed to the reactor:

$$CuO + H_2 \longrightarrow Cu + H_2O \quad \Delta_r H^\ominus = -86\,\mathrm{kJ\,mol^{-1}} \tag{R1.7}$$

$$CuO + CO \longrightarrow Cu + CO_2 \quad \Delta_r H^\ominus = -127\,\mathrm{kJ\,mol^{-1}} \tag{R1.8}$$

$$4\,CuO + CH_4 \longrightarrow 4\,Cu + CO_2 + 2\,H_2O \quad \Delta_r H^\ominus = -170\,\mathrm{kJ\,mol^{-1}} \tag{R1.9}$$

Absence of N_2 in the feed is required, otherwise the separation would get increasingly difficult. The reactor temperature during this stage should be maintained in between 800 °C and 90 °C to enhance $CaCO_3$ decomposition while limiting possible CaO deactivation or CuO undesired reactions (Abanades *et al.*, 2010). The pressure should be at most atmospheric to increase CO_2 desorption. The reducing mixture composition determines the Ca/Cu-ratio required for energy neutral operation. Ca/Cu ratios of 1.3, 1.5, 3.5 for CO, H_2 and CH_4 as reduction gases are suggested in the literature. These ratios were computed from reaction enthalpies (Berg, 2015). A major advantage of coupling reduction and calcination in the same bed matrix lies in the efficiency at which heat is transferred from Cu particles to $CaCO_3$. This makes operation at moderate temperature possible, allowing energy savings and avoidance of special materials in the reactor (Abanades *et al.*, 2010).

Besides calcination and reduction, also (R1.1) and (R1.2) occur, because the gas phase contains all the components needed for SMR and WGS and the bed is catalytic. Thus, the carbon content in the feed gas should be low to avoid CO_2 formation that could inhibit calcination. This stage is considered terminated when all CuO has been reduced. Due to differences in their kinetics, while complete reduction of Cu is rapidly attained, full calcination of $CaCO_3$ is not. Hence, some $CaCO_3$ is expected to be still present at the end of this stage.

The autothermal operation of all reaction stages allows to avoid the use of reactors with large heat transfer surfaces. This is a major advantage compared to traditional technologies for syngas production based on fired tubular reforming processes.

1.2.4 Example: industrial implementation of the process

A possible industrial implementation of the process is presented in this section. This flowsheet (Fig. 1.4) is expected to be part of a power production plant (Martini, Martínez, *et al.*, 2017). A battery of reactors operating in parallel is employed, to guarantee continuous H_2 supply downstream. Natural gas is pre-reformed with steam to decompose heavy hydrocarbons. This avoids coke formation in stage A. To minimize slips of CO_2 from SER, the stage was splitted into A' and A, with intermediate cooling. Feed gas at S/C = 5, $T_G^0 = 575\,^\circ\mathrm{C}$, $P_G^0 = 25$ bar is sent to the reactor for stage A', where

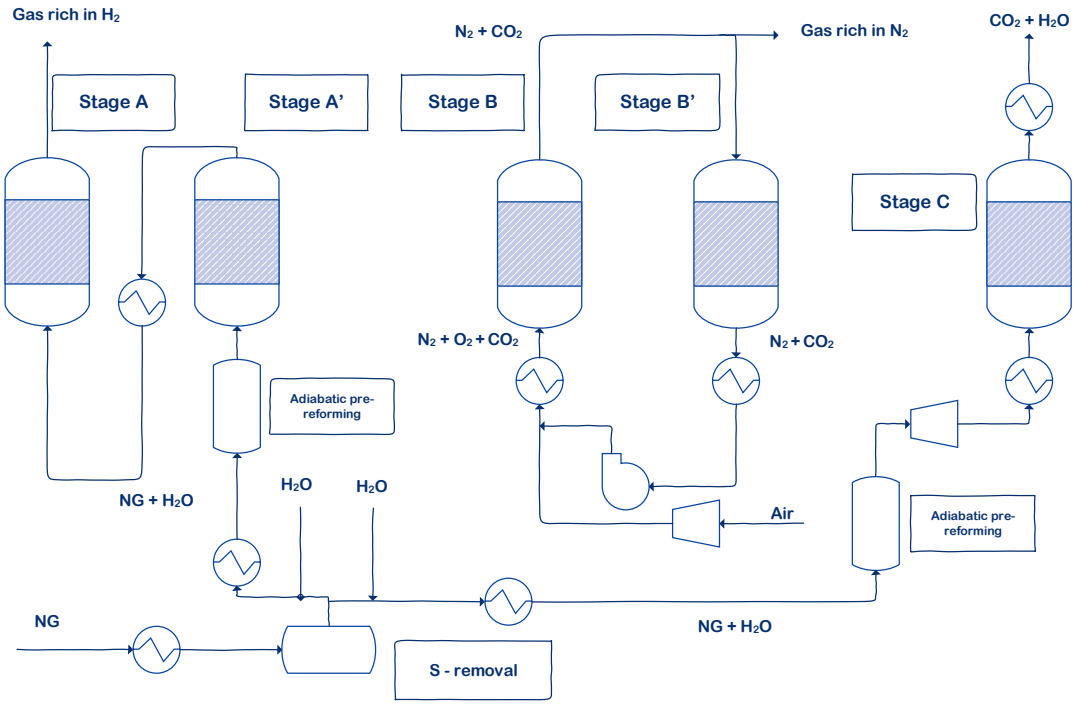


Figure 1.4: Schematic of the process stages involved in one possible industrial implementation of the Ca/Cu chemical looping reforming. Stages A and A': SER; Stage B: copper oxidation; Stage B': preheating; Stage C: solid regeneration.

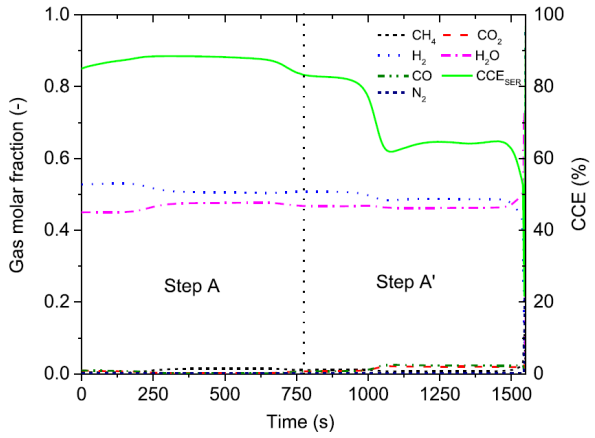


Figure 1.5: $\boldsymbol{x}^{\text{out}}$ and η_{CC} of stage A and A', simulated with a 1-dimensional pseudo-homogeneous model. Plots were taken from Martini, Martínez, *et al.*, 2017.

SER takes place. Effluent gases at about 850 °C are cooled down to 700 °C (carbonation is favoured at low temperature) and sent to stage A, where CO_2 slipped from stage A' is captured. Actually, some carbon losses are still present mainly due to the decoupling between reforming and carbonation, that do not entirely occur at the same reaction front. Indeed, the overall heat produced by the reactions is approximately null, but heat production and consumption do not occur at the same place. First SMR and WGS occur, forming CO_2, that reacts with CaO in a second moment. This causes formation of a high temperature heat plateau, that penalises carbonation, resulting in a decrease of the carbon capture. However, splitting SER into two stages allows a great improvement on this regard, as shown in Fig. 1.5.

In stage B, inlet air at 340 °C is pressurized and diluted with some effluent gases. As reported in Section 1.2, they contain some CO_2. This way, a O_2 dilution at 3% is attained. As displayed in Fig. 1.4, these effluent gases are first flowed through the bed B', at the double purpose of cooling down the (hot) gases before they are mixed with compressed air and heating the solid bed in view of the subsequent stage C for which a high T_s^0 is desired.

Analogously to A, a mixture of natural gas and steam (S/C = 1) is pre-reformed, before being fed to stage C. This configuration introduces some improvements to the process: the mixture of reducing species CH_4, H_2 and CO_2 is formed *in situ*. In addition, endothermic (R1.1) does not occur (except in a marginal way) in the reactor, making bed heating for the subsequent stage A unnecessary. Since the feed to this stage has some carbon content, residual $CaCO_3$ in the initial sections of the bed are unavoidable. The outflowing steam is separated by condensation. Concentrated CO_2 suitable for permanent geological storage is obtained. One major advantage of this technology is that CO_2 has no nitrogen dilution, because direct contact between air and fuel is circumvented through the Cu/CuO couple, that acts a oxygen carrier, being alternatively oxidized and reduced (Noorman, Sint Annaland, and Kuipers, 2007).

Performances of the Ca-Cu looping process are usually assessed in terms of carbon capture efficiency η_{CC}, defined as:

$$\eta_{CC} = \frac{\dot{m}_{\mathrm{CO_2}}^{captured}}{\dot{m}_{\mathrm{C}}^{0}}, \tag{1.1}$$

where $\dot{m}_{\mathrm{CO_2}}^{captured}$ ($\mathrm{kg\,m^{-2}\,s^{-1}}$) is the mass flow of CO_2 captured and $\dot{m}_{\mathrm{C}}^{0}$ ($\mathrm{kg\,m^{-2}\,s^{-1}}$) is the mass flow of carbon species entering the reactor. This process schemes yields an overall carbon capture efficiency (η_{CC}) of 88%. Estimated value for the same conditions without stage splitting yields η_{CC} =82%.

1.3 Summary

The description of the Ca-Cu looping process for H_2 production provided in this Chapter is mostly based on the results of several simulations reported in the literature. However, this process concept was also validated through an extensive experimental campaign by Druijff, (2018). All the data employed in this thesis work were collected during these experiments. They refer to a configuration based on stages A, B and C, with no auxiliary steps. The analysis of these data begins in the next Chapter.

Chapter 2

Data Analysis

THIS CHAPTER provides the reader with a description of the several data employed in the present work. This allows to get a useful insight to the physics of the process, expanding the description given in Chapter 1. Most important for the scope of this work, data analysis allows to identify the most representative variables for the determination of the optimal switching times. Chapter 1 gave a definition of the end time of each stage, based on the state assumed by the solid bed. As an example, stage A is commonly considered over when all CaO has been carbonated to $CaCO_3$. Similarly, stage B and C are conventionally assumed to be concluded when all Cu and CuO have been oxidised and reduced, respectively. However, on-line measurements of the solid bed composition are not available. Hence, the course of each stage should be monitored by means of other variables, such as temperature profiles or gas composition.

As a first step, data were pre-processed to discard the inconsistent ones. Datasets at different operating conditions were analysed for each process stage. As stated in the Introduction, all the experiments were carried out by the Process Intensification Group. A simplified schematic of the experimental set up employed is shown in Fig. 2.1. The reactor is a packed bed of 0.71 m height, 0.063 m diameter, filled with solid particles of approximately $D_p = 2\,\text{mm}$. The initial bed composition is as follows: 17.6% CuO, 54.4% CaO act, 4.0% NiO, and 24.0% Al_2O_3 on mass basis.

Section 2.1 enlists the operating conditions at which data were collected. Sections 2.2, 2.3 and 2.4 show representative profiles for step A, B and C, respectively.

2.1 Experimental Conditions

The following experimental conditions were investigated for the three stages:

A - SER: 62 datasets available at three different steam-to-carbon ratios (S/C): 3, 4 and 5 (mole basis) and inlet gas pressure of 2 bar and 7 bar. In all the experiments, the inputs were:

- feed composition ($\boldsymbol{x}^0$): steam and CH_4, at the given S/C;
- volumetric flow rate of CH_4 in the inlet of approximately 4 liter/min;

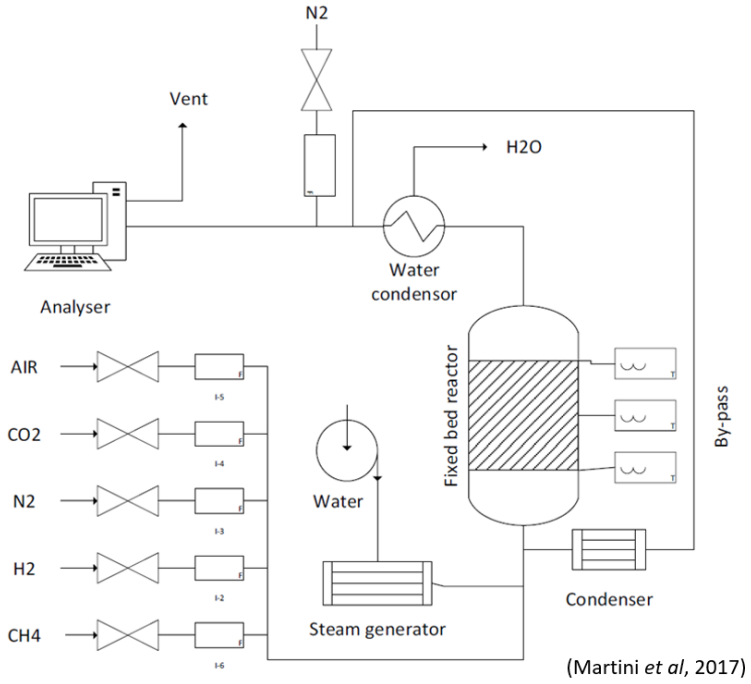


Figure 2.1: Simplified schematic of the experimental set up for the Ca-Cu looping process employed to collect the data used in the present work.

- mean temperature of the inlet gas (T_G^0) in the range 750 - 780 °C;

B - Oxidation: 145 datasets, with inlet O_2 mole fraction of approximately 3%, 5%, 10%. The other inputs were:

- $\boldsymbol{x}^0$: only O_2 and N_2 with the given O_2 mole fraction;
- total volumetric flow rate in the inlet of 100 and 113 L min^{-1}.
- mean temperature of the inlet gas (T_G^0) of approximately 830 °C;
- pressure of the inlet gas (P_G^0) of 3, 4 and 7.5 bar.

C - Reduction: 256 datasets, at different inlet H_2 mole fractions, namely 20%, 40%, 58%. The other selected inputs were:

- $\boldsymbol{x}^0$: only H_2 and N_2 at the given H_2 mole fraction;
- total inlet volumetric flow rate of ≈100 and 125 liter/min;
- temperature of the inlet gas (T_G^0) of 830 °C;
- pressure of the inlet gas (P_G^0) in the range 2.5 - 4 bar.

In all three stages, the following variables were monitored:

- inlet and outlet volumetric flow rates.
- outlet species mole fractions.

- temperature of the solid bed in 20 points along the axial coordinate of the reactor.

- inlet and outlet pressures of the gases.

- temperature of the feed gas.

2.2 SER

Fig. 2.2 shows the outlet profiles of CH_4, CO_2, H_2 and CO dry fractions (on mole basis) along time for three datasets at $P_G^0 = 2$ bar. Notice that a comparison among species mole (or mass) flow rates would allow a more immediate interpretation of the phenomena involved. However, due to the noise associated to flow measurements in this stage (especially for high S/C values), composition plots turned out to be more clear. Therefore, they are displayed to serve as the basis for the following analysis. As it can be noticed:

- $x_{\mathrm{CH_4}}^{out}$ grows along time, mainly due to the increase in the unreacted CH_4 flow rate. Indeed, CaO consumption in the solid bed undermines CO_2 capture. An increased amount of CO_2 within the gas phase comes together with more CO available (See (R1.2)), that pushes (R1.1) equilibrium towards CH_4.

- $x_{\mathrm{H_2}}^{out}$ decreases along time. A breakthrough at which $x_{\mathrm{H_2}}^{out}$ starts decreasing monotonically from an initial, constant value of approximately 95% is clearly noticeable (in the present plot after 4 min).

- $x_{\mathrm{CO_2}}^{out}$ is null at the beginning of the stage, then it starts increasing as CaO is consumed.

- the behaviour of x_{CO}^{out} is qualitatively similar to that of CO_2 in the first minutes of the stage, since the production (or consumption) of the two species is linked through (R1.2). However, unlike CO_2, after some time x_{CO}^{out} shows a maximum and decreases, due to the competition between (R1.2) and (R1.1).[1] It was ascertained that this behaviour is due to the increase in CO flow rate and not only to that in the total mole flow rate. The value of T_s^{out} also plays a role: its decrease after a sufficiently long time (Fig. 2.3 on page 23) slows down the direct SMR, simultaneously promoting the direct WGS, causing CO to diminish.

No direct correlation between the time profiles of all the species and the value of S/C can be drawn. Notice that it does not necessarily mean that S/C does not affect the time profile of x_i^{out}. Conversely, difficulty in finding such a direct correlation could

[1] Notice that a more rigorous description of this profile would require a quantitative comparison between the two reaction rates.

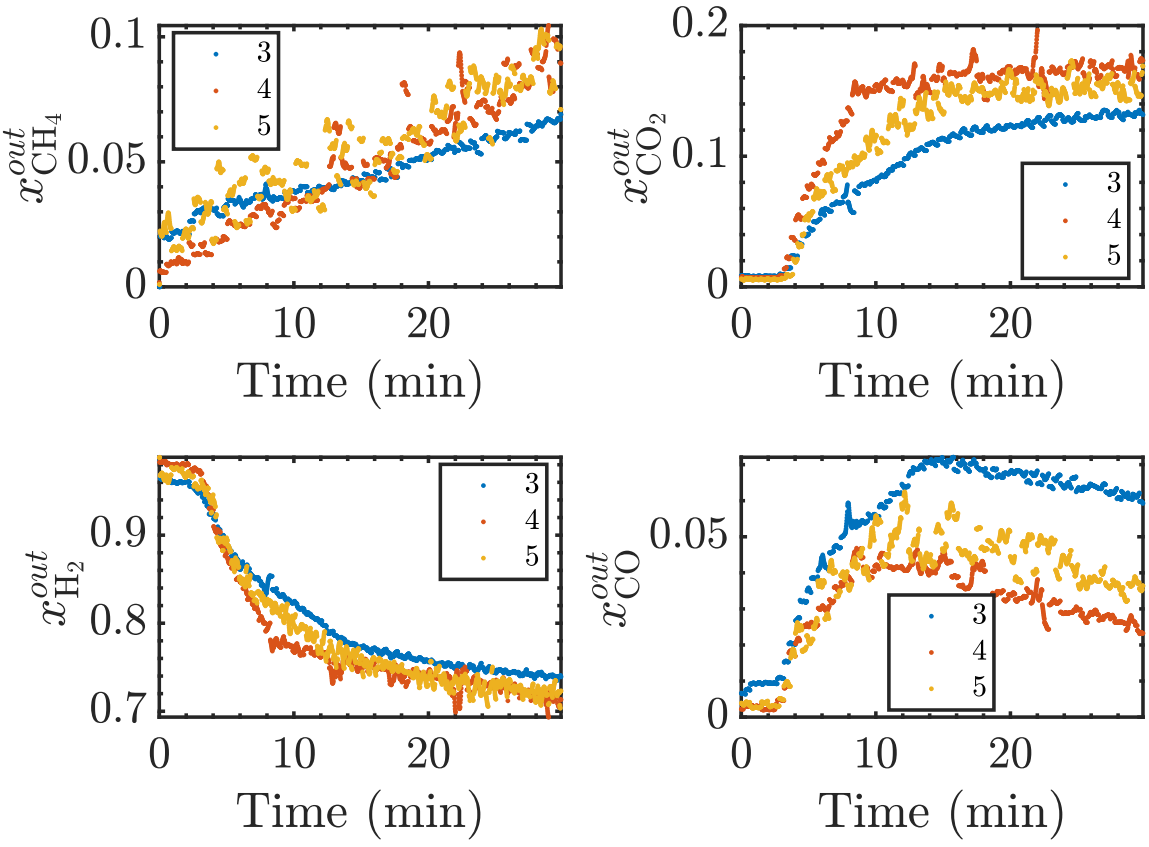


Figure 2.2: CH_4, CO_2, H_2 and CO outlet (dry) mole fractions along time in three datasets at different S/C, $P^0_G = 2$ bar. Color key: (blue) S/C = 3, (red) S/C = 4, (yellow) S/C = 5.

be due to the presence of disturbances (namely, fluctuations in unmeasured variables) entering the system, that may affect composition profiles to an extent greater than S/C does. Also notice that both profiles of CO_2 and CO align with that of H_2 as far the presence of the breakthrough is concerned.

Fig. 2.3 shows the time profile of the solid bed temperature at the outlet section for the different values of S/C. As it can be seen, at first T_s^{out} decreases because of the heat subtracted by the endothermic SMR. Subsequently, a temperature rise is observed, due to the exothermic carbonation of CaO. This profile is caused by the decoupling between the two reaction fronts, given that the reaction rate of (R1.1) is faster compared to that of carbonation (R1.3). As CO_2 sorption capacity of the bed undergoes saturation, (R1.3) becomes less effective, while (R1.1) keeps absorbing heat. This leads to the temperature decrease observed after the peak. No evident effect neither of S/C or P^0_G with respect to the temperature rise is observed. (The plot of this second input does not add anything to the discussion and hence is not reported here.)

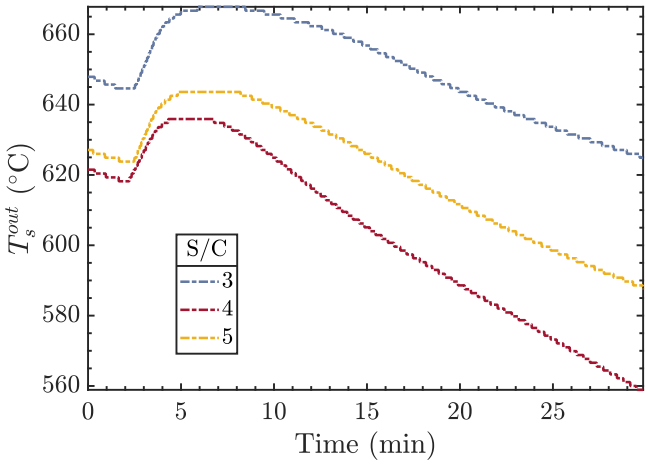


Figure 2.3: Time profile of the solid bed temperature at the outlet section for three datasets at different S/C, $P_G^0 = 2$ bar. Color key: (light blue) S/C = 3, (dark red) S/C = 4, (yellow) S/C = 5.

Besides x_i^{out} and T_s^{out}, the temperature of the solid bed along the axial coordinate was also investigated. As an example, Fig. 2.4 shows the axial profile parametric in time, for one dataset at S/C = 3, $P_G^0 = 2\,\text{bar}$. Temperature peaks are not clearly noticeable since all three reactions involved in this stage are equilibrium limited. Due to the decoupling between (R1.1) and (R1.3) mentioned in Chapter 1, T_s^{out} is expected to be not uniform along the axial coordinate but to present peaks. However, the position of these peaks is not clearly noticeable. For this reason, the sole time axial profile of T_s seemed not handy to be used to determine the switching time.

2.3 Oxidation

Fig. 2.5 shows the time profiles of CO_2 and O_2 flow rates (on mass base) along time at different $x_{O_2}^0$ in the inlet and same P_G^0. Ideal gas law was employed to convert volume to mass flow rates. As it can be seen:

- O_2 flow rate is approximately null for a given time interval during this stage, then it starts increasing. Indeed, at the beginning of the stage all O_2 flowed through the packed bed reacts with Cu. However, as soon as the Cu fraction available on the solid bed decreases, some O_2 is released, indicating that much of the bed has been oxidized. Clearly, as $x_{O_2}^0$ increases, keeping the other conditions constant, the time at which O_2 release begins is anticipated, because of the fastest Cu oxidation.

- CO_2 released is very low at the beginning, then it increases reaching a maximum, before it starts decreasing. This profile correlates with the one of T_s^{out}, displayed

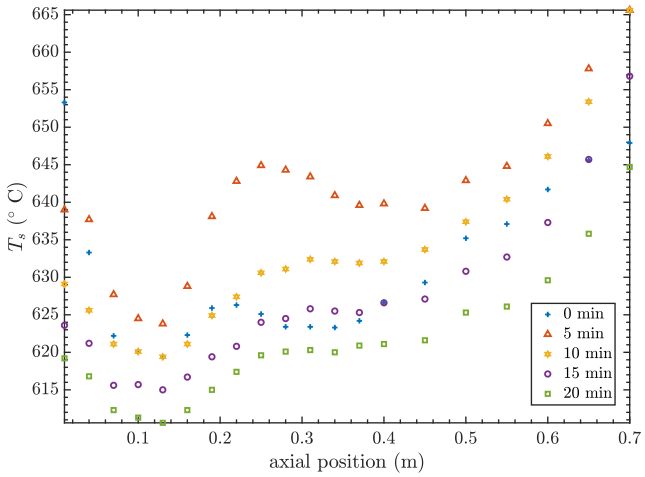


Figure 2.4: Temperature of the solid bed along the axial coordinate at different time instants for stage A (SER). Operating conditions: S/C = 3, $P_G^0 = 2$ bar.

in Fig. 2.6 (green curve): as T_s^{out} increases, calcination (undesired in this stage) becomes relevant. Indeed, the maximum CO_2 release occurs approximately when the maximum T_s^{out} is reached. After T_s^{out} decreases, losses in CO_2 diminish, accordingly.

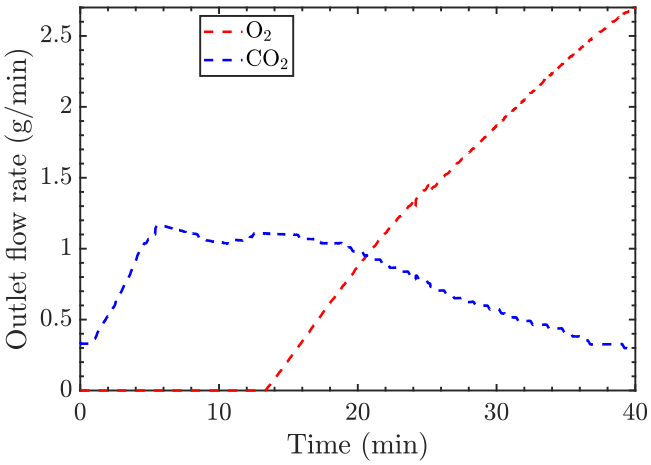
No evident correlation between x_i^0 profile and P_G^0 was found.

Fig. 2.6 shows the time-profile of T_s^{out} in two datasets at different $x_{O_2}^0$ and same P_G^0. As expected, both curves present a maximum; as $x_{O_2}^0$ increases, the value of $\max T_s^{\text{out}}$ augments, due to the faster heat released by Cu oxidation and it is anticipated, for the same cause. T_s^{out} decrease that follows the maximum is steeper for the higher $x_{O_2}^0$, because oxidation expires in a very short time, while for lower $x_{O_2}^0$ the reaction occurs more gradually. The same comparison was drawn for two datasets at same $x_{O_2}^0$, but different P_G^0 (here not reported). In this case, direct effect of P_G^0 on the curves was observed.

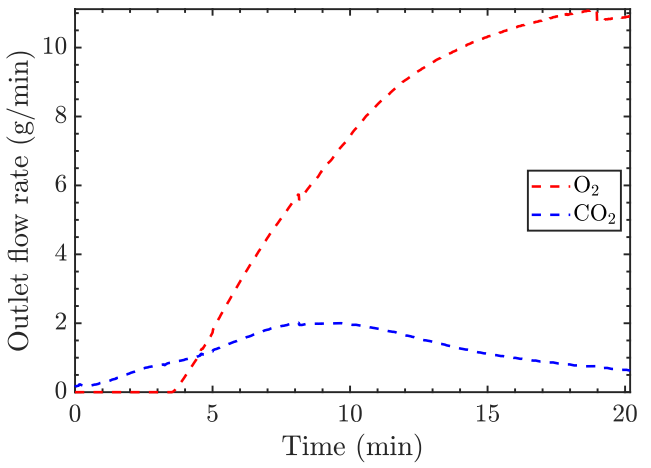
Fig. 2.7 shows T_s along the axial coordinate at different time values. Unlike stage A, where no clear temperature peaks are distinguishable, in this case a T_s peak due to the heat released by Cu oxidation is noticeable. The peak position moves along the axial coordinate as time progresses, indicating that the reaction front position changes along time as a consequence of Cu oxidation, for which the Cu which available to react is the one that has not come in contact with O_2 yet.

2.4 Reduction

Fig. 2.8 shows the time profile of the species mass flow rates in this stage. All the four profiles give useful insight into the stage: the fast decrease in CH_4, CO_2 and CO



(a) $x^0_{O_2} = 3\%$.



(b) $x^0_{O_2} = 10\%$.

Figure 2.5: O_2 and CO_2 outlet mass flow rate along time for stage B (oxidation), at $P^0_G = 3$ bar. (a) $x^0_{O_2} = 3\%$; (b) $x^0_{O_2} = 10\%$.

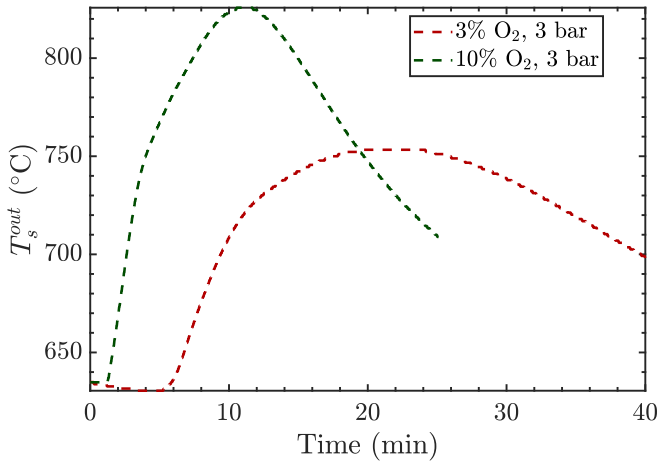


Figure 2.6: Time profile of the solid bed temperature at the outlet section during stage B (oxidatio n) at different $x^0_{O_2}$, same P^0_G.

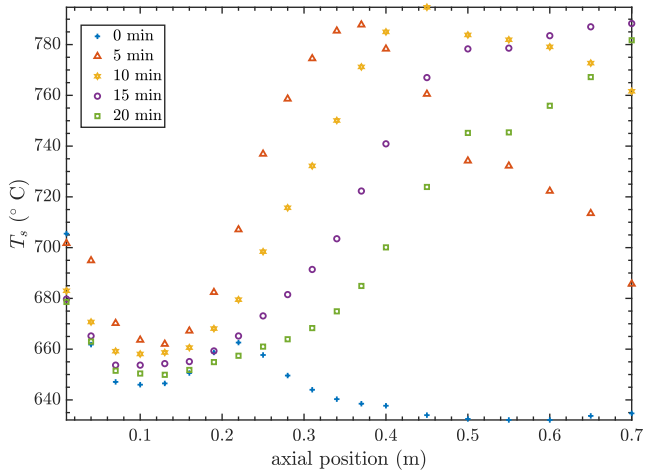


Figure 2.7: Temperature of the solid bed along the axial coordinate at different time instants for stage B (oxidation). Operating conditions: $x^0_{O_2} = 3\%$, $P^0_G = 2$ bar.

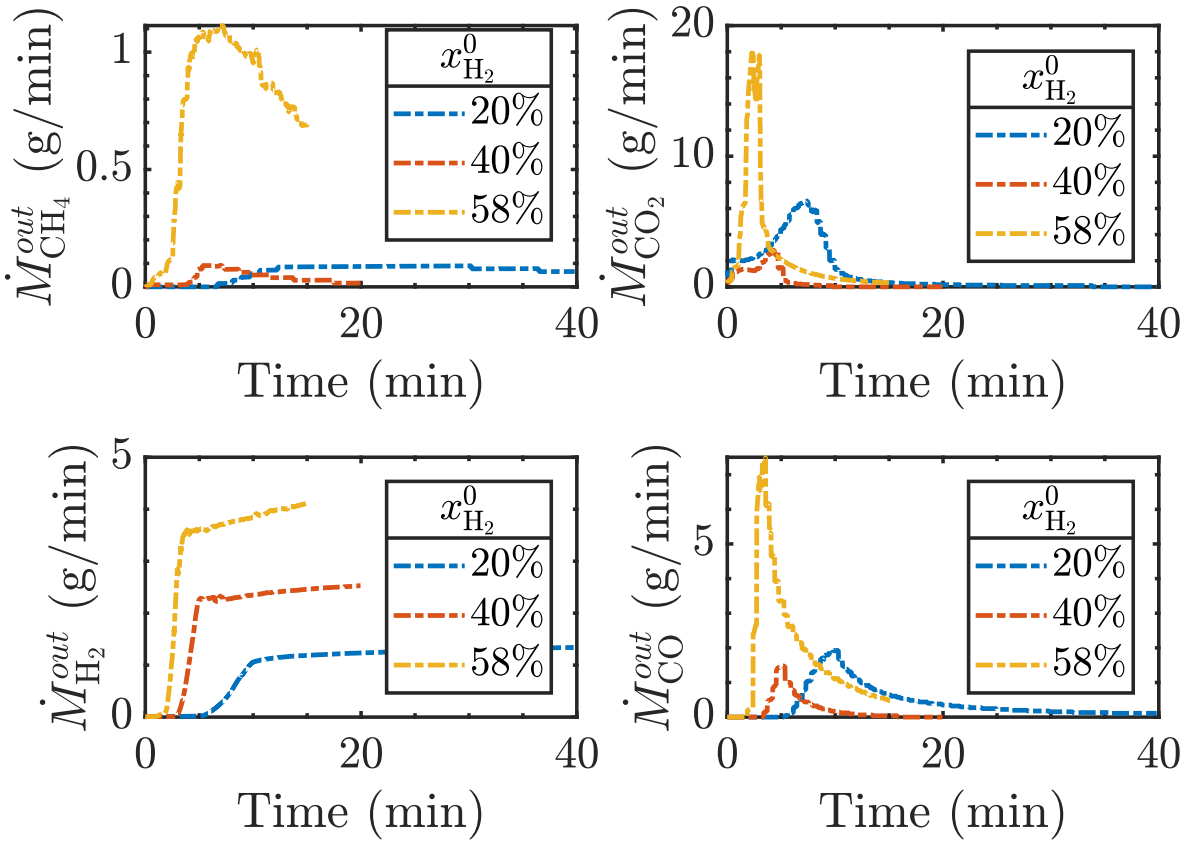


Figure 2.8: CH_4, CO_2, H_2, CO outlet mass flow rates along time at different $x^0_{H_2}$. Color key: (blue) $x^0_{H_2} = 20\%$, (red) $x^0_{H_2} = 40\%$, (yellow) $x^0_{H_2} = 58\%$.

together with the net increase in H_2 suggest that the stage is coming to an end. In details:

- $\dot{M}^{out}_{CH_4}$ is quite low, since CH_4 is formed only in limited amounts as a result of the inverse SMR taking place due to the (fed) H_2 reacting with CO formed by inverse WGS. A slight delay in CH_4 formation with respect to CO_2 is noticeable, especially in the slower processes, that is at $x^0_{H_2} = 20\%$ and $x^0_{H_2} = 40\%$. Indeed, formation of CH_4 molecules require some time for $CaCO_3$ to release CO_2 in the gas phase.

- $\dot{M}^{out}_{H_2}$: in the first minutes of the stage it is approximately null, indicating that all the H_2 fed is reducing CuO. In the subsequent minutes of the stage, it increases suggesting that the reduction stage is coming to the end. Clearly, by increasing $x^0_{H_2}$ the curve is stretched towards higher values of the outlet flow rates and squeezed to shorter time.

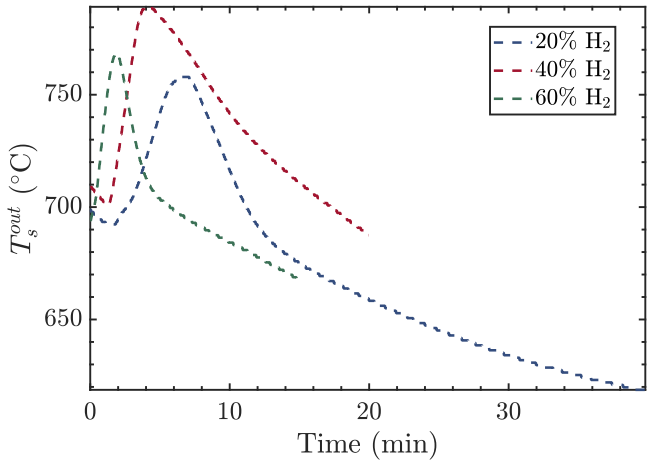


Figure 2.9: Time profile of the solid bed temperature at the outlet section during stage C (reduction) for three datasets at different $x^0_{H_2}$. Color key: (blue) $x^0_{H_2} = 20\%$, (red) $x^0_{H_2} = 40\%$, (dark green) $x^0_{H_2} = 58\%$.

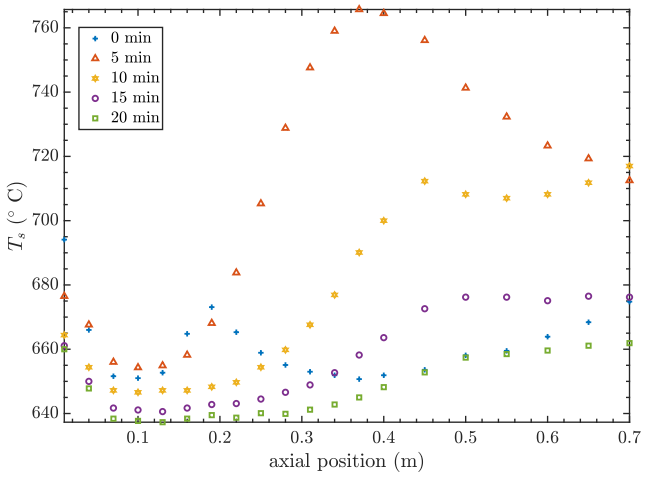
- $\dot{M}^{out}_{CO_2}$ increases fast in the very first minutes of stage C, as a result of the temperature increase that drives calcination of the solid bed. Again, higher $x^0_{H_2}$ cause the curve to be shifted towards shorter time intervals. Decrease in $x^0_{CO_2}$ is an evident sign that carbonation is coming to an end and that the stage is going to be over, since calcination has a slower kinetics compared to Cu reduction.

- Despite a delay in CO formation, which requires some time for WGS to occur, the profile of $\dot{M}^{out}_{CO}$ is qualitatively similar to that of CO_2. The decrease in $\dot{M}^{out}_{CO}$ indicates that the stage is coming to an end, as a consequence of the lower amount of CO_2 released. As for the other profiles, notice that for higher H_2 feeds, the peak is anticipated. No correlation between the value of $x^0_{H_2}$ and the height of the peak can be drawn.

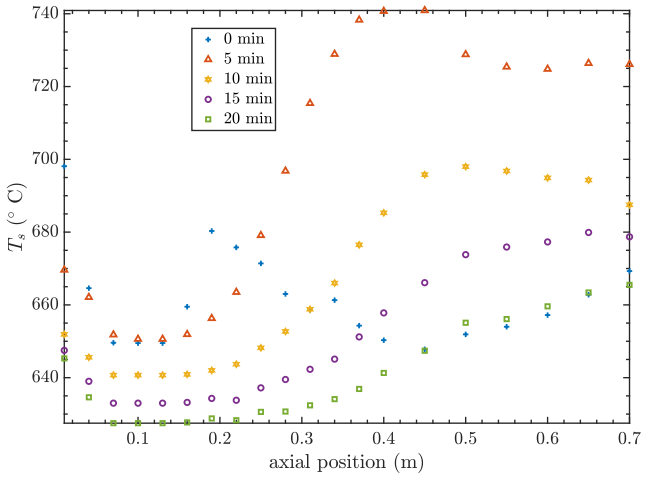
As a result, all four species profiles provide useful insight into the dynamics of the process; by comparison of the four plots a slight delay in CH_4 and CO formation with respect to CO_2 is noticeable, especially in the slower processes, that is at $x^0_{H_2} = 20\%$ and $x^0_{H_2} = 40\%$.

Fig. 2.9 shows T_s^{out} along time for three datasets at different $x^0_{H_2}$. As in stage B, the higher $x^0_{H_2}$, the first the maximum in T_s^{out} is found, suggesting that reduction is faster in reaching the final stages of the reactor. Notice that the highest peak does not coincide with the case at the greatest $x^0_{H_2}$.

Fig. 2.10 shows T_s^{out} along the axial coordinate at different times, for two values of $x^0_{H_2}$, namely 20% and 40%. As for the oxidation stage, a clear temperature peak, moving along the solid bed is noticeable; in this case, the temperature peak is due to



(a) $x^0_{H_2} = 20\%$



(b) $x^0_{H_2} = 40\%$

Figure 2.10: Temperature of the solid bed along the coordinate at different time instants for stage C (reduction). (a) $x^0_{H_2} = 20\%$; (b) $x^0_{H_2} = 40\%$.

the heat released by Cu reduction. Comparison of the two plots again confirms that higher amounts of H_2 in the inlet anticipate the time at which the maximum occurs.

2.5 Summary

The analysis of time profile of several variables shown in this Chapter allows to gain a useful insight into the process. Considerations made in the previous sections, led to the choice of the time profiles of T_s^{out} and composition as the most representative ones to monitor the course of the stages. Conversely, employing T_s seemed too difficult for the purpose, as a quantitative interpretation of this profile is not rapidly achievable.

Chapter 3

Offline Determination of the Switching Time

The analysis of the datasets displayed in the previous Chapter allowed to identify the most representative variables to monitor the course of each stage. This Chapter concerns the mathematical formalization of the employed switching criteria. Specifically, the optimal duration of each stage has been computed offline for all the experimental runs available, employing several methods. As mentioned in the Introduction, all the criteria rely on the time-profiles of one or more variables. Profiles obtained from measurements have been employed for the purpose. Each stage has been assumed to be independent from the others, in the sense that initial conditions of the solid bed are such that the reactions of each stage can occur.

Since this work follows an optimization-based approach, Section 3.1 provides the reader with a formal introduction of optimization theory. The following sections present the proposed criteria to compute optimal stages durations offline. For sake of clarity, all the combinations of objective functions and constraints employed are collected in Tab. 3.1.

3.1 Optimization-based approach

To determine the switching time, an optimization approach is followed. In general, the optimum of an objective function is found under some constraints. The choice of the objective functions and constraints is based on the data analysis presented in Chapter 2 and on the need to identify suitable criteria to match optimal operating targets for each

Table 3.1: Proposed combinations of objective functions and constraints for stages A, B and C. Letter A refers to criteria employed for stage A, B to those used for stages B and C. Subscript i stands for O_2 in stage B and H_2 in stage C, respectively.

Objective Function	Constraint	Criterion
$(1-\alpha)\, x^{out}_{H_2}(t) + \alpha \dfrac{\int_0^t P(\tau)\, d\tau}{P_{ref}}$	none	$A1$
$\dfrac{\int_0^t P(\tau)\, d\tau}{P_{ref}}$	$x^{out}_{H_2} \geq \varepsilon$	$A2$
$\alpha \dfrac{T^{out}_s(t)}{T_{ref}} \mp (1-\alpha) \dfrac{\int_0^t \dot{M}^{out}_{CO_2}(\tau)\, d\tau}{M_{ref}}$	$\dfrac{\dot{M}^{out}_i(t)}{\dot{M}^0_i(t)} \leq \varepsilon_{1,k}$	$B1$
$\dfrac{T^{out}_s(t)}{T_{ref}}$	$\dfrac{\dot{M}^{out}_i(t)}{\dot{M}^0_i(t)} \leq \varepsilon_{1,k}$	$B2$
	$\dfrac{\int_0^t \dot{M}^{out}_{CO_2}(\tau)\, d\tau}{M_{ref}} \leq (\geq)\varepsilon_{2,k}$	$B3$
	both	$B4$

process stage. In general, an optimization problem is formulated as:

$$\begin{aligned} \min_t \quad & \{\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_l\} && \text{objective function} \\ \text{subject to} \quad & f(\mathbf{x}) = 0 && \text{equality constraint} \\ & g(\mathbf{x}) \leq 0 && \text{inequality constraint} \\ & \underline{\mathbf{x}} \leq \mathbf{x} \leq \bar{\mathbf{x}} && \text{feasibility condition} \\ & \underline{\mathbf{u}} \leq \mathbf{u} \leq \bar{\mathbf{u}} && \text{feasibility condition} \\ & t_0 \leq t \leq t_{end} && \text{feasibility condition} \end{aligned} \tag{3.1}$$

where the variable t represents time, $\mathbf{x}$ the compositions and $\mathbf{u}$ the operating conditions. Therefore, each $\mathcal{F}_j$ can be expressed as:

$$\mathcal{F}_j = \mathcal{F}_j(t, \mathbf{x}, \mathbf{u}) \tag{3.2}$$

The time t which satisfies the optimization is the switching time, t_s. The specific formulation of the optimization problems for the three stages can be found in Section 3.2, 3.3, respectively.

3.2 Stage A

As described with details in Chapter 1 and 2, stage A is devoted to H_2 production via SER. The proposed criteria aim at satisfying the following requirements:

- the operation of stage A should last until it is economically profitable;
- H_2 purity should be high.

Since more than one objective has to be fulfilled, the problem can be categorized as a multi-objective optimization. Two alternative formulations were employed (Alhammadi and Romagnoli, 2004):

1. the weighted method: two objective functions, weighted through one parameter are optimized simultaneously. By varying the parameter values a set of solutions that constitute a Pareto front (Liuzzi, 1999) is obtained. A general advantage of the weighted method is that it yields to several solutions, thus providing the decision maker with a number of different possibilities among which making the choice.
2. the ε-method: the multi-objective optimization is converted into a single objective optimization, by turning the other objective functions into constraints.

The criteria employed are discussed below. All the functions are displayed in terms of time. No analytical forms were employed, but only their numerical values computed from data at any sampling time t.

3.2.1 Stage A: criterion 1 – A1

As a first step, the simultaneous optimization of two objective functions, namely profits and H_2 purity was sought through the weighted method. The problem was stated in the form:

$$A1 \quad \boxed{\max_t \quad F_{A1}(t) = (1-\alpha)\, F_1(t) + \alpha\, F_2(t),} \tag{3.3}$$

where α is the weighting parameter, varied within the interval $[0, 1]$. Specifically, the following values were set: $\alpha = \{0, 0.025, 0.050, 0.075, 0.1, 0.2, 0.5, 0.8, 1\}$. F_1 is defined as follows:

$$F_1(t) = x_{\mathrm{H_2}}^{out}(t), \tag{3.4}$$

where $x_{\mathrm{H_2}}^{out}$ is the H_2 dry fraction out of the reactor. F_2 is defined as:

$$F_2(t) = \frac{\int_0^t P(\tau)\, d\tau}{P_{ref}}, \tag{3.5}$$

where P_{ref} represent the maximum values of profits for the stage. This terms was introduced to rescale values, to make the comparison between F_1 and F_2 meaningful. $P(t)$ (€/time) represents the profits from operation of the stage. In details, profits were computed as:

$$\begin{aligned} P(t) &= \text{Revenues} - \text{Costs} \\ &= \underbrace{\dot{M}^{out}_{\mathrm{H_2}}(t)\, p_{\mathrm{H_2}}\, x^{out}_{\mathrm{H_2}}(t)}_{\text{revenues}} - \underbrace{\dot{M}^{0}_{\mathrm{NG}}(t)\, p_{\mathrm{NG}} - \dot{M}^{out}_{\mathrm{CO_2}}(t)\, CT}_{\text{costs}}, \end{aligned} \tag{3.6}$$

where p_i (€/kg$_i$) is the price of the i-th chemical to be bought or sold, $CT = €\,18$ per kg$_{\mathrm{CO_2}}$ is the carbon tax on CO_2 emissions (platts.com, 2017). Prices of p_{NG} and pure $p_{\mathrm{H_2}}$ were assumed to be €0.44 and €10, respectively. For sake of simplicity, the p_{NG} was approximated to that of pure CH_4 (Bezzo, 2018), because the employed data refer to experimental conditions for which pure CH_4 was used. Determining the price of pure H_2 is very complicated, since no global market price databases are available. Indeed, high price elasticity characterize hydrogen transactions between producers and consumers. Price depends on several parameters, like purity, type of contaminants, amount supplied and location. Among industrial players, it is known that common hydrogen prices vary between €10 and €60 per kg (Hinicio, 2015). At this point, two major approximations were done to estimate the value of H_2 according to its purity: the price was assumed to vary linearly with $x^{out}_{\mathrm{H_2}}$ and such that it is equal to zero for $x^{out}_{\mathrm{H_2}} < 0.80$. Such an estimation although unrealistic, suffices to the scope of this work. Otherwise, (3.6) would be likely monotonically increasing, due to its oversimplistic formulation that does not account for costs like steam generation, purification steps and other downstream operations. Since data available refer to a lab-scale set up and not to a real plant, all this information is missing. Hence, the need to employ very approximated cost figures. Though this is the grossest approximation made in the present study, notice that the aim of this work is not that of coming up with a result tailored to the current lab scale set up, but that of providing a suitable methodology to compute t_s in a generic process implementing SER. This is interesting in the view of a future up-scale of the process. As far as the other terms in (3.6) are concerned, $\dot{M}^0_i(t)$ and $\dot{M}^{out}_i(t)$ (kg min^{-1}) are the inlet and outlet flow rates of species i at the given time instant, namely NG, CO_2 and H_2, computed as:

$$\dot{M}_i(t) = \hat{\rho}_i(t)\dot{V}_i(t), \tag{3.7}$$

where $\dot{V}_i$ (m^3 min^{-1}) stands for the volume flow rate of i and $\hat{\rho}_i$ (kg m^{-3}) is the density of pure i. Due to the stage operating window (relatively low pressure and high temperature), the ideal gas law was employed as equation of state. Therefore, $\hat{\rho}_i$ was computed as:

$$\hat{\rho}_i(t) = \frac{P_G(t)\, MW_i}{R\, T_G(t)}, \tag{3.8}$$

where P_G (Pa) is the pressure of the gases in the inlet (or outlet) section, MW_i ($\mathrm{kg\,mol^{-1}}$) is the molar mass of the species i, R ($\mathrm{J\,mol^{-1}\,K^{-1}}$) is the gas constant and T_G (K) is the inlet (or outlet) gas temperature. The inlet temperature of the gases was measured. The outlet gas temperature was set as $T_G^{out} = T_s^{\mathrm{out}}$, given the assumption of thermal equilibrium between solid and gas phases, as stated in Chapter 1.

Application of criterion A1 to the experimental runs

Fig. 3.1a shows the values of F_{A1} along time obtained by varying α in the interval reported above. Three regions are distinguishable:

- for $\alpha = 0$, F_{A1} coincides with F_2, namely the time-profile of the H_2 outlet dry mole fraction. Optimization of F_2 yields to the t_s that maximizes H_2 purity.
- For $\alpha = 1$, F_{A1} coincides with F_1. Hence, t_s found at $\alpha = 1$ is the one that maximizes the economic profitability.

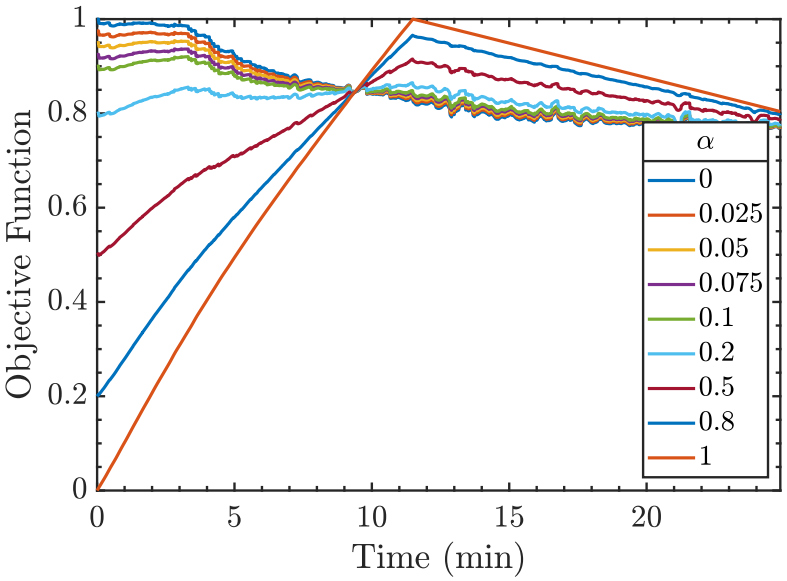
  At the beginning of the stage, F_1 increases, achieving its maximum at the given t_s, driven by the production of costly (high purity) H_2, that makes revenues higher than costs. Given the definition of (3.4) in integral terms, profits grow as more high purity H_2 is produced. After t_s, the profits decrease, even culminating in economic losses, if the stage protracts for a long time, because, based on the assumption made in Subsection 3.2.1, $x_{\mathrm{H_2}}^{out}$ is lower and it does not compensate for costs associated to raw materials and CO_2 release.
- For $0 < \alpha < 1$ (intermediate values) the optimization yields to a trade off between the highest economic profitability and H_2 purity.

Fig. 3.1b shows the optimal switching time vs α for five datasets, at S/C = 3, $P_G^0 = 2\,\mathrm{bar}$. As it can be seen, t_s decreases as α diminishes, because F_2, the term related to H_2 purity becomes dominant, when the highest H_2 purity is found in the very first minutes of the process, because most of CaO on the solid bed is available to allow CO_2 capture. As α slightly increases, the optimal switching time increases, due to the growing importance of F_1. Finally, t_s settles to an approximately constant value. This behaviour is in accordance with Fig. 3.1a: from a certain value of α, the position (that is t_s) at which $\max\{F_{A1}\}$ is found remains almost fixed.

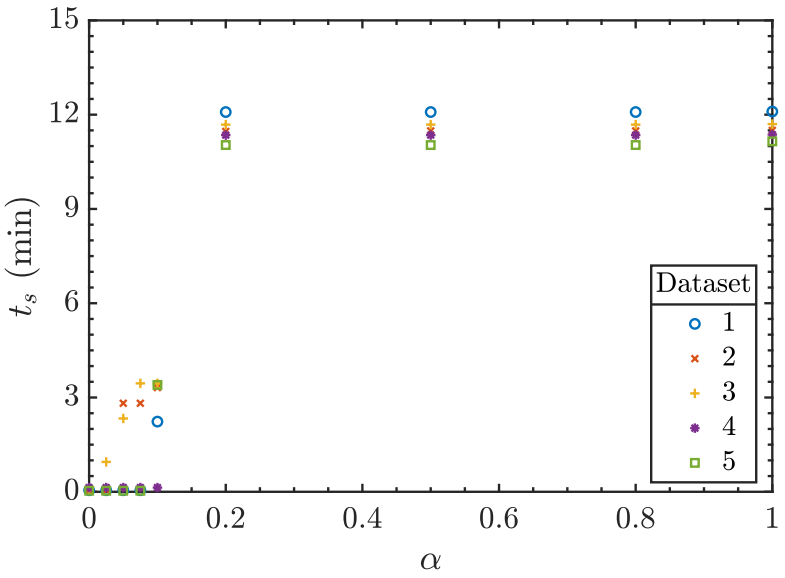
These considerations are common to all the analysed datasets. For further insights, see the plots in the Appendix.

3.2.2 Stage A: criterion 2 – A2

The ε-technique was also considered. Reasonably, most industrial plants will run the process by trying to maximize profits, imposing H_2 to reach at least a given degree of



(a) Objective functions



(b) Switching Time

Figure 3.1: Example of multi-objective optimization problem - criterion $A1$: (a) $F(t)$ parametric in α for one dataset; (b) t_s as a function of α for 5 different datasets. Operating conditions: S/C = 3, $P_G^0 =$ 2 bar in both plots.

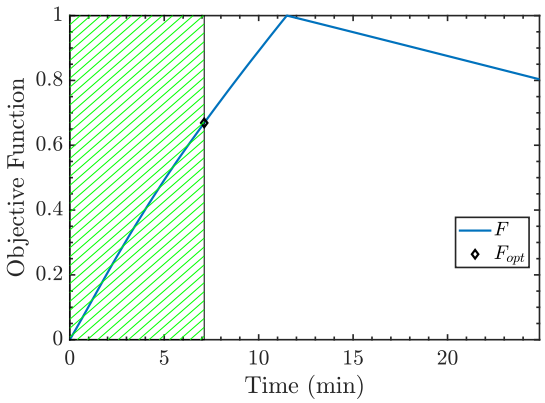


Figure 3.2: Example of constrained optimization problem applied to one dataset of stage A - criterion $A2$. The green area represents the feasible region. Nominal operating conditions: S/C = 3, $P_G^0 = 2$ bar.

purity. To account for that, the optimization problem was formulated as follows:

$$A2 \qquad \boxed{\begin{array}{c} \max\limits_{t} \quad F_{A2}(t) = \dfrac{\int_0^t P(\tau)\, d\tau}{P_{ref}} \\ \text{subject to} \quad x_{\mathrm{H_2}}^{out} \geq \varepsilon, \end{array}} \tag{3.9}$$

where $P(t)$ was computed as in the previous case and ε is the minimum acceptable H_2 purity, set equal to 0.85.

Application of criterion A2 to the experimental runs

Fig. 3.2 shows the objective function described in (3.9) for one examined dataset; the green area represents the feasible region, namely the time at which $x_{\mathrm{H_2}}^{out} \geq 0.85$. Due to definition of (3.6), the constraint turns out to be active, that is t_s is the time at which the value corresponding to the minimum acceptable H_2 purity is hit. Considerations about the shape of the function are the same described in Section 3.2.1 for $\alpha = 1$.

Table 3.2 summarizes some results of the investigation carried in the present work. Although the obtained values are highly dispersed, both S/C and P_G^0 seem to influence t_s. Precisely, their effects are opposite. As S/C increases, the average time to switch, $\bar{t}_s$, decreases, because increased amounts of H_2O boost production of CO_2 via (R1.1) and (R1.2), leading to a faster carbonation of CaO. Conversely, an increase in P_G^0 results in a longer t_s, because (R1.1) is equilibrium limited with products being favoured at low pressure. Therefore, by increasing P_G^0, the production of CO_2 is slowed down together with CaO carbonation. The results is that a longer time has to be taken before hitting the constraint.

Table 3.2: Mean ($\bar{t}_s$) and standard deviation (SD_{t_s}) of the optimal switching times computed with criterion $A2$ for each operating condition, defined in terms of S/C and P_G^0.

S/C (molar)	P_G^0 (bar)	$\bar{t}_s$ (min)	SD_{t_s} (min)	samples (n)
3	2	7.58	0.64	9
3	7	8.09	0.83	10
4	2	5.90	0.45	12
4	7	9.26	0.51	9
5	2	7.03	0.49	12
5	7	11.58	0.99	10

3.3 Stages B and C

Recalling Chapter 1, stages B and C constitute together the chemical looping reforming employed to regenerate the CaO on the solid bed. The purpose of stages B and C are:

- during stage B, Cu oxidation should be accomplished, while minimizing $CaCO_3$ calcination.
- during stage C, $CaCO_3$ calcination should be promoted. In addition, the heat released during this stage should guarantee the operation of this stage to be energy neutral.

Given their similarities, they are treated together in the following sections. Several combinations of objective functions and constraints are proposed. As for stage A, the choice of the most suitable ones are up to the designer. Motivations behind these proposals are explained with details in the following subsections. As a further step, these criteria were tested on 145 and 256 datasets for stages B and C, respectively.

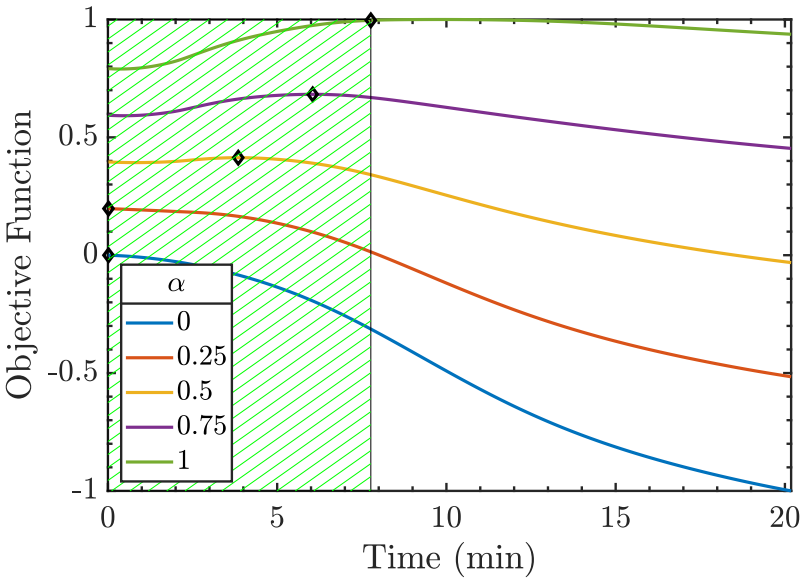
3.3.1 Stages B and C: criterion 1 – B1

The first criterion proposed is based on the weighted method:

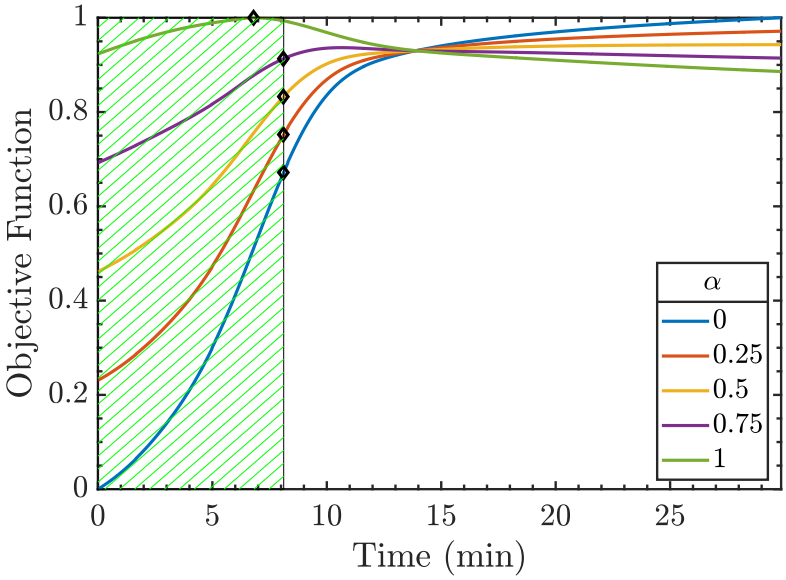
$$B2 \quad \boxed{\begin{array}{c} \max\limits_{t} \quad F_{B2}(t) = \alpha\, F_3(t) \mp (1-\alpha)\, F_4(t) \\ \text{subject to} \quad \dfrac{\dot{M}_i^{out}(t)}{\dot{M}_i^{0}(t)} \leq \varepsilon_{1,k}, \end{array}} \tag{3.10}$$

where the $-$ sign is employed for stage B and $+$ for stage C. The terms employed are:

$$F_3(t) = \frac{T_s^{out}(t)}{T_{ref}}, \tag{3.11}$$



(a) Stage B



(b) Stage C

Figure 3.3: Examples of multiobjective optimization - criterion $B1$ - applied to datasets of: (a) stage B and (b) stage C, respectively. The green area represents the feasible region, that is the time interval for which the constraint $\dot{M}_{\mathrm{i}}^{out}(t)/\dot{M}_{\mathrm{i}}^{0}(t) \leq \varepsilon_{1,k}$ is satisfied. Symbol key: ($\diamondsuit$) optimum of the objective function. Nominal operating conditions: (a) $x^0_{\mathrm{O}_2} = 10\%$, $P^0_G = 3$ bar; (b) $x^0_{\mathrm{H}_2} = 10\%$, $P^0_G = 3$ bar.

being $T_s^{out}(t)$ (K) the temperature of the solid bed at the outlet section. $T_{ref} = \max\{T_s^{out}(t)\}$ was employed as the scaling factor. The second function was defined as:

$$F_4(t) = \frac{\int_0^t \dot{M}_{\mathrm{CO_2}}^{out}(\tau)\, d\tau}{M_{ref}}, \tag{3.12}$$

where the numerator represents the total mass of CO_2 released from the beginning of the stage until time at which switching occurs and M_{ref} (kg) is the total mass of CO_2 that has flowed out of the reactor during the experiment. In the formulation of the constraint, i indicates the component O_2 in case of stage B and H_2 for stages C. The choice of $\varepsilon_{1,k}$ is up to the designer. $\varepsilon_{1,k}$ was set equal to 0.35 for both stages in the present work.

This criterion aims at finding a trade-off between the two conflicting objectives of B and C, recalled above. As described in Chapter 2, in both stages $T_s^{out}(t)$ shows a maximum; its subsequent decrease suggests that the entity of the oxidation (or reduction) are diminishing. The idea behind this proposal it that finding the maximum F_1 would result in a value of t_s for which most of Cu will be oxidized during stage B (and most of CuO will be reduced in C). Conversely, the presence of F_2 aims at answering the second requests for the two stages, that is minimizing CO_2 release during oxidation and maximizing it in reduction; hence the reason for $\mp$ signs in (3.10). About the constraint, since the release of considerable amounts of O_2 (and H_2, respectively) is a further indication of the Cu oxidation (and CuO reduction) coming to an end, setting it as a constraint seemed reasonable to allow one to spare amounts of O_2 and H_2.

Application of criterion B1 to the experimental runs

Application of this criterion to stages B and C was carried for $\alpha = \{0, 0.25, 0.50, 0.75, 1\}$ At low values of $x^0_{\mathrm{O_2}}$ and $x^0_{\mathrm{H_2}}$, F_2 is dominating for most values of α. For stage B, this yields to the unfeasible solution $t_s = 0$ min, while for stage C this results in hitting the constraint. If the problem were unconstrained, in stage C this optimization would yield to t_s equal to the time at which the stage is ended, because this corresponds to the time at which most of CO_2 is released. This behaviour depends on the fact that, based on its definition, F_2 decreases to an extent which is much greater compared to that at which F_1 increases. As P^0_G increases, in stage B the values of α at which unfeasible solutions are found tend to decrease, because higher pressure reduces calcination of $CaCO_3$ (hence, the module of F_2 at any point). The same trend was observed when $x^0_{\mathrm{O_2}}$ increases, because for higher fractions of oxidant T_s^{out} presents an anticipated spike as a result of the faster release of heat, as shown in Chapter 2. As an indication, for $x^0_{\mathrm{O_2}} \leq 3\%$, the unfeasible solution $t_s = 0$ is found for any $\alpha \leq 0.75$ in most experimental runs, while if $x^0_{\mathrm{O_2}} \leq 10\%$ the unfeasible solution appears only for $\alpha \leq 0.25$. As an

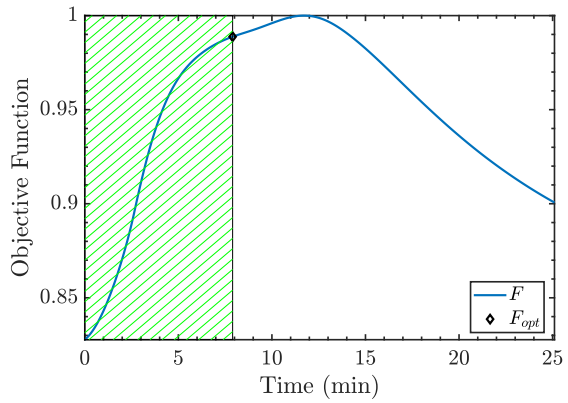


Figure 3.4: Example of constrained optimization - criterion $B2$ - applied to one dataset of stage B. The green area represents the feasible region for which the constraint $\dot{M}_{\mathrm{i}}^{out}(t)/\dot{M}_{\mathrm{i}}^{0}(t) \leq \varepsilon_{1,k}$ is satisfied. Symbol key: ($\diamondsuit$) optimum of the objective function. Nominal operating conditions: $x^0_{\mathrm{O}_2} = 10\%$, $P^0_G = 3$ bar.

example, Fig. 3.3a shows objective functions for several values of α for a dataset at $x^0_{\mathrm{O}_2} = 10\%$ and $P^0_G = 3$ bar. As it can be seen, in these conditions the solution $t_s = 0$ is found only at $\alpha \leq 0.25$, as a result of the relatively high O_2 fraction in the inlet. This is in accordance to the considerations reported in 3.3.1.

Conversely, as far as stage C is concerned, no particular effect of $x^0_{\mathrm{H}_2}$ on the values of α at which non feasible solutions are found are observed. Fig. 3.3b exhibits this behaviour, showing that the constraint is active for most values of α.

3.3.2 Stages B and C: criterion 2 – B2

Given the results obtained by employing criterion $B1$, a simplification of the objective function to the sole F_3 seemed reasonable. F_3 basically represents the time profile of the solid bed temperature. An advantage of this criterion is that the cost of thermocouples is limited and the dynamics associated to temperature measurements is usually negligible. Therefore, employing the sole F_3 as objective function seems very straightforward to be used in practice, especially in the light of a further real-time implementation. To account for the other objectives enlisted in Section 3.3, F_3 was combined with some constraints. At first the optimization problem was formulated as:

$$B2 \qquad \boxed{\begin{aligned} &\max_t F_{B2}(t) = F_3 = \frac{T_s^{out}(t)}{T_{ref}} \\ &\text{subject to} \quad \frac{\dot{M}_i^{out}(t)}{\dot{M}_i^0(t)} \leq \varepsilon_{1,k}. \end{aligned}} \tag{3.13}$$

Table 3.3: Mean ($\bar{t}_s$), standard deviation (SD_{t_s}) of the computed optimal switching times and number of samples for each operating condition of stage B, defined in terms of $x^0_{O_2}$ and P^0_G. Resulting values were obtained solving (3.13) - criterion $B2$.

$x^0_{O_2}$ (%, mole)	P^0_G (bar)	$\bar{t}_s$ (min)	SD_{t_s} (min)	samples (n)
3	3	21.94	0.20	15
3	8	14.90	0.29	14
4	$3-4$	12.62	1.48	47
5	3	15.73	0.18	14
5	$7-8$	11.98	0.54	20
8	7	10	0	1
10	3	7.91	0.10	20
10	7	6.65	0.08	14

Table 3.4: Mean ($\bar{t}_s$), standard deviation (SD_{t_s}) of the computed optimal switching times and number of samples for each operating condition of stage C, defined in terms of $x^0_{H_2}$, P^0_G. Resulting values were obtained solving (3.13) - criterion $B2$.

$x^0_{H_2}$ (%, mole)	$\bar{t}_s$ (min)	SD_{t_s} (min)	samples (n)
≤ 20	8.29	0.47	136
39	4.04	0.34	105
58	2.55	0.27	15

The mean of the optimal durations $\bar{t}_s$ for the two stages obtained employing this criterion are reported in Tab. 3.3 and Tab. 3.4, respectively. As it can be noticed, $\bar{t}_s$ always decreases with $x^0_{O_2}$ and $x^0_{H_2}$. Conversely, the effect of P^0_G in stage B is not really clear. Computed values of SD_{t_s} are relatively small for both stages, testifying a much lower variability of the results compared to those of stage A. This might be due to the fact that the network of reactions involved in stages B and C are much simpler compared to those of SER.

3.3.3 Stages B and C: criterion 3 – B3

To account for the second requirement underlined in Section 3.3, that is low (high) CO_2 releases in B (C), F_4 was then transformed into the constraint:

$$\frac{\int_0^t \dot{M}^{out}_{CO_2}(\tau)\, d\tau}{M_{cum,ref}} \leq (\geq)\, \varepsilon_{2,k}. \tag{3.14}$$

Table 3.5: Mean ($\bar{t}_s$), standard deviation (SD_{t_s}) of the computed optimal switching times and number of samples for each operating condition of stage B, defined in terms of $x^0_{O_2}$ and P^0_G. Resulting values were obtained solving (3.15) - criterion $B3$.

$x^0_{O_2}$ (%, mole)	P^0_G (bar)	$\bar{t}_s$ (min)	SD_{t_s} (min)	samples (n)
3	3	14.32	0.14	15
3	8	14.93	0.29	14
4	$3-4$	11.96	1.22	47
5	3	12.28	0.20	14
5	$7-8$	12.93	1.36	20
8	7	10.12	0	1
10	3	8.78	0.46	20
10	7	8.87	1.22	14

Then, the optimization problem was stated as:

$$B3 \quad \boxed{\begin{array}{l} \max_t F_{B3}(t) = F_3 = \dfrac{T_s^{out}(t)}{T_{ref}} \\ \text{subject to} \quad \dfrac{\int_0^t \dot{M}^{out}_{CO_2}(\tau)\, d\tau}{M_{cum,ref}} \leq (\geq)\, \varepsilon_{2,k}. \end{array}} \tag{3.15}$$

$\varepsilon_{2,k}$ was set equal to 0.50 for both stages B and C. Interestingly, as far as stage B is concerned, this constraint was observed not being active when the criterion was applied to the datasets corresponding to oxidations at $P^0_G = 7-8$ bar, while it is always active for $P^0_G = 3$ bar. This is reasonable because low pressure enhances calcination.

As far as stage C is concerned, this constraint was observed to become often active in datasets at relatively high $x^0_{H_2}$. This is due to the fact that when $x^0_{H_2}$ increases, the temperature peak is shifted to the left, as a result of the faster heat release, that speeds up calcination.

$\bar{t}_s$ obtained by applying this criterion to the experimental runs of B and C are summarized in Tab. 3.5 and Tab. 3.6. The considerations are analogous to those made for the criterion $B2$. However, this application of $B3$ yields to shorter stage durations for stage B, when $x^0_{O_2} = 3\%$, $P^0_G = 3$ bar. This indicates that the constraint in (3.15) is active and, precisely, it is hit before that of (3.13).

Table 3.6: Mean ($\bar{t}_s$), standard deviation (SD_{t_s}) of the computed optimal switching times and number of samples for each operating condition of stage C, defined in terms of $x^0_{\mathrm{H}_2}$, P^0_G. Resulting values were obtained solving (3.15) - criterion $B3$.

$x^0_{\mathrm{H}_2}$ (%, mole)	$\bar{t}_s$ (min)	SD_{t_s} (min)	samples (n)
≤ 20	9.05	1.09	136
39	4.30	0.43	105
58	3.27	0.14	15

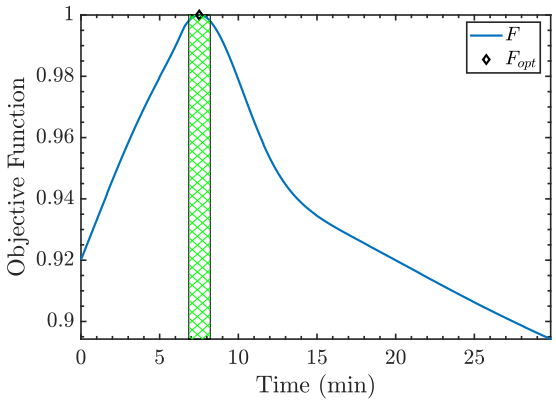


Figure 3.5: Example of constrained optimization - criterion $B4$ - applied to one dataset of stage C. The green area represents the feasible region for which the constraint $\dot{M}^{out}_{\mathrm{i}}(t)/\dot{M}^0_{\mathrm{i}}(t) \leq \varepsilon_{1,k}$ is satisfied. Symbol key: ($\diamondsuit$) optimum of the objective function. Nominal operating conditions: $x^0_{\mathrm{H}_2} = 10\%$, $P^0_G = 3$ bar.

3.3.4 Stages B and C: criterion 4 – B4

At the end, both the constraints were set simultaneously:

$$B4 \quad \boxed{\begin{aligned} &\max_t F_{B4}(t) = F_3 = \frac{T^{out}_s(t)}{T_{ref}} \\ &\text{subject to} \quad \frac{\dot{M}^{out}_i(t)}{\dot{M}^0_i(t)} \leq \varepsilon_{1,k} \\ &\qquad\qquad \frac{\int_0^t \dot{M}^{out}_{\mathrm{CO}_2}(\tau)\, d\tau}{M_{cum,ref}} \leq (\geq)\, \varepsilon_{2,k}. \end{aligned}} \tag{3.16}$$

As an example, Fig. 3.5 shows the constrained optimization problem for one dataset of stage C.

$\bar{t}_s$ for the two stages obtained with each of the criteria exposed are summarized in Tab.3.7 and Tab. 3.8. The obtained values are, again, close to those obtained by employing criteria $B2$ and $B4$, suggesting that all three criteria are equally suitable for

Table 3.7: Mean ($\bar{t}_s$), standard deviation (SD_{t_s}) of the computed optimal switching times and number of samples for each operating condition of stage B, defined in terms of $x^0_{O_2}$ and P^0_G. Resulting values were obtained solving (3.16) - criterion $B4$.

$x^0_{O_2}$ (%, mole)	P^0_G (bar)	$\bar{t}_s$ (min)	SD_{t_s} (min)	samples (n)
3	3	14.30	0.14	15
3	8	14.90	0.29	14
4	$3-4$	11.91	1.24	47
5	3	12.26	0.20	14
5	$7-8$	12.04	0.50	20
8	7	10	0	1
10	3	7.91	0.10	20
10	7	6.65	0.08	14

Table 3.8: Mean ($\bar{t}_s$), standard deviation (SD_{t_s}) of the computed optimal switching times and number of samples for each operating condition of stage C, defined in terms of $x^0_{H_2}$, $P^0_G = 3$ bar. Resulting values were obtained solving (3.16) - criterion $B4$.

$x^0_{H_2}$ (%, mole)	$\bar{t}_s$ (min)	SD_{t_s} (min)	samples (n)
≤ 20	8.34	0.47	136
39	4.10	0.32	105
58	2.63	0.21	15

the determination of the optimal stage duration.

3.4 Summary

Several criteria based on an optimization approach have been proposed, aimed at assessing the optimal stage duration for the three stages of the Ca-Cu looping process. Tab. 3.1 summarizes the objective functions and constraints employed. While the optimal switching times computed offline for the experimental runs of stage A are characterized by considerable variations, a very low dispersion has been observed for stages B and C. Chapter 5 will describe the online adaptation of some criteria. While in the offline determination time profiles of the variables of interest are obtained from measurements, the real-time approach makes it necessary to obtain some estimations. As stated in the Introduction, this is achieved via soft-sensing techniques.

As the employed soft-sensors have been realized within the framework of Gaussian Process Regression, the next Chapter constitute a parenthesis aimed at introducing the reader to this topic.

Chapter 4

Bayesian Inference and Gaussian Processes

THIS CHAPTER provides the reader with an introduction to Gaussian Processes (GPs), as they constitute the tool to realize the soft-sensors used in this work. GPs are currently employed in supervised learning to solve both *regression* and *classification* problems. The following discussion will focus on regression, which is the problem of finding a suitable smooth curve that interpolates a given set of noisy samples drawn from an unknown function.

Section 4.1 motivates the choice of GP as a tool to solve regression tasks. Section 4.2 gives the definition of GP. After that, Section 4.3 gives an introduction of Bayesian inference and its use to identify a distribution over functions. Section 4.4 concerns the principles of model selection and shows how GPR models are employed to make predictions.

4.1 Motivation

This work deals with regression problems of the type:

$$y_i = f(x_i) + \varepsilon_i, \tag{4.1}$$

where $f(\cdot)$ is a unknown smooth function to be reconstructed, x_i is a known point that belongs to the domain of $f(\cdot)$ and y_i is a measurement of $f(x_i)$ corrupted by additive Gaussian noise ε_i. The goal of Gaussian process regression (GPR) is to reconstruct $f(\cdot)$ from a set of n measurement points $\mathcal{D} = \{(\mathbf{x}_i, y_i) \,|\, i = 1, \ldots n\}$. For convenience, the definition $w_i = f(x_i)$ will be used in this Chapter.

Regression tasks are conventionally tackled by employing parametric models. The type of basis functions and the number of parameters, i.e. the number of basis functions, are specified upfront, assuming that once tuned, the function will be suitable to represent the data. One drawback of traditional parametric models, lies in possible lack of predictive power when dealing with complex datasets (Rasmussen, 2004).

One alternative approach is that of assigning a prior probability to any possible

function, such that higher probabilities are given to functions considered more likely (Bernardi, 2010). This is the scheme followed in GPR. Unlike conventional parametric models, the number of parameters does not have to be specified *a priori*, since GPs allow to consider a possibly infinite number of basis functions. This does not mean that the resulting estimated function will be a combination of an infinite number of basis functions. Indeed, it would be impossible compute them in a finite amount of time. In the framework of GPR, a function is conceived as a vector whose entries specify its values $f(\mathbf{x})$ at a given input $\mathbf{x}$. Seeking values of the function at a finite number of points (see Section 4.3) allows the algorithm to perform Bayesian inference only on those points, ignoring the (infinitely many) others.

4.2 Gaussian Processes

The definition of a GP is as follows (Rasmussen and Williams, 2006):

Definition 1. A GP is a collection of random variables, any finite number of which have a joint Gaussian distribution.

That is, all the random vectors $\mathbf{w}$ that can be obtained by taking an arbitrary number of random variables have the following joint probability density function (PDF):

$$p(\mathbf{w}) = \frac{1}{(2\pi)^N |K|} e^{-\frac{1}{2}(\mathbf{w}-\boldsymbol{\mu})^\top K^{-1}(\mathbf{w}-\boldsymbol{\mu})}, \tag{4.2}$$

where $\boldsymbol{\mu}$ and K are the mean and covariance matrix of the vector $\mathbf{w}$, respectively. Their entries are completely specified by the mean and covariance function (*kernel*) of the GP. For a real process $f(\mathbf{x})$, these are defined as:

$$m(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})], \tag{4.3}$$

and

$$k(\mathbf{x}, \mathbf{x}') = \mathbb{E}[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))], \tag{4.4}$$

respectively. Then, the $\mathcal{GP}$ in terms of m and k, is formally written as:

$$f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')), \tag{4.5}$$

that is the function f is distributed as a $\mathcal{GP}$ with mean function m and covariance function (kernel) k.

4.3 Bayesian Inference

In GPR, the general principle of Bayesian inference is applied to update the distribution over functions in the light of training data. Within the Bayesian framework, model selection takes place employing a hierarchical approach (Fig. 4.1): at the lowest level (level 1), Bayesian inference is employed to learn model parameters, $\mathbf{w}$, at level 2 the hyperparameters, $\boldsymbol{\theta}$ that controls the distribution of $\mathbf{w}$ and at top level (level 3) the set of possible model structures, $\mathcal{H}_i$.

4.3.1 Level 1

Consider the predictor matrix $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$, and the vector of targets $\mathbf{y} = [y_1, \ldots, y_n]^\top$. By fixing the hyperparameters $\boldsymbol{\theta}$ and the model $\mathcal{H}_i$, a joint Gaussian distribution of $\mathbf{w}$ and $\mathbf{y}$ is obtained. According to Bayes' rule, the inference posterior over the parameters at level 1 is given by:

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{marginal} \quad \text{likelihood}}, \quad p(\mathbf{w}|\mathbf{y}, X, \boldsymbol{\theta}, \mathcal{H}_i) = \frac{p(\mathbf{y}|X, \mathbf{w}, \mathcal{H}_i)p(\mathbf{w}|\boldsymbol{\theta}, \mathcal{H}_i)}{p(\mathbf{y}|X, \boldsymbol{\theta}, \mathcal{H}_i)}, \tag{4.6}$$

where the normalizing constant in the denominator of (4.6) is the *marginal likelihood* and it is independent of the parameters, according to:

$$p(\mathbf{y}|X, \boldsymbol{\theta}, \mathcal{H}_i) = \int p(\mathbf{y}|X, \mathbf{w}, \mathcal{H}_i)p(\mathbf{w}|\boldsymbol{\theta}, \mathcal{H}_i)d\mathbf{w}, \tag{4.7}$$

The terms at the numerator are:

- $p(\mathbf{y}|X, \mathbf{w}, \mathcal{H}_i)$ is the *likelihood*, that is the probability density of the observations given the parameters;

- $p(\mathbf{w}|\boldsymbol{\theta}, \mathcal{H}_i)$ is the parameters *prior*, the probability distribution of the parameters before seeing data.

Therefore, the posterior density combines the information from the prior and the data likelihood, and gives a probability distribution of the unknown function $\mathbf{w}$.

4.3.2 Level 2

An analogous procedure is carried out to compute the posterior on the hyperparameters, obtaining:

$$p(\boldsymbol{\theta}|\mathbf{y}, X, \mathcal{H}_i) = \frac{p(\mathbf{y}|X, \boldsymbol{\theta}, \mathcal{H}_i)p(\boldsymbol{\theta}|\mathcal{H}_i)}{p(\mathbf{y}|X, \mathcal{H}_i)}, \tag{4.8}$$

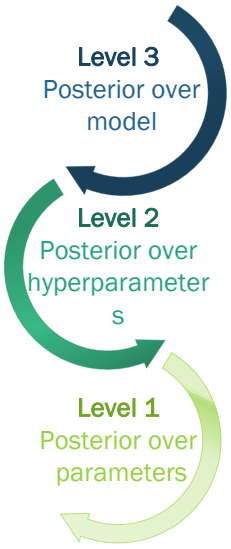


Figure 4.1: Hierarchical representation of the three levels of Bayesian inference.

where the marginal likelihood from level 1 is used as level 2 inference likelihood, $p(\boldsymbol{\theta}|\mathcal{H}_i)$ is the prior over the hyperparameters and the normalizing constant is given by:

$$p(\mathbf{y}|X, \mathcal{H}_i) = \int p(\mathbf{y}|X, \boldsymbol{\theta}, \mathcal{H}_i)p(\boldsymbol{\theta}|\mathcal{H}_i)d\boldsymbol{\theta}. \tag{4.9}$$

However, computing (4.8) can be computationally challenging, as in general there is no analytical expression for $p(\boldsymbol{\theta}|\mathbf{y}, X, \mathcal{H}_i)$. To circumvent this issue, techniques like marginal likelihood maximization and cross-validation (CV) are typically employed. A proper description of these techniques is given in Section 4.4.

4.3.3 Level 3

The posterior on the models is given by:

$$p(\mathcal{H}_i|\mathbf{y}, X) = \frac{p(\mathbf{y}|X, \mathcal{H}_i)p(\mathcal{H}_i)}{p(\mathbf{y}|X)}. \tag{4.10}$$

employing the sum extended to all the models as the marginal likelihood:

$$p(\mathbf{y}|X) = \sum_i p(\mathbf{y}|X, \mathcal{H}_i)p(\mathcal{H}_i) \tag{4.11}$$

Usually the prior on the models, $p(\mathcal{H}_i)$, is chosen such that no model is favored over another before observing data.

In many cases, approximation techniques are required to compute the integrals reported above.

4.4 Training a Gaussian process

Consider a training set $\mathcal{D} = \{(\mathbf{x}_i, y_i) \,|\, i = 1, \ldots n\}$, where n is the number of observations, each $\mathbf{x}_i$ is a vector of predictors, usually collected in a $n \times D$ matrix, X (as it will be refer throughout this Chapter), and $\mathbf{y}$ is the vector of targets y_i. Training a GP is performed according to the following steps:

1. choosing a covariance function (qualitative);

2. obtaining the values of the hyperparameters, $\boldsymbol{\theta}$, that parametrize the covariance function (quantitative);

3. obtaining the estimate of $\mathbf{w}$.

Note that these three steps correspond to the three steps outlined in the previous section, in an inverse order.

1. The kernel of the GP is parametrized in terms of $\boldsymbol{\theta}$. The choice of a proper covariance function is a crucial step in GPR models. It reflects prior knowledge on the physical process under investigation (Ebden, 2015), in the sense that it accounts for some properties of the underlying function (smoothness, possible periodicity etc).[1] Of major importance is the fact that kernels express the concept of *similarity*, that is how the response at one point $\mathbf{x}_i$ is affected by responses at other points $\mathbf{x}_j$, $i \neq j, i = 1, 2, \ldots n$. Points with similar predictor values $\mathbf{x}$ should be associated to close target values $\mathbf{y}$.

2. While setting several qualitative aspects of the underlying function f, the type of covariance function does not account for more quantitative aspect. For instance, it says that f is *smooth*, but not the extent of its smoothness; this depends on the values of $\boldsymbol{\theta}$, which are used to parametrize the structure of the covariance matrix in order to guarantee a more flexible choice of the prior of the GP. The hyperparameters are generally obtained using two different approaches:

(i) marginal likelihood maximization;

(ii) cross-validation.

The marginal likelihood is as in (4.7), namely:

$$p(\mathbf{y}|X) = \int p(\mathbf{y}|X, \mathbf{f}) p(\mathbf{f}) d\mathbf{f}. \tag{4.12}$$

From the assumption that both prior and likelihood are Gaussian distributed, i.e. $\mathbf{f}|X \sim \mathcal{N}(\mathbf{0}, K)$ and $\mathbf{y}|\mathbf{f} \sim \mathcal{N}(\mathbf{f}, \sigma_n^2 I)$, the marginal likelihood is $\mathbf{y} \sim \mathcal{N}(\mathbf{0}, K + \sigma_n^2 I)$.

[1] The covariance functions employed in this work are presented in Chapter 5

Passing to the logarithm yields:

$$\log p(\mathbf{y}|X,\boldsymbol{\theta}) = -\frac{1}{2}\log|K_y| - \frac{1}{2}\mathbf{y}^T K_y^{-1}\mathbf{y} - \frac{n}{2}\log 2\pi, \tag{4.13}$$

where $K_y = K + \sigma_n^2 I$ is the covariance matrix of the noisy targets $\mathbf{y}$ and K is the one of the noise-free latent $\mathbf{w}$. Three terms can be distinguished:

- the first term depends only on the model and acts as complexity penalty term. This term measures and penalizes the complexity of the model, i.e. the smoothness of the underlying function f.
- the second term is the data-fit one and it is the only one that depends on the observed targets. This term gives an indication about the capability of the model to explain the data.
- the third term acts as a normalizing constant and can be discarded when searching the maximum.

Maximization of (4.13) yields the optimal values of $\boldsymbol{\theta}$, according to:

$$\frac{\partial}{\partial \boldsymbol{\theta}_j}\log p(\mathbf{y}|X,\boldsymbol{\theta}) = \frac{1}{2}\mathbf{y}^\top K^{-1}\frac{\partial K}{\partial \boldsymbol{\theta}_j}K^{-1}\mathbf{y} - \frac{1}{2}\mathrm{tr}\big(K^{-1}\frac{\partial K}{\partial \boldsymbol{\theta}_j}\big). \tag{4.14}$$

The time required for model training is essentially the one due to inversion of the matrix K. This operation requires time $\mathcal{O}(n^3)$. Once computed the inversion of matrix K, calculating the j-th partial derivatives for each of the hyperparameters requires a computational time of $\mathcal{O}(n^2)$. Since computational time increases cubically with the size of the training set, problems might arise when dealing with large datasets. This is possibly the main limitation of GPR.

The second model selection strategy is CV. According to this technique, $\mathcal{D}$ is split into some subgroups, one of which is taken for training while the others are employed for testing. A model is fitted on the training set and evaluated on the test sets. By swapping training and tests subset until each of them has been employed as a training set, a number of models equal to the number of subsets is obtained. The resulting GPR model can be built by averaging the values of the hyperparameters obtained during each rotation. Common procedures employed for CV are:

- hold out method: the training set $\mathcal{D}$ is split into two disjoint sets of arbitrary size, one actually employed for training and the other for validation. This approach is relatively fast but it might be too inaccurate if the validation set is not big enough;

- k-fold cross validation: $\mathcal{D}$ is randomly partitioned into k disjoint, equally sized subsets; k - 1 sets are used for training and 1 for validation. The procedure is repeated k times, until all k are used for both training and validation. Then, the k results are often averaged to generate a single estimation. This procedure is usually good to diminish overfitting.

- leave-one-out cross-validation: it is a k-fold cross validation with k equal to the number of training points. Usually, this method is too computationally demanding.

3. The purpose of GPR models is to predict the output y_* generated by the system in response of a given point x_* that does not belong to the training set X. This corresponds to the third step of training GPs. Since data are often noisy, in general it holds that $y \neq w$. As a matter of fact, the covariance function of a noisy process is the sum of signal and noise covariances:

$$f \sim \mathcal{GP}(m, k), \quad y \sim \mathcal{GP}(m, k + \sigma_n^2 \delta_{ii'}), \tag{4.15}$$

being δ_{ii} is the Kronecker's delta. It follows that the expected values of y and f are identical, while their variances differ owing to the observational noise process (Rasmussen, 2004).

Once $\boldsymbol{\theta}$ of the covariance function have been computed with the procedure indicated in Section 4.4, prediction of y_* at input x_* is attained according to the following steps:

a. the kernel is computed for all the combination of points in the set X and x_*:

$$K = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \dots & k(x_1, x_n) \\ k(x_2, x_1) & k(x_2, x_2) & \dots & k(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_n, x_1) & k(x_n, x_2) & \dots & k(x_n, x_n) \end{bmatrix}, \tag{4.16}$$

$K_* = [k(x_*, x_1) \quad k(x_*, x_2) \quad \cdots \quad k(x_*, x_n)]$ and $K_{**} = k(x_*, x_*)$[2]

b. Data are assumed to be represented as a sample from a multivariate Gaussian distribution (this is the key assumption in GPR). The joint distribution is as follows:

$$\begin{bmatrix} \mathbf{y} \\ y_* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K & K_*^\top \\ K_* & K_{**} \end{bmatrix}\right). \tag{4.17}$$

The conditional probability of a certain prediction for y_* given the data, follows a

[2]Interestingly, notice that the diagonal elements of K are $\sigma_f^2 + \sigma_n^2$ and its extreme off-diagonal elements tend to zero when $\mathbf{x}$ spans a large enough domain.

Gaussian distribution:

$$y_*|\mathbf{y} \sim \mathcal{N}(K_* K^{-1}\mathbf{y}, K_{**} - K_* K^{-1} K_*^\top). \tag{4.18}$$

c. y_* is estimated as the mean of the previous distribution:

$$y_* = \bar{y}_* = K_* K^{-1}\mathbf{y} \tag{4.19}$$

and the uncertainty in the estimate is described by its variance:

$$var(y_*) = K_{**} - K_* K^{-1} K_*^\top. \tag{4.20}$$

Notice that $var(y_*)$ only depends on the covariance and not on the outputs. Moreover, the second quantity in this equation is always positive, suggesting that the posterior variance is anyway smaller than the prior variance, because of the addition of new data.

Equations (4.19) and (4.20) are those used to perform the predictions.

4.5 Summary

The purpose of this Chapter was that of introducing the reader to GPs, as they constitute the framework within which soft-sensors described in Chapter 5 have been realized. Model training for the specific case study faced in the present work as well as application of GPR models to infer the time profiles of several key variables as described in Section 4.4 is discussed in the next Chapter.

Chapter 5

Online Determination of the Switching Time

This chapter discusses the development of GPR-based soft-sensors for the real-time determination of the optimal switching times. As stated in the Introduction, an offline scheduling of the optimal duration of the stages is not sufficient to match satisfactory operating targets, as this approach does not allow to deal with disturbances and process uncertainties. Moreover, as shown in Chapter 3, experimental runs of stage A exhibit quite a large variability, making it necessary to employ a real-time approach.

Section 5.1 of this Chapter concerns the procedure employed to train the GPR models. Model selection and predictions quality are assessed. Section 5.2 shows the online adaptation of criteria $A2$ and $B2$, introduced in Chapter 3. Section 5.3 concerns the possibility of model retraining.

5.1 GPR models for soft-sensing

The adaptation of the criteria $A2$ and $B2$ for online use requires to predict the following outputs, according to (3.9) and (3.13):

- $\dot{V}_{tot}^{out}$, $x_{\mathrm{H_2}}^{out}$ and $x_{\mathrm{CO_2}}^{out}$ in stage A;
- $\dot{V}_{tot}^{out}$ and $x_{\mathrm{O_2}}^{out}$ in stage B;
- $\dot{V}_{tot}^{out}$ and $x_{\mathrm{H_2}}^{out}$ in stage C.

where, $\dot{V}_{tot}^{out}$ refers to the total *dry* volume flow rate flowing out from the reactor. Since all compositions are expressed on a dry basis, this quantity is needed to compute the flow rates of the single species. Dynamics of flow measurements are typically negligible, but as $\dot{V}_{tot}^{out}$ is on dry basis it cannot be sampled in real-time, because of the time delay associated to steam condensation. Besides these variables, GPR models to target $x_{\mathrm{CO_2}}^{out}$ have been trained also for stages B and C. As displayed in (3.15) and (3.16), this variable is employed in criteria $B3$ and $B4$. Although these two criteria have not been adapted online, assessing the predictive performances of GPR models when dealing with $x_{\mathrm{CO_2}}^{out}$ as an output is interesting in the view of their future implementation.

A separate GPR model has been trained for each primary variable, because the theory described in Chapter 4 holds only for scalar outputs. Therefore, defining GPR models with multiple outputs to predict all the targets of a given stage was not possible. Predictors have been chosen among those variables that satisfy the following requirements.

1. They should be influential on the outputs. It is not possible to employ all the sensed variables available, but only those that affect the outputs to a greater extent. Indeed, employing too many predictors makes model training troublesome, eventually leading to poor predictive performance. In addition, employing too many data points would make the training too demanding, as the computational time needed to train a GPR model is proportional to $\mathcal{O}(n^3)$ (see Chapter 4).

2. They should be secondary variables, that is easy-to-measure online, since online predictions of the outputs require the actual value of the input variables.

In the present work, t, P_G^0 and inlet compositions, specifically S/C, $x_{\mathrm{O}_2}^0$ and $x_{\mathrm{H}_2}^0$ in stage A, B and C, respectively, are employed as predictors to target all of the outputs. The choice of the most suitable predictors was motivated by the knowledge about the physical process, obtained via the investigation presented in Chapter 2. Among all the variables sensed in the lab-scale set up, these were found to be the most influential on the targets. An analogous behaviour is expected also in the industrial up-scale of the process. However, in industrial plants a larger number of variables, whose influence on the outputs is unknown, will be available for measurement. In these circumstances, statistical analysis (e.g. PLS and PCA) might be useful to identify their relevance. The employed combination of regressors is suitable to guarantee good accuracy in the predictions. Notice that several other variables, such as T_G^0, T_s^{out} and $\dot{V}_{tot}^0$ are expected to have great influence on the desired targets. However, since their variability in the available datasets is rather limited, including them would be of little use. The dimensions of the training sets were reduced by keeping one point every 15, which correspond to a sampling time of 15 s, to decrease the computational time needed for training. Predictor matrices, X_{A}, X_{B} and X_{C} for the three stages have been constructed as follows:

Stage A $X_{\mathrm{A}} = [\mathbf{t} \quad \mathbf{P_G^0} \quad \mathbf{S/C}]$, where $\mathbf{t}$ is the time series of the process stage, $\mathbf{P_G^0}$ collects P_G^0 along time and $\mathbf{S/C}$ the S/C. Data included in this matrix belong to 24 datasets, 4 for each set of operating conditions in terms of S/C and P_G^0, so that the predictors matrix has dimensions $\dim X_{\mathrm{A}} = 1464 \times 3$.

Stage B $X_{\mathrm{B}} = [\mathbf{t} \quad \mathbf{P_G^0} \quad \mathbf{x}_{\mathrm{O}_2}^{\mathbf{0}}]$, where $\mathbf{x}_{\mathrm{O}_2}^{\mathbf{0}}$ collects $x_{\mathrm{O}_2}^0$ along time. 49 datasets have been employed for training, 7 for each of the nominal operating conditions in terms of $x_{\mathrm{O}_2}^0$ and P_G^0. $\dim X_{\mathrm{B}} = 3927 \times 3$.

Stage C $X_{\mathrm{C}} = [\mathbf{t} \quad \mathbf{P_G^0} \quad \mathbf{x}_{\mathrm{H_2}}^{\mathbf{0}}]$, where $\mathbf{x}_{\mathrm{H_2}}^{\mathbf{0}}$ collects $x_{\mathrm{H_2}}^0$ along time. 36 datasets have been employed to train the model, 12 for each value of $x_{\mathrm{H_2}}^0$. $\dim X_{\mathrm{C}} = 1791 \times 3$.

Predictors have been standardized to avoid those having a greater order of magnitude dominating those having a smaller one and to reduce complexity during calculation. Great attention has been devoted to the choice of the most suitable kernels. The *automatic relevance determination* ARD-squared exponential kernel (k_{se}) has been employed to target $\dot{V}_{tot}^{out}$, as this covariance function typically yields to smooth functions, which is a desirable feature because flow rate data that are typically noisy. After testing several kernels, the ARD-rational quadratic (k_{rq}) has been chosen to model outlet species mole fractions. k_{se} and k_{rq} are defined as:

$$k_{se}(x, x'|\boldsymbol{\theta}) = \sigma_f^2 \exp\left(-\frac{1}{2}\sum_{m=1}^{d}\frac{(x_{i,m} - x_{i',m})^2}{\sigma_m^2}\right) + \delta_{ii'}\sigma_n^2 \tag{5.1}$$

and

$$k_{rq}(x, x'|\boldsymbol{\theta}) = \sigma_f^2\left(1 + \frac{1}{2\alpha_{rq}}\sum_{m=1}^{d}\frac{(x_{i,m} - x_{i',m})^2}{\sigma_m^2}\right)^{-\alpha_{rq}} + \delta_{ii'}\sigma_n^2, \tag{5.2}$$

respectively. The values of hyperparameters, $\boldsymbol{\theta}$, obtained by maximization of marginal likelihood, as described in Chapter 4 are reported in Tab. 5.1, 5.2 and 5.3. Commands from the MATLAB® Statistics and Machine Learning Toolbox have been employed for the purpose. The meanings of $\boldsymbol{\theta}$ in (5.1) and (5.2) are the following:

- σ_f^2 is the signal variance, indicating the extent of function variations from its mean value. As an example, consider the values collected in Tab. 5.1, 5.2 and 5.3. GPR models for $\dot{V}_{tot}^{out}$ in stage B and C allow quite broad amplitudes, while the one for stage A allows only limited ones. Analogously, notice that the models describing $x_{\mathrm{CO_2}}^{out}$ have been found to be less flexible than those describing the other compositions, in all three stages.

- σ_m is the length scale of the m-th predictor. Separate length scales have been employed for each predictor, using the approach of ARD. The value assumed by the σ_m when performing model training determines the extent at which a given predictor is influential on the response: small lengthscales refer to functions characterized by quick variations, while larger values are typical of very smooth functions with respect to the given predictor. Moreover, the lengthscale indicates how far a reliable extrapolation can be conducted: the bigger it is, the more reliable the extrapolation.

- σ_n is the noise variance, that expresses the amount of noise present in the data.

- α_{rq} is the shape parameter of the rational quadratic kernel, which determines the relative weighting between small and large scale variations.

Predictive capabilities of the GPR models for the primary variables have been tested on 38, 97 and 221 datasets, respectively. Tab. 5.4 displays the average normalized root mean squared error for the prediction. This quantity has been computed as:

$$\text{NRMSE} = \text{mean}\Big(\frac{\sqrt{\text{MSE}}}{|y_{max} - y_{min}|}\Big), \tag{5.3}$$

where y_{max} and y_{min} are the maximum and minimum values observed for the given target, the mean is computed over all the test sets of a given process stage and MSE is the mean squared error, namely:

$$\text{MSE} = \frac{1}{n_p}\sum_{l=1}^{n_p}(y_l - y_l^*)^2, \tag{5.4}$$

where n_p is the number of predictions generated from a sample of n_p data points, y_l is the l-th observation for one target variable and y_l^* is the corresponding prediction. Lower values of the average NRMSE, suggest that the prediction quality for key variables of stages A and C is in general superior to those of stage B, with $\dot{V}_{tot}^{out}$ being the only exception. In any case, quite reliable predictions with reasonable confidence intervals are obtained for all stages, with just a limited number of testing sets resulting in poor model performance. Fig. 5.1, Fig. 5.2 and Fig. 5.3 show time profiles of the key variables employed in stages A, B and C, respectively. These plots have been chosen because they are representative of most of the results. Symbols + indicate data points, while the continuous red line and the shaded region represent the predictions and the 95% confidence region of GPR models, respectively. A slight underestimation in $x_{\text{CO}_2}^{out}$ is noticeable in Fig. 5.3 before and around the peak. A similar behaviour has been found in several test sets. In general, $x_{\text{CO}_2}^{out}$ is the most difficult variable to target, but prediction quality is still acceptable. The predicted time-profiles of all variables are characterized by good degree of smoothness, confirming the effectiveness of GPR in acting as a smoothing device. This is a desirable feature since it allows to deal with noisy measurements, such as flow rates, as shown in Fig. 5.1a. Notice that a further improvement of the models employed in this work could be achieved by constructing kernel functions tailored to the given output. For example, with reference to the same Figure, an oscillatory pattern is slightly distinguishable. For the purpose of this work, catching this behaviour has not been considered relevant, but this could be achieved by summing k_{se} to a covariance function that includes a periodic term. This allows to grasp both the long-term trend and the short-term periodicity.

The use of these GPR models in the online determination of the switching time is

Table 5.1: Values of $\boldsymbol{\theta}$ obtained training the GPR models to predict variables of interest for stage A. σ_1, σ_2, σ_3 refer to length scales of the predictors t, P_G^0 and S/C, respectively.

	σ_1	σ_2	σ_3	α_{rq}	σ_f	σ_n
$\dot{V}_{tot}^{out}$	0.60	2.30	3.28	–	7.86	0.67
$x_{\mathrm{H_2}}^{out}$	14.8	4.7	42.3	0.002	0.88	0.014
$x_{\mathrm{CO_2}}^{out}$	2.14	1.92	4.39	0.108	0.08	0.007

Table 5.2: Values of $\boldsymbol{\theta}$ obtained training the GPR models to predict variables of interest for stage B. σ_1, σ_2, σ_3 refer to length scales of the predictors t, P_G^0 and $x_{\mathrm{O_2}}^0$, respectively.

	σ_1	σ_2	σ_3	α_{rq}	σ_f	σ_n
$\dot{V}_{tot}^{out}$	1.60	0.006	1.07	–	64.30	0.13
$x_{\mathrm{O_2}}^{out}$	3.07	0.017	3.44	0.043	0.066	0.001
$x_{\mathrm{CO_2}}^{out}$	1.38	0.004	1.90	0.43	0.006	0.001

discussed in the following Section.

5.2 Real-time determination of the switching time

The online determination of the switching times has been simulated for each process stage on several test sets. Objective function and constraints are evaluated at any time instant. The value of the objective function at any time is compared to the one at the previous instant, to determine whether to switch or not.

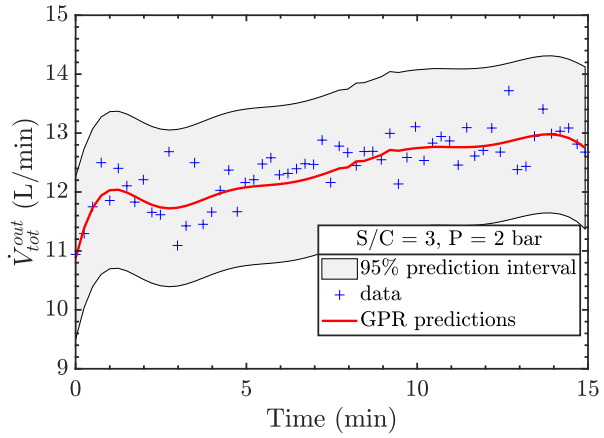
5.2.1 Stage A

As far as stage A is concerned, the real-time adaptation of criterion $A2$ has been coded as follows:

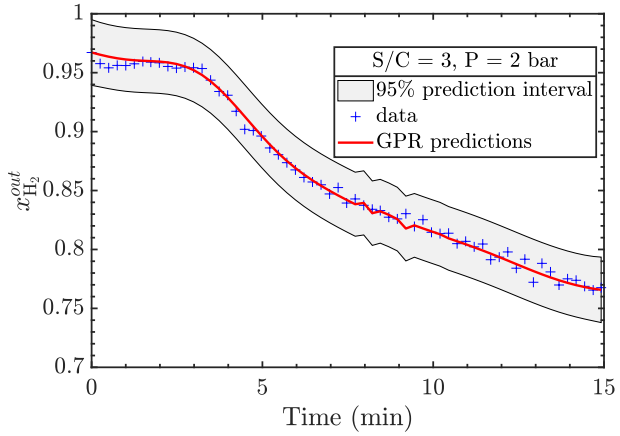
1: **while** $\Delta P \geq 0$ and $x_{\mathrm{H_2}}^{out*} \geq \varepsilon$ **do** ▷ Conditions to continue stage A
2: <body> ▷ Compute obj. function and constraints using GPR models
3: **end while**

Table 5.3: Values of $\boldsymbol{\theta}$ obtained training the GPR models to predict variables of interest for stage C. σ_1, σ_2, σ_3 refer to length scales of the predictors t, P_G^0 and $x_{\mathrm{H_2}}^0$, respectively.

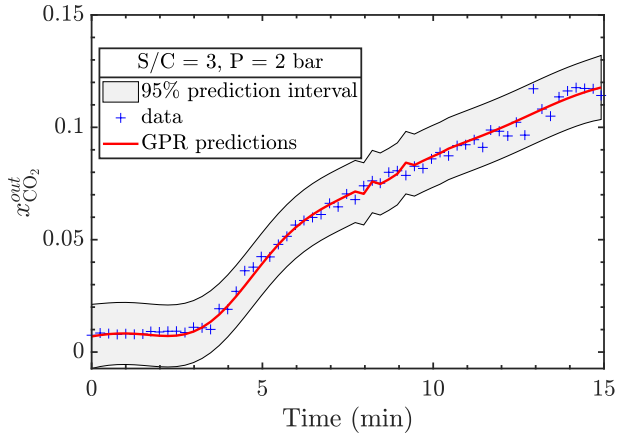
	σ_1	σ_2	σ_3	α_{rq}	σ_f	σ_n
$\dot{V}_{tot}^{out}$	0.47	4.81	0.73	–	54.48	2.89
$x_{\mathrm{H_2}}^{out}$	1.29	15.35	0.12	0.043	0.29	0.005
$x_{\mathrm{CO_2}}^{out}$	0.29	1.07	0.97	7.10	0.03	0.005



(a) $\dot{V}_{tot}^{out}$

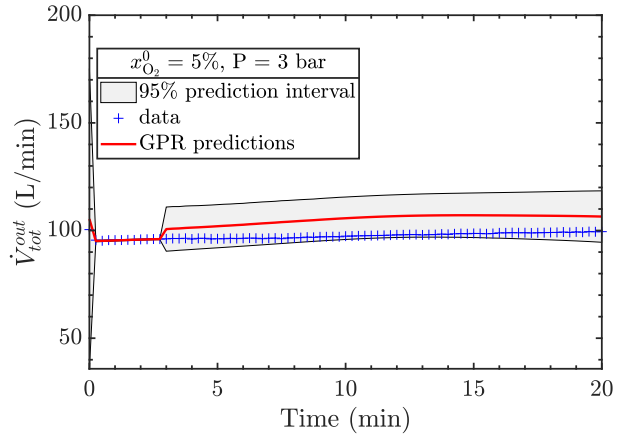


(b) $x_{H_2}^{out}$

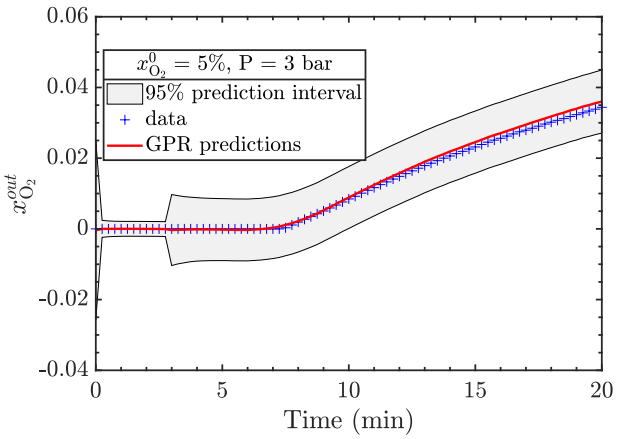


(c) $x_{CO_2}^{out}$

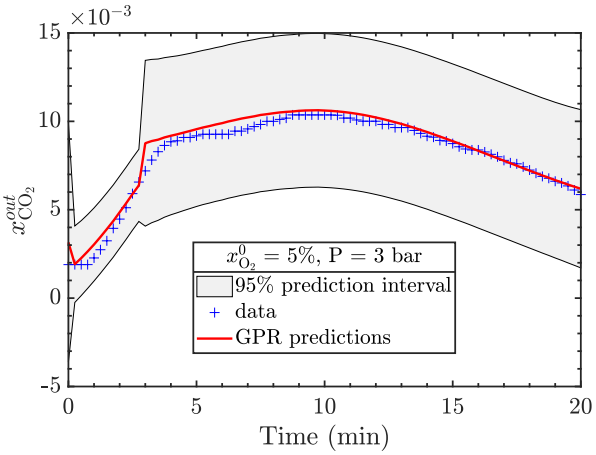
Figure 5.1: Time-profiles of the primary variables used to compute the optimal switching time one test set of stage A at S/C = 3, $P_G^0 = 2$ bar. Panels: (a) $\dot{V}_{tot}^{out}$; (b) $x_{H_2}^{out}$; (c) $x_{CO_2}^{out}$. Red lines represent predictions obtained via GPR with $\boldsymbol{\theta}$ in Tab. 5.1, shaded regions show the 95% confidence region, + the data points.



(a) $\dot{V}^{out}_{tot}$

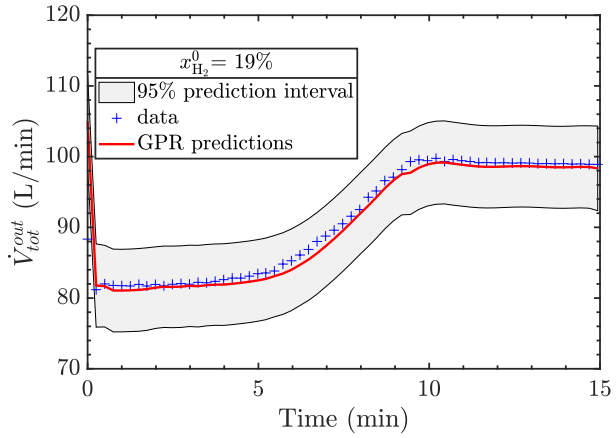


(b) $x^{out}_{O_2}$

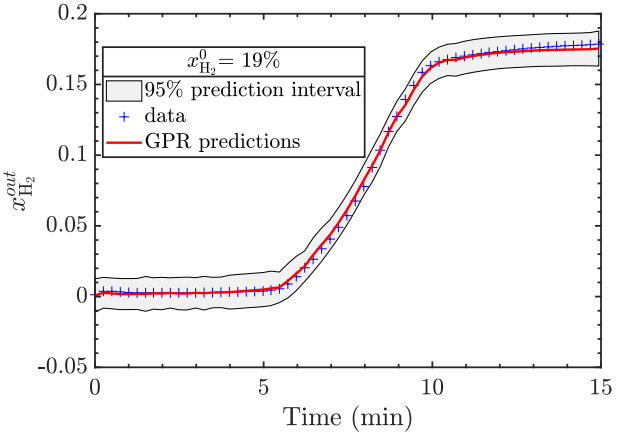


(c) $x^{out}_{CO_2}$

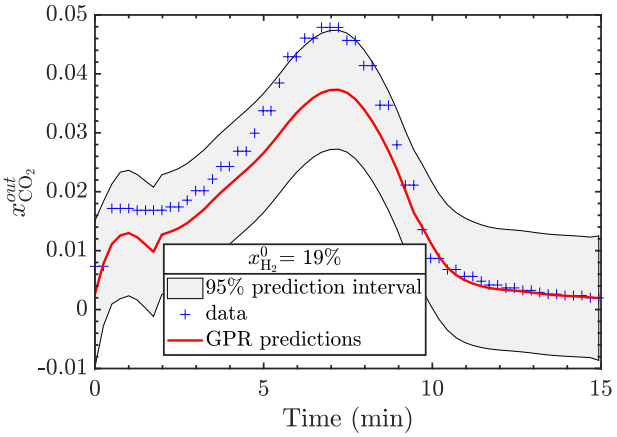
Figure 5.2: Time-profiles of the primary variables used to compute the optimal switching time one test set of stage B at $x^0_{O_2} = 5\%$, $P^0_G = 3$ bar. Panels: (a) $\dot{V}^{out}_{tot}$; (b) $x^{out}_{O_2}$; (c) $x^{out}_{CO_2}$. Red lines represent predictions obtained via GPR with $\boldsymbol{\theta}$ in Tab. 5.1, shaded regions show the 95% confidence region, + the data points.



(a) $\dot{V}^{out}_{tot}$



(b) $x^{out}_{H_2}$



(c) $x^{out}_{CO_2}$

Figure 5.3: Time-profiles of the primary variables used to compute the optimal switching time one test set of stage B at $x^0_{H_2} = 19\%$. Panels: (a) $\dot{V}^{out}_{tot}$; (b) $x^{out}_{O_2}$; (c) $x^{out}_{CO_2}$. Red lines represent predictions obtained via GPR with $\boldsymbol{\theta}$ in Tab. 5.1, shaded regions show the 95% confidence region, + the data points.

Table 5.4: Average NRMSE (%) of the predicted outputs in the test datasets for the datasets of three process stages.

Stage A	NRMSE	Stage B	NRMSE	Stage C	NRMSE
$\dot{V}_{tot}^{out}$	5.21	$\dot{V}_{tot}^{out}$	2.71	$\dot{V}_{tot}^{out}$	3.90
$x_{H_2}^{out}$	4.02	$x_{O_2}^{out}$	10.9	$x_{H_2}^{out}$	3.13
$x_{CO_2}^{out}$	5.22	$x_{CO_2}^{out}$	9.57	$x_{CO_2}^{out}$	4.71

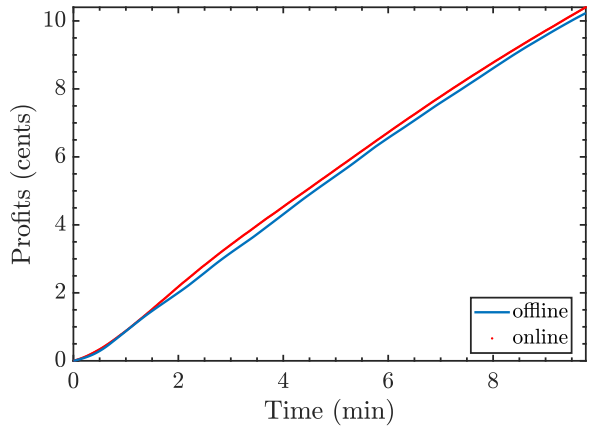


Figure 5.4: Example: Profits vs. time for one dataset of stage A (SER). Nominal operating conditions: SC = 5, $P_G^0 = 7$ bar. Color key: (red) online determination using predictions of a GPR model; (blue) offline determination using data.

where

$$\Delta P = P(t) - P(t-1) \tag{5.5}$$

is the delta of profits between two subsequent time instants. P is computed according to (3.6). The superscript * will be used hereinafter to distinguish the predicted values from the measured ones. The real-time approach yields to satisfactory performance. As an example, consider the following experimental run of SER at SC = 5, $P_G^0 = 7$ bar. Datasets at these operating conditions exhibited high variability in the optimal switching time. As stated in Chapter 3, $\bar{t}_s$ has been found to be 11.59 min with a SD_{t_s} of 0.99 min. Optimal switching time for this dataset has been found to be 9.43 min. t_s computed online employing GPR predictions resulted in 9.80 min, that is a variation of only 3.89% compared to the value obtained employing data, proving the reliability of the employed model. It should be noticed that an error of 22% would have been committed by setting *a priori* the switching time as the average of the values determined offline. In addition, by stopping the stage after 11.59 min would have not allowed to satisfy the constraint on H_2 purity.

To account for minor process-model mismatches, as soon as delayed data are acquired, the determination of the optimal switching time is computed offline, after the

batch is ended. Results are compared with predictions and an adjustment is employed in the subsequent experimental run. In the present work, to account for the slight underestimation in $x_{\mathrm{H_2}}^{out*}$ in some runs, a correction on the constraint ε has been done according to:

$$\varepsilon' = \varepsilon \times (1 - dk), \tag{5.6}$$

where

$$dk = \mathrm{mean}\left(\frac{|\boldsymbol{x}_{\mathrm{H_2}}^{out} - \boldsymbol{x}_{\mathrm{H_2}}^{out*}|}{\boldsymbol{x}_{\mathrm{H_2}}^{out}}\right). \tag{5.7}$$

Vectors of actual and predicted H_2 mole fractions contain values based on the last 75 s of the previous experimental run. With this correction a premature termination of stage A is avoided, since the constraint is typically active. As complexity of the employed profits function will grow when up-scaling the process by including more terms, analogous corrections are suggested for all the primary variables to avoid finding predicted optima that do not coincide with the actual process ones.

5.2.2 Stages B and C

As far as stages B and C are concerned, a modification of criterion $B2$ (3.13) has been adopted. As shown in Chapter 3, this criterion yields to results that are very similar to those of $B3$ and $B4$, being at the same time the simplest to be adapted online. Recalling (3.15) and (3.16), criteria $B3$ and $B4$ require to set an upper limit (or a lower limit, in the case of stage C) to the amount of CO_2 released. While in the offline case this was easily formulated as a ratio between the outlet CO_2 flow rate at a given time and the total CO_2 released during the complete time span explored, applying this criterion in real-time becomes much more complicated. One simple implementation would consist in setting the constraint values for $\dot{M}_{\mathrm{CO_2}}^{out}(t)$. However, the amount of CO_2 that might be released during stage B is affected by the amount of CO_2 captured during A. Similarly, the amount of CO_2 that can be released during stage C is dependent on the amount of CO_2 that has not been released during B. Therefore, setting reasonable values for these constraints would have not been trivial. A simple pseudo-code sketching the online adaptation of $B2$ is as follows:

1: **while** $\Delta T_s^{\mathrm{out}} \geq 0$ and $\dot{M}_{\mathrm{i}}^{out*}(t)/\dot{M}_{\mathrm{i}}^{0}(t) \leq \varepsilon_{1,k}$ **do** ▷ Conditions to continue stage B and C
2: <body> ▷ Compute constraints using GPR models
3: **end while**

where $\Delta T_s^{\mathrm{out}}$ represents the variation in the value of T_s^{out} measured in two subsequent sampling times:

$$\Delta T_s^{\mathrm{out}} = T_s^{\mathrm{out}}(t) - T_s^{\mathrm{out}}(t-1) \tag{5.8}$$

and $\dot{M}_{\mathrm{i}}^{out*}(t)$ stands for the value of the mass flow rate of O_2 and H_2 in stages B and C, respectively, computed with the ideal gas law employing predictions $\dot{V}_{tot}^{out*}$ and x_{i}^{out*}. Unlike stage A, the objective function is a variable that can be measured in real-time. Thermocouples dynamics is usually fast enough to make data acquisitions available almost in real-time. Conversely, the constraints still require the use of the model. The use of the real-time approach makes possible to deal with disturbances in those process variables that have been employed as predictors.

5.3 Model Retraining

In the previous Section, a correction on predictions has been made to tackle minor process-model mismatches. In general, there exist several reasons yielding to relevant process-model mismatches. Examples include changes to operating conditions differing from those on which training had been performed. Other possible causes are variations of catalyst activity or solid morphology, resulting from aging and poisoning after several cycles or presence of unmeasured disturbances entering the system. These side phenomena always determine variations in process performance, that would cause drifts in the time profiles of the primary variables, ultimately leading to the inability of GPR models to catch their profiles. When this is the case, simple corrections such (5.7) would be insufficient and model retraining becomes necessary to restore reliable predictions. In this work, a methodology for model retraining has been tested.

If substantial process-model mismatches are found for more than 3 cycles, application of the switching criterion is stopped for 3 runs and stage duration is prolonged to acquire data in a sufficiently broad time horizon. Recall from the Introduction that extrapolation capabilities of data-driven models are poor, as the information they provide is essentially the one embedded within the data employed for their construction. Therefore, stage prolongation is advisable to avoid unreliable predictions in the last moments of the stage, due to the use of a model trained only on data collected before the switching is carried out. Indeed, uncertainty in predictions grows in those regions that are far from training data. In the present work, predictions are considered unsatisfactory when the following condition is verified:

$$\mathrm{MSE} > (1 + d)\,\mathrm{MSE}_{ref}, \tag{5.9}$$

that is when the mean squared error for the predictions on the current run is bigger of a given percentage d (in this work, set equal to 0.3) than a reference MSE_{ref}. When this condition is verified, the dataset corresponding to the current experimental run is stored for the possible training of a new GPR model.

After 6 inaccurate predictions on some primary variables, the corresponding *shortcut*

GPR models are trained. These models were called like this, since they are trained just on a limited number of datasets. As stated in Chapter 4, the computational time needed to train a GPR is $\mathcal{O}(n^3)$. Therefore, a small number of datasets must be employed to quickly retrain a GPR model that is able to catch the profile of the variables, otherwise too many cycles would be conducted without employing reliable predictions. If the so trained shortcut GPR models are still unsuitable to represent the real process behaviour, as process cycles are repeatedly conducted, new GPR models are trained employing the last 30 datasets collected. Retraining is performed asynchronously in parallel with the assessment of stages duration.

However, such a retraining scheme shows several limitations. First, only data collected during the last runs in sequential order are employed. This is not necessarily the most effective choice, and other criteria for the selection of the most appropriate datasets should be investigated. Moreover, employing the sole MSE does not guarantee a reliable assessment of the predictions quality. Tuning a proper value for both MSE_{ref} and d is not trivial, since they affect the frequency at which model retraining is carried out. Too low values of MSE_{ref} and d would cause a too frequent model retraining, while too high values might cause a new model to be not trained when needed. With this mind, a more detailed study should be conducted to identify more robust retraining methods. Since GPR automatically provides the confidence intervals of the predictions, this feature could also be exploited. In any case, it is expected that the need for model retraining will disappear after several runs. The idea is that, as datasets spanning different operating conditions will become available, many *local* GPR models will be trained for the given set of operating conditions and stored in a model library. Then, output variables will be predicted in real-time by loading the model that corresponds to the given query state, thus that the need for model retraining becomes more and more rare.

5.4 Summary

A real-time adaptation of criteria $A2$ and $B2$ presented in Chapter 3 has been proposed. Since the formulation of these criteria lies on time-profiles of variable that cannot be measured online, GPR-based soft sensors have been developed. Implementation of real-time criteria resulted in excellent performance for most of the experimental runs; when this is not the case, model retraining might become necessary. Model retraining has been assessed only partially in the present work and deeper investigations are suggested for the future. However, it is expected that collecting datasets in a broad range of operating conditions will allow to train several *local* GPR models to perform predictions in a broad range of operating conditions, thus decreasing the need for model retraining.

Conclusions and Outlooks

The operability aspects of a novel concept for H_2 production with integrated CO_2 capture have been investigated in this work. The process comprises three stages, conducted sequentially in a packed bed reactor: (A) SER, (B) oxidation of Cu and (C) regeneration of CaO combined with reduction of CuO (Martini, Berg, *et al.*, 2016). As each stage is inherently dynamic, the assessment of the optimal duration of each of them is crucial: switching from one reactor mode to the following should be performed at a time such that satisfactory operating requirements, like process profitability and quality targets, are achieved. With this in mind, a suitable methodology to assess the optimal stage durations has been proposed.

Time profiles of several variables, obtained from measurements in a lab scale set up, have been investigated to identify the most appropriate way to monitor the course of each stage. Time-profiles of outlet compositions and outlet temperature of the solid bed resulted to be suitable for the purpose. Conversely, axial profiles of the solid bed temperature turned out to be unsuitable, especially for stage A.

As a further step, the optimal duration of each stage has been determined by means of an optimization-based approach. Two switching criteria have been proposed for stage A and four for stage B and C. All of them rely on time-profiles of the variables selected through the preliminary data analysis. These criteria aim at matching different and, sometimes, conflicting objectives. For stage A, the requirements are to maximize economic profitability, while keeping H_2 purity as high as possible. During stage B, the objective is to accomplish Cu oxidation and minimize undesired release of CO_2. Criteria for stage C aim at maximizing the release of CO_2 from $CaCO_3$ calcination, while guaranteeing the operation to be energy neutral. The offline application of the switching criteria to the available datasets resulted in a wide dispersion of the optimal switching times for stage A. Conversely, a low variability has been observed in the experimental runs for stages B and C.

Besides the offline determination, a real-time methodology has been proposed. Due to the variability among the batches, the real-time approach is mandatory to guarantee optimal operation during stage A. Although dispersion in the results of stages B and C was low, a real-time approach seemed necessary, to handle disturbances and process uncertainties, that are common during operation. Since the proposed switching criteria require the availability of time-profiles of the key variables, such as compositions, whose measurements are often delayed, an inferential model has been developed. Specifically, GPR models have been employed as soft sensors. The inferred primary variables were

$x^{out}_{\mathrm{H_2}}$, $x^{out}_{\mathrm{CO_2}}$ in stage A, $x^{out}_{\mathrm{O_2}}$ in stage B and $x^{out}_{\mathrm{H_2}}$ in stage C, respectively. The total *dry* volume flow rate exiting the reactor, $\dot{V}^{out}_{tot}$ has also been estimated. The time-profile of this variable is needed to compute species flow rates, since all the composition data in the present work were available on a dry basis. Measurements corresponding to $\dot{V}^{out}_{tot}$ cannot be available in real-time, due to delays associated to steam condensation, hence, the need for predicting it. GPR models have been trained to target $x^{out}_{\mathrm{CO_2}}$ also for steps B and C. This variable has not been employed in the online determination for these two process stages, but it is included in the formulations of other offline criteria. Therefore, assessing the performance of the proposed GPR-models to catch its profile seemed reasonable in the light of a future real-time adaptation of these criteria.

The following secondary variables have been employed as predictors: t, P^0_G, S/C for the outputs of stage A, t, P^0_G, $x^0_{\mathrm{O_2}}$ of stage B and t, P^0_G $x^0_{\mathrm{H_2}}$ of stage C. The number of observations employed in training for the three stages is 1464, 3927 and 1791 corresponding to 24, 49 and 12 datasets, respectively. Due to the relatively small number of variables that are sensed in the employed set up, the preparatory data analysis was sufficient to select the predictors. In situations wherein a big number of measurements will be available, such as a real plant, the use of statistical techniques to select the most suitable ones may be required.

The implementation of GPR models for soft-sensing yields to satisfactory performances. Minimal overfitting has been observed in most predictions. The real-time approach has guaranteed an accuracy of 5% in the calculation of the optimal switching times. However, process-model mismatches can be found. They might lead to predicted optima that differ from the real ones. Corrections of process-model mismatches have been only partially addressed in the present work, but a proper investigation is suggested. Useful techniques might be derived from the literature in the field of real-time optimization (Hille and Budman, 2018).

One issue expected when conducting the Ca-Cu chemical looping process concerns degradation of performances after a given number of cycles, because of phenomena like aging and poisoning. Variations in the time-profiles of the primary variables are expected. If this is the case, a model retraining has been proposed, according to two methodologies: (i) training of a *shortcut* model and (ii) training of a new GPR model. Shortcut models have been trained only on those datasets acquired in the last cycles. (Specifically 6 in the present work). Extrapolation capabilities of the shortcut models are low, because of the limited number of data used for training, but they allow to restore predictions while waiting for the sufficient amount of data needed to entirely retrain the GPR model (ii) to become available. In any case, the need for model retraining is expected to disappear as a sufficiently large number of datasets spanning several operating conditions will become available: all the *local* GPR models trained

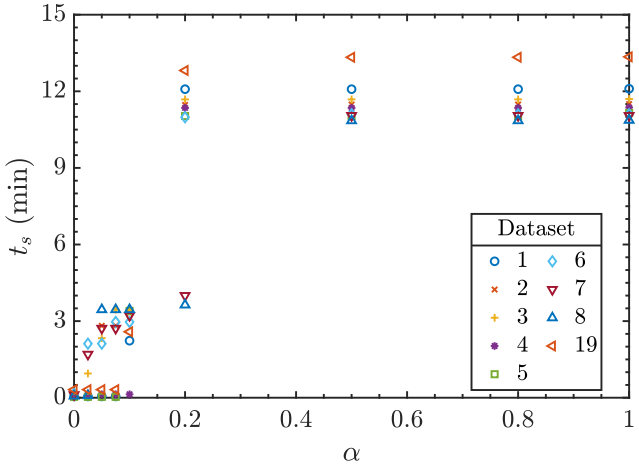
on a specific range of operating conditions will be stored in a library and the output variables will be predicted in real-time by loading the model corresponding to the given query state.

Appendix

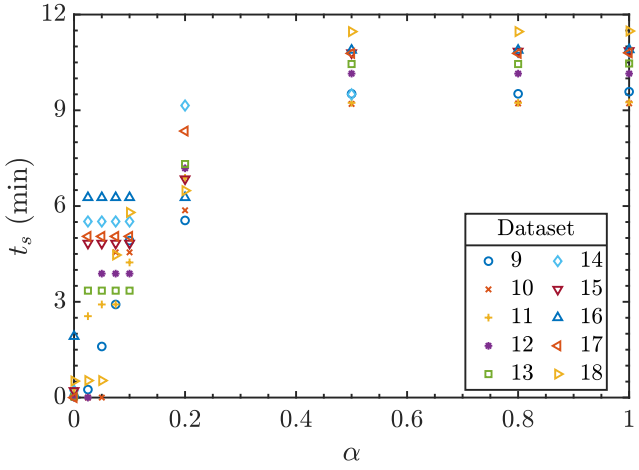
Appendix

THIS APPENDIX shows the results obtained by applying the criterion $A1$, as described in Chapter 3. Results are in the form of plots, showing t_s vs. α for all the datasets employed in this work. For sake of clarity, Figures referring to the same value of S/C at different P_G^0 are gathered together.

The following plots display t_s vs α for all the datasets at S/C = 3, with operating pressure $P_G^0 = 2\,\text{bar}$ and $P_G^0 = 7\,\text{bar}$, respectively.



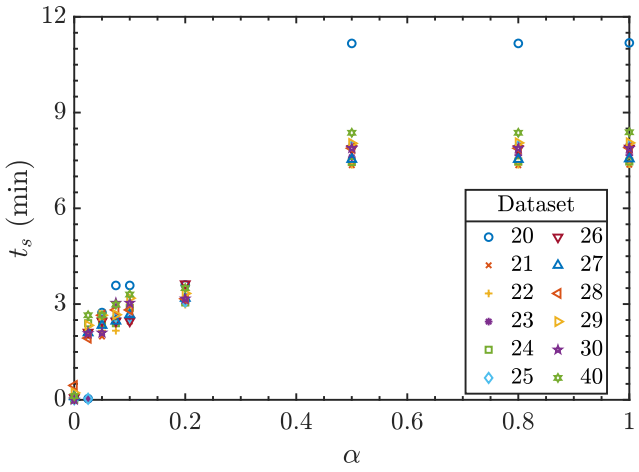
(a) S/C = 3, $P_G^0 =$ 2 bar.



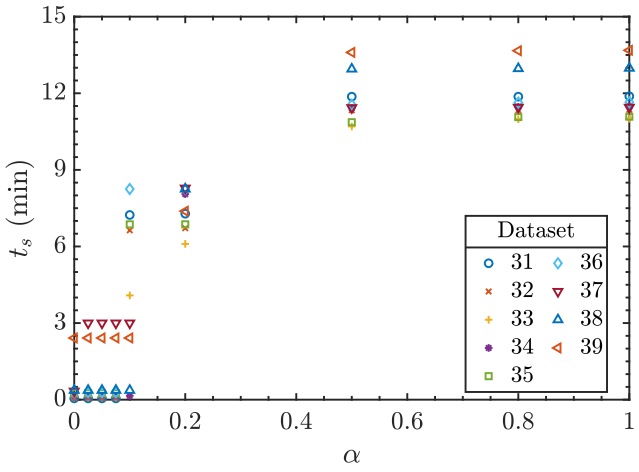
(b) S/C = 3, $P_G^0 =$ 7 bar.

Figure A.1: t_s as a function of α for all the employed datasets at S/C = 3: (a) $P_G^0 = 2$ bar; (b) $P_G^0 = 7$ bar.

The following plots show the values of t_s vs α for all the datasets at S/C = 4, with operating pressure $P_G^0 = 2\,\text{bar}$ and $P_G^0 = 7\,\text{bar}$, respectively.



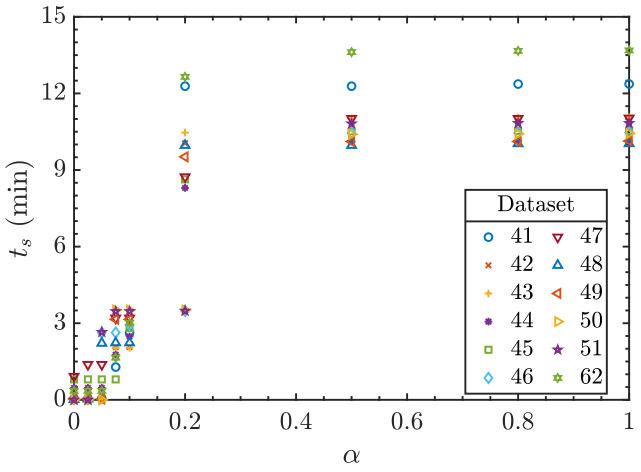
(a) S/C = 4, $P_G^0 =$ 2 bar.



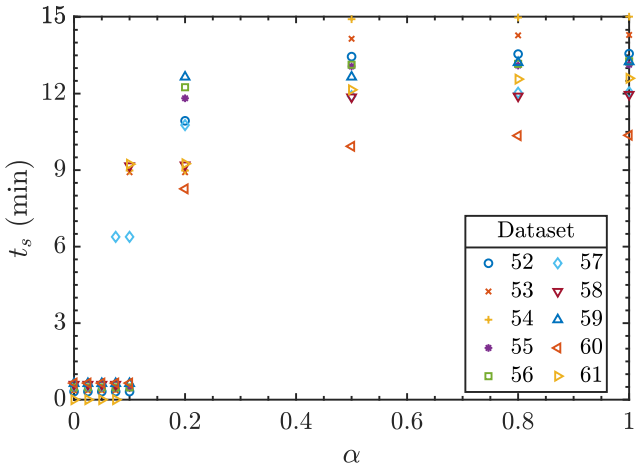
(b) S/C = 4, $P_G^0 =$ 7 bar.

Figure A.2: t_s as a function of α for all the employed datasets at S/C = 4: (a) $P_G^0 = 2$ bar; (b) $P_G^0 = 7$ bar.

The following plots shows the values of t_s vs α for all the datasets at S/C = 5, with operating pressure $P_G^0 = 2\,\text{bar}$ and $P_G^0 = 7\,\text{bar}$, respectively.



(a) S/C = 5, $P_G^0 =$ 2 bar.



(b) S/C = 5, $P_G^0 =$ 7 bar.

Figure A.3: t_s as a function of α for all the employed datasets at S/C = 5: (a) $P_G^0 = 2$ bar; (b) $P_G^0 = 7$ bar.

Notation

$\boldsymbol{x}$	vector containing species mole fractions in the gas phase
$\dot{M}$	mass flow rate ($\mathrm{kg\,min^{-1}}$)
$\dot{M}$	mass flow rate ($\mathrm{kg\,min^{-1}}$)
$\dot{m}$	mass flux ($\mathrm{kg\,m^{-2}\,s^{-1}}$)
$\dot{V}$	volume flow rate ($\mathrm{L\,min^{-1}}$) or ($\mathrm{m^3\,min^{-1}}$)
$\Delta_{\mathrm{r}}H^{\ominus}$	reaction enthalpy ($\mathrm{J\,mol^{-1}\,K^{-1}}$)
η_{CC}	carbon capture efficiency
$\mathcal{E}$	expected value
$\mathcal{F}$	objective function
$\mathcal{GP}$	Gaussian Process. $f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}'))$ means that the function f is distributed as a GP, whose mean function is $m($**x** and covariance function $k(\mathbf{x}, \mathbf{x}')$.
$\mathcal{N}$	normally distributed
t_s	switching time for a given stage, i.e. duration of that stage (min)
$\mathbf{f}$	vector of latent function values $\mathbf{f} = (f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)^\top)$
CT	carbon tax (€/$\mathrm{kg_{CO_2}}$)
D_p	diameter of a solid particle (mm)
F	objective function for a given switching criterion
I	identity matrix
K	covariance matrix
$k(\mathbf{x}, \mathbf{x}')$	covariance function (kernel) evaluated at $\mathbf{x}$ and $\mathbf{x}$
K_*	covariance matrix between training and test cases
K_{**}	covariance matrix for the test cases
$m(\mathbf{x})$	mean function of a Gaussian Process
MW	molar weight ($\mathrm{kg\,mol^{-1}}$)
P	pressure (bar, if not otherwise specified)
P	profits (€/min)
p	price (€/kg)
R	universal gas constant, $8.314\,\mathrm{J\,mol^{-1}\,K^{-1}}$
SD	standard deviation
T	temperature (°C) or (K)
t	time (min)

X	matrix of the training inputs $\{\mathbf{x}_i\}_{i=1}^n$
x	dry gas mole fraction
X_*	matrix of the test inputs
R	reaction
S/C	steam-to-carbon ratio

Greek Letters

α	weighting factor in multi-objective optimization
$\boldsymbol{\theta}$	vector of the hyperparameters of the covariance function
$\delta_{ii'}$	Kroeneker delta
$\hat{\rho}\rho$	density ($\mathrm{kg\,m^{-3}}$)
σ^2	variance
τ	variable of integration
ε	value set for the constraint

Superscripts

$*$	predicted value
out	outlet
0	inlet

Subscripts

$A1 - B4$	criteria for computing switching time
f	noise free signal
G	gas phase
i	species i
k	process stage
l	index of objective functions
n	noise
ref	reference
rq	rational quadratic
s	solid phase
se	squared-exponential

Acronyms

ARD	automatic relevance determination
CCS	carbon capture and storage
CV	cross-validation
GDP	gross domestic product

GHG	greenhouse gases
GP	Gaussian Process
GPR	Gaussian Process Regression
IPPC	Intergovernmental Panel on Climate Change
MSE	mean squared error
RMSE	root mean squared error
SER	Sorption Enhanced Reforming
SMR	steam methane reforming
WGS	water gas shift reaction

References

Abanades, J.C., R. Murillo, J.R. Fernandez, G. Grasa, and I. Martínez (2010). "New CO_2 capture process for hydrogen production combining Ca and Cu chemical loops". In: *Environmental Science and Technology* 44.17, pp. 6901–6904. DOI: 10.1021/es101707t.

Alhammadi, H.Y. and José Romagnoli (2004). "Process design and operation: Incorporating environmental, profitability, heat integration and controllability considerations". In: pp. 264–305.

Archer, David (2005). "Fate of fossil fuel CO_2 in geologic time". In: *Fate of Fossil fuel CO_2 in geologic time* 110, pp. 1–6. DOI: 10.1029/2004JC002625.

Berg, A. van den (2015). "A new CO_2 capture process for pure H_2 production combining Ca and Cu chemical loop". Technische Universiteit Eindhoven.

Bernardi, A. (2010). "Identificazione di sistemi non lineari tramite regressione Gaussiana e applicazione al Wiener-Hammerstein benchmark". Università degli Studi di Padova.

Bezzo, F. (2018). "Process Design". Course material for the a.y. 2017 - 2018.

Bonvin, D., C. Georgakis, C. C. Pantelides, M. Barolo, M. A. Grover, D. Rodrigues, R. Schneider, and D. Dochain (2016). "Linking Models and Experiments". In: *Industrial & Engineering Chemistry Research* 55.25, pp. 6891–6903. DOI: 10.1021/acs.iecr.5b04801.

carbontax.org, (2017). *Carbon Tax Center*. URL: https://www.carbontax.org/.

Corona, Francesco, Michela Mulas, Roberto Baratti, and Jose A. Romagnoli (2012). "Data-Derived Analysis and Inference for an Industrial Deethanizer". In: *Industrial & Engineering Chemistry Research* 51.42, pp. 13732–13742. DOI: 10.1021/ie202854b.

Druijff, P.M. (2018). "Proof of Concept of the Ca-Cu looping process". Technische Universiteit Eindhoven.

Ebden, M. (2015). "Gaussian Processes: A Quick Introduction". In: *ArXiv e-prints*. arXiv: 1505.02965 [math.ST].

Edgar, T.F., D.M. Himmelblau, and L.S. Lasdon (2001). *Optimization of chemical processes*. McGraw-Hill chemical engineering series. McGraw-Hill. ISBN: 9780070393592.

EIA (2016). *International Energy Outlook*. Report. U.S. Energy Information Administration.

Fraile, Daniel, Jean-Christophe Lanoix, Patrick Maio, Azalea Rangel, and Angelica Torres (2015). *Overview of the market segmentation for hydrogen across potential customer groups, based on key application areas*. Report. CertifHy.

Gupta, Amit (2018). "Introduction to Deep Learning: Part 1". In: *CEP*.

Hamers, H.P. (2015). "Packed bed chemical-looping combustion : experimental demonstration and energy analysis". PhD thesis. Technische Universiteit Eindhoven. ISBN: 978-90-386-3757-0. DOI: 10.6100/IR783193.

Helmi, Arash (2017). "Fluidized bed membrane reactors for water gas shift: mass transfer, hydrodynamics and experimental demonstration". PhD thesis. Department of Chemical Engineering and Chemistry. ISBN: 978-90-386-4243-7.

Hille, Rubin and Hector M. Budman (2018). "Experimental Design in Simultaneous Identification and Optimization of Batch Processes under Model-Plant Mismatch". In: *Preprints, 10th IFAC International Symposium on Advanced Control of Chemical Processes Shenyang, Liaoning, China*. IFAC, pp. 791–796.

Hinicio (2015). *internal Information*.

howstuffworks.com, (2018). *How Carbon Tax Works*. URL: https://science.howstuffworks.com/environmental/green-science/carbon-tax.htm.

Joglekar, Girish (2017). "Using Simulation for Scheduling and Rescheduling of Batch Processes". In: *Processes* 5.66. ISSN: 2227-9717. DOI: 10.3390/pr5040066.

Kikuchi, Eiichi (1995). "Palladium/ceramic membranes for selective hydrogen permeation and their application to membrane reactor". In: *Catalysis Today* 25.3. Catalysis in Membrane Reactors, pp. 333–337. ISSN: 0920-5861. DOI: https://doi.org/10.1016/0920-5861(95)00085-T.

Liuzzi, G (1999). "Appunti delle Lezioni di Ottimizzazione: Programmazione Multiobiettivo". Course material for the a.y. 1999 - 2000.

Mackay, David J.C. (1998). "Introduction to Gaussian Processes". In: *NATO ASI Series F Computer and Systems Sciences* 168, pp. 133–166.

Martini, M., A. van den Berg, F. Gallucci, and M. van Sint Annaland (2016). "Investigation of the process operability windows for Ca-Cu looping for hydrogen production with CO_2 capture". In: *Chemical Engineering Journal* 303, pp. 73–88. ISSN: 1385-8947. DOI: 10.1016/j.cej.2016.05.135.

Martini, M., I. Martínez, M.C. Romano, P. Chiesa, F. Gallucci, and M. van Sint Annaland (2017). "Increasing the carbon capture efficiency of the Ca/Cu looping process for power production with advanced process schemes". In: *Chemical Engineering Journal* 328, pp. 304–319. ISSN: 1385-8947. DOI: 10.1016/j.cej.2017.07.048.

Medrano, J.A. (2017). "Membrane-assisted chemical looping reforming: from fundamentals to experimental demonstration". PhD thesis. Department of Chemical Engineering and Chemistry. ISBN: 978-90-386-4329-8.

Meerman, J.C., E.S. Hamborg, T. van Keulen, A. Ramírez, W.C. Turkenburg, and A.P.C. Faaij (2012). "Techno-economic assessment of CO2 capture at steam methane reforming facilities using commercially available technology". In: *International Jour-*

nal of Greenhouse Gas Control 9, pp. 160–171. ISSN: 1750-5836. DOI: https://doi.org/10.1016/j.ijggc.2012.02.018.

Mei, C., M. Yang, D. Shu, H. Jiang, G. Liu, and Z. Liao (2016). "Soft sensor based on Gaussian process regression and its application in erythromycin fermentation process". In: *Chemical Industry and Chemical Engineering Quarterly* 22.2, pp. 127–135. DOI: 10.2298/CICEQ150125026M.

platts.com, (2017). *New Dutch coalition government plans CO_2 floor price of Eur18/mt: lawmaker*. URL: https://www.platts.com/latest-news/coal/london/new-dutch-coalition-government-plans-co2-floor-26822471.

Noorman, Sander, Martin van Sint Annaland, and Kuipers (2007). "Packed Bed Reactor Technology for Chemical-Looping Combustion". In: *Industrial & Engineering Chemistry Research* 46.12, pp. 4212–4220. DOI: 10.1021/ie061178i.

Paris Agreement (2015). United Nations.

Porru, M., J. Alvarez, and R. Baratti (2013). "A distillate composition estimator for an industrial multicomponent IC4-NC4 splitter with experimental temperature measurements". In: vol. 10. PART 1, pp. 391–396. DOI: 10.3182/20131218-3-IN-2045.00028.

Porru, M. and L. Özkan (2017). "Monitoring of Batch Industrial Crystallization with Growth, Nucleation, and Agglomeration. Part 2: Structure Design for State Estimation with Secondary Measurements". In: *Industrial and Engineering Chemistry Research* 56.34, pp. 9578–9592. DOI: 10.1021/acs.iecr.7b00243.

ICCP, (2007). *Projections of Future Changes in Climate*. URL: https://www.ipcc.ch/publications_and_data/ar4/wg1/en/spmsspm-projections-of.html.

Rasmussen, C. E. (2004). "Gaussian processes in machine learning". In: ed. by Springer Verlag, pp. 63–71.

Rasmussen, C. E. and C. Williams (2006). *Gaussian Processes for Machine Learning*. the MIT Press.

Reay, David, Colin Ramshaw, and Adam Harvey (2013). *Process Intensification. Engineering for Efficiency, Sustainability and Flexibility*. Butterworth-Heinemann. ISBN: 9780080983042.

Rostrup-Nielsen, Jens R. and Thomas Rostrup-Nielsen (2002). "Large-Scale Hydrogen Production". In: *CATTECH* 6.4, pp. 150–159. ISSN: 1572-8811. DOI: 10.1023/A:1020163012266.

Russell, S. and P. Norvig (2010). *Artificial Intelligence: A Modern Approach*. Third. Upper Saddle River, NJ: Prentice Hall. ISBN: 9780136042594.

scikit-learn, (2018). *Scikit Learn - Gaussian Processes*. URL: http://scikit-learn.org/stable/modules/gaussian_process.html.

Seborg, D.E., D.A. Mellichamp, T.F. Edgar, and F.J. Doyle (2010). *Process Dynamics and Control*. John Wiley & Sons. ISBN: 9780470128671.

Shimizu, T., T. Hirama, H. Hosoda, K. Kitano, M. Inagaki, and K. Tejima (1999). "A Twin Fluid-Bed Reactor for Removal of CO_2 from Combustion Processes". In: *Chemical Engineering Research and Design* 77.1, pp. 62–68. ISSN: 0263-8762. DOI: https://doi.org/10.1205/026387699525882.

Shokry, Ahmed, Montserrat Pérez-Moya, Moisès Graells, and Antonio Espuña (2017). "Data-Driven Dynamic Modeling of Batch Processes Having Different Initial Conditions and Missing Measurements". In: pp. 433–438. DOI: 10.1016/j.compchemeng.2017.03.016.

Stern, Nicholas (2007). *The Economics of Climate Change: The Stern Review*. Cambridge University Press. DOI: 10.1017/CBO9780511817434.

gaussianprocess.org, (2006). *The Gaussian Process Web Site*. URL: http://www.gaussianprocess.org/#gpr.

Tuinier, M.J. (2011). "Novel process concept for cryogenic CO_2 capture". PhD thesis. Department of Chemical Engineering and Chemistry. ISBN: 978-90-386-2900-1. DOI: 10.6100/IR719418.

Acknowledgements

THIS SECTION IS DEDICATED to all the people who have directly or indirectly supported me during these six months in Eindhoven.

First of all, I would like to thank my daily supervisor Dr. Marcella Porru for her warm welcome in Eindhoven (in such a cold day). I am very grateful for your support, patience, the tips you gave me and all the inspiring conversations we had in these months. I did my best for this work and, of course ... for your orchids!

I am grateful to Professor Martin van Sint Annaland and Professor Fausto Gallucci for giving me the opportunity to work on this project. Thanks are also due to Dr. Leyla Özkan who gave me the great possibility to join the Control Systems group and to interact with people from several different backgrounds. These months have constituted an important sample of professional life for me.

My gratefulness goes to Professor Massimiliano Barolo for always being available when I needed his opinion and for his useful suggestions. I would also like to acknowledge Professor Paolo Canu for encouraging me to embrace this beautiful experience.

Thanks to Dr. Giulio Bottegal for introducing me to topics such as Gaussian Processes, that were completely new for me. Thank you also for our coffee breaks (the sound of your cup hitting my desk still threatens me) and for your sense of humour. You really entertained me every day.

I also wish to thank Michela Martini and Mike Druijff for sharing their experimental data with me and for being always available for explanations and clarifications.

I want to thank all the students of the Control Systems group for the time we spent together, our lunch breaks and the conversations we had about our theses. I am grateful to you, for making me feel part of your group.

I owe special thanks to the Erasmus students and the people from SPI group for all the beautiful moments we spent together. It is amazing to think how close we became in such a short period of time. You really made me feel at home.

Finally, I am infinitely grateful to all my friends and my family for unconditionally

supporting me in all moments of my life, even from afar. You are special. Thank you so much.