RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE

Faculté de Technologie

Thèse

Présentée au département d'Electronique
Pour l'obtention du diplôme de

Doctorat LMD

Domaine: Sciences et Technologie
Filière: Électronique
Option: Instrumentation Électronique
par:

Abdelhani CHAABNA

Thème

Identification des paramètres des systèmes non linéaires pour l'optimisation de la production de biogaz par voie biologique

Soutenue le 07 Mars 2024

Devant le Jury:

Abdelhalim MAYOUF	Professeur	Univ. Sétif 1	Président
Samia SEMCHEDDINE	Professeur	Univ. Sétif 1	Directrice de thèse
Mouloud AYAD	Professeur	Univ. Sétif 1	Examinateur
Nacira DIFFELLAH	MCA	Univ BBA	Examinatrice
Jérome HARMAND	Directeur de recherche	INRAE	Invité

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

Faculty of Technology

Thesis

Presented to Department of Electronic

This thesis is submitted in fulfillment of the requirement for the degree

LMD Doctorate

Field: Sciences and Technology

Sector: Electronic

Specialty: Electronic Instrumentation

Produced by:

Abdelhani CHAABNA

Title

Parameter identification of non linear systems for the optimization of biogas production via biological way

Graduation Day: March, 7th, 2024

Examination committee:

Abdelhalim MAYOUF	Professor	Univ. Setif 1	President
Samia SEMCHEDDINE	Professor	Univ. Setif 1	Thesis Director
Mouloud AYAD	Professor	Univ. Setif 1	Examiner
Nacira DIFFELLAH	MCA	Univ BBA	Examiner
Jérome HARMAND	Research Director	INRAE	Invited

Acknowledgments

Firstly, I would like to thank ALLAH for granting me health and determination to complete this doctoral thesis.

I warmly thank all the people who helped me during the preparation of my thesis. A special thanks to my thesis director, Prof. Samia SEMCHEDDINE, who has been a guiding light and a motherly figure throughout this journey. Her support, encouragement, and great availability have been invaluable. Her numerous pieces of advice have significantly contributed to the realization of my thesis.

I would also like to sincerely thank the members of the jury for agreeing to evaluate this thesis:
Mr. Abdelhalim MAYOUF, Professor at the University of Setif 1, for agreeing to judge my work and preside over the defense jury of this thesis. Messrs. Mouloud AYAD, Professor at the University of Setif 1 and Madam Nacira DIFFELLAH, M.C.A at the University of BBA, for having honored me by accepting to be the examiners of this thesis. I would also like to express my heartfelt appreciation for the time you have dedicated to reviewing this thesis. I am fully aware that this process requires a significant investment of time and intellectual effort.

I would like to express my deepest thanks to Jérome HARMAND, research director at the National Institute for Agricultural, Food and Environmental Research (INRAE) for his invaluable advice and guidance which has been instrumental in shaping my research.

Finally, for all those who, near or far, have made this work possible.

Dedication

To my parents, for their unwavering belief in me.

To my mentors, for their guidance and wisdom.

To my peers, for their camaraderie and support.

To the pursuit of knowledge, for its own sake.

And to the future, may it be brighter for all Incha'allah.

Abstract

Abstract: This thesis explores the Anaerobic Digestion (AD) process model, focusing on optimizing productivity. The International Water Association's Anaerobic Digestion Model No.1 (ADM1) is a key model in this field. The main challenge is identifying uncertain parameters. To tackle this, a methodology using genetic algorithms (GA) is introduced to fine-tune the parameters of a simplified model, AM2HN. The GA minimizes the proposed objective function on two scales: linear and logarithmic. This methodology, validated through computer simulation, shows significant improvement over traditional techniques. The reduced model AM2 is used to formulate and test a robust control, demonstrating its efficacy and potential. This study contributes significantly to AD process modeling and control.

Keywords: *Anaerobic digestion, AM2HN, Genetic algorithm, Sliding mode control, Parameter Identification, ADM1.*

Table of Contents

List of publications

- **Journal paper:** Abdelhani Chaabna, Samia Semcheddine, Improvement of Biogas Production in Anaerobic Digestion Process. *Ecological Engineering and Environment Protection journal*, No 2, 2021, p. 26-31. Doi: 10.32006/eeep.2021.2.2631.
- **Journal paper:** Abdelhani Chaabna, Samia Semcheddine, Genetic algorithm based identification of biogas production model from wastewater via anaerobic digestion model no.1. *International journal of information technology*. 15, 1465–1472 (2023). Doi: https://doi.org/10.1007/s41870-023-01194-x
- **International Conference paper:** Abdelhani Chaabna, Samia Semcheddine, Improvement of biogas production by SMC: Application on AM2. *Seventh International Conference Ecological Engineering and Environment Protection (EEEP'2021) with Youth Scientific Session and MELiSSA Summer University*, September 30th - October 3rd, 2021, Varna, Bulgaria.
- **International Conference paper:** Abdelhani Chaabna, Samia Semcheddine, Identification of the AM2HN model parameters in the context of organic matter recycling, *19th international Multi-Conference on Systems, Signals and devices (SSD)- Conference on Systems, Automation and Control*, May 06th-10 2022, At Setif, Algeria. Doi: 10.1109/SSD54932.2022.9955922.
- **National Conference paper:**Abdelhani Chaabna, Samia Semcheddine, Oualid Messili, Integral sliding mode control of anaerobic digestion process for improved biogas production, *The 1st National Conference on Emergent Technologies in Electrical Engineering (NCETEE'23)*, December 16-17th 2023, At Setif, Algeria.
- **National Conference paper:** Abdelhani Chaabna, Samia Semcheddine,Genetic Algorithms and Multiple Regression: A Statistical comparison for Parameter Identification of biogas production model via ANOVA test, *The 1st National Conference on Emergent Technologies in Electrical Engineering (NCETEE'23)*, December 16-17th 2023, At Setif, Algeria.

- **Poster Session:** Abdelhani Chaabna, Samia Semcheddine, Robust Control of non linear systems for the optimization of biogaz production, Doctoral Day (Jd'23) - May, 16th, 2023, Ferhat Abbas University Setif1 , Setif, Algeria.

List of Figures

List of Tables

List of abbreviations

α: Biomass retention coefficient.
AD: Anaerobic digestion.
ADM1: Anaerobic Digestion Model N°01.
AM2: Acidogenesis Methanogenesis 2.
AM2HN: Acidogenesis Methanogenesis 2 Hydrolysis Nitrogen.
Al: Aluminum.
ANOVA: Analysis of Variance.
C/N ratio: Carbon to Nitrogen ratio.
C: Total inorganic carbon concentration(mM).
C_{in}: Total inorganic carbon concentration input value.
Ca: Calcium.
D: Dilution rate coefficient(D^{-1}).
GA: Genetic Algorithm.
HRT : hydraulic retention time (day).
ISMC: Integral sliding mode control.
k_1: Yield for substrate concentration.
k_2: Yield for VFA production ($mmolg^{-1}$).
k_3: Yield for VFA consumption ($mmolg^{-1}$).
k_4: Yield for CO_2 production by X_1 ($mmolg^{-1}$.
k_5: Yield for CO_2 production by X_2 ($mmolg^{-1}$).
k_6: Yield for CH_4 production ($mmolg^{-1}$).
k_{La}: Gas–liquid transfer coefficient.
k_{s1}: Half-saturation constant associated with S_1.
k_{s2}: Half-saturation constant associated with S_2.
K: Potassium.
MLR: Multiple linear regression.
Mg: Magnesium.
Na: Sodium.

RMSE: Root Mean Squared Error.

$\mu_1(S_1)$: Specific growth rate of acidogenic bacteria (D^{-1}).

$\mu_{1,max}$: maximum growth rate of acidogenic bacteria.

$\mu_2(S_2)$: Specific growth rate of methanogenic bacteria(D^{-1}).

$\mu_{2,max}$: maximum growth rate of methanogenic bacteria (D^{-1}).

S_1: Organic substrate concentration(gL^{-1}).

S_2: Volatile fatty acids concentration ($mmolL^{-1}$).

$S_{1,in}$: Organic substrate concentration input value.

$S_{2,in}$: Volatile fatty acids concentration input value.

SMC: Sliding Mode Control.

SRB: Sulfate Reduction Bacteria.

USBF:Upflow Sludge Blanket Filtration.

VFA: Volatile Fatty Acids.

LCFA: Long Chain Fatty Acids.

VSCS: Variable Structure Control Systems.

X_1: concentration of acidogenic bacteria ($gVSL^{-1}$).

X_2: concentration of methanogenic bacteria ($gVSL^{-1}$).

General introduction

Renewable energy sources are pivotal in the transition towards a sustainable energy future. Among these, anaerobic digestion stands out as a promising technology. It is a biological process that breaks down organic matter, such as agricultural waste, in the absence of oxygen. This process produces biogas, a mixture of methane and carbon dioxide, which can be used as a renewable energy source for heat and electricity generation. Moreover, the residual material from anaerobic digestion, known as digestate, can be used as a nutrient-rich fertilizer, contributing to a circular economy approach in energy production. Thus, anaerobic digestion not only provides a renewable energy source but also contributes to waste management and agricultural productivity [1].

In 1997, Gosh [2] projected that the recovery of organic waste and industrial effluents could potentially decrease global warming by 20 $\%$. In Algeria, the initial advancements in biodigestion took place around the late 1930s. This progress was driven by two professors from the National School of Agriculture in Algiers (El Harrach), Albert ISMAN and René DUCELLIER. They patented a system in the United States in 1938, which was designed to generate a pulsating flow in a device filled with gas-emitting products [3].

Anaerobic digestion technologies face challenges in terms of conversion efficiency, process stability, product quality, and economic feasibility [4]. The intrinsic complexity of the process makes parameter identification an extremely challenging task [5].

Current research is largely focused on optimizing anaerobic digestion under various digestion conditions. However, identifying the optimal parameters for these conditions can be challenging [6]. In the case of anaerobic co-digestion of different materials, such as chicken manure and corn residues, determining the microbial ecology of these systems and the optimal parameters for the digestion process remains a challenge [7].

Anaerobic digestion involves diverse bacterial communities and substrates. The advancements in scientific tools and computational technology have facilitated the creation of mathematical models that can predict the dynamic behavior of this process. This has been a significant area of focus in recent research. The Anaerobic Digestion Model No. 1 (ADM1) is a notable achievement in this field. However, the purpose of modeling dictates its design, and thus, numerous alternative models have been suggested in scholarly articles. These models vary based on their objectives, such as enhancing process comprehension, conducting dynamic simulations, optimization, or control. Models encompass uncertain constants (or parameters), such as initial conditions, stoichiometry, and kinetic parameters, which necessitate estimation from empirical data.

Parameter identification is a critical step in the modelling process. It involves determining the values of various parameters in the model, such as reaction rates, yield coefficients, and decay constants. These parameters are typically determined through experimental studies and are crucial for the accuracy of the model.

In control systems, parameter identification is equally important. The parameters of the system need to be accurately identified to design effective control strategies. For instance, in a feedback control system, the controller parameters need to be tuned based on the identified system parameters to ensure stability and optimal performance.

The process of parameter identification in AD models is intricate due to the potentially large number of parameters and the limited availability of informative experimental data. In scientific literature, there are instances where the values of parameters in models are provided without any detailed examination of the model's validity or the accuracy of these parameters. This makes it challenging to fully utilize the information that is published [8]. The primary focus of this thesis is the optimization of biogas production through biological way by optimizing parameters of non-linear system. The research employs a genetic algorithm, a heuristic search method that mimics the process of natural selection or survival of the fittest, to identify parameters that guarantee accurate predictions for the given model.

In this thesis, we have structured our work into four main chapters to provide a comprehensive understanding of the topic. The first chapter, "State of Art", provides an extensive overview of Biogas Technology and Anaerobic Digestion (AD). It begins with an introduction and definition, followed by a discussion on the positive impacts of Biogas Technology. The chapter then delves into the degradation pathways during anaerobic digestion and the physico-chemical parameters of the process. It concludes with a detailed look at the modeling of the AD process and the parameter identification of AD models.

The second chapter, "Sliding Mode Control", discusses the theoretical aspects of variable structure control systems (VSCS). It covers the fundamentals of Sliding Mode Control (SMC), including the sliding surface, equivalent control, discontinuous control, and the existence of the sliding mode. The section also addresses the issue of chattering reduction in SMC.

The third chapter, "Parameter Identification: Techniques and Approaches", delves into the topic of parameter identification. It explores the mathematical foundations of parameter identification and discusses various optimization techniques for parameter identification. It provides a detailed study of the Genetic Algorithm and discusses Multiple Linear Regression (MLR) and Analysis of P-value for parameter identification. It concludes with a discussion on the Root Mean Squared Error (RMSE) and the Coefficient of Determination (R squared).

The fourth chapter, "General Discussion", is a comprehensive discussion on the topics covered in the thesis, particularly focusing on parameter identification using Genetic Algorithm and

Sliding Mode Control (SMC).

The last section provides a summary of the entire work (General conclusion), followed by bibliography for further references.

CHAPTER 1

State of art

1.1 Introduction

Anaerobic digestion is a natural biological process of degradation of organic matter in the absence of oxygen. It produces biogas and digestates, which can be used as sources of renewable energy and natural fertilizers, respectively. In this chapter, we will discuss the fundamentals of anaerobic digestion, the biological processes involved, as well as the modeling of anaerobic digestion using the AM2, AM2HN and ADM1 models. We will also discuss the applications of anaerobic digestion in organic waste treatment and the benefits of using simulation models to optimize the anaerobic digestion process.

1.2 Definition

Anaerobic digestion, which is also known as methanization or fermentation, is a natural process in which microorganisms break down organic matter in the absence of oxygen, resulting in the production of biogas (primarily methane), carbon dioxide, and hydrogen. This renewable energy source has numerous benefits, including providing a sustainable solution for organic waste and being available to both energy-poor and energy-rich countries as a supplement to natural gas. The increased production control has made bio gas a popular alternative to and supplement for fossil fuels.

1.3 The Positive impacts of biogas technology

Biogas technology, when implemented and maintained effectively, has the capacity to yield a multitude of advantages at the individual, societal, and environmental levels. These advantages encompass the generation of energy in various forms such as thermal, luminous, and electrical energy, in addition to the transformation of organic waste into high-grade fertilizer [1].

The technology also contributes to improved sanitation by reducing the prevalence of pathogens, helminth eggs, and flies. Furthermore, it aids in the preservation of soil, water, air, and arboreal vegetation, thus conferring environmental benefits.

From an economic perspective, biogas technology provides both micro- and macro-economic advantages by replacing conventional energy and fertilizer sources, fostering decentralized energy production, and decreasing import reliance, all the while protecting the environment. In essence, biogas technology serves as a pivotal instrument in advancing conservation and sustainable development efforts.

1.4 Degradation pathways during anaerobic digestion

The anaerobic digestion process consists of a complex chain reaction that can be divided into four main stages: hydrolysis, acidogenesis, acetogenesis, and methanogenesis as exhibited in Figure 1.1.

Disintegration and hydrolysis are biological processes that occur outside of the cell. These processes are facilitated by enzymatic action and serve to break down complex molecules into their constituent monomers, which can then be utilized by the biomass. Disintegration is responsible for converting composite particulate substrates into inert particles, particulate carbohydrates, proteins, and lipids. Hydrolysis, on the other hand, converts particulate carbohydrates, proteins, and lipids into simple sugars (MS), amino acids (AA), and long-chain fatty acids (LCFA). Both disintegration and hydrolysis are modeled using first-order kinetics. The hydrolysis reaction can be written as follows:

$$C_6H_{10}O_4 + 2H_2O \Rightarrow C_6H_{12}O_6 + H_2$$

The process of acidogenesis or fermentation is characterized by the biological production of acids in the absence of external electron acceptors or donors under anaerobic conditions. This process involves the degradation of soluble sugars and amino acids by two distinct biomass components. It is generally fast due to the high growth rate of the bacteria involved. The metabolic reactions of these products are as follows:

$$C_6H_{12}O_6 + 2H_2O \Rightarrow 2CH_3COOH + 4H_2 + 2CO_2 \tag{1.1}$$

$$C_6H_{12}O_6 + 2H_2O \Rightarrow C_4H_7O_2^- + 2HCO_3^- + 3H^+ + 2H_2 \tag{1.2}$$

$$C_6H_{12}O_6 \Rightarrow 2C_3H_5O_3^-2H^+ \tag{1.3}$$

$$C_6H_{12}O_6 + HCO_3^- \Rightarrow CH_3COO^- + C_4H_4O_4^{2-} + CHO_2^- + 3H^+ + 2H_2O \tag{1.4}$$

$$C_6H_{12}O_6 + 2H_2O \Rightarrow 2C_2H_5OH^- + 2HCO_3^- + 2H^+ \tag{1.5}$$

The third step of digestion is called acetogenesis. During acetogenesis, volatile (VFA) and long-chain (LCFA) fatty acids are degraded by acetogenic bacteria through oxidation to produce acetate and hydrogen. However, it is important to maintain low levels of hydrogen production in order to ensure thermodynamic feasibility. The reactions involving these bacteria are as follows:

$$C_2H_5OH + 2H_2O \Rightarrow 2CH_3CH_3COO^- + 3H_2 + H^+ \tag{1.6}$$

$$2HCO_3^- + 4H_2 + H^+ \Rightarrow CH_3COO^- + 4H_2O \tag{1.7}$$

In the last step of digestion, the methanogenic biomass utilizes the hydrogen that is produced during this stage. The process allows for the transformation of hydrogen, carbon dioxide, and acetate into methane. The bacteria involved in these reactions are classified as strict anaerobes. They are divided into two bacterial populations: acetoclastic methanogens, which use acetates to produce methane, and hydrogenotrophic methanogens, which reduce carbon dioxide or formic acid with hydrogen. This biological process is a key part of the global carbon cycle and has significant implications for energy production and environmental sustainability. The following equations show the production of biogas:

$$CH_3COO^- + H_2O \Rightarrow CH_4 + HCO_3^- \tag{1.8}$$

$$CO_2^- + 4H_2 + H^+ \Rightarrow CH_4 + 2H_2O \tag{1.9}$$

$$CHOO^- + 3H_2 + H^+ \Rightarrow CH_4 + 2H_2O \tag{1.10}$$

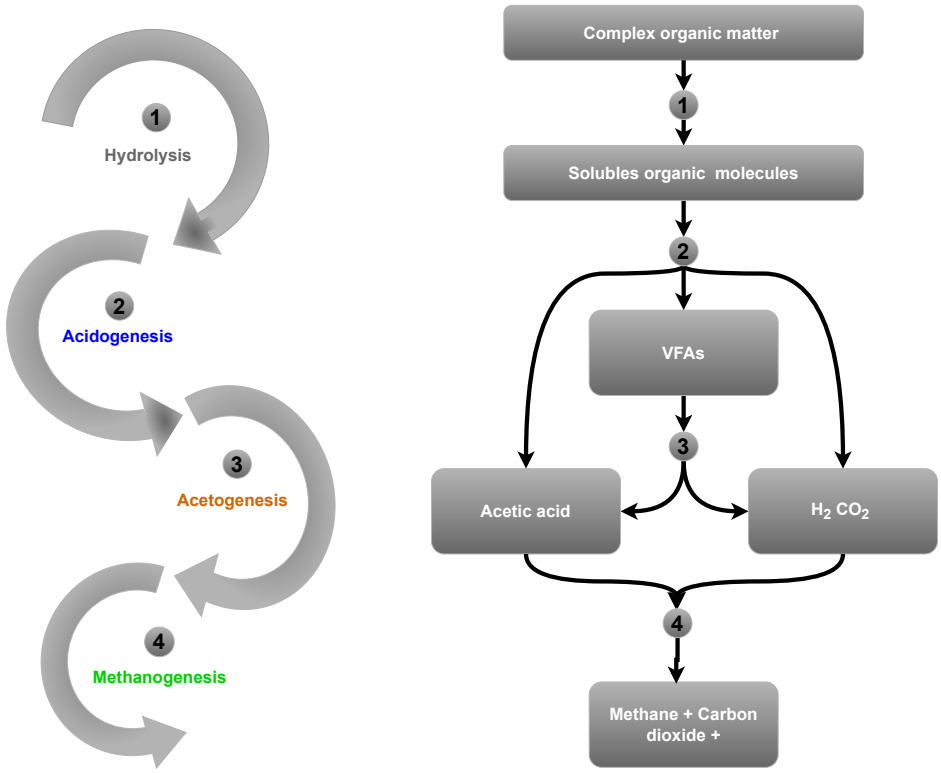


Figure 1.1: Degradation pathways in AD process

1.5 Physico-chemical parameters of the AD process

The AD process could be influenced by several factors. These include reactor operating conditions like temperature, pH, Organic Loading Rate (OLR), Hydraulic Retention Time (HRT), Sludge Retention Time (SRT), and upflow velocity, as well as the size distribution of particles in the influent. The following subsections will discuss these factors and their effects in more detail 1.2.

1.5.1 The temperature

The methanogenic bacteria responsible for anaerobic digestion are specific to the temperature of the biodigester. Organisms are divided into three categories based on their temperature range:
- The psychrophilic bacteria dominate between 4-20 ° C, with a growth optimum around 15° C.
- The mesophilic bacteria dominate between 20-45 ° C, with a growth optimum around 37 ° C.
- The thermophilic bacteria dominate between 55-70 ° C, with a growth optimum around 60 ° C.

1.5.2 Potential hydrogen (pH)

In AD, pH is a crucial process parameter that influences the dominant micro-organism and product yield. The pH regime in an operational reactor is determined by factors such as volatile fatty acid concentration, buffering capacity of influent alkalinity, ammonia concentration of

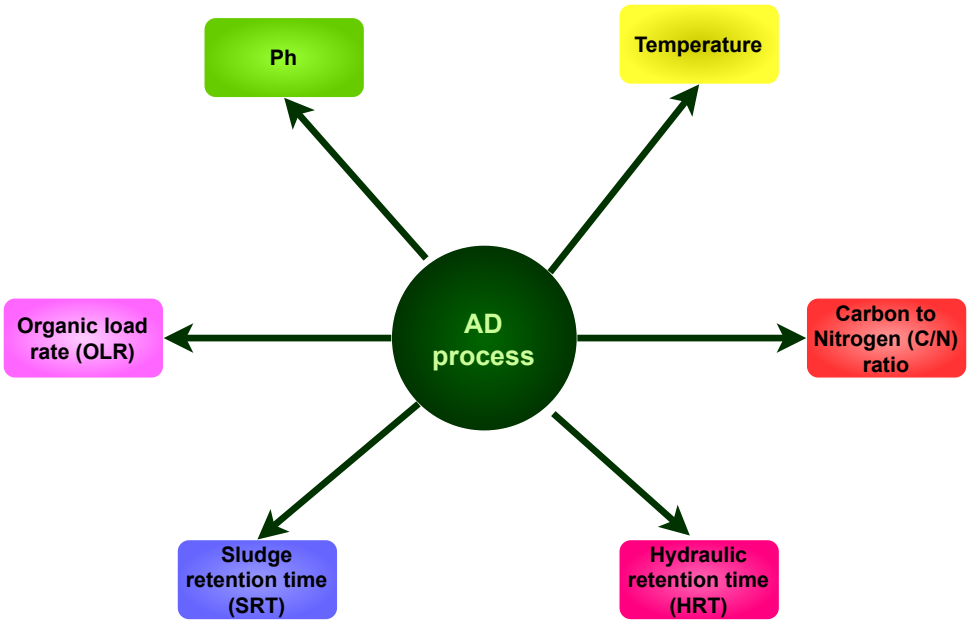


Figure 1.2: Main parameters that affect the AD process

ingestate, predominant bacteria, and the feedstock-to-inoculum ratio. These factors reflect the overall state of the reactor [9].

1.5.3 C/N ratio

The C/N ratio is an important parameter for high solids AD and refers to the amount of carbon and nitrogen in the feedstock. Both elements are essential for microbial growth and function. Nitrogen helps with the synthesis of amino acids, proteins, and nucleic acids, while carbon provides structure and energy for microbes. Some of the organic nitrogen in the feedstock is converted into ammonia, which helps neutralize volatile acids produced by fermentative bacteria and maintain a neutral pH.

1.5.4 Hydraulic Retention Time (HRT)

During the wastewater treatment process, a liquid or a soluble compound remains in a tank or a reactor for a certain period of time. This period is called the Hydraulic Retention Time (HRT) and it is measured by dividing the tank or reactor volume by the influent flow rate:

$$HRT = \frac{V}{Q_{in}} \tag{1.11}$$

where V is the volume in cubic meters and Q_{in} is the flow rate in cubic meters per hour.

1.5.5 Organic Load Rate (OLR)

The amount of substrate entering the digester over time is measured by the organic loading rate (OLR) in units of (g/L d). Given a specific substrate concentration and Hydraulic Retention Time (HRT) for the digester [10], one can use this formula to calculate the OLR:

$$OLR = \frac{S_0}{HRT} \tag{1.12}$$

where S_0 is the influent substrate concentration, $HRT = hydraulicretentiontime(d)$. High OLR can result in decrease the reactor efficiency, while an excessive OLR may precipitate VFA accumulation and consequent reactor overload.

1.5.6 Sludge Retention Time (SRT)

The SRT refers to the mean residence time of microorganisms in bioreactor tank. A crucial aspect in designing and operating anaerobic systems for treating mostly soluble industrial wastewaters is ensuring that the essential anaerobic bacteria are washed out of the system faster than they can reproduce. SRT is determined by dividing the total mass of bacteria present in the digester by the total mass of bacteria that is lost from the system per unit time. SRT can be calculated using the following formula:

$$SRT = \frac{Total\ \ mass\ \ of\ \ bacteria\ \ in\ \ the\ \ digester(g)}{Total\ \ mass\ \ of\ \ lost\ \ bacteria(g/day)} \tag{1.13}$$

1.6 Inhibitions of the AD

In AD, the degradation of organic matter is based on a sensitive balance between the different groups forming the trophic chain. It is between the acidogenic and methanogenic biomasses that this balance is the most precarious [11]. The inhibitor will have an impact on the bacterial growth of the species or group of microorganisms sensitive to the inhibitor, which may reduce the maximum production of methane, cause accumulations of reaction intermediates [12] and potentially impact the transformations of nitrogen. The inhibition thresholds are very difficult to set for certain substances because processes such as the adaptation of the biomass to the inhibitor or the use of another way of transformation of the trophic chain can come into play. The consequence will be better tolerance to the inhibitor and a return to a state of equilibrium. The pH, the temperature, the interactions between the different inhibitors or even the physical structure of the biomass will have a positive or negative influence on the effect of the inhibition. Certain substances can significantly reduce performance and influence stability of an anaerobic digester. In the literature, inhibition substances are defined as toxicants of the microbial communities which are responsible of the methane production,

1.6.1 Ammonia

The biological degradation of the nitrogenous matter produces Ammonia, mainly represented in proteins and urea, In aqueous solutions, inorganic ammonia nitrogen exists mainly as Ammonium ion (NH_4^+) and free ammonia (NH_3). It has been proposed that free ammonia is the primary cause of inhibition, as it can easily pass through the cell membrane due to its membrane-permeable nature (as suggested by [13]). This hydrophobic ammonia molecule can passively diffuse into the cell, leading to an imbalance in proton concentration and/or a deficiency of potassium.

1.6.2 Light metals (Na, K, Mg, Ca, and Al)

Light metal cations such as sodium, potassium, calcium, and magnesium are present in the substrate of anaerobic digesters. These cations can be derived from the organic matter (e.g., biomass) that undergoes anaerobic degradation, or from the alkaline agents that are applied to control the pH value. These cations are essential for microbial metabolism and, therefore, influence the specific growth rate as any other nutrient. They have a stimulatory effect on microbial growth at moderate concentrations, but a detrimental effect at excessive concentrations. At very high concentrations, they can cause severe inhibition or toxicity to the microorganisms.

1.6.3 Heavy metals

The heavy metals that pose a particular threat include chromium, iron, cobalt, copper, zinc, cadmium, and nickel. Heavy metals have a unique characteristic that sets them apart from many other toxic substances: they are not biodegradable and can build up to potentially toxic levels. A comprehensive study of anaerobic digester performance revealed that heavy metal toxicity is one of the main reasons for digester disturbance or failure [14].

1.6.4 Sulfide

Sulfate is a ubiquitous constituent of many industrial wastewaters.In anaerobic digesters, sulfate is converted to sulfide by mean of the Sulfate Reduction Bacteria (SRB). Sulfate reduction is carried out by two main types of SRB: incomplete oxidizers, which deoxidize compounds like lactate to acetate and CO_2, and complete oxidizers, which fully convert acetate to CO_2 and HCO_3. Sulfate reduction affects methane production in two manners. The first way is that SRB use up some of the same substances that methane producers need. This reduces the amount of methane that can be made.

1.6.5 Organics

Anaerobic processes can be affected by many kinds of organic compounds. Some of these compounds do not dissolve well in water or stick to the sludge solids. They can build up to high

levels in anaerobic digesters. When these nonpolar pollutants get into the bacterial membranes, they make the membrane expand and leak. This messes up the ion balance and can cause the cell lysis.

1.7 Modes of the AD

Anaerobic digester can operate in three different modes, namely continuous mode, batch mode or fed-batch mode

1.7.1 Continuous mode

This type of operation implies a permanent feeding of the bioreactor (Figure 1.3c), also called nominal mode, generally associated with a constant flow rate. The volume of effluent or solid waste in the reactor remains constant, because an amount equivalent to that entering each day must be evacuated.

1.7.2 Batch mode (discontinuous)

The bioreactor operates without any interaction with the external environment (Figure 1.3a). Indeed, once filled, the reactor is no longer supplied with charge. At the end of the digestion, the bioreactor is emptied to allow a new fermentation.

1.7.3 Fedbatch mode (sequential batch)

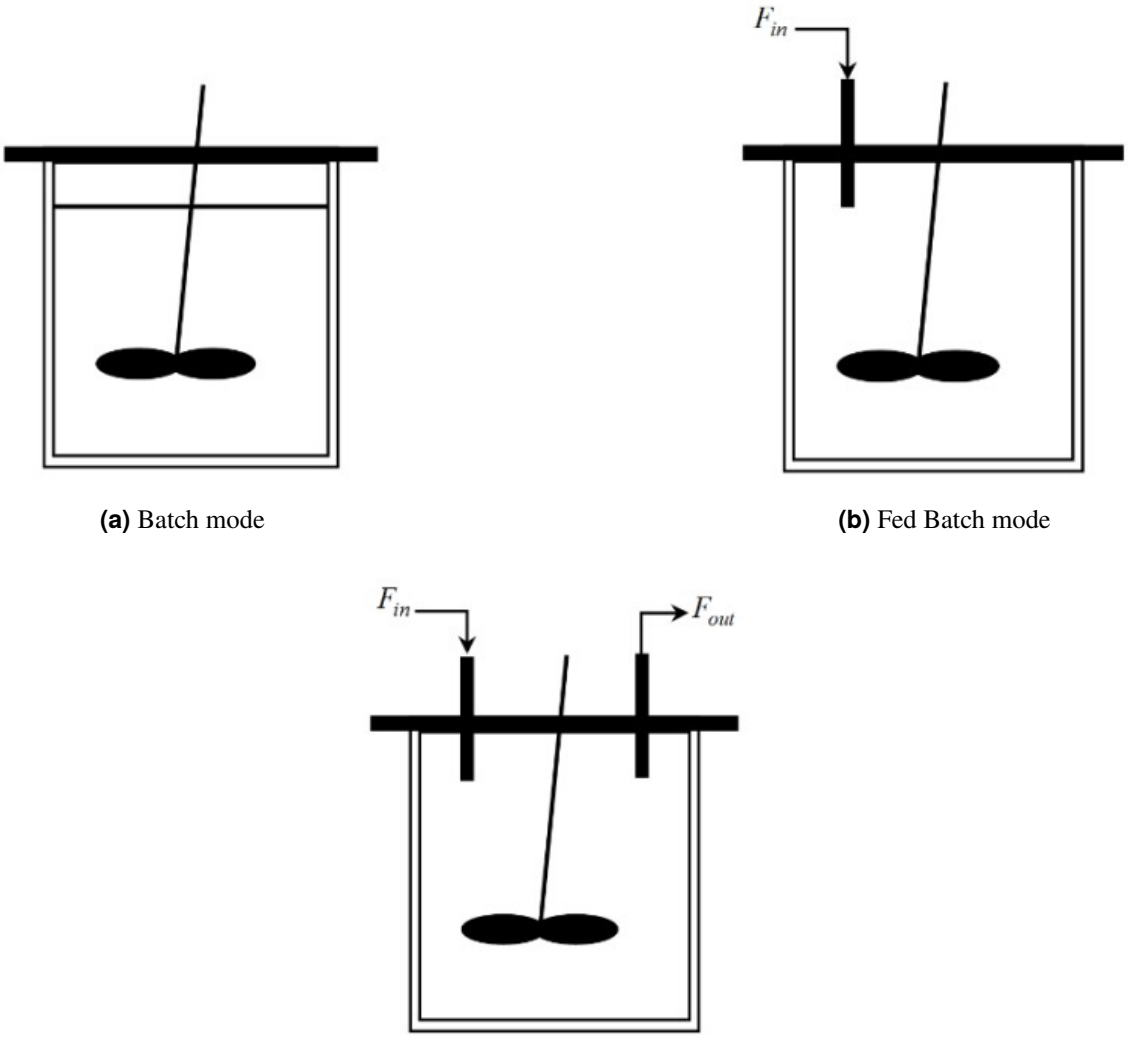
It is the combination of the continuous and batch modes. The reactor operates with a cycle alternating filling, reaction, settling, and emptying (Figure 1.3b). The reactor is not completely emptied, this is to keep some of the microorganisms to form a new cycle. Generally, anaerobic digestions on a small scale operate in sequential batch mode.

1.8 Modeling of the AD process

To optimize biogas production in anaerobic digestion processes, process modeling is essential. It leads to a better understanding of the complex interactions between the microorganisms involved and environmental parameters such as temperature, pH, and nutrients. Mathematical models also help predict system performance and identify means to improve process efficiency.

1.8.1 The ADM1 model

The ADM1 (Anaerobic Digestion Model No.1) is a complex mathematical model based on the Petersen matrix description [15, 16]. It was developed within the framework of the



(a) Batch mode

(b) Fed Batch mode

(c) Continuous mode

Figure 1.3: Operational modes of AD bioreactor

International Water Association (IWA) task group for mathematical modeling in 2002. There are 32 state variables in the ADM1. It describes accurately the 19 biochemical processes, 3 gas-liquid transfer kinetic processes and 7 bacterial populations that are involved in anaerobic digestion (1.4). Likewise, ADM1 is undoubtedly one of the leading models in simulating the AD process. It represents a common basis for many engineering and research applications. The first-order kinetics characterize the extracellular steps, whereas the intracellular steps involve growth, consumption, and death processes. The kinetics of biomass death follow first-order kinetics. The substrate consumption is modeled using a Monod equation that is correlated to the yield of biomass for the specific substrate and its growth.

The inert fractions (I index), which are brought in by the influent or are a result of bacterial lysis and hydrolysis of complex molecules, are incapable of degradation and thus do not partake in biomass proliferation. Soluble inert compounds (SI) act as tracers within the wastewater treatment plant, while particulate inert compounds (XI) get trapped in the sludge. The degradable fractions

consist of substrates that the bacteria use for growth. These fractions are further divided into two categories, i.e., soluble and particulate. Soluble compounds are the ones that penetrate the biomass cell wall and are immediately utilized. They are designated by S. Particulate compounds, on the other hand, represent biomass and compounds that are unable to pass through the bacterial cell wall and require a limiting hydrolysis step. They are designated by X.
The model considers three different groups of fatty acids: propionate, butyrate, and valerate, as well as long-chain fatty acids. The model considers two different pathways of methanogenesis: one from hydrogen and the other from acetate. Additionally, two different biomass components are considered.

In the literature, three main iterations of this model have been identified: the standard description [15], the ADM1 COST [17], and the ADM1 CEIT [18]. the principle change made in the COST ADM1 are The incorporation of sludge treatment into the IWA Benchmark Simulation Model (BSM) to develop a plant-wide or "within-the-fence" model. The CEIT ADM1 is distinguished by the incorporation of some model components and the estimation of the chemical Oxygen Demand as a function of the redox equations associated with these components via elemental mass fractions [18] .

A. Extenions of ADM1

Since the release of ADM1, there has been an increasing demand to add other necessary processes to the ADM1 framework. The Task Group has recognized this and divided these demands into three main categories: phosphorus metabolism, sulfur metabolism, and precipitation/dissolution of minerals. There have also been efforts to incorporate other important biochemical and physicochemical processes into the ADM1 framework. The table 1.1 depicts a list of studies whose incorporate ADM1 extensions have been summarized and reviewed.

B. Applications of the ADM1

In the literature, the ADM1 model has been widely applied and has proven its efficiency in several works. These can be categorized into the following areas:

a. Prediction of effluent characteristics in AD bioreactors

Y. Huang et al. [30] conducted a lab-scale experiment to simulate the effluent performance of an anaerobic reactor using the Increasing-Size Continuous Stirred-Tank Reactors (ISC) model. They also integrated the ISC model with the Anaerobic Digestion Model No. 1 (ADM1) to simulate the effluent COD (COD_{eff}) under steady and over-loading conditions. The authors were the pioneers in taking into account the hydraulic dynamic of the reactor when implementing ADM1. M.E Ersahin [31] used the Anaerobic Digestion Model No. 1 (ADM1) to simulate the dynamic behavior of a full-scale primary sludge digester in a municipal wastewater treatment

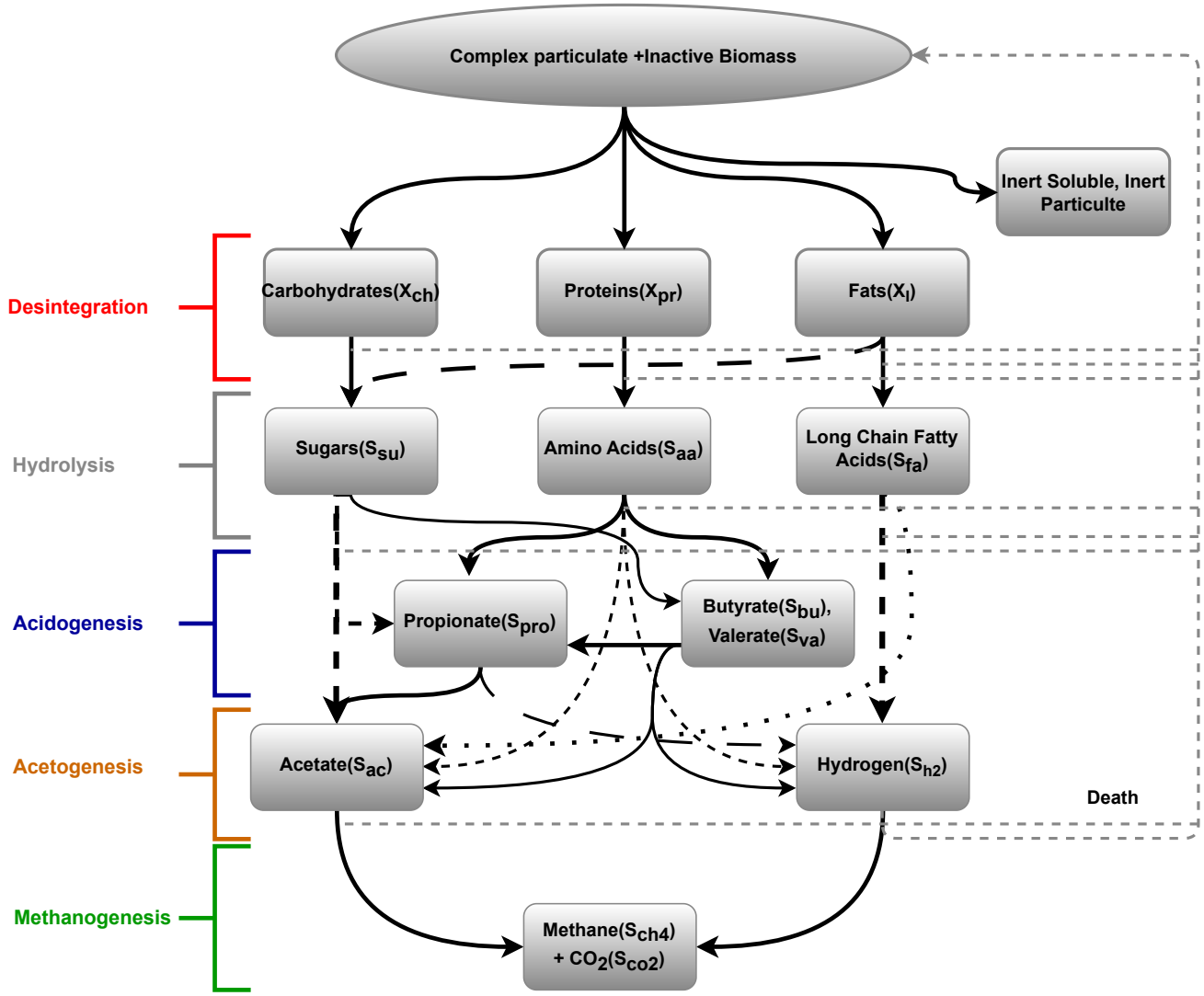


Figure 1.4: The anaerobic digestion process described by the ADM1 model

Table 1.1: Extensions of the ADM1

Reference	Goal	Findings
Bastone et al. (2003a) [19]	Consideration of CaCO3 precipitation	Model avoided process modifications that will be expensive, time consuming and unlikely to get succeeded.
Batstone et al. (2003b)[20]	Investigation of the degradation kinetics of organic acids in AD	Assesement of I valerate stoichiometry allows higher energy yields
Fedorovich et al. (2003) [21]	Removal efficiency representation of the of acetate, butyrate, propionate, and sulfate	a discrepancy between the model and experiment regarding methane production is caused by The inhibition effect of H2S
Zaher et al. (2009)[22]	Development of a general transformer model to interface ADM1 to different solid waste streams (co-digestion)	it is important to consider separate hydrolysis rates for each particulate component from each waste stream
I. Ramirez et al. (2009) [23]	Integration of the hydrolysis and disintegration steps of thermophilic anaerobic digestion of thermally pretreated Waste activated Sudge into ADM1	Accurate prediction of VFAs, Ph and methane from four different batch thermophilic AD of untreated and thermally pretreated sludges.
Palatsi et al. (2010) [24]	modelling the adsorption effect of LCFA on the cell walls	important insights into the reversible inhibition of anaerobic digestion by long-chain fatty acids (LCFA) and the recovery capacity of $\beta-$oxidizing bacteria and syntrophic methanogens have been provided
Mairet et al. (2011) [25]	the addition of a Contois model for hydrolysis and Ph inhibition to represent the AD of microalgae	Providing Insights into the potential of microalgae as a sustainable source of biodiesel and renewable energy.
Ernesto et al. (2015) [26]	Extending ADM1 by Development of a model structure for the sulfate reduction process in the AD of cane-molasses vinasse	increasing the applicability of ADM1 to specific industrial wastewaters (vinasse)
Xavier F.A et al. [27] (2016)	Inclusion of a model for water resource recovery facilities (WRRFs) that can simulate multiple mineral precipitation and phosphorus recovery	model can be easily linked with standard activated sludge models
Hangyu sun et al. (2021a) [28]	development of the gas–liquid mass transfer, the inhibition equations and biochemical reaction of H2 and CO	syngas conversion efficiency into methane was increased
Hangyu sun et al. (2021b) [29]	Inclusion of formate utilizing methanogenesis	Formate-utilizing and hydrogen-utilizing methanogenesis had stronger ammonia resistance than acetate-utilizing methanogenesis

plant. The author reported that the model could predict the methane flow and effluent COD concentration in good agreement with the measured data. R.K Dereli et al. [32] performed a ADM1 based influent characterization. They calibrated the ADM1 for the treatment of a specific industrial wastewater and validated it with experimental data. They reported that the calibrated ADM1 could predict the effluent COD, pH, methane and biogas flows accurately. They demonstrated a useful approach for the implementation of the ADM1 for this type of wastewater.

b. Substrate characterization

A realistic description of the input substrate is crucial, even for a sophisticated model like ADM1 that differentiates between various organic fractions such as carbohydrates, proteins, and lipids. Zaher et al. [33] presented one of the first applications of ADM1 in their work. By using experimental measurements, the model, and data on the composition of wastewater, they were able to expand knowledge of the characteristics of the influent and describe the reactor's behavior. D. Poggio et al. [34] introduced and evaluated a detailed methodology for substrate characterization to be used with ADM1. This methodology combines biochemical and kinetic fractionation approaches. They found that by modifying ADM1 to include different particulate fractions with varying degradation kinetics, the prediction of methane production from complex substrates such as green waste and food waste can be enhanced.

c. Control of waste water treatment plants

H. Zhou and colleagues [35] suggest a new multi-objective control strategy for a USBF reactor that treats winery wastewater. It uses a cascade control strategy and takes advantage of the differences between the liquid and gas phases. The ADM1 model was used to validate its performance. K.V Bharat and colleagues [36] modified the ADM1 using a variable stoichiometry approach, which accurately predicted experimental data from a bio-hydrogen reactor fed with sucrose. E Nordlander and colleagues [37] used a model based on the ADM1 to investigate a full-scale digester co-digesting various wastes. The model was best at predicting biogas flow, methane content/flow, and ammonia nitrogen. Results showed that increasing influent VS concentration increased biogas and methane outflow but decreased biogas/methane per unit of volatile solids.

1.8.2 The AM2 model

The AM2 model was developed as part of the AMOCO (Advanced Monitoring and Control System for anaerobic processes, European FAIR project No. ERB-FAIR-CT96-1198) project that aimed to develop advanced control strategies for industrial bioprocesses, including fermentation processes [38]. With it's six state variables, the model is considered as one of the most suitable models for control applications [39]. It is a two-step model (corresponding to two biological reactions in cascade, hence its name) describing in particular the processes of acidification and methanization.

The acidogenic biomass X_1 consumes organic substrate S_1 and produces VFAs (S_2) and CO_2. In the second step of methanization, the methanogenic biomass X_2 consumes VFAs and produces methane and carbon dioxide. The biological reactions are described in this model are as follows: Acidogenesis (with a reaction rate μ_1)

$$k_1 S_1 \Rightarrow X_1 + k_2 S_2 + k_4 CO_2$$

Methanogenesis (with a reaction rate μ_1)

$$k_3 S_2 \Rightarrow X_2 + k_5 CO_2 + k_6 CH_4$$

$$\begin{cases} \frac{dX_1}{dt} = (\mu_1(S_1) - \alpha D)X_1 & \text{(1.14a)} \\ \frac{dX_2}{dt} = (\mu_2(S_2) - \alpha D)X_2 & \text{(1.14b)} \\ \frac{dS_1}{dt} = D(S_{1,in} - S_1) - k_1\mu_1(S_1)X_1 & \text{(1.14c)} \\ \frac{dS_2}{dt} = D(S_{2,in} - S_2) + k_2\mu_1(S_1)X_1 - k_3\mu_2(S_2)X_2 & \text{(1.14d)} \\ \frac{dC}{dt} = D(C_{in} - C) - q_c + k_4\mu_1(S_1)X_1 + k_5\mu_2(S_2)X_2 & \text{(1.14e)} \\ \frac{dZ}{dt} = D(Z_{in} - Z) & \text{(1.14f)} \end{cases}$$

where k_i are stoichiometric coefficients, homogeneous with yield coefficients. The bacterial growth rate $\mu_1(j^{-1})$ of acidogenic biomass is of Monod type:

$$\mu_1(S_1) = \mu_{1,max} \times \frac{S_1}{S_1 + K_{s_1}} \tag{1.15}$$

Whereas the methanogenic bacteria is of the Haldane type:

$$\mu_2(S_2) = \mu_{2,max} \times \frac{S_2}{S_2 + K_{s_2} + \frac{S_2^2}{K_I}} \tag{1.16}$$

1.8.3 The AM2HN model

The AM2HN devlopped by Hassam et al. as part of the European project COADVISE [40]. It is an improved version of the AM2. The improvement was initiated by the inclusion of two relevant processes: hydrolysis and the concomitant release of ammoniacal nitrogen. The following differential equations (ode) describes the mass balance for the six state variables.

$$
\begin{cases}
\frac{dX_1}{dt} = (\mu_1(S_1) - \alpha D)X_1 & \text{(1.17a)} \\
\frac{dX_2}{dt} = (\mu_2(S_2) - \alpha D)X_2 & \text{(1.17b)} \\
\frac{dS_1}{dt} = D(S_{1,in} - S_1) - k_1\mu_1(S_1)X_1 + k_{hyd}X_T & \text{(1.17c)} \\
\frac{dS_2}{dt} = D(S_{2,in} - S_2) + k_2\mu_1(S_1)X_1 - k_3\mu_2(S_2)X_2 & \text{(1.17d)} \\
\frac{dC}{dt} = D(C_{in} - C) - q_c + k_4\mu_1(S_1)X_1 + k_5\mu_2(S_2)X_2 & \text{(1.17e)} \\
\frac{dX_T}{dt} = D(X_{T,in} - X_T) - k_{hyd}X_T & \text{(1.17f)} \\
\frac{dZ}{dt} = D(Z_{in} - Z) + (k_1 \times N_{S1} - N_{bac}) \times \mu_1(S_1) \times X_1 - N_{bac} \times \mu_2(S_2) \times X_2 & \\
\quad + k_{d,1} \times N_{bac} \times X_1 + k_{d,2} \times N_{bac} \times X_2 & \text{(1.17g)}
\end{cases}
$$

Where X_T The particulate substrate concentration ($gVSm^{-3}$).
k_{hyd} Maximum specific hydrolysis rate.

1.9 Challenges in AD and potential solutions

While Anaerobic Digestion (AD) holds significant promise due to its ability to convert organic materials into biogas, it is not without its challenges. These include:

1.9.1 Conversion Efficiency:

The process of converting organic matter into biogas is not always efficient. This means that a significant portion of the potential energy in the feedstock is not captured, reducing the overall effectiveness of the AD process.

1.9.2 Process Stability:

The AD process involves a complex series of biochemical reactions. Maintaining stability throughout these reactions can be difficult, leading to inconsistent results and potential process failures.

1.9.3 Product Quality:

The quality of the biogas produced by AD can vary significantly. This can be due to a variety of factors, including the composition of the feedstock and the specific conditions under which the AD process is conducted.

1.9.4 Economic Feasibility:

Despite the potential benefits of AD, it can be expensive to implement and maintain. This can make it economically unfeasible in certain situations.
In addition to these challenges, specific issues such as foaming and over-acidification can occur during the AD process. Foaming can cause operational issues, while over-acidification can lead to a decrease in the pH of the system, inhibiting the activity of the methanogenic bacteria and potentially leading to process failure.

Furthermore, the feedstock used in the AD process can also pose challenges. If the feedstock has a high solid content, it can be difficult to maintain process stability and performance. This is because solids can accumulate in the system, leading to blockages and other operational issues. Therefore, careful management of the feedstock is crucial for the successful operation of an AD system.

1.10 Parameter identification of AD models

Since AD models have limitations in their universality, some parameters must be identified for each individual case study. Traditionally, these models were calibrated through a process of trial and error, which can be quite time-consuming and does not offer any insight into the uncertainty of the parameter values, nor does it ensure their uniqueness.
Dynamic bioprocess models are complex because they reflect the complexity of the processes and the detailed knowledge acquired over time. These models have many parameters that need to be assigned numerical values based on prior knowledge or experimental data. The accuracy of parameter estimation depends on the quantity and quality of available data for calibration. However, another challenge is that bioprocess models are nonlinear and their parameters can be strongly correlated. Therefore, studying the identifiability of model parameters before estimation is crucial.

1.10.1 Why we identify model parameters?

The identification of model parameters allows us to make better predictions, design optimal controllers and perform accurate simulations. Model parameters could help in understanding the

relationships between different variables in the model are identified, They provide insight into how changes in one variable might affect another. Once we identify the parameters of a model, we can use them to make predictions about future outcomes. This is particularly useful in fields like economics, finance, and machine learning. In many cases, we aim to find the best possible values for the parameters of a model to optimize some objective. This could be minimizing error in predictions, maximizing production, or any number of other objectives. The parameters of a model can also be used to evaluate the model's performance. One can compare different set of parameters to determine which set best fits the observed data.

1.10.2 Structural identifyability

In the case of structural identifiability, one will check if all the parameters of the model can be identified in an ideal academic scenario with noise-free and continuous measurements. This may result in finding that only a combination of parameters is identifiable. If the number of combinations is less than the number of model parameters or if there isn't a one-to-one relationship between them, then prior knowledge of certain parameters is necessary for one to solve the identifiability problem [38]. In other words, this refers to the theoretical ability to estimate unknown parameters of a model uniquely from perfect output data. Structural identifiability is well understood for linear systems and there are many tests available, such as those mentioned in [6]. However, Testing the identifiability of nonlinear models is more complicated, with only a limited number of methods available so far. These methods often only highlight the necessary and/or sufficient conditions of local identifiability, which is valid only for certain domains of the model, as mentioned in [41].

1.10.3 Practical identifiability

Practical identifiability is concerned with the impact of available experimental data on the identifiability of a model's parameters, as opposed to ideal measurements in structural identifiability. For example, Monod's model for simple microbial growth is structurally identifiable [42], but often non-identifiable in practice due to low-quality and limited experimental data [43]. This means that variations in one parameter can be compensated by proportional variations in another, while still maintaining a good fit between experimental data and numerical predictions. Hence, The analysis of practical identifiability can be carried out by examining how well an existing model reproduces experimental data Additionally, algorithms for estimating nonlinear model parameters may have poor convergence properties or be numerically ill-conditioned, resulting in estimated values that are sensitive to estimation conditions and difficult to interpret physically [44].

1.11 Conclusion

In short, anaerobic digestion is a proven and effective technique to transform organic waste into biogas and fertilizer. It represents an increasingly popular solution for sustainable management of organic waste while respecting the environment. Despite its many benefits, some challenges remain, such as the need for more efficient and cost-effective digester designs, improved pre-treatment techniques for certain waste streams, and increased public awareness of the use of anaerobic digestion.

Research and development continues to make significant advances in anaerobic digestion, including technologies such as high efficiency digesters, membranes and microbial communities. In addition, the growing demand for renewable energy and the need to reduce greenhouse gas emissions are encouraging investment in anaerobic digestion and driving innovation.

As the world faces major challenges such as climate change and resource depletion, anaerobic digestion is set to play a key role in sustainable waste management, energy production and agriculture. With continued research and development, anaerobic digestion has the potential to become even more efficient and cost-effective, making it an indispensable tool for a more sustainable future.

CHAPTER 2

Sliding Mode Control

2.1 Introduction

This chapter delves into the intricate world of Variable Structure Control Systems (VSCS), with a particular focus on Sliding Mode Control (SMC). It begins by laying out the theoretical aspects of VSCS, setting the stage for a deeper exploration of the subject matter.
The chapter then presents the problem statement, providing a clear understanding of the challenges and objectives associated with VSCS. This sets the foundation for the subsequent sections, which delve into the fundamentals of SMC. Here, key concepts such as the sliding surface, equivalent control, discontinuous control, and the existence of the sliding mode are discussed in detail.
The chapter further addresses the issue of chattering in SMC, a common phenomenon that can affect the performance of control systems. It presents boundary layer solutions as a method to reduce chattering, providing a comprehensive understanding of this approach.
The chapter then transitions into the topic of Integral Sliding Mode Control, a variant of SMC that offers certain advantages. The characteristics of this control method are discussed, providing insights into its functionality and benefits.
Finally, the chapter concludes with a summary of the key points discussed, offering a holistic view of the subject matter. This chapter serves as a comprehensive guide to understanding the complexities and nuances of Variable Structure Control Systems and Sliding Mode Control.

2.2 Control Strategies in Anaerobic Digestion: literature review

The main purpose of controlling the AD process is to juggle two contrasting goals: boosting the AD process's efficiency and maintain it stable. In simpler terms, we aim to track the AD system to consistently manage a high loading rate without compromising its efficiency. However, this situation necessitates an appropriate control strategy, as raising the organic loading rate frequently causes instabilities to the system. The complexity of control algorithms is directly linked to the number of objectives that need to be achieved. Consequently, a wide array of process control strategies have been formulated. These range from basic proportional integral derivative (PID) control to more complex strategies like model-adaptive control, fuzzy logic, artificial neural network schemes, and even combinations of these. In simpler terms, a more advanced control algorithm could make up for the information gathered from observing the AD process.

2.2.1 Feedback control

In feedback control, the controlled variable is assessed and juxtaposed with a set value (Figure 2.1). The variance, termed as error, is minimized by adjusting system inputs. PID control is a commonly utilized feedback control scheme for automated process control [45]. PID control, encompassing proportional, integral, and derivative elements, is predicated on current, historical, and predicted errors, respectively. The proportional element is derived by scaling the error with a constant, known as the proportional gain. This, however, engenders a persistent deviation between the set-point and output, even in a steady state, termed as the steady state error. To mitigate this deviation, an integral component is integrated into the algorithm. The integral control is derived by scaling the cumulative error over time with a constant, termed as the integral gain. In contrast, the derivative control hinges on the rate of alteration of the error, also known as the time derivative of the error. As a result, the outputs from the integral and derivative controls aid the system in rapidly reaching the desired set point with minimal fluctuation. In simpler terms, the PID control algorithm has been effectively used in AD systems, particularly for controlling non-biological parameters like feeding rate [35], pH [46] and temperature [47]. Sliding Mode Control (SMC) is a control method that primarily operates on feedback principles as well. The unique aspect of SMC is that it doesn't use a constant control law that operates continuously over time. There are many works that apply Sliding Mode Control (SMC) in the field of AD [39, 48]

2.2.2 Feedforward control

The fundamental principle of feedforward control that was implemented in the three-element level control loop of the boiler drums (Figure 2.2). Its origins can be dated back to as early as

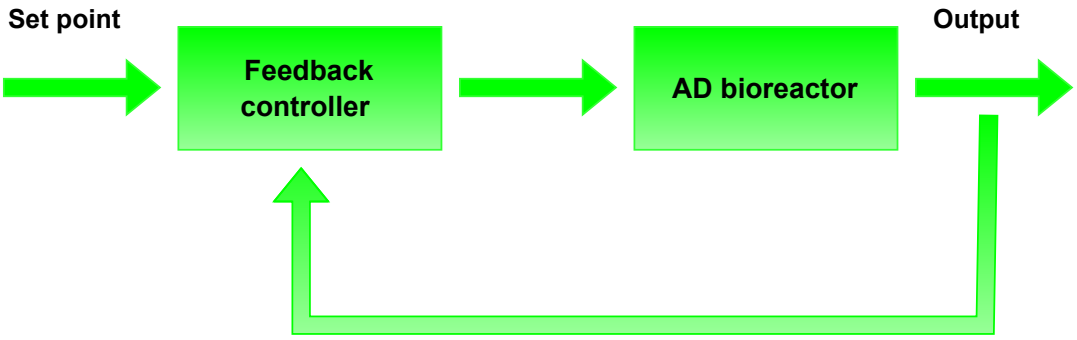


Figure 2.1: Feedback control diagram

1925. However, it wasn't broadly utilized in industrial processes until the 1960s [49]. Since then, feedforward control has emerged as one of the most frequently employed control algorithms in the industrial sector. A fully feedforward-operated system should have the capability to identify disruptions and act preemptively. However, it won't be able to modify the performance of the system. Therefore, feedforward control systems typically require the addition of certain types of feedback control for optimal performance. In AD, given the process's high complexity, feedforward controllers are typically used alongside feedback controllers. A heating method that integrates a Smith predictor with a fractional order PID controller has been established in [50]. Another form of feedforward controllers for AD bioreactor, known as disturbance accommodating control, has been utilized. This control strategy has been developed to regulate the output gas flow rate by controlling the quantity of wastewater to be treated, as detailed in [51]. A linear reference-feedforward/output-feedback methodology has been formulated specifically for an up-flow fixed bed anaerobic digester in [52]. The digester was employed in the treatment of industrial wine distillery wastewater. Remarkably, this approach facilitates the regulation of Chemical Oxygen Demand (COD), notwithstanding the modeling inaccuracies associated with the kinetic terms.

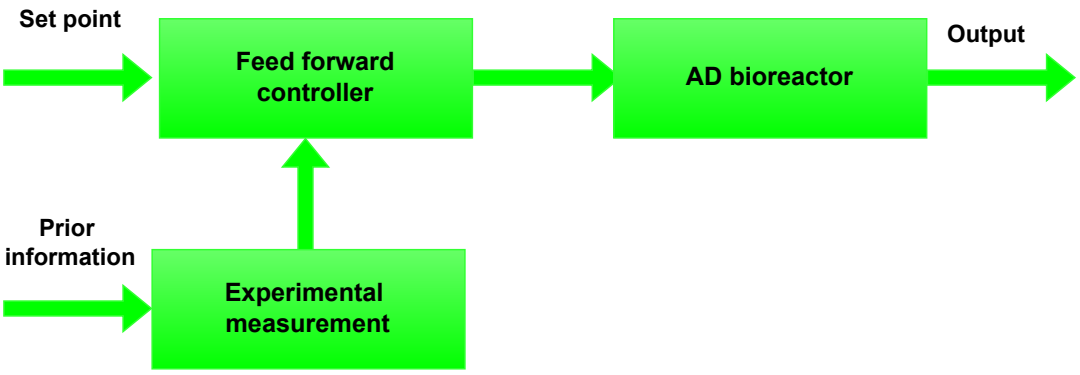


Figure 2.2: Feedforward control diagram

2.2.3 Adaptive control

Given the inherent characteristics of the AD process, which is both dynamic and non-linear, the optimal operating condition is not a fixed point. Instead, it exhibits temporal variations, reflecting the system's non-static nature influenced by its dynamic and non-linear properties. Consequently, a constant-gain control strategy such as PID control may not perform optimally when the system undergoes state transitions. In such scenarios, adaptive controllers prove to be more effective alternatives [53]. Their inherent ability to auto-tune their parameters enables them to accommodate and counteract the process variations dynamically (Figure 2.3). An adaptive controller was designed and successfully implemented in multiple AD processes to regulate the VFA concentration [54], pollution level control [55], improving the effectiveness of total nitrogen and total phosphorus removal [56] and control the ratio of total alkalinity over partial alkalinity as well [57].

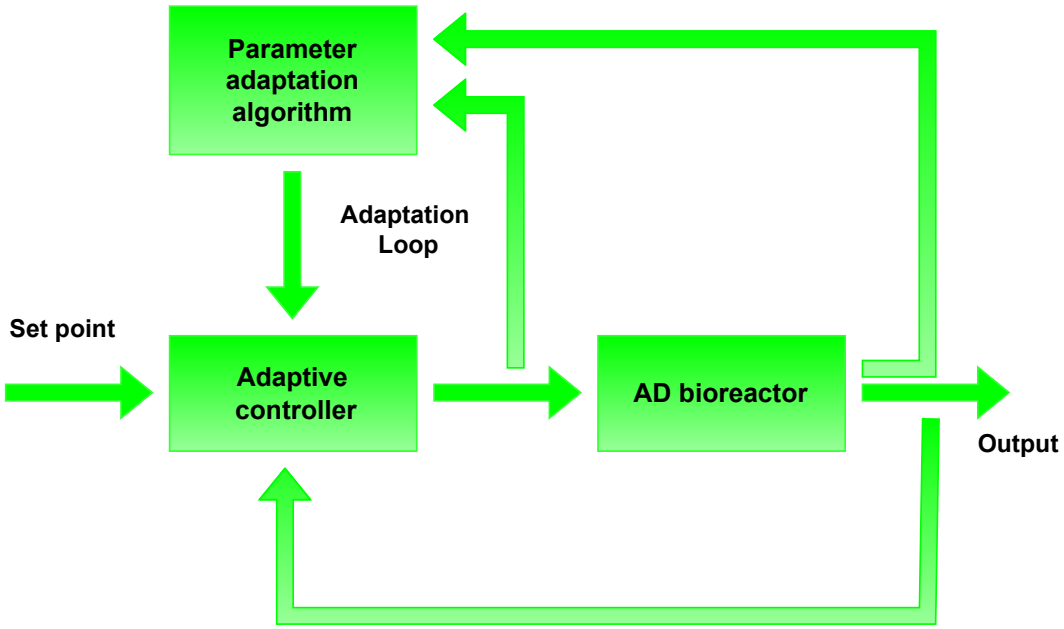


Figure 2.3: Adaptive control diagram

2.3 Theoretical aspects of Variable Structure Control Systems (VSCS)

The theory of VSCS and associated sliding modes has been the subject of detailed studies over the recent thirty years [58, 59], researchers from the Soviet Union as well as other countries were involved. In 1960, Neimark and Philipov had a discussion that led to the identification of a new problem in the field of sliding mode theory [60].
Sliding mode control (SMC) is a is a specific form of VSCS. It is widely recognized technique that involves guiding the state trajectory of a system towards a sliding surface via discontinuous

control actions. Once on the surface, the system is made to switch using an appropriate switching logic until it reaches an equilibrium point, resulting in the sliding phenomenon.

In VSCS, the controller's structure alternates between two distinct configurations. The transition between these configurations is determined by the sign of a hyperplane in the system's state space. This forces the system's state trajectory to remain on the hyperplane, resulting in what is known as an ideal sliding mode [59].

In recent years, the successful implementation of SMC in various practical applications has demonstrated the significance of sliding mode theory, which has been extensively developed over the past three decades. This is further evidenced by many special issues of academic journals dedicated to the topic of sliding mode control and its industrial applications [60, 61].

2.3.1 Problem statement

The mathematical modeling of variable structure control (formalization in the context of differential equation theory) leads to differential equations of the form:

$$\begin{aligned} \dot{x} &= f(x,t) + g(x,t)u(x,t) \\ y(t) &= h(t) \end{aligned} \tag{2.1}$$

where x: state vector n $\in R^n$,
f: vector of functions of x and t n $\in R^n$,
g: matrix of functions of x and t $\in R^{n-m}$,
u: control vector $\in R^m$,
m and n: rank of the function with m $\neq$ n.
y: the output of the system.
The objective of the control is to accurately track a desired input $y_d(t)$, despite the presence of uncertainties, model imprecision and parametric variations in f(x) and g(x), i.e. e(t) = y(t) - $y_d(t)$ tends towards zero.

2.4 Fundamentals of SMC

The objective of the control system is to ensure that the system reaches and remains on a specific surface, known as the sliding surface, until it reaches equilibrium. This control process is carried out in two stages: first, the system converges towards the sliding surface, and then it slides along it until it reaches equilibrium. The advantages of SMC are significant and numerous, including high precision, stability and robustness, ease of implementation, invariance, etc., making it particularly suitable for systems with an imprecise model. Often, it is preferable to specify the system dynamics during the reaching mode. In this case, the structure of a

controller has two parts: continuous, representing the system dynamics during sliding mode, and discontinuous, representing the system dynamics during the reaching mode. In summary, sliding mode control is divided into two parts [59]:

$$U = U_{slide} + U_{eq} \tag{2.2}$$

U_{eq}:The equivalent control vector. It is determined by using the convergence condition, we can accurately describe the ideal sliding motion of a system without uncertainties. This is achieved through the derivation of the sliding surface, which are given by: S =$\dot{S}$ = 0.
U_{slide}: Sliding is useful for compensating for model uncertainties. It consists of the sign function of the sliding surface s, multiplied by a constant K slide. The sliding surface is defined in the state space of errors to ensure the convergence of states.

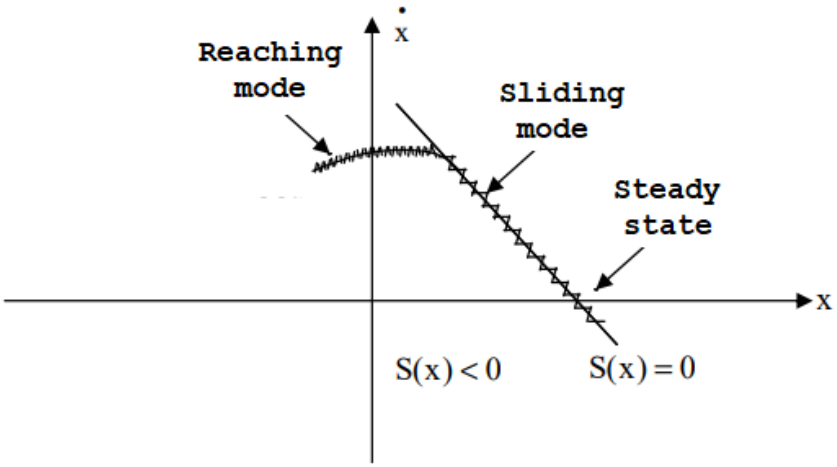


Figure 2.4: Operational mechanism of the SMC

2.4.1 Sliding surface

The sliding surface S(x) can have the form of a Hurwitz polynomial. Hence, S(x) will be a linear combination of the system states:

$$S(x) = e^{n-1} + ... + \lambda_2\ddot{e} + \lambda_1\dot{e} + \lambda_0 e \tag{2.3}$$

where n is the relative degree of the system. Thus the surface S(x) represents the desired dynamic behavior of the system. J.J Stoline [62] proposes a general equation form to determine the sliding surface that ensures the convergence of a variable towards its desired value.

$$S(x) = (\frac{d}{dt} + \lambda_i)^{n-1} e \tag{2.4}$$

With $e(x) = y_d - y$.
Where n is the relative degree of the system.
λ_i Positive constant that interprets the bandwidth of the desired control.
The change in value of the control vector u depends on the sign of the surface S(x) [59]:

$$\begin{cases} u(x) = u^+ & \text{if } S(x) > 0 \\ u(x) = u^- & \text{if } S(x) < 0 \end{cases} \tag{2.5}$$

The surface S(x)=0 is known as the switching surface and the control has a variable structure on this surface.

2.4.2 Equivalent control

The equivalent control method (Utkin method [59]) is a means to determine the movement on the sliding surface. An equivalent control vector U_{eq} is defined as the equations of the ideal sliding regime.
In sliding regime we have:

$$\dot{S}(x) = \frac{dS}{dt} = \frac{\partial S}{\partial x}\frac{\partial x}{\partial t} = \frac{\partial S}{\partial x}\{f(x,t) + g(x,t)U_{eq}(t)\} + \frac{\partial S}{\partial x}\{g(x,t)U_{slide}\} \tag{2.6}$$

In sliding mode and in steady state, the derivative of the surface is zero (because S(x)= 0 and $U_{slide} = 0$). Thus we obtain:

$$U_{eq} = -\{\frac{\partial S}{\partial x}g(x,t)\}^{-1}\{\frac{\partial S}{\partial x}f(x,t)\} \tag{2.7}$$

Substituting U_{eq} in 2.3.1 by its expression in 2.7, we obtain:

$$\dot{x} = f(x,t) - g(x,t)\{\frac{\partial S}{\partial x}^{-1}\}\{\frac{\partial S}{\partial x}f(x,t)\} \tag{2.8}$$

This equation represents the dynamics of the equivalent system in the sliding surface.
The equivalent control vector U_{eq} could be interpreted as the average value of the control vector U during rapid switching between) u^+ and u^-.

2.4.3 Discontinuous control

The addition of the U_{slide} term to the control law U ensures the attractivity of the sliding surface S(x). This is true if and only if $\dot{S}(x)S(x) <$0. This condition defines the region in which the sliding mode exists. During the convergence mode, we replace the U_{eq} term with its value given by equation (2.3.1) in equation 2.7 . We therefore obtain a new expression for the derivative of the surface, namely:

$$\dot{S}(x) = \frac{\partial S}{\partial x}\{g(x,t)U_{slide}\} \tag{2.9}$$

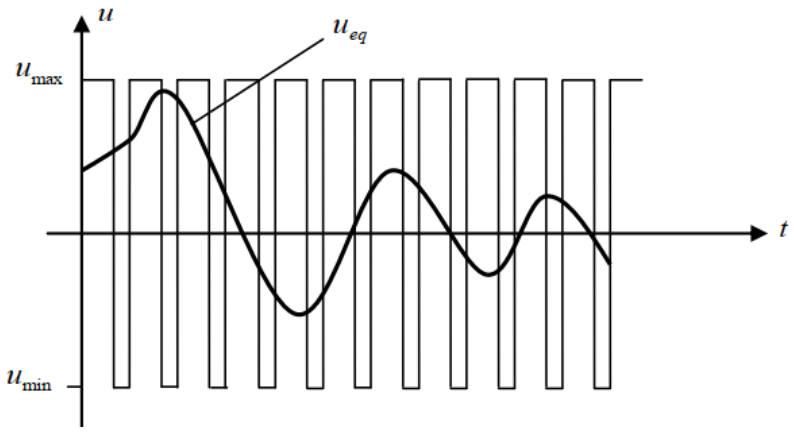


Figure 2.5: Equivalent control

The challenge is to find U_{slide} that satisfies the following condition:

$$\dot{S}(x)S(x) = S(x)\frac{\partial S}{\partial x}\{g(x,t)U_{slide}\} < 0 \tag{2.10}$$

One of the simplest approaches is to choose U_{slide} whereas it takes the form of a relay. With this approach, the control vector U_{slide} can be expressed as follows:

$$U_{slide} = K.sign(S(x)) = K\frac{|S(x)|}{S(x)} \tag{2.11}$$

where K is a positive gain.
Replacing 2.11 in 2.10, we obtain:

$$\dot{S}(x)S(x) = \frac{\partial S}{\partial x}\{g(x,t)K|S(x)|\} < 0 \tag{2.12}$$

Where the term $\frac{\partial S}{\partial x}g(x,t)$ is always negative for the class of systems we are considering. The gain K must be positive in order to meet the conditions of attractiveness and stability.

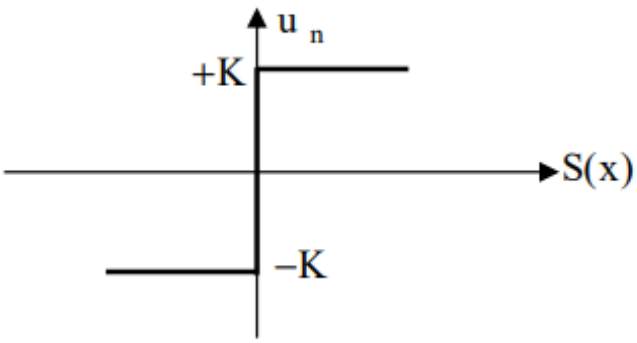


Figure 2.6: The discontinuous sign function

2.4.4 Existence of the sliding mode

The second important aspect, after the sliding surface design, is to ensure the existence of the sliding mode. The existence problem represents a generalized stability problem. The second Lyapunov method provides us with a natural environment for analysis. Specifically, the stability of the sliding surface requires the choice of a positive definite generalized scalar Lyapunov function V(x, t). This is the first convergence condition that allows the system dynamics to converge towards the sliding surfaces.
The Lyapunov function is defined as follows:

$$V = \frac{1}{2}S^2(x) \tag{2.13}$$

For the Lyapunov function to decrease, it is sufficient to ensure that its derivative is negative. This is verified if:

$$\dot{S}(x)S(x) < 0 \tag{2.14}$$

This fundamental inequality, known as an existence condition, is used to solve the problem of synthesizing variable structure systems. It therefore allows us to determine the adjustment parameters. As long as Eq. 2.14 is verified, the dynamics of the system on S(x), as well as its stability, are independent of the function f(x,t). They depend only on the parameters of the chosen surface. This explains the robustness of these control laws with respect to disturbances acting on the control part.

2.5 Chattering reduction in SMC

When discontinuities are introduced to the command, it results in oscillations at the system's output, a phenomenon known as chattering (Fig.3.2). This is a significant drawback of the traditional sliding mode. In the 1980s, several strategies were suggested to mitigate chattering. Authors in [63] suggested substituting the discontinuous "sign" function with a continuous command like the saturation function, pseudo-sign function, arc tangent function, or hyperbolic tangent function. While these functions render the control continuous, the trajectories only converge towards the vicinity of the sliding surface [64]. An observer was proposed in [65] to eliminate chattering, although this method is sensitive to uncertainties. The concept of second-order sliding mode control was introduced in [66], ensuring the convergence of the sliding variable and its first derivative to zero in finite time. This approach was further developed by [60] to reduce chattering and ensure system robustness of any relative degree. Recent studies have proposed adaptive SMC commands of order 1 and higher [67, 68]. These control laws employ dynamic gains that adapt to disturbances and uncertainties in the controlled system.

2.5.1 Addressing Chattering in SMC with Boundary layer solutions

It is known as the “boundary layer solution”, which involves replacing the sign function with a continuous approximation, of the high gain type, only in the vicinity of the surface. Among the functions used, we will mention the saturation function,

Saturation function

It consists of approximating the discontinuous control by a continuous law in the vicinity of the surface, this function is named $'sat'$ (Figure 2.7). In this case, the control becomes:

$$u = -ksat(S) \tag{2.15}$$

$$sat(x) = \begin{cases} \dfrac{S}{\phi} & \text{if } |S| \leq \phi \\ sign(S) & \text{else} \end{cases} \tag{2.16}$$

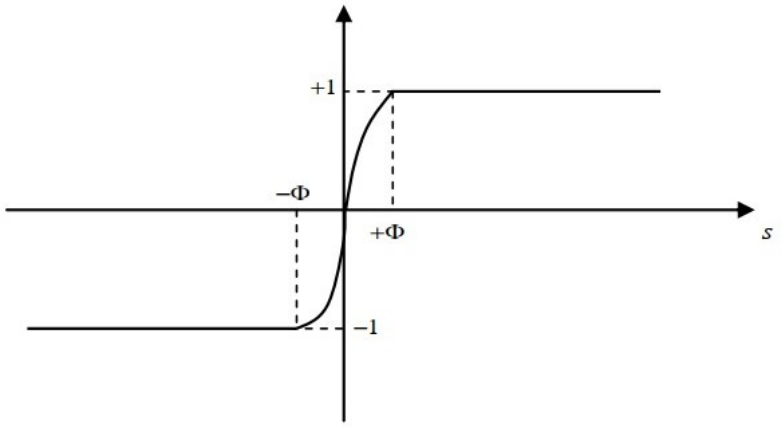


Figure 2.7: The saturation function

Hyperbolic Tangent

The hyperbolic tangent function, often denoted as $\tanh(S)$, is defined as follows:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{2.17}$$

The tanh(S) function has a similar shape to the sign function, but it transitions smoothly between -1 and 1, rather than abruptly. However, it is important to note that the use of tanh(S) may also result in slower convergence to the sliding surface compared to the sign function. Therefore, this choice depends on the specific requirements of the controlled system.

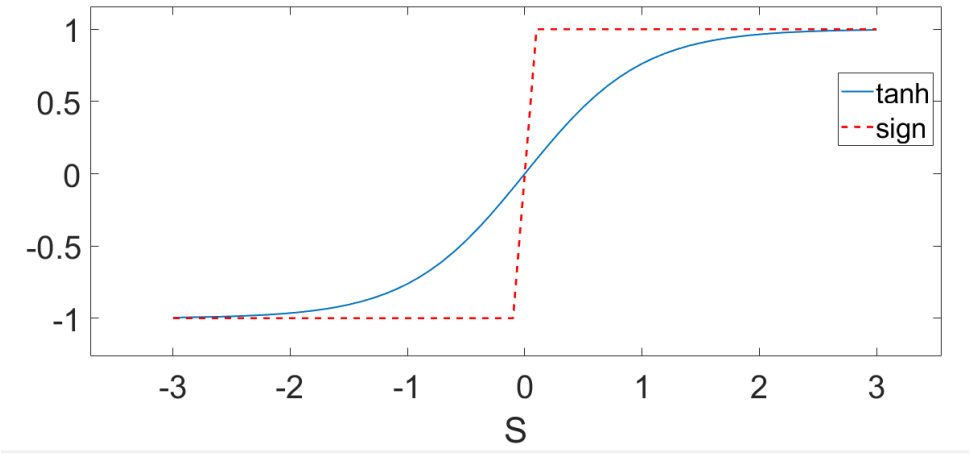


Figure 2.8: The hyperbolic tangent function

The solution of Ambrosino [69]

The function of the equation 2.18 demonstrated its effectiveness in the works of Ambrosino et al. [69] in the synthesis of the adaptive control signal, resulting in a significant reduction in chattering.

$$\frac{S}{|S| + \epsilon} \tag{2.18}$$

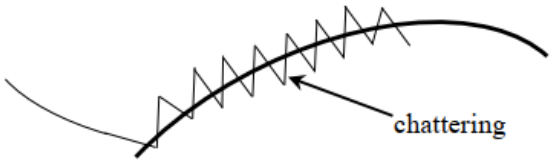


Figure 2.9: Chattering phenomenon in SMC

2.6 Integral sliding mode control

Integral Sliding Mode Control (ISMC) is a robust control method that combines the principles of Sliding Mode Control (SMC) and integral control. The main idea behind ISMC is to design a control law that forces the system to reach a sliding surface in finite time and stay on it indefinitely. Integral Sliding Mode Control (ISMC) is an advanced control strategy that combines the robustness of Sliding Mode Control (SMC) with the accuracy of integral control. The key feature of ISMC is the design of the sliding surface, which incorporates the integral of the system error.

The sliding surface in ISMC is defined as:

$$S = e + \int edt \tag{2.19}$$

where: $(e) = y_d - y$. The integral action is added to improve the robustness of the system against uncertainties and disturbances. It helps to eliminate the steady-state error that might occur due to model uncertainties or external disturbances.
The sliding variable s is usually defined as a function of the system states and the desired trajectory.

2.6.1 Characteristics of Integral Sliding Mode Control

The characteristics of Integral Sliding Modes (ISMC) can be encapsulated as follows:

- The sliding mode is enforced throughout the entire system response, eliminating the reaching phase.
- The order of motion during the sliding mode is identical to that of the original system.
- The impact of unmatched uncertainty can be mitigated through an appropriate selection of the sliding surface.
- The system motion is invariant to matched uncertainties during the sliding mode.
- The ISMC approach can be retrofitted to an existing feedback controller, demonstrating its adaptability.

2.7 Conclusion

In conclusion, this chapter provides a comprehensive exploration of Variable Structure Control Systems (VSCS) and Sliding Mode Control (SMC). It presents the theoretical aspects, addresses the challenges, and delves into the fundamentals of these control systems. The chapter also discusses the issue of chattering in SMC and presents solutions to mitigate it. Furthermore, it introduces the concept of Integral Sliding Mode Control, highlighting its unique characteristics and benefits. This chapter serves as a valuable resource for anyone seeking to understand and apply these advanced control systems in various fields.

CHAPTER 3

Parameter Identification: techniques and approaches

3.1 Introduction

This chapter delves into the process of parameter identification in mathematical models. It begins with an introduction and rationale for parameter identification, followed by a discussion on the mathematical foundations, including optimization basics and variables.
The chapter explores various optimization techniques, with a detailed focus on the Genetic Algorithm. It also discusses Multiple Linear Regression (MLR) and Analysis of Variance (ANOVA) as tools for parameter identification, and introduces measures of model fit like Root Mean Squared Error (RMSE), the Coefficient of Determination (R squared) and the P-value test as well. The chapter concludes with a summary of the key points discussed.

3.2 Introduction to parameter identification

Identifying the parameters of a reduced AD model is a challenging and delicate task. This difficulty arises from the well-known issue of parameters being poorly identifiable [70]. Thus, the selection of an appropriate identification method becomes paramount in these instances. However, due to the inherent complexity and non-linearity of the AD process, the identification procedure for AD models is not uniform and may vary on a case-by-case basis. Traditional identification techniques, predominantly those that are gradient-based search methods (for instance, regression-based methods), are extensively employed in the identification of AD models [38, 40]. These methods are characterized by a local search scope, which significantly increases the likelihood of the algorithm becoming trapped in local minima/maxima. To address this issue, the implementation of a multi-start strategy has been suggested [71]. Numerous studies

have indicated that methods with a local search scope tend to be more dependable when the identification variables are converted to a logarithmic scale [72]. However, these methods are not without their limitations. Recently, computational intelligence methods such as Genetic Algorithms (GA) have been utilized to tackle a wide array of optimization problems. These techniques have also been employed in the identification of AD models [73, 74]. One of the key advantages of these methods over others is their effectiveness and robustness in handling uncertainty and insufficient noisy information. This is largely due to their global search scope, which seeks solutions across an entire domain space and avoids the issue of local minima/maxima. In fact, these global search identification methods may exhibit greater accuracy when utilizing the logarithmic scale [75].

3.3 Mathematical foundations

The field of parameter identification is a large and diverse area of research, with numerous techniques developed for estimating model parameters. These techniques range from purely statistical methods to optimization-based approaches, each with its own strengths and limitations. Statistical methods, such as Maximum Likelihood Estimation (MLE) and Maximum A Posteriori (MAP), provide a solid theoretical foundation for parameter estimation. These methods are based on probability theory and are particularly useful when we have a probabilistic model of the data. However, in many practical situations, the model is complex and non-linear, with many parameters to estimate. In such cases, optimization techniques often prove to be more effective. These methods view parameter identification as an optimization problem, where the goal is to find the set of parameters that minimize or maximize a certain objective function, often representing the error between the predicted and observed data. In the following subsections, we will explain the fundamentals of optimization techniques for parameter estimation.

3.3.1 Optimization basics and terminology

First, we will define the most common concepts of any optimization method:

3.3.2 Optimization variables

The parameters of the objective function are adjusted during the optimization process to find the best solution(s). These parameters are also known as optimization variables or design variables.

3.3.3 Objective function

To determine the best fit of a model to experimental data, an appropriate criterion must be chosen for the optimal solution of the model parameter vector. Various objective functions, mostly in the form of output-error criteria, have been used for parameter identification in AD models. These

criteria measure the deviation between the model and real system outputs. The selected function can affect the behavior of the optimization procedure and how it adjusts the parameters. The sum of least squares is the most commonly used objective function :

$$S = (a_0, a_1, ..., a_m) = \sum_{i=0}^{m} [y_i - f(x_i)]^2 \tag{3.1}$$

3.3.4 Search space

Defined by the set of combinations of parameter values (all possible solutions to the problem). It corresponds to the solution space. The dimension of the search space is defined by the number of parameters involved in the solutions for example, if each solution is defined by four parameters, the search space is four-dimensional. It is also called the parameter space.

3.3.5 Constraints

A constraint is a strict restriction imposed on the value of a variable, which prevents it from extending indefinitely in certain directions. A constrained optimization problem involves finding the minimum of an objective function, denoted as $f : R^n \rightarrow R$, over a set of variables $z \in R^n$. This must be done under the condition that a set of constraints, denoted as C, are satisfied between the variables. In other words, we are looking for the values of z that minimize f while also satisfying all the constraints in C. This is a common task in many fields such as economics, engineering, and machine learning.

3.4 Optimization Techniques for Parameter Identification

Optimizing a multivariable function is a common problem in many research fields, and there is significant expertise on the subject. However, it is important to note that the minimum of the criterion can be either local or global, as shown in Figure 3.4. A local minimum is limited to a specific region in the space, while a global minimum encompasses the entire parameter space. The convergence to a minimum, whether local or global, is highly dependent on the initial values chosen for the parameters. This can be understood by examining curve 3.4; if the initial values are to the left of the local minimum, it is likely that the minimization algorithm will converge to the local minimum and not reach the global minimum. To avoid converging only to a local minimum, it may be necessary to perform multiple iterations of the minimization algorithm with different sets of initial parameter values to cover as wide a range of values as possible.

3.4.1 Local optimization algorithms

There are numerous procedures for local minimization, many of which rely on the derivatives of the objective function, either directly or through numerical calculation. The derivative based algorithms are called gradient based methods. These methods search for a direction in the parameter space to find the minimum. Once the direction is chosen, the algorithm determines the magnitude of the change in parameter values. However the convergence may be slow [76]. Another alternative is Newton's method, which can converge more quickly but may also diverge or oscillate around the solution. The Levenberg-Marquardt algorithm [77] is a well-known modification that combines the advantages of both gradient and Newton's methods. By choosing appropriate algorithm parameters, it is possible to achieve both convergence and rapid convergence.

To improve convergence, the fact that most functions are well approximated by a quadratic function near the optimum can be utilized. The method of Fletcher [78] is among the best known in this class of methods. However, author in [79] has pointed out that these techniques may be susceptible to local minima if they converge quickly. The Powell method, with improvements proposed by Brent [80], is a local minimization method that does not rely on the calculation or approximation of derivatives. It involves repeated one-dimensional searches along multiple directions and is one of the best methods in terms of balancing convergence speed and insensitivity to local minima.

The Powell method, with improvements proposed by Brent [81], is a local minimization method that does not rely on the calculation or approximation of derivatives. It involves repeated one-dimensional searches along multiple directions and is considered one of the best methods in terms of balancing convergence speed and insensitivity to local minima. Another well-known local minimization method that does not require derivative information is the Simplex method, proposed by Nelder and Mead (not to be confused with the simplex method in linear programming). A simplex is a set of (p + 1) parameter vectors in the parameter space. The method involves comparing the objective function with these edges and replacing the edge with the highest value with another point in the p-dimensional parameter space. The method is appreciated for its robustness against local minima, ease of implementation, and reasonable convergence speed [81].

3.4.2 Global optimization algorithms

Global minimization methods can be divided into two groups [82]. The first group includes purely deterministic methods, such as the gridding method, which evaluates the objective function at many predefined points on a grid covering the parameter space. With enough function evaluations, the minima may be found. However, this method is not very effective unless the grid is refined after a series of evaluations. The second group of global methods, known as random probing methods, incorporates random decisions in the search for the optimum. Adaptive methods in this group use information from previous evaluations. One such method is simulated

Table 3.1: Review of the recent studies in the field of the optimization of the AD process

Reference	Model	N° of Identified parameters	Identification method	Aim of the study/Findings
This work [84]	AM2HN model	12 parameters	GA	- GA based identification of the AM2HN model. - Ga are more accurate when introducing the optimization variables in the log scale
Bernard et al.[38]	Two reaction model (AM2)	13 parameters	Multiple linear regression	Two-step AD model was developed. Multiple linear regression based identification procedure was proposed
Hassam et al.[40]	Two reaction model derived from ADM1 (AM2HN)	13 parameters	Multiple linear regression	Development of three stages AD model. The same identification procedure proposed by Bernard et al.[38] was employed.
Biernacki et al.[85]	ADM1 model	4 parameters	Downhill Simplex algorithm	- The ADM1 is very reliable model to simulate the biogas production process. -New parameters for hydrolysis step were optimized.
K.Yoshida et al. [86]	Two reaction model from ADM1	7 parameters	Least squares estimator	Dynamical model development An adaptive parameter estimation system was proposed, applicable to AD processes with various conditions.
E.Chorukova et al. [87]	Newly created mathematical model	11 parameters	Sequential quadratic programming (SQP) algorithm/ GA	New AD model of wheat straw was created. The GA outperforms the SQP algorithm as an effective identification method.
Z.fatolahi et al. [88]	ADM1 model	13 parameters	GA	The Shanon entropy was employed as a sensitivity index. Parameters with high sensitivity index were optimized via GA
A.D. Bona [89]	modified AMOCO model	6 parameters	Maximum likelihood estimator	A reduced model was developed. Reformulation of the model by mean of the linear fractional transformation in order to identify the unknown parameters.
R.J Ashraf et al . [90]	ADM1 model	2 parameters	GA	-Minimisation of the flared and unmet gas in AD system. -Multi-objective optimization is mandatory for off grid AD systems.
D.Barik et al. [91]	ANN model	4 parameters	GA	The quality of Biogas produced differs from one substrate to another. ANN simulations was performed and optimized by the GA.

annealing [81], which allows the search to occasionally move in a different direction rather than always towards a potential solution (which could be a local minimum). This method preceded popular methods like genetic algorithms (GA) [83], which start with an initial population of randomly sampled potential solutions in the parameter space. New solutions are generated by simulating biological evolution through crossbreeding, mutation, and selection among the parameter populations. Properly defining the algorithm's parameters is crucial for successful implementation.

3.5 Genetic Algorithm

The GA was introduced by J.H. Holland and his team at Michigan. Their work began in 1960, and in 1975, he published the first book titled "Adaptation in Natural and Artificial Systems." The objectives of the GA are to improve understanding of natural adaptation processes and to design artificial systems with properties similar to natural systems. The genetic algorithm (GA) is a search algorithm based on the mechanisms of natural selection and genetics. It combines a "survival of the fittest" strategy with a random but structured exchange of information. For a problem for which a solution is unknown, a set of possible solutions is randomly created. This set is called the population. The characteristics (or variables to be determined) are then

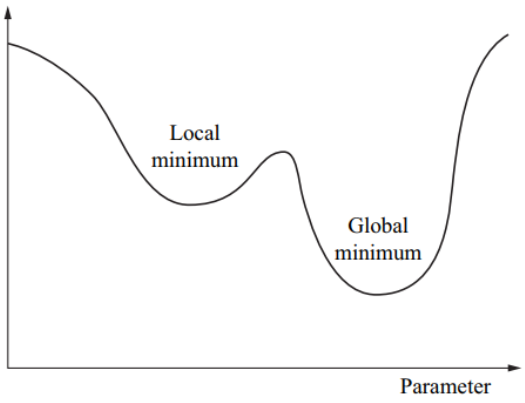


Figure 3.1: Identification problem: Global and local minima/maxima

used in gene sequences that will be combined with other genes to form chromosomes and later individuals. Each solution is associated with an individual, and this individual is evaluated and ranked according to its resemblance to the best, but still unknown, solution to the problem. It can be shown that using a natural selection process inspired by Darwin, this method will gradually converge to a solution.
As in biological systems subject to constraints, the best individuals in the population are those with the best chance of reproducing and transmitting part of their genetic heritage to the next generation. A new population, or generation, is then created by combining the genes of the parents. It is expected that some individuals in the new generation will possess the best characteristics of their two parents, and therefore they will be better and will be a better solution to the problem. The new group (the new generation) is then subjected to the same selection criteria, and later generates its own offspring. This process is repeated several times until all individuals have the same genetic heritage. The members of this last generation, who are usually very different from their ancestors, possess genetic information that corresponds to the best solution to the problem.

3.5.1 Encoding and initialization

First, the different possible values of the variable whose optimal value is being sought must be represented in a form that can be used for a GA: this is called encoding. This encoding process allows for a connection to be established between the value of the variable and the individuals within the population. This is done in a way that mimics the genotype-phenotype transcription that occurs in biological organisms.
There are primarily two types of encoding methods used in genetic algorithms:

Binary Encoding:

This is one of the most common methods of encoding where chromosomes are represented as strings of one (1) or/and zero (0). Each position in the chromosome represents a particular characteristic of the solution. However,the performance of the algorithm degrades in the face

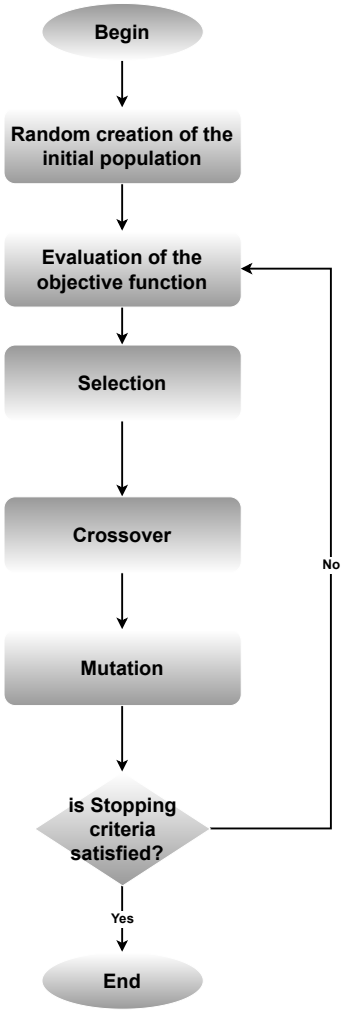


Figure 3.2: Genetic Algorithm flowchart

of high-dimensional optimization problems with high numerical precision. For such problems, genetic algorithms based on binary strings have poor performance [92].

Real Encoding:

This method is used when binary encoding is not sufficient, especially in problems where complex values, such as real numbers, are used. For example, if we are seeking the optimum of a function of n variables, we can simply use a chromosome containing the n variables. With this type of encoding, the chromosome evaluation procedure is faster due to the absence of the transcoding step (from binary to real). Results given by Michalewicz [92] show that real representation often leads to better accuracy and significant gains in terms of execution time

3.5.2 Selection

Selection in genetic algorithms serves the same purpose as natural selection. It adheres to Darwin's principle of survival of the fittest, determining which individuals will survive and which

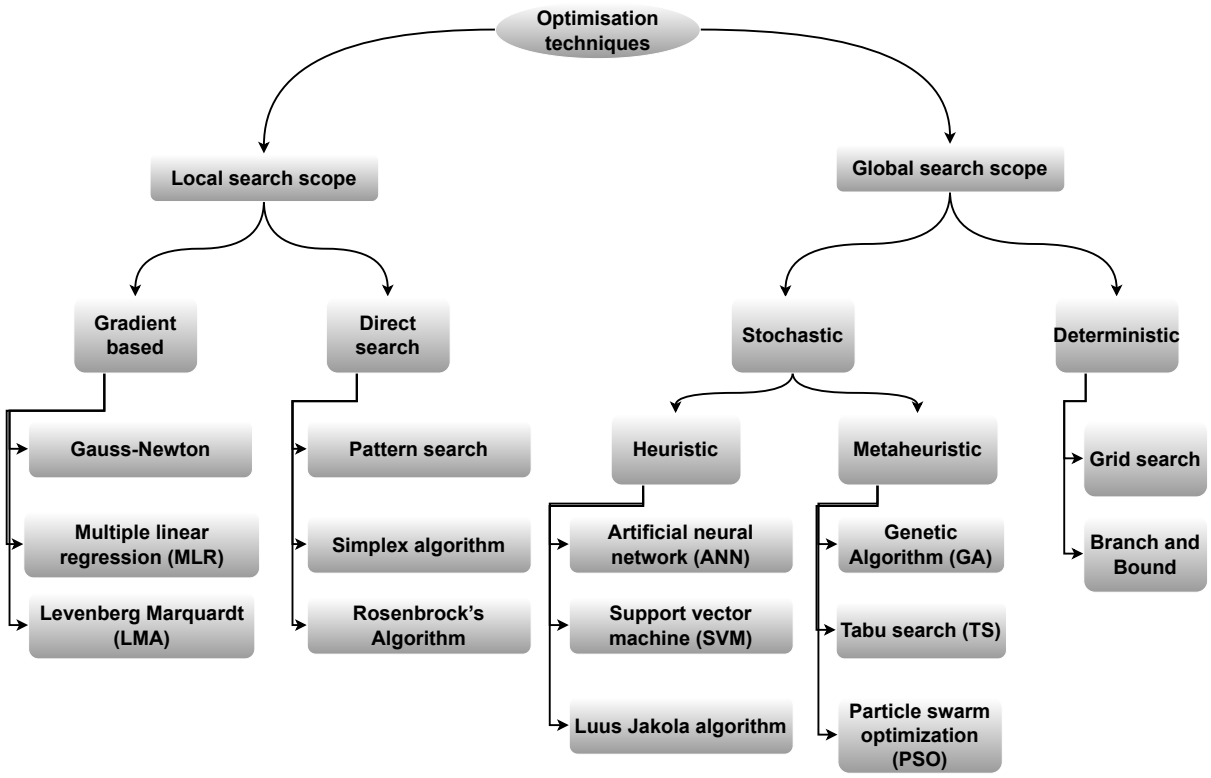


Figure 3.3: The main families of optimization techniques

will disappear based on the objective functions. The selected individuals form an intermediate population. There are various methods of selection available:

Roulette wheel

One common approach to implementing the selection operator in genetic algorithms is through the use of a method known as the biased roulette or roulette wheel selection. This method is based on the concept of a casino roulette wheel, where the probability of an individual (chromosome) being selected is proportional to its fitness value. As a result, individuals with higher fitness values have a greater chance of being selected for reproduction, allowing their genetic information to be passed on to future generations. The figure 3.4 is adapted from [93] and depicts the mechanism of the roulette wheel selection method.

Tournament

In Tournament selection, a pressure is applied by conducting a competition among 'S' participants, where 'S' represents the size of the tournament. The individual with the highest fitness among the 'S' competitors is declared the winner of the tournament and is then added to the mating pool. The mating pool, which consists of tournament winners, has a higher average fitness than the overall population. This difference in fitness levels provides the selection pressure that propels the Genetic Algorithm to enhance the fitness of each subsequent generation.

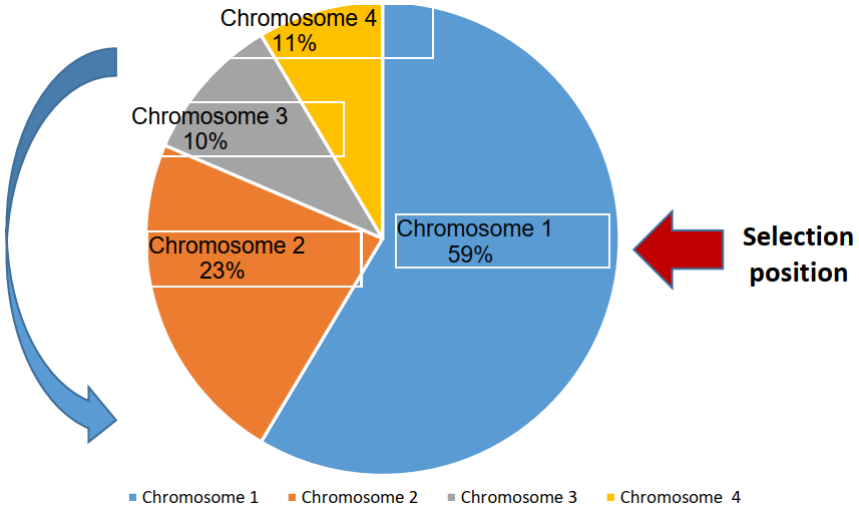


Figure 3.4: Depiction of Roulette Wheel Mechanism in Genetic Algorithms: Proportional Representation of Best and Worst Individuals

By simply increasing the size of the tournament, we can intensify the selection pressure [94]. This is because, on average, a winner from a larger tournament will have a higher fitness level than a winner from a smaller tournament.

Ranking selection

In 1985, Baker brought forward the concept of ranking selection in the context of genetic algorithms. The principle is simple. Organize the population from the highest to the lowest, allocate the quantity of copies each individual should get based on a decreasing assignment function, and then execute proportionate selection as per that assignment. Grefenstette and Baker (1989) provided some theoretical insights about such methods, but these theories offer no assistance in assessing anticipated performance. As per the methodology proposed by Baker et al. [95], the calculation of the probability P for rank position R_i in linear ranking is as follows:

$$P(R_i) = \frac{1}{n}(sp - (2sp - 2)\frac{i-1}{i+1}) \tag{3.2}$$

with $1 \leq i \geq n$; $1 \leq sp \geq 2$; $P(R_i) \geq 0$; $\sum_{i=1}^{n} P(R_i) = 1$ where sp is the selection pressure and ”n” is the total number of individuals.

3.5.3 Crossover

The process of natural selection, as discussed in the previous section, enables the choice of individuals to serve as parents for the crossover phase. This phase facilitates the exchange of genes among individuals to generate new solutions. Various crossover methods exist in scientific literature. The simplest method involves splitting chromosomes into two (single-point) or three parts (double-point) . This is followed by an exchange of genes between two chromosomes, as

illustrated in Fig. 3.5.
In single-point crossover, the chromosomes of two parent solutions are interchanged before and after a specific point. On the other hand, double-point crossover involves two crossover points, and only the chromosomes located between these points are interchanged.

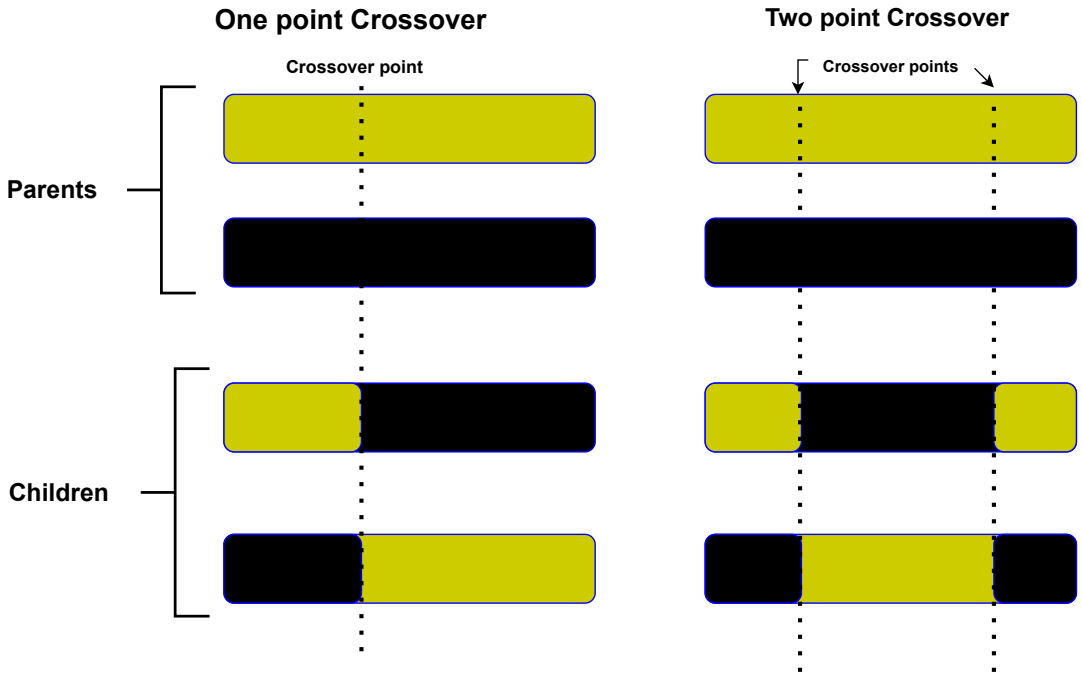


Figure 3.5: Depiction of one point and two point crossover

Crossover for real valued individuals

Recombination operators of this kind, used for floating-point representations, are referred to as discrete recombination. They possess a characteristic where, when creating an offspring 'z' from parents 'x' and 'y', the allele value for gene 'i' is determined by either 'x_i' or 'y_i', with both having an equal probability of being chosen.

Another alternative is commonly referred to as intermediate or arithmetic recombination, an operator is used that generates a new allele value in the offspring at each gene position, which falls between the parent values. Using the previously mentioned terminology, we get $z_i = \alpha x_i + (1-\alpha)y_i$ for some α in the range [0,1]. This method allows recombination to produce new gene material. However, it has a downside: the averaging process reduces the range of allele values in the population for each gene.

Thirdly, an operator is utilized that generates a new allele value in the offspring at each position. This value is close to one of the parent values but may exceed them (i.e., it could be larger than the highest of the two values, or smaller than the smallest). Operators of this kind can produce

new genetic material without limiting the range. These operators are commonly referred to as blend recombination [96].

3.5.4 Mutation

The process of the crossover operator involves the swapping of genes between chromosomes. However, a potential drawback of this method is that it does not allow for the introduction of new genes. Furthermore, if all solutions deteriorate and become trapped in locally optimal solutions, the crossover operation fails to yield diverse solutions with new genes that are distinct from those in the parent chromosomes. In order to tackle this issue, the genetic algorithm (GA) also incorporates the mutation operator. During mutation, genetic alteration occur randomly. Every gene in a child chromosome that is created in the crossover stage has a chance to mutate, which is determined by a parameter called Probability of Mutation (P_m)and it decides whether it will mutate or not. The parameter P_m is a number between 0 and 1. For each gene, a number P is randomly chosen in the same interval as P_m. If $P > P_m$ the gene stays the same and does not undergo any mutation, otherwise if $P \leq P_m$ mutation occurs and a random gene from the range of potential values will substitute the original gene. The literature also presents various techniques for mutation, such as Uniform mutation [97], Power [98], Non-uniform [99], Shrink [100], Gaussian [101], etc. The mutation operator introduces the concept of ergodicity to genetic algorithms used in space exploration. Ergodicity ensures that the algorithm has the potential to access every point within the state space, even though it may not visit all of them during the problem-solving process. Consequently, the theoretical convergence characteristics of genetic algorithms are heavily reliant on the mutation operator [83].

3.5.5 Elitism

The processes of crossover and mutation result in changes to the genetic material of chromosomes. Based on the Pm value, there's a possibility that offspring could completely replace their parent generation. This scenario might result in the loss of optimal solutions from the present generation.A potential solution to this issue could be the implementation of a new operator known as "Elitism" [102]. he initial version of the Genetic Algorithm did not incorporate this operator. However, Recently, a growing number of researchers have begun to utilize it due to the substantial enhancement it brings to the performance of Genetic Algorithm. The functionality of this operator is quite straightforward. A segment of the top-performing chromosomes in the current population is preserved and carried forward into the next generation without undergoing any alterations. This safeguards these solutions from being adversely affected by crossover and mutation during the generation of new populations.

3.5.6 Stopping criteria

In any iterative algorithm, it's necessary to establish a stopping condition. This can be expressed in several ways, including:
- Halting the algorithm once a satisfactory solution has been achieved.
- Ceasing operation if no enhancements are observed over a specified number of generations.
- Terminating the process if the count of generations exceeds a certain threshold.

3.6 Multiple linear regression (MLR)

3.6.1 Definition 1

The variable you aim to predict is known as the response variable. The variables utilized to make this prediction are referred to as predictor variables.
Notation: In the following text, the response variable, which we want to predict, is typically represented by Y. The predictor variables, which are used to predict the response variable, are often represented by $x_1, \ldots, x_p$. These predictor variables are also known as independent variables, explanatory variables, carriers, or covariates. The response variable is also referred to as the dependent variable. Frequently, the predictor variables are gathered in a vector x. In such cases, x^T denotes the transpose of x.

3.6.2 Definition 2

Regression analysis involves examining the conditional distribution of the response variable $Y|x$, given the predictor variables represented by the vector x = $(x_1, \ldots, x_p)^T$.

3.6.3 Definition 3

Assuming that the response variable is Y and there is at least one predictor variable xi, then the model for multiple linear regression (MLR) is defined as follows:

$$Y = \beta_1 x_{i,1} + \beta_2 x_{i,2} + ... + \beta_p x_{i,p} + \epsilon_i \tag{3.3}$$

For i ranging from 1 to n, where n represents the sample size and the random variable ϵ_i denotes the ith error. . These n equations can be succinctly expressed in matrix notation as $Y = X\beta + e$. In this context, Y is a dependent variable vector of size $n \times 1$, X is a matrix of predictors of size $n \times p$, β is a vector of unknown coefficients of size $p \times 1$, and ϵ is a vector of unknown errors of size $n \times 1$. This can also be expressed as follows:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,p} \\ x_{2,1} & x_{2,2} & \dots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \dots & x_{n,p} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \tag{3.4}$$

It's common for the first column of X to be X_1 = 1, which is a vector of ones with a size of $n \times 1$. The ith instance, represented as $(x_i^T, Yi_i) = (x_{i,1}, x_{i,2}, ..., x_{i,p}, Y_i)$, corresponds to the ith row, x_i^T , of X and the ith element of Y. If $x_{i,1}$ is always equal to 1, then it could be excluded.
with the following assumptions:
- Variable types: The fundamental premise of a linear regression is that the dependent variable is continuous. In other words, it can assume any value within a certain range, from the smallest to the largest possible value.
-Linearity: Linear regression modeling operates under the assumption that there's a linear relationship between the outcome and each explanatory variable. However, this might not always hold true.
- Normal distribution of the residuals: The assumption of residuals' normality suggests that they are symmetrically distributed around zero, without any skewness or kurtosis. This implies that the model has successfully captured the primary patterns and sources of variation in the data, and that any errors are random and independent. This normal distribution embodies the residual variability in your dependent variable, over and above the variability elucidated by the model itself.
- Homoscedasticity: Homoscedasticity pertains to the dispersion of residuals or error terms. If this assumption is valid, then the error terms exhibit a uniform variance. Put differently, the error associated with each observation is independent of any variable included in the model. A different way to express this is that the standard deviation of the error terms remains constant and is not influenced by the values of the explanatory variables.
- Multicolinearity: When variables are collinear, the model can lose its stability. This is often signaled by a large standard error surrounding the estimation of the β coefficients, and the coefficients may undergo significant changes when variables are added or removed from the model. The model is unable to differentiate between the impacts of different variables, thereby violating one of the assumptions of linear regression.

Parameter identification can effectively utilize multiple linear regression. The objective function in this context is typically represented by the sum of squared residuals. This function essentially quantifies the discrepancy between the predicted and actual values, indicating the extent to which the model's outcomes deviate from reality. The ultimate aim is to identify the coefficients that minimize this sum.
In Fig. 3.6, There are no noticeable pattern within the variance: these residuals appear to be fulfilling the assumption of homoscedasticity.

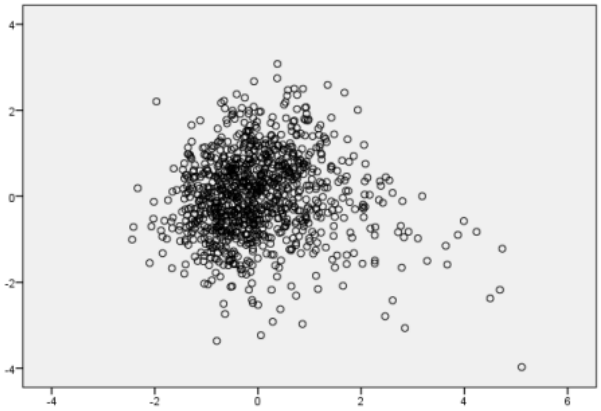


Figure 3.6: Illustrative example of Visualizing Residuals to Validate Homoscedasticity in Multiple Regression

3.7 Analysis of Variance (ANOVA) for parameter identification

ANOVA is a vital statistical tool in parameter identification. It allows for the comparison of multiple groups, which is particularly useful when dealing with several parameters. One of the key outputs of ANOVA is the p-value, which plays a significant role in determining the statistical significance of the parameters being compared.
The p-value helps us understand if the differences observed between the parameters are due to random chance or if they are statistically significant. If the p-value is less than a predetermined significance level, we can conclude that the parameters have a significant effect on the dependent variable.
This ability to discern significant differences between parameters makes ANOVA an essential tool in parameter identification. It provides a quantitative measure of the strength of evidence against the null hypothesis, thereby aiding in data-driven decision-making.

3.7.1 P-value

A two-tailed hypothesis test could be conducted to ascertain whether the parameter values deviate significantly from zero. Hence, a p-value could be calculated for each parameter that was fitted. The p-value is indicative of the likelihood of achieving the estimated value, Assuming the validity of the null hypothesis, denoted as $H0 : \beta_r = 0$, we are postulating that the true value of the parameter β_r is indeed zero. When the p-value falls below a predetermined significance level, typically set at 0.05, we reject the null hypothesis. This leads us to accept the alternative hypothesis, which proposes that the parameter differs from zero. If the p-value exceeds the significance level, it indicates that we cannot reject the null hypothesis. A similar process can be applied to determine if a parameter value statistically deviates from a specific non-zero value. The Matlab function "$linhyptest$" is developed for this purpose [103].

3.7.2 Confidence interval of the identified parameters

Estimating confidence intervals (CIs) is a common statistical technique used to evaluate the uncertainty associated with a parameter's mean value [104]. However, many published studies report fitted values of kinetic parameters without including CIs [105, 106]. The reason for not including statistical analysis in parameter estimation may be that it is not always straightforward and often requires advanced numerical methods. This lack of statistical analysis has two drawbacks. First, the evaluation of the model's quality becomes biased because there are no quantitative indicators to assess its performance. Second, when results are published, other researchers may draw incorrect conclusions when comparing reported parameters with those obtained in their own studies. This can lead to a biased perception of a fermentation's reproducibility, understood as achieving similar kinetic parameters in different trials. Given the complexity and variability of biological processes, it is crucial to calculate confidence intervals for kinetic parameters and propagate their variability to simulated predicted variables to minimize biased conclusions. A confidence level of 95 $\%$ is frequently used, although other levels like 90% and 99% are also possible. With a 95% confidence level, it is believed that the estimate will be within the confidence interval's range 95 out of 100 times. In Matlab, one can utilize the "nlparci" function to calculate the 95% confidence intervals (CIs). The inputs required for this function are the estimated parameters, the residuals of the model, and the Jacobian matrix of the kinetic model [103].

3.8 Root Mean Squared Error (RMSE)

The RMSE or Erreur Quadratique Moyenne (EQM) in French, is a statistical measure that evaluates the difference between the values predicted by a model and the actual observed values. It is commonly used in many fields, including anaerobic digestion [107, 84], where it serves quantify the error of a prediction model and to assess the accuracy of the identified parameters. The RMSE is calculated by taking the square root of the average of the squares of the errors. The "errors" are calculated as the difference between each predicted value and the corresponding actual value. By squaring these errors, we ensure that all errors are positive and that larger errors have a greater weight in the final calculation.
In sum, the RMSE is a useful measure for comparing the performance of different prediction models. A model with a lower RMSE is generally considered better, as it tends to make predictions closer to the actual values.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2} \tag{3.5}$$

3.9 Coefficient of determination (R squared)

R-squared (R^2) is a statistical metric that signifies the percentage of the dependent variable's variance that the independent variable in a regression model accounts for. This statistic is utilized in the realm of mathematical models, primarily for predicting future behavior of complex systems or testing hypotheses based on related information. It offers a gauge of the model's effectiveness in replicating observed outcomes, measured by the proportion of total outcome variation that the model accounts for.

Consider a dataset with "n" observations denoted as $y_1, ..., y_n$ (collectively referred to as y_1 or as a vector $y = [y_1, ..., y_n]^T$). Each of these observations is associated with a predicted value $f_1, ..., f_n$ (referred to as f_1, or sometimes $\hat{y}_i$, as a vector "f").

The residuals are defined as: $e_i = y_i - f_i$. If $\bar{y}$ is the mean of the observed data, then we can represent these as:

$y = [y1, ..., yn]^T$ $f = [f1, ..., fn]$ $e_i = y_i - f_i$ $\bar{y} =$ mean of y

The sum of squares of residuals, also known as the residual sum of squares, is given by:

$$SS_{res} = \sum_i (y_i - f_i)^2 = \sum_i e_i^2 \tag{3.6}$$

The total sum of squares, which is proportional to the variance of the data, is given by:

$$SS_{tot} = \sum_i (y_i - \bar{y})^2 \tag{3.7}$$

The most general definition of the coefficient of determination is:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \tag{3.8}$$

3.10 Conclusion

In this chapter, we delved into parameter identification techniques and approaches. The chapter covered a range of optimization techniques, from local to global optimization algorithms, and introduced the genetic algorithm as a potential solution. The chapter also discussed MLR technique, analysis of variance (ANOVA), and other statistical tools for parameter identification.

CHAPTER 4

General Discussion

4.1 Introduction

In this chapter, we delve into two significant aspects of our study.
Firstly, in section 4.2, we explore the concept of Parameter Identification using GA. Here, we will discuss how these algorithms can be used for parameter identification, a crucial step in the modeling of complex systems. We will walk through the process, explaining how the GA searches through the solution space and optimizes the parameters.
An important part of this section is a comparison between GA and Multiple Linear Regression for parameter identification. We will examine the strengths and weaknesses of both methods, and discuss scenarios where one might be preferred over the other. This comparative analysis aims to provide a deeper understanding of these techniques and their applicability in different situations.
Following this, in sections 4.3 and 4.4, we turn our attention to Sliding Mode Control (SMC) and the Integral Sliding Mode Control (ISMC) respectively. We will discuss the principles of SMC, ISMC and how it can be applied to control systems. We will also delve into the mathematics behind both techniques, providing a comprehensive understanding of its workings.
This chapter aims to provide a thorough understanding of these methodologies, demonstrating their practical applications and effectiveness in dealing with complex systems.

4.2 Parameter Identification using GA

The model parameters were identified three times: initially through the multiple regression method, and subsequently twice using the proposed Genetic Algorithm (GA) approach. Consequently, three sets of parameters were derived, as depicted in Tables 4.4 and 4.3. A comparison is

Table 4.1: GA MATLAB configuration

GA parameter	Value/function
Selection function	stochastic uniform
Generation size	1200
Population size	200
Mutation function	Adaptive feasible
Crossover	Scattered
Function tolerance	1e-6

provided. The accuracy of the AM2HN model prediction was verified by comparing the model response with the ADM1 data. The search space is delineated by the upper bounds (Ub) and lower bounds (Lb) of the twelve parameters [k_1, k_2, k_3, k_4, k_5, k_6, $\mu_{1,max}$, k_{s_1}, $\mu_{2,max}$, k_{s_2}, k_i, k_{hyd}] as follows:
Lb = [5, 300, 300, 100, 200, 10, 0.1, 0.1, 0.1, 1, 20, 1],
Ub = [150, 550, 600, 300, 500, 300, 5, 8, 5, 10, 300, 10]
The choice of these bounds was based on parameter values reported in the literature [38, 40]. The simulations are carried out using Matlab 2017on a PC/Intel Core i5-1145G7 CPU@2.6 GHz, 8 GB Memory (RAM), Windows 10 (64 bit) operating system.

$$obj.1 = Min \sum_{p=1}^{n} |X_{1m,p} - X_{1e,p}| + |X_{2m,p} - X_{2e,p}| + |S_{1m,p} - S_{1e,p}| + |S_{2m,p} - S_{2e,p}| + |C_{m,p} - C_{e,p}| + |X_{Tm,p} - X_{Te,p}| + |q_{ch_4m,p} - q_{ch_4e,p}| \quad (4.1)$$

$$obj.2 = Min \sum_{p=1}^{n} |lnX_{1,m,p} - lnX_{1,e,p}| + |lnX_{2,m,p} - lnX_{2,e,p}| + |lnS_{1m,p} - lnS_{1e,p}| + |lnS_{2m,p} - lnS_{2e,p}| + |lnC_{m,p} - lnC_{e,p}| + |lnX_{Tm,p} - lnX_{Te,p}| + |lnq_{ch_4m,p} - lnq_{ch_4e,p}| \quad (4.2)$$

The term m refers to the measured (observed) data, e refers to the estimated data, and n is the total number of observations. The Obj. 1 and Obj. 2 are graphically represented by Figures 4.1 and 4.2, respectively. These two figures were obtained by varying the AM2HN model parameters between the lower and upper bounds. Consequently, 100 iterations were performed. The impact of data standardization is clearly evident as it smooths the objective function and eliminates saddle points, as well as local maxima and minima.

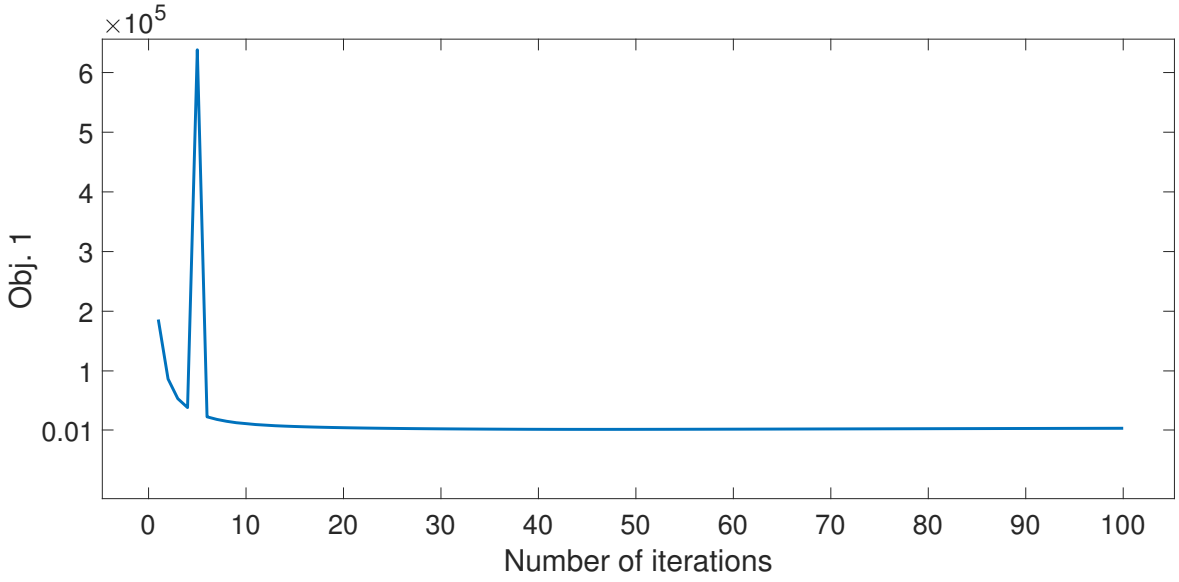


Figure 4.1: Graphical representation of the linear objective function obj. 1

Table 4.2: Steady state data generated by the ADM1 virtual platform

HRT	$S_1(KgCODm^{-3})$	$S_2(KgCODm^{-3})$	$X_1(KgCODm^{-3})$	$X_2(KgCODm^{-3})$	$X_T(KgCODm^{-3})$	$C(kmoleCm^{-3})$	$q_c(molm^{-3}d^{-1})$	$q_{ch4}(kmolem^{-3}d^{-1})$
5	0.92	86.65	1.42	1.11	1.295	71.6	46.98	3.96
10	0.25	15.2	1.35	1.19	0.78	139.5	26.16	2.55
15	0.15	5.36	1.27	1.13	0.58	150	18.8	1.777
20	0.12	3.37	1.19	1.06	0.47	153	15.04	1.355
25	0.1	2.56	1.12	1	0.4	154	12.77	1.098
30	0.08	2.12	1.05	0.94	0.35	155.5	11.21	.9216
35	0.07	1.84	1	0.89	0.31	156.5	10.12	.798
40	0.07	1.66	0.95	0.85	0.28	157	9.3	.7047
45	0.06	1.51	0.9	0.8	0.26	158	8.6	.6265
50	0.06	1.42	0.86	0.77	0.24	158.5	8.13	.5714

4.2.1 Parameter identifiability assessment

In the context of model identification, a parameter is said to be identifiable if it consistently assumes a unique value each time the identification process is performed using the same data set. This means that, given the same data, the identification process should yield the same value for the parameter each time it's run. If a parameter is identifiable, it suggests that the parameter can be reliably estimated from the data, which is a crucial aspect of model validation and prediction.

Prior to proceeding with system control, it is imperative to verify the identifiability of the parameters. The GA is executed 150 times, each iteration utilizing the same data derived from ADM1. The outcomes of these identifications are subsequently depicted using a histogram (refer to Figures 4.3 and 4.4).

The histogram effectively illustrates the distribution of the identified parameter values. Each bar in the histogram signifies a range of parameter values, with the height of the bar representing the frequency of a parameter value within that range being identified.
The tallest peak in the histogram, which corresponds to the value most frequently identified, is

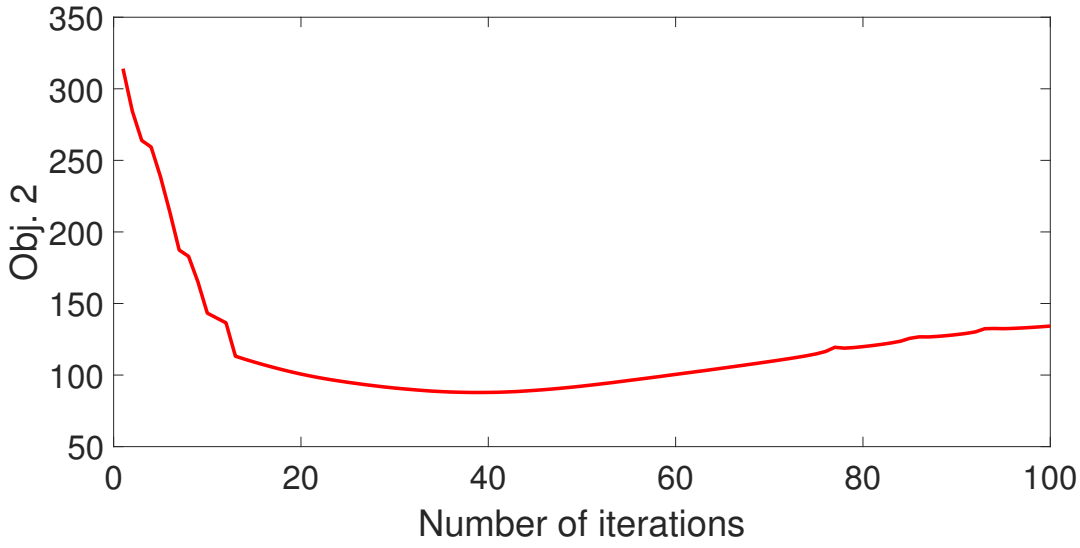


Figure 4.2: Graphical representation of the logarithmic objective function obj. 2

Table 4.3: The optimized stoichiometric parameters

Identification method	Stoichiometric parameters					
	$\mu_{1,max}$	k_{s1}	$\mu_{2,max}$	k_{s2}	k_i	k_{hyd}
Regression	0.33	0.39	0.29	9.6	2.6	5
GA (obj.1)	0.4	5.78	0.25	4.92	278	1.08
GA (obj.2)	0.314	0.328	0.326	9.86	98.5	3.09

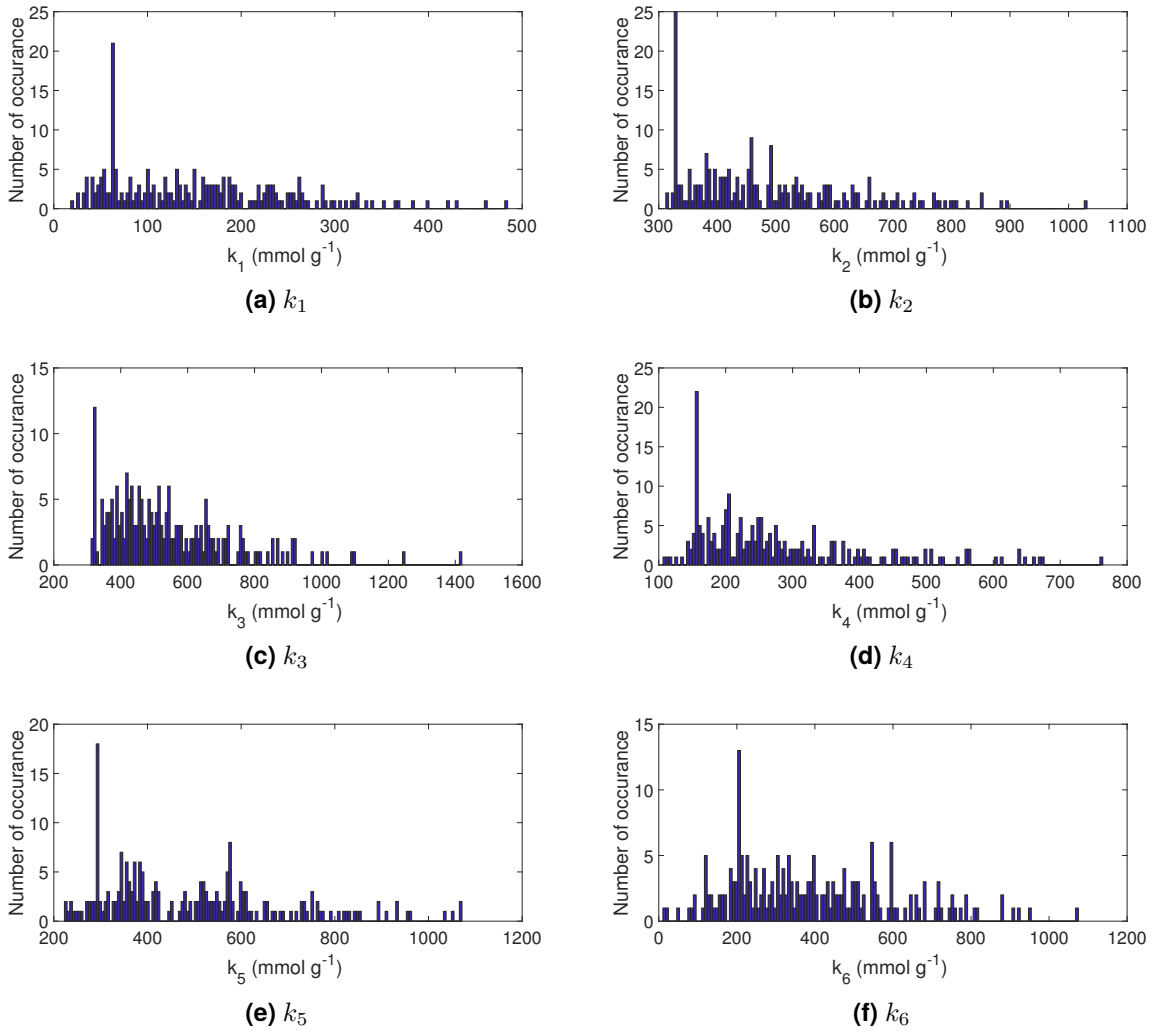
considered to represent the true value of the parameter. This is predicated on the assumption that the most reliable estimate of the parameter will be the one that is identified most frequently and therefore the parameter is said identifiable.

The figure 4.6 and the table 4.2 is adapted from [84]. It represents the AM2HN model's response using the parameters identified by the regression method, GA(obj.1), and GA(obj.2) and is compared to the ADM1 data shown in table 4.2.

Figures 4.6. a, b, d, e, and f demonstrate that the model with parameters obtained by GA yields superior results compared to the traditional method (multiple linear regression). In Figures 4.6. d and g, the linear objective function (GA(obj.1)) provides more precise results than the logarithmic objective function (GA(obj.2)).

Table 4.4: The optimized kinetic parameters

Identification method	The kinetic parameters					
	k_1	k_2	k_3	k_4	k_5	k_6
Regression	19	455	504	-255	535	15.75
GA (obj.1)	93	319	409.5	299	495	61.5
GA (obj.2)	62	331	323	157.7	293.7	204.56



(a) k_1 **(b)** k_2 **(c)** k_3 **(d)** k_4 **(e)** k_5 **(f)** k_6

Figure 4.3: Identifiability of kinetic parameters

Conversely, 4.6. a, b, c, e, and f indicate that the logarithmic objective function outperforms the linear objective function.

Only one figure (Figure 4.6. c) out of seven suggests that the regression method produces better results than the proposed methods with both objective functions. In Figures 4.6. a, e, and f, it is evident that the model parameters related to X_1, X_T, and C align closely with the experimental data when the natural logarithm is incorporated into the objective function. Upon computation, the Root Mean Square Error (RMSE) for the majority of the state variables (all state variables excluding the case of q_{ch4} and S_1) was significantly smaller than the RMSE associated with multiple regression. This underscores the effectiveness of the proposed method, as illustrated in Table 4.5.

However, a notable limitation of the multiple regression method is the negative value yielded for the estimated parameter k_4, which lacks physical significance. As a result, Figure 4.6. e, which pertains to the parameter k_4, has not been plotted. Furthermore, it is observed that data normalization using log transformation significantly enhances the precision and accuracy of

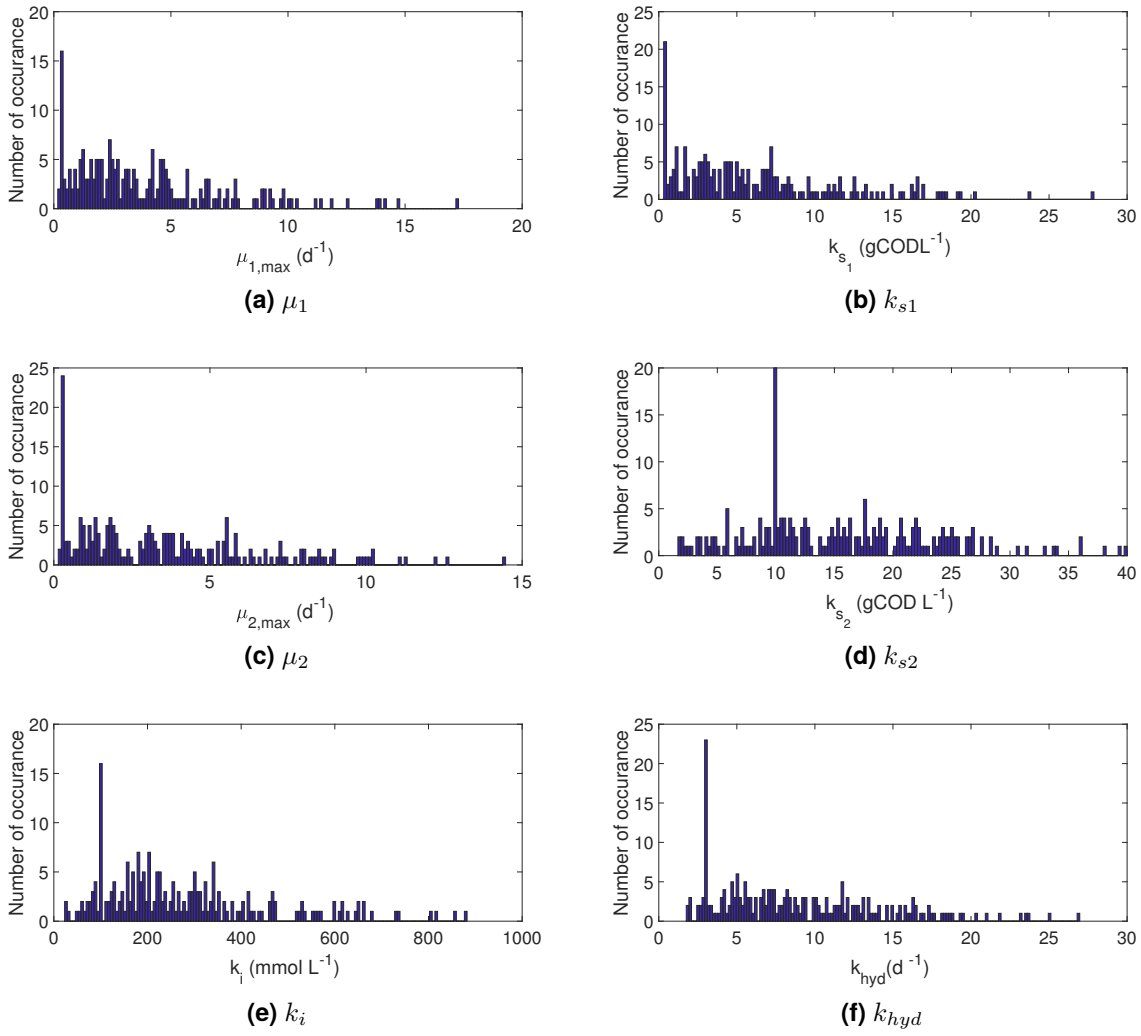


Figure 4.4: Identifiability of stoichiometric parameters

the GA. Similarly, based on the figures presented earlier, it can be inferred that GA is a more precise and efficient methodology compared to traditional methods for identifying AD models. The data modeled and measured align well, and the results derived from the GA approach are consistent with findings reported in existing literature, attesting to the effectiveness of the proposed method. The suggested algorithm holds promise as a tool for optimizing parameters of analogous bioprocess models, particularly when the objective function is minimized on a logarithmic scale.

4.2.2 P-value

In order to verify whether the estimated parameters statistically significant (different from zero), a two-sided statistical hypothesis test was conducted. Consequently, a p-value was calculated for each set of parameter that was estimated. The Matlab function $linhyptest$ is available for this purpose. The function needs the following arguments: the parameter estimates, the covariance matrix, the degree of freedom of the covariance matrix and hypothesis matrix H. The $linhyptest$

Table 4.5: RMSE related to each state variable of the AM2HN model using data acquired from the virtual platform as reference

Method	RMSE related to each State variable						
	X_1	X_2	S_1	S_2	X_T	C	q_{ch_4}
Regression	1.8549	0.8541	0.0324	1.2760^3	0.1878	1.0917^3	1.0061
GA (obj.1)	0.4032	0.4294	0.3814	1.7521	0.5637	6.9546	0.1737
GA (obj.2)	0.1083	0.0407	0.0499	1.2648	0.0565	17.2809	6.1723

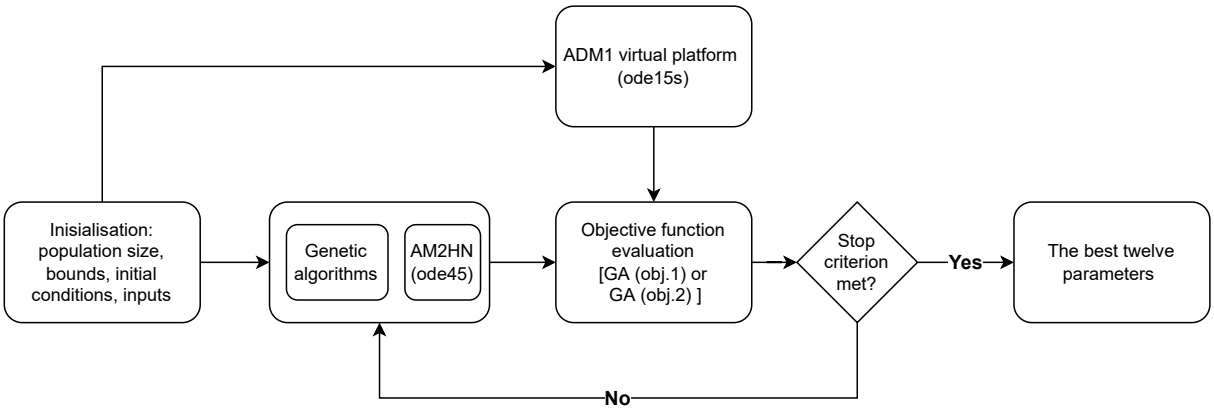


Figure 4.5: Methodology flowchart

function in Matlab conducts a statistical test that assumes the parameter estimates follow an asymptotic normal distribution. This function can be utilized after any estimation method. When used in the context of linear regression, the p-values obtained are exact. However, for other methods, the p-values are approximations and might not be as precise as those obtained from methods based on a likelihood ratio. In other words, the accuracy of the p-values may vary depending on the estimation procedure used [103]. The hypothesis test was conducted at the steady state, where the system becomes linear. We modeled the equations 4.3-4.9 using a multiple linear regression models.

$$\overline{S_1} = \frac{\alpha}{0.9\mu_{1,max}} \times D \times \overline{S_1} + \frac{\alpha K_{s_1}}{0.9\mu_{1,max}} \times D + O.11K_{S_1} \tag{4.3}$$

$$\overline{S_2} = \frac{\alpha}{0.9\mu_{2,max}} \times D \times \overline{S_2} + \frac{\alpha K_{s_2}}{0.9\mu_{2,max}} \times D + 0.11K_{S_1}$$

$$+\frac{\alpha}{0.9\mu_{2,max}} \times D \times \overline{S_2^2} + \frac{0.11}{K_I}\overline{S_2^2} \tag{4.4}$$

$$D(X_{T,in} - X_T) = k_{hyd}X_T \tag{4.5}$$

$$D(S_{1,in} - S_1) - k_{hyd}X_T = k_1 \times \alpha \times D \times X_1 \tag{4.6}$$

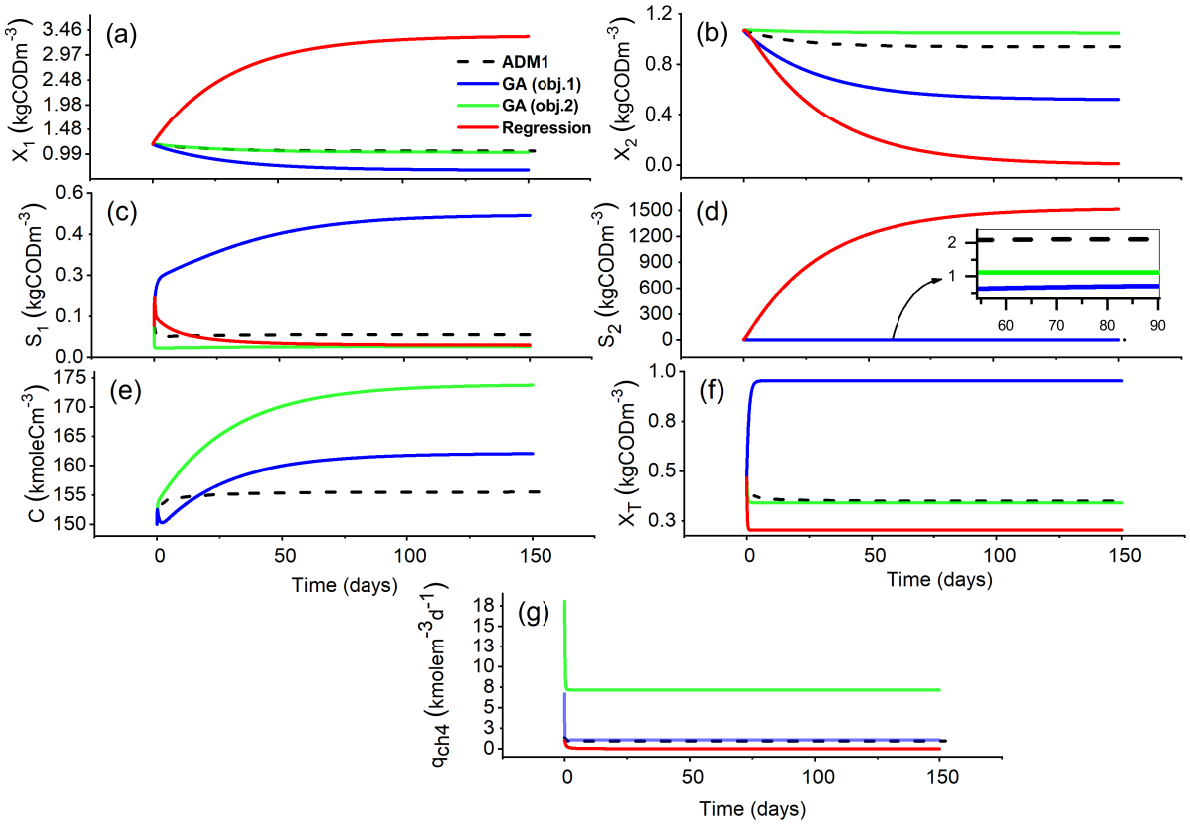


Figure 4.6: Model validation using the ADM1 data presented in table 3: results of solving the AM2HN model using the optimized parameters HRT equal to 30 days, (a): X_1, (b): X_2, (c): S_1, (d): S_2, (e): C, (f): X_T, (g): q_{ch_4} .

$$\overline{q_{ch4}} = k_6 \times \alpha \times D \times \overline{X_2} \tag{4.7}$$

$$D(S_{2,in} - S_2) = k_3 \times \alpha \times D \times X_2 - k_2 \times \alpha \times D \times X_1 \tag{4.8}$$

$$D(C_{in} - C) - q_c = k_4 \times \alpha \times D \times X_1 + k_5 \times \alpha \times D \times X_2 \tag{4.9}$$

For instance, when we need to calculate the p-value of k_2 and k_3, the MATLAB code is written as follows:

$$p = linhyptest(beta, COVB, c, H, dfe) \tag{4.10}$$

The regression model corresponding to the steady state equation 4.8 is:

$$y = k_2 \times x_1 - k_3 \times x_2 \tag{4.11}$$

Thus, $yfit = beta(1) \times x(:,1) - beta(2) \times x(:,2)$, where x is the observation vector of the independent variables and y is the response variable and $beta$ is the estimates vector. The degrees

Table 4.6: Statistical Significance of Estimated Parameters by using GA

Parameter	$[\mu_{1,max} + k_{s1}]$	$[\mu_{2,max} + k_{s2} + k_i]$	$[k_1]$	$[k_2 + k_3]$	$[k_4 + k_5]$	$[k_6]$	$[k_{hyd}]$
P-value	0.016	2.5^{-07}	4.31^{-05}	0.19	4.5^{-05}	4.9^{-05}	4.6^{-07}

Table 4.7: Statistical Significance of Estimated Parameters by using MLR

Parameter	$[\mu_{1,max} + k_{s1}]$	$[\mu_{2,max} + k_{s2} + k_i]$	$[k_1]$	$[k_2 + k_3]$	$[k_4 + k_5]$	$[k_6]$	$[k_{hyd}]$
P-value	0.016	3^{-08}	1.8^{-05}	0.06	8^{-10}	4.5^{-09}	1.5^{-09}

of freedom are calculated as follows:

$$dfe = \text{Number of observations} - \text{Number of estimates} \tag{4.12}$$

Afterward, we calculate the residuals:

$$residuals = y - yfit \tag{4.13}$$

and the Mean Squared Error (MSE):

$$MSE = \frac{\sum(residuals^2)}{\text{Total number of observations}} \tag{4.14}$$

We calculate the Jacobian matrix J of the proposed regression model: $J = [x_1 \quad -x_2]$.
Since the covariance matrix COVB can be calculated using the Jacobian and the MSE , we calculate it as follows:

$$CovB = (J' \times J)^{-1} \times MSE \tag{4.15}$$

Finally, we perform a hypothesis test using $linhyptest$, where $H = [1 \quad 0; 0 \quad 1]$ is the null hypothesis that assumes both coefficients are zero, and $c = [0; 0]$ are the expected values under the null hypothesis ($H * beta = c$).
The p-value is calculated as $p_value = linhyptest(beta, CovB, c, H, dfe)$. Results are shown in table 4.6 and 4.7. Despite GA recording less steady state error, the MLR approach has lower p-values. This is interesting because a lower p-value typically indicates a statistically significant. This could suggest that while MLR finds statistically significant relationships, GA might be capturing some complex patterns in the model that MLR is missing.

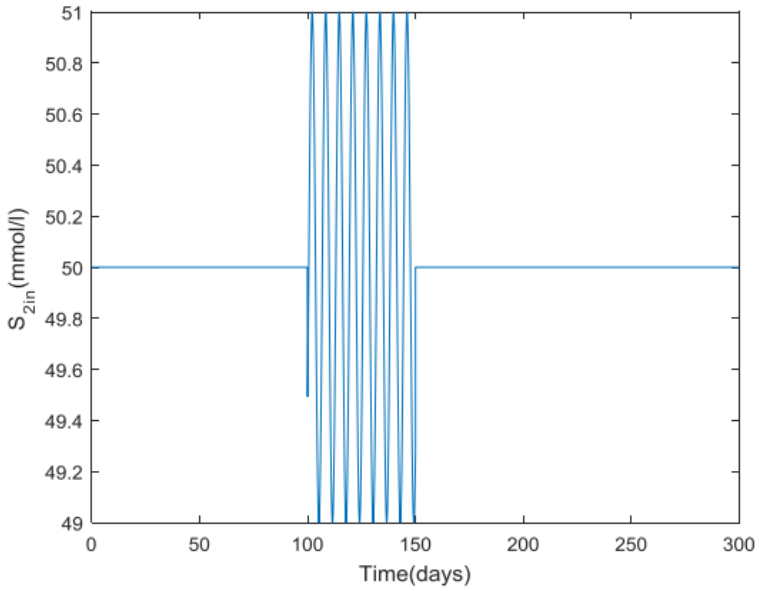


Figure 4.7: Disturbance injected between 100 and 150 days

4.3 Sliding Mode Control (SMC): Application on AM2 bioreactor

Sliding mode control has seen significant advancement, primarily due to its rapid convergence property and robustness against modeling errors and certain types of external disturbances. However, it's important not to overlook certain limitations. Sliding mode controls operate in a discontinuous manner. The discontinuities in the control algorithm directly impact the actuator. If the actuator is not designed for this type of operation, it could lead to premature wear and tear. Furthermore, the system is constantly subjected to high control to ensure its convergence to the desired state.

$$Q = k_6 X_2 \mu_2 \tag{4.16}$$

The tracking error is defined as :

$$e = Q_d - Q \tag{4.17}$$

and the sliding surface is defined as [62] :

$$S = (\frac{d}{dt} + \lambda)^{n-1} e \tag{4.18}$$

where n is the relative degree of the system.
For the system above we have n=1 (its relative degree is 1), hence : $S = e = Q_d - Q$

$$\dot{S} = \dot{Q}_d - \dot{Q} \tag{4.19}$$

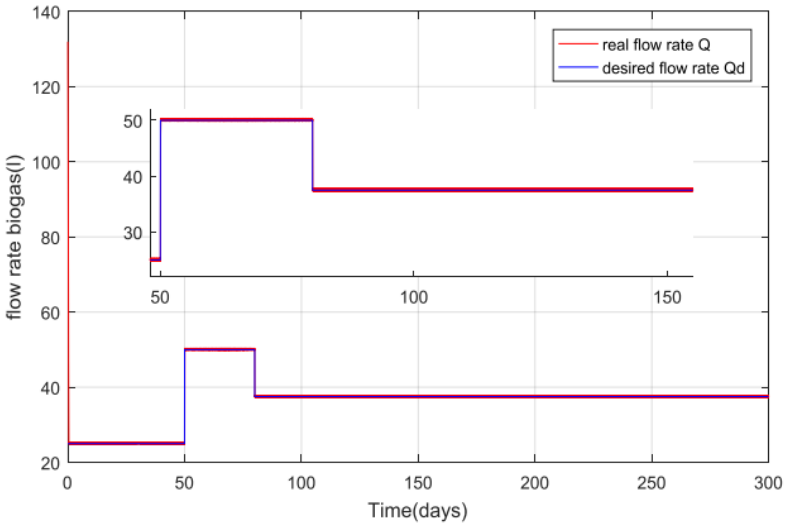


Figure 4.8: Biogas flow rate under classical SMC

with $\dot{Q}_d = 0$ (constant), $\dot{S} = -\dot{Q}$

$$\dot{S} = -k_6 \dot{\mu_2} X_2 - k_6 \mu_2^2 X2 + k_6 \mu_2 \alpha D X_2 \tag{4.20}$$

$$\dot{\mu_2} = \frac{\mu_{2,max} \dot{S}_2 (k_{s2} + \frac{S_2^2}{k_i})}{(k_{s2} + S_2 + \frac{S_2^2}{k_i})^2} \tag{4.21}$$

for $\dot{S} = 0$

$$D = \frac{\mu_{2max}(k_{s_2} - \frac{S_2^2}{k_i})(k_2 \mu_1(S_1) X_1 - k_3 \mu_2(S_2) X_2) + (k_{s_2} + S_2 + \frac{S_2^2}{k_i})^2 \mu_2^2}{\mu_{2max}(k_{s_2} - \frac{S_2^2}{k_i})(S_2 - S_{2,in}) + (k_{s_2} + S_2 + \frac{S_2^2}{k_i})^2 \mu_2 \alpha} \tag{4.22}$$

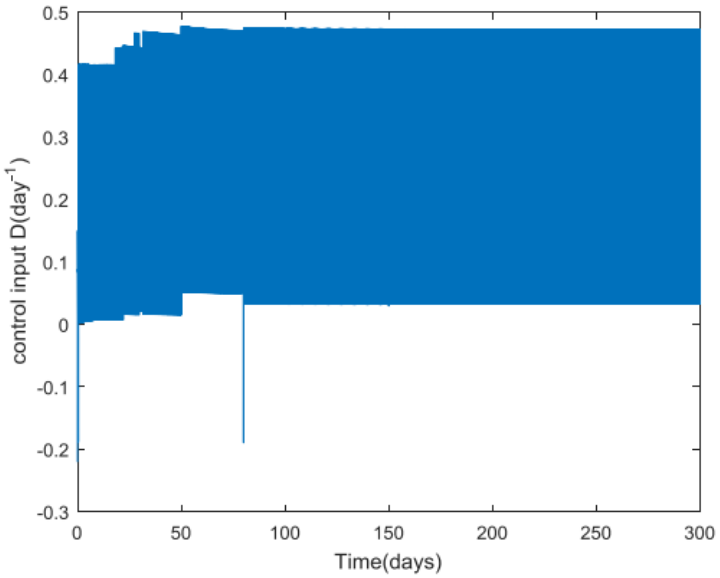


Figure 4.9: Control input D

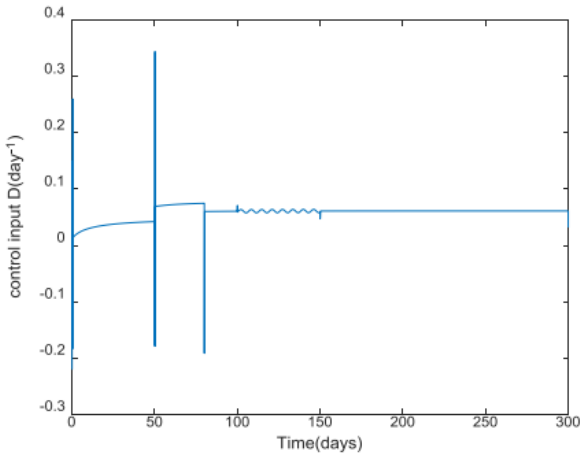


Figure 4.10: Control input D under the proposed SMC

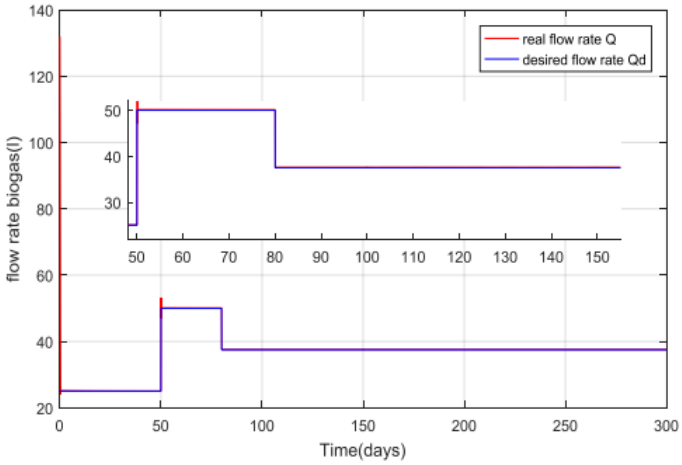


Figure 4.11: The biogas produced under the proposed SMC

$$D_{eq} = \frac{\mu_{2max}(k_{s_2} - \frac{S_2^2}{k_i})(k_2\mu_1(S_1)X_1 - k_3\mu_2(S_2)X_2) + (k_{s_2} + S_2 + \frac{S_2^2}{k_i})^2\mu_2(S_2)^2}{-\mu_{2max}(k_{s_2} - \frac{S_2^2}{k_i})(S_2 - S_{2,in}) + (k_{s_2} + S_2 + \frac{S_2^2}{k_i})^2\mu_2(S_2)\alpha} + \lambda sign(S) \tag{4.23}$$

with λ is a positive constant.

Figures 4.7, 4.8, 4.9, 4.10, 4.11, 4.12 are adapted from [39, 108]. The robustness of the proposed control law is evident in Figure 4.8, which was derived from a Simulink simulation. The classical SMC scheme demonstrated superior tracking performance. The system responded instantaneously to changes in the desired methane flow rate and disturbances in the S2in input, exhibiting no overshoots and minimal fluctuation. Furthermore, the response time was notably brief. In Figure 4.9, it is observed that the control input D remains consistently positive, albeit with considerable chattering. However, a significant reduction in this chattering phenomenon is demonstrated in Figures 4.10 and 4.11, as revealed by the simulation results. The results depicted in Figures 4.10 and 4.11 validate the superior tracking ability of the proposed control scheme in comparison to the traditional method, particularly in response to changes in the desired methane flow rate and disturbances in the $S2_{in}$ input (Figure 4.7). The application of the proposed control scheme effectively regulates the methane flow rate to match its reference, while also reducing the chatter phenomenon. The control input (D) displays significantly less fluctuation under the proposed control scheme, as illustrated in Figure 4.10 and in contrast to Figure 4.9. Notably, the proposed control scheme outperforms other methods in terms of steady-state fluctuation, underscoring its performance superiority.

4.4 Integral Sliding mode Control (ISMC)

In this section we consider a second order prototype model [109]:

$$\dot{X} = \mu X - DX \tag{4.24}$$

$$\dot{S} = -k_1\mu X - DS + F \tag{4.25}$$

Values considered for the process parameters: $k_1 = 12g/L, ks = 10g/L, D = 0.2h^{-1}, \mu_m = 2.1h^{-1}$

The initial conditions preceding control activation are given as: $X_0 = 20g/L, S_0 = 40g/L$

the gas produced is given by: $Q = k_2\mu X$. Thus, Q is the main output of interest. It's the biogas flow rate. Hence, any optimization of induces an optimization on Q. For the integral sliding

mode control (ISMC), the sliding surface (8) is selected as to include the integral error:

$$Sur = a_1 e + a_2 \dot{e} + a_3 \int e dt \tag{4.26}$$

with e is the error: $e = X_d - X$.
Taking the time derivative of the sliding surface and making use of evident simplifications leads to:

$$F_{ISMC} = B\frac{a_1}{a_2}(D - \mu) + k_1 \mu X + DS - B(\mu - D)^2 - B\frac{a_3}{a_2}(1 - \frac{X_d}{X}) \tag{4.27}$$

in which: $B = \dfrac{(S + ks)^2}{\mu_m + S}$. Parameters to be optimized: a_1, a_2 and a_3Lower bounds [0.001,0.001,0.001] and upper bounds [5,2,5] Optimal values obtained at iteration 100: $a_1 = 0.99, a_2 = 0.3, a_3 = 1$.
Simulation results for biomass and substrate concentration obtained under ISMC approach are given in Figure 4.12 showing good overall tracking performance.

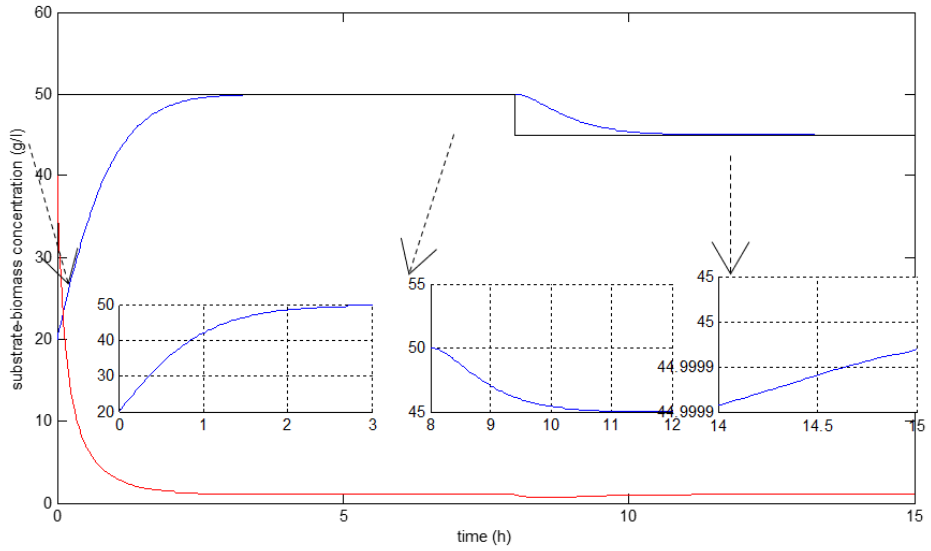


Figure 4.12: Biomass- Substrate concentration and reference under ISMC

4.5 Conclusion

In this chapter, we examined two key methodologies in control and optimization: Parameter Identification using Genetic Algorithms (GA) and control using the Sliding Mode Control (SMC) and the integral sliding mode control as well.
Section 4.2 focused on the use of GA for parameter identification in complex systems. A comparative analysis was conducted between GA and Multiple Linear Regression, elucidating the advantages and limitations of both techniques.

Section 4.3 discussed the principles and applications of SMC, a robust control design technique effective in managing system uncertainties and disturbances.
The chapter aimed to provide a comprehensive understanding of these methodologies and their practical applications in system modeling and control. The insights gained should prove beneficial for future research and applications in this field.

General conclusion and perspectives

This thesis represents a significant journey of research that has culminated in a comprehensive exploration of biogas technology, with a particular emphasis on the anaerobic digestion (AD) process. The AD process plays a pivotal role in energy production and waste management, and this work has highlighted its importance. It has also shed light on the various physico-chemical parameters that influence the process, thereby providing a deeper understanding of the intricacies involved.

The research has ventured into the realm of parameter identification techniques and approaches, traversing through a spectrum of optimization techniques. It has introduced the genetic algorithm as a potential solution, thereby broadening the scope of possibilities in this field. The use of Multiple Linear Regression (MLR), Analysis of P-value, and other statistical tools for parameter identification have also been discussed in detail, providing a robust analytical framework for future research.

The study has further delved into the control aspect of biogas technology, investigating the sliding mode control (SMC) and its theoretical aspects to regulate biogas production. The role of variable structure control systems (VSCS) and strategies for chattering reduction in SMC were discussed, contributing to the optimization of biogas production. This exploration of control mechanisms has opened up new avenues for enhancing the efficiency of biogas production.

This work is the culmination of four conference papers [110, 108, 107, 111] and two journal papers [39, 84], which is a testament to the academic rigor and relevance of this research. It has provided an opportunity to familiarize with the well-known ADM1 model, learning all its variables and gaining the ability to make different manipulations related to control and optimization. This has significantly enhanced the practical applicability of the research.

An enriching internship at the Laboratory of Biotechnology and Environment (LBE) in Narbonne, France, has further augmented this research journey. It has facilitated collaboration with researchers there and reinforced my practical knowledge and expertise in anaerobic digestion, especially in ADM1. This hands-on experience has provided invaluable insights and has greatly contributed to the depth of this research.

In conclusion, this research provides a comprehensive understanding of the techniques for parameter identification in biogas production models its control mechanisms. This knowledge is crucial for optimizing biogas production and developing more efficient and sustainable energy solutions. Future work should continue to explore these areas and address the challenges identified in this research gap. The findings of this research have the potential to significantly contribute to the field of biogas technology and pave the way for more sustainable energy solutions. This thesis, therefore, represents a significant step forward in the quest for sustainable energy solutions. It is hoped that this work will inspire further research in this field and contribute to a more sustainable

future.

From perspectives, we aim to explore and compare new techniques for parameter identification, such as the Grey Wolf Optimizer and Whale Optimization Algorithm, among others. Our goal is to continually refine our approach, enhancing the efficiency and effectiveness of biogas production. This research not only contributes to parameter identification but also opens up new avenues for the application of advanced control techniques like Takagi–Sugeno Fuzzy Inference System (FIS) in more complex AD systems. Furthermore, we are excited about the prospect of integrating these parameter identification techniques with system controllers. This integration could potentially lead to more efficient and precise control mechanisms, thereby optimizing the biogas production process.

Bibliography

[1] D. Kashyap, K. Dadhich, S. Sharma, Biomethanation under psychrophilic conditions: a review, *Bioresource Technology* 87 (2) (2003) 147–153. doi:https://doi.org/10.1016/S0960-8524(02)00205-5.

[2] S. Ghosh, Anaerobic digestion for renewable energy and environmental restoration, in: *The 8th International Conference on Anaerobic Digestion, Sendai International Center, Sendai, Japan, Ministry of Education Japan*, 1997.

[3] D. G. L. Rene, I. Marcel Albert, System for producing a pulsating circulation in apparatus containing gas evolving products, *US Patent 2,174,203* (Sep. 26 1939).

[4] M. Khan, H. H. Ngo, W. Guo, Y. Liu, L. D. Nghiem, F. I. Hai, L. Deng, J. Wang, Y. Wu, Optimization of process parameters for production of volatile fatty acid, biohydrogen and methane from anaerobic digestion, *Bioresource technology* 219 (2016) 738–748.

[5] N. Noykova, T. G. Mller, M. Gyllenberg, J. Timmer, Quantitative analyses of anaerobic wastewater treatment processes: Identifiability and parameter estimation, *Biotechnology and Bioengineering* 78 (2002) 89–103. doi:10.1002/bit.10179.

[6] K. R. Godfrey, J. J. DiStefano, Identifiability of model parameters., *IFAC Proceedings Series* 18 (1985) 89–114. doi:10.1016/s1474-6670(17)60544-5.

[7] E. G. Aklilu, Y. A. Waday, Optimizing the process parameters to maximize biogas yield from anaerobic co-digestion of alkali-treated corn stover and poultry manure using artificial neural network and response surface methodology, *Biomass Conversion and Biorefinery* 13 (14) (2023) 12527–12540.

[8] A. Donoso-Bravo, C. Retamal, M. Carballa, G. Ruiz-Filippi, R. Chamy, Influence of temperature on the hydrolysis, acidogenesis and methanogenesis in mesophilic anaerobic digestion: parameter identification and modeling application., *Water science and technology* 60 (2009) 9–17. doi:10.2166/wst.2009.316.

[9] J. Shi, Z. Wang, J. A. Stiverson, Z. Yu, Y. Li, Reactor performance and microbial community dynamics during solid-state anaerobic digestion of corn stover at mesophilic and thermophilic conditions, *Bioresource Technology* 136 (2013) 574–581.

[10] R. A. Labatut, J. L. Pronto, Chapter 4 - sustainable waste-to-energy technologies: Anaerobic digestion, in: *Thomas A. Trabold and Callie W. Babbitt* (Ed.), *Sustainable Food Waste-To-energy Systems*, *Academic Press*, 2018, pp. 47–67. doi:https://doi.org/10.1016/B978-0-12-811157-4.00004-8.

[11] B. Demirel, O. Yenigün, Two-phase anaerobic digestion processes: a review, *Journal of Chemical Technology & Biotechnology* 77 (7) (2002) 743–755. doi:https://doi.org/10.1002/jctb.630.

[12] E. Kroeker, D. Schulte, A. Sparling, H. Lapp, Anaerobic treatment process stability, *Journal of Water Pollution Control Federation* (1979) 718–727.

[13] L. De Baere, M. Devocht, P. Van Assche, W. Verstraete, Influence of high nacl and nh4cl salt levels on methanogenic associations, *Water research* 18 (5) (1984) 543–548.

[14] J. Swanwick, D. Shurben, S. Jackson, A survey of the performance of sewage sludge digesters in great britain, *Journal of Water Pollution Control Federation* 68 (1969) 639–653.

[15] D. J. Batstone, J. Keller, I. Angelidaki, S. V. Kalyuzhnyi, S. G. Pavlostathis, A. Rozzi, W. T. M. Sanders, H. Siegrist, V. A. Vavilin, The iwa anaerobic digestion model no 1 (adm1)., *Water science and technology* 45 (2002) 65–73.

[16] E. E. Petersen, Chemical reaction analysis, *Prentice-Hall*, 1965.

[17] C. Rosen, D. Vrecko, K. V. Gernaey, M. N. Pons, U. Jeppsson, Implementing adm1 for plant-wide benchmark simulations in matlab/simulink, *Water Science and Technology* 54 (2006) 11–19. doi:10.2166/wst.2006.521.

[18] M. de Gracia, L. Sancho, J. L. García-Heras, P. Vanrolleghem, E. Ayesa, Mass and charge conservation check in dynamic models: application to the new adm1 model, *Water Science and Technology* 53 (2006) 225–240. doi:10.2166/wst.2006.025.

[19] D. J. Batstone, J. Keller, Industrial applications of the iwa anaerobic digestion model no. 1 (adm1), *Water Science and Technology* 47 (2003) 199–206. doi:10.2166/wst.2003.0647.

[20] D. J. Batstone, P. F. Pind, I. Angelidaki, Kinetics of thermophilic, anaerobic oxidation of straight and branched chain butyrate and valerate, *Biotechnology and Bioengineering* 84 (2) (2003) 195–204.

[21] V. Fedorovich, P. Lens, S. Kalyuzhnyi, Extension of enaerobic digestion model no. 1 with processes of sulfate reduction, *Applied biochemistry and biotechnology* 109 (2003) 33–45.

[22] U. Zaher, R. Li, U. Jeppsson, J.-P. Steyer, S. Chen, Giscod: general integrated solid waste co-digestion model, *Water research* 43 (10) (2009) 2717–2727.

[23] I. Ramirez, A. Mottet, H. Carrère, S. Déléris, F. Vedrenne, J.-P. Steyer, Modified adm1 disintegration/hydrolysis structures for modeling batch thermophilic anaerobic digestion of thermally pretreated waste activated sludge, *Water research* 43 (14) (2009) 3479–3492.

[24] J. Palatsi, J. Illa, F. X. Prenafeta-Boldú, M. Laureni, B. Fernandez, I. Angelidaki, X. Flotats, Long-chain fatty acids inhibition and adaptation process in anaerobic thermophilic digestion: batch tests, microbial community structure and mathematical modelling, *Bioresource technology* 101 (7) (2010) 2243–2251.

[25] F. Mairet, O. Bernard, M. Ras, L. Lardon, J.-P. Steyer, Modeling anaerobic digestion of microalgae using adm1, *Bioresource Technology* 102 (13) (2011) 6823–6829.

[26] E. L. Barrera, H. Spanjers, K. Solon, Y. Amerlinck, I. Nopens, J. Dewulf, Modeling the anaerobic digestion of cane-molasses vinasse: Extension of the anaerobic digestion model no. 1 (adm1) with sulfate reduction for a very high strength and sulfate rich wastewater, *Water research* 71 (2015) 42–54.

[27] X. Flores-Alsina, K. Solon, C. K. Mbamba, S. Tait, K. V. Gernaey, U. Jeppsson, D. J. Batstone, Modelling phosphorus (p), sulfur (s) and iron (fe) interactions for dynamic simulations of anaerobic digestion processes, *Water Research* 95 (2016) 370–382.

[28] H. Sun, Z. Yang, G. Shi, S. G. Arhin, V. G. Papadakis, M. A. Goula, L. Zhou, Y. Zhang, G. Liu, W. Wang, Methane production from acetate, formate and h2/co2 under high ammonia level: modified adm1 simulation and microbial characterization, *Science of the Total Environment* 783 (2021) 147581.

[29] H. Sun, Z. Yang, Q. Zhao, M. Kurbonova, R. Zhang, G. Liu, W. Wang, Modification and extension of anaerobic digestion model no. 1 (adm1) for syngas biomethanation simulation: From lab-scale to pilot-scale, *Chemical Engineering Journal* 403 (2021) 126177.

[30] Y. Huang, Y. Ma, J. Wan, Y. Wang, Mathematical modelling of the internal circulation anaerobic reactor by anaerobic digestion model no. 1, simultaneously combined with hydrodynamics, *Scientific Reports* 9 (1) (2019) 6249.

[31] M. E. Ersahin, Modeling the dynamic performance of full-scale anaerobic primary sludge digester using anaerobic digestion model no. 1 (adm1), *Bioprocess and biosystems engineering* 41 (2018) 1539–1545.

[32] R. K. Dereli, M. E. Ersahin, H. Ozgun, I. Ozturk, A. F. Aydin, Applicability of anaerobic digestion model no. 1 (adm1) for a specific industrial wastewater: Opium alkaloid effluents, *Chemical Engineering Journal* 165 (1) (2010) 89–94.

[33] U. Zaher, J. Rodríguez, A. Franco, P. Vanrolleghem, Application of the iwa adm1 model to simulate anaerobic digester dynamics using a concise set of practical measurements, in: *IWA Conference on Env. Biol.–Advancement on Water and Wastewater Applications in the Tropics. Kuala Lumpur, Malaysia*, 2003.

[34] D. Poggio, M. Walker, W. Nimmo, L. Ma, M. Pourkashanian, Modelling the anaerobic digestion of solid organic waste–substrate characterisation method for adm1 using a combined biochemical and kinetic parameter estimation approach, *Waste management* 53 (2016) 40–54.

[35] C. García-Diéguez, F. Molina, E. Roca, Multi-objective cascade controller for an anaerobic digester, *Process Biochemistry* 46 (4) (2011) 900–909.

[36] B. K. Penumathsa, G. C. Premier, G. Kyazze, R. Dinsdale, A. J. Guwy, S. Esteves, J. Rodríguez, Adm1 can be applied to continuous bio-hydrogen production using a variable stoichiometry approach, *Water research* 42 (16) (2008) 4379–4385.

[37] E. Nordlander, E. Thorin, J. Yan, Investigating the possibility of applying an adm1 based model to a full-scale co-digestion plant, *Biochemical engineering journal* 120 (2017) 73–83.

[38] O. Bernard, Z. Hadj-Sadok, D. Dochain, A. Genovesi, J. P. Steyer, Dynamical model development and parameter identification for an anaerobic wastewater treatment process., *Biotechnology and bioengineering* 75 (2001) 424–438. doi:10.1002/bit.10036.

[39] A. Chaabna, S. Semcheddine, Improvement of biogas production in anaerobic digestion process, *Ecological Engineering and Environment Protection* 2021 (2021) 26–31. doi:10.32006/eeep.2021.2.2631.

[40] S. Hassam, E. Ficara, A. Leva, J. Harmand, A generic and systematic procedure to derive a simplified model from the anaerobic digestion model no. 1 (adm1), *Biochemical Engineering Journal* 99 (2015) 193–203. doi:10.1016/j.bej.2015.03.007.

[41] E. Walter, Y. Lecourtier, Global approaches to identifiability testing for linear and nonlinear state space models, *Mathematics and Computers in Simulation* 24 (6) (1982) 472–482.

[42] S. Aborhey, D. Williamson, State and parameter estimation of microbial growth processes, *Automatica* 14 (5) (1978) 493–498.

[43] A. Holmberg, On the practical identifiability of microbial growth models incorporating michaelis-menten type nonlinearities, *Mathematical Biosciences* 62 (1982) 23–43. doi:https://doi.org/10.1016/0025-5564(82)90061-X.

[44] D. Dochain, Automatic control of bioprocesses, *1st* Edition, *John Wiley & Sons*, 2013.

[45] K. J. AAstrom, T. Hagglund, The future of pid control, *Control engineering practice* 9 (11) (2001) 1163–1175.

[46] M. Heredia-Molinero, J. Sanchez-Prieto, J. Briongos, M. Palancar, Feedback pid-like fuzzy controller for ph regulatory control near the equivalence point, *Journal of Process Control* 24 (7) (2014) 1023–1037.

[47] R. Aguilar, A. Poznyak, R. Martinez-Guerra, R. Maya-Yescas, Temperature control in catalytic cracking reactors via a robust pid controller, *Journal of Process Control* 12 (6) (2002) 695–705.

[48] M. C. Colantonio, A. C. Desages, J. A. Romagnoli, A. Palazoglu, Nonlinear control of a cstr: disturbance rejection using sliding mode control, *Industrial & engineering chemistry research* 34 (7) (1995) 2383–2392.

[49] W. S. Levine, *Control system applications*, *CRC press*, 1999.

[50] K. Anand, A. P. Mittal, B. Kumar, Dual heating of anaerobic digester substrate using variable speed drive and fractional order pid controller tuned with ant colony optimization, *Biomass Conversion and Biorefinery* (2023) 1–12.

[51] J. Harmand, A. Manh, J.-P. Steyer, Identification and disturbance accommodating control of a fluidized bed anaerobic digester, *Bioprocess Engineering* 23 (2000) 177–185.

[52] H. Méndez-Acosta, D. U. Campos-Delgado, R. Femat, V. González-Alvarez, A robust feedforward/feedback control for an anaerobic digester, *Computers & chemical engineering* 29 (7) (2005) 1613–1623.

[53] K. J. Åström, B. Wittenmark, Computer-controlled systems: theory and design, 3rd Edition, Courier Corporation,, 2013.

[54] A. Rincón, C. Erazo, F. Angulo, A robust adaptive controller for an anaerobic digester with saturated input: guarantees for the boundedness and convergence properties, *Journal of Process Control* 22 (9) (2012) 1785–1792.

[55] E. Petre, D. Selişteanu, D. Şendrescu, Adaptive and robust-adaptive control strategies for anaerobic wastewater treatment bioprocesses, *Chemical Engineering Journal* 217 (2013) 363–378. doi:https://doi.org/10.1016/j.cej.2012.11.129.

[56] T. Luo, M. Yang, J. Han, P. Sun, A novel model-based adaptive control strategy for step-feed sbrs dealing with influent fluctuation, *Bioresource technology* 167 (2014) 476–483.

[57] O. Bernard, M. Polit, Z. Hadj-Sadok, M. Pengov, D. Dochain, M. Estaben, P. Labat, Advanced monitoring and control of anaerobic wastewater treatment plants: software sensors and controllers for an anaerobic digester, *Water Science and Technology* 43 (7) (2001) 175–182.

[58] S. V. Emelianov, S. K. Korovin, I. Mamedov, Variable-structure control systems: discrete and digital, *CRC Press, Inc.*, 1995.

[59] V. Utkin, Variable structure systems with sliding modes, *IEEE Transactions on Automatic control* 22 (2) (1977) 212–222.

[60] K. D. Young, V. I. Utkin, U. Ozguner, A control engineer's guide to sliding mode control, *IEEE transactions on control systems technology* 7 (3) (1999) 328–342.

[61] A. Bartoszewicz, O. Kaynak, V. I. Utkin, Special section on sliding-mode control in industrial applications–part i-guest editorial, *IEEE transactions on industrial electronics (1982)* 55 (11) (2008) 3806.

[62] J.-J. E. Slotine, Sliding controller design for non-linear systems, *International Journal of control* 40 (2) (1984) 421–434.

[63] J. Burton, A. S. Zinober, Continuous approximation of variable structure control, *International journal of systems science* 17 (6) (1986) 875–885.

[64] S. Laghrouche, Commande par modes glissants d'ordre supérieur: théorie et applications, Ph.D. thesis, *École doctorale sciences et technologies de l'information et des matériaux (Nantes)* (2004).

[65] V. Utkin, J. Guldner, J. Shi, Sliding mode control in electro-mechanical systems, *2nd Edition* Edition, *CRC press*, 2017. doi:https://doi.org/10.1201/9781420065619.

[66] G. Bartolini, A. Ferrara, A. Levant, E. Usai, On second order sliding mode controllers, *Variable structure systems, sliding mode and nonlinear control* (1999) 329–350.

[67] F. Plestan, Y. Shtessel, V. Bregeault, A. Poznyak, New methodologies for adaptive sliding mode control, *International journal of control* 83 (9) (2010) 1907–1919.

[68] Y. Shtessel, M. Taleb, F. Plestan, A novel adaptive-gain supertwisting sliding mode controller: Methodology and application, *Automatica* 48 (5) (2012) 759–769.

[69] G. Ambrosino, G. Celektano, F. Garofalo, Variable structure model reference adaptive control systems, *International Journal of Control* 39 (6) (1984) 1339–1349.

[70] S. Sharifi, S. Murthy, I. Takács, A. Massoudieh, Probabilistic parameter estimation of activated sludge processes using markov chain monte carlo, *Water Research* 50 (2014) 254–266.

[71] L. Kocsis, A. György, Efficient multi-start strategies for local search algorithms, *Journal of Artificial Intelligence Research* (2009) 705–720.

[72] A. Raue, M. Schilling, J. Bachmann, A. Matteson, M. Schelke, D. Kaschek, S. Hug, C. Kreutz, B. D. Harms, F. J. Theis, et al., Lessons learned from quantitative dynamical modeling in systems biology, *PloS one* 8 (9) (2013) e74335.

[73] A. Zhu, J. Guo, B.-J. Ni, S. Wang, Q. Yang, Y. Peng, A novel protocol for model calibration in biological wastewater treatment, *Scientific reports* 5 (1) (2015) 8493.

[74] M. Asadi, K. McPhedran, Biogas maximization using data-driven modelling with uncertainty analysis and genetic algorithm for municipal wastewater anaerobic digestion, *Journal of Environmental Management* 293 (2021) 112875.

[75] A. F. Villaverde, F. Fröhlich, D. Weindl, J. Hasenauer, J. R. Banga, Benchmarking optimization methods for parameter estimation in large kinetic models, *Bioinformatics* 35 (5) (2019) 830–838.

[76] N. R. Draper, H. Smith, Applied regression analysis, Vol. 326, *John Wiley & Sons*, 1998.

[77] D. W. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, *Journal of the society for Industrial and Applied Mathematics* 11 (2) (1963) 431–441.

[78] R. Fletcher, Practical methods of optimization, *Hoboken, NJ, USA, John Wiley & Sons*, 2000.

[79] P. A. Vanrolleghem, K. J. Keesman, Identification of biodegradation models under model and data uncertainty, *Water Science and Technology* 33 (2) (1996) 91–105.

[80] K. R. Gegenfurtner, Praxis: Brent's algorithm for function minimization, *Behavior Research Methods, Instruments, & Computers* 24 (1992) 560–564.

[81] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Numerical recipes 3rd edition: The art of scientific computing, *Cambridge university press*, 2007.

[82] T. Weise, Global optimization algorithms-theory and application, *Self-Published Thomas Weise* 361 (2009).

[83] D. E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning, 1st Edition, *Addison-Wesley Longman Publishing Co., Inc.*, 1989.

[84] A. Chaabna, S. Semcheddine, Genetic algorithm based identification of biogas production model from wastewater via anaerobic digestion model no. 1, *International Journal of Information Technology* 15 (3) (2023) 1465–1472.

[85] P. Biernacki, S. Steinigeweg, A. Borchert, F. Uhlenhut, Application of anaerobic digestion model no. 1 for describing anaerobic digestion of grass, maize, green weed silage, and industrial glycerine, *Bioresource Technology* 127 (2013) 188–194. doi:https://doi.org/10.1016/j.biortech.2012.09.128.

[86] K. Yoshida, K. Kametani, N. Shimizu, Adaptive identification of anaerobic digestion process for biogas production management systems, *Bioprocess and Biosystems Engineering* 43 (2020) 45–54. doi:10.1007/s00449-019-02203-9.

[87] E. Chorukova, L. Kabaivanova, V. Hubenov, I. Simeonov, O. Roeva, Mathematical model of a thermophilic anaerobic digestion for methane production of wheat straw, *Processes* 10 (2022). doi:10.3390/pr10040742.

[88] Z. Fatolahi, G. Arab, V. Razaviarani, Calibration of the anaerobic digestion model no. 1 for anaerobic digestion of organic fraction of municipal solid waste under mesophilic condition, *Biomass and Bioenergy* 139 (2020) 105661. doi:https://doi.org/10.1016/j.biombioe.2020.105661.

[89] A. Della Bona, G. Ferretti, E. Ficara, F. Malpei, Lft modelling and identification of anaerobic digestion, Control Engineering Practice 36 (2015) 1–11.

[90] R. J. Ashraf, J. D. Nixon, J. Brusey, Using multi-objective optimisation with adm1 and measured data to improve the performance of an existing anaerobic digestion system, *Chemosphere* (2022) 134523doi:https://doi.org/10.1016/j.chemosphere.2022.134523.

[91] D. Barik, S. Murugan, An artificial neural network and genetic algorithm optimized model for biogas production from co-digestion of seed cake of karanja and cattle dung, *Waste and Biomass Valorization* 6 (2015) 1015–1027. doi:10.1007/s12649-015-9392-1.

[92] Z. Michalewicz, GAs: Why Do They Work?, *Genetic Algorithms + Data Structures = Evolution Programs*, *Springer Berlin Heidelberg*, *Berlin, Heidelberg*, 1996, pp. 45–55. doi:10.1007/978-3-662-03315-9_4.

[93] S. Mirjalili, J. Song Dong, A. S. Sadiq, H. Faris, Genetic algorithm: Theory, literature review, and application in image reconstruction, *Nature-Inspired Optimizers: Theories, Literature Reviews and Applications* (2020) 69–85.

[94] B. L. Miller, D. E. Goldberg, et al., Genetic algorithms, tournament selection, and the effects of noise, *Complex systems* 9 (3) (1995) 193–212.

[95] D. E. Goldberg, K. Deb, A comparative analysis of selection schemes used in genetic algorithms, in: *Foundations of genetic algorithms*, Vol. 1, *Elsevier*, 1991, pp. 69–93.

[96] A. E. Eiben, J. E. Smith, Introduction to evolutionary computing, *2nd edition* Edition, *Springer Berlin, Heidelberg*, 2015. doi:https://doi.org/10.1007/978-3-662-44874-8.

[97] M. Srinivas, L. M. Patnaik, Adaptive probabilities of crossover and mutation in genetic algorithms, *IEEE Transactions on Systems, Man, and Cybernetics* 24 (4) (1994) 656–667.

[98] K. Deep, M. Thakur, A new mutation operator for real coded genetic algorithms, *Applied mathematics and Computation* 193 (1) (2007) 211–230.

[99] A. Neubauer, A theoretical analysis of the non-uniform mutation operator for the modified genetic algorithm, in: *Proceedings of 1997 IEEE International Conference on Evolutionary Computation (ICEC'97)*, IEEE, 1997, pp. 93–96.

[100] S. Tsutsui, Y. Fujimoto, et al., Forking genetic algorithm with blocking and shrinking modes (fga)., in: ICGA, 1993, pp. 206–215.

[101] R. Hinterding, Z. Michalewicz, T. C. Peachey, Self-adaptive genetic algorithm for numeric functions, in: *Parallel Problem Solving from Nature—PPSN IV: International Conference on Evolutionary Computation—The 4th International Conference on Parallel Problem Solving from Nature Berlin, Germany, September 22–26, 1996 Proceedings 4*, Springer, 1996, pp. 420–429.

[102] C. W. Ahn, R. S. Ramakrishna, Elitism-based compact genetic algorithms, *IEEE Transactions on Evolutionary Computation* 7 (4) (2003) 367–385.

[103] M. C. Sadino-Riquelme, J. Rivas, D. Jeison, R. E. Hayes, A. Donoso-Bravo, Making sense of parameter estimation and model simulation in bioprocesses, *Biotechnology and bioengineering* 117 (5) (2020) 1357–1366.

[104] A. Nicholls, Confidence limits, error bars and method comparison in molecular modeling. part 2: comparing methods, *Journal of Computer-Aided Molecular Design* 30 (2) (2016) 103–126.

[105] C. T. Goudar, K. Joeris, K. B. Konstantinov, J. M. Piret, Logistic equations effectively model mammalian cell batch and fed-batch kinetics by logically constraining the fit, *Biotechnology progress* 21 (4) (2005) 1109–1118.

[106] H. Shokrkar, S. Ebrahimi, M. Zamani, Enzymatic hydrolysis of microalgal cellulose for bioethanol production, modeling and sensitivity analysis, *Fuel* 228 (2018) 30–38.

[107] C. Abdelhani, S. Samia, Identification of the am2hn model parameters in the context of organic matter recycling, in: *2022 19th International Multi-Conference on Systems, Signals and Devices (SSD)*, 2022, pp. 367–372. doi:10.1109/SSD54932.2022.9955922.

[108] A. Chaabna, S. Semcheddine, M. Oualid, Integral sliding mode control of anaerobic digestion process for improved biogas production, in: *The 1st National Conference on Emergent Technologies in Electrical Engineering (NCETEE'23) At: Setif, Algeria*, 2023.

[109] D. Selicsteanu, E. Petre, V. B. Ruasvan, Sliding mode and adaptive sliding-mode control of a class of nonlinear bioprocesses, *International Journal of Adaptive Control and Signal Processing* 21 (8-9) (2007) 795–822.

[110] A. Chaabna, S. Semcheddine, Genetic algorithms and multiple regression: A statistical comparison for parameter identification of biogas production model via anova test, in: *The 1st National Conference on Emergent Technologies in Electrical Engineering (NCETEE'23) At: Setif, Algeria*, 2023.

[111] A. Chaabna, S. Semcheddine, Improvement of biogas production by smc: application on am2, in: *Seventh international conference on Ecological Engineering and Environment Protection (EEEP'2021) with Youth Scientific Session and MELiSSA Summer University 30 September - 3 October 2021At: Varna, Bulgaria*, 2021.

ملخص: هذه الدراسة تقدم استكشافًا شاملاً لنموذج عملية الهضم اللاهوائي، وهو عنصر حاسم لتطوير قوانين التحكم لتعزيز الإنتاجية. يُتوقع أن يكون نموذج الهضم اللاهوائي رقم ١ (ADM1) من الجمعية الدولية للمياه IWA هو النموذج الرائد في هذا المجال وبالتالي يُستخدم كمنصة افتراضية. التحدي مع هذه النماذج يكمن في ضبط معلماتها غير المؤكدة. لمعالجة هذا، يتم تقديم منهجية تعتمد على خوارزميات الجينات لحساب وضبط معلمات نموذج هضم لاهوائي مبسط رقم ٢ هضم النيتروجين (AM2HN). الدالة الهدف المقترحة يتم تصغيرها على سلمين مختلفين: النطاق الخطي والنطاق اللوغاريتمي. مقارنة المنهجية المقترحة مع التقنيات التقليدية من الأدبيات، التي تم إجراؤها من خلال محاكاة الكمبيوتر، تُثبت الصلة والتحسين الذي يقدمه النهج المقترح. يتم الحصول على البيانات التجريبية من ADM1 بالإضافة إلى ذلك، تم استخدام النموذج المصغر AM2 لصياغة واختبار قانون التحكم القوي SMC من خلال المحاكاة. النتائج تؤكد فعالية الطريقة التحكم المقترحة في هذه الدراسة. فحص متانة التتبع و مقاومة التشويش يبرز الفعالية الكبيرة لهذه الطريقة. تقدم هذه الدراسة بالتالي، مساهمة كبيرة في مجال نمذجة عملية والتحكم.
كلمات مفتاحية: الهضم اللاهوائي ، خوارزمية الجينات ، التحكم في الوضع الانزلاقي SMC ، تحديد المعلمات ، نموذج ADM1 .

Résumé: Cette étude montre une analyse complète de la processus de digestion anaérobie (AD), essentiel pour développer des lois de commande optimisant la productivité du biogas. Le défi est l'identification des paramètres incertains. Une méthodologie basée sur les algorithmes génétiques (GA) est introduite pour ajuster les paramètres d'un modèle simplifié, AM2HN. Le GA minimise la fonction objectif sur deux échelles : linéaire et logarithmique. Cette méthodologie, validée par simulation Matlab/Simulink, montre une amélioration significative par rapport aux techniques traditionnelles. Le modèle réduit AM2 est utilisé pour formuler et tester une technique de commande robuste: Commande par Mode Glissant (SMC). Les résultats confirment l'efficacité de cette méthode. Cette étude contribue de manière significative à la modélisation et au contrôle des processus AD.
Mots-clés: *Digestion anaérobie, AM2HN, Algorithmes génétiques, Commande par Mode Glissant (SMC), Identification des paramètres, ADM1.*

Abstract: This thesis explores the Anaerobic Digestion (AD) process model, focusing on optimizing productivity. The International Water Association's Anaerobic Digestion Model No.1 (ADM1) is a key model in this field. The main challenge is identifying uncertain parameters. To tackle this, a methodology using genetic algorithms (GA) is introduced to fine-tune the parameters of a simplified model, AM2HN. The GA minimizes the proposed objective function on two scales: linear and logarithmic. This methodology, validated through computer simulation, shows significant improvement over traditional techniques. The reduced model AM2 is used to formulate and test a robust control technique (Sliding Mode Control), demonstrating its efficacy and potential. This study contributes significantly to AD process modeling and control.
Keywords: *Anaerobic digestion, AM2HN, Genetic algorithm, Sliding mode control, Parameter Identification, ADM1.*