

**POLITECNICO DI MILANO**  
Facoltà di Ingegneria Industriale e dell'Informazione  
Corso di Laurea Magistrale in Ingegneria dell'Automazione



**IDENTIFICAZIONE PARAMETRICA  
DI MODELLI DI DIGESTIONE  
ANAEROBICA**

Relatore: Prof. Gianni Ferretti  
Correlatore: Ing. Alessandro Della Bona

Autore: Stefano Civardi  
Matricola: 798862

Anno Accademico 2013-2014



*Alessia, Antonella, Claudio, Valerio.*  
*Senza retorica, il vostro prezioso aiuto è ben visibile in quest'opera.*



# Abstract

La produzione di biogas mediante digestione anaerobica, benché sia ormai una tecnologia matura, necessita tuttora di ulteriori approfondimenti, riguardanti la modellistica del processo, l'identificazione dei parametri significativi, e il controllo e l'ottimizzazione dell'impianto.

Nello specifico, quest'opera si concentrerà in particolare sul problema dell'identificazione parametrica applicata a diverse scelte modellistiche.

Innanzitutto, il capitolo 1 presenta che cosa sia la digestione anaerobica, nonché quali vantaggi e svantaggi essa possa offrire. Vengono quindi illustrate le principali fasi del processo in esame.

Il capitolo 2 argomenta sull'impossibilità di ottenere un modello adatto per qualsiasi scopo, presentando quindi le due scelte modellistiche maggiormente studiate e utilizzate: AMOCO, compatto e sintetico, adatto al monitoraggio e al controllo dell'impianto, e di cui è fornita anche una versione alternativa; e ADM1, più complesso ma anche più dettagliato, maggiormente adatto a simulare accuratamente il processo di digestione.

Il capitolo 3 si occupa del problema dell'identificazione dei parametri di modello. In particolare, la sezione 3.2 offre una panoramica dello stato dell'arte per quanto riguarda l'identificazione applicata al modello ADM1, mentre il resto del capitolo si concentra sull'identificazione del modello AMOCO mediante una preliminare trasformazione in forma LFT.

Il capitolo 4 prosegue idealmente il compito del capitolo precedente, con l'applicazione di un nuovo paradigma, detto Two Stages, per l'identificazione del modello ADM1. I test eseguiti su di esso mostrano però come finora non si siano raggiunti i risultati sperati.



# Abstract

Biogas production through anaerobic digestion, although it is already a mature technology, needs further study about process modelling, identification of the main parameters and plant control and optimization.

This work will deal mainly on the issue of parametric identification applied to different model choices.

First of all, chapter 1 presents what anaerobic digestion is, explaining what are the advantages and the drawbacks it offers. The main steps of the process are then shown.

Chapter 2 argues about the impossibility of obtaining one model suitable for every possible aim; therefore it shows the two most studied and exploited model choices. AMOCO, concise and compact, is suitable for plant control and monitoring; another version of this model is provided, too. ADM1 is more complex, but also more detailed; it is suitable for an accurate simulation of the digestion process.

Chapter 3 deals with identification of the model parameters. Section 3.2 notably recalls the latest developments about identification of ADM1 model, while the rest of the chapter focuses on the identification of AMOCO model through a preliminary reduction to LFT form.

Finally, chapter 4 proposes the application of a new paradigm, known as Two Stages, to ADM1 identification. However, tests show that it does not lead to good results yet.





# Indice

<b>1</b>	<b>Digestione</b>	<b>1</b>
1.1	Introduzione . . . . .	1
1.2	Vantaggi e svantaggi della digestione anaerobica . . . . .	3
1.3	Fasi del processo . . . . .	4
<b>2</b>	<b>Modellistica</b>	<b>7</b>
2.1	Generalità . . . . .	7
2.2	Modello AMOCO . . . . .	8
2.3	Modello AMOCO modificato . . . . .	11
2.4	Modello ADM1 . . . . .	12
<b>3</b>	<b>Identificazione</b>	<b>17</b>
3.1	Importanza dell'identificazione . . . . .	17
3.2	Identificazione del modello ADM1: lo stato dell'arte . . . . .	17
3.3	Trasformazione LFT (generalità) . . . . .	19
3.4	Identificazione del processo di digestione anaerobica mediante trasformazione LFT . . . . .	21
<b>4</b>	<b>TwoStages</b>	<b>25</b>
4.1	Il paradigma Two Stages: gli obiettivi . . . . .	25
4.2	Notazione e considerazioni preliminari . . . . .	26
4.3	La prima fase . . . . .	27
4.4	La seconda fase . . . . .	28
4.5	Identificazione Two-Stages applicata alla digestione anaerobica	28
4.6	Primo test: identificazione dei parametri originali . . . . .	30
4.7	Secondo test: identificazione con nuovi valori dei parametri . .	32
4.8	Terzo test: identificazione in presenza di rumore di misura . .	32
4.9	Valutazione dei risultati e possibili sviluppi . . . . .	33
	<b>Bibliografia</b>	<b>36</b>
<b>A</b>	<b>Matrici di Petersen</b>	<b>41</b>

<b>B</b>	<b>Script di Matlab utilizzati</b>	<b>45</b>
B.1	Applicazione del paradigma TS in assenza di rumore . . . . .	45
B.2	Applicazione del paradigma TS in presenza di rumore . . . . .	49

# Elenco delle figure

1.1	Impianto a biomasse per la produzione di biogas . . . . .	1
1.2	Digestore anaerobico a singolo stadio . . . . .	2
1.3	Esempio di processo di digestione anaerobica . . . . .	5
2.1	Cinetiche di Monod e di Haldane . . . . .	10
3.1	Schema LFT . . . . .	20
3.2	Schema LFT con calcolo della derivata dell'errore . . . . .	23
3.3	Confronto fra AMOCO, in condizioni nominali dei parametri, e AMOCO in forma LFT, ma con valori non nominali dei parametri . . . . .	24
3.4	Confronto fra AMOCO, in condizioni nominali dei parametri, e AMOCO in forma LFT, con i valori identificati dei parametri	24
4.1	Sistema SISO . . . . .	26
4.2	Onda a tre gradini . . . . .	29
4.3	Andamento della portata di metano in uscita nel caso $k_{dis}=0.1$	29
4.4	Errore quadratico medio . . . . .	30
4.5	Fitting della rete neurale . . . . .	31
4.6	Test su dati originali . . . . .	32
4.7	Test su nuovi dati . . . . .	33
4.8	Portata di metano con misura di rumore . . . . .	34
4.9	Test in presenza di rumore di misura . . . . .	34
A.1	Matrice di Petersen per le componenti solubili . . . . .	42
A.2	Matrice di Petersen per le componenti particolate . . . . .	43



# Elenco delle tabelle

2.1	Variabili di stato nel modello ADM1 . . . . .	13
2.2	Processi biochimici considerati nel modello ADM1 . . . . .	14
2.3	Altre variabili e parametri del modello ADM1 . . . . .	15
3.1	Risultati dell'identificazione LFT . . . . .	22



# Capitolo 1

## Digestione anaerobica

### 1.1 Introduzione

La produzione di biogas mediante *digestione anaerobica* è una realtà sempre più diffusa nelle nostre campagne<sup>1</sup>.

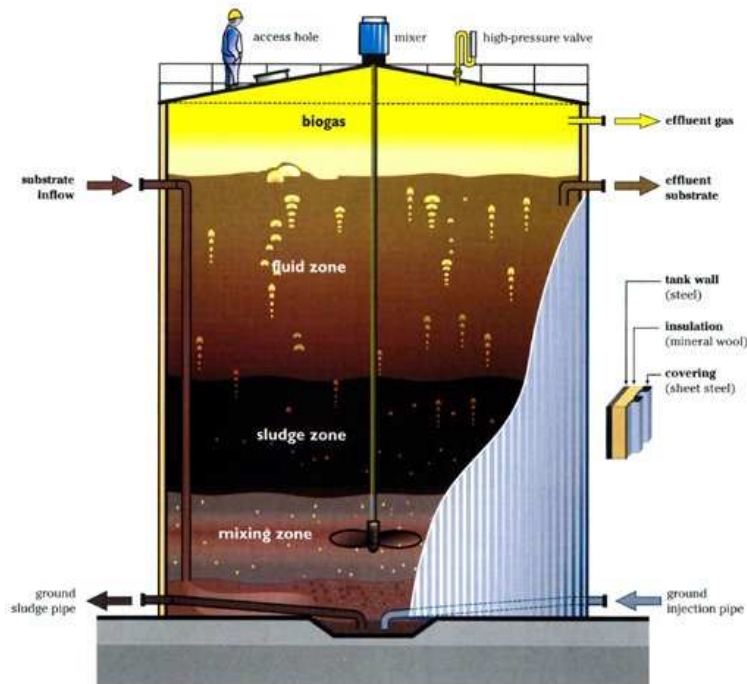


**Figura 1.1:** Impianto a biomasse per la produzione di biogas.

---

<sup>1</sup>Per questo capitolo si faccia riferimento a [1] e [2]. Essi sono un buon riferimento anche per il capitolo 2.

Essa avviene in impianti chiamati *digestori* (o *reattori*, figura 1.2), nei quali viene immesso un *substrato* d'ingresso (scarti agroindustriali, liquami zootecnici, fanghi di depurazione), insieme a dell'acqua, e da cui vengono estratti il *biogas* e il *materiale digestato*. Ciò è possibile grazie all'azione di diverse famiglie batteriche presenti all'interno del digestore.



**Figura 1.2:** Digestore anaerobico a singolo stadio

Questo processo si differenzia dalla digestione aerobica (compostaggio) per l'assenza di ossigeno all'interno delle reazioni coinvolte.

Il *biogas* è una miscela di vari gas:

- metano ( $\text{CH}_4$ ), fra il 50 e il 70%;
- anidride carbonica ( $\text{CO}_2$ ), fra il 25 e il 50%;
- azoto ( $\text{N}_2$ ), fra lo 0 e il 10%;
- idrogeno ( $\text{H}_2$ ), fra lo 0 e l' 1%;
- acido solfidrico ( $\text{H}_2\text{S}$ ), fra lo 0 e il 3%.

Questa ampia variabilità è dovuta a vari fattori, primo fra tutti la diversa composizione della biomassa in ingresso. Ciò che è importante è però il fatto



che questa miscela può essere utilizzata per la produzione di energia elettrica o calore, anche in impianti di teleriscaldamento. Se raffinato, il biogas può diventare *biometano*, ed essere utilizzato come carburante per veicoli a motore o immesso nelle reti di distribuzione gas.

Il *materiale digestato*, invece, può essere utilizzato per il compostaggio (*digestato acidogenico*) o come fertilizzante (*digestato metanogenico*).

## **1.2 Vantaggi e svantaggi della digestione anaerobica**

È evidente che l'utilizzo di un digestore anaerobico può portare diversi benefici:

- produzione di combustibile, utilizzabile come detto per produrre energia o calore. Si noti come si tratta di una fonte completamente rinnovabile, e quindi in linea con l'impegno a dipendere sempre meno dai combustibili fossili.
- smaltimento di rifiuti, tipicamente di origine agricola (sia vegetale che animale) o industriale (settore alimentare). Anche questo è notoriamente un punto critico per il nostro Paese.
- produzione di fertilizzanti di origine organica.
- minor richiesta di nutrienti rispetto alla digestione aerobica, bassa richiesta di energia, alta capacità digestiva.

È interessante notare, in aggiunta a ciò, come l'anidride carbonica emessa dalla combustione del biogas sia proprio la stessa assorbita dai vegetali e dagli animali i cui residui hanno costituito la biomassa. Il processo viene cioè definito "a breve ciclo del carbonio", e al netto il suo apporto di CO<sub>2</sub> nell'atmosfera è quindi nullo.

Fra gli svantaggi possiamo invece citare:

- tentazione di utilizzare suolo agricolo per colture da utilizzarsi direttamente come biomassa (non limitandosi quindi a riutilizzare gli scarti), con conseguente calo della superficie destinata alle colture per uso alimentare. Si badi che si tratta di un'obiezione "politica", più che scientifica; al legislatore e alla comunità spetta giudicare e regolamentare quest'aspetto.
- possibilità di cattivi odori nei pressi dell'impianto.
- tempi più lunghi rispetto alla digestione aerobica.

- complessità di gestione del processo.

L'ultimo aspetto è quello che più interessa ai fini di questo lavoro. La principale causa di complessità è il fatto che il processo in sé è estremamente complicato, con diverse famiglie batteriche che intervengono in momenti diversi della reazione. Per questo, non esiste un modello semplice in grado di rappresentare dettagliatamente tutti i passaggi del processo; bisogna scegliere se utilizzare modelli molto complicati, e quindi poco maneggevoli, o modelli più agili, che però perdono in capacità descrittiva. Questo discorso sarà approfondito nel capitolo 2, dedicato alla modellistica.

Inoltre, anche avendo trovato il modello perfetto, resterebbero da identificare i parametri di questo modello: operazione senz'altro non semplice, come si può apprezzare nei capitoli 3 e 4.

La grande variabilità nella composizione del materiale in ingresso è un altro scoglio nel prevedere l'esito del processo (quantità di biogas prodotto, composizione dello stesso...)

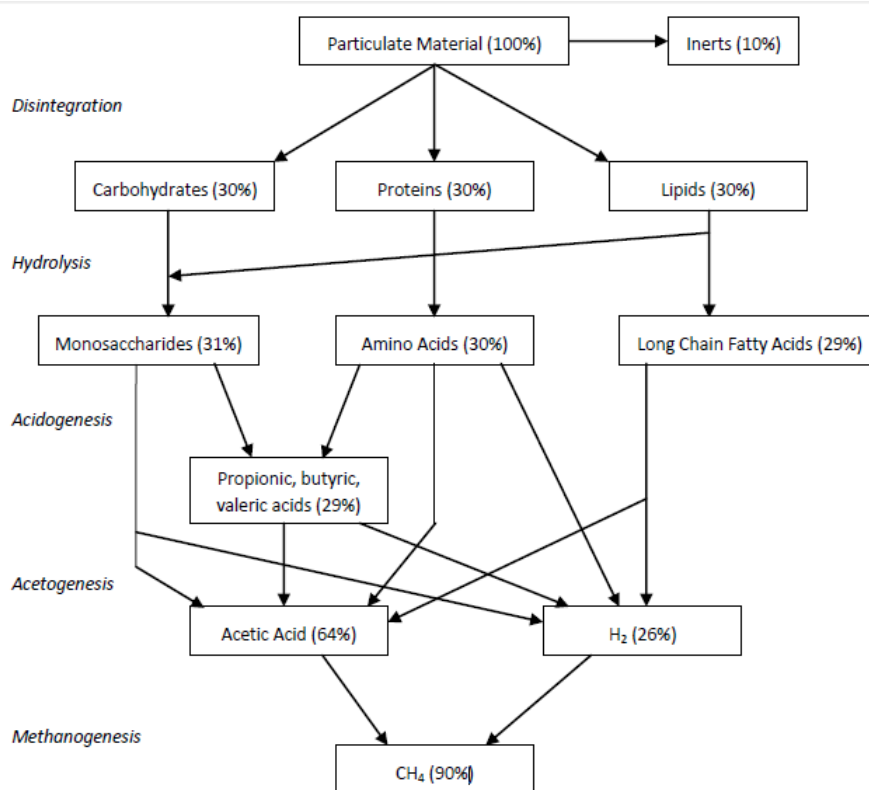
In aggiunta, si consideri che è opportuno adottare una qualche forma di ottimizzazione, in maniera da migliorare la resa del processo. Esso, fra l'altro, richiede particolari condizioni di temperatura e pH per poter avere successo. Infine, ma non meno importante, la stabilità del processo non è garantita: tuttavia, l'insorgere di fenomeni di instabilità va assolutamente evitato, perché porterebbe a lunghi periodi di inattività forzata (nell'ordine dei mesi) e alla formazione di scorie tossiche da smaltire in maniera opportuna.

Il fatto che la digestione anaerobica sia una tecnologia matura non implica quindi ormai una conoscenza perfettamente dettagliata del processo; resta molto lavoro da fare, in tutti i campi: modellistica, identificazione e controllo.

### 1.3 Fasi del processo

Il processo può essere schematicamente suddiviso in quattro fasi: *disintegrazione e idrolisi*, *acidogenesi*, *acetogenesi* e *metanogenesi*. In ogni passaggio, un particolare substrato viene elaborato da un'appropriata famiglia batterica (la *biomassa* vera e propria), andando a generare un nuovo substrato, pronto per essere dato in pasto a una nuova biomassa, e così via, andando a formare la cosiddetta *catena biologica*.

Nella fase di **disintegrazione e idrolisi** il substrato di ingresso è degradato in sostanze più semplici, adatte a essere digerite dalle biomasse: *monosaccaridi*, *amminoacidi* e *acidi grassi a catena lunga (LCFA)*. La durata di questo processo dipende enormemente dalla composizione del materiale inserito nel digestore: sostanze già semplici non hanno nemmeno bisogno di questo stadio iniziale, laddove sostanze particolarmente complesse necessitano di più



**Figura 1.3:** Esempio di processo di digestione anaerobica. I valori numerici sono relativi a un caso particolare, e non vanno intesi come generali; inoltre, alcuni passaggi sono raggruppati per semplicità.

tempo.

Durante la fase di **acidogenesi** il materiale organico viene fermentato, producendo *acidi grassi a catena corta* (o *acidi grassi volatili, VFA*), in particolare acido propionico, butirico, valerico e acetico, oltre ad altre sostanze quali etanolo, metanolo, lattato, idrogeno e anidride carbonica.

I VFA e gli LCFA rimanenti vengono convertiti in acido acetico durante la fase di **acetogenesi**. Ancora una volta, idrogeno e anidride carbonica sono sottoprodotti della reazione.

L'ultima fase della digestione, la **metanogenesi**, porta alla formazione di metano (e CO<sub>2</sub>) a partire dall'acido acetico.

Ancora una volta, si tenga conto che si tratta di una descrizione semplificata di ciò che avviene nell'impianto.

Normalmente - almeno in Italia - il processo è mantenuto in condizioni mesofile (intorno ai 35°C), ritenute un buon compromesso fra consumo di energia per il riscaldamento e velocità di reazione (e quindi resa del proces-

so). Anche il pH va tenuto sotto controllo, tenendo conto che popolazioni batteriche differenti, agenti in fasi diverse, richiedono livelli di acidità differenti: ad esempio, per l'idrolisi è necessario un pH basso (5 - 5,5), mentre per la metanogenesi è richiesto un pH neutro/alto (fra i 6 e gli 8.2).

Per quanto riguarda la velocità delle reazioni, l'acidogenesi e l'acetogenesi rispondono velocemente a fronte di una variazione del proprio influente, mentre la metanogenesi risponde lentamente a una tale evento. La velocità della fase di disintegrazione e idrolisi, come già notato, è invece estremamente variabile. La prima fase funge quindi da collo di bottiglia per l'intero processo. Se il materiale in ingresso è complesso, la reazione è lenta, e gli stadi successivi ricevono il loro influente in maniera ugualmente lenta. Al contrario, un'alta concentrazione di sostanze semplici porta a una rapida fornitura di materiale pronto per l'acidogenesi e acetogenesi. Queste due fasi non hanno problemi a elaborare tali sostanze, producendo un accumulo di sostanza acida che la metanogenesi non riesce a smaltire. Diretta conseguenza di ciò è un calo del pH, che inibisce la metanogenesi stessa; ciò porta a una grave crisi che può compromettere l'intero processo (*instabilità per inibizione da pH*) e va assolutamente evitata. Paradossalmente, quindi, il substrato d'ingresso più problematico è quello più semplice da digerire.

## Capitolo 2

# Modellistica

### 2.1 Generalità

Per poter descrivere la realtà, la scienza e l'ingegneria hanno bisogno di rappresentarla in modelli. Essi descrivono l'interazione fra le varie grandezze fisiche (massa, velocità, temperatura, energia, numero di moli di una sostanza...) tramite un certo numero di equazioni, contenenti diversi parametri. Questi modelli possono essere *a scatola bianca*, cioè basati su equazioni già note, descrittivi fenomeni già conosciuti (come, ad esempio, i principi della dinamica); oppure essere costruiti per via matematica in maniera da “interpretare” al meglio i dati misurati (modellistica *a scatola nera*). Comunque venga eseguita, obiettivo della modellistica è realizzare un “buon” modello del processo. Ciò è utile per diversi scopi, quali:

- ottenere una descrizione quantitativa (e non solo qualitativa) del fenomeno, appagando il desiderio umano di conoscere meglio la realtà che ci circonda;
- sfruttare la conoscenza acquisita per spiegare fenomeni simili o derivati da esso;
- avere a disposizione uno strumento utile per prevedere il comportamento del sistema in esame, analiticamente o tramite simulazioni;
- applicare la conoscenza del modello in maniera da migliorarlo (rendendolo più veloce, più efficiente, più sicuro...)

Ovviamente, cosa si intenda per modello “buono” dipende dal caso in esame: si vuole descrivere una situazione la più generale possibile, o ci si sta concentrando su un caso particolarmente specifico? Si vuole ottenere la massima accuratezza, descrivendo ciò che avviene “per davvero”, o si punta a una maggiore semplicità ed efficienza di calcolo?

Nel caso della digestione anaerobica, in particolare, gli obiettivi principali che ci spingono a cercare un modello sono:

- studiare il comportamento del sistema tramite simulazioni, così da risparmiare tempo (ricordiamo i lunghissimi tempi del processo in esame), risparmiare denaro, e senza correre il rischio di danneggiare l'impianto o i batteri ivi contenuti.
- il controllo, il monitoraggio e l'ottimizzazione dell'impianto, operazioni impossibili da eseguire senza un modello.

In letteratura si possono trovare diversi modelli del processo di digestione anaerobica. Essi infatti devono trovare un compromesso fra due esigenze opposte: descrivere il comportamento del sistema in maniera accurata; e allo stesso tempo fornire uno strumento effettivamente utilizzabile nella pratica. Modelli “troppo” complessi, difatti, pur descrivendo in maniera approfondita ciò che avviene all'interno del digestore, da un lato richiedono importanti campagne di identificazione dei parametri prima di poter essere utilizzati, dall'altro non sono adatti né per il progetto di un controllore, né per il monitoraggio in tempo reale. Al contrario, modelli “troppo” semplificati risultano molto più maneggevoli, ma perdono in capacità espressiva.

In questo studio verranno presentati due modelli diversi: l'*Anaerobic Digestion Model No.1 (ADM1)*, considerato un modello complesso, e l'*Advanced MONitoring and COntrol System for anaerobic processes (AMOCO)*, che invece punta a un maggiore compromesso fra semplicità e accuratezza (ma con uno sbilanciamento a favore della prima).

Entrambi i modelli sono a parametri concentrati; ciò significa che sono validi solo se è mantenuta una opportuna miscelazione all'interno del reattore. Ciò è possibile grazie a una portata di ricircolo, che però ha il difetto di portare fuori dal digestore una certa quota di biomassa e substrato.

## 2.2 Modello AMOCO

Il modello AMOCO è stato sviluppato in [3]. Esso trascura le fasi di disintegrazione, idrolisi e acetogenesi, considerando solo le popolazioni batteriche responsabili di acidogenesi e metanogenesi. Nonostante ciò, è in grado di rappresentare la destabilizzazione da eccesso di acidi volatili, che come abbiamo detto è una delle principali criticità del processo.

In dettaglio, il modello è rappresentato dalle seguenti equazioni:

$$\begin{aligned}
 \frac{dX_1}{dt} &= -\alpha DX_1 + \rho_1 \\
 \frac{dX_2}{dt} &= -\alpha DX_2 + \rho_2 \\
 \frac{dS_1}{dt} &= D(S_{1,in} - S_1) - k_1\rho_1 \\
 \frac{dS_2}{dt} &= D(S_{2,in} - S_2) + k_2\rho_1 - k_3\rho_2
 \end{aligned}
 \tag{2.1}$$

$X_1$  e  $X_2$  rappresentano le *popolazioni di batteri*, acidogeni e metanigeni, mentre  $S_1$  e  $S_2$  rappresentano il *substrato organico* e i *VFA*.  $S_{1,in}$  e  $S_{2,in}$  sono le *quantità in ingresso* di queste variabili. Popolazioni batteriche e substrati sono misurati in  $[KgCODm^{-3}]$ , dove COD sta per *Chemical Oxygen Demand*: essa rappresenta la misura della quantità di ossigeno richiesta per la completa ossidazione dei campioni organici e inorganici presenti in un campione di liquido. Si tratta quindi di una misura indiretta della grandezza d'interesse per il nostro caso.

$\alpha$  rappresenta la frazione di biomassa in fase liquida che può essere trasportata dalla portata di ricircolo: può variare da 0 (reattore a letto fisso ideale) a 1 (reattore CSTR ideale).

$D = Q/V, [d^{-1}]$  è il *rapporto di diluizione*, pari al rapporto fra la *portata di ricambio*  $Q [m^3d^{-1}]$  e un *volume di controllo*  $V [m^3]$ . Il suo inverso è pari al *tempo di ritenzione idraulica*  $t_{HR} [d]$ , che può essere inteso come tempo di ricambio del reattore.

$\rho_1$  e  $\rho_2$  sono i *ratei di produzione/consumo* delle sostanze:

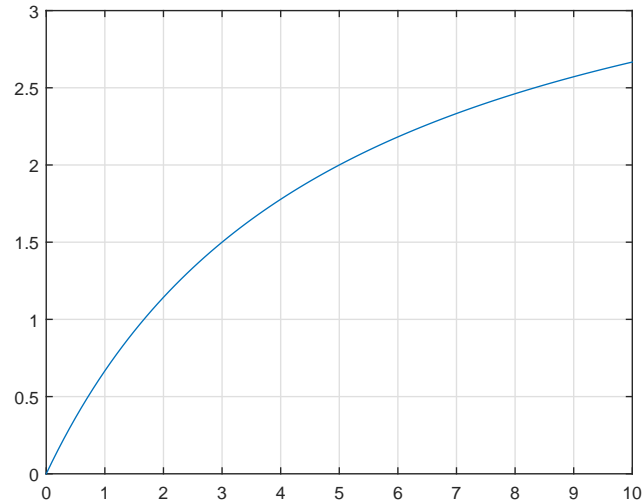
$$\rho_1 = \mu_{1max} X_1 \frac{S_1}{S_1 + K_{S1}} \quad \rho_2 = \mu_{2max} X_2 \frac{S_2}{S_2 + K_{S2} + \frac{S_2^2}{K_{I2}}} \quad (2.2)$$

Il primo è modellato secondo una *cinetica di Monod*, dove  $\mu_{1max} [d^{-1}]$  è il *massimo tasso di crescita specifico*, e  $K_{S1} [dm^{-3}]$  è la *costante di semisaturazione*. Il secondo rateo è invece basato su una *cinetica di Haldane*, in cui interviene un nuovo parametro,  $K_{I2}$ , chiamato *costante d'inibizione*.

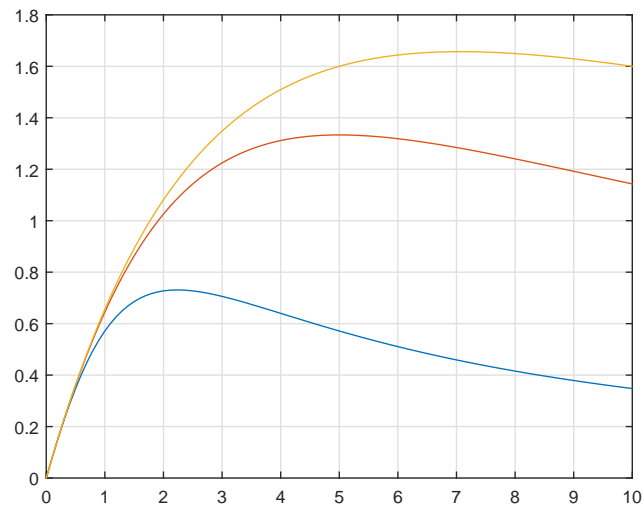
Oltre alle equazioni di bilancio della massa sopra esaminate (2.1), possiamo considerare altre equazioni:

$$\begin{aligned} \frac{dZ}{dt} &= D(Z_{in} - Z) & (2.3) \\ \frac{dC}{dt} &= D(C_{in} - C) - q_C + k_4\rho_1 + k_5\rho_2 \\ q_C &= k_L a [C + S_2 - Z - K_H P_C] \\ P_C &= \frac{\phi - \sqrt{\phi^2 - 4K_H P_T (C + S_2 - Z)}}{2K_H} \\ \phi &= C + S_2 - Z + K_H P_T + \frac{k_6}{k_L a} \rho_2 \\ q_M &= k_6 \rho_2 \\ pH &= -\log \left( K_b \frac{C - Z + S_2}{Z - S_2} \right) \end{aligned}$$

Esse modellano l'*alcalinità* ( $Z$ ), la *concentrazione di  $CO_2$*  ( $C$ ), la *portata di  $CO_2$*  ( $q_C$ ), la *portata di metano* ( $q_M$ ), la *pressione parziale di  $CO_2$*  ( $P_C$ ) e l'*acidità* ( $pH$ ). Si noti che le equazioni (2.3) dipendono dalle (2.1), ma non le influenzano. In particolare, per lo studio della sola catena biologica ci si



(a) Cinetica di Monod (l'ascissa è  $S_1$ ) con  $\mu = 4$  e  $K_{S1} = 5$



(b) Cinetica di Haldane (l'ascissa è  $S_2$ ) con  $\mu = 4$ ,  $K_{S2} = 5$ ,  $K_I$  che assume i valori di 1 (in blu), 5 (rosso) e 10 (arancione).

**Figura 2.1:** Esempi di cinetica di Monod e di Haldane. Si noti come, nella prima figura, la costante di semisaturazione rappresenta il valore dell'ascissa a cui corrisponde un valore di ordinata pari a metà di  $\mu$ , a cui la funzione tende asintoticamente. Nel secondo caso, invece, la funzione arriva a un picco, che cresce al crescere di  $K_I$  (vale  $\sqrt{K_I K_{S2}}$ ).



può concentrare unicamente sulle (2.1).

## 2.3 Modello AMOCO modificato

Il modello AMOCO è stato realizzato avendo come obiettivi principali il monitoraggio e il controllo, ma non un'accurata simulazione del comportamento del processo. In particolare, l'eliminazione della fase di disintegrazione e idrolisi presuppone che il substrato di ingresso abbia un contenuto trascurabile di particolato; ciò è giustificabile se pensiamo all'obiettivo di prevenire la destabilizzazione da eccesso di acidi volatili (problema che insorge se la fase di idrolisi è troppo veloce), ma non permette di prevedere accuratamente l'evolversi delle variabili in un caso più generale.

In [4] sono state proposte tre modifiche al modello:

1. La *fase di disintegrazione e idrolisi* è presa in considerazione grazie all'aggiunta di una nuova variabile  $X_0$ , che rappresenta il materiale organico particolato, e della sua equazione di stato.
2. Si è tenuto conto del *decadimento nel tasso di crescita* delle biomasse, tramite i parametri  $k_{d1}$  e  $k_{d2}$ .
3. Il calcolo dell'alcalinità è stato modificato, in maniera da non trascurare il *contributo dell'azoto inorganico*.

Il modello risultante è quindi descritto dalle seguenti equazioni:

$$\begin{aligned}
 \frac{dX_0}{dt} &= D(X_{0,in} - X_0) - \rho_0 & (2.4) \\
 \frac{dX_1}{dt} &= -\alpha DX_1 + \rho_1 - k_{d1}\mu_{1,max}X_1 \\
 \frac{dX_2}{dt} &= -\alpha DX_2 + \rho_2 - k_{d2}\mu_{2,max}X_2 \\
 \frac{dS_1}{dt} &= D(S_{1,in} - S_1) + k_0\rho_0 - k_1\rho_1 \\
 \frac{dS_2}{dt} &= D(S_{2,in} - S_2) + k_2\rho_1 - k_3\rho_2 \\
 \rho_0 &= \mu_0 X_0 \\
 \rho_1 &= \mu_{1max} X_1 \frac{S_1}{S_1 + K_{S1}} \\
 \rho_2 &= \mu_{2max} X_2 \frac{S_2}{S_2 + K_{S2} + \frac{S_2^2}{K_{I2}}}
 \end{aligned}$$

$$\begin{aligned}
\frac{dZ}{dt} &= D(Z_{in} - Z) + (k_1 N_{S1} - N_{bac})\rho_1 \\
&\quad - N_{bac}\rho_2 + k_{d1}N_{bac}\mu_{1,max}X_1 \\
&\quad + k_{d2}N_{bac}\mu_{2,max}X_2 \\
\frac{dC}{dt} &= D(C_{in} - C) - q_C + k_4\rho_1 + k_5\rho_2 \\
q_C &= k_L a [C + S_2 - Z - K_H P_C] \\
P_C &= \frac{\phi - \sqrt{\phi^2 - 4K_H P_T (C + S_2 - Z)}}{2K_H} \\
\phi &= C + S_2 - Z + K_H P_T + \frac{k_6}{k_L a} \rho_2 \\
q_M &= k_6 \rho_2 \\
pH &= -\log \left( K_b \frac{C - Z + S_2}{Z - S_2} \right)
\end{aligned}$$

## 2.4 Modello ADM1

ADM1 è un modello proposto in [6], su impulso della IWA (*International Water Association*). Il suo scopo è quello di rappresentare in modo più dettagliato ciò che avviene nel reattore, in maniera da avere uno strumento utile in fase di simulazione del processo.

Si tratta di un modello DAE (Differential Algebraic Equations) con 35 equazioni differenziali e 1 equazione algebrica. Le 35 variabili di stato rappresentano tutte delle concentrazioni, presentate in tabella 2.1, mentre la variabile algebrica è la concentrazione di ione idrogeno  $S_H$ , espressa in  $kmolm^{-3}$ .

Il numero di processi biochimici considerati è ben superiore a quello del modello AMOCO, ed è visibile in tabella 2.2.

Le equazioni di stato corrispondono allora ai bilanci di massa delle variabili di stato (che descrivono i 19 processi biochimici), a 3 processi di trasferimento gas-liquido, e a 6 addizionali processi di dissociazione acido-base. Sono riportate in (2.5) fino a (2.10).

$C_1^l$	$S_{su}$	Zuccheri	$kgCODm^{-3}$
$C_2^l$	$S_{aa}$	Amminoacidi	$kgCODm^{-3}$
$C_3^l$	$S_{fa}$	Acidi grassi a catena lunga (LCFA)	$kgCODm^{-3}$
$C_4^l$	$S_{va}$	Acido valerico	$kgCODm^{-3}$
$C_5^l$	$S_{bu}$	Acido butirrico	$kgCODm^{-3}$
$C_6^l$	$S_{pro}$	Acido propionico	$kgCODm^{-3}$
$C_7^l$	$S_{ac}$	Acido acetico	$kgCODm^{-3}$
$C_8^l$	$S_{H_2}$	Idrogeno disciolto	$kgCODm^{-3}$
$C_9^l$	$S_{CH_4}$	Metano disciolto	$kgCODm^{-3}$
$C_{10}^l$	$S_{IC}$	Carbonio inorganico totale	
$C_{11}^l$	$S_{IN}$	Azoto inorganico totale	
$C_{12}^l$	$S_i$	Inerti solubili	$kgCODm^{-3}$
$C_{13}^l$	$X_c$	Materiale composito	$kgCODm^{-3}$
$C_{14}^l$	$X_{su}$	Carboidrati	$kgCODm^{-3}$
$C_{15}^l$	$X_{pr}$	Proteine	$kgCODm^{-3}$
$C_{16}^l$	$X_{li}$	Lipidi	$kgCODm^{-3}$
$C_{17}^l$	$X_{su}$	Degradatori degli zuccheri	$kgCODm^{-3}$
$C_{18}^l$	$X_{aa}$	Degradatori degli amminoacidi	$kgCODm^{-3}$
$C_{19}^l$	$X_{fa}$	Degradatori degli LCFA	$kgCODm^{-3}$
$C_{20}^l$	$X_{C_4}$	Degradatori del valerato e del butirrato	$kgCODm^{-3}$
$C_{21}^l$	$X_{pro}$	Degradatori del propionato	$kgCODm^{-3}$
$C_{22}^l$	$X_{ac}$	Degradatori dell'acetato	$kgCODm^{-3}$
$C_{23}^l$	$X_{H_2}$	Degradatori dell'idrogeno	$kgCODm^{-3}$
$C_{24}^l$	$X_i$	Inerti particolati	$kgCODm^{-3}$
$C_{25}^l$	$S_{cat}$	Cationi	
$C_{26}^l$	$S_{an}$	Anioni	
$C_1^g$	$S_{CO_2}$	Anidride carbonica	$kmolm^{-3}$
$C_2^g$	$S_{CH_4}$	Metano	$kmolm^{-3}$
$C_3^g$	$S_{H_2}$	Idrogeno	$kmolm^{-3}$
$C_{27}^l$	$S_{vam}$	Valerato	$kmolm^{-3}$
$C_{28}^l$	$S_{bum}$	Butirrato	$kmolm^{-3}$
$C_{29}^l$	$S_{prom}$	Propionato	$kmolm^{-3}$
$C_{30}^l$	$S_{acm}$	Acetato	$kmolm^{-3}$
$C_{31}^l$	$S_{HCO_3}$	Bicarbonato	$kmolCm^{-3}$
$C_{32}^l$	$S_{NH_3}$	Ammoniaca libera	$kmolNm^{-3}$

Tabella 2.1: Variabili di stato nel modello ADM1

1	Disintegrazione	11	Assorbimento dell'acetato
2	Idrolisi dei carboidrati	12	Assorbimento dell'idrogeno
3	Idrolisi delle proteine	13	Decadimento di $X_{su}$
4	Idrolisi dei lipidi	14	Decadimento di $X_{aa}$
5	Assorbimento degli zuccheri	15	Decadimento di $X_{fa}$
6	Assorbimento degli amminoacidi	16	Decadimento di $X_{C_4}$
7	Assorbimento degli LCFA	17	Decadimento di $X_{pro}$
8	Assorbimento del valerato	18	Decadimento di $X_{ac}$
9	Assorbimento del butirrato	19	Decadimento di $X_{H_2}$
10	Assorbimento del propionato		

**Tabella 2.2:** Processi biochimici considerati nel modello ADM1

$$\frac{dC_i^l}{dt} = \frac{1}{t_{HR}}(C_{i,in}^l - C_i^l) + \sum_{j=1}^{19} \rho_j \nu_{i,j} \quad i = 1, \dots, 9, 12, \dots, 24 \quad (2.5)$$

$$\frac{dC_{10}^l}{dt} = \frac{1}{t_{HR}}(C_{10,in}^l - C_{10}^l) + \sum_{j=1}^{19} \rho_j \nu'_{10,j} - \rho_{T,10} \quad (2.6)$$

$$\frac{dC_{11}^l}{dt} = \frac{1}{t_{HR}}(C_{11,in}^l - C_{11}^l) + \sum_{j=1}^{19} \rho_j \nu'_{11,j} \quad (2.7)$$

$$\frac{dC_i^l}{dt} = \frac{1}{t_{HR}}(C_{i,in}^l - C_i^l) \quad i = 25, 26 \quad (2.8)$$

$$\frac{dC_i^l}{dt} = -\rho_{A,i} \quad i = 27, \dots, 32 \quad (2.9)$$

$$\frac{dC_i^g}{dt} = -\frac{q_g}{V_g} C_i^g + \rho_{T,i} \frac{V_l}{V_g} \quad i = 1, 2, 3 \quad (2.10)$$

Esse rappresentano:

- i bilanci del materiale solubile e particolato (2.5);
- i bilanci di carbonio (2.6) e azoto (2.7) nei vari processi;
- la dinamica di anioni e cationi (2.8), dovuta alla sola diluizione del reattore;
- la dinamica degli acidi dissociati e dell'ammoniaca libera (2.9);
- la dinamica della fase gassosa (2.10).

A seguire, troviamo:

- il bilancio della cariche, unica equazione algebrica del sistema DAE (2.11);
- equazioni addizionali per il calcolo del flusso di biogas in uscita (2.12) e della sua pressione (2.13).

$$S_{cat} + S_{NH_4} - S_{HCO_3} - \frac{S_{acm}}{64} - \frac{S_{prom}}{112} - \frac{S_{bum}}{160} - \frac{S_{vam}}{208} - S_{an} + S_H - \frac{K_w}{S_H} = 0 \quad (2.11)$$

$$q_g = k_p(p_g - p_{atm}) \frac{p_g}{p_a t m} \quad (2.12)$$

$$p_g = p_{H_2} + p_{CH_4} + p_{CO_2} + p_{H_2O} = S_{H_2} \frac{RT_g}{16} + S_{CH_4} \frac{RT_g}{64} + S_{CO_2} RT_g + p_{H_2O} \quad (2.13)$$

Le variabili e i parametri non ancora presentati sono i seguenti (tabella 2.3):

$t_{HR}$	tempo di ritenzione idraulica ( <i>giorni</i> )
$\rho_j$	tasso di reazione per il processo $j$ ( $kgCODm^{-3}d^{-1}$ o $kmolm^{-3}d^{-1}$ )
$\nu_{i,j}$	coefficiente stechiometrico del componente $i$ nel processo $j$
$\nu_{10,j}, \nu_{11,j}$	contenuti di carbonio e azoto
$\rho_{T,10}$	tasso di trasferimento del carbonio alla fase gassosa
$\rho_{A,i}$	tasso di reazione acido-base relativo al componente $i$ ( $kgCODm^{-3}d^{-1}$ )
$V_g$	volume totale di gas nel reattore ( $m^3$ )
$\rho_{T,i}$	tasso di trasferimento alla fase gassosa del componente $i$ ( $kmolm^{-3}d^{-1}$ )
$S_{NH_4}$	concentrazione di ammonio ( $S_{NH_4} = S_{IN} - S_{NH_3}$ )
$K_w$	costante di dissociazione dell'acqua ( $kmol^2L^{-2}$ )
$q_g$	flusso di biogas in uscita ( $m^3d^{-1}$ )
$k_p$	coefficiente del flusso in uscita ( $m^3bar^{-1}d^{-1}$ )
$p_g$	pressione del biogas ( <i>bar</i> )
$p_{atm}$	pressione atmosferica ( <i>bar</i> )
$p_x$	pressione parziale del gas $x$ : $H_2$ , $CH_4$ , $CO_2$ o $H_2O$ ( <i>bar</i> )
$R$	costante dei gas ideali
$T_g$	temperatura del gas ( $K$ )

**Tabella 2.3:** Altre variabili e parametri del modello ADM1

I coefficienti stechiometrici sono riportati nella cosiddetta *matrice di Petersen*, in allegato (A).

Il modello presentato è decisamente più complicato di un modello come l'A-MOCO. Anche in questo caso, però, sono possibili alcune modifiche o semplificazioni a seconda della situazione d'interesse, tipicamente trascurando alcuni termini numericamente non significativi, o semplificando i termini di inibizione [1]. Per maggiori approfondimenti sulla simulazione di ADM1, si veda [7].

## Capitolo 3

# Identificazione: generalità, stato dell'arte e trasformazione LFT

### 3.1 Importanza dell'identificazione

Una volta eseguita la modellistica, il lavoro chiaramente non è finito: bisogna dare un valore numerico a tutti le grandezze che vi compaiono (parametri, condizioni iniziali delle variabili...), altrimenti il modello trovato è inutilizzabile ai fini pratici.

In casi fortunati, questo non è un problema: le grandezze sono direttamente misurabili, oppure si fa uso di costanti universali. In altri casi, ciò non è possibile: le grandezze non sono fisicamente accessibili per una misura, o addirittura non rappresentano nemmeno qualcosa di “fisico”. In tali situazioni, bisogna dedurre il loro valore a partire dai dati sperimentali che si hanno a disposizione.

Nelle prossime sezioni ci si concentrerà sull'identificazione nell'ambito della digestione anaerobica.

### 3.2 Identificazione del modello ADM1: lo stato dell'arte

ADM1 (Anaerobic Digestion Model) è attualmente il modello più usato per descrivere in maniera accurata i fenomeni di digestione anaerobica [8]. Come visto, sono in uso anche modelli alternativi, più semplici o anche più complessi; ma allo stesso tempo si sta cercando di ampliare la nostra conoscenza del modello, sia andando a stimarne parametri e stato iniziale, sia proponendo vere e proprie modifiche alla sua struttura.

A causa della complessità della situazione oggetto di studio, molti autori si

soffermano su casi particolari: in concreto, riducendo l'ingresso al modello (la sostanza da fermentare) a substrati predeterminati, quali letame bovino [8], rifiuti alimentari [9], acque reflue di origine suina [10], funghi [11] e così via; esistono comunque studi dedicati proprio agli effetti dell'imprevedibilità nel substrato di ingresso [12]. Anche il punto di vista può cambiare molto da articolo ad articolo, potendo essere orientato più agli aspetti biologici, biochimici, energetici, o addirittura medici [13]. A titolo di esempio, una delle modifiche proposte al modello consente di tenere meglio in considerazione i fenomeni di inibizione dovuti alla presenza di acidi grassi a catena lunga (LCFA), a cui la popolazione acetoclastica è particolarmente sensibile [14]; oppure, si cerca di valutare l'impatto di un non perfetto mescolamento all'interno del reattore [8]. A questo proposito, va notato come l'ipotesi di omogeneità può portare a errori sperimentali solitamente attribuiti alla stima dei parametri o al modello cinetico, magari invece corretti [ibidem].

Un obiettivo comune [8, e molti altri qui citati] è comunque quello di ottenere stime affidabili del modello senza doversi basare su stime accurate di tutti i parametri in gioco. Le varianti di ADM1 vanno di solito in questa direzione (modello più accurato, ma meno sensibile ad errori di calibrazione). È chiaro che ciò potrebbe portare a modelli sempre più complessi; da qui varianti tese invece a ridurre la complessità del modello [15],[16]. Mentre alcune modifiche hanno obiettivi generici, altre sono più specifiche: ad esempio [11], ottenere buone stime sia della produzione di metano che del pH, ma in una situazione particolare come la presenza di un determinato substrato di ingresso.

Una costante che si ritrova in vari studi [8][17][18] è che i parametri più importanti da identificare accuratamente sono quelli di idrolisi, sia perché questa fase è considerata il "collo di bottiglia" della digestione anaerobica, sia perché essi sono considerati i parametri più sensibili da determinare.

Per quanto riguarda il monitoraggio dello stato del processo, un nuovo studio [19] propone come indicatore una combinazione delle concentrazioni di acetato e propionato e di biogas prodotto, ritenuta in grado di stimare sia le prestazioni del processo, sia i bilanci al suo interno, mentre altri indicatori tradizionalmente utilizzati (pH e altri acidi grassi volatili, come il butirrato) sono indicati come meno affidabili. Va tuttavia specificato che in questo caso l'esperimento era condotto su un digestore con alta capacità (buffering capacity); in altre situazioni (buffering capacity bassa) il pH può ancora essere considerato un utile indicatore. Un altro aspetto interessante di questa metodologia è la facilità di misura (automatica e online) del biogas prodotto e del pH, laddove l'analisi dei VFAs deve ancora essere eseguita manualmente. A questo proposito, si rimanda a [20] per un riassunto delle principali tecniche di monitoraggio.

Per l'analisi del substrato di ingresso, invece, i test più usati [17],[18],[21] sono i BMP (*biochemical methane potential test*); tuttavia, studi recenti [22] suggeriscono un metodo alternativo per la stima della biodegradabilità e del-



la bioaccessibilità. Questo metodo, considerato molto più breve (impiega 5 giorni invece che 1-2 mesi), si basa sull'accoppiamento di estrazioni chimiche sequenziali con la spettroscopia a fluorescenza. Si noti che la conoscenza del substrato di ingresso, nonché delle sue caratteristiche, può avere ricadute ampiamente positive sia per il progetto di nuovi reattori, sia per il loro monitoraggio e controllo. Inoltre, già nel 2008 Batstone notava [17] che, per quanto riguarda l'identificazione dei parametri di idrolisi (che come detto sono i più interessanti da stimare) i test BMP danno buoni risultati solo se si sta svolgendo un'analisi di fattibilità di un progetto, ma non per la modellistica della sua dinamica. Tuttavia, va notato come altri autori [18] riescano ad ottenere buone stime dei parametri di idrolisi proprio grazie ai test BMP. Ai test fisico-chimici si possono poi abbinare procedure numeriche per meglio identificare i vari valori di ingresso per il modello ADM1, come in questo studio di Girault e Steyer [23] in cui si è utilizzata, invece che BMP, la respirometria anaerobica.

Va qui rimarcato che per quanto riguarda il controllo del reattore, invece, normalmente si utilizzano modelli ridotti rispetto all'ADM1 [24][25]. Per la fase di progetto del digestore, però, si può citare anche uno studio che utilizza metodi statistici (metodo Montecarlo) [26].

### 3.3 Trasformazione LFT (generalità)

Molte delle principali tecniche di identificazione sono adatte a casi lineari; ma abbiamo visto come tutti i modelli in esame siano decisamente non lineari. Può essere allora utile riformulare le equazioni del modello in maniera da rendere lineare la loro dipendenza dai parametri da identificare: è quello che si intende fare con la *trasformazione in forma lineare frazionaria*, o *LFT (linear fractional transformation)* [4].

In generale, consideriamo un sistema dinamico tempoinvariante con equazioni di vincolo (sistema *DAE*, *differential algebraic equation*), descritto dalle equazioni:

$$\begin{aligned} F_0(x_0, \dot{x}_0, v_0, u, p_0) &= 0 \\ p_0 &= G_0(p_0) \end{aligned} \tag{3.1}$$

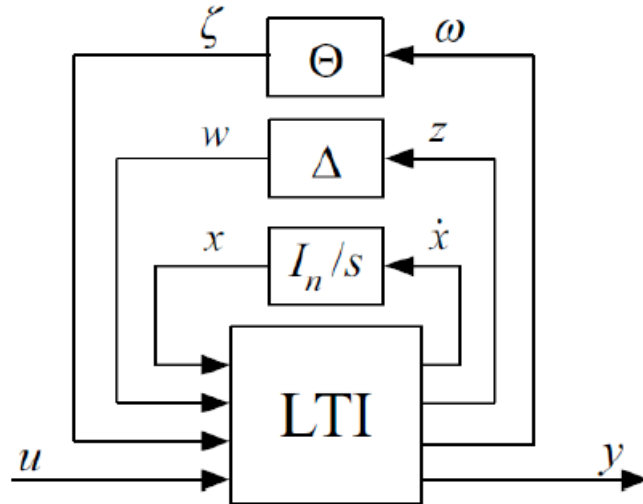
$F_0$  è una funzione vettoriale,  $x_0$ ,  $v_0$ ,  $u$  e  $p_0$  sono vettori rappresentanti le variabili di stato, le variabili algebriche, gli ingressi del sistema e i suoi parametri, e  $G_0$  è una funzione vettoriale che esprime i vincoli dei parametri con valori numerici oppure come funzioni di altri parametri (*binding equations*).

La rappresentazione LFT del sistema (3.1) è allora:

$$\begin{aligned}
 \dot{x} &= Ax + B_1 w + B_2 \zeta + B_3 u & (3.2) \\
 z &= C_1 x + D_{11} w + D_{12} \zeta + D_{13} u \\
 \omega &= C_2 x + D_{21} w + D_{22} \zeta + D_{23} u \\
 y &= C_3 x + D_{31} w + D_{32} \zeta + D_{33} u \\
 w &= \text{diag}\{\delta_1 I_{r_1}, \dots, \delta_q I_{r_q}\} z = \Delta z \\
 \zeta &= \Theta(\omega)
 \end{aligned}$$

che corrisponde alla figura 3.1.

L'idea è di avere in linea di andata un sistema lineare e tempo invariante



**Figura 3.1:** Generica forma LFT di un sistema DAE

(LTI), corrispondente (da un puro punto di vista ingresso-uscita!) alle prime quattro equazioni di (3.2), e legato ai parametri noti o ai loro valori medi. I parametri non noti (o la loro parte da identificare), indicati con  $\delta$ , sono invece normalizzati e inseriti nella matrice  $\Delta$  (si noti che  $I_{r_i}$  rappresenta la matrice identità di ordine  $r_i$ );  $\delta_i$  quindi varierà da -1 (valore minimo del parametro) a +1 (valore massimo), passando per 0 (valore medio).  $\Theta$  contiene funzioni non lineari in parametri noti, mentre la dinamica del sistema è concentrata nel blocco  $I_n/s$ , con  $n$  ordine della parte ODE.

Nonostante le apparenze, questo artificio permette davvero una più facile identificazione dei parametri del sistema. Quello che può risultare difficile è la conversione del sistema in forma LFT; tuttavia la procedura può essere automatizzata mediante impiego di software.

### 3.4 Identificazione del processo di digestione anaerobica mediante trasformazione LFT

In [1] e [4] è proposta l'applicazione della trasformazione LFT per l'identificazione dei parametri del modello AMOCO. L'intenzione è infatti quella di un'identificazione funzionale al controllo del digestore; in particolare, si desiderano cogliere in tempo utile le variazioni nei parametri della fase di idrolisi, in maniera tale da prevenire la destabilizzazione da eccesso di acidi. Visto che nelle applicazioni pratiche non è possibile conoscere in anticipo le sostanze contenute nell'ingresso del digestore, esso è da ritenersi ignoto.

In breve, per l'identificazione si è scelto un metodo output-error, quello della *massima verosimiglianza*. In poche parole, si punta a calcolare il valore dei parametri affinché sia minimizzato l'errore fra il valore misurato dell'uscita e quello simulato.

Più precisamente, si punta a minimizzare la cifra di merito

$$J(\delta) = \frac{1}{2N} \sum_{t=1}^T e(t, \delta)' e(t, \delta) \quad (3.3)$$

dove  $e(t, \delta) = y_{mis}(t, \delta_0) - y(t, \delta)$  è l'errore sopra definito,  $\delta$  è il vettore dei parametri da identificare (il cui valore esatto è  $\delta_0$ ),  $y_{mis}(t, \delta_0)$  è l'uscita misurata (con  $1 \leq t \leq T$ ),  $y(t, \delta)$  è l'uscita simulata, e  $N = 1, 2, \dots, T$  è la lunghezza del periodo d'osservazione. Minimizzando tale cifra di merito, si ottiene  $\hat{\delta}$ , cioè il valore dei parametri per cui la simulazione approssima più accuratamente i risultati sperimentali.

La minimizzazione è ottenuta sfruttando gli strumenti dell'analisi matematica: si approssima la cifra di merito con il proprio sviluppo in serie di Taylor, e se ne calcola il minimo. La trattazione analitica dei passi da compiere non verrà qui riportata; dopotutto, si parla di tecniche di uso comune. Il problema sorge nel momento in cui bisogna calcolare la storia della derivata dell'errore:

$$\frac{\partial e(t, \delta)}{\partial \delta}$$

valutata in  $\delta^{(r)}$  a ogni iterazione ( $r$ ).

Il problema non è teorico, ma puramente computazionale; e la trasformazione in forma LFT ci permette proprio di risolverlo.

La chiave sta nel derivare la forma generale LFT (3.2) rispetto al vettore dei parametri incogniti  $\delta$ , ottenendo:

$$\begin{aligned}
\dot{x}_\delta &= Ax_\delta + B_1w_\delta + B_2\zeta_\delta & (3.4) \\
z_\delta &= C_1x_\delta + D_{11}w_\delta + D_{12}\zeta_\delta \\
\omega_\delta &= C_2x_\delta + D_{21}w_\delta + D_{22}\zeta_\delta \\
y_\delta &= C_3x_\delta + D_{31}w_\delta + D_{32}\zeta_\delta \\
w_\delta &= \text{diag}\{\delta_1 I_{r_1}, \dots, \delta_q I_{r_q}\}z_\delta = \Delta z_\delta \\
\zeta_\delta &= \frac{\partial \Theta}{\partial \delta} = \frac{\partial \Theta}{\partial \omega} \frac{\partial \omega}{\partial \delta} = \frac{\partial \Theta}{\partial \omega} \omega_\delta
\end{aligned}$$

Il pedice  $\delta$  (come in  $x_\delta$ ) rappresenta proprio la derivata rispetto a  $\delta$ . Si noti come si sia sfruttato il fatto che l'ingresso  $u(t)$  non dipende dal vettore dei parametri.

Si ottiene così lo schema in figura 3.2.

Sempre in [1] è presentato un esperimento di test.

Si è costruita una forma LFT del modello AMOCO, e la si è utilizzata per identificare i parametri  $\mu_{1max}$ ,  $K_{S1}$ ,  $\mu_{2max}$  e  $K_{S2}$ . Per farlo, si è sollecitato il sistema (ingresso  $S_{1,in}$ ) con l'onda a 3 gradini visibile in figura 4.2. L'ampiezza è stata scelta tale da non innescare il comportamento di destabilizzazione, mentre la forma d'onda è stata scelta in maniera da sollecitare adeguatamente il sistema e fornire l'informazione necessaria per identificare i 4 parametri incerti del modello [27]. La costante di tempo dominante è il tempo di ritenzione idraulica  $t_{HR}$ . La figura 3.3 mostra un confronto fra le variabili di stato del modello AMOCO, simulato in condizioni nominali, e il modello AMOCO in forma LFT, ma con valori perturbati dei parametri. La figura 3.4, invece mostra gli stessi grafici ma dopo aver proceduto all'identificazione: si vede come ora le due simulazioni coincidano. La tabella 3.1 conferma come l'identificazione sia andata a buon fine.

Parametro		Valore identificato	Valore esatto
$\mu_{1max}$	$[d^{-1}]$	0.2062	0.206
$\mu_{2max}$	$[d^{-1}]$	0.2011	0.2
$K_{S1}$	$[dm^{-3}]$	1.0976	1.096
$K_{S2}$	$[dm^{-3}]$	6.8806	6.86

**Tabella 3.1:** Risultati dell'identificazione LFT

Questo test mostra come effettivamente questa tecnica applicata al modello AMOCO possa dare buoni risultati. Tuttavia, ciò non deve illuderci di aver risolto il problema dell'identificazione una volta per tutte: a questo proposito, nel capitolo 4 verrà messo alla prova un nuovo metodo.

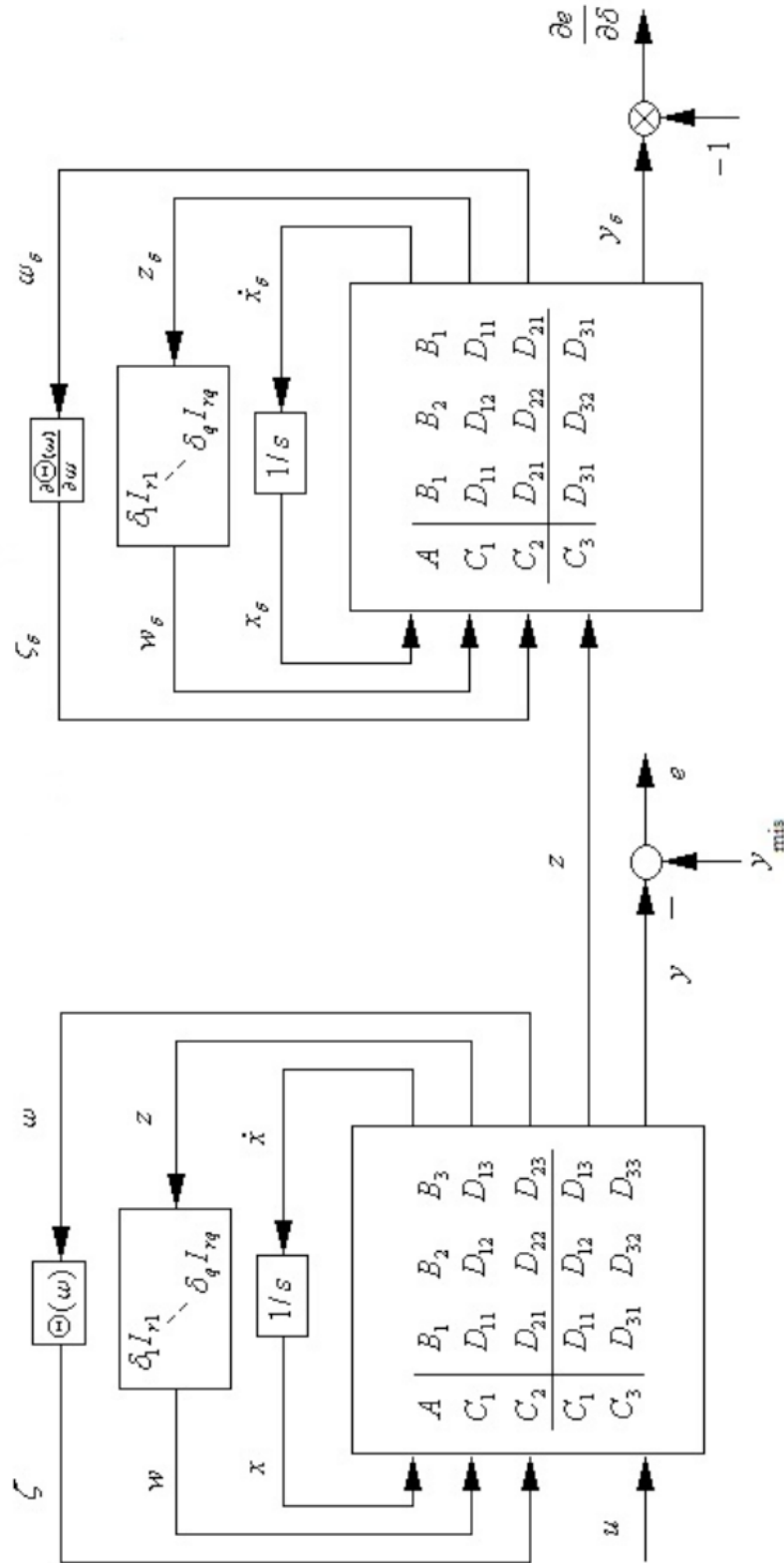
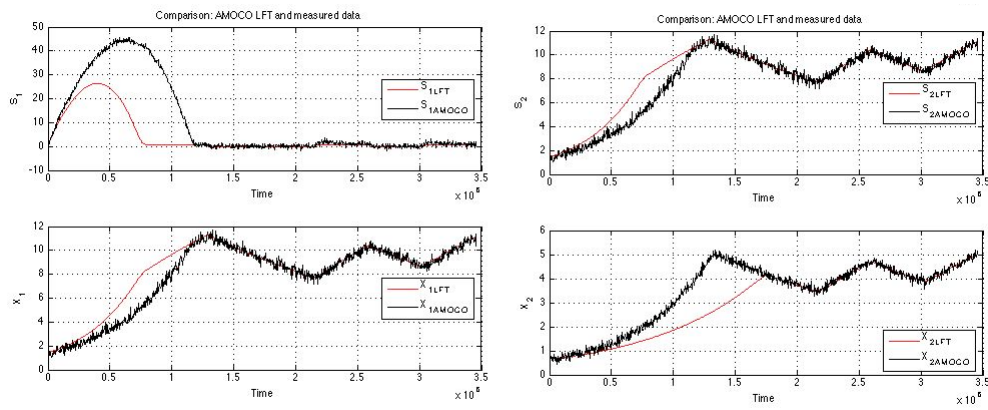
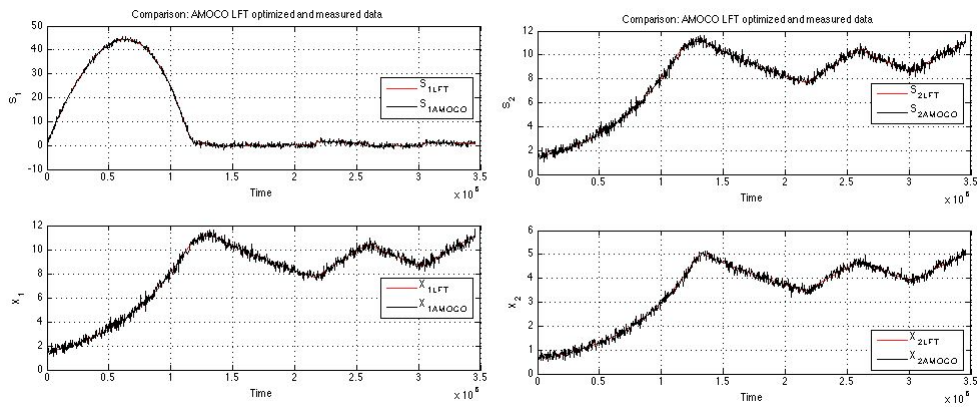


Figura 3.2: Schema di modello LFT, con anche calcolo della derivata dell'errore



**Figura 3.3:** Confronto fra AMOCO, in condizioni nominali dei parametri, e AMOCO in forma LFT, ma con valori non nominali dei parametri.



**Figura 3.4:** Confronto fra AMOCO, in condizioni nominali dei parametri, e AMOCO in forma LFT, con i valori identificati dei parametri.

## Capitolo 4

# Identificazione con paradigma Two Stages

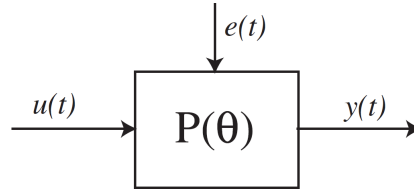
### 4.1 Il paradigma Two Stages: gli obiettivi

Il paradigma Two-Stages (Due Passaggi), proposto in [28], ha come scopo quello di diminuire il carico computazionale da svolgersi in tempo reale, ma senza perdere in accuratezza e senza bisogno di una continua supervisione umana. Si tratta ovviamente di un obiettivo ambizioso, motivato dall'eccessiva complessità di altri metodi di identificazione, che spesso porta a scegliere metodi meno accurati, ma più leggeri. Il costo da pagare per ottenere questo obiettivo è una maggiore complessità iniziale, ovvero nella ricerca dello stimatore del parametro (o vettore di parametri) da identificare; si tratta però di un carico computazionale da svolgersi offline, e una volta soltanto.

I vari metodi di identificazione cercano uno stimatore, cioè una funzione analitica che mappi i dati sperimentali in una stima dei valori dei parametri da identificare. Questa espressione analitica, basata su criteri di stima rispetto a cui è definito lo stimatore, deve essere ricalcolata ogni volta che si desidera una stima, operazione non sempre facile (tipicamente richiede la ricerca di un minimo o di un massimo, come nel caso della massima verosimiglianza). L'idea fondamentale alla base di questo paradigma è invece la rinuncia a ottenere l'espressione analitica di uno stimatore tramite un criterio di stima; esso è invece ottenuto empiricamente a seguito di una estensiva campagna di simulazione del modello in esame. Lo stimatore ottenuto è in grado di descrivere bene il parametro da identificare nelle simulazioni svolte, a patto che esse siano in numero sufficiente a rappresentare in maniera adeguata il comportamento del sistema. L'espressione analitica ottenuta può ora essere utilizzata per mappare in tempo reale il valore del parametro cercato: ma si tratta semplicemente della valutazione in un punto della funzione ottenuta, operazione computazionalmente banale.

## 4.2 Notazione e considerazioni preliminari

Consideriamo un sistema  $P(\theta)$ , dipendente da un vettore di parametri  $\theta \in \mathbb{R}^q$ . Esso riceve un ingresso misurabile  $u(t)$ , un ingresso non misurabile  $e(t)$ , e ha un'uscita misurabile  $y(t)$ . Per semplicità espositiva, supponiamo un sistema SISO ( $e(t)$  e  $u(t)$  reali).



**Figura 4.1:** Il sistema in esame

Abbiamo a disposizione  $N$  campioni ingresso-uscita:

$$\bar{D}^N = \{\bar{y}(1), \bar{u}(1), \bar{y}(2), \bar{u}(2), \dots, \bar{y}(N), \bar{u}(N)\}. \quad (4.1)$$

L'obiettivo è ottenere uno stimatore  $\hat{f} : \mathbb{R}^{2N} \rightarrow \mathbb{R}^q$ , che dia una stima dei parametri

$$\hat{\theta} := \hat{f}(\bar{y}(1), \bar{u}(1), \dots, \bar{y}(N), \bar{u}(N)) \quad (4.2)$$

La situazione presentata finora è perfettamente generale, ed esistono diversi metodi per approssimare il problema. Cosa suggerisce il paradigma TS?

Si selezionano  $m$  valori del parametro  $\theta$ , in maniera da coprire densamente il range di interesse. Il modello è simulato in ognuna di queste  $m$  situazioni, ottenendo il *diagramma dei dati da simulazione*:

$$\begin{aligned} \theta_1 &\rightarrow \bar{D}_1^N = \{y^1(1), u^1(1), \dots, y^1(N), u^1(N)\} \\ \theta_2 &\rightarrow \bar{D}_2^N = \{y^2(1), u^2(1), \dots, y^2(N), u^2(N)\} \\ &\dots \\ \theta_m &\rightarrow \bar{D}_m^N = \{y^m(1), u^m(1), \dots, y^m(N), u^m(N)\} \end{aligned} \quad (4.3)$$

Idealmente, si potrebbe ora ottenere  $\hat{f}$  come la funzione che minimizza l'errore di stima:

$$\hat{f} \leftarrow \min_f \frac{1}{m} \sum_{i=1}^m \|\theta_i - f(y^i(1), u^i(1), \dots, y^i(N), u^i(N))\|^2 \quad (4.4)$$

Invece di definire un criterio di stima, si vede quindi che lo stimatore è ottenuto in maniera da adattarsi alla campagna di simulazioni svolte.



Questo passaggio, così pulito nella teoria, è invece irrealizzabile nella pratica. Un problema di minimizzazione con un tale numero di variabili ed equazioni in gioco non può essere realisticamente eseguito: bisogna trovare il modo di limitare la complessità di quest'operazione. Ecco perché si è scelto di suddividerla in due passaggi - da cui il nome del paradigma.

### 4.3 La prima fase

La prima fase ha lo scopo di ridurre la dimensionalità del problema: invece che  $m$  sequenze di  $2N$  dati, dove  $N$  è tipicamente un numero molto alto, si comprime l'informazione in  $m$  sequenze di  $n$  dati, dove  $n \ll N$ . Otteniamo dunque il *diagramma dei dati compressi*:

$$\begin{aligned}\theta_1 &\rightarrow \tilde{D}_1^n = \{\alpha_1^1, \alpha_2^1, \dots, \alpha_n^1\} \\ \theta_2 &\rightarrow \tilde{D}_2^n = \{\alpha_1^2, \alpha_2^2, \dots, \alpha_n^2\} \\ &\dots \\ \theta_m &\rightarrow \tilde{D}_m^n = \{\alpha_1^m, \alpha_2^m, \dots, \alpha_n^m\}\end{aligned}\quad (4.5)$$

D'accordo, ma come comprimere i dati? Ognuna delle  $m$  sequenze viene rappresentata con un modello ARX (autoregressivo con ingresso esogeno - *Auto Regressive with exogenous input*) di ordine  $n = n_y + n_u$ :

$$\begin{aligned}y^i(t) &= \alpha_1^i y^i(t-1) + \dots + \alpha_{n_y}^i y^i(t-n_y) + \\ &\alpha_{n_y+1}^i u^i(t-1) + \dots + \alpha_n^i u^i(t-n_u)\end{aligned}\quad (4.6)$$

I dati compressi non sono altro che i parametri del modello ARX, calcolabili con il metodo dei minimi quadrati:

$$\begin{aligned}\bar{n} &= \max\{n_y, n_u\} \\ \varphi^i(t) &= [y^i(t-1), \dots, y^i(t-n_y), u^i(t-1), \dots, u^i(t-n_u)]^T \\ [\alpha_1^i, \dots, \alpha_n^i]^T &= \left[ \sum_{t=\bar{n}+1}^N \varphi^i(t) \varphi^i(t)^T \right]^{-1} \cdot \sum_{t=\bar{n}+1}^N \varphi^i(t) y^i(t)\end{aligned}\quad (4.7)$$

Servono ora alcune precisazioni:

- i modelli ARX *non* vanno intesi come modelli del sistema, e non hanno alcuna interpretazione fisica;
- è fondamentale che i modelli ARX catturino la variabilità dei dati, ma non è necessario che li rappresentino in maniera accurata;
- la scelta dell'ordine dell'ARX (cioè il parametro  $m$ ), che dev'essere lo stesso per tutte le sequenze di dati, non è particolarmente critica, e può essere semplicemente frutto del confronto fra più tentativi;

- di fatto, si è ottenuta una *funzione di compressione*  $\hat{g} : \mathbb{R}^{2N} \rightarrow \mathbb{R}^n$  che rimappa le sequenze ottenute in simulazione nelle rispettive sequenze compresse.

#### 4.4 La seconda fase

La seconda fase consiste nella minimizzazione che si intendeva svolgere in (4.4), ma con una complessità notevolmente ridotta:

$$\hat{h} \leftarrow \min_h \frac{1}{m} \sum_{i=1}^m \|\theta_i - h(\alpha_1^i, \dots, \alpha_n^i)\|^2 \quad (4.8)$$

dove  $\hat{h} : \mathbb{R}^n \rightarrow \mathbb{R}^q$  è una *funzione di fitting*.

Possiamo scegliere varie classi di funzioni per  $h$ : ci si può limitare a funzioni lineari, oppure passare al caso non lineare (modelli NARX, reti neurali...) da cui possiamo aspettarci migliori risultati. La minimizzazione deve essere eseguita con il metodo adeguato alla classe funzionale scelta.

In definitiva, abbiamo trovato due funzioni,  $\hat{f}$  e  $\hat{g}$ . L'operazione è stata probabilmente lunga (scelta dei valori dei parametri per la campagna di simulazione, campagna di simulazione vera e propria, compressione dei dati, scelta della funzione di fitting, calcolo di quest'ultima), ma ora la stima di  $\theta$  può essere eseguita in maniera banale:

$$\hat{\theta} = \hat{h}(\hat{g}(\bar{D}^N)) \quad (4.9)$$

Ogni volta che si ottiene una nuova serie di dati  $\bar{D}^N$ , basta valutare nuovamente  $\hat{\theta}$  e si ha immediatamente una stima aggiornata del valore dei parametri.

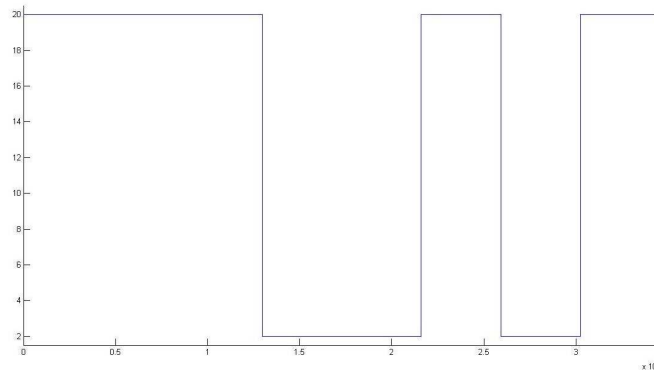
Per maggiori approfondimenti (ad esempio sulla scelta del periodo di campionamento) si rimanda a [28].

#### 4.5 Identificazione Two-Stages applicata alla digestione anaerobica

Per mettere alla prova il metodo, si è considerato ADM1 come in grado di descrivere esattamente quello che accade nel sistema; di esso si ha a disposizione un modello Dymola, usato per generare i dati.

Si è quindi scelto un parametro ritenuto significativo, `k_dis` (legato alla fase di idrolisi), e si sono eseguite 50 simulazioni, da 40 punti ciascuna, variando il valore del parametro da 0.1 a 9.802. Lo script matlab `twostageadm1.m`, riportato in appendice, carica i dati ottenuti dalla simulazione Dymola e seleziona il vettore di input (`aDM1_1.x.c.in`), di output (`aDM1_1.r.ch4_tilde`)

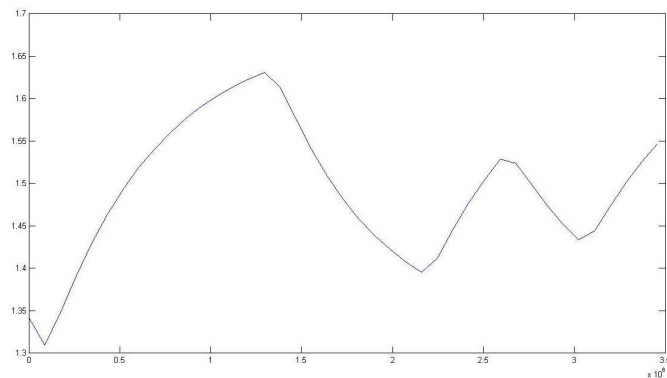
e il valore del parametro ( $k_{dis}$ ) per ognuna delle simulazioni. Per stimolare il sistema, si è adottato lo stesso ingresso descritto nella sezione 3.4 per il test sul paradigma LFT, ovviamente riadattato a un modello di tipo AMD1 (l'originale era AMOCO): si tratta di una variazione nel contenuto dell'influente in ingresso, modellata secondo una cosiddetta *onda a tre gradini*, come quella mostrata in figura 4.2. Essa, infatti, è considerata in grado di stimolare il modello in maniera adeguata [27].



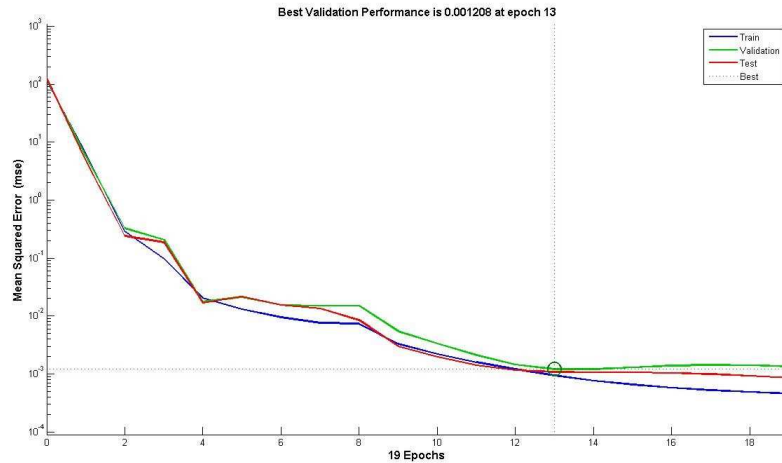
**Figura 4.2:** Onda a tre gradini

Tutte le altre variabili di ingresso sono state mantenute a un livello costante.

Come uscita, invece, è stata presa in considerazione la variabile relativa al flusso di metano in uscita, in quanto misurabile senza difficoltà, di reale interesse, e facilmente confrontabile con i risultati di una modellistica AMOCO. Un tipico grafico della variabile di uscita (in questo caso, relativa alla simulazione con  $k_{dis}=0.1$ ) è visibile in figura 4.3.



**Figura 4.3:** Andamento della portata di metano in uscita nel caso  $k_{dis}=0.1$



**Figura 4.4:** Errore quadratico medio a ogni iterazione

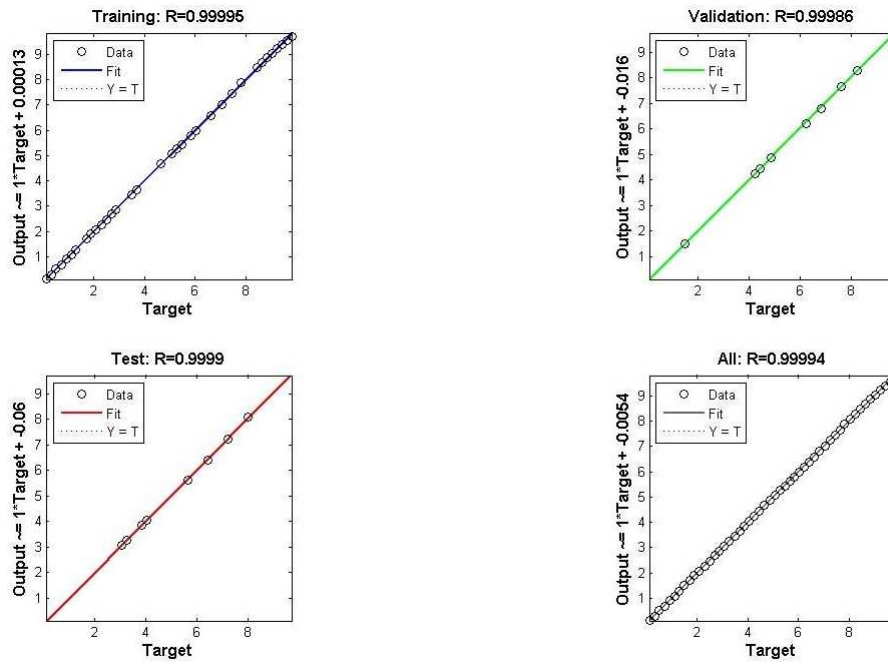
I dati vengono ora “compressi” (*prima fase* del paradigma Two Stages) mediante l’identificazione a minimi quadrati di un modello ARX per ogni serie di dati (cioè per ogni valore del parametro); in tutto, quindi, abbiamo 50 modelli ARX di ordine 10 ( $N_y = 0$ ,  $N_u = 5$ , ritardo nullo). Per far ciò, si è utilizzata la funzione di matlab  $arx()$ . Questa operazione ha richiesto 7.426820 secondi.

Facendo uso del *Neural Network Toolbox* di Matlab, è stata quindi addestrata una rete neurale di tipo feedforward in maniera da imitare il comportamento del sistema (*seconda fase* del paradigma Two Stages). Questa operazione ha richiesto 2.345404 secondi.

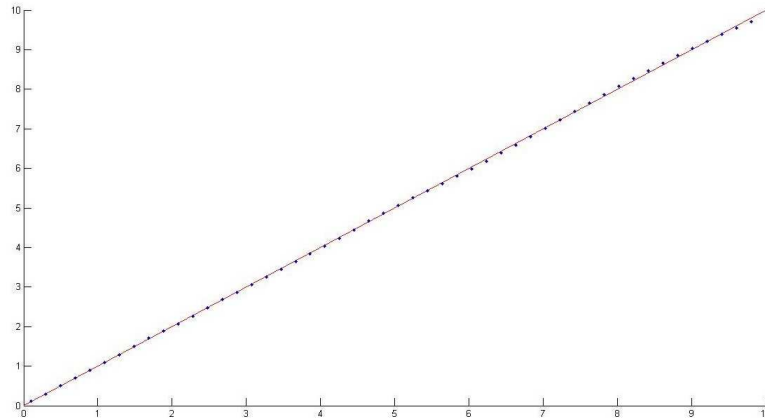
La bontà dei risultati ottenuti può essere apprezzata dalle figure 4.4 e 4.5. Va notato come non sempre l’addestramento dà buoni risultati già al primo tentativo; può infatti capitare che l’errore quadratico medio non si abbassi sufficientemente. In tal caso, è sufficiente ripetere l’addestramento; trattandosi di un’operazione da eseguirsi offline, ciò è senz’altro possibile.

## 4.6 Primo test: identificazione dei parametri originali

Un primo test è eseguito sugli stessi dati input-output utilizzati per l’addestramento: la rete neurale è in grado di fornire una buona stima dei parametri utilizzati? È chiaro che un risultato positivo non può essere considerato significativo (la rete è stata addestrata proprio su quei dati), mentre un risultato negativo porrebbe immediatamente fine all’applicabilità di que-



**Figura 4.5:** Test sulla bontà dei risultati: desideriamo che i dati stiano sulla diagonale in ogni riquadro, e che il coefficiente di correlazione R sia 1.



**Figura 4.6:** Primo test: capacità della rete neurale di identificare il valore dei parametri utilizzati per l'addestramento

sto metodo. Come previsto, invece, i parametri sono stimati accuratamente (figura 4.6)

## 4.7 Secondo test: identificazione con nuovi valori dei parametri

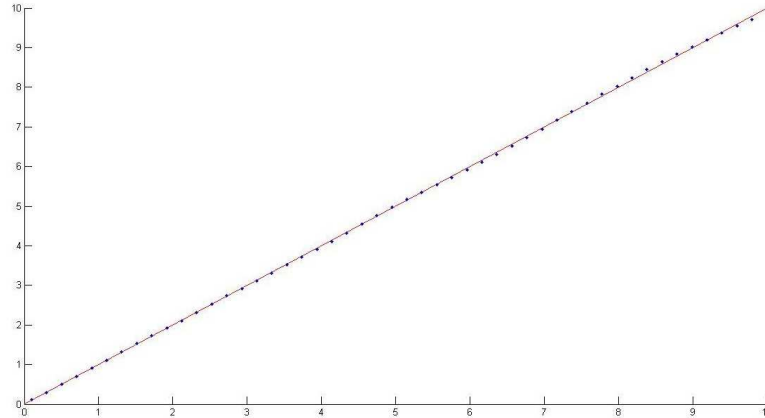
Sono stati eseguite 49 nuove simulazioni in ambiente Dymola, con valori di  $k_{dis}$  variabili fra 0.1 e 9,798. Ciò serve a valutare il comportamento dello stimatore a seguito di valori dei parametri diversi da quelli utilizzati per l'addestramento della rete. I valori ottenuti sono stati caricati in Matlab, compressi in modelli arx e processati dalla rete neurale ottenuta nella sezione 4.5. Ne risulta una stima dei parametri con errore relativo compreso fra  $5,26e-5$  e 0,0785 (figura 4.7).

Il tempo impiegato è risultato pari a 5.232063 secondi, con una media di 0.1068 secondi a stima.

Il test è poi stato ripetuto su una singola simulazione ( $k_{dis}=5.9592$ ). Il risultato della stima è pari a  $k_{dis\_stimato}= 5.9115$ , con un errore relativo pari a 0.0080, ottenuto in 0.141324 secondi.

## 4.8 Terzo test: identificazione in presenza di rumore di misura

In realtà, non possiamo davvero supporre che il nostro modello sia perfetto, né che le nostre misure siano perfettamente accurate. Ciò non è mai



**Figura 4.7:** Secondo test: capacità della rete neurale di identificare valori dei parametri differenti da quelli utilizzati per l'addestramento

possibile, tanto più in una situazione complessa come è la digestione anaerobica. Per questo motivo, l'esperimento è stato ripetuto (nello script matlab *twostageadm1rumore.m*) con l'aggiunta di un rumore gaussiano bianco alla variabile misurata, con rapporto segnale/rumore pari a 50. Un esempio si ha in figura 4.8. Come si nota in figura, e come evidenziato dal valore del rapporto S/N, pari a 50, non si tratta di un rumore notevole. Eppure, i risultati sono visibili in figura 4.9: la rete così addestrata perde completamente in capacità identificativa.

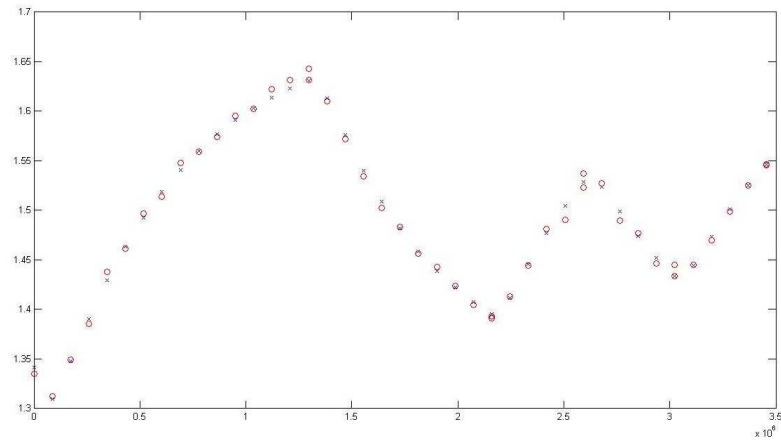
## 4.9 Valutazione dei risultati e possibili sviluppi

Il paradigma Two Stages conferma il suo principale punto di forza, ovvero la velocità di esecuzione una volta processati i dati iniziali; si vedano i tempi impiegati nelle sezioni 4.6 e 4.7 per identificare ben 50 valori dei parametri.

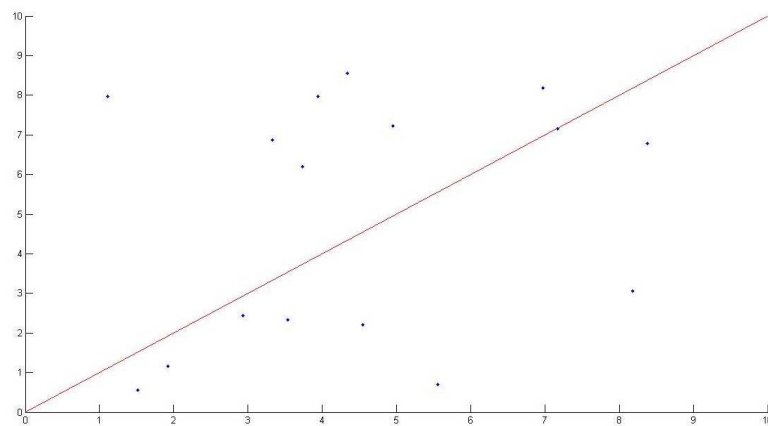
Tuttavia, si trattava di un caso semplificato (si supponeva che un solo parametro cambiasse), e l'aggiunta di un rumore poco influente si rivela completamente deleterio per la buona riuscita dell'identificazione. Così com'è, quindi, non è impiegabile nella nostra situazione.

Chiaramente, il lettore interessato potrebbe provare a indagare in questi ambiti:

- valutare come influiscono l'ordine del modello arx, ma soprattutto la struttura della rete neurale, su accuratezza dei risultati, velocità di addestramento e velocità di stima;
- valutare come cambiano i risultati in caso di ingressi alternativi;



**Figura 4.8:** Portata di metano vera ('x' blu) e misurata in presenza di rumore ('o' rosse)



**Figura 4.9:** Terzo test: capacità della rete neurale di identificare valori dei parametri in presenza di rumore



- valutare se il parametro considerato, nonché l'ingresso e l'uscita utilizzati, siano davvero significativi (e non ci siano, ad esempio, problemi di identificabilità);
- valutare che accade se il parametro varia fuori dall'intervallo considerato, o se la distribuzione che segue non è uniforme;
- capire se l'applicabilità del metodo in questo contesto presenta falle teoriche prima ancora che pratiche.



# Bibliografia

- [1] G. Invernizzi. *Modellistica e identificazione del processo di digestione anaerobica*. Tesi di laurea magistrale, Politecnico di Milano, A.A. 2011-12. <http://hdl.handle.net/10589/72365>
- [2] A. Allegrini. *Anaerobic Digestion Modelling: a Comparison between ADM1 and AMOCO*. Tesi di laurea magistrale, Politecnico di Milano, A.A. 2009-2010. <http://hdl.handle.net/10589/16362>.
- [3] O. Bernard, Z. Hadj-Saddok, D. Dochain, A. Genovesi, J. Steyer. *Dynamical Model Development and Parameter Identification for an Anaerobic Wastewater Treatment Process*. Biotechnology and Bioengineering, 2001.
- [4] A. Della Bona, G. Ferretti, E. Ficara, F. Malpei. *Parameter identification of a reduced order LFT model of anaerobic digestion. Control-oriented Modelling and Identification: Theory and Practice*. M. Lovera Ed., The Institution of Engineering and Technology, Stevenage, 2015, pp. 301-323.
- [5] A. Della Bona, G. Ferretti, E. Ficara, F. Malpei. *LFT modelling and identification of anaerobic digestion*. Control Engineering Practice. Vol. 36, March 2015, pp. 1-11.
- [6] 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 & Technology, 2002.
- [7] C. Rosen, U. Jeppson. *Aspects on ADM1 implementation within the BSM2 framework*. Dept. of Industrial Electrical Engineering and Automation, Lund University, 2006.
- [8] Van Hulle, Stijn W.H.; Vesvikar, Mehul; Poutiainen, Hannu; Nopens, Ingmar. *Importance of scale and hydrodynamics for modeling anaerobic digester performance*. Chemical Engineering Journal, anno:2014 volume:255 pp. 71-77.

- [9] Ariunbaatar, J., Panico, A., Frunzo, L., Esposito, G., Lens, P.N.L., Pirozzi, F. *Enhanced anaerobic digestion of food waste by thermal and ozonation pretreatment methods*. Journal of Environmental Management, Volume 146, 15 December 2014, pp. 142-149.
- [10] Deng, L., Yang, H., Liu, G., Zheng, D., Chen, Z., Liu, Y., Pu, X., Song, L., Wang, Z., Lei, Y. *Kinetics of temperature effects and its significance to the heating strategy for anaerobic digestion of swine wastewater*. Applied Energy, Volume 134, 1 December 2014, pp. 349-355.
- [11] Xiao-Shuang Shi, Xian-Zheng Yuan, Yu-Ping Wang, Shu-Juan Zeng, Yan-Ling Qiu, Rong-Bo Guo, Li-Sheng Wang. *Modeling of the methane production and pH value during the anaerobic co-digestion of dairy manure and spent mushroom substrate*. Chemical Engineering Journal 244 (2014) pp. 258-263.
- [12] U. G. Ozkan-Yucel, C. F. Gökçay. *Application of ADM1 model to a full-scale anaerobic digester under dynamic organic loading conditions*. Environmental technology, 31(6):633-640, 2010.
- [13] Anne Motelica Wagenaar, Arjen Nauta, Ellen G. H. M. van den Heuvel, Robbert Kleerebezem. *Flux analysis of the human proximal colon using Anaerobic Digestion Model 1*. Anaerobe, 28, pp. 137-148, 2014.
- [14] T.Zonta, Ž. Zonta, M. M. Alves, X. Flotats, J. Palatsi. *Modelling inhibitory effects of long chain fatty acids in the anaerobic digestion process*. Water research, 47(3):1369-1380, 2013.
- [15] S. Hassam, B. Cherki, E. Ficara, J. Harmand. *Towards a systematic approach to reduce complex bioprocess models - Application to the ADM1*. In 2012 20th Mediterranean Conference on Control and Automation, MED 2012 - Conference Proceedings, pp. 573-578, 2012.
- [16] Astrid Bornhöft. *Steady-state analysis of the anaerobic digestion model no. 1 (ADM1)*. Nonlinear Dynamics, 73(1):535-549, 2013.
- [17] D. J. Batstone. *Estimation of hydrolysis parameters in full-scale anaerobic digesters*. Biotechnology and bioengineering, 102(5):1513-1520, 2009.
- [18] Theo S. O. Souza. *ADM1 calibration using BMP tests for modeling the effect of autohydrolysis pretreatment on the performance of continuous sludge digesters*. Water research, 47(9):3244-3254, 2013.
- [19] Kanokwan Boe. *State indicators for monitoring the anaerobic digestion process*. Water research, 44(20):5973-5980, 2010.

- [20] Michael Madsen, Jens Bo Holm-Nielsen, Kim H. Esbensen. *Monitoring of anaerobic digestion processes: A review perspective*. Renewable and Sustainable Energy Reviews, 15(6):3141-3155, 8 2011.
- [21] Alexis Mottet. *Estimating anaerobic biodegradability indicators for waste activated sludge*. Chemical Engineering Journal, 160(2):488-496, 2010.
- [22] J. Jimenez, E. Gonidec, J. A. Cacho Rivero, E. Latrille, F. Vedrenne, J.-P. Steyer. *Prediction of anaerobic biodegradability and bioaccessibility of municipal sludge by coupling sequential extractions with fluorescence spectroscopy: Towards ADM1 variables characterization*. Water research, 50:359-372, 2014.
- [23] R. Girault, G. Bridoux, F. Nauleau, C. Poullain, J. Buffet, J.-P. Steyer, A. G. Sadowski, F. Béline. *A waste characterisation procedure for ADM1 implementation based on degradation kinetics*. Water research, 46(13):4099-4110, 2012.
- [24] Cesar-Arturo Aceves Lara. *Optimal control of hydrogen production in a continuous anaerobic fermentation bioreactor*. International Journal of Hydrogen Energy, 35(19):10710-10718, 2010.
- [25] J. Lauwers, L. Appels, I. P. Thompson, J. Degreve, J. F. Van Impe, R. Dewil. *Mathematical modelling of anaerobic digestion of biomass and waste: Power and limitations*. Progress in Energy and Combustion Science, 39(4):383-402, 2013.
- [26] Xavier Flores Alsina. *Assessing the use of activated sludge process design guidelines in wastewater treatment plant projects: A methodology based on global sensitivity analysis*. Environmental modelling & software, 38:50-58, 2012.
- [27] V. Klein, E. A. Morelli. *Aircraft system identification: theory and practice*, America Institute of Aeronautics and Astronautics, 2006
- [28] S.Garatti, S.Bittanti. *A new paradigm for parameter estimation in system modeling*. International journal of adaptive control and signal processing, 2013; 27(8), pp. 667-687.



## Appendice A

# Matrici di Petersen

Sono ora presentati le matrici di Petersen del modello ADM1, che riporta i vari coefficienti stechiometrici (vedi capitolo 2, sezione 2.4). Le colonne si riferiscono alle variabili di stato, mentre le file ai processi; una colonna aggiuntiva contiene l'equazione cinetica corrispondente a ogni processo.

Component	1	2	3	4	5	6	7	8	9	10	11	12	
Process	$S_{su}$	$S_{sa}$	$S_b$	$S_{va}$	$S_{su}$	$S_{pro}$	$S_{ac}$	$S_{H_2}$	$S_{CH_4}$	$S_{ic}$	$S_{N}$	$S_i$	
1 Disintegration												$k_{dis} \cdot X_c$	
2 Hydrolysis Carbohydrates	1											$k_{hydro} \cdot X_{ch}$	
3 Hydrolysis of Proteins	1											$k_{hydro} \cdot X_{pr}$	
4 Hydrolysis of Lipids	$1-f_{su,li}$		$f_{su,li}$									$k_{hydro} \cdot X_{li}$	
5 Uptake of Sugars	-1				$(1-Y_{su}) \cdot f_{su,sa}$	$(1-Y_{su}) \cdot f_{pro,sa}$	$(1-Y_{su}) \cdot f_{ac,sa}$	$(1-Y_{su}) \cdot f_{H_2,sa}$		$-\sum_{i=1}^{24} C_i \cdot Y_{i,1-24}$	$-(Y_{su}) \cdot N_{bac}$		$k_{m,sa} \cdot \frac{S_{su}}{K_S + S} \cdot X_{su} \cdot I_1$
6 Uptake of Amino Acids		-1			$(1-Y_{su}) \cdot f_{su,sa}$	$(1-Y_{su}) \cdot f_{pro,sa}$	$(1-Y_{su}) \cdot f_{ac,sa}$	$(1-Y_{su}) \cdot f_{H_2,sa}$		$-\sum_{i=1}^{24} C_i \cdot Y_{i,6}$	$(1-Y_{su}) \cdot N_{aa} - (Y_{su}) \cdot N_{bac}$		$k_{m,sa} \cdot \frac{S_{su}}{K_S + S} \cdot X_{su} \cdot I_1$
7 Uptake of LCFA			-1				$(1-Y_b) \cdot 0.7$	$(1-Y_b) \cdot 0.3$			$-(Y_b) \cdot N_{bac}$		$k_{m,fa} \cdot \frac{S_b}{K_S + S_b} \cdot X_{fa} \cdot I_2$
8 Uptake of Valerate				-1		$(1-Y_v) \cdot 0.54$	$(1-Y_v) \cdot 0.31$	$(1-Y_v) \cdot 0.15$			$-(Y_v) \cdot N_{bac}$		$k_{m,v} \cdot \frac{S_{va}}{K_S + S_{va}} \cdot X_{va} \cdot \frac{I}{1 + S_{va}/S_{va}}$
9 Uptake of Butyrate					-1		$(1-Y_b) \cdot 0.8$	$(1-Y_b) \cdot 0.2$			$-(Y_b) \cdot N_{bac}$		$k_{m,b} \cdot \frac{S_{bu}}{K_S + S_{bu}} \cdot X_{bu} \cdot \frac{I}{1 + S_{bu}/S_{bu}}$
10 Uptake of Propionate						-1	$(1-Y_p) \cdot 0.57$	$(1-Y_p) \cdot 0.43$		$-\sum_{i=1}^{24} C_i \cdot Y_{i,10}$	$-(Y_p) \cdot N_{bac}$		$k_{m,pr} \cdot \frac{S_{pro}}{K_S + S_{pro}} \cdot X_{pro} \cdot I_2$
11 Uptake of Acetate							-1		$(1-Y_a)$	$-\sum_{i=1}^{24} C_i \cdot Y_{i,11}$	$-(Y_a) \cdot N_{bac}$		$k_{m,ac} \cdot \frac{S_{ac}}{K_S + S_{ac}} \cdot X_{ac} \cdot I_3$
12 Uptake of Hydrogen								-1	$(1-Y_H)$	$-\sum_{i=1}^{24} C_i \cdot Y_{i,12}$	$-(Y_H) \cdot N_{bac}$		$k_{m,H} \cdot \frac{S_{H_2}}{K_S + S_{H_2}} \cdot X_{H_2} \cdot I_4$
13 Decay of $X_{su}$													$k_{dec,su} \cdot X_{su}$
14 Decay of $X_{sa}$													$k_{dec,sa} \cdot X_{sa}$
15 Decay of $X_b$													$k_{dec,b} \cdot X_b$
16 Decay of $X_{va}$													$k_{dec,va} \cdot X_{va}$
17 Decay of $X_{pro}$													$k_{dec,pro} \cdot X_{pro}$
18 Decay of $X_{ac}$													$k_{dec,ac} \cdot X_{ac}$
19 Decay of $X_{H_2}$													$k_{dec,H_2} \cdot X_{H_2}$

Figura A.1: Matrice di Petersen per le componenti solubili



j	Component →	i	Rate (n, g COD·L <sup>-1</sup> ·d <sup>-1</sup> )																		
			13	14	15	16	17	18	19	20	21	22	23	24							
1	Disintegration	$X_c$	$X_{ch}$	$X_{gr}$	$X_i$	$X_{su}$	$X_{aa}$	$X_b$	$X_{c4}$	$X_{pro}$	$X_{ac}$	$X_{h2}$	$X_i$	$f_{lisc}$							
2	Hydrolysis of Carbohydrates	-1	$f_{h,nc}$	$f_{p,nc}$	$f_{l,nc}$																
3	Hydrolysis of Proteins		-1																		
4	Hydrolysis of Lipids			-1																	
5	Uptake of Sugars					$Y_{su}$															
6	Uptake of Amino Acids						$Y_{aa}$														
7	Uptake of LCFA							$Y_b$													
8	Uptake of Valerate								$Y_{c4}$												
9	Uptake of Butyrate									$Y_{c4}$											
10	Uptake of Propionate										$Y_{pro}$										
11	Uptake of Acetate											$Y_{ac}$									
12	Uptake of Hydrogen												$Y_{h2}$								
13	Decay of $X_{su}$	1				-1															
14	Decay of $X_{aa}$	1					-1														
15	Decay of $X_b$	1						-1													
16	Decay of $X_{c4}$	1							-1												
17	Decay of $X_{pro}$	1								-1											
18	Decay of $X_{ac}$	1										-1									
19	Decay of $X_{h2}$	1											-1								

Figura A.2: Matrice di Petersen per le componenti particolante



## Appendice B

# Script di Matlab utilizzati

### B.1 Applicazione del paradigma TS in assenza di rumore

L'applicazione del paradigma Two Stages in assenza di rumore, con l'addestramento della rete e il suo utilizzo ai fini dell'identificazione, è stato svolto grazie allo script `twostageadm1.m`.

```
%twostageadm1
clear all
close all
clc
home

%%
%caricamento dati

[s,n]=
tload('C:\MieiProgetti\Dymola\RisultatiADM1\
SimulaSimula50x40.mat');

time = s(:,1); %è una colonna
nSim=s(1,2);

indice1y=tnindex(n,'Condizioni[1].aDM1_1.r_ch4_tilde');
indice2y=tnindex(n,'Condizioni[2].aDM1_1.r_ch4_tilde');
diffindy=indice2y-indice1y;
indice1u=tnindex(n,'Condizioni[1].aDM1_1.x_c_in');
indice2u=tnindex(n,'Condizioni[2].aDM1_1.x_c_in');
diffindu=indice2u-indice1u;
indice1k=tnindex(n,'Condizioni[1].aDM1_1.k_dis');
```

```

indice2k=tnindex(n,'Condizioni[2].aDM1_1.k_dis');
diffindk=indice2k-indice1k;
%così gli indici vengono identificati in automatico.
%Le prime cinque colonne di s rappresentano
%le variabili Time, nSim,kdismin, kdismax, passo.
%Seguono poi gli nSim valori dei parametri.
%Ogni simulazione ha poi 475 variabili.

puntipersim=
length(s(:,tnindex(n,'Condizioni[2].aDM1_1.r_ch4_tilde')));

u=zeros(puntipersim,nSim);
y=zeros(puntipersim,nSim);
par_sim=zeros(1,nSim);
for i=1:nSim
u(:,i)=s(:,indice1u+(i-1)*diffindu);
y(:,i)=s(:,indice1y+(i-1)*diffindy);
par_sim(i)=s(1,indice1k+(i-1)*diffindk);
dati{i}=iddata(y(:,i),u(:,i));
end

figure; hold on;
axis([0 time(end) (min(u(:,1))-0.5) (max(u(:,1))+0.5)]);
plot(time,u(:,1))
figure;
plot(time,y(:,1))

%%
%generazione arx

tic

N_a = 5;
N_b = 5;
N_k = 0;
par_arx = zeros(N_a+N_b,nSim);

for k=1:nSim
funzg{k} = arx(dati{k},[N_a N_b N_k]);
par_arx(:,k) = [funzg{k}.a(2:N_a+1)';
funzg{k}.b(N_k+1:N_k+N_b)'];
end

toc %questo è il tempo per addestrare gli arx.

```

```
%%
%addestramento rete

tic

funzh = newff(par_arx,par_sim,10);
funzh.trainParam.epochs = 100;
[funzh,datiadd] = train(funzh,par_arx,par_sim);

toc %questo è il tempo per addestrare la rete

figure
plotperform(datiadd)

%%
%test su stessi dati
par_arx_test_vecchi = zeros(N_a+N_b,nSim);
tic
for k=1:nSim
arx_test_vecchi = arx(dati{k},[N_a N_b N_k]);
par_arx_test_vecchi(:,k) = [arx_test_vecchi.a(2:N_a+1)';
arx_test_vecchi.b(N_k+1:N_k+N_b)'];
end

par_stimati_vecchi = sim(funzh,par_arx_test_vecchi);
toc

figure; hold on;
axis([0 10 0 10]); plot([0 10],[0 10],'r');
plot(par_sim,par_stimati_vecchi,'.')
```

```
%%
%test su nuovi dati

%dati di test
[stest,ntest]=
tload('C:\MieiProgetti\Dymola\RisultatiADM1\
SimulaSimula50x40test.mat');
```

```
timet = stest(:,1); %è una colonna
nSimt=stest(1,2);
```

```

indice1yt=tnindex(nctest,'Condizioni [1].aDM1_1.r_ch4_tilde');
indice2yt=tnindex(nctest,'Condizioni [2].aDM1_1.r_ch4_tilde');
diffindyt=indice2yt-indice1yt;
indice1ut=tnindex(nctest,'Condizioni [1].aDM1_1.x_c_in');
indice2ut=tnindex(nctest,'Condizioni [2].aDM1_1.x_c_in');
diffindut=indice2ut-indice1ut;
indice1kt=tnindex(nctest,'Condizioni [1].aDM1_1.k_dis');
indice2kt=tnindex(nctest,'Condizioni [2].aDM1_1.k_dis');
diffindkt=indice2kt-indice1kt;

puntipersimt=
length(stest(:,tnindex(nctest,
'Condizioni [2].aDM1_1.r_ch4_tilde'))));

ut=zeros(puntipersimt,nSimt);
yt=zeros(puntipersimt,nSimt);
par_test=zeros(1,nSimt);

for i=1:nSimt
ut(:,i)=stest(:,indice1ut+(i-1)*diffindut);
yt(:,i)=stest(:,indice1yt+(i-1)*diffindyt);
par_test(i)=stest(1,indice1kt+(i-1)*diffindkt);
dati_test{i}=iddata(yt(:,i),ut(:,i));
end

par_arx_test=zeros(N_a+N_b,nSimt);
tic
for k=1:nSimt
arx_test= arx(dati_test{k},[N_a N_b N_k]);
par_arx_test(:,k) = [arx_test.a(2:N_a+1)'; arx_test.b(N_k+1:N_k+N_b)'];
end

par_stimati = sim(funzh,par_arx_test);
toc

figure; hold on;
axis([0 10 0 10]); plot([0 10],[0 10],'r');
plot(par_test,par_stimati,'.')

differenza_parametri=par_test-par_stimati;
errore_relativo=differenza_parametri./par_test;
errore_min=min(abs(errore_relativo));
errore_max=max(abs(errore_relativo));

```

```

%%
%Test su singolo parametro

indice_selezionato = randi(51,1);
u_sel=stest(:,indice1ut+(indice_selezionato-1)*diffindut);
y_sel=stest(:,indice1yt+(indice_selezionato-1)*diffindyt);
par_sel=stest(1,indice1kt+(indice_selezionato-1)
*diffindkt) %#ok<NOPTS>
dati_sel=iddata(y_sel,u_sel);

tic
arx_sel= arx(dati_sel,[N_a N_b N_k]);
par_arx_sel = [arx_sel.a(2:N_a+1)';
arx_sel.b(N_k+1:N_k+N_b)'];
par_sel_stimato = sim(funzh,par_arx_sel);
toc

par_sel_stimato %#ok<NOPTS>

errore_sel=par_sel-par_sel_stimato;
errore_rel=errore_sel/par_sel %#ok<NOPTS>

```

## B.2 Applicazione del paradigma TS in presenza di rumore

L'applicazione del paradigma Two Stages in presenza di rumore, con l'addestramento della rete e il suo utilizzo ai fini dell'identificazione, è stato svolto grazie allo script `twostageadm1rumore.m`.

```

%twostageadm1
clear all
close all
clc
home

%%
%caricamento dati

[s,n]=
tload('C:\MieiProgetti\Dymola\RisultatiADM1\
SimulaSimula50x40.mat');

```

```

time = s(:,1); %è una colonna
nSim=s(1,2);

indice1y=tnindex(n,'Condizioni[1].aDM1_1.r_ch4_tilde');
indice2y=tnindex(n,'Condizioni[2].aDM1_1.r_ch4_tilde');
diffindy=indice2y-indice1y;
indice1u=tnindex(n,'Condizioni[1].aDM1_1.x_c_in');
indice2u=tnindex(n,'Condizioni[2].aDM1_1.x_c_in');
diffindu=indice2u-indice1u;
indice1k=tnindex(n,'Condizioni[1].aDM1_1.k_dis');
indice2k=tnindex(n,'Condizioni[2].aDM1_1.k_dis');
diffindk=indice2k-indice1k;
%così gli indici vengono identificati in automatico.
%Le prime cinque colonne di s rappresentano
%le variabili Time, nSim,kdismin, kdismax, passo.
%Seguono poi gli nSim valori dei parametri.
%Ogni simulazione ha poi 475 variabili.

puntipersim=
length(s(:,tnindex(n,'Condizioni[2].aDM1_1.r_ch4_tilde')));

u=zeros(puntipersim,nSim);
y=zeros(puntipersim,nSim);
par_sim=zeros(1,nSim);
for i=1:nSim
u(:,i)=s(:,indice1u+(i-1)*diffindu);
y(:,i)=s(:,indice1y+(i-1)*diffindy);
par_sim(i)=s(1,indice1k+(i-1)*diffindk);
dati{i}=iddata(y(:,i),u(:,i));
end

figure; hold on;
axis([0 time(end) (min(u(:,1))-0.5) (max(u(:,1))+0.5)]);
plot(time,u(:,1))
figure;
plot(time,y(:,1))

%%
%generazione arx

tic

N_a = 5;

```



```

N_b = 5;
N_k = 0;
par_arx = zeros(N_a+N_b,nSim);

for k=1:nSim
funzg{k} = arx(dati{k},[N_a N_b N_k]);
par_arx(:,k) = [funzg{k}.a(2:N_a+1)';
funzg{k}.b(N_k+1:N_k+N_b)'];
end

toc %questo è il tempo per addestrare gli arx.

%%
%addestramento rete

tic

funzh = newff(par_arx,par_sim,10);
funzh.trainParam.epochs = 100;
[funzh,datiadd] = train(funzh,par_arx,par_sim);

toc %questo è il tempo per addestrare la rete

figure
plotperform(datiadd)

%%
%test su stessi dati
par_arx_test_vecchi = zeros(N_a+N_b,nSim);
tic
for k=1:nSim
arx_test_vecchi = arx(dati{k},[N_a N_b N_k]);
par_arx_test_vecchi(:,k) = [arx_test_vecchi.a(2:N_a+1)';
arx_test_vecchi.b(N_k+1:N_k+N_b)'];
end

par_stimati_vecchi = sim(funzh,par_arx_test_vecchi);
toc

figure; hold on;
axis([0.9 10.1 0.9 10.1]); plot([0.9 10.1],[0.9 10.1],'r');
plot(par_sim,par_stimati_vecchi,'.')

%%

```

```

%test su nuovi dati

%dati di test
[stest,ntest]=
tload('C:\MieiProgetti\Dymola\RisultatiADM1\
SimulaSimula50x40test.mat');

timet = stest(:,1); %è una colonna
nSimt=stest(1,2);

indice1yt=tnindex(ntest,'Condizioni [1].aDM1_1.r_ch4_tilde');
indice2yt=tnindex(ntest,'Condizioni [2].aDM1_1.r_ch4_tilde');
diffindyt=indice2yt-indice1yt;
indice1ut=tnindex(ntest,'Condizioni [1].aDM1_1.x_c_in');
indice2ut=tnindex(ntest,'Condizioni [2].aDM1_1.x_c_in');
diffindut=indice2ut-indice1ut;
indice1kt=tnindex(ntest,'Condizioni [1].aDM1_1.k_dis');
indice2kt=tnindex(ntest,'Condizioni [2].aDM1_1.k_dis');
diffindkt=indice2kt-indice1kt;

puntipersimt=
length(stest(:,tnindex(ntest,
'Condizioni [2].aDM1_1.r_ch4_tilde')));

ut=zeros(puntipersimt,nSimt);
yt=zeros(puntipersimt,nSimt);
ym=zeros(puntipersimt,nSimt); %y del test con rumore aggiunto
par_test=zeros(1,nSimt);

for i=1:nSimt
ut(:,i)=stest(:,indice1ut+(i-1)*diffindut);
yt(:,i)=stest(:,indice1yt+(i-1)*diffindyt);
ym(:,i)=awgn(yt(:,i),50,'measured'); %50 è il rapporto s/n,
%measured indica di misurare la potenza del segnale
%prima di applicare il rumore
%ym(:,i)=yt(:,i);
par_test(i)=stest(1,indice1kt+(i-1)*diffindkt);
dati_test{i}=iddata(ym(:,i),ut(:,i));
end

figure
plot(timet,yt(:,1),'bx',timet,ym(:,1),'ro');
%%
par_arx_test=zeros(N_a+N_b,nSimt);

```

```

tic
for k=1:nSimt
arx_test= arx(dati_test{k},[N_a N_b N_k]);
par_arx_test(:,k) = [arx_test.a(2:N_a+1)';
arx_test.b(N_k+1:N_k+N_b)'];
end

par_stimati = sim(funzh,par_arx_test);
toc

figure; hold on;
axis([0 10 0 10]); plot([0 10],[0 10],'r');
plot(par_test,par_stimati,'.')
```

```

differenza_parametri=par_test-par_stimati;
errore_relativo=differenza_parametri./par_test;
errore_min=min(abs(errore_relativo));
errore_max=max(abs(errore_relativo));

%%
%Test su singolo parametro

indice_selezionato = randi(51,1);
u_sel=stest(:,indice1ut+(indice_selezionato-1)*diffindut);
y_sel=stest(:,indice1yt+(indice_selezionato-1)*diffindyt);
par_sel=stest(1,indice1kt+(indice_selezionato-1)
*diffindkt) %#ok<NOPTS>
dati_sel=iddata(y_sel,u_sel);

tic
arx_sel= arx(dati_sel,[N_a N_b N_k]);
par_arx_sel = [arx_sel.a(2:N_a+1)';
arx_sel.b(N_k+1:N_k+N_b)'];
par_sel_stimato = sim(funzh,par_arx_sel);
toc

par_sel_stimato %#ok<NOPTS>

errore_sel=par_sel-par_sel_stimato;
errore_rel=errore_sel/par_sel %#ok<NOPTS>
```