UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI INGEGNERIA INDUSTRIALE

CORSO DI LAUREA MAGISTRALE IN INGEGNERIA CHIMICA E DEI PROCESSI INDUSTRIALI

Tesi di Laurea Magistrale in

Ingegneria Chimica e dei Processi Industriali

STATE ESTIMATION TECHNIQUES FOR ON-LINE MODEL ADAPTATION: A CASE STUDY IN THERMAL CRACKING

Relatore: Prof. Massimiliano Barolo

Correlatore: Dr. Ying Sheng Cheng

Laureando: **GABRIELE BANO**

ANNO ACCADEMICO 2014-2015

Allo zio Esto e a Nonna Maria

Abstract

This thesis deals with the implementation of a state estimation technique (EKF, Extended Kalman Filter) on a preexisting dynamic model of a steam cracker (PDAE, 227 differential variables, 14268 algebraic variables), using the dynamic process simulator gPROMS®.

Given few and reliable on-line measurements coming from the real plant (pressure drop along the coil, fuel consumption, mass fraction of ethane at the outlet), the estimator is able to predict the state of the furnace, in particular its state of coking. The knowledge of the state of coking is fundamental to perform a cyclic whole-plant optimization without assuming clean-tube conditions. Some model parameters, like the coking kinetics constants, are adjusted in real-time by the estimator to reconcile the model prediction with the available plant observations. The performance of the estimator has been tested off-line using four different sets of plant data. For a given set of specifications (initial/model/measurement error covariances), results have proven to be satisfactory independently of the feed and set of data considered. Furthermore, the computational effort required has proven to be compatible with an on-line implementation of the state estimator in the real plant.

Riassunto

Introduzione e obiettivo L'utilizzo in linea di stimatori per predire lo stato di un sistema è una tecnica consolidata nell' ambito del controllo di processo e in ambito metereologico. Tuttavia, il loro utilizzo per l'aggiustamento in linea di modelli dinamici di impianti dell'industria chimica-petrolchimica trova ancora scarsa applicazione, prevalentemente a causa delle difficoltà numeriche e computazionali che si riscontrano quando tali tecniche vengono applicate a modelli con un numero elevato di variabili (decine di migliaia).

L'obiettivo del seguente lavoro di Tesi è l'applicazione di una tecnica di analisi di stato (filtro di Kalman esteso) and un preesistente modello dinamico di uno steam cracker (227 variabili differenziali, 14268 algebriche) usando il simulatore dinamico gPROMS® della Process Systems Enterprise Ltd. Fornendo allo stimatore pochi ed affidabili dati di impianto (perdite di carico lungo il serpentino, portata di combustile, frazione massiva di etano/propano in uscita), si desidera ottenere una corretta previsione dello stato del sistema (ed in particolare dello stato di *coking* della fornace). Il modello dello steam cracker analizzato appartiene ad un più articolato modello di un intero impianto per la produzione di etilene situato in Arabia Saudita. Tale modello viene attualmente utilizzato per condurre una ottimizzazione globale delle prestazioni dell'impianto. Tuttavia, data l'incapacità del modello di predire correttamente lo stato di *coking* della fornace, tale ottimizzazione viene condotta senza considerare la deposizione di coke nel serpentino della stessa. La deposizione di coke nel serpentino influenza fortemente le prestazioni della fornace (e quindi dell'intero impianto), rendendo i risultati ottenuti inaffidabili. L'utilizzo di uno stimatore in linea per l'ottenimento di una corretta stima dello stato di coking del sistema può quindi risultare estremamente utile per poter condurre una ottimizzazione ciclica delle prestazioni dell' impianto. Il tempo computazionale richiesto dallo stimatore deve chiaramente risultare compatibile con una sua possibile implementazione in linea. Inoltre, le prestazioni dello stimatore devono risultare robuste indipendentemente dal tipo e numero di misurazioni provenienti dall'impianto, garantendo la massima flessibilità possibile nel suo utilizzo.

Materiali e metodi Il lavoro è stato condotto presso la sede centrale della Process Systems Enterprise Ltd a Londra, Regno Unito. Lo sviluppo della Tesi si è articolato in cinque fasi.
Nella prima fase, l'obiettivo è stato quello di raggiungere una solida compresione del modello dinamico dello steam cracker. Data una serie di input (composizione dell'alimentazione, portata di idrocarburi e vapore al serpentino, temperature in ingresso e pressione in uscita) il modello consente di calcolare una serie di variabili di output (conversione di etano/propano, resa in etilene, perdite di carico lungo il serpentino, portata di combustibile richiesta, quantità di coke depositata nei tubi, temperatura in uscita, temperatura di pelle del serpentino). Le previsioni di tale modello risultano, con entità diversa a seconda della variabile di output presa in considerazione, in disaccordo con le osservazioni sperimentali disponibili in impianto. In particolare, la previsione sullo stato di coking del sistema risulta sensibilmente inaccurata. Pertanto, l'utilizzo di una tecnica in linea di analisi di stato per ottenere una corretta predizione dello stato di coking della fornace è stata presa in considerazione.

Nella seconda fase si sono acquisite le conoscenze di base sulle diverse tecniche di analisi di stato (filtro di Kalman esteso/Unscented, *Ensemble Kalman filter*, *Particle Filter*). I diversi stimatori sono stati dapprima implementati in linguaggio MATLAB®, quindi testati su alcuni semplici esempi per verificarne i punti di forza e debolezza. Si è quindi deciso di testare sul modello dello steam cracker il filtro di Kalman esteso, data la sua facilità di implementazione e testata conoscenza.
Nella terza fase il filtro di Kalman esteso è stato testato in gPROMS® su alcuni semplici esempi (sistemi ODE con max. 10 variabili di stato) per valutarne prestazioni e robustezza.
Nella quarta fase lo stimatore è stato implementato nel modello originale dello steam cracker. Le sue prestazioni sono state valutate in questa fase *off-line* generando dati sperimentali fittizi da modello e provvedendo a perturbare lievemente alcuni dei parameteri e/o condizioni iniziali di alcune delle variabili di stato. La procedura è stata adottata usando sia input invarianti nel tempo che input variabili nel tempo. Date le ottime prestazioni dimostrate dallo stimatore, è stato possibile procedere alla fase successiva del progetto.
Nella quinta fase lo stimatore è stato testato *off-line* utilizzando dati reali di impianto, sia per quanto concerne gli input del modello che per quanto riguarda le variabili di output i cui dati d'impianto risultano affidabili e disponibili in linea durante l'operazione della fornace. Il modello è stato reso più flessibile consentendo allo stimatore di aggiustare in tempo reale alcuni parameteri dello stesso (es. i fattori pre-esponenziali delle costanti cinetiche per la descrizione della velocità di *coking*).

Risultati I risultati ottenuti sono stati soddisfacenti sia in termini di capacità dello stimatore di identificare lo stato reale del sistema sia in termini computazionali. Le prestazioni sono risultate robuste per tutti e quattro i gruppi di dati d'impianto disponibili, mantenendo invariate le specifiche da fornire (matrice di covarianza per l'errore iniziale, matrice di covarianza per l'incertezza sul modello e matrice di covarianza per l'incertezza sulle misurazioni). La predizione di tutte le variabili di stato, sia quelle per le quali venivano fornite misurazioni in linea che quelle non misurabili, è risultata in ottimo accordo con i dati d'impianto disponibili. A livello computazionale, escluso lo step iniziale dove è richiesto un tempo di poco superiore al minuto, il tempo di predizione dello stimatore per simulare un'ora di operazione della fornace è risultato di circa 5 secondi. Tale tempistica risulta chiaramente compatibile con una sua possibile implementazione in linea.

I risultati ottenuti sono stati esposti ai responsabili tecnici del Cliente per il quale il modello della fornace è stato sviluppato. La qualità dei risultati, ed in particolare la corrispondenza fra la predizione dello stato di coking dello stimatore e quanto viene effettivamente osservato in impianto, ha convinto il Cliente a commissionare un nuova serie di test su un'altra fornace di un altro impianto di sua proprietà . Se le prestazioni dello stimatore verranno confermate anche in tale fornace, si procederà con la sua implementazione in linea.

Contents

Introduction **1**

1 Background **3**
1.1 Ethylene production . . . 3
1.2 Steam cracking of hydrocarbons . . . 4
1.2.1 Cracking Furnace . . . 6
1.2.2 Transfer-line heat exchanger . . . 6
1.2.3 Recovery section . . . 7
1.3 Coke formation . . . 8
1.4 Decoking . . . 9
1.5 Motivation . . . 9

2 The Furnace Model **11**
2.1 General model topology . . . 11
2.2 1D tube model . . . 12
2.2.1 Mass and energy balances . . . 14
2.2.2 Pressure drop correlation . . . 15
2.2.3 Heat transfer model . . . 16
2.2.4 Firebox model . . . 16
2.2.5 Cracking kinetics . . . 17
2.2.6 Coking kinetics . . . 18
2.3 USX model . . . 20
2.4 Downstream piping model . . . 20
2.5 Weaknesses of the Furnace model . . . 21

3 State estimation techniques **23**
3.1 Problem description for pure ODE systems . . . 23
3.2 The Extended Kalman Filter . . . 24
3.2.1 Extension to PDAE systems . . . 26
3.3 The Unscented Kalman Filter . . . 29
3.4 The Ensemble Kalman Filter . . . 32
3.5 The Particle filter . . . 34
3.6 Comparison and choice of the estimator . . . 37
3.7 Parameter Estimation using State Estimation . . . 37

4 Implementation of the state estimator **41**
4.1 The gPROMS® software . . . 41
4.1.1 Performing a simulation with gPROMS® . . . 41
4.2 Implementation issues . . . 42
4.2.1 Mathematical structure of the Furnace model . . . 43
4.2.2 Input-Output structure . . . 44
4.2.3 Plant data and reliability . . . 45
4.2.4 Model parameters . . . 47
4.2.5 Specifications to the estimator . . . 48

4.2.6 Update frequency and computational issues 50
4.3 Project organization 50
4.4 Need for a state estimator 51

5 Results with model-based data **55**
5.1 Overall procedure 55
5.1.1 Model-based generation of noisy measurements 56
5.1.2 Model modification 57
5.1.3 State estimation 57
5.2 Results 58
5.2.1 No real time adjustment of model parameters 58
5.2.2 Real-time adjustment of model parameters 63
5.3 Remarks 69

6 Results with real plant data **73**
6.1 Introduction 73
6.2 Model weaknesses and real-time adjustment 74
6.2.1 Firebox model adjustment 75
6.2.2 Coking model adjustment 76
6.3 General structure 79
6.4 Targets 80
6.5 Results for a pure ethane feed 81
6.5.1 Set #1 81
6.5.2 Set #2 84
6.6 Results for a pure propane feed 86
6.6.1 Set #3 89
6.6.2 Set #4 91
6.7 Conclusions 91

Conclusions **93**

A Codes **95**

Acknowledgments **103**

List of Figures

1.1 *Ratio of the domestic consumption to the global consumption of ethylene in Western Europe, the U.S. and BRIC countries. Adapted from Eramo (2013).* 4
1.2 *Principal uses of ethylene in the global market. Reference period: 2011-2012.* 4
1.3 *General topology of a steam cracker. Adapted from Zhang & Evans (2012).* 5
1.4 *Split radiant coil arrangement. Adapted from Zimmermann & Walzi (2009).* 6
1.5 *Simplified block flow diagram of the recovery section of an ethylene plant.* 7

2.1 *Topology of the Furnace model. The main units are the coil, the firebox, the TLE and the downstream pipe.* 12
2.2 *Schematic overview of the model of the radiant section of the Furnace. The overall structure is composed by a model of the firebox and a specific model of the coil . Source and sink models are used to describe the inlet and outlet streams. The material connections are represented in blue and the distributed thermal contact in red.* 12
2.3 *Coil topology. There are 6 different passes with different geometry; each pass is modelled with the Tube 1D model.* 13
2.4 *General structure of the Tube 1D model. Models marked with different colours belong to different libraries.* 14
2.5 *Heat-transfer resistances: bulk gas, coke layer and tube wall.* 16
2.6 *Simplified firebox model implemented in gCRACKER®.* 17

3.1 *A priori and a posteriori state estimates and estimation error covariances.* 25
3.2 *Extended Kalman Filter: graphical overview.* 27
3.3 *Unscented Kalman Filter: graphical overview.* 32
3.4 *Ensemble Kalman Filter: graphical overview.* 35

4.1 *Input-output structure of the Furnace model.* 44
4.2 *Hierarchical organization of the project. The red-path represents how the project evolved.* 51
4.3 *Comparison between the plant observations (black lines) and the model predictions (red lines) for some of the output variables of the Furnace model. Set of data #1.* 52

5.1 *Results obtained for some of the most important output variables. Red lines: wrong model prediction. Black lines: orignal model prediction. Blue lines: estimator prediction. Specifications given to the estimator:* $\mathbf{P}_0 = \mathbf{0}$, $\mathbf{Q}$=*diag(0.01),* $\mathbf{R}$ = *diag(1e-4; 1e-2;1e-6), WM: -2% on* $\log_{10} k_{c1}(T_{ref})$ *and +2 % on* $\log_{10} k_{c2}(T_{ref})$. 61
5.2 *Effect of the measurement error variance for pressure drop* ($r^2_{\Delta_p}$) *on the estimate of the total coke deposited in all passes. Results are not satisfactory even with very low values of* $r^2_{\Delta_p}$. 62

5.3 *Effect of the deviation on model parameters on the total coke deposited in all passes. The estimation of the state of coking is not satisfactory independently of the modification made on the coking parameters.* 62
5.4 *Real-time estimation of coking parameters. Black: OM nominal value. Red: WM modified value. Blue: estimator prediction.* 66
5.5 *Results obtained with real-time parameter adjustment for some of the most important output variables of the Furnace model. Red lines: wrong model prediction. Black lines: orignal model prediction. Blue lines: estimator prediction. Specifications given to the estimator:* $\mathbf{P}_0 = diag(0.05; 0.05; 0; \ldots; 0)$, $\mathbf{Q} = diag(0.05; 0.05; 0.01; \ldots; 0.01)$, $\mathbf{R} = diag(1e-4; 1e-2; 1e-6)$, *WM: +10% on* $\log_{10} k_{c1}(T_{ref})$ *and -10 % on* $\log_{10} k_{c2}(T_{ref})$. 67
5.6 *Estimation of the composition of the outlet stream exiting the coil. Red lines: wrong model prediction. Black lines: orignal model prediction. Blue lines: estimator prediction. Specifications given to the estimator:* $\mathbf{P}_0 = diag(0.05; 0.05; 0; \ldots; 0)$, $\mathbf{Q} = diag(0.05; 0.05; 0.01; \ldots; 0.01)$, $\mathbf{R} = diag(1e-4; 1e-2; 1e-6)$, *WM: +10% on* $\log_{10} k_{c1}(T_{ref})$ *and -10 % on* $\log_{10} k_{c2}(T_{ref})$. 68
5.7 *Results obtained using time-varying control inputs. Red lines: wrong model prediction. Black lines: orignal model prediction. Blue lines: estimator prediction. Specifications given to the estimator:* $\mathbf{P}_0 = diag(0.05; 0.05; 0; \ldots; 0)$, $\mathbf{Q} = diag(0.05; 0.05; 0.01; \ldots; 0.01)$, $\mathbf{R} = diag(1e-4; 1e-2; 1e-6)$, *WM: +10% on* $\log_{10} k_{c1}(T_{ref})$ *and -10 % on* $\log_{10} k_{c2}(T_{ref})$. 70
5.8 *Normalised time-varying supplied to the estimator. Set of data #1.* . . . 71

6.1 *Comparison between the OM approach and the new approach adopted using the EKF with real-time adjustment of the* COT_{offset}. 78
6.2 *General structure of the simulations performed using real plant measurements.* 80
6.3 *Results obtained with the first set of experimental data available. Black lines: plant data. Red lines: OM prediction. Blue lines: estimator prediction. Specifications like in eq.* (6.15), (6.16) *and* (6.17). 83
6.4 *Total coke deposited in all the 6 passes as predicted by the OM (red lines) and by the estimator (blue lines).* 85
6.5 *Computational time required by the estimator during each prediction and update step. The slowest step is the first prediction step.* 86
6.6 *Results obtained with the second set of experimental data available. Black lines: plant data. Red lines: OM prediction. Blue lines: estimator prediction. Specifications like in eq.* (6.15), (6.16) *and* (6.17). 87
6.7 *Total coke deposited in all the 6 passes as predicted by the OM (red lines) and by the estimator (blue lines). Set of data #2.* 88
6.8 *Results obtained with the thirdset of experimental data (propane feed). Black lines: plant data. Red lines: OM prediction. Blue lines: estimator prediction. Specifications like in eq.* (6.15), (6.16) *and* (6.17). 90

List of Tables

2.1 *List of the components involved in the cracking mechanism.* 14
2.2 *Kinetic model for cracking reactions. The model can be used with mixed feedstock (ethane+propane) of every composition.* 18

3.1 *Comparison between the different estimators considered.* 38

4.1 *Mathematical structure of the Furnace model* 43
4.2 *Model inputs to be collected in the control vector* $\mathbf{u}$. 45
4.3 *Most important outputs of the Furnace model.* 45
4.4 *Different sets of plant data available.* 46
4.5 *Measurements available per each data-set and their mathematical characterization in the Furnace model.* 46
4.6 *Reliability of the plant data available.* 47
4.7 *Most important parameters of the Furnace model* 47

5.1 *Time invariant control inputs used in the first runs of the estimator.* . . . 58
5.2 *Quality of the estimates obtained using different sets of measurements. The specifications for the estimator are the same in all cases:* $\mathbf{P}_0 = \mathbf{0}$*;* $\mathbf{Q} = diag(0.01)$*;* $\mathbf{R}$ *diagonal with* $r^2_{\Delta_p} = 10^{-4}$, $r^2_{TMT} = 10^{-3}$, $r^2_{w_{fuel}} = 10^{-6}$, $r^2_{x_{C_2H_6}} = 10^{-6}$*; -2% deviation on* $\log_{10} k_{c1}(T_{ref})$ *and +2 % on* $\log_{10} k_{c2}(T_{ref})$. 63

6.1 *Feed composition - pure ethane.* 81
6.2 *Comparison between the OM and estimator prediction of TMT with respect to the real plant observations. Set of data #1.* 82
6.3 *Comparison between the OM and estimator prediction of TMT with respect to the real plant observations. Set of data #2.* 86
6.4 *Feed composition for a pure propane feed.* 89
6.5 *From literature (Froment, 1992).* 89
6.6 *From historical data of another furnace.* 89
6.7 *Comparison between the OM and estimator prediction of TMT with respect to the real plant observations. Set of data #3.* 89
6.8 *Comparison between the OM and estimator prediction of TMT with respect to the real plant observations. Set of data #4.* 91

A.1 *gPROMS® process entities and case files related to the second chapter.* . 95
A.2 *MATLAB® codes used in the third chpater.* 96
A.3 *gPROMS® process entities and case files used in the fifth chapter (Part 1).* 97
A.4 *gPROMS® process entities and case files used in the fifth chapter (Part 2).* 98
A.5 *gPROMS® process entities and case files used in the fifth chapter (Part 3).* 99
A.6 *gPROMS® process entities and case files used in the fifth chapter (Part 4).* 100
A.7 *gPROMS® process entities and case files used in the sixth chapter.* . . . 101

Nomenclature

Acronyms

CIP Coil inlet pressure

COP Coil outlet pressure

NG Natural gas

NC Number of components

NR Number of cracking reaction

Nu Nusselt number

OM Original model

Pr Prandtl number

Re Reynolds number

WM Wrong model

Greek Symbols

ϵ Air excess

α Coil roughness parameter

λ_{coke} Coke thermal conductivity

$\rho^{coke\,layer}(z)$ Density of the coke layer

β Firebox model adjustment parameter

$\Delta\dot{h}_L$ Heat loss through the firebox walls

$\phi_{coke}(z)$ Thickness of coke deposited along the coil

Other Symbols

$c_{cat}(z)$ Concentration of catalitically active sites along the coil

$\mathbf{u}$ Control input vector

$\hat{H}(z)$ Convective heat flux inside the tube

$A(z)$ Cross sectional area of the tube

$K_{eq,i}$ Equilibrium constant of reaction i

$\hat{\mathbf{z}}$ Estimate of state vector $\mathbf{z}$

$q_{ext}(z)$ External heat flux profile

R_0 External tube radius

U Global heat transfer coefficient

$\tilde{T}_i^{in/out}$ Guessed inlet/output temperature of i

$\mathbf{P}_0$ Initial error or *surrogate* covariance

R_w Internal radius of the tube

h_{jacket} Jacket-side heat transfer coefficient

$\mathbf{K}$ Kalman gain

I Least squares functional

c_{cat}^{max} Maximum concentration of catalitically active sites

$\mathbf{v}$ Measurement noise

$\mathbf{R}$ Measurement uncertainty covariance

$\mathbf{y}$ Measurement vector

$\mathbf{Q}$ Model uncertainty covariance

$\mathbf{w}$ Model uncertainty

$F_i(z)$ Molar flowrate of component i

$N_i(z)$ Molar flux of component i

$h(z)$ Molar specific enthalpy

$R_I(z)$ Radius at the interface between the coke layer and the gas

$r_{catalytic}(z)$ Rate of catalytic coking

$r_{pyrolytic}(z)$ Rate of pyrolytic coking

$\mathbf{x}$ State vector

T_w Temperature at the tube wall

$T_{gas\,phase}(z)$ Temperature of the gas phase inside the coil

$\dot{Q}_R$ Total heat power generated in the firebox

$\dot{Q}_I$ Total heat power input to the coil

w_{coke} Total mass of coke deposited

p Vector of model parameters

$v(z)$ Velocity of the gas inside the coil

$\overleftarrow{k}_i$ *Backward* rate of reaction i

$\overrightarrow{k}_i$ *Forward* rate of reaction i

Introduction

State estimation is a widespread and well-established technique in control engineering and weather forecasting. However, its usage combined with typical process simulation activities, like whole-plant optimization, is not so common, especially in the chemical and petrochemical industry. Steady-state process simulation and steady-state whole plant optimization are the most common tools used by chemical-petrochemical manufacturers to optimize their processes and to plan their production activities.

In recent years, with the spread of dynamic process simulators, state estimation applied to typical chemical engineering problems has regained attention. It is well known, in fact, that a suitable design of a state estimator requires a representative dynamic model to capture the plant behavior. To date, however, state estimation has always been applied to relative simple chemical engineering-related models (ODE/DAE systems with a small number of state variables) because of the computational effort required when a large number of state variables is involved and because of the numerous numerical instability problems that can arise during its implementation.

An example of a chemical engineering-related problem for which the usage of a state estimator can be extremely beneficial is the determination of the state of coking of a cracking furnace.
Coking is a well-known phenomenon that occurs during the operation of a steam cracker and deeply influences the key performance indicators (KPI: yield, conversion) of the cracker. Coke slowly deposits on the internal walls of the coil, thus reducing the residence time of reactants and increasing the pressure drop, up to a point at which the coke build-up is such that the furnace must be shut down for decoking. Therefore, coking is an intrinsically dynamic phenomenon that can theoretically be modeled once a deep understanding of its mechanism has been obtained. However, because of its complexity and because of the enormous number of variables that can affect it, obtaining a reliable estimation of the state of coking of a steam cracker using a first-principles modelling approach is non-feasible and, in most cases, not reliable. In view of this, a state estimator, given a few of on-line measurements coming from the furnace, can be used to update the state of the system and improve the prediction of the state of coking and of all the other important state variables. Once a reliable estimate of the state of coking is obtained, this can be used to perform a cyclic whole-plant optimization to maximize profits and increasing the whole-plant productivity.

The aim of this study is the implementation of a state estimation technique (EKF, Extended Kalman Filter) on a preexisting dynamic model of a steam cracker, using the dynamic process simulator gPROMS®. The model considered is a highly nonlinear partial differential-algebraic system (PDAE) with 227 differential and 14268 algebraic variables and belongs to a whole-plant model developed for an important ethylene manufacturer. What is expected is to analyze the possibility of implementing this estimator on-line in the real plant by testing it off-line using available historical data of the furnace. A reliable estimate of the state of coking of the furnace is expected to be obtained in order to perform a cyclic whole-plant optimization without assuming clean-tube conditions.

The structure of the thesis is as follows.
In the first chapter, a brief overview of the steam cracking process is given. The main unit operations, with a particular focus on the cracking furnace, are quickly reviewed and

the problem of coking is introduced.
In the second chapter, the gPROMS® model of the steam cracker analyzed in this study is presented. The main input and output variables, the main parameters, the main equations and the main assumptions involved in the model are explained. Particular emphasis is given to the relevant weaknesses of the model and to how they can affect its prediction.
In the third chapter, an exhaustive introduction to the mathematical formalism behind the most important state-estimation algorithms available is performed, and their pro's and con's analyzed. The reason why the Extended Kalman Filter has been chosen as a first option in this study is explained.
In the fourth chapter, the implementation issues that need to be faced when implementing the state estimator in gPROMS® are discussed. The hierarchical approach used to test the estimator is presented and the mismatch between the original model prediction and the available plant observations is considered.
In the fifth chapter, the results obtained by testing the estimator with model-based measurements are shown. The need for a real-time adjustment of some of the model parameters is explained and the effect of the specifications given to the estimator (initial/model/measurement uncertainty covariances) on its performance is analyzed.
Finally, in the sixth chapter the results obtained by testing the estimator with real plant data are presented. The actions that have been taken to fix the model weaknesses and to ease the effort of the estimator are explained. The computational performance of the estimator is considered and its compatibility with an on-line implementation is discussed.

Chapter 1

Background

In this chapter, a brief overview of the pyrolysis of hydrocarbons for the production of ethylene is presented. First, a general introduction of the global market of ethylene is given. Then, a brief description of the main process for ethylene production (steam cracking of hydrocarbons) is carried out. The most important unit operations of this process are discussed and briefly reviewed. Finally, the problem of coking on the industrial scale is presented and the decoking strategies adopted are reviewed.

1.1 Ethylene production

Ethylene is a large volume petrochemical and one of the most important building blocks in the chemical industry (Zimmermann & Walzi, 2009). In 2013, the ethylene capacity in the world was approximately 150 Mton/y with an average yearly growing capacity of 3.5% (Laugier, 2013). This continuous increase is mainly due to existing plant expansions, but also new grassroots plants have been built, especially in the Middle East, where profits are higher due to the lower feedstock costs. The expected global capacity in 2020 is 200 Mton/y .

In the last few years, the shale gas boom in the US has considerably affected the ethylene market (Eramo, 2013). The strong decline of prices for ethane has led to a number of new ethane crackers being built in the US. Since ethane gives higher ethylene yields than naphta, which is the main feedstock for the European ethylene production, several European manifactures have announced to either close their crackers or to change over to using imported ethane as feedstock. In Fig. 1.1 the ratio of the domestic consumption to the global consumption of ethylene in Western Europe, the U.S. and BRIC countries[1] is compared. It is clear that the growth of ethylene demand in the next years will be localized in the emerging economies.

The leading technology for the production of ethylene from ethane, propane or naphta is steam cracking, a high temperature pyrolysis in the presence of steam. Alternative technologies, such as the dehydratation of ethanol or the production of ethylene from methanol, have been studied, but none of them exhibits the economics to be a challenge for the well-established steam-cracking process.

In Fig. 1.2 the principal uses of ethylene in the global market are presented. As can be noticed, more than half of the production of ethylene is used for the production of polyethylene; other applications are for the production of monoethylene gycol (MEG), ethylene oxide (EO) and 1-2 dichloroethane (EDC).

In recent times, since the profitability of an ethylene plant is strongly related to productivity and the cost for product manifacturing, sophisticated model-based approaches have been developed to optimise the plant operation, thus reducing the final cost of the product. Process Systems Engineering (PSE) techniques have been deployed, especially

[1] Brazil, Russia, India and China.

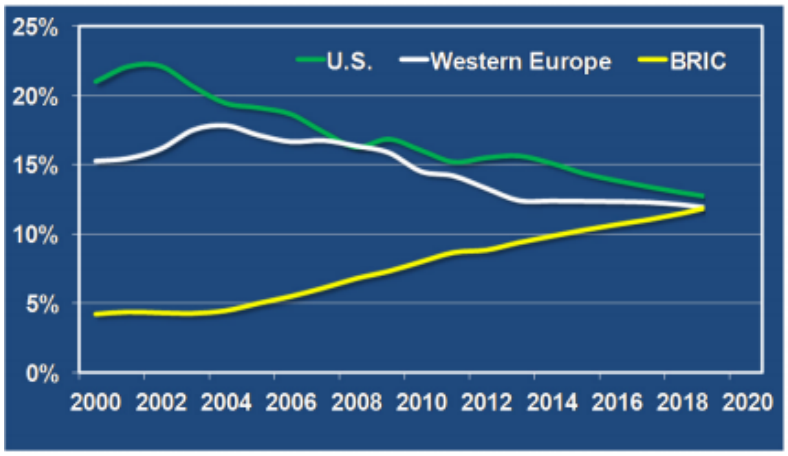


Figure 1.1: *Ratio of the domestic consumption to the global consumption of ethylene in Western Europe, the U.S. and BRIC countries. Adapted from Eramo (2013).*

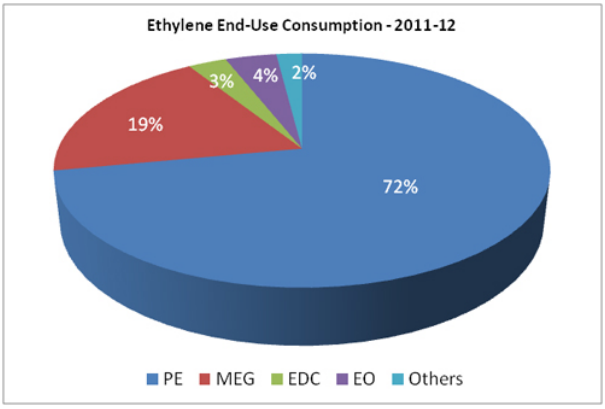


Figure 1.2: *Principal uses of ethylene in the global market. Reference period: 2011- 2012.*

for the optimization of the operation of the steam cracker, which is the *core* of an ethylene plant[2].

1.2 Steam cracking of hydrocarbons

Steam cracking of hydrocarbons is the most important route to produce ethylene in the world (Zimmermann & Walzi, 2009). Different types of feedstock can be used: in the U.S the most common are ethane and propane, while in Europe naphta is the preferred one (Zimmermann & Walzi, 2009).

The general arrangement of a cracking reactor is presented in Fig.1.3. The furnace can be divided into three different sub-units: the convection section, where a hydrocarbon stream is heated by heat exchange against flue gas and mixed with steam, the radiant section, a fired tubular reactor (*radiant coil*) where cracking reactions occur, and the transfer line heat exchanger (TLE), where the exit stream is quenched to prevent degradation of the highly reactive products by secondary reactions. This cooling is carried out by vaporizing high-pressure boiler feed water (BFW), which is separated in the steam drum and then superheated in the convection section of the furnace.

[2]The software SPYRO® is an example of the application of such techniques.

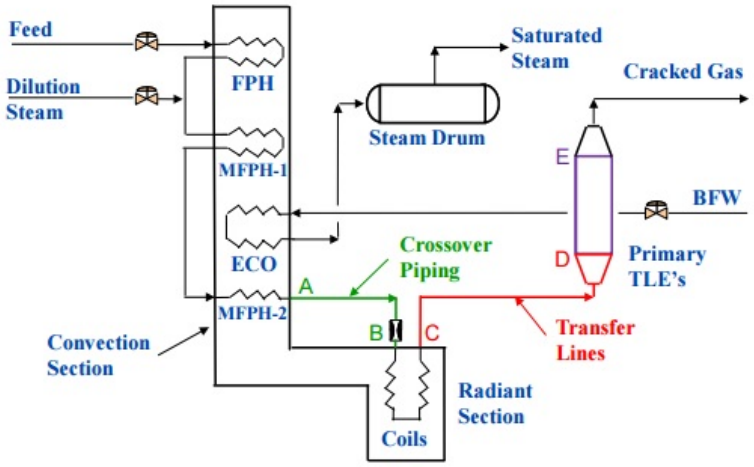


Figure 1.3: *General topology of a steam cracker. Adapted from Zhang & Evans (2012).*

The cracking of the feedstock that occurs in the radiant coil can be described by a complex network of radical reactions. However, for the pyrolysis of simple compounds (ethane/propane), simplified molecular schemes can be applied with some success in order to describe the product distribution. These molecular schemes are usually composed by no more than 20 reactions and the kinetic parameters are estimated using available experimental data of product composition.

Steam is mixed in the stream of hydrocarbons in order to accomplish these three effects:

i) a strict control of the temperature inside the coil is obtained. Temperature in the coil is a key-factor to maximize the yield of ethylene production;

ii) dilution of reactants: by decreasing their partial pressures, the selectivity towards primary products (ethylene/propylene) increases;

iii) reduction of the coking rate. Steam reacts with coke forming gaseous products such as CO and CO_2 (in a limited extent), thus reducing the deposition of coke on the tube wall.

The furnace design is strictly related to the residence time of the reactants in the coil, the temperature profile along the coil and the partial pressure of hydrocarbons in the main stream. Their influence can be explained as follows:

- **a) Residence time:** long residence times favor the secondary reactions, thus decreasing the selectivity towards ethylene. Typical residence times in the coil can vary between 0.1 and 0.5 seconds, according to the feedstock composition.

- **b) Partial pressure of reactants**: decreasing the partial pressure of reactants results in a higher selectivity towards the primary products, as previously discussed. Steam is used to decrease the partial pressure of reactants in order to enhance selectivity.

- **Temperature profile**: secondary reactions are favored by lower temperatures. Long residence times at low temperature must be avoided. A typical inlet temperature for the coil (CIT) is between 500-50 degC and products leave the coil at 800-850 degC (COT: coil outlet temperature).

A brief description of the Furnace, the transfer line heat exchanger and the separation section is presented in the next sections.

1.2.1 Cracking Furnace

Different types of commercial pyrolysis furnace exist; they typically have one or two rectangular fireboxes with vertical or horizontal radiant coils located between two radiant refractory walls. Heat is transferred mainly by radiation using wall- or floor-mounted radiant burners, which typically use natural gas as fuel. If short residence times are required, many more individual coils must be used with respect to longer residence times for the same production capacity. The number of coils required for a given ethylene capacity is determined by the radiant coil surface.

The radiant coil can have different configurations. However, most of these configurations have two features in common: coils are hung vertically and they are both fired from both sides of the radiant coil. One of the most common coil configuration is the so called split radiant coil, presented in Fig.1.4. Parallel small-diameter coils are combined into a larger diameter outlet coil. Each pass of the larger coil can have a different internal diameter with respect to the other ones. With this arrangement an advantageous temperature profile is obtained, with a rapid temperature increase at the inlet of the coil which allows higher yields than a uniform diameter coil.

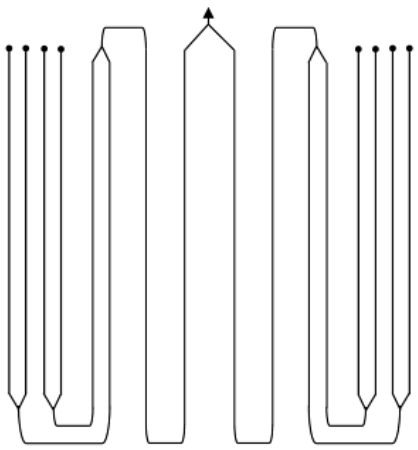


Figure 1.4: *Split radiant coil arrangement. Adapted from Zimmermann & Walzi (2009).*

1.2.2 Transfer-line heat exchanger

The cracked gas leave the coil at 800-850 degC: a rapid reduction of the gas temperature is required to avoid losses of valuable products by secondary reactions. The quenching of the cracked gas is accomplished by a transfer-line heat exchanger (TLE), which allows the recovery of the heat released by the cracked gas by vaporizing high

pressure boiling feed water (BFW). The TLE can be a conventional or a linear heat exchanger[3]; it is usually mounted on the top of the furnace since in modern designs the radiant coils exit the firebox at the top. As with the coils, the deposition of coke on the TLE walls requires a periodic *decoking* procedure simultaneously with the furnace *decoking*.

1.2.3 Recovery section

The temperature of the cracked gas leaving the TLE can vary between 300 and 420 degC, according to the type of feedstock. In a typical plant, different cracking furnaces work in parallel and all the cracked gases leaving each furnace are collected in a large cracked gas line, cooled to 200 degC in appropriate heat exchangers (oil quenching is performed with heavy feedstock) and then sent to the recovery section. This section consists in the removal of the heat contained in the cracked gas, the condensation of water and heavy hydrocarbons, the compression of the gaseous products, the separation of the relevant products using high-pressure distillation columns and the the hydrogenation of certain multiple unsaturated components. A simplified block flow diagram of this section is shown in Fig.1.5.

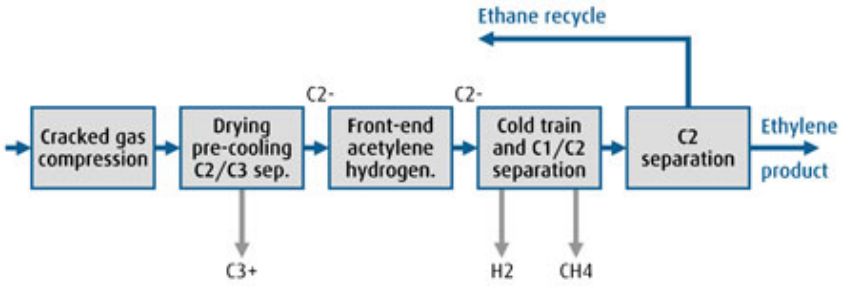


Figure 1.5: *Simplified block flow diagram of the recovery section of an ethylene plant.*

Apart from ethylene, the co-products that can be obtained from the cracking reactions are:

- *propylene*: negligible if pure ethane is used as a feedstock. It is usually recovered and sold like ethylene.
- *C_4 fraction*: this can be refined for butadiene, butene, isobutene or hydrogenated and recycled to cracking.
- *Hydrogen*: it is usually recovered and used for the hydrogenation steps involved in the plant. The hydrogen in excess is used as a fuel.
- *Methane*: it is usually recovered and used as a fuel for the Furnace.
- *Others*: they can include aromatics, acetylenes, C_5 olefins, naphtalene, tars.

The variety of products that can be obtained makes the separation particularly difficult and often most of the secondary products are not recovered since their separation costs are higher than the profits that can be obtained from their sale.

[3]This last solution is usually preferred because it avoids erosion problems due to coke deposition.

1.3 Coke formation

The pyrolysis of hydrocarbons produces different compounds (acetylenic, diolefinic, aromatic) which act as coke precursors on the inside surface of the radiant coil. The deposition of coke on the tube wall has two main impacts on the furnace operation:

i) inhibition of heat transfer from the tube to the process gas, with a consequent increase of the tube wall temperature (TMT);

ii) reduction of the cross sectional area available for the gas flow. This causes:

   a) an increase of the pressure drop along the radiant coil;

   b) a decrease of the cracking yield caused by the decrease of the residence time of the reactants inside the coil.

When temperature limits for the tube wall are reached, the furnace must be shut down and a decoking procedure (strict-controlled combustion with a mixture of steam and air of the deposited coke) must be carried out. The period of operation of the furnace before decoking is the so called *run-length* of the furnace. As suggested by Goosens *et al.* (1978), the coke deposition on the tube wall can lead to an increase of the fuel consumption up to 5%.

There is no general agreement on the mechanism of coke deposition. However, it is now well accepted that *at least* two mechanisms are involved during the furnace operation: a catalytic mechanism, due to the presence of iron and nickel on the tube skin that catalyze the coking reactions, and a pyrolytic mechanism, i.e. a non-catalytic coke deposition due to the presence of coke precursors (especially aromatic compounds) in the gas phase. The presence of these two different mechanisms can explain the initial sustained increase of pressure drop along the coil (due to catalytic coking) that is observed in real plants. Wauters & Marin (2002) tried to obtain a coherent kinetic scheme to describe pyrolytic coking. They tried to describe the kinetics of the surface reactions using prototype gas-phase reactions and taking into account the solid phase using a correction factor derived from collision theory, and the kinetics of the gas-phase using the group-contribution theory. However, their work was strongly criticized by Albright (Albright, 2002). Unless the mechanism of pyrolytic coking is a radical one, simple molecular mechanisms, particularly effective for gas-phase feedstock, have been proposed by Sundaram & Froment (1977) and refined by Froment (1992). To reduce catalytic coking, sulfur-based components (typically DMDS) are added to the main stream. However, the effect of sulfur on coke deposition is complex and high amounts can eventually lead to an increase instead of a decrease of coke formation (Wang *et al.*, 2007).

Coke deposition is not confined only to the steam cracker. In fact, coke also accumulates on the walls of the transfer line heat exchanger (TLE), to such an extent that the equipment must be cleaned. The mechanisms involved in the TLE for coke formation are different from the ones involved in the furnace because of the different conditions in the two equipments. In general, the mechanism of coke deposition in the TLE is even less understood than in the furnace (Zimmermann & Walzi, 2009).

1.4 Decoking

As already discussed, due to the deposition of coke on the inside surface of the radiant coil, when tube temperature limits are reached the furnace must be shut down to remove the coke deposited. The procedure adopted to remove this coke is called *decoking*.

Since the mechanical removal of coke is usually not feasible (Zimmermann & Walzi, 2009), *decoking* is carried out by burning out the coke with a mixture of air and steam. The furnace is first taken off-line and the residual hydrocarbons are purged downstream using steam. Then, a mixture of steam and air is introduced in the coil to burn out the coke. The air flowrate is gradually increased to avoid the overheating of the coil. Usually, the CO_2 produced by the combustion is continuously monitored and the air flowrate is adjusted accordingly. The *decoking* of the radiant coil takes approximately 20 hours. Once the radiant coil is clean, the removal of the coke deposited in the TLE is carried out. In modern plants, the off-gases leaving the radiant coil during *decoking* are directly sent to the TLE which is then decoked as a result. However, it is always necessary, at least once a year, to perform a mechanical cleaning of the heat exchanger(s).

An increase of the furnace length means higher productivity and, consequently, higher profits. Because of this, the optimization of the operating conditions of the furnace plays a key role to increase the general process profitability.

1.5 Motivation

The aim of this Thesis is the implementation of a state estimation technique on a pre-existing dynamic model of a steam cracker, using the dynamic process simulator gPROMS®. The model belongs to a whole model of an ethylene plant developed by Process Systems Enterpise Ltd.. The objective is to obtain an on-line robust and reliable estimate of the state of the furnace (and particularly of its state of coking). The estimator, using few and reliable measurements coming from the real plant, should be able to reconcile the model prediction with the available plant observations.

There is at least one immediate application and several possible future applications that can be found for the state estimator. As regards its immediate application, it should be considered that, as just said, the model of the steam cracker belongs to an overall preexisting model of an ethylene plant. In this model, three different furnaces run in parallel [4], and all the complex network of unit operations involved in the separation section (including recycles) is modelled using gPROMS® flowsheeting capability. This model is currently used to perform a whole-plant optimization in order to improve the plant's productivity and its profits; however, this optimization is carried out considering clean-tube conditions in the three furnaces since a reliable estimate of their state of coking is not available. Clearly, the results obtained from this optimization cannot be considered as accurate: in fact, the state of coking of the furnace deeply influences its key performance indicators (yield, conversion), thus modifying the optimal conditions to maximize its productivity. The implementation of the state estimator should hopefully give a real-time estimate of the state of coking of the furnace and this could help to perform a cyclic whole-plant optimization removing the assumption of clean-tube condi-

[4] The three furnaces are all modeled using the same model that will be presented in chapter §2: what changes is just the feedstock (1 works with pure ethane, 1 with pure propane and 1 with mixed feed).

tions. From the other hand, the possible future applications of the state estimator are enormous: between these, the usage of the state estimator to develop a model predictive control (MPC) strategy for the furnace and its usage for production planning purposes seem the most promising.

Chapter 2

The Furnace Model

As discussed in the previous chapter, the most important section of an ethylene plant is the cracking furnace. The optimization of its operation is then crucial to improve the overall process profitability. - In this chapter, a detailed description of a gPROMS® model of a steam cracker is presented. The model belongs to a pre-existing detailed model of an ethylene plant developed by Process Systems Enterprise Ltd. for a world's leading ethylene manufacturer.

First, a general overview of the model is given. Then, the model for the coil of the furnace is analyzed. The model of the firebox is then discussed, and the modelling of coke formation is considered. Finally, the most important weaknesses and assumptions of the model are discussed and their influence on model prediction is presented.

2.1 General model topology

The general structure of a typical cracking furnace was presented in the last chapter (Fig.1.3). As already discussed, the furnace is composed by a convective section, where the stream of hydrocarbons is pre-heated and mixed with steam, and a radiant section, where the gas stream is cracked inside a fired tubular reactor (radiant coil). A transfer-line heat exchanger is used to quench the reaction products in order to prevent their degradation by secondary reactions.

Building a model of such a structure is not an easy task. In Fig. 2.1 the general topology of the model of the steam cracker investigated in this study is shown.

The feed is first mixed with steam using a mixer: the feedstock that can be specified by the user are pure ethane (ID=0), pure propane (ID=1) or mixed feed (ethane+propane, ID=2). The convective section is modelled using a simple heat exchanger: once the exit temperature of the main stream is specified, the heat duty can be computed and so the amount of boiling feed water that can be vaporized in the real plant. The radiant section is modelled using a specific model for the coil, a specific model for the firebox and the related sink/source material models. Heat is exchanged between the two units as graphically shown in Fig.2.2.

The model of the coil is case-dependent according to the type of furnace installed in the real plant. In this study, the coil is a split radiant coil as the one presented in Fig.1.4. It is modelled using six different passes of different internal diameters, with two additional passes used to represent the entrance and the exit of the gas stream. Each pass is modelled using a 1-dimensional distributed model, assuming a plug flow for the gas stream with uniform temperature and composition across the tube cross-section. Each section of the coil receives an external energy input supplied by the firebox. The general topology of the coil is presented in Fig.2.3.

The TLE is modelled using the same 1D tube model of the coil, coupled with a simple cooling jacket model to describe the gas stream cooling. Cracking reactions and coking are not modelled in the TLE. Finally, the downstream piping is modelled using the same

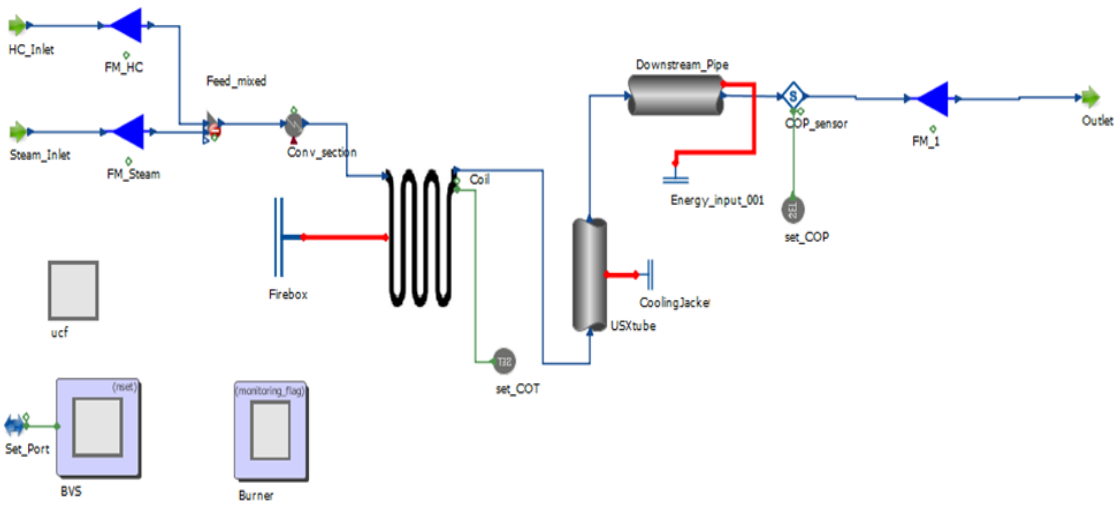


Figure 2.1: *Topology of the Furnace model. The main units are the coil, the firebox, the TLE and the downstream pipe.*

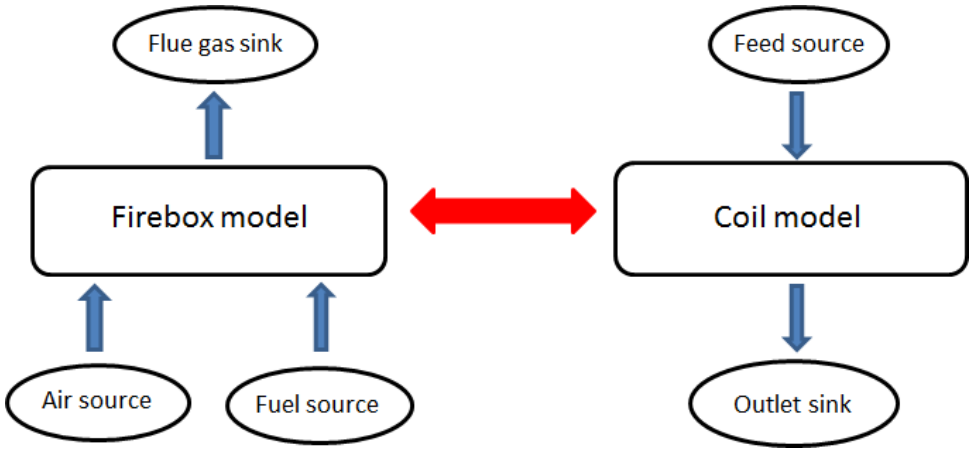


Figure 2.2: *Schematic overview of the model of the radiant section of the Furnace. The overall structure is composed by a model of the firebox and a specific model of the coil . Source and sink models are used to describe the inlet and outlet streams. The material connections are represented in blue and the distributed thermal contact in red.*

Tube 1D model of the coil.

In the next few sections the main equations and the main assumptions of the models previously described are discussed.

2.2 1D tube model

The one-dimensional tube model is the key-model for the description of the furnace operation. This model is used to represent each pass of the coil, the transfer- line heat exchanger (coupled with a cooling jacket model) and the downstream piping. It assumes plug flow of the gas within the tubes, with no variation of temperature/composition

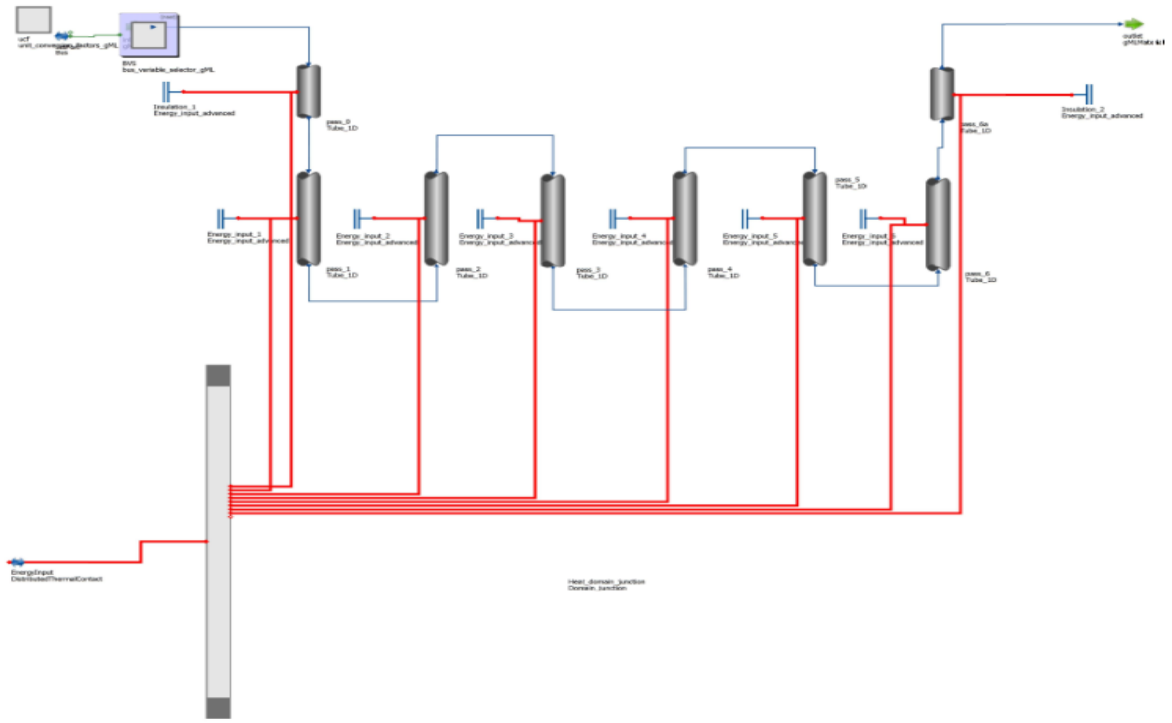


Figure 2.3: *Coil topology. There are 6 different passes with different geometry; each pass is modelled with the Tube 1D model.*

along the radial direction. It will be shown that, given the following (time-invariant or time-varying) inputs:

1. mass flowrate, temperature and composition of the inlet gas;
2. outlet pressure;
3. radiative heat flux profile received by the gas stream along the tube section,

this model is able to compute the following output variables:

1. inlet pressure and outlet composition and temperature;
2. wall temperatures along the tube and coke deposition.

The general structure of the Tube 1D model is shown in Fig.2.4. The most important equations implemented in the main model are the mass balances, the energy balance and the pressure drop correlation. Three sub-models are linked to the main model: one describes the cracking kinetics, another one is used to describe the coking kinetics and in the last one the heat transfer coefficient correlation is implemented. The main assumption of the model is that, since coke deposition occurs on a much longer time-scale than cracking reactions, a dynamic model should be used for the former and a steady-state model should be sufficient for the latter. Thus, coking is intrinsically considered as a dynamic process. The build-up of coke inside the tubes over time affects the key performance indicators of the system (yields, conversion): its prediction is crucial to optimize the furnace operation. The most important equations involved in the main model and in the three submodels are discussed in the next few sections.

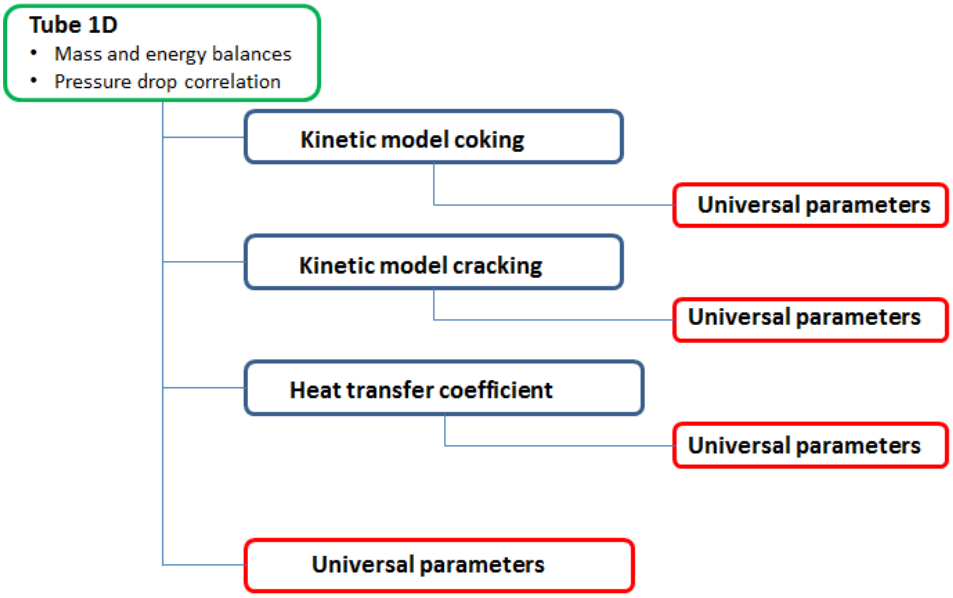


Figure 2.4: *General structure of the Tube 1D model. Models marked with different colours belong to different libraries.*

2.2.1 Mass and energy balances

The components involved in the cracking mechanism (§2.2.5) are reported in Tab.2.1.

Table 2.1: *List of the components involved in the cracking mechanism.*

Component #	Component	Type
1	H_2	Byproduct
2	CH_4	Byproduct
3	C_2H_2	Byproduct
4	C_2H_4	Main product
5	C_2H_6	Reactant
6	C_3H_6	Byproduct
7	C_3H_8	Reactant
8	C_4H_6	Byproduct
9	C_4H_8	Byproduct
10	$C_{5,plus}$	Byproducts

The steady-state mass balance for each component is given by:

$$\frac{\partial N_i(z)}{\partial z} = \sum_{j=1}^{NR} \nu_{i,j} r_{cracking,j}(z), \quad i = 1, \ldots, NC, z \in (0, L] \tag{2.1}$$

with boundary conditions:

$$N_i(0)A(0) = F_i. \tag{2.2}$$

The steady-state energy balance equation for the gas phase is given by:

$$\frac{\partial \left(\hat{H}(z)A(z) \right)}{\partial z} = 2\pi R_0 q_{ext}, \quad z \in (0, L] \tag{2.3}$$

where $\hat{H}(z)$ is the convective heat flux along the tube defined by:

$$\hat{H}(z) = \frac{F(z)h(z)}{A(z)}. \tag{2.4}$$

The following boundary condition is used for eq. (2.3):

$$T_{gas\,phase}(0) = T^{inlet}_{gas\,phase}. \tag{2.5}$$

The external heat flux profile along the coil $q_{ext}(z)$ that appears in eq.(2.3) should theoretically be obtained once an accurate model of the firebox is available. However, as will be discussed in §2.2.4, such a detailed model is difficult to build and only a simplified model, which is not able to describe the heat flux profile along the external skin of the coil, can be derived. Because of this, in the Furnace model an *a priori* heat flux profile, derived from the knowledge of the coil and firebox configurations in the real plant, is used to describe the external heat flux.

In a general form, the *a priori* heat flux profile can be described by:

$$\frac{q_{ext,i}(z)}{q_{max,i}} = f(z) \quad z \in [0, L_i] \tag{2.6}$$

where L_i is the i-th pass of the coil and $q_{max,i}$ is the maximum heat flux to the i-th pass. Each pass of the coil is exposed to the same heat flux profile and it can be noticed that the heat flux reaches its maximum at the centre of each pass. Coke deposition is then expected to be at its maximum in this point, since higher temperatures (and then coking rates) are involved.

2.2.2 Pressure drop correlation

The pressure drop along each pass (tube section) of the coil is calculated using a momentum balance equation according to the work of Froment & Bischoff (1980):

$$A(z)\frac{\partial p(z)}{\partial z} = F(z)\frac{\partial v(z)}{\partial z} - A(z)\rho(z)v(z)\left[\frac{2}{R_I(z)}f_{tube}(z) + \frac{NB}{L}f_{bend}(z)\right], \quad z \in [0, L) \tag{2.7}$$

with the boundary condition at the outlet:

$$p(L) = p^{outlet}. \tag{2.8}$$

From eq. (2.7) it can be noticed that both the straight part of the coil and the bends contribute to the pressure drop along the coil. The contribution of the bends is taken into account implicitly considering it as distributed along the entire tube section. The loss of accuracy due to this simple approach is minimal and the friction factors in eq. (2.7) for the straight tube contribution (R.H & D.W., 2009) and the bend contribution (Nekrasov, 1969) can be calculated according to the following equations:

$$\frac{1}{\sqrt{4f_{tube}}} = -2\log\left[\frac{\frac{K(z)}{2R_I(z)}}{3.71} + \left(\frac{7}{Re}\right)^{0.9}\right],\ K(z) = K_{clean\,tube} + \alpha r_{coking}(z) \tag{2.9}$$

$$f_{bend}(z) = \left(0.7 + \frac{\alpha}{90}\cdot 0.35\right)\left(0.051 + 0.19\frac{2R_I(z)}{R_B}\right), \quad z \in [0, L]. \tag{2.10}$$

Eq. 2.7 can be used to build the pressure profile along each pass of the coil, given the pressure at the coil outlet (COP), allowing to compute the coil inlet pressure (CIP) and so the pressure drop.

2.2.3 Heat transfer model

Heat is transferred from the firebox to the process gas inside the tubes via a series of resistances, as shown in Fig. 2.5. No energy hold up is assumed in the tube wall/coke

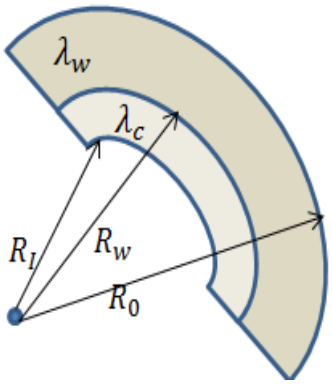


Figure 2.5: *Heat-transfer resistances: bulk gas, coke layer and tube wall.*

layers, so that the temperature at various interfaces across the wall can be obtained from equations for steady-state heat conduction through a material of constant heat conductivity:

$$q_{ext}R_0 = U(z)R_I(z)\left[T_I(z) - T_G(z)\right] = \frac{\lambda_c\left[T_w(z) - T_I(z)\right]}{\log\left(\frac{R_w}{R_I}\right)} = \frac{\lambda_w\left[T_0 - T_w(z)\right]}{\log\left(\frac{R_0}{R_w}\right)}. \tag{2.11}$$

The heat transfer coefficient is calculated according to the adimensional correlation suggested by Sundaram & Froment (1980):

$$Nu(z) = 2.43 \cdot 10^{-2} Re(z)^{0.8} Pr(z)^{0.4} + \frac{429.2}{l(z)} - \frac{544.3}{\left(\frac{X_{\%}(z)}{100}\right)^{0.0437}}, \quad z \in [0, L] \tag{2.12}$$

where $l(z)$ is a dimensionless length and is calculated according to the equation:

$$l(z) = \frac{2z}{Re(z)Pr(z)R_I(z)}, \quad z \in [0, L] \tag{2.13}$$

while the heat transfer coefficient $U(z)$ is obtained using the definition of the Nusselt number:

$$U(z) = \frac{Nu(z)\lambda_G(z)}{2R_I(z)}, \quad z \in [0, L]. \tag{2.14}$$

The thermal conductivities for the coke layer λ_c and the wall λ_w which appear in eq.(2.11) can be directly specified by the user.

2.2.4 Firebox model

As already discussed, an accurate model of the firebox should be able to describe the heat radiation to the coil, so that the external heat flux profile $q_{ext}(z)$ that appears in the previous equations could be obtained. However, building a complex firebox model is a difficult task which requires a lot of assumptions that can deeply affect the model

prediction. For this reason, in this study only a simplified firebox model is implemented, coupled with an *a priori-known* external heat flux profile. This simplified model allows to predict the fuel consumption, which is an important plant measurement that can be used to validate the model. The firebox is modelled as a conversion reactor (Fig. 2.6): the fuel considered is natural gas (CH_4 + H_2) with fixed composition. It is assumed the complete combustion of methane and hydrogen according to the two reactions:

$$CH_4 + 2\,O_2 \longrightarrow CO_2 + 2\,H_2O \tag{2.15}$$

$$H_2 + \frac{1}{2}\,O_2 \longrightarrow H_2O \tag{2.16}$$

with no external heat loss through the firebox wall. The air excess is fixed and can be changed by the user.

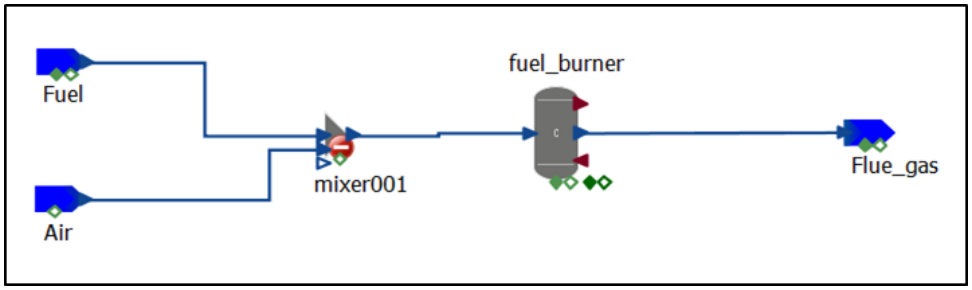


Figure 2.6: *Simplified firebox model implemented in gCRACKER®.*

Given the previous assumptions, the energy balance for the system can be written, in a general form, as:

$$F^{in}_{fuel} h^{in}_{fuel}\left(T^{in}_{fuel}\right) + F^{in}_{air} h^{in}_{air}\left(T^{in}_{air}\right) - F^{out}_{flue\,gas} h^{out}_{flue\,gas}\left(T^{out}_{flue\,gas}\right) = \dot{Q}_R \tag{2.17}$$

where $\dot{Q}_R$ is the total heat power generated by the two combustion reactions (2.15) and (2.16). $\dot{Q}_R$ is set equal to the total energy input to the coil, calculated according to the equation:

$$\dot{Q}_I = \sum_{i}^{all\,passes} \left(\int_0^1 2\pi R_{0,i} q_{ext}(z) \mathrm{d}z \right) L_i \tag{2.18}$$

where $R_{0,i}$ is the external radius of the i-th pass and L_i its total length. Once T^{in}_{fuel}, T^{in}_{air}, the air excess, and $T^{out}_{flue\,gas}$ are specified, eq. (2.17) can be used to predict the fuel consumption. The model prediction of fuel consumption can then be compared with the available plant data to validate the model prediction.

2.2.5 Cracking kinetics

The pyrolysis of hydrocarbons is a radical-based mechanism. Different radical intermediates are formed according to a complex radical scheme, which usually involves hundreds or thousands of reactions.

The implementation of a radical scheme on a pre-existing model is usually a modelling and computational burden that can, in some cases, be avoided by using some simple molecular schemes that can substantially reduce the model complexity. These

mechanisms usually fail to predict the by-products composition, but accurate prediction of the yield of reaction and of the conversion of the main reactant can be obtained.

The kinetic model used in this study is based on a simple molecular scheme first proposed by Sundaram & Froment (1977). Some improvements to the original mechanism have been added to make it suitable to a mixed feed (ethane + propane) of every composition, including pure ethane or pure propane. The full kinetic model is presented in Tab. 2.2.

Table 2.2: *Kinetic model for cracking reactions. The model can be used with mixed feedstock (ethane+propane) of every composition.*

Reaction #	Reaction	Reaction rate expression
1	$C_3H_8 \rightarrow C_2H_4 + CH_4$	$R_1 = k_1 [C_3H_8]$
2	$C_3H_8 \rightleftharpoons C_3H_6 + H_2$	$R_2 = \overrightarrow{k_2} [C_3H_8] - \frac{\overrightarrow{k_2}}{K_{eq,2}} [C_3H_6] [H_2]$
3	$C_2H_4 + C_3H_8 \rightarrow C_3H_6 + C_2H_6$	$R_3 = k_3 [C_3H_8] [C_2H_4]$
4	$2C_3H_6 \rightarrow 3C_2H_4$	$R_4 = k_4 [C_3H_6]$
5	$2C_3H_6 \rightarrow 2CH_4 + \frac{2}{3}C_{5,plus}$	$R_5 = k_5 [C_3H_6]$
6	$C_3H_6 \rightleftharpoons C_2H_2 + CH_4$	$R_6 = \overrightarrow{k_6} [C_3H_6] - \frac{\overrightarrow{k_6}}{K_{eq,6}} [C_2H_2] [CH_4]$
7	$C_2H_6 + C_3H_6 \rightarrow CH_4 + C_4H_8$	$R_7 = k_7 [C_3H_6] [C_2H_6]$
8	$C_2H_6 \rightleftharpoons C_2H_4 + H_2$	$R_8 = \overrightarrow{k_8} [C_2H_6] - \frac{\overrightarrow{k_8}}{K_{eq,8}} [C_2H_4] [H_2]$
9	$C_2H_4 + C_2H_2 \rightleftharpoons C_4H_6$	$R_9 = \overrightarrow{k_9} [C_2H_4] [C_2H_2] - \frac{\overrightarrow{k_9}}{K_{eq,9}} [C_4H_6]$
10	$2C_2H_6 \rightarrow CH_4 + C_3H_8$	$R_{10} = k_{10} [C_2H_6]$
11	$C_2H_6 + C_2H_4 \rightarrow CH_4 + C_3H_6$	$R_{11} = k_{11} [C_2H_4] [C_2H_6]$
12	$3C_2H_4 \rightarrow C_{5,plus} + 3H_2$	$R_{12} = k_{12} [C_2H_4]$
13	$C_2H_4 + C_4H_6 \rightarrow C_{5,plus} + 2H_2$	$R_{13} = k_{13} [C_2H_4] [C_4H_6]$
14	$C_2H_4 \rightleftharpoons C_2H_2 + H_2$	$R_{14} = \overrightarrow{k_{14}} [C_2H_4] - \frac{\overrightarrow{k_{14}}}{K_{eq,14}} [C_2H_2] [H_2]$
15	$2C_2H_6 \rightarrow 2CH_4 + C_2H_4$	$R_{15} = k_{15} [C_2H_6]$

The mechanism is composed by 15 reactions, all of them being considered first order with respect to each of their reactants. The *forward* reaction rates are calculated according the Arrhenius law:

$$k_i(T) = k_{i,0} \exp\left(-\frac{E_a}{R}\left(\frac{1}{T} - \frac{1}{T_{ref}}\right)\right) \tag{2.19}$$

while the *backward* reaction rates are obtained assuming equilibrium reversible reactions:

$$\overleftarrow{k}_i = \frac{\overrightarrow{k}_i}{K_{eq,i}} \tag{2.20}$$

with

$$K_{eq,i} = K_{i,0} \exp\left(-\frac{\Delta G_i}{RT}\right). \tag{2.21}$$

The material balances for each of the components involved in this cracking mechanism are solved as discussed in §2.2.1.

2.2.6 Coking kinetics

As already discussed, the formation of coke during steam cracking is the subject of a lot of different studies (some examples are those of Albright (2002), Mahulkar *et al.*

(2014), Sundaram & Froment (1977), Wauters & Marin (2002)). However, it is hard to find a univocal consensus on the way coke is formed during the furnace operation. Different simplistic kinetic models are available in the open literature to describe the coke build-up, but they are feedstock-specific and work properly only on certain circumstances (Wauters & Marin, 2002).

The first thing that should be observed is that the output variable which is directly related to the coke formation is pressure drop: due to the deposition of coke at the inner wall surface, pressure drop along the coil increases. Plant data suggest that the rate of increase of pressure drop decreases consistently during the furnace operation; this decrease is related to a decreasing rate of coke formation. As suggested also in Albright (2002) and Wauters & Marin (2002), the reduced rate of pressure-drop build-up is related to a shifting mechanism on the coke formation. In this study, the kinetic model for coking assumes the existence of two different mechanisms taking place in parallel:

i) *catalytic coking*: this mechanism is particularly important during the start-up of the furnace and is responsible for the quick increase of pressure drop during the first hours of operation. This is due to the presence of catalytically active sites on the clean surface of the tubes that enhance the coke deposition. Different models have been proposed (Wauters & Marin, 2002) to take into account this coking mechanism. In the model considered, catalytic coking is assumed to be described by a first order reaction with respect to ethylene with a rate constant that decreases with decreasing concentration of catalytically active sites:

$$r_{catalytic}(z) = k_{cat}\left(\frac{c_{cat}(z)}{c_{cat}^{max}}\right) c_{C_2H_4}(z), \quad z \in [0, L]. \tag{2.22}$$

ii) *pyrolitic coking*: this mechanism is responsible for the coke formation during the entire run length of the furnace. The model used to describe this mechanism is the one proposed by Plehiers (Plehiers *et al.*, 1977):

$$r_{pyrolitic}(z) = k_{c1} c_{C_2H_4}(z) + k_{c2}\left(\frac{c_{C_2H_4}^2}{c_{C_3H_6}}\right). \tag{2.23}$$

The mass balance equation for coke is given by:

$$\frac{\mathrm{d}\rho^{coke\,layer}(z)}{\mathrm{d}t} = \frac{2R_I(z)}{(R_w)^2} r_{coking}(z) \tag{2.24}$$

where $\rho^{coke\,layer}(z)$ is the mass of coke deposited per unit volume of clean tube, and the coking rate is given by the combination of eq.(2.22) and eq.(2.23):

$$r_{coking}(z) = k_{cat}\left(\frac{c_{cat}(z)}{c_{cat}^{max}}\right) c_{C_2H_4}(z) + k_{c1} c_{C_2H_4}(z) + k_{c2}\left(\frac{c_{C_2H_4}^2}{c_{C_3H_6}}\right). \tag{2.25}$$

$\rho^{coke\,layer}(z)$ is related to the radius at the gas-coke interface according to the equation:

$$\rho^{coke\,layer}(z) = \left[1 - \left(\frac{R_I(z)}{R_w}\right)^2\right] \hat{\rho}_{coke} \tag{2.26}$$

In the expression of the coking rate (2.25) the surface concentration of the active sites is needed. The amount of active sites available for catalytic coking decreases with time as a result of their blockage by the pyrolytic coke being formed. Because of this, their concentration decreases proportionally to the concentration of the pyrolytic coking precursor (ethylene) according to the expression:

$$\frac{\mathrm{d}c_{cat}(z)}{\mathrm{d}t} = -k_{c1}c_{C_2H_4}(z)\left(\frac{c_{cat}(z)}{c_{cat}^{max}}\right) \quad z \in [0, L] \tag{2.27}$$

with the initial condition:

$$c_{cat}(z) = c_{cat}^{max} \tag{2.28}$$

where c_{cat}^{max} is the maximum surface concentration of active sites. The rate constants in eq. (2.25) and (2.27) change with temperature according to the Arrhenius law:

$$k_i = k_i^{ref} \exp\left(\frac{-E_i}{R}\left(\frac{1}{T} - \frac{1}{T_{ref}}\right)\right), \quad i = c_1, c_2, cat \tag{2.29}$$

with E_i = activation energy for reaction i.

The deposition of coke on the tubes can be monitored using two other variables, the thickness of coke deposited $\phi_{coke}(z)$ and the total mass of coke deposited w_{coke}, defined respectively by the two equations:

$$\phi_{coke}(z) = R_w - R_I(z) \tag{2.30}$$

$$\frac{\mathrm{d}w_{coke}}{\mathrm{d}t} = 2\pi L \int_0^1 r_{coking}(z^{'})R(z^{'}) \tag{2.31}$$

with the usual meaning for the symbols involved.

2.3 USX model

The transfer line heat exchanger used to quench the products of the cracking reactions is modelled using the Tube 1D model coupled with a cooling jacket model. It is important to notice that, since the kinetic parameters of cracking and coking reactions are not accurate at the temperature range of operation of the heat exchanger, these reactions are not considered in the USX. A fixed value is assigned to its length and internal diameter. The cooling jacket is modelled using a simple heat-transfer expression:

$$Q_{jacket}(z) = h_{jacket}\left(T_c - T_{wall}\right) \tag{2.32}$$

where $Q_{jacket}(z)$ is the heat flux from the tube wall to the jacket, h_{jacket} is the heat transfer coefficient for the jacket side, T_c is the cooling medium temperature and T_{wall} the temperature at the wall of the tubes. The heat transfer coeeficient h_{jacket} has been tuned to match the plant data of the USX outlet temperature.

2.4 Downstream piping model

According to eq.(2.7), given to the model the coil outlet pressure (COP) as an input, the pressure profile along the coil is calculated, the coil inlet pressure (CIP) is then computed and so the pressure drop.

The location of the COP measuring instrument in the plant is downstream of the USX heat exchanger. Thus, in order to use the COP data as a model input, the modelling of the piping network downstream of the USX till the location of the pressure measurement is required. However, a detailed model of the downstream piping does not contribute much to the validation of the furnace model, while increasing its size and complexity. Thus, a simplified approach has been used. The downstream piping is modelled using an instance of the Tube1D model (cracking and coking reactions turned off) with an effective length tuned in order to match the pressure drop along the downstream piping.

2.5 Weaknesses of the Furnace model

In the previous paragraphs the most important equations and assumptions of the Furnace model were presented.

When state estimation techniques need to be tested on a preexisting dynamic model, it is of fundamental importance to be able to understand what assumptions have been made in the model and how they can affect the model prediction. In fact, an *a priori* knowledge of the model weaknesses can deeply help the implementation of a state estimation algorithm and make it well-suited for the specific model considered.

From the previous discussions, it should be clear that the following weaknesses can be detected in the Furnace model:

1. the pyrolysis is described by a simplified molecular scheme instead of a more rigorous radical scheme. This has a deep effect on the prediction of the product composition;

2. the firebox model is oversimplified and cannot predict the heat flux profile on the external surface of the coil. Furthermore, the prediction of the fuel consumption is affected by the hypotheses of (i) total combustion of the fuel and (ii) no heat loss to the surroundings. This last assumption can deeply affect the prediction of fuel consumption;

3. the coking rate is described by a simplistic model and assumptions on both the coking mechanism and the rate expression have been made. It will be shown that the coking model represents the weakest part of the Furnace model. However, a correct prediction of the state of coking of the furnace is fundamental to carry out a global optimization strategy for the process and implement an effective control scheme for the furnace. The inability of the model to give an accurate description of the state of coking is the most important driver to apply state estimation techniques to improve the coking prediction;

4. there is a high uncertainty on the numerical values that should be given to some of the model parameters, like the roughness parameter α of eq.(2.9) or the coke thermal conductivity λ_c. It will be proved (§4.2.4) that these parameters have little influence on the model prediction, so no particular attention should be driven to this weakness.

All of these aspects have been carefully taken into account when applying state estimation to the system and their knowledge has proved to be fundamental for the success of the project.

Chapter 3

State estimation techniques

There is a wide variety of different applications, ranging from process control to on-line model adjustment, that require on-line estimates and predictions of an evolving set of variables, given uncertain data and dynamics (Ikonen, 2013).

In this chapter, the most promising state-estimation techniques and the algorithms for their implementation are presented.
First, the widely-used Extended Kalman Filter (EKF) is presented for discrete-time ordinary differential systems. Since most of the systems involved in chemical engineering-related applications are, from a mathematical perspective, nonlinear PDAE systems, the extension of the EKF to these systems is discussed.
The Unscented Kalman Filter (UKF), the Ensemble Kalman Filter (EnKF) and the Particle Filter (PF) are then discussed and their advantages and disadvantages with respect to EKF are emphasized.
Finally, the possibility to estimate model parameters using state estimation is discussed.

3.1 Problem description for pure ODE systems

The behavior of many physical systems can be described using a mathematical dynamic model (Simon, 2006). Given a set of present and future inputs, the model allows to deduce the present and future state of the system.

Every model is characterized by a certain degree of uncertainty: if few on-line data of some output variables are available, a proper state estimator can adjust the model prediction using these measurements to obtain a better estimate of the state. This is the most important application of on-line state estimation (Simon, 2006).

Let $\aleph$ be the physical system of interest, $\mathbf{x} \in \mathbb{R}^n$ the state vector, $\mathbf{u} \in \mathbb{R}^m$ the input vector and $\mathbf{y} \in \mathbb{R}^p$ the measurement vector. In order to use a state-space estimator on this system, the following requirements are needed:

- a *dynamic model* for the system: this can be a linear model or, most often, a nonlinear model;
- a *measurement model*: the model which describes the relation between the measurement output variables with the other state variables;
- a proper way to describe uncertainty on model prediction;
- a proper way to describe measurement noise.

In the previous considerations, no uncertainty is assumed for the control input vector.

The model uncertainty $\mathbf{w}$ and the measurement noise $\mathbf{v}$ are assumed as white, zero-mean, uncorrelated random variables with covariance $\mathbf{Q}$ and $\mathbf{R}$ respectively.

If the dynamic model for the system is a discrete-time linear dynamic model, then the following equations hold:

$$\mathbf{x}_k = \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} + \mathbf{w}_{k-1} \tag{3.1}$$

$$\mathbf{y}_k = \mathbf{C}_k\mathbf{x}_k + \mathbf{v}_k \tag{3.2}$$

where $\mathbf{A}$ is called the system matrix, $\mathbf{B}$ the input matrix and $\mathbf{C}$ the output matrix.

Almost all relevant physical systems are described by a *nonlinear* dynamic model instead of a linear one. In this case, equations (3.1) and (3.2) become:

$$\mathbf{x}_k = \mathbf{f}_{k-1}(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}, \mathbf{w}_{k-1}) \tag{3.3}$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{v}_k) \tag{3.4}$$

where $\mathbf{f}$ collects the system model equations and $\mathbf{h}$ the measurement equations. The notation used to describe the model and measurement noise as white zero-mean normal distributed noises is as follows:

$$\mathbf{w}_k \sim (0, \mathbf{Q}_k) \tag{3.5}$$

$$\mathbf{v}_k \sim (0, \mathbf{R}_k). \tag{3.6}$$

Given the model and measurement equations (3.1),(3.2) for the linear case and (3.3), (3.4) for the nonlinear case, and a set of on-line available measurements collected in $\mathbf{y}$, different algorithms can be adopted to obtain an optimal estimate of the state of the system considering both model prediction and available data. The most important ones are discussed in the next sections.

3.2 The Extended Kalman Filter

The extended Kalman filter (EKF) is one of the simplest and most important tools for state estimation purposes (Simon, 2006). This estimator updates the mean and covariance of the distribution of the state (assumed as normal distributed) according to a prediction-correction approach: in the prediction step, model equations are taken into account, while in the correction step available measurements are used to correct the predicted state estimate. The mean of the state *pdf*[1] is considered as the best state estimate and its covariance as a measure of the error (spread) around the mean. If the system is linear, the distribution of the state is completely characterized by its mean and covariance, hence the EKF is called an *optimal* state estimator. With nonlinear systems, the role of the system matrix is played by the Jacobian of the dynamics matrix: a *surrogate covariance*, which fails to account for the fully nonlinear dynamics of the system, is propagated in the same way as the error covariance in the linear case (Gillijns *et al.*, 2006).

At time step k, the *a priori* estimate of $\mathbf{x}_k$ is defined as the expected value of $\mathbf{x}_k$ conditioned on all the measurements before time k:

$$\hat{\mathbf{x}}_k^- = E\left[\mathbf{x}_k \middle| \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{k-1}\right] \tag{3.7}$$

and the *a posteriori* estimate as the the expected value of $\mathbf{x}_k$ conditioned on all the measurements up to time k:

$$\hat{\mathbf{x}}_k^+ = E\left[\mathbf{x}_k \middle| \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k\right]. \tag{3.8}$$

[1] Probability density function.

The covariance of the estimation error of $\hat{\mathbf{x}}_k^-$ is defined as:

$$\mathbf{P}_k^- = E\left[(\mathbf{x}_k - \hat{\mathbf{x}}_k^-)(\mathbf{x}_k - \hat{\mathbf{x}}_k^-)^T\right] \tag{3.9}$$

while for the *a posteriori* estimate $\hat{\mathbf{x}_k^+}$ as:

$$\mathbf{P}_k^+ = E\left[(\mathbf{x}_k - \hat{\mathbf{x}}_k^+)(\mathbf{x}_k - \hat{\mathbf{x}}_k^+)^T\right]. \tag{3.10}$$

The EKF works according to a prediction-correction approach: at time step k, *before* the measurement $\mathbf{y}_k$ is processed, an estimate of the state of the system ($\hat{\mathbf{x}}_k^-$) and of its error covariance ($\mathbf{P}_k^-$) is obtained. Then, the state estimate is refined after $\mathbf{y}_k$ is processed obtaining $\hat{\mathbf{x}}_k^+$, and so the error covariance. Fig. 3.1 helps understanding these relationships.

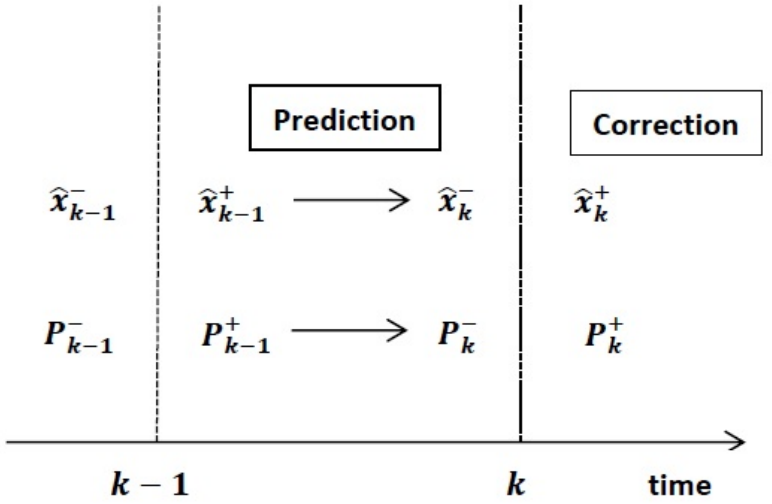


Figure 3.1: *A priori and a posteriori state estimates and estimation error covariances.*

If the system is linear (i.e it is described by eq, (3.1) and (3.2)) the algorithm for the EKF is as follows (Simon, 2006):

1. Filter initialization:

$$\hat{\mathbf{x}_0}^+ = E(\mathbf{x}_0) \tag{3.11}$$
$$\mathbf{P}_0^+ = E\left[(\mathbf{x}_0 - \hat{\mathbf{x}}_0^+)(\mathbf{x}_0 - \hat{\mathbf{x}}_0^+)^T\right] \tag{3.12}$$

2. (prediction step) Time update equations:

$$\hat{\mathbf{x}}_k^- = \mathbf{A}_{k-1}\hat{\mathbf{x}}_{k-1}^+ + \mathbf{B}_{k-1}\mathbf{u}_{k-1} \quad \textit{a priori} \text{ estimate} \tag{3.13}$$
$$\mathbf{P}_k^- = \mathbf{A}_{k-1}\mathbf{P}_{k-1}^+\mathbf{A}_{k-1}^T + \mathbf{Q}_{k-1} \tag{3.14}$$

3. Kalman gain computation:

$$\mathbf{K}_k = \mathbf{P}_k^-\mathbf{C}_k^T\left(\mathbf{C}_k\mathbf{P}_k^-\mathbf{C}_k^T + \mathbf{R}_k\right)^{-1} \tag{3.15}$$

4. (correction step) Measurement update equations:

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k\left(\mathbf{y}_k - \mathbf{C}_k\hat{\mathbf{x}}_k^-\right) \quad \textit{a posteriori} \text{ estimate} \tag{3.16}$$

$$\mathbf{P}_k^+ = \left(\mathbf{I} - \mathbf{K}_k\mathbf{C}_k\right)\mathbf{P}_k^-. \tag{3.17}$$

The previous equations allow to propagate the mean and the covariance of the state distribution if the system is a pure ODE linear system. Eq. (3.14) is known as the discrete-time Riccati equation.

If the system is nonlinear (i.e is described by eq.(3.3) and (3.4)) , eq.(3.14) cannot be used to propagate the error covariance (Gillijns *et al.*, 2006). However, if a linearization procedure is applied, the same equations of the linear Kalman filter can be used if the appropriate Jacobian matrices are computed. At time step $k-1$, the following Jacobians need to be computed:

$$\mathbf{F}_{k-1} = \left.\frac{\partial \mathbf{f}_{k-1}}{\partial \mathbf{x}}\right|_{\hat{\mathbf{x}}_{k-1}^+} \tag{3.18}$$

$$\mathbf{L}_{k-1} = \left.\frac{\partial \mathbf{f}_{k-1}}{\partial \mathbf{w}}\right|_{\hat{\mathbf{x}}_{k-1}^+} \tag{3.19}$$

and the time-update equations (prediction step) become:

$$\mathbf{P}_k^- = \mathbf{f}_{k-1}\mathbf{P}_{k-1}^+\mathbf{F}_{k-1}^T + \mathbf{L}_{k-1}\mathbf{Q}_{k-1}\mathbf{L}_{k-1}^T \tag{3.20}$$

$$\hat{\mathbf{x}}_k^- = \mathbf{f}_{k-1}(\hat{\mathbf{x}}_{k-1}^+, \mathbf{u}_{k-1}, \mathbf{0}). \tag{3.21}$$

In order to compute the correction step, two additional Jacobians need to be calculated:

$$\mathbf{H}_k = \left.\frac{\partial \mathbf{h}_k}{\partial \mathbf{x}}\right|_{\hat{\mathbf{x}}_k^-} \tag{3.22}$$

$$\mathbf{M}_k = \left.\frac{\partial \mathbf{h}_k}{\partial \mathbf{v}}\right|_{\hat{\mathbf{x}}_k^-}. \tag{3.23}$$

The measurement update equations become:

$$\mathbf{K}_k = \mathbf{P}_k^-\mathbf{H}_k^T\left(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^T + \mathbf{M}_k\mathbf{R}_k\mathbf{M}_k^T\right)^{-1} \tag{3.24}$$

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k\left[\mathbf{y}_k - \mathbf{h}_k(\hat{\mathbf{x}}_k^-, \mathbf{0})\right] \tag{3.25}$$

$$\mathbf{P}_k^+ = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\mathbf{P}_k^- \tag{3.26}$$

Eq. (3.21) represents the state estimate before the measurement refinement and eq. (3.25) the state estimate after the measurement refinement. In Fig 3.2 a graphical resume of the algorithm is presented.

3.2.1 Extension to PDAE systems

Many chemical processes are described by a set of nonlinear algebraic and (partial) differential equations. With these systems, in principle, only the differential state of the system could be propagated by the estimator and the algebraic state obtained from the

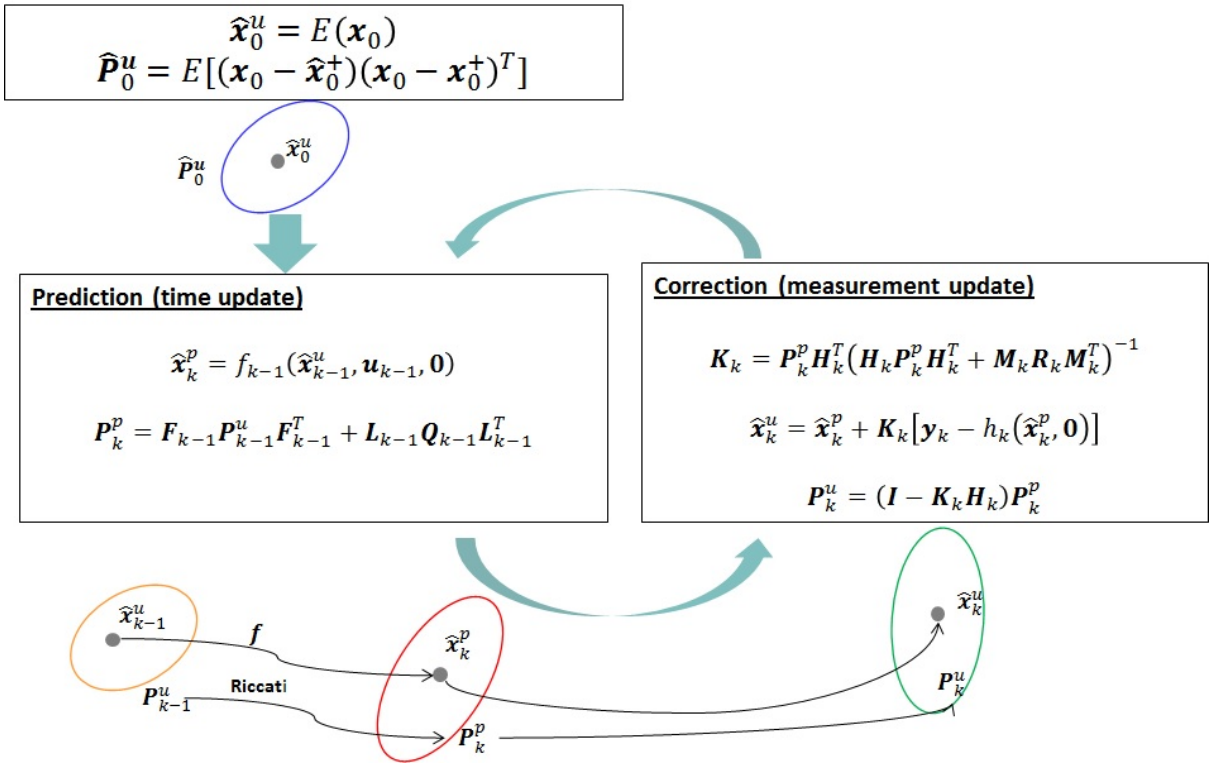


Figure 3.2: *Extended Kalman Filter: graphical overview.*

differential state using model equations. However, this is usually not convenient or not feasible (Cheng *et al.*, 1997).

Different alternative criteria can be used for state estimation purposes on nonlinear (P)DAE systems: minimal least squares, maximum likelihood and minimum maximum are the most common. The criterion that has been employed in this study and that is here briefly reviewed is the minimal least squares criterion.

With PDAE systems, the state vector at time k $\mathbf{x}^k \in \mathbb{R}^n$ is first decomposed into its differential part $\mathbf{x}_1^k \in \mathbb{R}^{n_1}$ and its algebraic part $\mathbf{x}_2^k \in \mathbb{R}^{n_2}$. According to this, the model uncertainty vector $\mathbf{w}^k$ is split into $\mathbf{w}_1^k$ and $\mathbf{w}_2^k$, and the error covariance according to:

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix}. \tag{3.27}$$

The model equations can then be written as:

$$\begin{bmatrix} \dot{\mathbf{x}}_1(t) \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_1(\mathbf{x}_1(t), \mathbf{x}_2(t), \mathbf{u}(t)) \\ \mathbf{f}_2(\mathbf{x}_1(t), \mathbf{x}_2(t), \mathbf{u}(t)) \end{bmatrix} + \begin{bmatrix} \mathbf{w}_1(t) \\ \mathbf{w}_2(t) \end{bmatrix} \tag{3.28}$$

in the continuous form or

$$\begin{bmatrix} \mathbf{x}_1^k \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_1(\mathbf{x}_1^{k-1}, \mathbf{x}_2^{k-1}, \mathbf{u}^{k-1}) \\ \mathbf{f}_2(\mathbf{x}_1^{k-1}, \mathbf{x}_2^{k-1}, \mathbf{u}^{k-1}) \end{bmatrix} + \begin{bmatrix} \mathbf{w}_1^{k-1} \\ \mathbf{w}_2^{k-1} \end{bmatrix} \tag{3.29}$$

in the discrete-time form. The measurement equations, with the assumption of additive noise, are given by:

$$\begin{aligned} \mathbf{y}(t) &= \mathbf{h}(\mathbf{x}_1(t), \mathbf{x}_2(t), t) + \mathbf{v}(t) \quad \text{continuous form} && (3.30) \\ \mathbf{y}^k &= \mathbf{h}(\mathbf{x}_1^k, \mathbf{x}_2^k) + \mathbf{v}^k \quad \text{discrete form.} && (3.31) \end{aligned}$$

In the interval $0 \leq t_k \leq T$, with $k = 1, \ldots, M$ the estimation problem is to determine the values of $\mathbf{x}_1^k$ and $\mathbf{x}_2^k$ that minimize the following objective function:

$$\begin{aligned} I = \frac{1}{2} \left[\mathbf{x}_1(0) - \mathbf{x}_{1,0}\right]^T \mathbf{P}_{11}^{-1}(0) \left[\mathbf{x}_1(0) - \mathbf{x}_{1,0}\right] \\ + \frac{1}{2} \int_0^T \left(\begin{bmatrix} \dot{\mathbf{x}}_1(t) \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{f}_1(\mathbf{x}_1(t), \mathbf{x}_2(t), t) \\ \mathbf{f}_2(\mathbf{x}_1(t), \mathbf{x}_2(t), t) \end{bmatrix} \right)^T \mathbf{R}^{-1}(t) \\ \left(\begin{bmatrix} \dot{\mathbf{x}}_1(t) \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{f}_1(\mathbf{x}_1(t), \mathbf{x}_2(t), t) \\ \mathbf{f}_2(\mathbf{x}_1(t), \mathbf{x}_2(t), t) \end{bmatrix} \right) \mathrm{d}t \\ + \frac{1}{2} \int_0^T \left[\mathbf{y}(t) - \mathbf{h}(\mathbf{x}_1(t), \mathbf{x}_2(t), t)\right]^T \mathbf{Q}^{-1}(t) \\ \left[\mathbf{y}(t) - \mathbf{h}(\mathbf{x}_1(t), \mathbf{x}_2(t), t)\right] \mathrm{d}t \end{aligned} \quad (3.32)$$

where the first term minimizes the square error of the initial estimate of $\mathbf{x}_1$, the second term minimizes the integral square system modelling error and the third term minimizes the integral square measurement error.

Using variational calculus, it is possible to prove (Cheng *et al.*, 1997) that the minimization of the objective function (3.32) for discrete-time nonlinear PDAE system yields to the following recursive procedure:

1. Filter initialization:

$$\begin{aligned} \hat{\mathbf{x}}_1^+(0) &= E(\mathbf{x}_1(0)) && (3.33) \\ \mathbf{P}_{11}(0)^+ &= E\left[(\mathbf{x}_1^+(0) - \mathbf{x}_1(0))(\mathbf{x}_1^+(0) - \mathbf{x}_1(0))^T\right] && (3.34) \end{aligned}$$

2. (prediction step) Time update for the state estimate:

$$\begin{bmatrix} \hat{\mathbf{x}}_1^-(t_k) \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{f}}_1(\hat{\mathbf{x}}_1^+(t_{k-1}), \hat{\mathbf{x}}_2^+(t_{k-1}), \mathbf{u}(t_{k-1})) \\ \hat{\mathbf{f}}_2(\hat{\mathbf{x}}_1^+(t_{k-1}), \hat{\mathbf{x}}_2^+(t_{k-1}), \mathbf{u}(t_{k-1})) \end{bmatrix} \quad (3.35)$$

while the error covariance is predicted by solving:

$$\begin{aligned} \begin{bmatrix} \dot{\mathbf{P}}_{11}(t) & \mathbf{P}_0 \\ \mathbf{0} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \frac{\partial \hat{\mathbf{f}}_1}{\partial \mathbf{x}_1} & \frac{\partial \hat{\mathbf{f}}_1}{\partial \mathbf{x}_2} \\ \frac{\partial \hat{\mathbf{f}}_2}{\partial \mathbf{x}_1} & \frac{\partial \hat{\mathbf{f}}_2}{\partial \mathbf{x}_2} \end{bmatrix} \begin{bmatrix} \mathbf{P}_{11}(t) & \mathbf{0} \\ \mathbf{P}_{21}(t) & \mathbf{P}_{22}(t) \end{bmatrix} \\ + \begin{bmatrix} \mathbf{P}_{11}(t) & \mathbf{P}_{12}(t) \\ \mathbf{0} & \mathbf{P}_{22}(t) \end{bmatrix} \begin{bmatrix} \frac{\partial \hat{\mathbf{f}}_1}{\partial \mathbf{x}_1} & \frac{\partial \hat{\mathbf{f}}_1}{\partial \mathbf{x}_2} \\ \frac{\partial \hat{\mathbf{f}}_2}{\partial \mathbf{x}_1} & \frac{\partial \hat{\mathbf{f}}_2}{\partial \mathbf{x}_2} \end{bmatrix}^T + \mathbf{Q}(t). \end{aligned} \quad (3.36)$$

3. (correction step) The state estimate is updated after $\mathbf{y}(t_k)$ is processed according

to the equation:

$$\begin{aligned}\begin{bmatrix} \hat{\mathbf{x}}_1^+(t_k) \\ \mathbf{0} \end{bmatrix} &= \begin{bmatrix} \hat{\mathbf{x}}_1^-(t_k) \\ \mathbf{0} \end{bmatrix} \\ &+ \begin{bmatrix} \mathbf{P}_{11}^+(t_k) & \mathbf{P}_{12}^+(t_k) \\ \mathbf{0} & \mathbf{P}_{22}^+(t_k) \end{bmatrix} \begin{bmatrix} \frac{\partial \hat{\mathbf{h}}}{\partial \hat{\mathbf{x}}_1} & \frac{\partial \hat{\mathbf{h}}}{\partial \hat{\mathbf{x}}_2} \end{bmatrix}^T \mathbf{Q}_k^{-1} \left[\mathbf{y}(t_k) - \hat{\mathbf{h}}(\hat{\mathbf{x}}_1^-(t_k), \hat{\mathbf{x}}_2^-(t_k), t_k) \right] \\ &- \begin{bmatrix} \mathbf{P}_{11}^-(t_k) & \mathbf{P}_{12}^-(t_k) \\ \mathbf{0} & \mathbf{P}_{22}^-(t_k) \end{bmatrix} \begin{bmatrix} \frac{\partial \hat{\mathbf{h}}}{\partial \hat{\mathbf{x}}_1} & \frac{\partial \hat{\mathbf{h}}}{\partial \hat{\mathbf{x}}_2} \end{bmatrix}^T \mathbf{Q}_k^{-1} \begin{bmatrix} \frac{\partial \hat{\mathbf{h}}}{\partial \hat{\mathbf{x}}_1} & \frac{\partial \hat{\mathbf{h}}}{\partial \hat{\mathbf{x}}_2} \end{bmatrix} \begin{bmatrix} \delta \hat{\mathbf{x}}_1(t_k) \\ \delta \hat{\mathbf{x}}_2(t_k) \end{bmatrix} \end{aligned} \tag{3.37}$$

while the error covariance is updated according to:

$$\begin{aligned}\begin{bmatrix} \mathbf{P}_{11}^+(t_k) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} &= \begin{bmatrix} \mathbf{P}_{11}^-(t_k) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{P}_{11}^+(t_k) & \mathbf{P}_{12}^+(t_k) \\ \mathbf{0} & \mathbf{P}_{22}^+(t_k) \end{bmatrix} \begin{bmatrix} \frac{\partial \hat{\mathbf{h}}}{\partial \hat{\mathbf{x}}_1} & \frac{\partial \hat{\mathbf{h}}}{\partial \hat{\mathbf{x}}_2} \end{bmatrix}^T \\ &\mathbf{Q}_k^{-1} \begin{bmatrix} \frac{\partial \hat{\mathbf{h}}}{\partial \hat{\mathbf{x}}_1} & \frac{\partial \hat{\mathbf{h}}}{\partial \hat{\mathbf{x}}_2} \end{bmatrix} \begin{bmatrix} \mathbf{P}_{11}^-(t_k) & \mathbf{0} \\ \mathbf{P}_{21}^-(t_k) & \mathbf{P}_{22}^-(t_k) \end{bmatrix} - \begin{bmatrix} \mathbf{P}_{11}^-(t_k) & \mathbf{P}_{12}^-(t_k) \\ \mathbf{0} & \mathbf{P}_{22}^-(t_k) \end{bmatrix} \\ &\begin{bmatrix} \frac{\partial \hat{\mathbf{h}}}{\partial \hat{\mathbf{x}}_1} & \frac{\partial \hat{\mathbf{h}}}{\partial \hat{\mathbf{x}}_2} \end{bmatrix}^T \mathbf{Q}_k^{-1} \begin{bmatrix} \frac{\partial \hat{\mathbf{h}}}{\partial \hat{\mathbf{x}}_1} & \frac{\partial \hat{\mathbf{h}}}{\partial \hat{\mathbf{x}}_2} \end{bmatrix} \begin{bmatrix} \delta \mathbf{P}_{11}(t_k) & \mathbf{0} \\ \delta \mathbf{P}_{12}(t_k) & \delta \mathbf{P}_{22}(t_k) \end{bmatrix} \end{aligned} \tag{3.38}$$

with

$$\begin{bmatrix} \delta \hat{\mathbf{x}}_1(t_k) \\ \delta \hat{\mathbf{x}}_2(t_k) \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{x}}_1^+(t_k) - \hat{\mathbf{x}}_1^-(t_k) \\ \hat{\mathbf{x}}_2^+(t_k) - \hat{\mathbf{x}}_2^-(t_k) \end{bmatrix} \tag{3.39}$$

$$\begin{bmatrix} \delta \mathbf{P}_{11}(t_k) & \delta \mathbf{P}_{12}(t_k) \\ \delta \mathbf{P}_{21}(t_k) & \delta \mathbf{P}_{22}(t_k) \end{bmatrix} = \begin{bmatrix} \mathbf{P}_{11}^+(t_k) - \mathbf{P}_{11}^-(t_k) & \mathbf{P}_{12}^+(t_k) - \mathbf{P}_{12}^-(t_k) \\ \mathbf{P}_{21}^+(t_k) - \mathbf{P}_{21}^-(t_k) & \mathbf{P}_{22}^+(t_k) - \mathbf{P}_{22}^-(t_k) \end{bmatrix}. \tag{3.40}$$

Equations (3.35) and (3.36) represent the prediction estimate of the state and of the error covariance respectively, while equation (3.37) and (3.38) the state estimate and the error covariance estimate after the measurement update. It can be noticed that these equations reduce to the standard EKF prediction-correction equations (3.21), (3.20), (3.25) and (3.26) when applied to ordinary differential systems. It should also be emphasized that the covariance matrix $\mathbf{P}(t)$ corresponds to the covariance of the estimated error only for linear PDAE systems; for nonlinear systems, although $\mathbf{P}(t)$ can be associated with the uncertainty of the estimated state, it has no clear statistical meaning. For sake of simplicity, in the previous discussion $\mathbf{P}(t)$ has been defined as the estimated error covariance, independently of the system considered (linear/nonlinear).

3.3 The Unscented Kalman Filter

As discussed so far, the Extended Kalman Filter attempts to propagate the mean and covariance of the distribution of the state (assumed as normal distributed) according to a predictor-corrector procedure. If the system is linear, the distribution is univocally characterized by its mean and covariance, otherwise a linearization of the model equations is required to propagate the covariance matrix. As a matter of facts, for nonlinear systems, the following observations can be made if the EKF approach is considered:

- a first order linearization procedure is adopted: the estimator cannot account for the full nonlinear nature of the system. This often leads to an incorrect representation of the error probability density;

- the Jacobians (3.18), (3.19), (3.22), (3.23) need to be computed. This is not an easy task, especially for large systems (Curn, 2014). If the model and measurement noises are considered as linear additive noises with respect to the state and the measurement vectors respectively, the computation of (3.19) and (3.23) can be avoided.

A different approach which allows to overcome these limitations is the Unscented Kalman Filter (UKF). This filter is based on two fundamental principles (Simon, 2006):

1. it is easier to perform a nonlinear transformation on a single point in the state space rather than on an entire *pdf*;

2. it is not difficult to find a set of deterministic individual points (called *sigma points*) whose sample *pdf* approximates the true *pdf* of a state vector.

Based on these two principles, the algorithm for the UKF is as follows:

1. First, the initial state and the initial error covariance are initialized:

$$\hat{\mathbf{x}}_0^+ = E(\mathbf{x}_0) \tag{3.41}$$

$$\mathbf{P}_0^+ = E\left[\left(\mathbf{x}_0 - \hat{\mathbf{x}}_0^+\right)\left(\mathbf{x}_0 - \hat{\mathbf{x}}_0^+\right)^T\right]; \tag{3.42}$$

2. at time step $k-1$ $2n$ sigma points are chosen to approximate the state distribution:

$$\hat{\mathbf{x}}_{k-1}^i = \hat{\mathbf{x}}_{k-1}^+ + \tilde{\mathbf{x}}^i \qquad i = 1, \ldots, 2n \tag{3.43}$$

$$\tilde{\mathbf{x}}^i = \left(\sqrt{n\mathbf{P}_{k-1}^+}\right)_i^T \qquad i = 1, \ldots, n \tag{3.44}$$

$$\tilde{\mathbf{x}}^{n+i} = -\left(\sqrt{n\mathbf{P}_{k-1}^+}\right)_i^T \qquad i = 1, \ldots, n; \tag{3.45}$$

3. the *sigma points* are then propagated to time step k using the model nonlinear equations:

$$\hat{\mathbf{x}}_k^i = f(\hat{\mathbf{x}}_{k-1}^i, \mathbf{u}_k); \tag{3.46}$$

4. the predicted estimate of the state vector is then computed as the mean of the forecasted *sigma points*:

$$\bar{\hat{\mathbf{x}}}_k^- = \frac{1}{2n}\sum_{i=1}^{2n}\hat{\mathbf{x}}_k^i; \tag{3.47}$$

5. the predicted error covariance, with the assumption of additive noise, is calculated according to the expression:

$$\mathbf{P}_k^- = \frac{1}{2n}\sum_{i=1}^{2n}\left(\hat{\mathbf{x}}_k^i - \bar{\hat{\mathbf{x}}}_k^-\right)\left(\hat{\mathbf{x}}_k^i - \bar{\hat{\mathbf{x}}}_k^-\right)^T + \mathbf{Q}_{k-1}. \tag{3.48}$$

6. The forecast state estimate (3.47) and the forecast error covariance (3.48) need to be updated according to the measurement vector supplied to the estimator. The first step is to choose other $2n$ *sigma points* as follows:

$$\hat{\mathbf{x}}_{k-1}^{i} = \hat{\mathbf{x}}_{k}^{-} + \tilde{\mathbf{x}}^{i} \qquad i = 1, \ldots, 2n \tag{3.49}$$

$$\tilde{\mathbf{x}}^{i} = \left(\sqrt{n\mathbf{P}_{k}^{-}}\right)_{i}^{T} \qquad i = 1, \ldots, n \tag{3.50}$$

$$\tilde{\mathbf{x}}^{n+i} = -\left(\sqrt{n\mathbf{P}_{k}^{-}}\right)_{i}^{T} \qquad i = 1, \ldots, n. \tag{3.51}$$

   This step can be avoided and the same *sigma points* of the forecast step can be used.

7. The state *sigma points* are then transformed into predicted measurements using the measurement equations:

$$\hat{\mathbf{y}}_{k}^{i} = h(\hat{\mathbf{x}}_{k}^{i}); \tag{3.52}$$

8. the forecast measurement vector is then computed as the mean of the predicted measurements:

$$\bar{\hat{\mathbf{y}}}_{k}^{-} = \frac{1}{2n}\sum_{i=1}^{2n}\hat{\mathbf{y}}_{k}^{i}. \tag{3.53}$$

9. The covariance of the predicted measurements is then computed taking into account the measurement noise uncertainty (additive noise is considered):

$$\mathbf{P}_{yy} = \frac{1}{2n}\sum_{i=1}^{2n}\left(\hat{\mathbf{y}}_{k}^{i} - \bar{\hat{\mathbf{y}}}_{k}^{-}\right)\left(\hat{\mathbf{y}}_{k}^{i} - \bar{\hat{\mathbf{y}}}_{k}^{-}\right)^{T}; \tag{3.54}$$

10. the cross covariance $\mathbf{P}_{xy}$ is computed:

$$\mathbf{P}_{xy} = \frac{1}{2n}\sum_{i=1}^{2n}\left(\hat{\mathbf{x}}_{k}^{i} - \bar{\hat{\mathbf{x}}}_{k}^{-}\right)\left(\hat{\mathbf{y}}_{k}^{i} - \bar{\hat{\mathbf{y}}}_{k}^{-}\right)^{T}. \tag{3.55}$$

11. The measurement updates of the state estimate and the error covariance are calculated with the usual Kalman-gain expression:

$$\mathbf{K}_{k} = \mathbf{P}_{xy}\mathbf{P}_{yy}^{-1} \tag{3.56}$$

$$\hat{\mathbf{x}}_{k}^{+} = \hat{\mathbf{x}}_{k}^{-} + \mathbf{K}_{k}\left(\mathbf{y}_{k} - \bar{\hat{\mathbf{y}}}_{k}^{-}\right)^{T} \tag{3.57}$$

$$\mathbf{P}_{k}^{+} = \mathbf{P}_{k}^{-} - \mathbf{K}_{k}\mathbf{P}_{yy}\mathbf{K}_{k}^{T}. \tag{3.58}$$

In Fig. 3.3 a graphical resume of the algorithm is presented.

The main differences between the UKF and the EKF for nonlinear systems can be summarized as follows:

i) in the UKF, since every single sigma point is propagated using the original nonlinear transformation, no linearization errors are introduced. In the EKF, instead, a linearization procedure is adopted in order to propagate the error covariance using the Riccati equation.

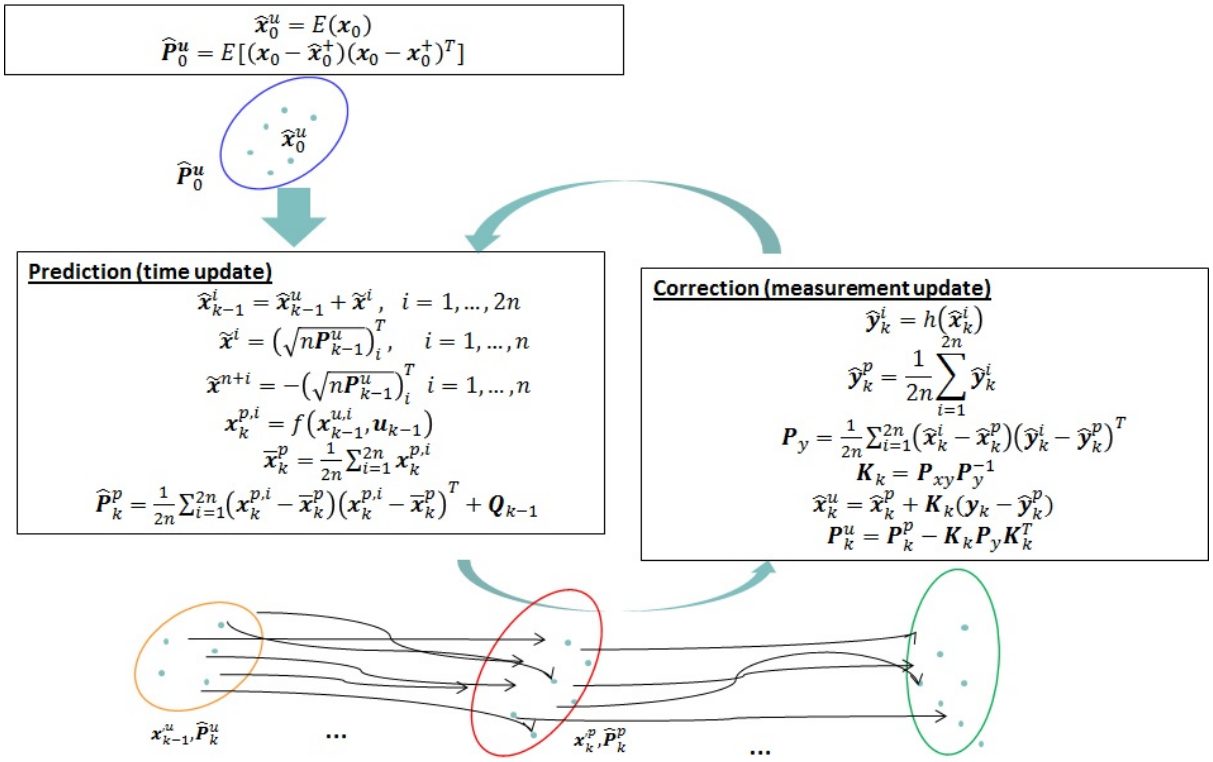


Figure 3.3: *Unscented Kalman Filter: graphical overview.*

ii) While in the EKF a single integration is required during the prediction step, in the UKF $2n$ integrations are required in order to propagate every single sigma point. This results in a higher computational effort.

iii) However, while in the EKF four Jacobians need to be computed, the UKF is a *derivative-free* algorithm.

iv) It can be proved (Simon, 2006) that the mean and covariance approximation using the EKF has a first order accuracy, while with the UKF a higher order accuracy (up to 3) can be obtained.

The Unscented Kalman filter has been tested (Romanenko & Castro, 1980) on a number of different case studies; it seems to perform better than the EKF with systems that present nonlinearities, but in some situations the *pdf* of the state cannot be properly represented by a discrete number of sigma points and the EKF can be a better choice (Laviola, 2003).

3.4 The Ensemble Kalman Filter

The Ensemble Kalman filter (EnKF) is a state estimation algorithm that can be considered a Monte-Carlo implementation of the Extended Kalman Filter (Evensen, 2003).

This estimator is widely used for systems with a huge number of state variables ($\sim$ millions) like in weather forecasting or ocean dynamics (Gillijns *et al.*, 2006). The basic idea - it is easier to perform a nonlinear transformation on a sample of points instead of on an entire *pdf* - is the same as in the UKF. However, while in the UKF the choice of the sigma points used to represent the distribution of the state is deterministic, in the EnKF it is completely stochastic. The starting point is choosing a set of heuristic sampling points, i.e. an ensemble of state estimates that captures the initial probability distribution of the state. Then, these sample points are propagated using the nonlinear model equations and the *pdf* of the actual state is approximated by the ensemble of the estimates. Finally, the measurement update step is applied with the same procedure as in the other estimation algorithms. This estimator can give consistent improvements with respect to the other estimators especially with large systems, since the number of sampling points needed to obtain a good estimate (q) is usually much smaller than the state vector dimension (n).

The EnKF algorithm can be summarized as follows:

1. at time $k = 0$, an ensemble of q state estimates is chosen in a stochastic way to represent the initial distribution of the state.

2. At time step (k) every member of the ensemble of the updated state estimates from the previous time-step:

$$X_{k-1}^{+} = \left(\mathbf{x}_{k-1}^{+,1}, \mathbf{x}_{k-1}^{+,2}, \ldots, \mathbf{x}_{k-1}^{+,q}\right) \tag{3.59}$$

   is updated using the nonlinear model equations:

$$\mathbf{x}_k^{-,i} = \mathbf{f}(\mathbf{x}_{k-1}^{+,i}, \mathbf{u}_k) + \mathbf{w}_k^i, \qquad i = 1, \ldots, q. \tag{3.60}$$

3. The predicted estimate of the actual state is then calculated as the ensemble mean:

$$\bar{\mathbf{x}}_k^{-} = \frac{1}{q}\sum_{i=1}^{q} \mathbf{x}_k^{-,i}. \tag{3.61}$$

4. The ensemble error matrix is computed according to:

$$\mathbf{E}_k^{-} = \left[x_k^{-,1} - \bar{\mathbf{x}}_k^{-}, \ldots, x_k^{-,p} - \bar{\mathbf{x}}_k^{-}\right] \tag{3.62}$$

   and the true state error covariance is approximated using the ensemble covariance matrix:

$$\mathbf{P}_k^{-} \sim \hat{\mathbf{P}}_k^{-} = \frac{1}{q-1}\mathbf{E}_k^{-}\left(\mathbf{E}_k^{-}\right)^T. \tag{3.63}$$

5. In the measurement update step, each member of the ensemble is updated using the usual Kalman gain expression:

$$\mathbf{x}_k^{+,i} = \mathbf{x}_k^{-,i} + \mathbf{K}_k\left(\mathbf{y}_k^i - h(\mathbf{x}_k^{-,i})\right) \tag{3.64}$$

   where the perturbed observations $\mathbf{y}_k^i$ are given by:

$$\mathbf{y}_k^i = \mathbf{y}_k + \mathbf{v}_k^i \tag{3.65}$$

with $\mathbf{v}_k^i$ = random noise with zero mean and covariance $\mathbf{R}_k$. The Kalman gain is computed according to the expression:

$$\mathbf{K}_k = \hat{\mathbf{P}}^-_{xy_k} \left(\hat{\mathbf{P}}^-_{yy_k} \right)^{-1} \tag{3.66}$$

where:

$$\hat{\mathbf{P}}^-_{xy_k} = \frac{1}{q-1} \mathbf{E}^-_k \left(\mathbf{E}^-_{\mathbf{y}_k} \right)^T \tag{3.67}$$

$$\hat{\mathbf{P}}^-_{yy_k} = \frac{1}{q-1} \mathbf{E}^-_{\mathbf{y}_k} \left(\mathbf{E}^-_{\mathbf{y}_k} \right)^T \tag{3.68}$$

$$\mathbf{E}^-_{\mathbf{y}_k} = \left[\mathbf{y}_k^{-,1} - \bar{\mathbf{y}}^-_k, \ldots, \mathbf{y}_k^{-,q} - \bar{\mathbf{y}}^-_k \right]. \tag{3.69}$$

According to the previous equations, it can be noticed that in the EnKF the evaluation of the filter gain does not involve a linearization of the model and measurement equations like in the EKF. The computation of Jacobians is avoided reducing the computational burden. To obtain accurate results, usually an ensemble size much smaller than the state dimension is required ($q << n$) (Gillijns *et al.*, 2006), thus reducing the number of integrations required at each time step in order to propagate the ensemble member with respect to the UKF. It should also be emphasized that the computation of the Kalman gain in the EnKF (eq.(3.66)) requires the evaluation of $\hat{\mathbf{P}}^-_{xy_k} \in \Re^{n \times p}$ and $\hat{\mathbf{P}}^-_{yy_k} \in \Re^{p \times p}$, which is an $O(pqn)$ operation, while in the EKF it requires the computation of $\hat{\mathbf{P}}^-_k \in \Re^{n \times n}$, which is a $O(n^3)$ computation. Thus, if $q << n$, the computational burden is consistently reduced. In Fig.3.4 a graphical overview of the EnKF is reported.

3.5 The Particle filter

The particle filter (PF) is a numerical implementation of the Bayesian estimator (Arulampalam *et al.*, 2002). The Ensemble Kalman Filter, discussed in the previous section, can be considered a particular example of a particle filter. The idea is always the same: a set of state vectors (called *particles* in this case) is randomly generated, transformed via known nonlinear equations and then combined to obtain the estimate of the state and the error covariance. The main feature of this filter is that the relative likelihoods of the transformed state vectors are computed, so that only the state vectors which are close to the real state of the system are retained, while the others are discarded.

The general algorithm for a particle filter is as follows:

1. N initial particles are randomly generated given the initial distribution of the state (usually a Gaussian distribution with mean $\mathbf{x}_0$ and covariance $\mathbf{P}_0$ is assumed).

2. For $k = 1, 2, \ldots$ each particle is propagated using the model equations:

$$\mathbf{x}^-_{k,i} = \mathbf{f}(\mathbf{x}^,_{k-1} \mathbf{w}^i_k) \quad i = 1, \ldots, N \tag{3.70}$$

where $\mathbf{w}^i_k$ is randomly generated once the model uncertainty covariance $\mathbf{Q}$ is known.

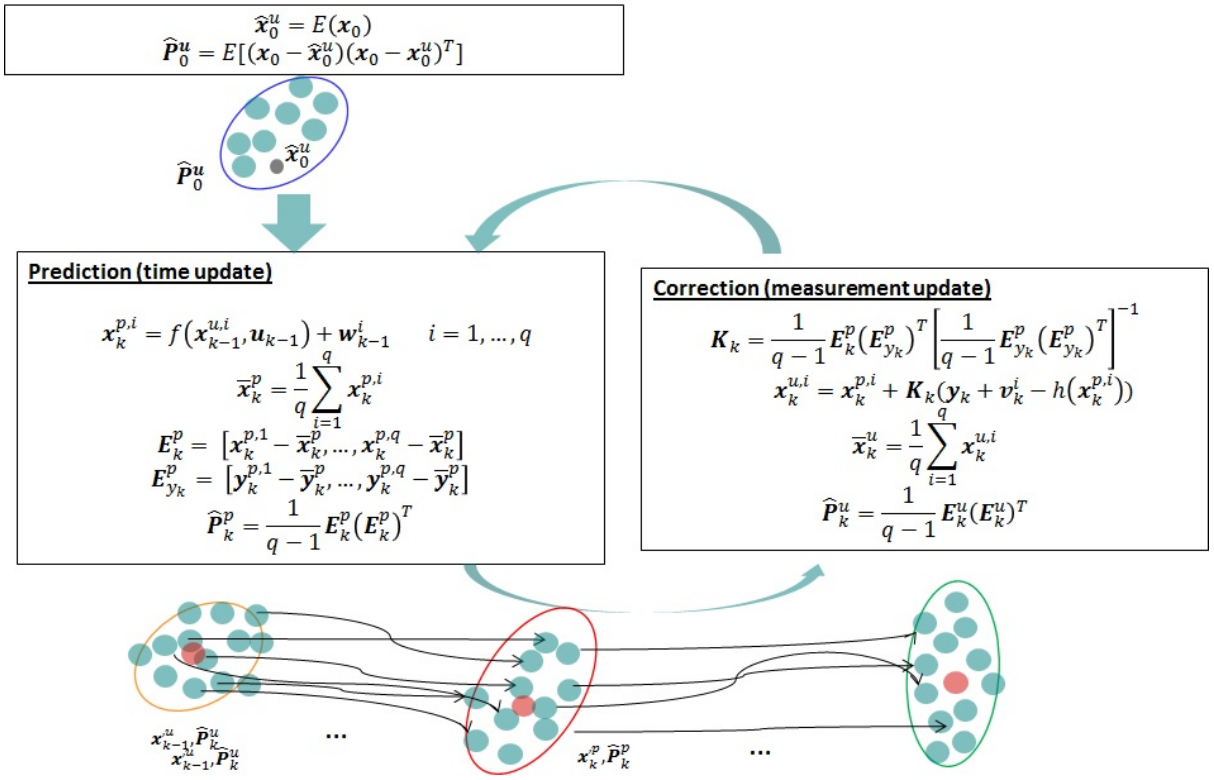


Figure 3.4: *Ensemble Kalman Filter: graphical overview.*

3. The relative likelihood q_i of each particle $\mathbf{x}_{k,i}^-$ conditioned on the measurement $\mathbf{y}_k$ is computed. In other terms, the pdf $p(\mathbf{y}_k|\,\mathbf{x}_{k,i}^-)$ need to be computed. In the case of a p-dimensional measurement vector $\mathbf{y}_k$ and a normal distributed measurement noise $\mathbf{v}_k \sim (\mathbf{0}, \mathbf{R})$ it can be proven(Simon, 2006) that the relative likelihood of each particle is proportional to the right side of the following expression:

$$q_i \sim \frac{1}{(2\pi)^{p/2}\det(\mathbf{R})^{\frac{1}{2}}} \exp\left(-\frac{\left[\mathbf{y_k} - \mathbf{h}(\mathbf{x}_{k,i}^-)\right]^T \mathbf{R}^{-1} \left[\mathbf{y_k} - \mathbf{h}(\mathbf{x}_{k,i}^-)\right]}{2}\right). \tag{3.71}$$

4. The relative likelihoods computed according to eq.(3.71) are normalized so that their sum is equal to one:

$$q_i = \frac{q_i}{\sum_{j=1}^{N} q_j}. \tag{3.72}$$

5. The particles $\mathbf{x}_{k,i}^-$ are then re-sampled according to the relative likelihoods q_i. This step is crucial and can be done in different ways. One straightforward approach is the one proposed by Ristic *et al.* (2004):

   i) a random number r uniformly distributed on [0,1] is generated;

ii) For $i = 1, \ldots, N$, the relative likelihoods q_i are accumulated until their sum is greater than r. That is, the value of j which satisfies the condition $\sum_{m=1}^{j-1} q_m < r$ but $\sum_{m=1}^{j} q_m \geq r$ is determined.

iii) The updated particle $\mathbf{x}_{k,i}^{+}$ is then set equal to the old particle $\mathbf{x}_{k,j}^{-}$.

Another resampling procedure, more sophisticated than the one discussed, is the following:

i) the mean and the covariance of the a priori set of particles are computed as follows:

$$\mu = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_{k,i}^{-} \tag{3.73}$$

$$S = \frac{1}{N-1} \sum_{i=1}^{N} \left(\mathbf{x}_{k,i}^{-} - \mu\right) \left(\mathbf{x}_{k,i}^{-} - \mu\right)^{T}; \tag{3.74}$$

ii) the matrix A such that $AA^T = S$ is computed (Cholesky factorization of S);

iii) The volume of the n-dimensional sphere $v_n = 2\pi v_{n-2}/2$ is computed starting from $v_1 = 2$, $v_2 = \pi$ and $v_3 = 4\pi/3$;

iv) The optimal kernel bandwidth h is computed according to:

$$h = \frac{1}{2} \left[\frac{8}{v_n}(n+4)(2\sqrt{\pi})^n\right]^{\frac{1}{(n+4)}} \frac{1}{N^{\frac{1}{n+4}}}; \tag{3.75}$$

v) An approximation of the pdf $p(\mathbf{x}_k|\,\mathbf{y}_k)$ is obtained according to:

$$\hat{p}(\mathbf{x}_k|\,\mathbf{y}_k) = \sum_{i=1}^{N} q_i K_h\left(\mathbf{x}_k - \mathbf{x}_{k,i}\right) \tag{3.76}$$

with

$$K_h(\mathbf{x}) = \frac{1}{\det(\mathbf{A})\, h^n} K\left(\mathbf{A}^{-1}\mathbf{x}/h\right) \tag{3.77}$$

and

$$K(x) = \begin{cases} \frac{n+2}{2v_n}\left(1 - ||\mathbf{x}||_2^2\right) & \text{if } ||\mathbf{x}||_2 < 1 \\ 0 & \text{otherwise} \end{cases} \tag{3.78}$$

6. Once the new set of particles $\mathbf{x}_{k,i}^{+}$ has been determined, the *a posteriori* state estimate $\mathbf{x}_k^{+}$ and the *a posteriori* error covariance P_k^{+} can be calculated as the mean and the covariance of the ensemble of particles:

$$\mathbf{x}_k^{+} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_{k,i}^{+} \tag{3.79}$$

$$P_k^{+} = \frac{1}{N-1} \sum_{i=1}^{N} \left(\mathbf{x}_{k,i}^{+} - \mathbf{x}_k^{+}\right) \left(\mathbf{x}_{k,i}^{+} - \mathbf{x}_k^{+}\right)^{T}. \tag{3.80}$$

The number of particles that must be chosen in order to obtain a good estimate of the state is usually larger than the dimension of the state vector ($N > n$). This means that the computational effort in the particle filter is a point of concern. However, it can be proved (Simon, 2006) that in the particle filter the estimation error converges to zero as the number of particles approach infinity, while in the UKF the estimation error does not converge to zero in any sense.

3.6 Comparison and choice of the estimator

All the algorithms presented in this chapter (EKF/UKF/EnKF/PF) have been implemented in MATLAB ® and tested using three simple text examples (pure ODE systems with max. 3 state variables) to understand their performance. A qualitative comparison of the pro's and con's of each algorithm is reported in Tab. 3.1.

After a careful analysis, it has been decided to test in on the gPROMS® Furnace model of chapter §2 the PDAE-extended version of the EKF presented in §3.2.1. The main reason is that its implementation is the easiest between all the estimation algorithms: if bad results wil lbe obtained using this estimation technique, the implementation of one of the other state estimators (e.g. the UKF) will be taken into account.

3.7 Parameter Estimation using State Estimation

State estimation can be used not only to estimate the state of a dynamic system, but also to estimate some (or all) of the unknown parameters of a dynamic model. The two things are not exclusive: state estimation can also be used to estimate *both* the state of the system and its unknown parameters.

Let $\mathbf{p}$ be a set of z model parameters that need to be estimated. The model and measurement equations can be expressed as:

$$\mathbf{x}_k = \mathbf{f}_{k-1}(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}, \mathbf{p}_{k-1}, \mathbf{w}_{k-1}) \tag{3.81}$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{p}_k, \mathbf{v}_k). \tag{3.82}$$

In order to estimate the vector of parameters $\mathbf{p}$, an augmented state vector $\mathbf{x}^{'}$ is defined:

$$\mathbf{x}_k^{'} = \begin{bmatrix} \mathbf{x}_k \\ \mathbf{p}_k \end{bmatrix} \tag{3.83}$$

Since the members of $\mathbf{p}$ are constant, the following z equations are added to the model equations:

$$\mathbf{p}_{k+1} = \mathbf{p}_k + \mathbf{w}_{pk} \tag{3.84}$$

where $\mathbf{w}_{pk}$ is a small artifical noise that allows the estimator to change its estimate of $\mathbf{p}_k$. The augmented model can then be rewritten as:

$$\mathbf{x}_k^{'} = \mathbf{f}_{k-1}(\mathbf{x}_{k-1}^{'}, \mathbf{u}_{k-1}, \mathbf{w}_{k-1}, \mathbf{w}_{p,k-1}) \tag{3.85}$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k^{'}, \mathbf{v}_k). \tag{3.86}$$

and any nonlinear filter can be used to estimate the state and the unknown parameters of the system.

Table 3.1: *Comparison between the different estimators considered.*

	EKF	**UKF**	**EnKF**	**PF**
What is propagated	First 2 moments of the state pdf	First 2 moments of the state pdf	First 2 moments of the state pdf	All particles
Procedure	Linearization	Approximation of the *pdf* with $2n$ sigma points	Approximation of the *pdf* with a random ensemble	Statistics derived from the ensemble of particles
Computation	• Single integration at each time step • Computation of Jacobians required	• $2n$ integrations at each time-step are required • Derivative free	• q integrations required at each time-step • Derivative free	• N integrations required at each time step • Derivative free
Performance and applications	• Extensively tested • Problems may occur with systems with high nonlinearities	• Tested especially for tracking problems • Usually performs better than EKF	• Used for weather forecasting or for studies on ocean dynamics • Gives great advantages with large systems	• Used especially for tracking problems and pattern recognition • Computational effort can be an issue

It should be noticed that, for a continuous-time system, the previous procedure just translates into adding the following set of differential equations to the system:

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = 0, \quad i = 1, \ldots, z \tag{3.87}$$

with the dimension of the model error covariance changed accordingly.

Chapter 4

Implementation of the state estimator

This chapter deals with all the issues that had to be faced during the implementation of the state estimator to the Furnace model of chapter 2.
First, the software and the estimation algorithm used in this study are presented. Then, a brief overview of all the aspects that need to be considered when using state estimation are discussed. The reasons why the usage of a state estimator with the Furnace model could be extremely beneficial are analyzed, and the hierarchical procedure adopted to test the estimator is given. Finally, the possible outcomes from the estimator are considered.

4.1 The gPROMS® software

The sofware used to test the state estimator on the model of the steam cracker presented in chapter 2 is gPROMS® Model Builder v.4.2.0. The estimator that has been used in all the simulations is the EKF, implemented using its generalized formulation for PDAE systems as discussed in §3.2.1.

gPROMS® Model Builder is a platform for high-fidelity predictive modelling developed and commercialized by Process Systems Enterprise Ltd. (London,UK). It can be used to build steady-state and dynamic process models of any complexity. Its structure is equation-oriented: all the model equations are solved simultaneously, thus increasing the speed and robustness of the simulation if compared with the traditional sequential-modular process simulators.

gPROMS® Model Builder allows the usage of external components, named Foreign Objects (FOs), which provide certain computational services to the original gPROMS® model. This FOs can be physical properties packages, external unit operation modules, complete computational fluid dynamics (CFD) software packages or, like in this study, external set of data to be used by the state estimator.

4.1.1 Performing a simulation with gPROMS®

Performing a simulation in gPROMS® requires a four-step procedure that can be summarized as follows:

1. variable types declaration: in this step, all the types of variables specified in the model (e.g. `Mass`, `Length`, ...) are defined with their units of measurement, default value, upper and lower bounds;

2. model entity creation: this new gPROMS® entity collects the mathematical description of the system. In the `Parameter` section, all the model parameters are defined. In the `Variable` section, all the variables involved in the model are declared. In the `Equation` section, all the equations of the model are implemented;

3. process entity creation: this entity is used to declare the simulation activity to be perfomed. The model to be used, the definition of the variables and parameters to saturate the degrees of freedom of the system, the intial conditions and the type (steady-state/dynamic) and length of the simulation are defined in this step;

4. run of the simulation with the execution of the process entity created in the previous step.

To apply state estimation in gPROMS®, the same procedure need to be carried out, with the following differences:

a) the solver to be defined in the process entity is the desired state-estimation solver instead of the standard DAE solver `DASOLV`;

b) three new foreign objects need to be created: the first contains the control vector $\mathbf{u}$, the second the measurement vector $\mathbf{y}$ and the third the specifications for the initial error covariance, the model uncertainty covariance and the measurement uncertainty covariance.

It is usually convenient to create a new process entity completely identical to the original process entity except for the two differences previously reported. The original process entity can then be executed to obtain the original model prediction, while the second process entity can be executed to obtain the estimator prediction. The improvements in the estimate of the state can then be easily traced comparing the results obtained from the two simulations.

4.2 Implementation issues

The implementation of a state estimator to an existing dynamic model requires a deep understanding of the general structure of the system to be considered. In particular, it is necessary to understand:

i) what is the mathematical structure of the system (ODE/ DAE/ PDAE), how many variables and equations are involved (differential and algebraic) and how many parameters;

ii) which variables represent an *input* for the model, in order to build the control vector $\mathbf{u}$;

iii) which variables represent an *output* for the model, in order to identify the state-vector $\mathbf{x}$ and in particular its differential part;

iv) what *on-line* plant measurements are available and which of them can be considered reliable, in order to build the measurement vector $\mathbf{y}$;

v) how the initial error covariance $\mathbf{P_0}$, the model uncertainty covariance $\mathbf{Q}$ and the measurement uncertainty covariance $\mathbf{R}$ should be specified and how this influences the results from the estimator[1];

[1] $\mathbf{P_0}, \mathbf{Q}$ and $\mathbf{R}$ need to be specified only for the differential state variables involved in the model , not for the algebraic ones.

vi) which are the parameters that have the highest influence on model prediction, in order to improve the estimator performance by their real-time adjustment;

vii) how frequently the state of the system should be updated and if this is compatible with the computational time required by the estimator.

In the following paragraphs all these aspects will be discussed.

4.2.1 Mathematical structure of the Furnace model

In chapter 2 the main equations involved in the Furnace model were presented. From an overall perspective, the system can be defined as a non-linear partial-differential algebraic system (PDAE) composed by 14495 equations, 227 of them differential and 14268 algebraic. In Tab. 4.1 the mathematical structure of the system is summarized.

Table 4.1: *Mathematical structure of the Furnace model .*

Type	PDAE
# of differential variables	227
# of algebraic variables	14268
# of model equations	14495
# of model parameters	13754

The state vector $\mathbf{x}$ is composed by 227 differential state variables and 14268 algebraic variables. From now on, the symbol $\mathbf{x}$ will always be associated with the differential part of the state vector. According to this, $\mathbf{x}$ collects all the differential variables x_i of the system :

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ x_{227} \end{bmatrix}. \tag{4.1}$$

The initial error covariance $\mathbf{P_0}$ and the model uncertainty covariance $\mathbf{Q}$ are then both 227×227 matrices:

$$\mathbf{P}_0 = \begin{pmatrix} \sigma_1^2|_0 & \sigma_1|_0\,\sigma_2|_0 & \cdots & \sigma_1|_0\,\sigma_{227}|_0 \\ \sigma_2|_0\,\sigma_1|_0 & \sigma_2^2|_0 & \cdots & \sigma_2|_0\,\sigma_{227}|_0 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{227}|_0\,\sigma_1|_0 & \sigma_{227}|_0\,\sigma_2|_0 & \cdots & \sigma_{227}^2|_0 \end{pmatrix} \tag{4.2}$$

$$\mathbf{Q} = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2 & \cdots & \sigma_1\sigma_{227} \\ \sigma_2\sigma_1 & \sigma_2^2 & \cdots & \sigma_2\sigma_{227} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{227}\sigma_1 & \sigma_{227}\sigma_2 & \cdots & \sigma_{227}^2 \end{pmatrix} \tag{4.3}$$

where the diagonal element σ_i^2 represents the (initial) model uncertainty variance on the i-th differential variable and the non-diagonal element $\sigma_i\sigma_j$ the (initial) model uncertainty covariance between the i-th and the j-th differential variables. Since $\sigma_i\sigma_j = \sigma_j\sigma_i$, the structure of $\mathbf{P}_0$ and $\mathbf{Q}$ is lower triangular.

4.2.2 Input-Output structure

A key information which is needed before applying state-estimation to an existing model is what is given to the model (its *inputs*) and what is calculated by the model (its *outputs*). This task is not trivial, especially with models with a large number of variables, as the one considered.

In Fig.4.1 the general structure of the model is presented. As can be noticed, the model is built in such a way that time-invariant or time-varying inputs can be supplied. In this last case, the time-varying model inputs are measurements coming from the real plant. The choice of time-invariant and/or time-varying inputs will be discussed in the next chapter.

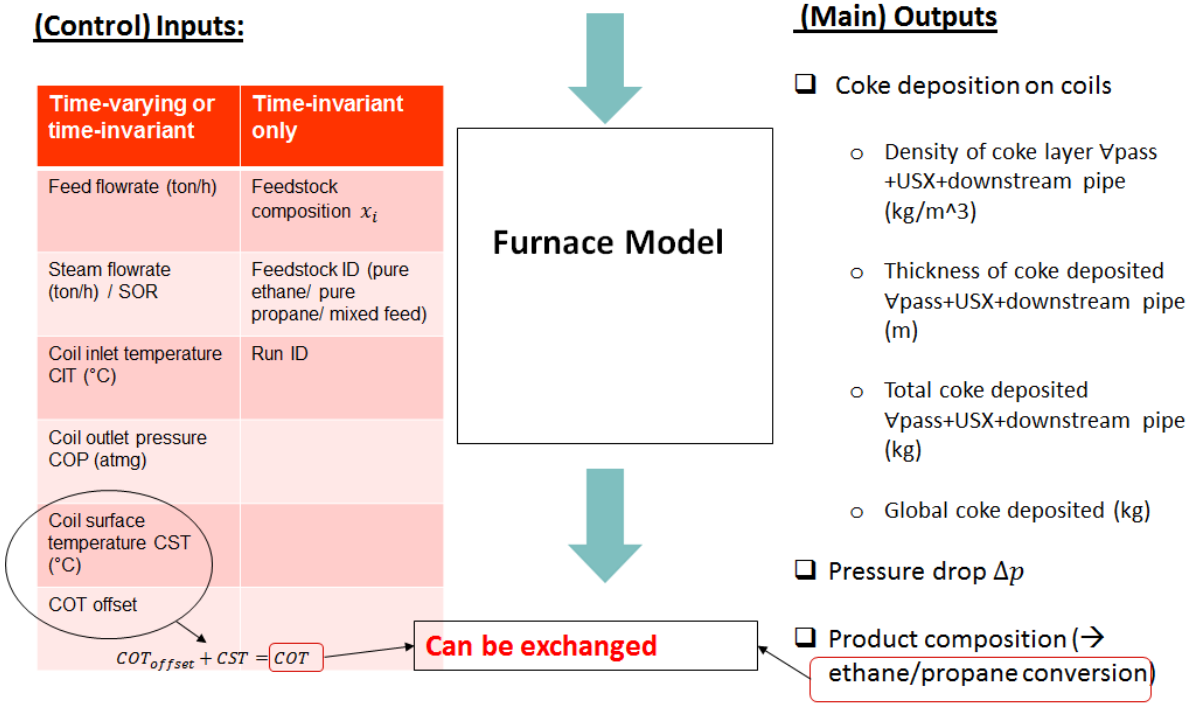


Figure 4.1: *Input-output structure of the Furnace model.*

The first input is the feed flowrate, expressed in ton/h. Three types of feedstock can be used, each one identified by a feedstock ID: pure ethane (ID:0), pure propane (ID:1) and mixed feed (propane+ethane, ID:2). A fixed value or time-varying data coming from the real plant can be used for this input, as suggested by Tab. 4.2. Another time-invariant or time-varying input that must be supplied to the model is the steam flowrate, eventually substituted by the steam to hydrocarbons ratio (SOR). The coil inlet temperature (CIT) and the coil outlet pressure (COP) are other two inputs. The model calculates the coil inlet pressure (CIP) using the pressure drop correlation [eq.(2.7)] given the pressure at the outlet of the coil. Other two inputs for the model are the external surface temperature at the outlet of the coil (CST) and the COT offset, defined as the difference between the gas-phase temperature at the outlet COT and the CST:

$$\mathrm{COT}_{offset} = \mathrm{COT} - \mathrm{CST}. \tag{4.4}$$

Measurements of CST, but not COT, are available from the real plant; this means that COT can be supplied only as a time-invariant input, while with time-varying inputs CST and a guessed or estimated COT offset must be specified. The model is also built in such a way that, if conversion of the main reactant (ethane/propane) is given as an input, COT is then calculated as a model output and *viceversa*.

Table 4.2: *Model inputs to be collected in the control vector* **u**.

Input	**UOM**	**Plant data available?**
Feed flowrate	ton/h	Yes
Steam flowrate	ton/h	Yes
Coil inlet temperature (CIT)	degC	Yes
Coil outlet pressure (COP)	atmg	Yes
Coil surface temperature (CST)	degC	Yes
COT offset	degC	No

The most important model output is the coke deposition on the coil. Different variables are related to coke deposition: the mass of coke deposited per unit volume of clean tube [eq.(2.24)], the thickness of coke deposited on the tube wall [eq.(2.30)], the total mass of coke deposited along the coil [eq.(2.31)]. Another important output is the so called tube metal temperature (TMT), omitted in Fig. 4.1, defined as the temperature at the tube surface at the axial position $z/L = 0.88$. Only few manual[2] measurements of TMT are available. As discussed earlier, another variable that is calculated by the model is the pressure drop along the coil: given COP as an input, the model calculates the coil inlet pressure (CIP) and so the pressure drop. In addition, the model allows to calculate the product composition at the outlet, in particular the yield of ethylene production and the conversion of the main reactant (if this is used as an input, COT becomes an output). Another important output is the fuel flowrate, which is calculated as discussed in §2.2.4. The most important outputs that have been monitored during the simulations are summarized in Tab.4.3.

Table 4.3: *Most important outputs of the Furnace model.*

Ouput	**UOM**	**Plant data available?**
Density of the coke layer along each pass	kg/m^3	No
Total mass of coke deposited on each pass	kg	No
Thickness of coke deposited along each pass	m	No
Tube metal temperature (TMT)	degC	Yes but not on-line
Pressure drop	atmg	Yes
Fuel flowrate	kg/s	Yes
Product composition	(-)	Yes

4.2.3 Plant data and reliability

As discussed in chapter 3, a state estimator can update the state of a system given a set of few and reliable real-time measurements coming from that system. Therefore, its intrinsic use is for *on-line* applications. However, the robustness of the estimator can be tested *off-line* if a predetermined set of data is available. In this project, four different

[2]TMT is measured by the plant operators using an optical pyrometer at the axial poistion specified.

sets of hourly data were available: two of them for pure ethane feed, the other two for pure propane. Each set of data refers to a different run-length of the furnace, for a total of $\sim$ 55 days of operation (the run-length of the furnace depends on the set of data considered). The general overview is presented in Tab. 4.4.

Table 4.4: *Different sets of plant data available.*

Feedstock	**Furnace run #1**	**Furnace run #2**
Pure ethane	Set of data #1	Set of data #2
Pure propane	Set of data #3	Set of data #4

The measurements available for each data-set are presented in Tab. 4.5. It can be noticed that measurements are available for some of the model inputs and some of the model outputs. As already discussed, the measurements of TMT are only occasionally taken by the plant operators with a pyrometer, so only a limited number of TMT measurements is available per each data set.

Table 4.5: *Measurements available per each data-set and their mathematical characterization in the Furnace model.*

Measurement	**UOM**	**Variable type**
Feed flowrate	ton/h	Input
Steam flowrate	ton/h	Input
CIT	degC	Input
CST	degC	Input
COP	atmg	Input
Pressure drop	atmg	Output
Mass fraction of hydrogen at the outlet	(-)	Output
Mass fraction of methane at the outlet	(-)	Output
Mass fraction of ethylene at the outlet	(-)	Output
Mass fraction of ethane at the outlet	(-)	Output if COT is an input
Mass fraction of propylene at the outlet	(-)	Output
Mass fraction of propane at the outlet	(-)	Output if COT is an input
Fuel flowrate	ton/h	Output
TMT	degC	Output

If time-varying control inputs are used, the measurements available for the input variables are collected in the control vector $\mathbf{u}$. These measurements are subject to uncertainty: however, as discussed in the previous chapter, one of the first assumption that is made when deriving the recursive equations for the EKF[3] is that no uncertainty on the control vector $\mathbf{u}$ is considered (§3.2). This problem can be overcome by implicitly taking into account the uncertainty on the model inputs by increasing the model uncertainty covariance $\mathbf{Q}$. This solution is much simpler than deriving a completely new algorithm for the EKF assuming a noisy control vector. This point will be explained with more detail in §5.2.2.

The choice of the output variables to be collected in the measurement vector $\mathbf{y}$ and supplied to the estimator must be careful. The target is to give to the estimator the lowest number of measurements available which allows to obtain a correct estimate of the state of the system. Because of this, the reliability of the measurements collected in the

[3]The same is done for the other state estimators discussed in the previous chapter.

measurement vector is a fundamental point of concern. Deeply inaccurate measurements would cause wrong correction on the state of the system by the estimator.
The easiest way to obtain information on the on-line measurements coming from the real plant is to rely on experience. The availability and reliability of measurements strongly depend on the conditions and state of repair of the measuring devices in the real plant. In Tab. 4.6 a general overview of the reliability of the plant measurements is presented. It is important to notice that what is written in Tab. 4.6 should be considered only as a rough guideline since the actual availability of the measurements is subject to a series of circumstances that may vary during the plant operation.

Table 4.6: *Reliability of the plant data available.*

Measurement	**Reliable?**
Pressure drop	Yes
Mass fraction of hydrogen at the outlet	No
Mass fraction of methane at the outlet	No
Mass fraction of ethylene at the outlet	No
Mass fraction of ethane at the outlet	No
Mass fraction of propylene at the outlet	No
Mass fraction of propane at the outlet	No
Fuel flowrate	Yes
TMT	Yes but manual

4.2.4 Model parameters

The number of parameters involved in the Furnace model is enormous (13754). Luckily, not all of these parameters have the same influence on model prediction. Some of them can deeply affect the results that can be obtained from the simulations, while others have little influence on the final results. Understanding which are the most influential parameters is of fundamental importance, since they can be slightly adjusted in real-time by the estimator (§3.7) so as to 'adapt' the model to the on-line data coming from the real plant. In other words, the real-time adjustment of these parameters can help the estimator to obtain a good estimate of the state with minimal effort. In Tab. 4.7 some of the most important parameters of the system are presented. Their relative importance has been obtained with a trial-and-error sensitivity analysis on the original model. In the same table, it is also shown the specific model in which they are involved and their relative importance on model prediction. In the original model, all the parameters reported in Tab. 4.7 have been estimated using a gPROMS® Parameter Estimation entity and a set of 2 averaged daily data. From Tab. 4.7 it can be noticed that the two

Table 4.7: *Most important parameters of the Furnace model .*

Symbol	**Description**	**Model involved**	**Influence**
$\log_{10} k_{c1}(T_{ref})$	Pyrolytic coking rate pre-exponential factor	Coking model [eq.(2.25)]	High
$\log_{10} k_{c2}(T_{ref})$	Pyrolytic coking rate pre-exponential factor	Coking model [eq.(2.25)]	Very high
$\log_{10} k_{cat}(T_{ref})$	Catalytic coking rate pre-exponential factor	Coking model [eq.(2.25)]	Low
α	Coil roughness parameter	Pressure drop [eq.(2.7)]	Low
λ_{coke}	Coke thermal conductivity	Heat transfer model	Low
$L_{down.\,pipe}$	Downstream pipe length	Tube 1D	Low

most influential parameters for the model are the pre-exponential factors of the kinetic constants involved in the pyrolytic coking-rate expression [eq.(2.25)]. Instead, the pre-exponential factor of the catalytic-coking kinetic constant, the coil roughness parameter, the coke thermal conductivity and the length of the downstream pipe show little influence on model prediction. The relative high importance of the first two parameters can be understood considering that the pyrolytic mechanism is the main responsible for coke deposition on the coil and that coking deeply influences the key performance indicators (KPI: yield, conversion) of the furnace. In particular, the model reveals to be extremely sensitive to the pre-exponential factor of the kinetic constant k_{c2}: a small adjustment of its value can deeply modify the state of the system. This aspect has been taken into account when testing the EKF on the model.

4.2.5 Specifications to the estimator

As already discussed, when a state-estimation algorithm need to be applied to an existing dynamic model, the following specifications must be given *a priori* to the estimator:

- the initial state vector $\mathbf{x}_0$. Considering only the differential state variables, $\mathbf{x}_0$ represents the set of initial conditions for the state;
- the initial error covariance $\mathbf{P}_0$;
- the model uncertainty covariance $\mathbf{Q}$;
- the measurement uncertainty covariance $\mathbf{R}$.

The choice of the initial error covariance is related to the uncertainty in the knowledge of the initial state of the system. If the initial state of the system is not well known, high values for the diagonal (and eventually non-diagonal) elements of $\mathbf{P}_0$ should be used. Viceversa, if the initial state of the system is perfectly known, $\mathbf{P}_0$ should be chosen such that $\mathbf{P}_0 = \mathbf{0}$. Generally (Cheng *et al.*, 1997) the choice of the initial error covariance has little effect on the estimator performance.

The model uncertainty covariance $\mathbf{Q}$ is a measure of the degree of confidence in the model to describe the behavior of the dynamic system. High values for the elements of $\mathbf{Q}$ should be used if the confidence in the model is very low, and low values should be used if there is high confidence in the model. The choice of $\mathbf{Q}$ greatly influences the performance of the estimator : caution is required in order to avoid wrong results.

The measurement uncertainty covariance $\mathbf{R}$ weights the uncertainty related to the measurements supplied to the estimator. In theory, $\mathbf{R}$ should be a measure of the measuring device errors. High values for the elements of $\mathbf{R}$ should be related to very noisy measurements, and viceversa. In practice, what is really important is the relative value between the model uncertainty covariance and the measurement uncertainty covariance. In fact, suppose x_i is a state variable for the system, σ_{ii}^2 is the diagonal element of $\mathbf{Q}$ related to that state variable[4] and $\mathbf{R}$ is a diagonal matrix with equal diagonal elements $r_i^2 = r^2$. If $\sigma_{ii}^2 < r^2$, that means that, for the given state variable, the confidence in the model is higher than in the real-time measurements supplied to the estimator. Thus, during the update step, the estimator will perform only a '*soft*' correction on the value

[4]It is implicitly assumed that $\mathbf{Q}$ is diagonal, so that the model uncertainty on the state variable i does not affect the uncertainty on the state variable j.

predicted by the model equations for x_i. If $\sigma_{ii}^2 > r^2$, the confidence in the available measurements is higher than the confidence in the model, so a strong correction will be performed by the estimator in the update step to the predicted value of x_i.

From a mathematical perspective, $\mathbf{P}_0$ and $\mathbf{Q}$ are both $n \times n$ matrices, while $\mathbf{R}$ is a $p \times p$ matrix. If the uncertainty on the state (or measurement) variable i does not affect the uncertainty on the state (or measurement) variable j, a diagonal structure can be used for $\mathbf{P}_0$, $\mathbf{Q}$ and $\mathbf{R}$:

$$\mathbf{P}_0 = \begin{pmatrix} \sigma_{11}^2|_0 & 0 & \cdots & 0 \\ 0 & \sigma_{22}^2|_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \sigma_{nn}^2|_0 \end{pmatrix} \tag{4.5}$$

$$\mathbf{Q} = \begin{pmatrix} \sigma_{11}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{22}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \sigma_{nn}^2 \end{pmatrix} \tag{4.6}$$

$$\mathbf{R} = \begin{pmatrix} r_{11}^2 & 0 & \cdots & 0 \\ 0 & r_{22}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & r_{pp}^2 \end{pmatrix}. \tag{4.7}$$

If there is a correlation between the uncertainty on a state (or measurement) variable and another one a lower triangular structure should be used for $\mathbf{P}_0$ or $\mathbf{Q}$ or $\mathbf{R}$[5]:

$$\mathbf{P}_0 = \begin{pmatrix} \sigma_{11}^2|_0 & 0 & \cdots & 0 \\ \sigma_2|_0\,\sigma_1|_0 & \sigma_{22}^2|_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ \sigma_n|_0\,\sigma_1|_0 & \sigma_n|_0\,\sigma_2|_0 & \cdots & \sigma_{nn}^2|_0 \end{pmatrix} \tag{4.8}$$

$$\mathbf{Q} = \begin{pmatrix} \sigma_{11}^2 & 0 & \cdots & 0 \\ \sigma_2\sigma_1 & \sigma_{22}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ \sigma_n\sigma_1 & \sigma_n\sigma_2 & \cdots & \sigma_{nn}^2 \end{pmatrix} \tag{4.9}$$

$$\mathbf{R} = \begin{pmatrix} r_{11}^2 & 0 & \cdots & 0 \\ r_2r_1 & r_{22}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ r_pr_1 & r_pr_2 & \cdots & r_{pp}^2 \end{pmatrix}. \tag{4.10}$$

Introducing a correlation between the uncertainties on different state or measurement variables is usually a burden for the estimator and should be avoided if there is no clear evidence of this correlation (Gillijns *et al.*, 2006). The choice of the specifications to give to the estimator is case-dependent and no general rules can be applied. In general, having a deep knowledge of the dynamic model and of the available measurements can help to

[5] $\sigma_i\sigma_j = \sigma_j\sigma_i$.

choose the best structure and values for $\mathbf{P}_0$, $\mathbf{Q}$ and $\mathbf{R}$, thus improving the performance of the estimator. In this study, the structures to be used for $\mathbf{P}_0$, $\mathbf{Q}$, $\mathbf{R}$ have been considered as additional degrees of freedom to be tuned in order to obtain the best possible results from the estimator.

4.2.6 Update frequency and computational issues

It has already been pointed out that a state estimator, independently of the type considered (EKF/UKF/EnKF/PF), works according to a two-step procedure: in the prediction step, the state of the system at time k is forecasted, given the state at time $k-1$, using model equations. In the update step, the state of the system is corrected taking into account the on-line measurements coming from the real plant. The 'amount' of correction that is performed by the estimator to the state vector is determined by the relative values of the model and measurement uncertainty covariances.

One choice that must be done when applying state estimation is how frequently the state of the system should be updated to obtain a reasonable estimate. In other words, it must be decided how frequently the on-line measurements should be used by the estimator to update its prediction. The estimator update frequency does not necessarily need to coincide with the frequency at which the on-line measurements are available in the real plant. In fact, the state vector should be updated only when a considerable 'knowledge' of the state is brought in by the measurement vector. If the update frequency is too high, the computational burden increases without improving the performance of the estimator. In this study, the data are collected on-line in the real plant every minute. However, since coking is a process which occurs over a longer time-scale (§2.2), there is no need to update the state of the system (i.e the state of coking) with this frequency. A reasonable choice seems that of updating the state vector every hour : when installed on-line, the estimator will receive the desired measurements every hour and will update its prediction according to these data. According to this, in all the simulations the estimator has been tested using different sets of hourly plant data.

Another important aspect that must be taken into account is the computational time required by the estimator to propagate the state vector during the prediction step. It should be clear that, if this time is higher than the frequency at which measurements are supplied to the estimator, its on-line implementation is not feasible. This point has been considered in all the simulations and the results obtained will be shown in the next chapters.

4.3 Project organization

All the activities carried out during this project have been organized according to the hierarchical structure presented in the flowchart of Fig. 4.2.

First of all, different runs of the original model have been performed using time-invariant values for the control inputs of Tab. 4.2. These runs were mainly meant to gain experience with the model and to understand the time-behavior of the main outputs of Tab 4.3. Then, real plant data have been used for the model inputs and the predicted outputs have been compared with the available data. Considering the substantial mismatch between the model prediction and these data (§), the choice to

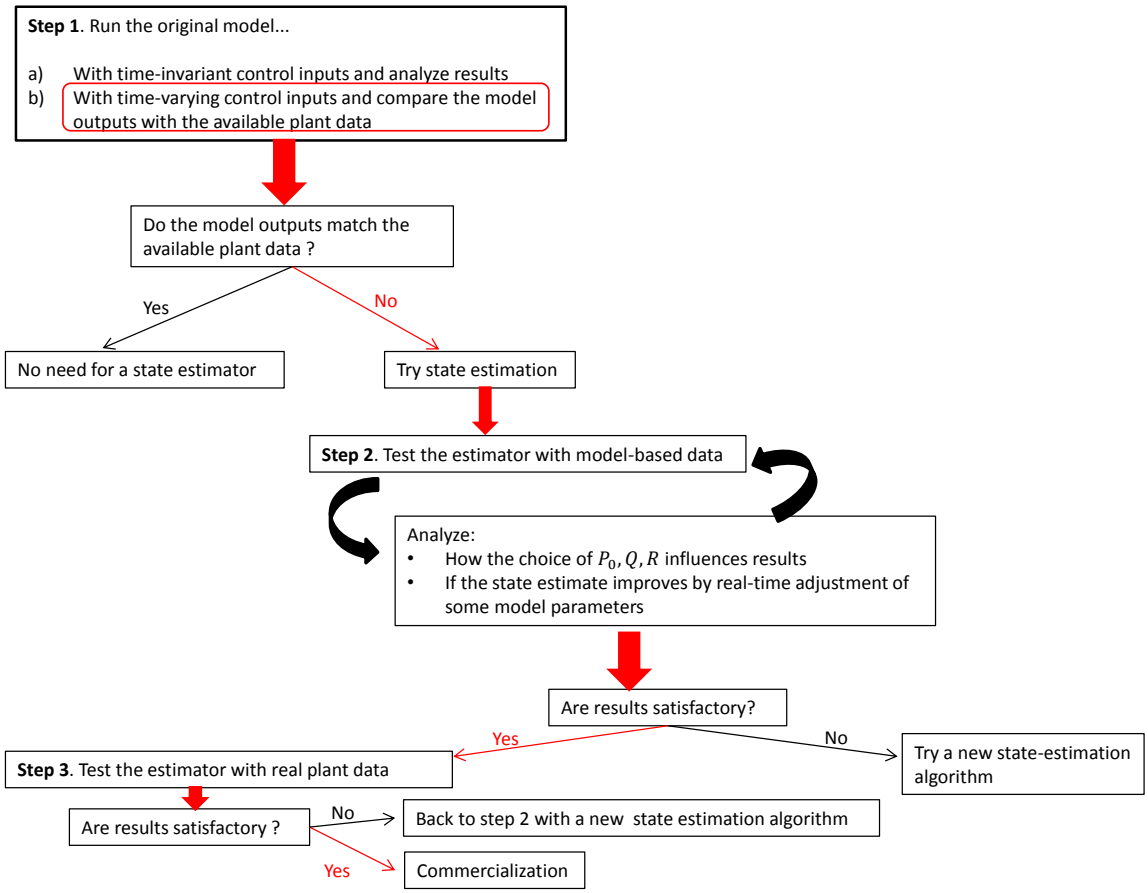
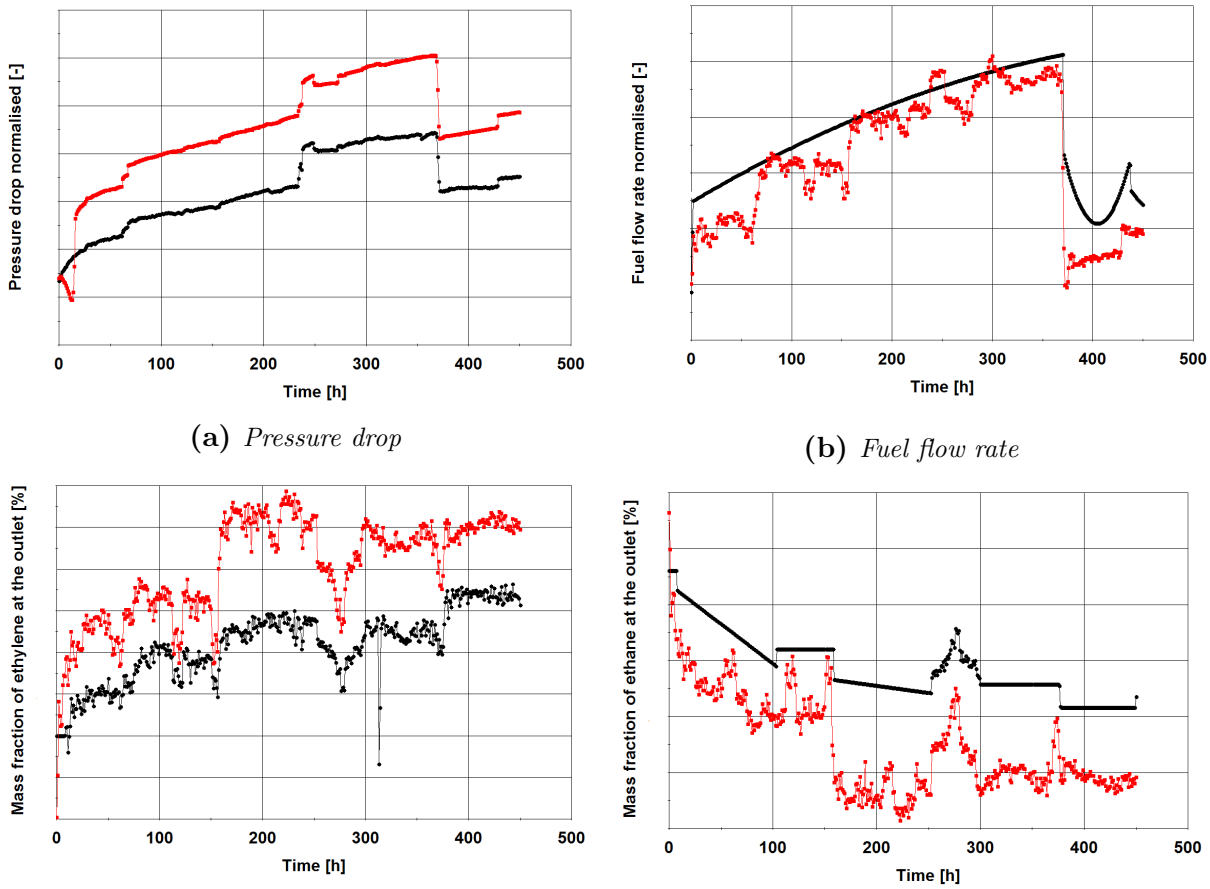


Figure 4.2: *Hierarchical organization of the project. The red-path represents how the project evolved.*

test a state estimator to improve the prediction of the state of the system, especially its state of coking, has been done. As already specified, the estimator that has been tested is the EKF. First, noisy measurements of the output variables of interest have been generated using the model prediction and adding a white gaussian random noise to simulate the real measuring device noise. The performance of the estimator have then been analyzed using these model-based measurements. The systematic procedure used to the test the estimator at this stage of the project is exhaustively explained in the next chapter. Extremely important aspects, such as the influence of the structure of $\mathbf{P}_0$, $\mathbf{Q}$ and $\mathbf{R}$ on the estimator performance and the importance of the real-time adjustment of some of the model parameters have been addressed at this stage. Considering the promising results obtained, the next step has been that of testing the estimator using real plant data to be collected in the measurement vector $\mathbf{y}$. The agenda and the results obtained at this stage are the subject of chapter §6.

4.4 Need for a state estimator

The main weaknesses of the model described in chapter §2 have already been *qualitatively* discussed in §2.5. However, to understand why state estimation can be beneficial for this model, it is necessary to *quantitatively* asses its performance, i.e. the gap between its predictions and the observations coming from the real plant need to be determined.



(a) *Pressure drop*

(b) *Fuel flow rate*

(c) *Mass fraction of ethylene at the outlet*

(d) *Mass fraction of ethane at the outlet*

Figure 4.3: *Comparison between the plant observations (black lines) and the model predictions (red lines) for some of the output variables of the Furnace model. Set of data #1.*

In Fig. 4.3 the model predictions for some of the most important output variables of Tab. 4.3 are compared with their observations in the real plant. These results refer to the first set of plant data available (ethane feed), but analogous considerations can be done for the other sets of data. From the same figure it can be noticed that the model predictions of pressure drop, fuel flow rate and product composition (ethane and ethylene) do not show a good agreement with the experimental data: the general trend is tracked, but there is an off-set between the predicted values and the expected values of these state variables. In particular, the prediction of pressure drop is not satisfactory. Given this and the fact that pressure drop is the key-variable that can be tracked to deduce the state of coking of the furnace, it can be claimed that the model is not able to give a reliable description of the state of the system, especially of the coke deposition on the coil. Furthermore, it must be said that using a first-principles approach to improve the coking prediction (i.e. by increasing the complexity of the coking mechanism and thus of the coking rate) can be a time-consuming task that could not necessarily give the expected improvements. Thus, the usage of a state estimator to obtain a real-time accurate description of the state of coking of the furnace seems the best possible solution for this problem.

Looking at Fig. 4.3a, it can be noticed that the model predicts a decrease in pressure

drop during the first hours of operation of the furnace. This non-physical behavior derives from a numerical problem that arises only when a pure ethane feed is considered and is related to the structure of the coking rate expression:

$$r_{coking}(z) = k_{cat}\left(\frac{c_{cat}(z)}{c_{cat}^{max}}\right)c_{C_2H_4}(z) + k_{c1}c_{C_2H_4}(z) + k_{c2}\left(\frac{c_{C_2H_4}^2}{c_{C_3H_6}}\right). \tag{2.25}$$

When the feed is pure ethane, due to the very low concentration of propylene during the first hours of operation, the last term of (2.25) becomes dominant and can lead to a wrong prediction of the available cross sectional area for the gas-phase in the tube [eq.(2.7)], which translates into a wrong prediction of the pressure drop along the coil. This non-physical behavior is however confined only to the very first hours of the furnace run, and the correct trend is rapidly recovered after few hours (Fig. 4.3a).
From the previous discussions, it should be clear that the main goal of the usage of state estimation on the model presented in chapter §2 is to obtain a reliable and robust estimate of the state of coking of the furnace, since its prediction is not satisfactory. To be more accurate, the general purpose of state estimation is to improve the prediction of all the state variables involved in the model (coking, fuel consumption, TMT etc. . .), given the consistent mismatch between the original model prediction and the available plant data for these state variables (§4.4).

Chapter 5

Results with model-based data

In this chapter the methods and the results obtained in the Step 2 of the organization chart of Fig. 4.2 are presented. At this stage, the estimator has been tested using measurements generated with the model described in chapter 2 and its performance has been analyzed.

First, the overall strategy that has been followed to test the estimator using model-based measurements is given. Then, the results obtained without real-time adjustment of some of the model parameters are presented. A discussion on how these results could be improved is then performed. Finally, the results obtained with a real-time estimation of some model parameters are shown for both time-invariant and time-varying inputs.

5.1 Overall procedure

Testing the estimator with model-based measurements is a preliminary task that is extremely useful to obtain some key pieces of information that are needed before using real plant data. Some of them are:

a) how many and what measurements are needed to obtain a good an robust estimate of the state of the system (independently of the fact that they could or could not be available in the real plant);

b) what is the effect of the structure of $\mathbf{P}_0$, $\mathbf{Q}$ and $\mathbf{R}$ on the estimator prediction;

c) which are the main weaknesses of the model and what model parameters can improve the results if adjusted in real-time by the estimator.

To give an answer to these questions, a systematic procedure has been adopted that can be summarized as follows:

1. first, noisy measurements are generated using the original model prediction and adding a random white gaussian noise to each measurement variable;

2. then, a 'wrong' model is created by slightly modifying the initial conditions (ICs) of some of the state variables or the numerical value of some of the parameters of the original model;

3. finally, since measurements are generated using the original unmodified model, state estimation is applied using the 'wrong' model to verify if the estimator is able to follow the correct state of the system despite using a wrong model with modified ICs and/or parameters.

The practical implementation of all these points in gPROMS® is discussed in the next subsections.

5.1.1 Model-based generation of noisy measurements

The first step to be carried out is the generation of noisy measurements to be collected in the measurement vector $\mathbf{y}$ and then supplied to the estimator. The assumption that has been made is that the measurement noise is a white gaussian additive noise with zero mean and covariance $\mathbf{R}$. In other terms, if $y_i^{pred.}$ is the model prediction of the i-th measurement variable, the noisy measurement $y_i^{meas.}$ to be collected in $\mathbf{y}$ is generated as follows:

$$y_{k,i}^{meas.} = y_{k,i}^{pred.} + v_i \tag{5.1}$$

$$v_i \sim (0, r_i^2). \tag{5.2}$$

v_i should represent the measuring device noise for the measurement variable i and its choice should be done accordingly. For example, if $y_i^{pred.}$ is the predicted value of the fuel consumption, a high-frequency and low-intensity noise should be added since flow rate measurements are usually characterized by this type of noise. However, if $y_i^{pred.}$ is the tube metal temperature, a low-intensity and low-frequency noise should be added to the model prediction since temperature measurements are usually smoother and less noisy then flow rate measurements.

To generate noisy measurements using the original model in gPROMS®, a new process entity called `Generate_Measurements` has been created. This entity has the same structure of the process entity used to simulate the original Furnace model, but in the `Schedule` section noisy measurements of the desired output variables are generated adding a white gaussian distributed noise. An example of the `Schedule` section used to generate, in this specific case, model-based measurements of pressure drop, fuel consumption and TMT is reported below.

```
1  SCHEDULE
2   WHILE TIME ≤ 450.0 DO
3    SEQUENCE
4     CONTINUE FOR 1
5      RESET
6      Flowsheet.Furnace_USX_1.pressure_drop_m_kgf_cm2:=
7      OLD(Flowsheet.Furnace_USX_1.pressure_drop_kgf_cm2)+  NORMAL(0,1E-2);
8
9      Flowsheet.Furnace_USX_1.TMT_m_degC:=
10     OLD(Flowsheet.Furnace_USX_1.TMT_degC)+NORMAL(0,1E-3);
11
12     fuel_mass_flowrate_m:=
13     OLD(Flowsheet.Furnace_USX_1.Burner(1).Fuel.SP_F_w.mass_flowrate)+ ...
           NORMAL(0,1E-2);
14    END
15   END   #WHILE
16  END  # SEQUENCE
```

All or some of the noisy measurements obtained by executing this process entity can then be collected in the measurement vector $\mathbf{y}$ and passed to the estimator using an external foreign object (§4.1.1). Measurements of any model output variable can be created using this approach independently of their availability or unavailability in the real plant. The estimator can then be tested using different measurement variables to understand which are the most influential to obtain a good estimate of the state of the system.

5.1.2 Model modification

Once the noisy measurements have been generated using the original model (OM), the next step is to slightly modify some of its ICs or parameters, thus creating a *wrong* model (WM) whose predictions will be different from the ones of the original model. This WM will then be used by the estimator to propagate the state of the system from a given time-step to the next one: since measurements are generated using the correct OM, the state of the system should be 'driven' towards its correct value by the measurements supplied to the estimator during the update step.

As already said, the initial state of the system (i.e. the ICs for some or all the state differential variables) and/or some of the model parameters can be modified to create the WM. In the last case, to obtain an appreciable deviation between the OM prediction and the WM prediction, only the most influential parameters (§4.2.4) should be modified. Typical deviations are between -10% to +10% with respect to their nominal value. Higher deviations should be avoided to prevent very steep corrections during the update step which can cause the estimator to diverge. An example of a 2 % deviation on $\log_{10} k_{c1}(T_{ref})$ and $\log_{10} k_{c2}(T_{ref})$[1] (§4.2.4) is reported below.

```
1  # Model parameters   gabrieleb
2  Flowsheet.Furnace_USX_1.Coil.log10kc2_nominal_value := nom_value*0.98;
3  Flowsheet.Furnace_USX_1.Coil.log10kinetic_constant_coking("c1") ...
       :=nom_value;
4  Flowsheet.Furnace_USX_1.Coil.log10kc4_nominal_value := nom_value *1.02;
5  Flowsheet.Furnace_USX_1.Coil.log10kinetic_constant_coking("cat1") ...
       :=nom_value;
6  Flowsheet.Furnace_USX_1.Coil.roughness_par:= nom_value;
7  Flowsheet.htc_adjustment_C2_feed := nom_value;
8  Flowsheet.htc_adjustment_C3_feed := nom_value;
9  Flowsheet.Furnace_USX_1.Downstream_Pipe.tube_length:= nom_value;
10 Flowsheet.lambda_coke_Tref_C2_feed:= nom_value;
11 Flowsheet.lambda_coke_Tref_C3_feed:= nom_value;
12 Flowsheet.Furnace_USX_1.Coil.n_C2H4:= nom_value;
13 Flowsheet.Furnace_USX_1.COT_offset_nominal_value:= nom_value;
```

A new process entity called `Furnace_wrong` has been created to simulate the WM and collect its prediction. The results obtained from this process entity can then be compared with the ones obtained from the original model and with the ones obtained from the estimator.

5.1.3 State estimation

Once the model-based measurements to be collected in the measurement vector have been created (§5.1.1) and some of the original model ICs or parameters have been modified (§5.1.2), state estimation can be performed.

A new process entity called `Furnace_S_E` has been created: it has the same structure of the process entity `Furnace_wrong` (with modified ICs and/or parameters), but it uses a different solver (the state-estimation algorithm desired, in this case the EKF) and is combined with three external FOs. The first, called `Furnace_S_E.control`,

[1] In the model, the nominal value of $\log_{10} k_{c1}(T_{ref})$ is identified as `log10kc2_nominal_value` and $\log_{10} k_{c2}(T_{ref})$ as `log10kc4_nominal_value`.

contains the control vector **u**; in the second, called `Furnace_S_E.measurement`, the measurement vector **y** is stored; the third, called `Furnace_S_E.covariance`, contains the initial error covariance, the model uncertainty covariance and the measurement uncertainty covariance.

The results obtained with this process entity represent the estimator prediction and can then be compared with the original model prediction and the wrong model prediction. The estimator performance is satisfactory if its prediction, for all the relevant state variables, is in good agreement with the original model prediction [2].

5.2 Results

To give an answer to the questions discussed in §5.1, the estimator has been tested, following the procedure described in the previous sections, using different model-based measurements, different deviations on model parameters and/or ICs and different specifications for $\mathbf{P}_0$, $\mathbf{Q}$ and $\mathbf{R}$. The first runs have been performed using time invariant control inputs; time-varying control inputs have been used only at a later stage.

In the following sections, the main results obtained at this stage of the project are presented. First, the results obtained without any real-time adjustment of some of the model parameters are shown. Then, how these results change with a real-time adjustment of some of the model parameters is discussed.

5.2.1 No real time adjustment of model parameters

The time-invariant control inputs used during the first runs are reported in Tab. 5.1. It should be noticed that COT is directly passed as an input to the estimator, since its time-invariant value can be obtained as the sum of the time-invariant value of CST and the fixed value chosen for the COT offset (§4.2.2).

Table 5.1: *Time invariant control inputs used in the first runs of the estimator.*

Input	Symbol	UOM
Feed flowrate	w_{feed}	ton/h
Steam flowrate	w_{steam}	ton/h
Coil inlet temperature	CIT	degC
Coil outlet pressure	COP	atmg
Coil outlet temperature	COT	degC

The time-invariant inputs have been collected in the control vector **u** and passed to the estimator using the FO `Furnace_S_E.control`:

$$\mathbf{u} = \begin{bmatrix} w_{feed} \\ w_{steam} \\ \text{CIT} \\ \text{COP} \\ \text{COT} \end{bmatrix}. \tag{5.3}$$

[2] It is again emphasized that the measurements supplied to the estimator are generated using the original model with unmodified ICs and/or parameters.

The structures of **Q** and **R** and their relative values have been changed to understand their influence on the estimator prediction taking into account the remarks discussed in §4.2.5. As regards the initial error covariance $\mathbf{P}_0$, it is important to remember that the dynamic simulation starts from clean-tube conditions: there is no uncertainty related to the initial state of the system, so the correct choice for $\mathbf{P}_0$ is the null matrix:

$$\mathbf{P}_0 = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix}. \tag{5.4}$$

This is true only when the WM is created without changing the ICs of all the state variables (i.e. only some model parameters are modified); if not, the initial error covariance should take into account the uncertainty on the modified initial conditions of the selected state variables. However, all the simulations performed at this stage have been carried out by modifying only some of the model parameters to create the WM. The ICs for all the state variables have not been modified, thus the structure used in all simulations for $\mathbf{P}_0$ is that of eq. (5.4).

As already said, the model uncertainty covariance **Q** and the measurement uncertainty covariance **R** have been changed to understand their influence on the estimator performance. In all the simulations, a diagonal structure for both **Q** and **R** has been maintained, since no clear correlation between the different state/measurement variables could be recognized. The WM has always been generated by modifying only the pre-exponential factors $\log_{10} k_{c1}(T_{ref})$ and $\log_{10} k_{c2}(T_{ref})$ in the coking rate expression, given their strong influence on the model prediction (§4.2.4).

In Fig. 5.1 an example of the results obtained in one of the simulations performed is shown. Black lines represent the OM prediction, red lines the WM prediction, blue lines the estimator prediction. These results have been obtained simulating 450 hours of furnace operation ($\sim$ 19 days) and using the following specifications:

- model-based measurements of fuel flow rate, TMT and pressure drop have been generated and collected in the measurement vector **y**:

$$\mathbf{y} = \begin{bmatrix} w_{fuel} \\ \text{TMT} \\ \Delta p \end{bmatrix}; \tag{5.5}$$

- a -2% variation on the nominal value of $\log_{10} k_{c1}(T_{ref})$ and a +2% variation on the nominal value of $\log_{10} k_{c2}(T_{ref})$ have been used to build the WM;

- the following structure has been used for **Q**:

$$\mathbf{Q} \begin{pmatrix} 0.01 & 0 & \cdots & 0 \\ 0 & 0.01 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 0.01 \end{pmatrix}. \tag{5.6}$$

and for $\mathbf{R}$:

$$\mathbf{R} = \begin{array}{c} \begin{array}{ccc} \Delta p & \text{TMT} & w_{fuel} \\ \downarrow & \downarrow & \downarrow \end{array} \\ \left(\begin{array}{ccc} 10^{-4} & 0 & 0 \\ 0 & 10^{-3} & 0 \\ 0 & 0 & 10^{-6} \end{array} \right) \end{array} . \tag{5.7}$$

$\mathbf{P}_0$ has been chosen as in eq. (5.4).

The output variables reported in Fig. 5.1 are pressure drop, TMT, fuel flow rate, the mass fraction of ethane at the coil outlet, the density of coke layer on pass #2 and the total mass of coke deposited in all passes.

From Fig. 5.1, it can be noticed that the estimates of the three output variables for which model-based measurements are supplied (pressure drop, TMT, fuel flow rate) are in good agreement with the OM prediction. In other words, the estimator, despite using a wrong model with modified parameters, is able to 'adjust' the estimates of these state variables to match the OM prediction. This is not surprising (§4.2.5), as a lower value for the measurement uncertainty covariance has been used with respect to the model uncertainty covariance, as suggested by eq.(5.6) and (5.7).
Two comments need to be made on Fig. 5.1a and Fig. 5.1c:

- pressure drop decreases during the first hours of furnace operation according to the OM and WM prediction. This is clearly a non-physical behavior due to a model mismatch, as explained in §4.4;

- according to the OM and the WM, fuel flow rate decreases with time. This is only due to the fact that COT is given to the model as a fixed input: in real plants, fuel flow rate is usually gradually increased to obtain a nearly steady ethylene yield. This correct behavior will be seen when using real plant data (§6).

Looking at Fig. 5.1d,e,f it can be seen that the estimates of the state variables for which measurements are not supplied to the estimator are not satisfactory. This is especially true for the total coke deposited in all passes, which is a direct measure of the state of coking of the system[3]. In Fig. 5.1e only the density of coke layer on pass #2 is reported, but results for all the other passes are similar and the estimation is still not satisfactory. Different runs have been performed trying to obtain a better estimate of the state of coking by:

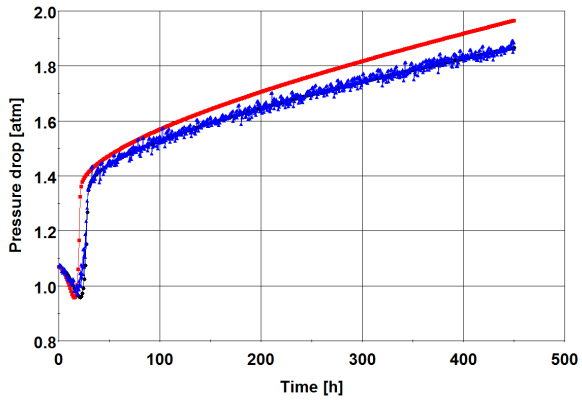
1. decreasing the measurement uncertainty on pressure drop[4] and/or increasing the model uncertainty on all the state variables involved in the coking model;

2. giving different deviations (in both magnitude and sign) on the coking model parameters to test the sensibility of the estimator performance on the state/covariance propagation in the forecast step using model equations

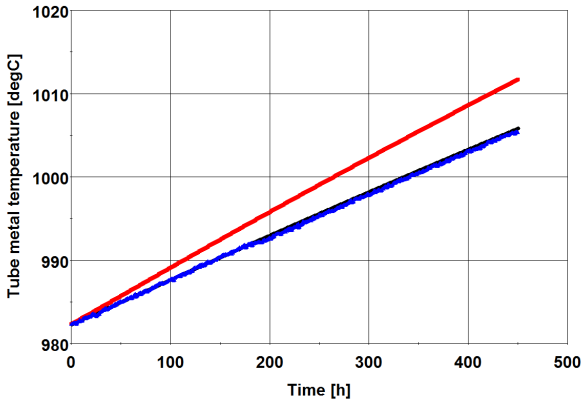
but little improvements on the results have been obtained. As an example, a comparison between the results obtained for thetotal coke deposited in all passes with the first and

[3]It should be remembered that obtaining an accurate prediction of the state of coking of the furnace is the main objective of this Project.

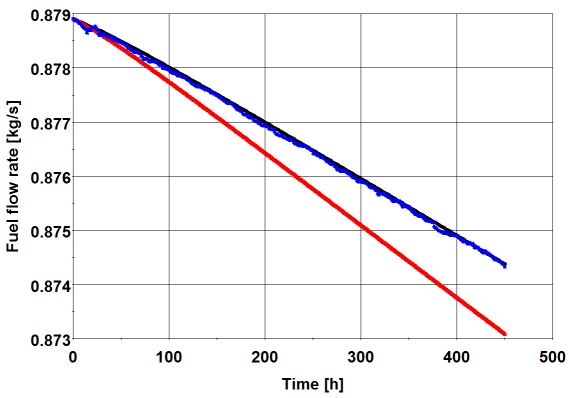
[4]Pressure drop is the output variable which is directly linked to coke deposition.



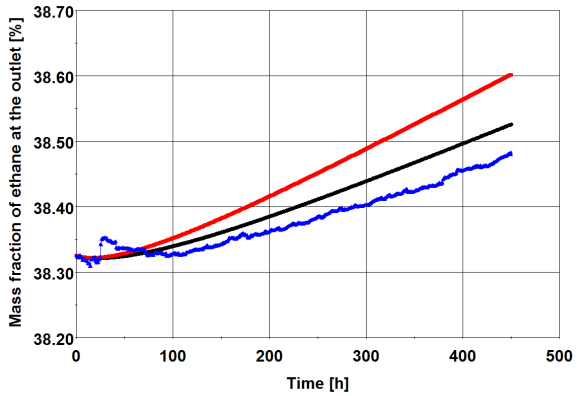
(a) *Pressure drop*



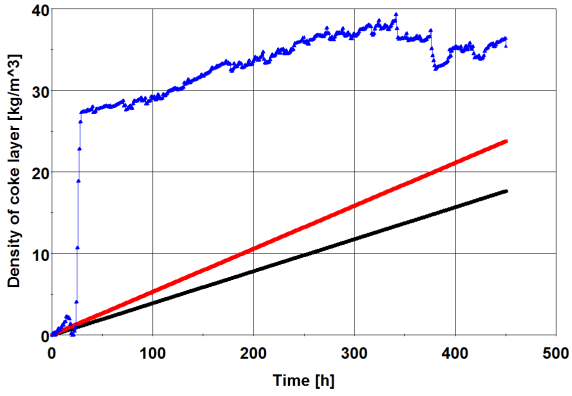
(b) *Tube metal temperature*



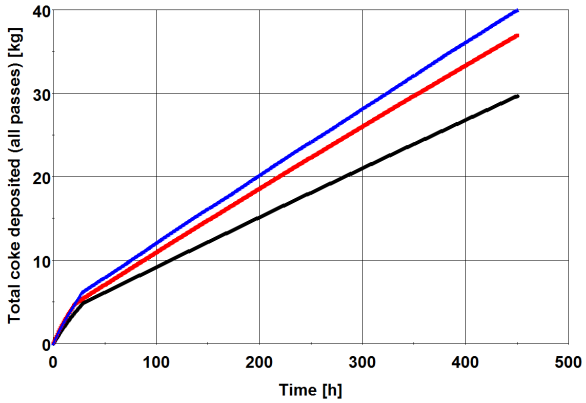
(c) *Fuel flow rate*



(d) *Mass fraction of ethane at the outlet*

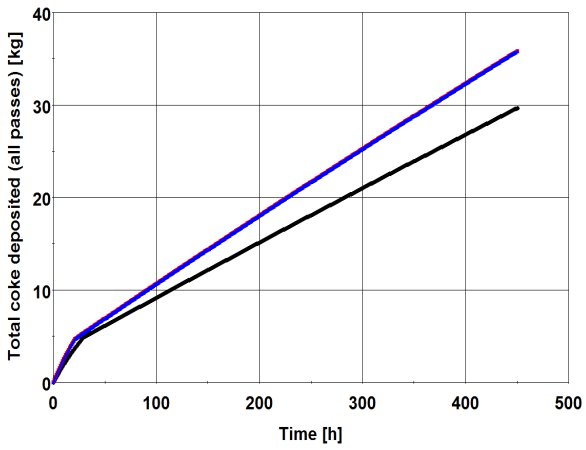


(e) *Density of coke layer on pass #2*

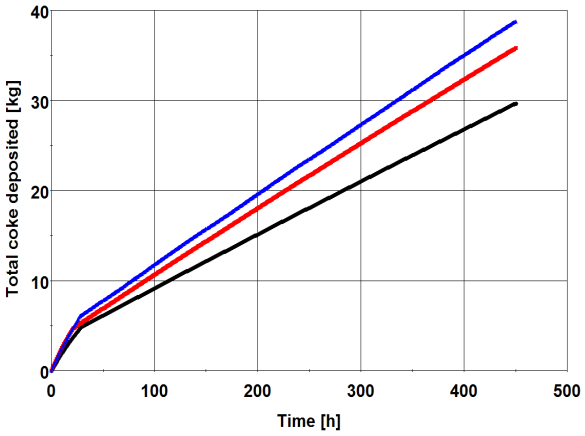


(f) *Total coke deposited (all passes)*

Figure 5.1: *Results obtained for some of the most important output variables. Red lines: wrong model prediction. Black lines: orignal model prediction. Blue lines: estimator prediction. Specifications given to the estimator:* $\mathbf{P}_0 = \mathbf{0}$, $\mathbf{Q}$=*diag(0.01)*, $\mathbf{R}$ = *diag(1e-4; 1e-2;1e-6), WM: -2% on* $\log_{10} k_{c1}(T_{ref})$ *and +2 % on* $\log_{10} k_{c2}(T_{ref})$.

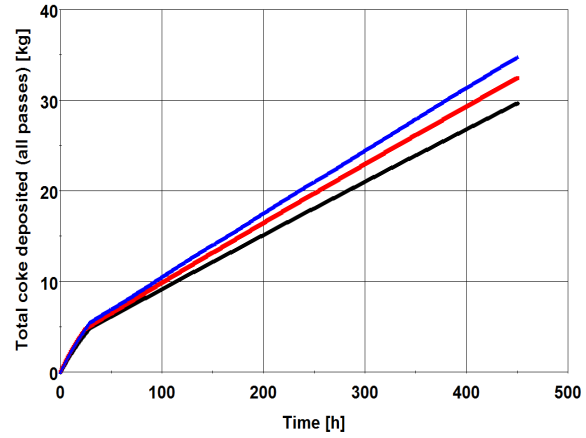


(a) $r^2_{\Delta_p} = 10^{-4}$

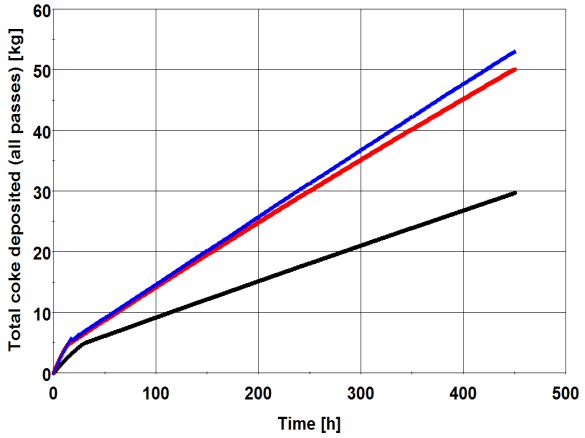


(b) $r^2_{\Delta_p} = 10^{-8}$

Figure 5.2: *Effect of the measurement error variance for pressure drop ($r^2_{\Delta_p}$) on the estimate of the total coke deposited in all passes. Results are not satisfactory even with very low values of $r^2_{\Delta_p}$.*



(a) *-1% deviation on* $\log_{10} k_{c1}(T_{ref})$ *and +1 % on* $\log_{10} k_{c2}(T_{ref})$.



(b) *-5% deviation on* $\log_{10} k_{c1}(T_{ref})$ *and +5 % on* $\log_{10} k_{c2}(T_{ref})$.

Figure 5.3: *Effect of the deviation on model parameters on the total coke deposited in all passes. The estimation of the state of coking is not satisfactory independently of the modification made on the coking parameters.*

second strategy is shown in Fig. 5.2 and Fig. 5.3. Results for small deviations on the coking parameters or with lower measurement error variances for Δp are even worse than with high deviations.

The estimator has also been tested with different sets of model-based measurements to understand what measurements should be used to obtain the best possible estimate. The different sets of measurements that have been used and a qualitative description of the results obtained is reported in Tab. 5.2. In all cases, the estimator prediction is satisfactory only for those state variables that are supplied as measurements. In particular, the estimate of the state of coking is very poor independently of the measurements supplied to the estimator. All the cases of Tab. 5.2 have been obtained using the same structure and value for $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{P}_0$ and the same deviation on coking parameters.

The two most convincing arguments that can be put forward to improve the poor

Table 5.2: *Quality of the estimates obtained using different sets of measurements. The specifications for the estimator are the same in all cases:* $\mathbf{P}_0 = \mathbf{0}$*;* $\mathbf{Q} = \mathrm{diag}(0.01)$*;* $\mathbf{R}$ *diagonal with* $\mathrm{r}^2_{\Delta\mathrm{p}} = 10^{-4}$, $\mathrm{r}^2_{TMT} = 10^{-3}$, $\mathrm{r}^2_{\mathrm{w_{fuel}}} = 10^{-6}$, $\mathrm{r}^2_{\mathrm{x_{C_2H_6}}} = 10^{-6}$*; -2% deviation on* $\log_{10}\mathrm{k_{c1}}(\mathrm{T_{ref}})$ *and +2 % on* $\log_{10}\mathrm{k_{c2}}(\mathrm{T_{ref}})$.

y	Δp **estimate**	**TMT estimate**	$\mathrm{w_{fuel}}$ **estimate**	$\mathrm{x_{C_2H_6}}$ **estimate**	$\Delta\mathrm{R_{\%}}$ **estimate**
$\begin{bmatrix} \Delta\mathrm{p} \\ \mathrm{TMT} \end{bmatrix}$	Excellent	Excellent	Poor	Poor	Very poor
$\begin{bmatrix} \Delta\mathrm{p} \\ \mathrm{TMT} \\ \mathrm{w_{fuel}} \end{bmatrix}$	Excellent	Excellent	Excellent	Poor	Very poor
$\begin{bmatrix} \Delta\mathrm{p} \\ \mathrm{TMT} \\ \mathrm{x_{C_2H_6}} \end{bmatrix}$	Excellent	Excellent	Poor	Excellent	Very poor
$\begin{bmatrix} \Delta\mathrm{p} \\ \mathrm{x_{C_2H_6}} \\ \mathrm{w_{fuel}} \end{bmatrix}$	Excellent	Poor	Excellent	Excellent	Very poor

results obtained from the estimator are:

1. a higher number of measurements should be supplied to the estimator to increase its *a posteriori* knowledge of the system and enhance the correction on the state during the update step;

2. the options for action should be strengthened by letting the estimator adjust some of the model parameters to upgrade its effectiveness during the correction step.

The first option is clearly non desirable considering that the final goal is to obtain a robust and reliable on-line estimate of the state of the furnace using the lowest possible number of available plant measurements. The second argument, instead, looks promising, given that a right choice of the parameters to be adjusted is made. In the next section the results obtained using this strategy are presented.

5.2.2 Real-time adjustment of model parameters

State estimation can be used not only to determine the correct state of a system, but also for parameter estimation (§3.7). From the results discussed in the previous section, it is clear that, independently of the measurements supplied to the estimator and of its specifications ($\mathbf{P}_0$, $\mathbf{Q}$, $\mathbf{R}$), a satisfactory estimation of the state of coking of the furnace can never be reached also for small deviations on the coking parameters.

If a real-time adjustment of the coking parameters that have been modified in the WM is performed, it is expected that the estimator, receiving measurements generated using the OM, will slowly adjust the wrong value of these parameters until their OM value is obtained. By adjusting these parameters, the estimator should reconcile its prediction with the OM prediction, thus giving a correct estimate of the state of the system (including coking).

The theory behind parameter estimation using state estimation was presented in §3.7.

The set of parameters to be adjusted in this case is:

$$\mathbf{p} = \begin{bmatrix} \log_{10} k_{c1}(T_{ref}) \\ \log_{10} k_{c2}(T_{ref}) \end{bmatrix}. \tag{5.8}$$

The practical implementation in gPROMS® consists of four steps:

1. two new differential equations are added to the OM and WM:

$$\frac{\mathrm{d} \log_{10} k_{c1}(T_{ref})}{\mathrm{d}t} = 0 \tag{5.9}$$

$$\frac{\mathrm{d} \log_{10} k_{c2}(T_{ref})}{\mathrm{d}t} = 0; \tag{5.10}$$

```
#Parameter estimation using the EKF   gabrieleb
$log10kc2_nominal_value    =0;
$log10kc4_nominal_value     =0;
```

2. $\log_{10} k_{c1}(T_{ref})$ and $\log_{10} k_{c2}(T_{ref})$ become two new state differential variables for the system, so the dimension of $\mathbf{P}_0$ and $\mathbf{Q}$ changes accordingly (229×229) and their values for the two new differential variables need to be specified in the FO `Furnace_S_E.covariance`;

3. In the OM process entity the original nominal values of the two coking parameters are supplied as initial conditions for eq.(5.9) and (5.10):

```
INITIAL
#Initial nominal values of coking pre-exponential factors gabrieleb
Flowsheet.Furnace_USX_1.Coil.log10kc2_nominal_value =  nom_value;
Flowsheet.Furnace_USX_1.Coil.log10kc4_nominal_value =nom_value;
```

4. in the WM process entity `Furnace_wrong` the initial conditions for the two coking parameters are modified (up to ± 10%). For example:

```
INITIAL

#Modified initial conditions - coking parameters #gabrieleb

Flowsheet.Furnace_USX_1.Coil.log10kc2_nominal_value = ...
    nom_value*0.98;
Flowsheet.Furnace_USX_1.Coil.log10kc4_nominal_value ...
    =-nom_value*1.02;
```

5. in the estimator process entity `Furnace_S_E` the same wrong initial conditions are used.

In Fig. 5.4 and Fig. 5.5 the results obtained with real-time adjustment of the coking parameters are reported. The specifications used to obtain these results are:

- initial error covariance:

$$\mathbf{P}_0 = \begin{array}{c} \begin{matrix} \log_{10} k_{c1} & \log_{10} k_{c2} & & & \\ \downarrow & \downarrow & & & \end{matrix} \\ \begin{pmatrix} 0.05 & 0 & 0 & \cdots & 0 \\ 0 & 0.05 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \end{array}; \tag{5.11}$$

- model and measurement uncertainty covariances:

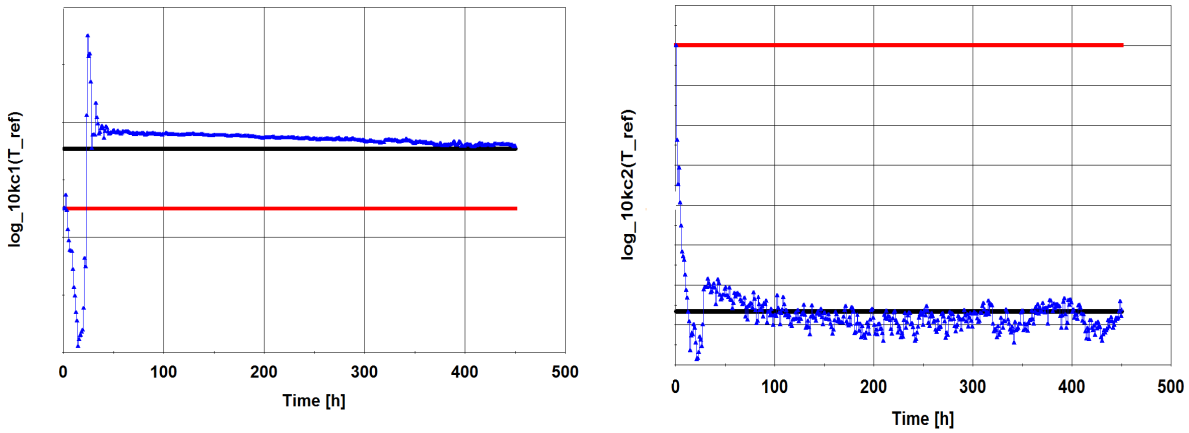
$$\mathbf{Q} = \begin{array}{c} \begin{matrix} \log_{10} k_{c1} & \log_{10} k_{c2} & & & \\ \downarrow & \downarrow & & & \end{matrix} \\ \begin{pmatrix} 0.05 & 0 & 0 & \cdots & 0 \\ 0 & 0.05 & 0 & \cdots & 0 \\ 0 & 0 & 0.01 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0.01 \end{pmatrix} \end{array}; \tag{5.12}$$

$$\mathbf{R} = \begin{array}{c} \begin{matrix} \Delta p & \text{TMT} & w_{fuel} \\ \downarrow & \downarrow & \downarrow \end{matrix} \\ \begin{pmatrix} 10^{-4} & 0 & 0 \\ 0 & 10^{-3} & 0 \\ 0 & 0 & 10^{-6} \end{pmatrix} \end{array}; \tag{5.13}$$

- +10% deviation on the IC of $\log_{10} k_{c1}(T_{ref})$ with respect to its nominal value and -10% deviation on $\log_{10} k_{c2}(T_{ref})$.

In Fig. 5.4 the real-time estimation of the two coking parameters is shown. Black lines represent the nominal values of the two parameters in the OM; red lines represent their modified values given as initial conditions to the estimator. As expected, the estimator prediction (blue lines) starts from the wrong initial values and rapidly converges to the correct original values. The speed at which the correct value of the two parameters is reached strongly depends on the choice of the initial error uncertainty for $\log_{10} k_{c1}(T_{ref})$ and $\log_{10} k_{c2}(T_{ref})$: convergence is very fast if a high initial uncertainty on the two parameters is specified to the estimator. The fluctuations on the estimates of the two parameters (as in Fig. 5.4b for $\log_{10} k_{c2}(T_{ref})$) can be reduced by decreasing their model uncertainty variance and/or increasing their initial error variance. However, to avoid a steep correction during the first update step of the estimator, it is usually not desirable increasing too much the initial error variance of the two parameters: a good compromise must be found.

In Fig. 5.5 the estimator prediction for all the state variables already presented in Fig. 5.1 is shown. It should be emphasized that these results have been obtained using the same specifications for the initial and model error uncertainties for all the 227 differential state variables (excluding the two coking parameters) and the same structure for $\mathbf{R}$ as the results shown in Fig. 5.1. The only difference is that a deeper deviation on the



(a) $\log_{10} k_{c1}(T_{ref})$ *(+10% deviation on ICs).* **(b)** $\log_{10} k_{c2}(T_{ref})$ *(-10% deviation on ICs).*

Figure 5.4: *Real-time estimation of coking parameters. Black: OM nominal value. Red: WM modified value. Blue: estimator prediction.*

coking parameters has been performed (±10% in Fig. 5.5 against ±2% in Fig. 5.1) and a real-time estimation of these parameters has been carried out.

The estimation of the three state variables (Δp, TMT, w_{fuel}) for which measurements are supplied is as good as in Fig. 5.1. However, a big difference in the estimator performance for the other state variables can be noticed in Fig. 5.5 with respect to Fig. 5.1. A marked improvement can be seen on both the prediction of the mass fraction of ethane at the outlet and the state of coking of the system[5]. This is clearly related to the added flexibility given by the real-time adjustment of the coking parameters performed by the estimator.

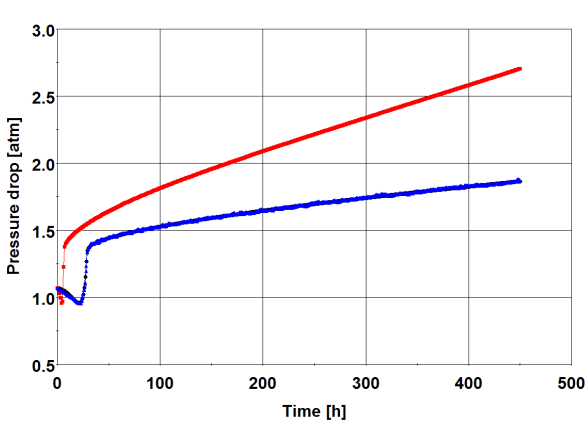
Different runs have been done to test the estimator performance by using different deviations on the coking parameters, different sets of model-based measurements and different specifications for $\mathbf{P}_0$, $\mathbf{Q}$ and $\mathbf{R}$. Results have proved satisfactory for all the runs performed and for all the output variables of the model. As an example, in Fig. 5.6 the results obtained in one of these runs for the outlet composition of all the components involved in the cracking mechanism is presented.

Given the good results obtained with real-time adjustment of the coking parameters using time-invariant control inputs, the next step that has been taken is testing the estimator prediction using time-varying control inputs. The overall procedure (measurement generation, model modification and state estimation) has been repeated using the available experimental data for the control inputs of Tab. 4.2[6] .

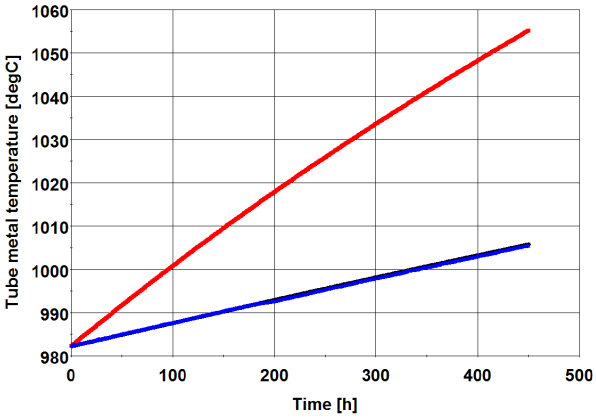
It has already been pointed out (§4.2.3) that one of the main assumption that is made when deriving the governing equations for the EKF (but also for the other state-estimation algorithms) is that there is no uncertainty related to the control vector $\mathbf{u}$. All the filter equations are obtained based on this assumption (§3.2.1).
In the case of this project, this hypothesis is clearly not satisfied since plant measure-

[5]Only the density of the coke layer on pass #2 and the total mass of coke deposited in all passes are reported in Fig. 5.5. However, the estimation is satisfactory for all the other state variables that describe the coke deposition on the coil.

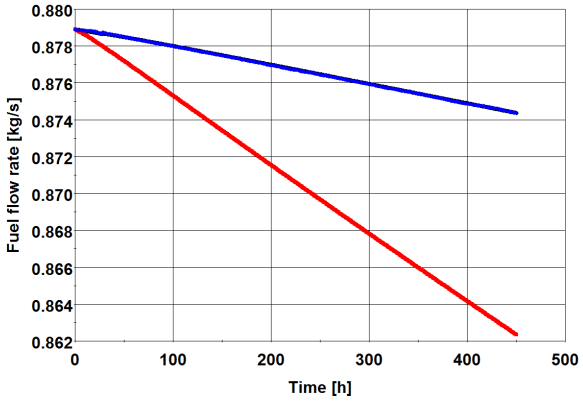
[6]It should be emphasized that all the measurements given to the estimator at this stage of the project are still obtained using the OM prediction. The only difference is that measurements are now generated using in the original model the available plant data for the control variables.



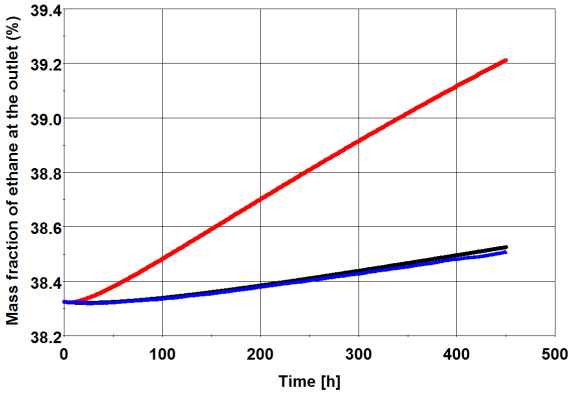
(a) *Pressure drop*



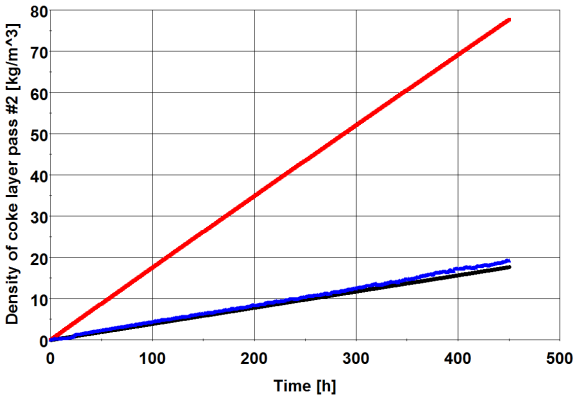
(b) *Tube metal temperature*



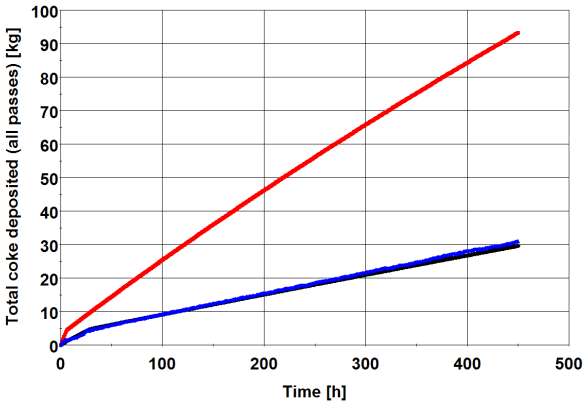
(c) *Fuel flow rate*



(d) *Mass fraction of ethane at the outlet*

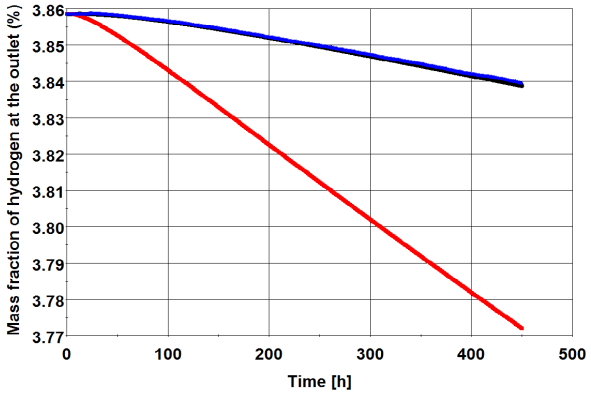


(e) *Density of coke layer on pass #2*

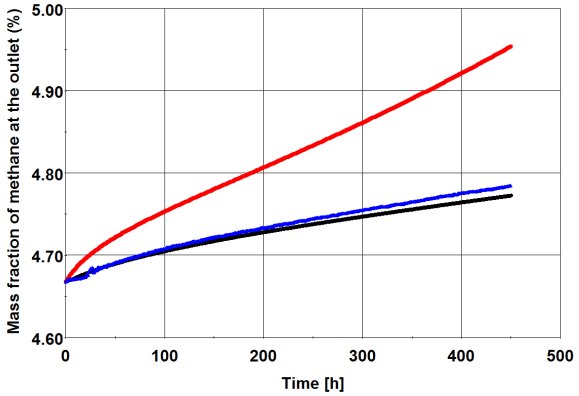


(f) *Total coke deposited (all passes)*

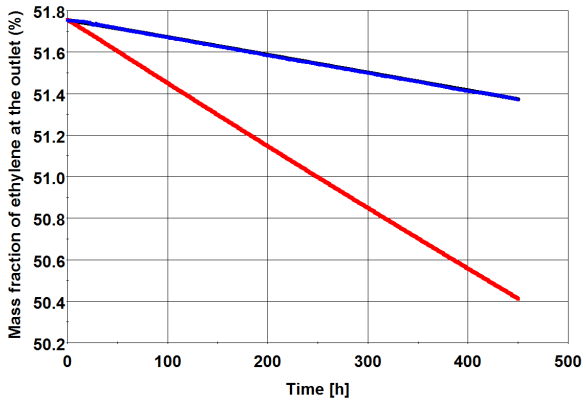
Figure 5.5: *Results obtained with real-time parameter adjustment for some of the most important output variables of the Furnace model. Red lines: wrong model prediction. Black lines: orignal model prediction. Blue lines: estimator prediction. Specifications given to the estimator:* $\mathbf{P}_0 = diag(0.05; 0.05; 0; \ldots; 0)$, $\mathbf{Q} = diag(0.05; 0.05; 0.01; \ldots; 0.01)$, $\mathbf{R} = diag(1e-4; 1e-2; 1e-6)$, *WM: +10% on* $\log_{10} k_{c1}(T_{ref})$ *and -10 % on* $\log_{10} k_{c2}(T_{ref})$.



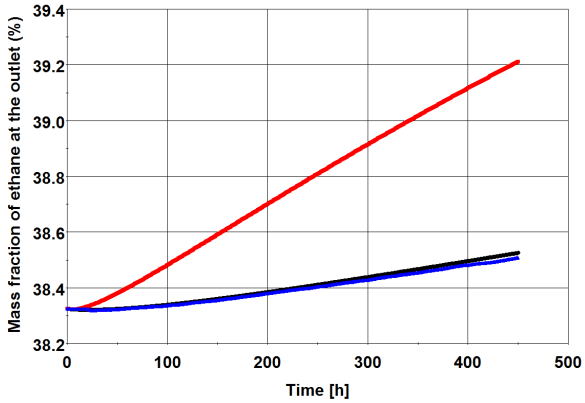
(a) *Mass fraction of hydrogen at the outlet*



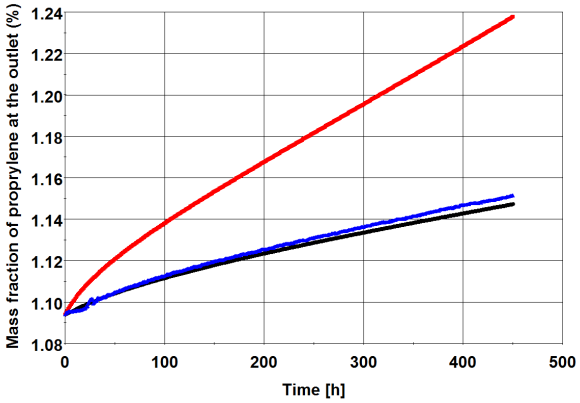
(b) *Mass fraction of methane at the outlet*



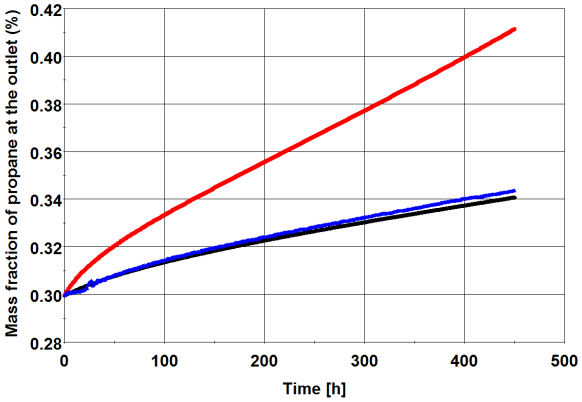
(c) *Mass fraction of ethylene at the outlet*



(d) *Mass fraction of ethane at the outlet*



(e) *Mass fraction of propylene at the outlet*



(f) *Mass fraction of propane at the outlet*

Figure 5.6: *Estimation of the composition of the outlet stream exiting the coil. Red lines: wrong model prediction. Black lines: orignal model prediction. Blue lines: estimator prediction. Specifications given to the estimator:* $\mathbf{P}_0 = diag(0.05; 0.05; 0; \ldots; 0)$, $\mathbf{Q} = diag(0.05; 0.05; 0.01; \ldots; 0.01)$, $\mathbf{R} = diag(1e-4; 1e-2; 1e-6)$, *WM: +10% on* $\log_{10} k_{c1}(T_{ref})$ *and -10 % on* $\log_{10} k_{c2}(T_{ref})$.

ments (thus subject to noise/uncertainty) are used as time-varying inputs to be passed to the estimator. Deriving the filter equations considering the uncertainty on the control variables is tedious and often reveals to be useless (Simon, 2006). In fact, the uncertainty on the control inputs can be implicitly taken into account by considering it as part of the uncertainty on the model (i.e. $\mathbf{Q}$). This is not rigorous, but it can reduce the effort required to implement the estimator. All the simulations have been performed by using higher values for the model error variances for all the state variables with respect to the case with time-invariant inputs. Good results have been obtained using this strategy, as can be noticed in Fig. 5.7. The time-varying behavior of the control inputs is shown in Fig. 5.8. These results refer to the first set of experimental data available (feed: pure ethane) but similar satisfactory results have been obtained also for the other sets of data. Given the good response of the estimator, it has been decided to move to the next step of the project (step 3 of Fig. 4.2) by testing the estimator using real off-line plant data. The results obtained at this stage are the subject of the next chapter.

5.3 Remarks

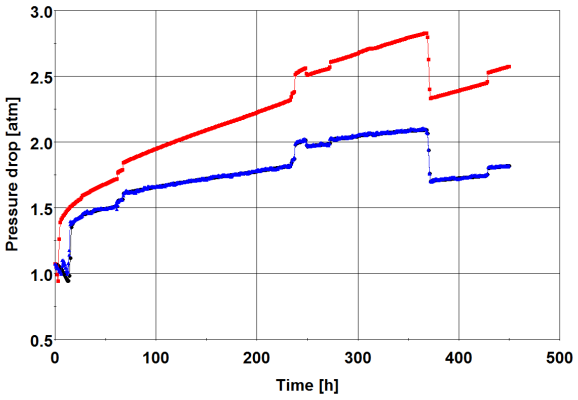
At the beginning of this chapter, the reasons why testing the estimator with model-based measurements before using real plant data can reveal its potential applicability on a specific problem and the key pieces of information that can be obtained at this stage were presented.
Considering the results obtained with model-based measurements, it can be said that the following points should be at least taken into account when using the estimator with real plant data:

- to improve the estimation of the state of coking of the system, the coking parameters should be re-calibrated in real-time increasing the flexibility of the estimator. In fact, given the high dimensionality of the state vector $\mathbf{x}$ if compared with the measurement vector $\mathbf{y}$, it is not possible to obtain a good state estimate by just updating the state vector using the available measurements. It is then of fundamental importance ensure that these measurements will be available in the on-line implementation of the estimator;

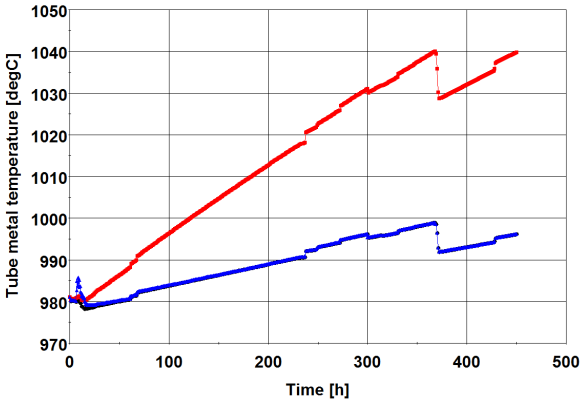
- the plant measurements to be supplied to the estimator should be chosen according to their actual availability and their reliability. However, by testing the estimator using model-based measurements, it has turned out that the availability of pressure drop measurements is essential to obtain a good estimate of the state of coking of the furnace;

- the effect of $\mathbf{P}_0$ on the results is important only when some ICs for the differential variables involved in the model are modified. If not, since the simulation starts from clean tube conditions, there is no uncertainty on the initial state of the system;

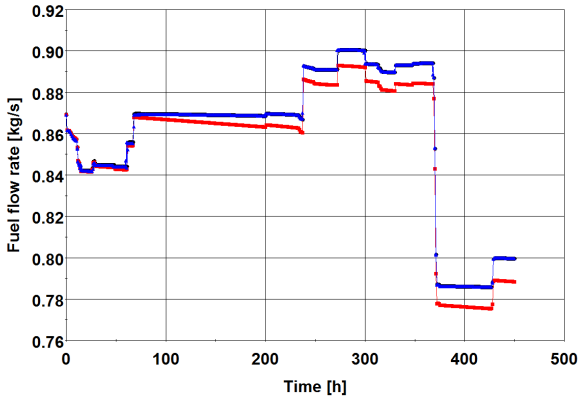
- the choice of $\mathbf{Q}$ and $\mathbf{R}$ deeply affects the quality of the results that can be obtained from the estimator. When using real-plant data, a fixed structure for $\mathbf{Q}$ and $\mathbf{R}$ should be identified in order to ease the on-line implementation of the estimator (i.e. its specification should be independent of the on-line measurements coming from the furnace).



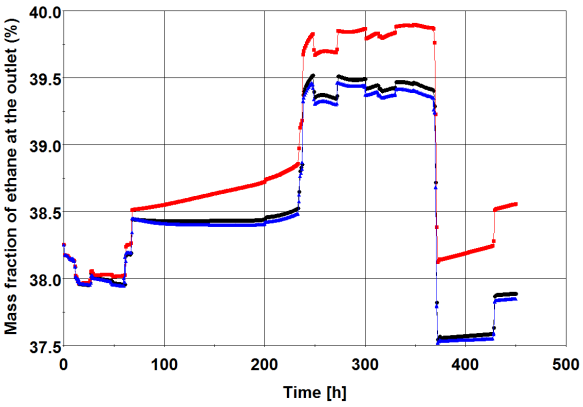
(a) *Pressure drop*



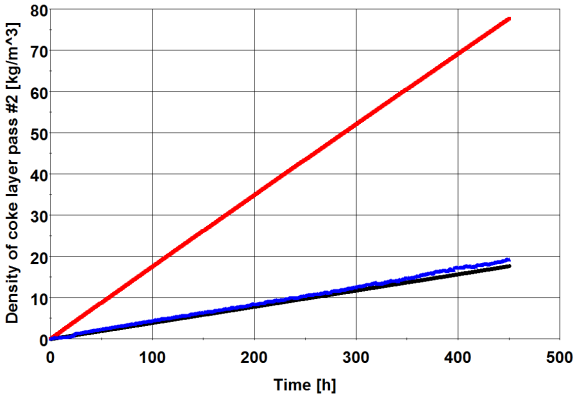
(b) *Tube metal temperature*



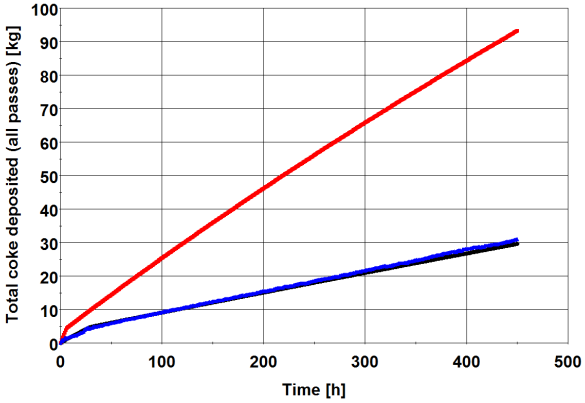
(c) *Fuel flow rate*



(d) *Mass fraction of ethane at the outlet*

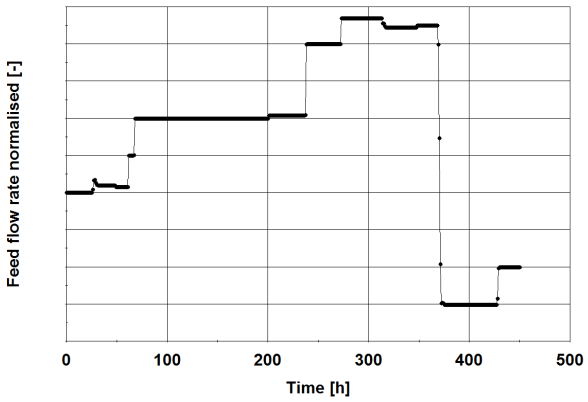


(e) *Density of coke layer pass #2*

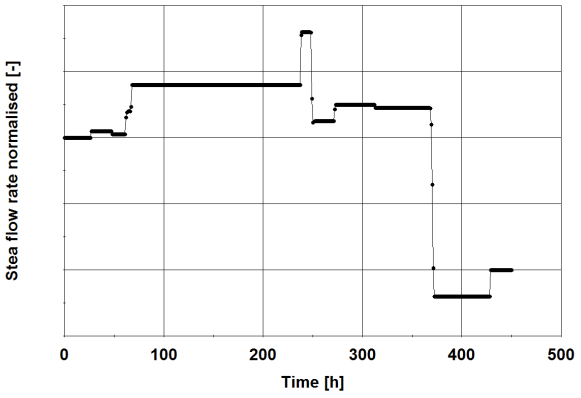


(f) *Total coke deposited (all passes)*

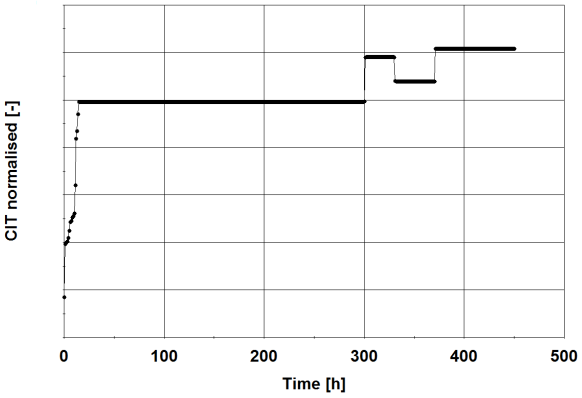
Figure 5.7: *Results obtained using time-varying control inputs. Red lines: wrong model prediction. Black lines: orignal model prediction. Blue lines: estimator prediction. Specifications given to the estimator:* $\mathbf{P}_0 = diag(0.05; 0.05; 0; \ldots; 0)$, $\mathbf{Q} = diag(0.05; 0.05; 0.01; \ldots; 0.01)$, $\mathbf{R} = diag(1e-4; 1e-2; 1e-6)$, *WM: +10% on* $\log_{10} k_{c1}(T_{ref})$ *and -10 % on* $\log_{10} k_{c2}(T_{ref})$.



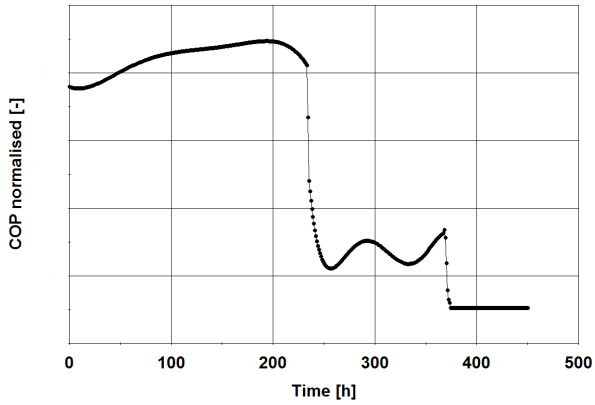
(a) *Feed flowrate*



(b) *Steam flow rate*



(c) *Coil inlet temperature*



(d) *Coil outlet pressure*

Figure 5.8: *Normalised time-varying supplied to the estimator. Set of data #1.*

Additional problems, such as noise and structural inconsistencies, should be considered when switching to real experimental data. However, knowing *a priori* the previous remarks has proved to be extremely beneficial for the final success of the project.

Chapter 6

Results with real plant data

This chapter deals with the third step of the organizational chart of Fig. 4.2.
First, a brief introduction to the main aspects that need to be considered when using real plant data is given. Then, the weaknesses of the original model and how they have been adjusted using the estimator are discussed. The general structure of the model is then presented and the results obtained for all the four sets of data available are shown. The results that involve real plant data have been normalized and sanitized (i.e numbers are removed from axes) for confidentiality reasons.
Finally, a brief discussion on the results obtained is performed.

6.1 Introduction

Testing the estimator using model-generated measurements has proved to be extremely beneficial to obtain a preliminary overview of the potential and the main limitations of the EKF on the Furnace model. Some important hints for the on-line implementation of the estimator with real plant data have been obtained at this stage (§5.3). However, lots of complications can arise when real data are supplied to the estimator.
First, the availability and reliability of data become a major point of concern (§4.2.3). It has already been pointed out that the availability and reliability of measurements in the real plant strongly depends on the state of repair of the installed measuring devices. The possibility that one or more measurements could not be available at some point of the furnace operation should be considered.
Second, structural inconsistencies in the data can deeply influence the estimator performance. It will be shown (§6.2) that, in some cases, available data are not consistent with conservation laws (e.g. they do not respect the conservation of mass). Luckily, the degree of confidence in the available data can be tuned modifying the measurement uncertainty covariance, thus reducing the 'weight' of these inconsistencies on the estimator prediction.
Third, when using plant measurements, being able to fix the weaknesses of the original model becomes a critical task to obtain a robust and reliable estimate of the state of the system. All the assumptions made in the original model contribute to the mismatch between its prediction and the plant observations: the task of the estimator is to offset these assumptions with the additional information brought in by the measurement vector.
Finally, the computational performance of the estimator becomes important when its on-line implementation need to be considered. It has already been said (§4.2.6) that the computational time for every prediction step should never be higher than the frequency at which measurements are supplied to the estimator. This should be taken into account when testing the estimator.

As discussed in §4.2.3, the intrinsic use of the estimator is for on-line applications; however, its performance has been tested off-line using four different sets of plant data of the furnace. In the next few sections, the adjustments made to fix the model weaknesses

and the results obtained from the estimator are discussed.

6.2 Model weaknesses and real-time adjustment

The main weaknesses of the Furnace model have been described in §2.5. When using model-based measurements, these weaknesses have clearly no effect on the estimator performance. However, when using real data, they are the main responsible for the mismatch between the model prediction and the available plant measurements. One of the tasks of the estimator is to somewhat fix these weaknesses to adapt the model prediction to the plant observations.

Following the discussion in §2.5, it seems clear that these three weaknesses should be considered before testing the EKF on the OM:

- **coking model**: the expression for the coking rate is over-simplified and the mechanism of coke deposition is not clearly understood. This is the critical aspect of the model since a stand-alone first-principles approach to obtain an accurate description of coke deposition is not viable;

- **firebox model**: this model is extremely simplified and makes some assumptions (e.g. no external heat loss or total combustion of reactants) that are not verified in the real plant;

- **cracking model**: a simplified molecular mechanism is used to describe the cracking reactions and only few components are considered .

One important remark must be made regarding this last weakness. It has already been said (§6.1) that, when using plant data, structural inconsistencies in the data can always be present and can deeply influence the results from the estimator. In the case of this study, looking at the data of outlet composition, it has been noticed that they are not consistent with the law of conservation of mass. This is likely to be due to the presence of minor components in the outlet stream that are not tracked by the composition measuring device and to the high uncertainty related to the measurements of the minor components (H_2, CH_4, C_3H_6, C_3H_8 (for ethane feed) or C_2H_6 (for propane feed)). Luckily, there is not a urgent need to obtain a reliable estimate of the composition of the minor components at the outlet of the coil. What is required is to obtain an accurate estimate of the global conversion (thus of the composition of ethane/propane at the outlet) and of the yield (thus of the composition of ethylene at the outlet) of the cracking reactions. In fact, the separation section of the real plant is modelled using guessed inlet compositions for the minor components that do not depend on what is predicted by the cracking mechanism implemented in the Furnace model. In other terms: the outlet composition of the minor components predicted by the Furnace model is not used as an input for the separation section of the plant. Thus, considering also that the only equations in which the minor components are involved are the species material balances (2.1), no particular attention should be given to these measurements and to their predicted values. According to this, only the weaknesses of the firebox model and of the coking model should be considered. In the next paragraph, the adjustments made to fix these weaknesses are presented.

6.2.1 Firebox model adjustment

The model of the firebox has been exhaustively explained in §2.2.4. The firebox is modelled as a conversion reactor: complete combustion of the fuel (NG:natural gas) is assumed and no external heat loss is considered. Once a fixed value is assigned to the following variables:

- the fuel composition;
- the air excess ϵ;
- the air inlet temperature $\tilde{T}^{in}_{air}$ and the fuel inlet temperature $\tilde{T}^{in}_{fuel}$;
- the outlet temperature of the flue gas $\tilde{T}^{out}_{flue\,gas}$

the model computes the fuel consumption using eq. (2.17) and the closure equation (2.18). As already shown in Fig. 4.3, the prediction of fuel consumption is not in good agreement with the plant data available. This is true for all the four sets of data considered. In particular, it can be seen from Fig. 4.3 that a systematic offset seems to exist between the model prediction and the plant observations. This is mainly due to the assumption of no external heat loss and, to a lesser extent, to the fact that guessed inlet and outlet temperatures are fixed for the reactants and the flue gas and to the assumption of total combustion of reactants. To better understand this point, the energy balance that is implemented in the firebox model is here recalled[1] :

$$F^{in}_{fuel}h^{in}_{fuel}\left(\tilde{T}^{in}_{fuel}\right) + F^{in}_{air}h^{in}_{air}\left(\tilde{T}^{in}_{air}\right) - F^{out}_{flue\,gas}h^{out}_{flue\,gas}\left(\tilde{T}^{out}_{flue\,gas}\right) = -\dot{Q}_I. \tag{2.17}$$

However, the actual energy balance for the system is given by[2]:

$$F^{in}_{fuel}h^{in}_{fuel}(T^{in}_{fuel}) + F^{in}_{air}(T^{in}_{air}) - F^{out}_{flue\,gas}h^{out}_{flue\,gas}(T^{out}_{flue\,gas}) = -\dot{Q}_I - \Delta\dot{h}_L \tag{6.1}$$

where $\Delta\dot{h}_L$ is the external heat loss and $T^{in/out}_i$ is the actual inlet/outlet temperature of i (not to be confused with the guessed temperature $\tilde{T}^{in/out}_i$).

The energy balance implemented in the model (2.17) can be adjusted in real-time by adding an adjustment parameter β as follows:

$$F^{in}_{fuel}h^{in}_{fuel}(\tilde{T}^{in}_{fuel}) + F^{in}_{air}(\tilde{T}^{in}_{air}) - F^{out}_{flue\,gas}h^{out}_{flue\,gas}(\tilde{T}^{out}_{flue\,gas}) = -(1+\beta)\dot{Q}_I. \tag{6.2}$$

This parameter is adjusted in real-time by the estimator to match the available experimental data of fuel consumption. To do this, the following differential equation need to be added to the model:

$$\frac{\mathrm{d}\beta}{\mathrm{d}t} = 0 \tag{6.3}$$

with IC:

$$@t = 0: \qquad \beta = 1 \tag{6.4}$$

[1]The specific enthalpies are considered as functions of temperature only.

[2]The convention adopted is that heat released to the external environment is negative: this is why the heat loss term is negative.

One new differential variable is then added to the system and the dimensions of $\mathbf{P}_0$ and $\mathbf{Q}$ change accordingly. To understand the meaning of this adjustment parameters, the 'adjusted' energy balance (6.2) is subtracted to the actual energy balance (6.1), obtaining:

$$\beta\dot{Q}_I - \Delta\dot{h}_L = F^{in}_{fuel}\left[h^{in}_{fuel}(T^{in}_{fuel}) - h^{in}_{fuel}(\tilde{T}^{in}_{fuel})\right] + F^{in}_{air}\left[h^{in}_{air}(T^{in}_{air}) - h^{in}_{air}(\tilde{T}^{in}_{air})\right] - F^{out}_{flue\,gas}\left[h^{out}_{flue\,gas}(T^{out}_{flue\,gas}) - h^{out}_{flue\,gas}(\tilde{T}^{out}_{flue\,gas})\right]. \quad (6.5)$$

Let ϕ be the right term of eq. (6.5):

$$\phi = F^{in}_{fuel}\left[h^{in}_{fuel}(T^{in}_{fuel}) - h^{in}_{fuel}(\tilde{T}^{in}_{fuel})\right] + F^{in}_{air}\left[h^{in}_{air}(T^{in}_{air}) - h^{in}_{air}(\tilde{T}^{in}_{air})\right] - F^{out}_{flue\,gas}\left[h^{out}_{flue\,gas}(T^{out}_{flue\,gas}) - h^{out}_{flue\,gas}(\tilde{T}^{out}_{flue\,gas})\right]. \quad (6.6)$$

From eq. (6.5) the following expression can be obtained for β:

$$\beta = \frac{\phi + \Delta\dot{h}_L}{\dot{Q}_I}. \quad (6.7)$$

In other terms, β represents the fraction of non-modelled energy (NME) with respect to the total energy input to the coil. This NME is the sum of two contributions : the heat loss through the walls of the furnace and the enthalpy-mismatch due to the guessed inlet and outlet temperatures for the reactants and the flue gas. The biggest uncertainty on these guessed temperatures is on the flue gas outlet temperature $\tilde{T}^{out}_{flue\,gas}$: the fuel and air inlet temperatures are usually approximately known so that it can be written:

$$\phi \approx -F^{out}_{flue\,gas}\left[h^{out}_{flue\,gas}(T^{out}_{flue\,gas}) - h^{out}_{flue\,gas}(\tilde{T}^{out}_{flue\,gas})\right]. \quad (6.8)$$

In other terms, β is a parameter that accounts for both the uncertainty on the guessed value for the flue gas temperature and the heat loss through the furnace walls. Its real-time adjustment is expected to seriously improve the prediction of the fuel consumption if on-line data of this variable are supplied to the estimator.

6.2.2 Coking model adjustment

The weaknesses of the coking model have already been pointed out in many parts of this thesis. It is worth remembering that coking, since it occurs over a much longer time- scale than cracking reactions, is the only phenomenon that is modelled using a dynamic model: its correct prediction is crucial to obtain a correct estimate of the state of the system. The coking rate expression (2.25) is over-simplified and, despite trying to give a simplistic representation of the two different mechanisms that contribute to coke build-up, is not able to give a reasonable prediction of its deposition on the coil. In this regard, the output variable that is directly linked to coking is the pressure drop, and its prediction is far from being in good agreement with the available plant data (Fig. 4.3).

The simplest way to intervene on the coking model is to let the estimator adjust in real-time the coking parameters that are most influential on its prediction: $\log_{10} k_{c1}(T_{ref})$ and $\log_{10} k_{c2}(T_{ref})$. These parameters can be modified in real-time in order to match the

available plant data supplied to the estimator. As already said in §5.2.2, to do this the following differential equations need to be added to the model:

$$\frac{\mathrm{d}\log_{10} k_{c1}(T_{ref})}{\mathrm{d}t} = 0 \tag{6.9}$$

$$\frac{\mathrm{d}\log_{10} k_{c2}(T_{ref})}{\mathrm{d}t} = 0 \tag{6.10}$$

with ICs:

$$@t = 0: \quad \log_{10} k_{c1}(T_{ref}) = [\log_{10} k_{c1}(T_{ref})]_{nominal} \tag{6.11}$$

$$@t = 0: \quad \log_{10} k_{c2}(T_{ref}) = [\log_{10} k_{c2}(T_{ref})]_{nominal}\,. \tag{6.12}$$

It is worth remembering the the initial nominal values of the two coking pre-exponential factors are the fixed-values used for these parameters in the OM and have been estimated in the OM using a gPROMS® Parameter Estimation entity and four sets of daily plant data. Considering the three new differential equations added to the model (6.3), (6.11) and (6.12) the number of differential variables involved in the model becomes 230.

In theory, the real-time adjustment of the coking parameters could be sufficient to obtain a good estimate of the state of coking of the system. However, it should be remembered that:

1. since these parameters are extremely influential on the model prediction, a slight modification on their values performed by the estimator during the update step can result in a substantial modification of the output distribution predicted by the model;

2. considering the huge mismatch between the model prediction and the plant observations (Fig. 4.3), it is reasonable to expect a substantial adjustment of these parameters by the estimator.

The combination of (1) and (2) can represent a major threat for the numerical stability of the estimator. In fact, the estimation can rapidly diverge if a wrong correction is performed on these parameters at some point during the simulation. To avoid this problem, two actions can be taken:

i) low model error variances should be chosen for these two differential variables in the model uncertainty covariance $\mathbf{Q}$;

ii) the flexibility of the estimator to adjust the coking model should be enhanced by increasing the number of parameters that can be adjusted in real-time to match the plant observations.

The first point has been taken into account when performing the estimation. As regards the second remark, it should be remembered (§4.2.2) that, in the OM:

1. if COT is given as an input, the mass fraction of ethane/propane at the outlet (i.e. conversion) is calculated as an output and *viceversa*;

2. since COT measurements are not available in the real plant, the skin temperature at the outlet of the coil (CST) plus an *a priori* fixed offset are supplied to the model instead of COT:

$$\mathrm{COT} = \mathrm{CST} + \mathrm{COT}_{offset}. \tag{6.13}$$

The value of the COT_{offset} is a fixed guessed value that changes according to the feed considered and that has been estimated to improve the agreement between the model prediction and the plant data.

When using state estimation, however, a new approach can be used to improve the numerical stability of the estimator. According to this approach, CST is still an input for the model and is thus collected in the control vector $\mathbf{u}$. However, the plant observations for the mass fraction of ethane at the outlet are supplied to the estimator (and thus collected in the measurement vector $\mathbf{y}$) and the COT_{offset} is adjusted in real-time like the firebox adjustment parameter β and the two coking parameters:

$$\frac{\mathrm{dCOT}_{offset}}{\mathrm{d}t} = 0; \quad @t = 0: \quad \mathrm{COT}_{offset} = [\mathrm{COT}_{offset}]_{nominal}\,. \tag{6.14}$$

In this way, the number of parameters that the estimator can adjust is increased by one and the 'amount' of correction that need to be performed during the update step can be shared between four instead of three parameters. This new approach is graphically compared with the approach adopted in the OM in Fig. 6.1.

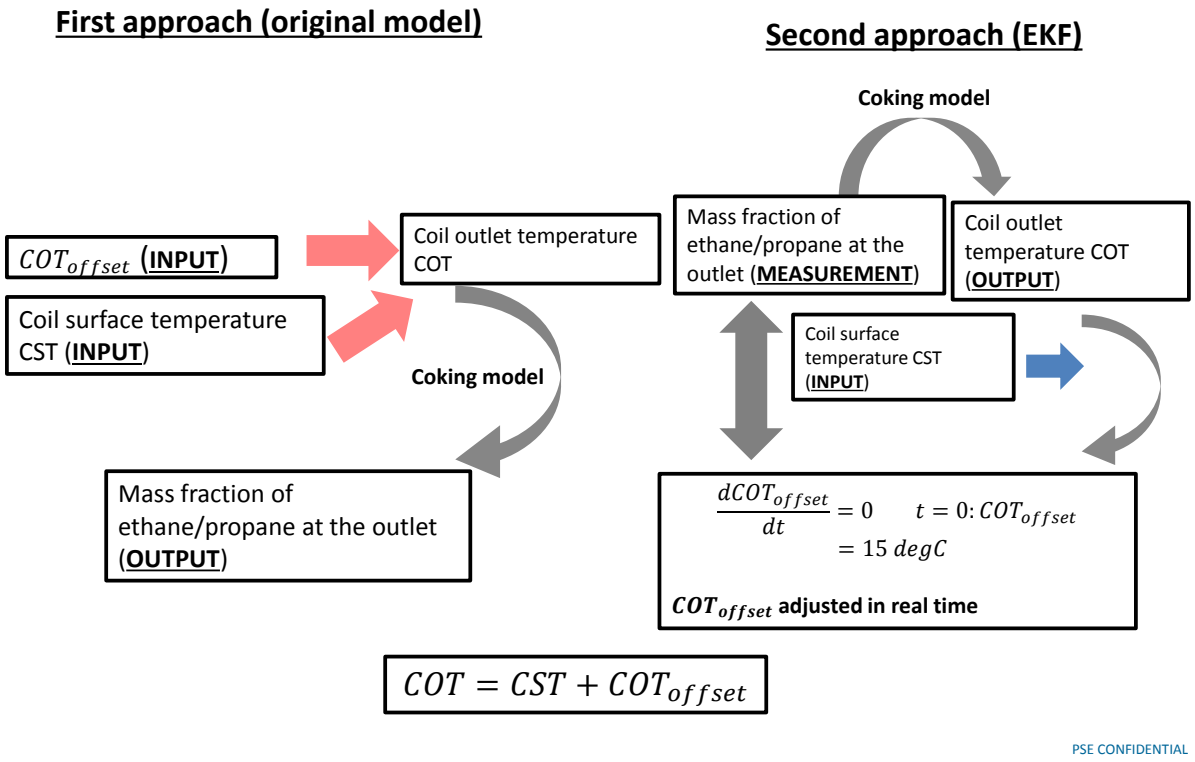


Figure 6.1: *Comparison between the OM approach and the new approach adopted using the EKF with real-time adjustment of the* COT_{offset}.

There is one clear objection that can be put forward against this approach. In fact, it requires the on-line availability of measurements of the mass fraction of ethane/propane at the outlet and, remembering Tab. 4.3, these measurements are usually considered

not reliable[3]. To be more precise, this approach can be used also if measurements of COT are available. In this case, the estimator would adjust the COT_{offset} to match the available experimental data of COT; the mass fraction of ethane/propane at the outlet, however, should be supplied as an input. Both these two measurements seem to be non-available and/or non reliable according to Tab. 4.3. However, it must be said that what is written in Tab. 4.3 should just be considered as an indication. It has however been agreed that the availability and reliability of these measurements (especially COT) could be obtained (e.g. by installing new measuring devices) if a substantial improvement in the prediction of the state of the system is performed by the estimator. Thus, just for testing purposes, the available measurements of $w_{ethane/propane}$ have been used to verify the estimator response.

6.3 General structure

Considering what has been explained in the previous sections, it is important to give a general overview of the system before discussing the results obtained from the estimator. In Fig. 6.2 the key pieces of information to be remembered are collected. The main inputs and outputs of the model are shown. Following the previous discussions and remembering Tab. 4.3, the measurements supplied to the estimator in all the simulations are:

- pressure drop;
- fuel flow rate;
- mass fraction of ethane/propane at the outlet **or** COT.

TMT measurements are not given to the estimator, despite being occasionally available and being considered reliable. This is just because the current implementation of the estimator in gPROMS® is not able to handle occasional measurements like the ones for TMT. This does not mean that the EKF solver cannot be modified to handle partial measurements[4], but, as will be seen in the next sections, the measurements of TMT has been used to validate the estimator prediction. Thus, there is no need to supply these measurements.

According to the previous discussions, 4 parameters are adjusted in real-time by the estimator. These are:

- the firebox adjustment parameter β;
- the two coking pre-exponential factors $\log_{10} k_{c1}(T_{ref})$ and $\log_{10} k_{c2}(T_{ref})$;
- the COT offset.

Thus, the final state vector dimensionality is 231 (227+4). It follows that $\mathbf{P}_0, \mathbf{Q} \in \mathbb{R}^{231\times231}$. All the output variables have been monitored to analyze the estimator performance. Some of them, like TMT and the mass fraction of ethylene at the outlet (i.e. the yield of cracking reactions) have been used to validate the estimator prediction.

[3]On-line composition measurements are usually taken with a much lower frequency than, for example, temperature or flow rate measurements. However, it should be remembered that, in a potential on-line application of the estimator, the frequency at which the state of the system would be updated is every hour (§4.2.6). Thus, this problem does not affect the potential usage of composition measurements to perform the estimation.

[4]Actually, its implementation is work in progress.

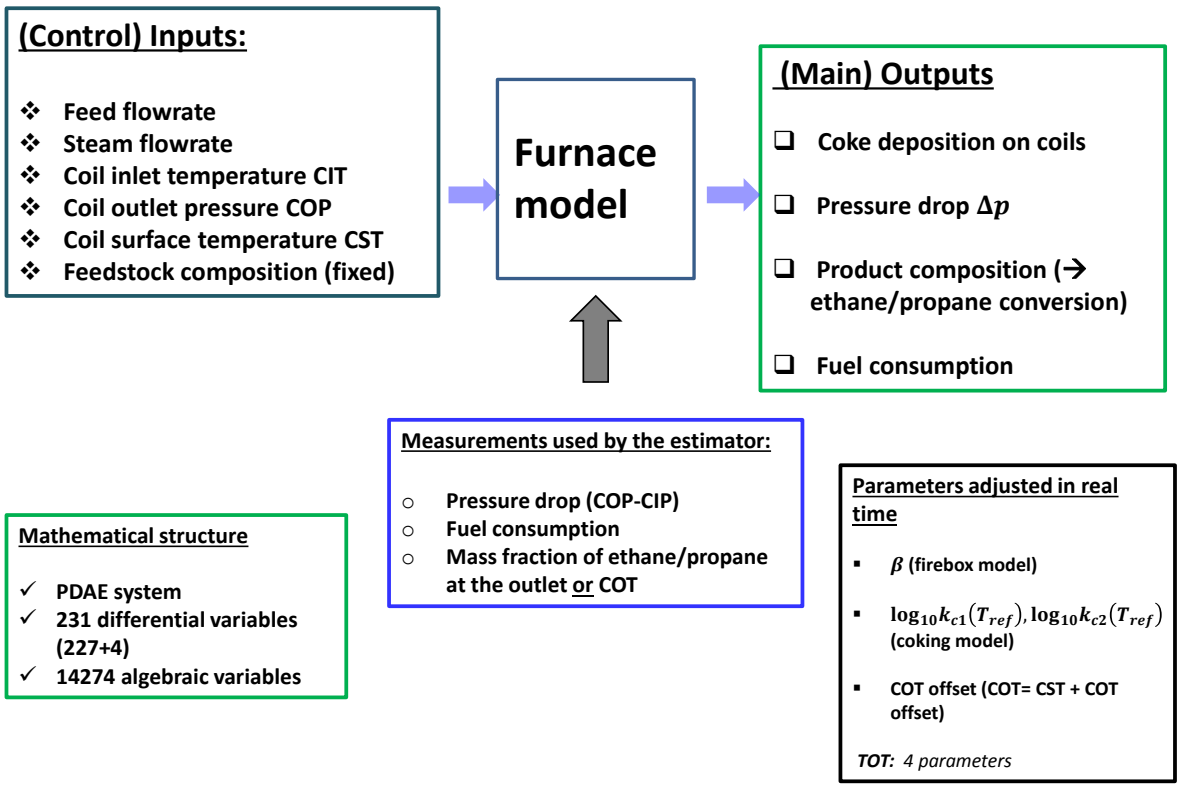


Figure 6.2: *General structure of the simulations performed using real plant measurements.*

6.4 Targets

Before discussing the results obtained from the estimator, it is important to briefly recall the main targets to be achieved:

1. obtain a robust and reliable estimate of the state of the furnace using the lowest possible number of on-line measurements. To validate the estimator prediction, some plant measurements that are **not** supplied to the estimator should be used. The coking prediction cannot be directly validated but pressure drop can be used as a reliable indicator;

2. obtain a computational performance suitable for an on-line application of the estimator;

3. obtain a good estimator performance while using the same structure of $\mathbf{P}_0$, $\mathbf{Q}$ and $\mathbf{R}$ independently of the type and number of on-line measurements supplied. In other terms, the estimator should work as a *black box* independently of the type of data supplied. This requirement is mainly a commercial one rather than a technical one.

Keeping in mind these targets, the results obtained using the different sets of data available are presented in the next sections.

6.5 Results for a pure ethane feed

It has already been said (§4.2.3) that the estimator has been tested using four different sets of historical data of the furnace. The first two sets of data refer to a pure ethane feed and two different runs of the furnace.

The feed composition is usually not constant during the furnace operation. However, in the model, a fixed averaged time-invariant composition is used to simplify the modelling burden. This composition for a pure ethane feed is reported in Tab. 6.1.

Table 6.1: *Feed composition - pure ethane.*

Component	**Mass fraction (%)**
Methane	0.46
Ethylene	1.05
Ethane	98.45
Propylene	0
Propane	0.01
1- Butene	0
C_5^+	0

In the next subsections the results obtained using the two sets of data are reported.

6.5.1 Set #1

The first runs have been performed using the first set of data available. The control inputs collected in the control vector $\mathbf{u}$ are the ones reported in Fig. 6.2. The measurements collected in the measurement vector $\mathbf{y}$ are, as discussed in §6.3, pressure drop, fuel flow rate and the mass fraction of ethane at the outlet of the coil.

First, the four parameters adjusted in real-time by the estimator have all been normalized by their nominal value. Then, after a subsequent refinement of the specifications for $\mathbf{P}_0$, $\mathbf{Q}$ and $\mathbf{R}$, the following structure has been chosen:

$$\mathbf{P}_0 = \begin{array}{c} \begin{array}{ccccccc} \mathrm{COT}_{offset} & \beta & \log_{10} k_{c1} & \log_{10} k_{c2} & & & \\ \downarrow & \downarrow & \downarrow & \downarrow & & & \end{array} \\ \begin{pmatrix} 0.01 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0.01 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0.01 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0.01 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \end{array}; \tag{6.15}$$

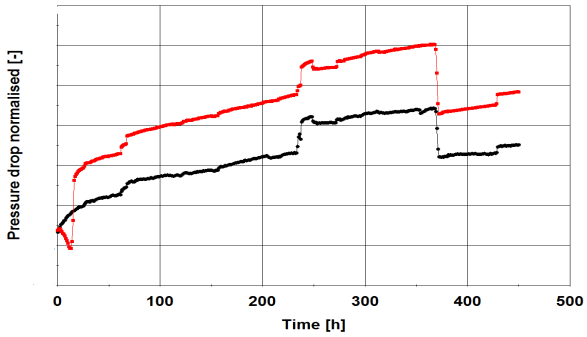
$$\mathbf{Q} = \begin{array}{c} \begin{array}{ccccccc} \mathrm{COT}_{offset} & \beta & \log_{10} k_{c1} & \log_{10} k_{c2} & & & \\ \downarrow & \downarrow & \downarrow & \downarrow & & & \end{array} \\ \begin{pmatrix} 0.01 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0.01 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0.01 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0.01 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0.1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0.1 \end{pmatrix} \end{array}; \tag{6.16}$$

$$\mathbf{R} = \begin{array}{c} \begin{array}{ccc} \Delta p & w_{fuel} & w_{C_2H_6} \\ \downarrow & \downarrow & \downarrow \end{array} \\ \begin{pmatrix} 10^{-7} & 0 & 0 \\ 0 & 10^{-5} & 0 \\ 0 & 0 & 10^{-6} \end{pmatrix} \end{array}. \tag{6.17}$$

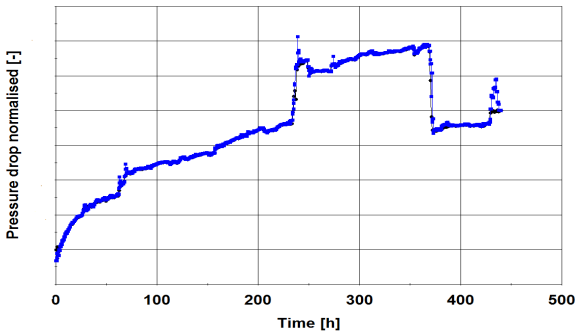
In Fig. 6.6 the available plant data for some of the most important output variables are compared with the OM prediction and the estimator prediction. Black lines are used for the plant data, red lines for the OM prediction and blue lines for the estimator prediction.
As regards the three output variables for which measurements are supplied to the estimator (pressure drop, fuel flow rate and the mass fraction of ethane at the outlet), there is a very good agreement between the estimator prediction and the plant observations. The improvement in the estimation of pressure drop with respect to the OM prediction is mainly due to the real-time adjustment of the two coking parameters; the estimation of the fuel consumption is greatly improved by the adjustment of the NME parameter β and the improvement in the estimation of the mass fraction of ethane at the outlet is related to the real-time estimation of the COT offset. The good agreement between the estimator prediction and the plant observations for these state variables could be expected, considering that the measurement uncertainty variances that have been chosen for these state variables are much lower than their model error variances (eq. (6.16) and (6.17)). However, the reliability and robustness of the estimator prediction can be detected looking at Fig. 6.3h and Tab. 6.2.

Table 6.2: *Comparison between the OM and estimator prediction of TMT with respect to the real plant observations. Set of data #1.*

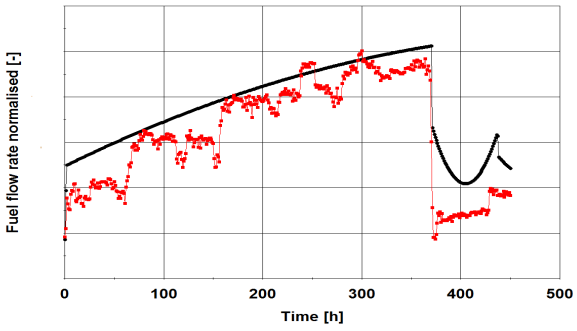
Time [h]	**TMT^{normalized} data**	**TMT^{normalized} OM**	**TMT^{normalized} estimated**
73	0.98995	0.981407	0.990754
144	0.998995	0.990754	0.997186
278	1.011055	1.000201	1.010352
333	1.021106	1.011558	1.021206
362	1.023116	1.012362	1.024221
431	1.015075	1.017186	1.016181
	Std. dev.	**0.009695**	**0.001172**



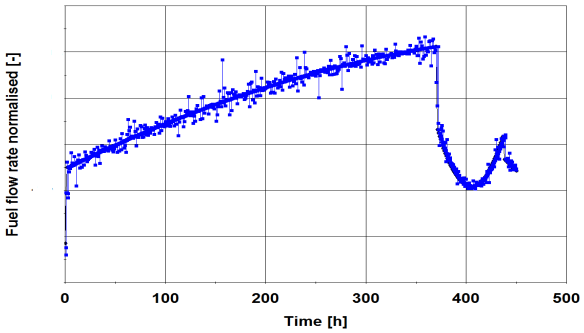
(a) *Pressure drop - OM -*



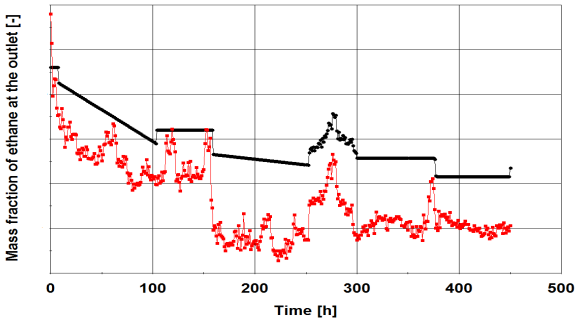
(b) *Pressure drop - estimator-*



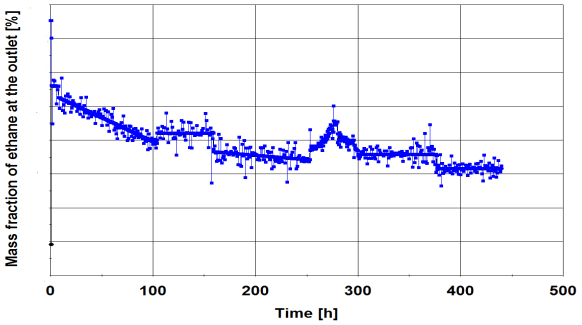
(c) *Fuel flow rate - OM -*



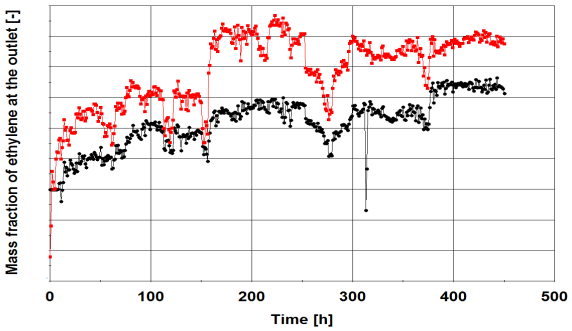
(d) *Fuel flow rate - estimator -*



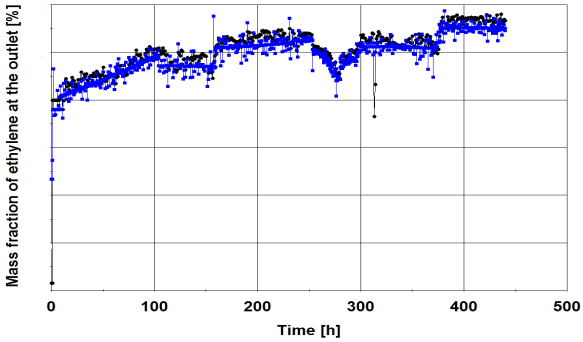
(e) *Mass fraction of ethane at the outlet - OM -*



(f) *Mass fraction of ethane at the outlet - estimator -*



(g) *Mass fraction of ethylene at the outlet - OM -*



(h) *Mass fraction of ethylene at the outlet - estimator -*

Figure 6.3: *Results obtained with the first set of experimental data available. Black lines: plant data. Red lines: OM prediction. Blue lines: estimator prediction. Specifications like in eq.* (6.15), (6.16) *and* (6.17).

In Fig. 6.3h the estimator prediction of the mass fraction of ethylene at the outlet of the coil is compared with the available plant observations. It can be noticed by comparison with Fig. 6.3g that the prediction of this state variable deeply improves when state estimation is performed. It is worth remembering that this variable, despite the availability of its plant observations, has not be given to the estimator as a measurement. The good agreement between its plant measurements and the estimator prediction suggests that the estimator is able to update the state of the system in a way that is consistent with the plant observations. This statement is enhanced if the estimator prediction of TMT is compared with the OM prediction and the occasional plant measurements available (Tab. 6.2). The standard deviation of the values predicted by the estimator is much lower than the standard deviation of the values predicted by the OM. In Tab. 6.2 the comparison is made between the normalized values of TMT in the two cases; in practical terms, the estimation error is decreased from $\sim \pm 10$ degC in the OM case to $\sim \pm 1$ degC by using the EKF. Again, this suggests that the estimator prediction is reliable for all the state variables involved in the model, not only for those for which measurements are supplied.

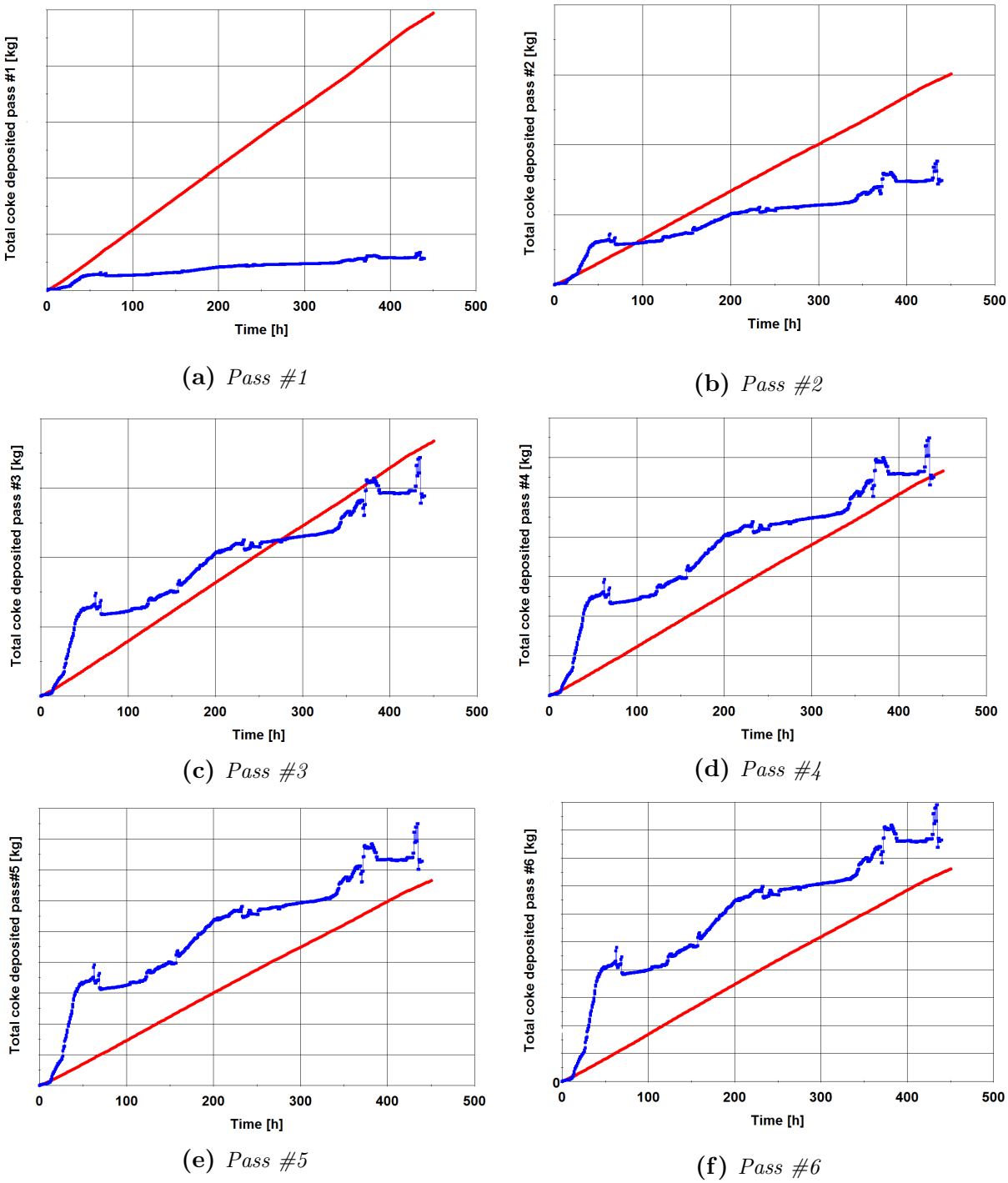
In Fig. 6.4 the prediction of the state of coking of the furnace in the OM and using the EKF are presented. The coking prediction cannot be directly validated since no measurements of the mass of coke deposited on the coil are available. The good estimation of pressure drop is a reason to think that the estimation of coking is satisfactory. However, this is not the only reason. In fact, it has been pointed out that the trend shown in Fig. 6.4, with a lower coke deposition on the first passes and a higher coke deposition on the last passes of the coil with respect to the OM prediction, is in agreement with what is experienced during the operation of the real furnace. This is an additional motivation to believe in the estimator robustness.

Recalling the targets discussed in §6.4, the computational performance of the estimator has been analyzed to understand its feasibility for on-line applications. As suggested by Fig. 6.5, the computational time required by the estimator is much lower than the frequency at which measurements are supplied (every hour). The slowest step is the first prediction step: it requires 72 seconds to predict one hour of furnace operation. The other prediction steps requires ~ 5 seconds to simulate 1 h of operation and each update step is extremely fast (~ 1 s.). It should be noticed that this computational performance has been obtained using a . The computational performance is excellent even on this machine.

6.5.2 Set #2

Given the excellent results obtained using the first set of experimental data, the estimator has been tested using the data for the second run of the furnace. The same feed composition of Tab. 6.1 has been used (pure ethane), and, remembering the third target of §6.4, the same specifications for $\mathbf{P}_0$, $\mathbf{Q}$ and $\mathbf{R}$ have been given to the estimator. The results obtained for the most important output variables are reported in Fig. 6.6. The TMT prediction is presented in Tab. 6.3. The coking prediction given by the estimator compared with the OM prediction is shown in Fig. 6.7.

As in the previous case, the state estimator deeply improves the prediction of the most important output variables of the model. The good agreement between the available plant data of the mass fraction of ethylene and TMT with their estimated values validates the



(a) *Pass #1*

(b) *Pass #2*

(c) *Pass #3*

(d) *Pass #4*

(e) *Pass #5*

(f) *Pass #6*

Figure 6.4: *Total coke deposited in all the 6 passes as predicted by the OM (red lines) and by the estimator (blue lines).*

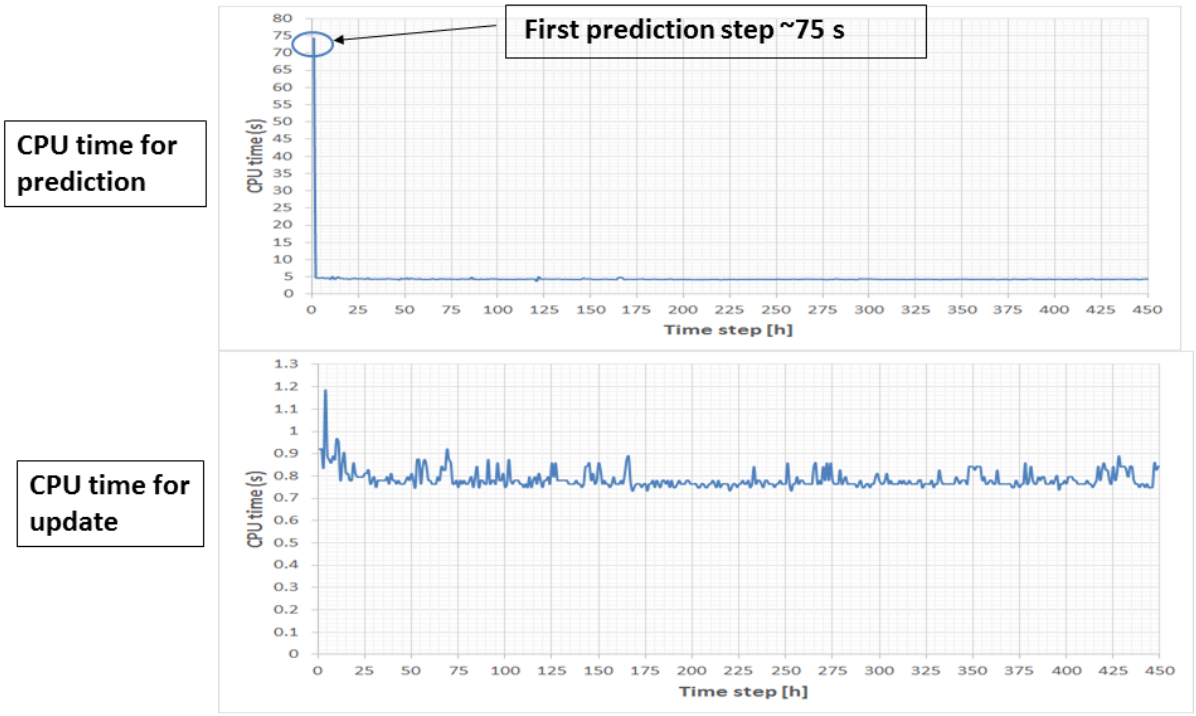


Figure 6.5: *Computational time required by the estimator during each prediction and update step. The slowest step is the first prediction step.*

Table 6.3: *Comparison between the OM and estimator prediction of TMT with respect to the real plant observations. Set of data #2.*

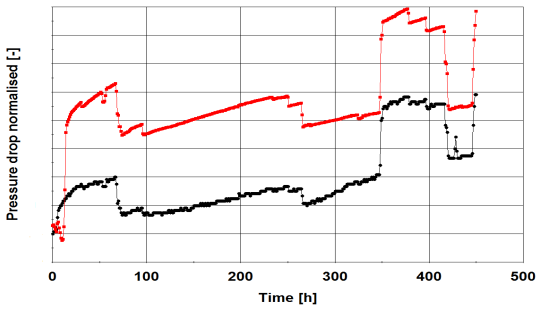
Time [h]	$TMT^{normalized}$ data [-]	$TMT^{normalized}$ OM [-]	$TMT^{normalized}$ estimated [-]
27	0.98794	0.993266	0.985025
165	1.00402	0.994171	1.002211
251	1.013065	1.008141	1.012362
309	1.013065	1.00794	1.013869
370	1.021106	1.009548	1.018291
424	0.99196	1.003216	0.991256
	Std. Dev.	**0.00934**	**0.002065**

good performance of the estimator. The estimated trend of coke deposition on the 6 passes of the coil is in agreement with what is experienced in the real plant. The fact that these good results have been obtained using the same specifications for $\mathbf{P}_0$, $\mathbf{Q}$ and $\mathbf{R}$ as in the previous case fulfills the target #3 of §6.4.

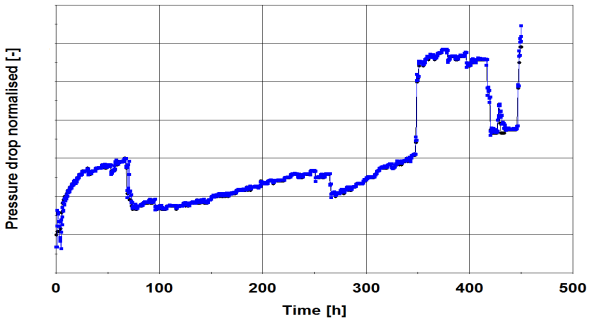
6.6 Results for a pure propane feed

Given the good results obtained for the two sets of experimental data that refer to a pure ethane feed, the estimator has been tested using the other two sets of available data for a pure propane feed. The only two differences that arise when using a pure propane feed with respect to the previous case are:

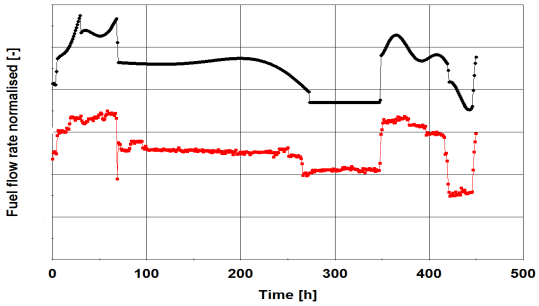
i) an appropriate time-invariant feed composition should be used instead of the one



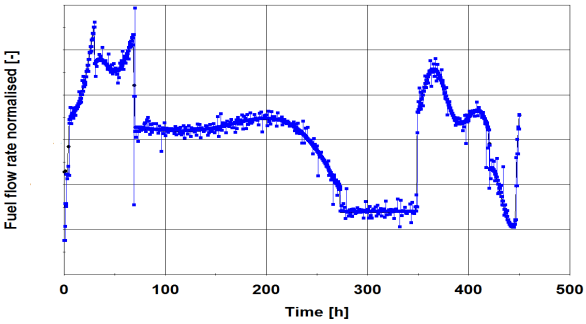
(a) *Pressure drop - OM -*



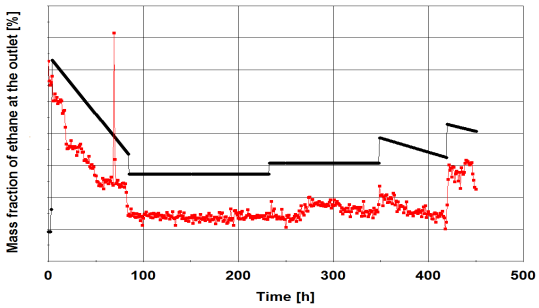
(b) *Pressure drop - estimator-*



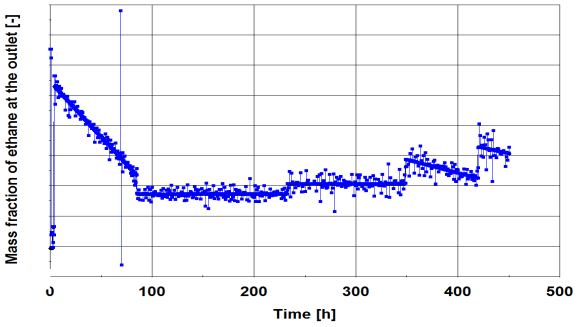
(c) *Fuel flow rate - OM -*



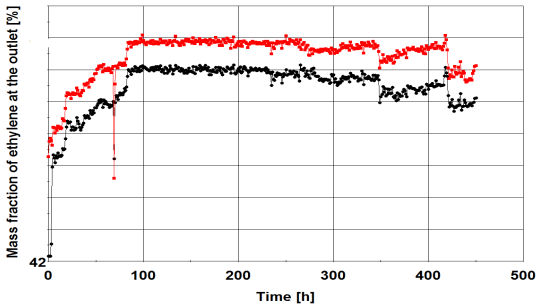
(d) *Fuel flow rate - estimator -*



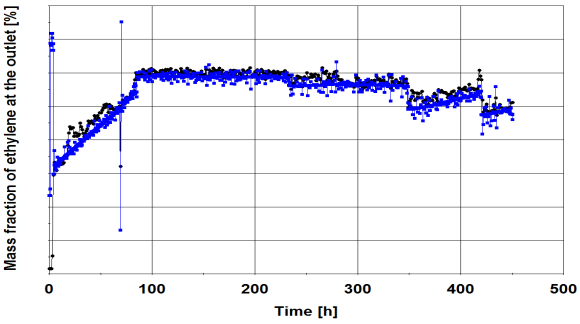
(e) *Mass fraction of ethane at the outlet - OM -*



(f) *Mass fraction of ethane at the outlet - estimator -*

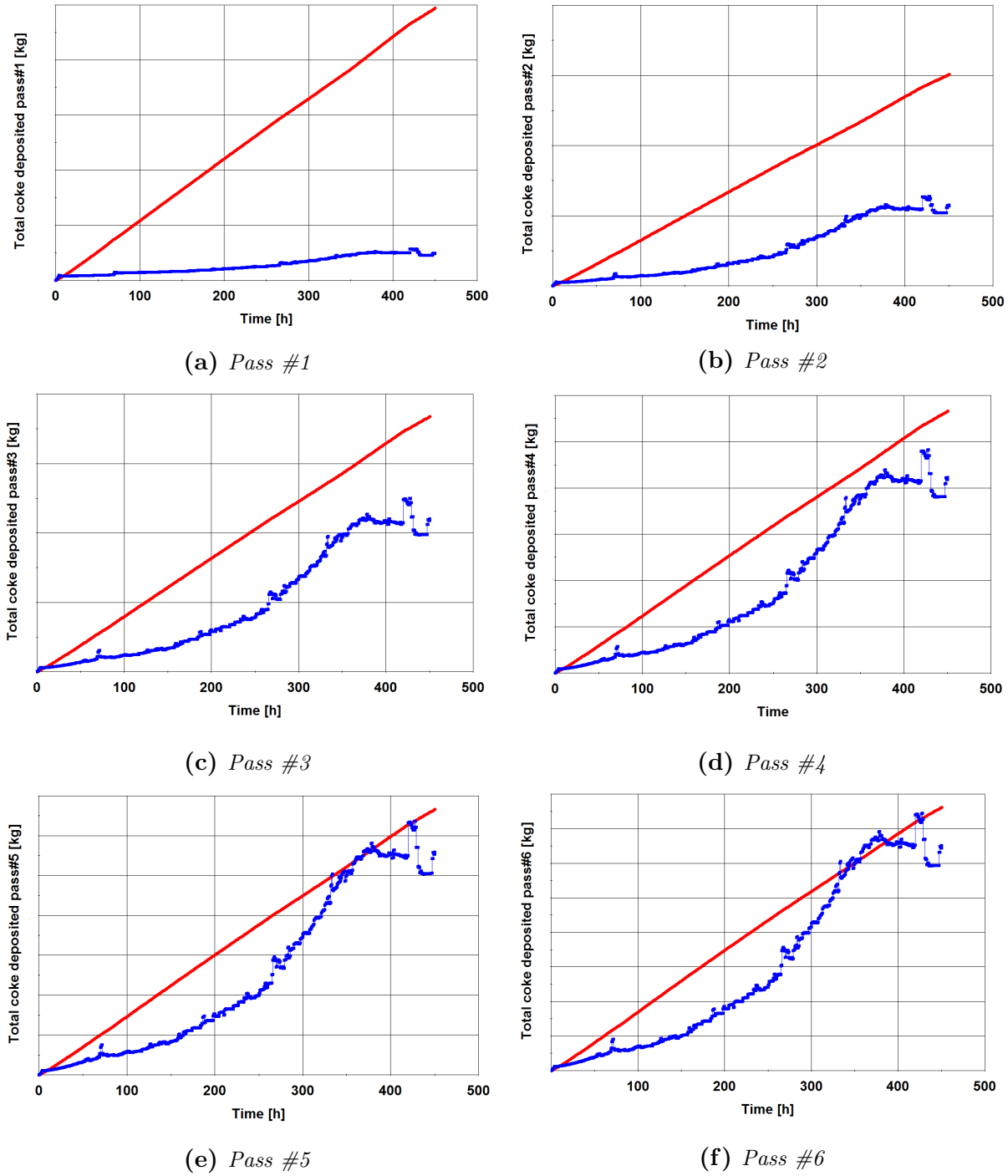


(g) *Mass fraction of ethylene at the outlet - OM -*



(h) *Mass fraction of ethylene at the outlet - estimator -*

Figure 6.6: *Results obtained with the second set of experimental data available. Black lines: plant data. Red lines: OM prediction. Blue lines: estimator prediction. Specifications like in eq.* (6.15), (6.16) *and* (6.17).



(a) *Pass #1* **(b)** *Pass #2*

(c) *Pass #3* **(d)** *Pass #4*

(e) *Pass #5* **(f)** *Pass #6*

Figure 6.7: *Total coke deposited in all the 6 passes as predicted by the OM (red lines) and by the estimator (blue lines). Set of data #2.*

reported in Tab. 6.1;

ii) the composition output variable that needs to be monitored is the mass fraction of propane at the outlet of the coil, not the mass fraction of ethane at the outlet.

As regards points (i), it must be said that the feed composition used for pure ethane of Tab. 6.1 has been derived as an average of plant data and it can be considered as reliable. For a pure propane feed, however, no plant measurements of feed composition has been supplied for the furnace of this study. Therefore, it has been decided to test the estimator using two different time-invariant feed compositions: one derived from literature and another one obtained as an average of available plant data for a different furnace of a different plant. These two feed compositions are reported in Tab. 6.4.

Table 6.4: *Feed composition for a pure propane feed.*

Table 6.5: *From literature (Froment, 1992).*

Component	**Mass fraction [%]**
Methane	0
Ethylene	0
Ethane	2.11
Propylene	0
Propane	94.42
1- Butene	3.47
C_5^+	0

Table 6.6: *From historical data of another furnace.*

Component	**Mass fraction [%]**
Methane	0
Ethylene	0
Ethane	1.85
Propylene	0.33
Propane	88.55
1- Butene	4.06
C_5^+	5.21

It is worth noticing that the uncertainty on the feed composition is an added element of uncertainty for the model prediction: the task of the estimator to adjust the model prediction is then even more complicate than in the previous case.

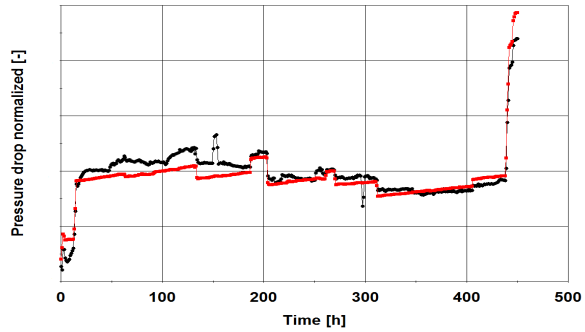
6.6.1 Set #3

The results obtained for the first set of data for a pure propane feed are reported in Fig. 6.8 and Tab. 6.7. Only the results obtained using the feed composition derived from literature (Tab. 6.5) are reported: the estimator has also been tested using the feed composition of Tab. 6.6 and similarly good results have been obtained. This confirms again the robustness of the estimator performance.

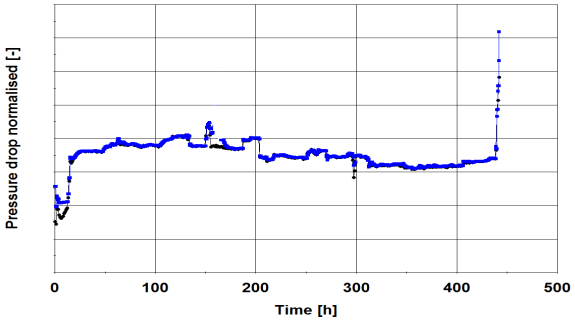
Table 6.7: *Comparison between the OM and estimator prediction of TMT with respect to the real plant observations. Set of data #3.*

Time [h]	**TMT^{normalized} data [-]**	**TMT^{normalized} OM [-]**	**TMT^{normalized} estimated [-]**
17	1.017085	1.014774	1.016482
73	1.00804	0.998392	1.009347
151	1.020101	1.006231	1.018693
279	1.01809	1.011558	1.017387
413	1.020101	1.012462	1.019598
	Std. Dev.	**0.008852**	**0.00098**

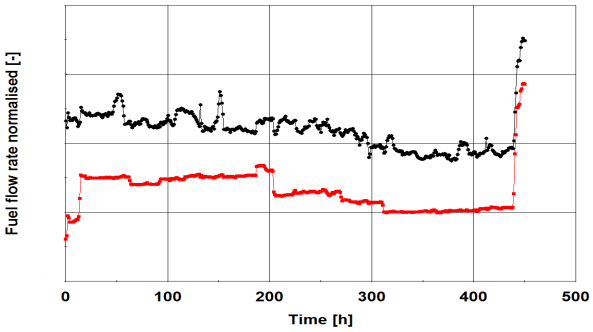
It can be noticed from Fig. 6.8 that the results are satisfactory for all the most important output variables of the model. The results obtained for the mass fraction of



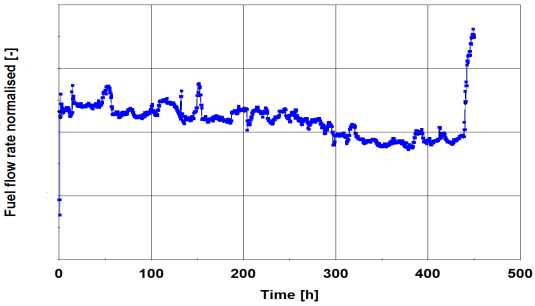
(a) *Pressure drop - OM -*



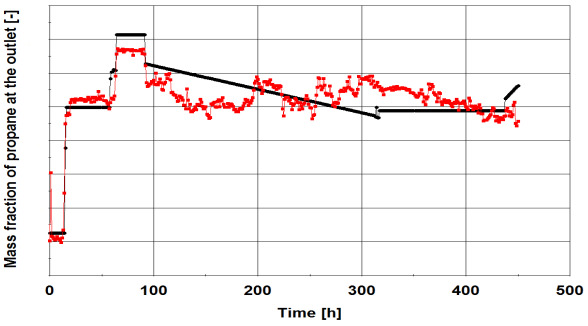
(b) *Pressure drop - estimator-*



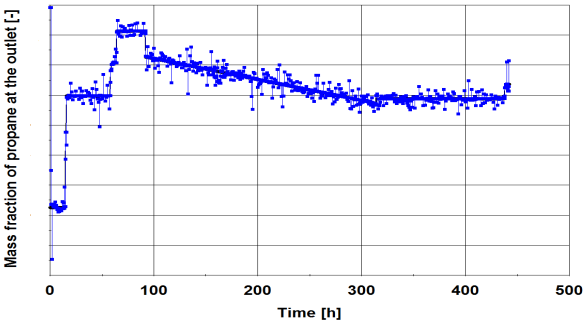
(c) *Fuel flow rate - OM -*



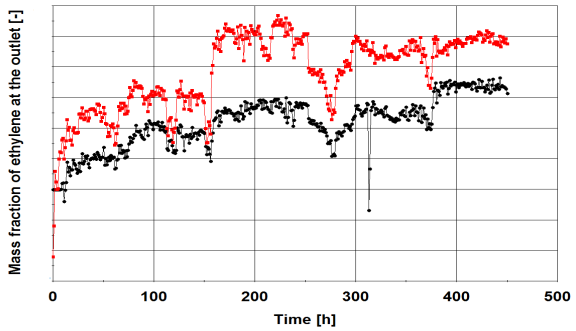
(d) *Fuel flow rate - estimator -*



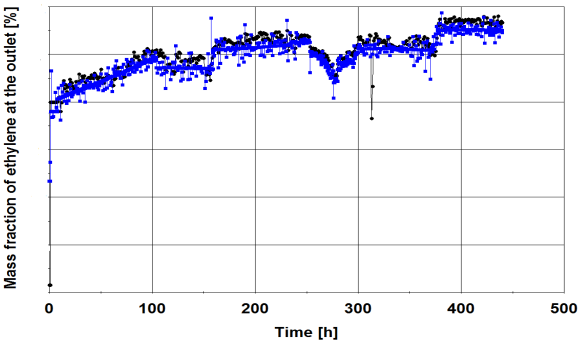
(e) *Mass fraction of propane at the outlet - OM -*



(f) *Mass fraction of propane at the outlet - estimator -*



(g) *Mass fraction of ethylene at the outlet - OM -*



(h) *Mass fraction of ethylene at the outlet - estimator -*

Figure 6.8: *Results obtained with the thirdset of experimental data (propane feed). Black lines: plant data. Red lines: OM prediction. Blue lines: estimator prediction. Specifications like in eq.* (6.15), (6.16) *and* (6.17).

ethylene and TMT (Tab. 6.7) are a proof of the reliaibility of the estimator prediction. It is worth remembering that these results have been obtained using exactly the same specifications of $\mathbf{P}_0$, $\mathbf{Q}$ and $\mathbf{R}$ as with a pure ethane feed. This is in line with target #3 of §6.4. The coking prediction has proven to be in this case, again, in agreement with what is experienced in the real plant.

6.6.2 Set #4

Like for set #3, the results obtained using the last set of data available are satisfactory. They have been obtained using the feed composition taken from literature of Tab. 6.5. Same results have been obtained using the feed composition of Tab. 6.6. The prediction of the mass fraction of ethylene and TMT (Tab. 6.8) has validated the quality of the estimator prediction.

Table 6.8: *Comparison between the OM and estimator prediction of TMT with respect to the real plant observations. Set of data #4.*

Time [h]	**TMT^{normalized} data [-]**	**TMT^{normalized} OM [-]**	**TMT^{normalized} estimated [-]**
10	1.017085	0.984422	1.018191
127	1.037186	0.999698	1.034975
283	1.025126	1.016985	1.026231
380	1.022111	1.009447	1.022613
448	1.024121	1.016985	1.02402
	Std. Dev	**0.023451**	**0.001233**

6.7 Conclusions

The performance of the estimator using the four sets of historical data of the furnace has proven to be satisfactory independently of the type of data considered. A good prediction of all the state variables has been obtained despite using the same specifications for $\mathbf{P}_0$, $\mathbf{Q}$, $\mathbf{R}$ and despite the uncertainty on the feed composition (especially for the composition of the propane feed). The time required by the estimator to perform every prediction step is much smaller than the update frequency at which measurements are supplied (the slowest prediction step is the first one and it takes $\sim$ 72 s, while the measurement update frequency is 1 hour). The results confirm the robustness of the estimator and its compatibility with an on-line implementation.

Conclusions

The objective of this study was to implement a robust and reliable state estimation technique on a preexisting dynamic model of a steam cracker using gPROMS ® ModelBuilder. The model considered is a PDAE system composed by 227 state differential variables and 14268 state algebraic variables. The targets to be reached were:

i) obtaining a correct estimate of all the state variables involved in the model and particularly of the state of coking of the pyrolysis furnace;

ii) finding a set of specifications for the estimator (initial/model/measurement error covariances) that can guarantee a good performance independently of the type of feed considered and the type and number of measurements supplied to the estimator;

iii) obtaining a computational performance compatible with a future on-line implementation of the estimator in the real plant. In other terms, every prediction step performed by the estimator should never take longer than the frequency at which measurements are supplied.

The final goal is to implement on-line the estimator in the real plant in order to obtain a real-time prediction of the state of coking of the furnace. If the state of coking is known, a cyclic whole plant optimization can be performed without assuming (as it is currently done) clean-tube conditions for the coil of the cracker.

In order to reach these targets, the project has been organized as follows.
First, a solid background on the original model of the steam cracker has been obtained, trying to understand its general structure, its main variables and equations, its most influential parameters. The weaknesses and assumptions of the model have been critically discussed and the mismatch between the model prediction and the plant observations analyzed.
Second, a deep understanding of the mathematical formalism behind the most important state-estimation algorithms currently available has been obtained. These algorithms have been implemented in MATLAB ® and their performances have been compared using some simple test examples. The pro's and con's of each algorithm have been studied. As a first choice, it has been decided to test the EKF (Extended Kalman Filter) on the furnace model.
Third, the practical issues to be considered when implementing the EKF in gPROMS® have been faced, so as to define an effective estimation procedure based on the usage of three foreign objects (FOs) to store the control vector, the measurement vector and the initial/model/measurement covariances to be passed to the estimator.
Then, the estimator has been tested off-line using measurements generated with the original model prediction. A slight modification on some of the model parameters and/or initial conditions (ICs) has been carried out and the ability of the estimator to adjust its prediction despite using a wrong model (i.e. with modified parameters and/or ICs) has been tested. Given the good results obtained for both time-invariant and time varying inputs, it has been decided to move to the next step.
In the last step, the estimator has been tested off-line using four set of historical data of

the furnace, two for a pure ethane feed, the others for a pure propane feed. 450 hours of furnace operation have been estimated with an hourly measurement update frequency.

The results obtained using real plant data have proved to be satisfactory independently of the set of historical data being considered. The prediction of the estimator is in good agreement with the plant observations not only for those state variables for which measurements are supplied (pressure drop, fuel flow rate, mass fraction of ethane at the outlet) but also for other state variables (mass fraction of ethylene at the outlet, tube metal temperature). The prediction of the state of coking of the furnace is in agreement with what is actually experienced in the real plant. From a computational point of view, the estimator requires 72 seconds for the first prediction step and ~ 5 s for the other prediction steps to simulate 1 hour of furnace operation. Thus, its on-line implementation is feasible and no computational problems can arise.

The EKF is usually considered a low-performance estimator when applied to highly nonlinear systems like the one analyzed in this study. Its bad performance is usually justified in the open literature by considering the linearized expression of the Riccati equation used to propagate the error covariance during its prediction step. However, there are at least two reasons that can be put forward to justify the good results obtained from the EKF in this study despite the highly nonlinear model:

- a clever implementation of the estimator (§3.2.1) is adopted in gPROMS® with respect to the usual EKF implementation;
- the real-time adjustment of some of the model parameters (two pre-exponential factors of the coking rate expression, a correction parameter for the firebox model, the COT offset) can ease the effort of the estimator to reconcile the model prediction with the plant observations. In other terms, these real-time adjusted parameters give to the estimator a prompt and specific knowledge of the hidden model weaknesses thus enhancing the effectiveness of its intervention.

The results have been obtained using the same structure for the initial/model/measurement uncertainty covariances for all the four sets of data considered, in line with target ii).

Considering the potential benefits that could be obtained by the on-line implementation of the estimator, it has been decided to test the estimator using the historical data of another steam cracker situated in a different ethylene's plant. If good results were obtained, the on-line implementation of the estimator will be performed on this furnace.

Furthermore, future work will aim at implementing the following improvements:

1. modification of the original EKF code to make it suitable for partial measurements (i.e. measurements supplied randomly without a predetermined frequency). This could be extremely useful if measurements of TMT (tube metal temperature), that are taken only occasionally in the real plant, need to be supplied to the estimator;
2. modification of the EKF code to handle constraints on some or all the state variables. This can be useful to avoid divergence issues related to bounds violation for some of the state variables.

Another possibile development is the implementation in gPROMS® of different state estimation algorithms (UKF/PF/EnKF) to assess their performances with complex models like the one considered in this study.

Appendix A

Codes

In this Appendix all the gPROMS® and MATLAB® codes used for the simulations in the different chapters are presented.

Table A.1: *gPROMS® process entities and case files related to the second chapter.*

Process entity	Case file	Description
Furnace_OM_TI_Run_1	Furnace_OM_TI_Run_1.gCS	Original model prediction with time-invariant inputs for the first run of the furnace (feed: pure ethane)
Furnace_OM_TI_Run_2	Furnace_OM_TI_Run_2.gCS	Original model prediction with time-invariant inputs for the second run of the furnace (feed:pure ethane)
Furnace_OM_TI_Run_3	Furnace_OM_TI_Run_3.gCS	Original model prediction with time-invariant inputs for the third run of the furnace (feed: pure propane)
Furnace_OM_TI_Run_4	Furnace_OM_TI_Run_4.gCS	Original model prediction with time invariant inputs for the fourth run of the furnace (feed: pure propane)
Furnace_OM_TV_Run_1	Furnace_OM_TV_Run_1.gCS	Original model prediction with time-varying inputs for the first run of the furnace (feed: pure ethane)
Furnace_OM_TV_Run_2	Furnace_OM_TV_Run_2.gCS	Original model prediction with time-varying inputs for the second run of the furnace (feed:pure ethane)
Furnace_OM_TV_Run_3	Furnace_OM_TV_Run_3.gCS	Original model prediction with time-varying inputs for the third run of the furnace (feed: pure propane)
Furnace_OM_TV_Run_4	Furnace_OM_TV_Run_4.gCS	Original model prediction with time-varying inputs for the fourth run of the furnace (feed: pure propane)

Table A.2: *MATLAB® codes used in the third chpater.*

Code	Description
Solvers	
`ekf.m`	Extended Kalman filter
`ukf.m`	Unscented Kalman filter
`Enkf.m`	Ensemble Kalman filter
`pf.m`	Particle filter
Test examples	
`batch_homogeneous.m`	Test example #1: homogeneous batch reactor, single reaction (ODE 2×2)
`batch_reversible.m`	Test example #2: homogeneous batch reactor, two reversible reactions (ODE 3×3)
`plane_flying.m`	Test example #3: tracking the trajectory of a plane flying using state estimation (ODE×3)

Table A.3: *gPROMS® process entities and case files used in the fifth chapter (Part 1).*

Process entity	Case file	Description
Furnace_OM_TI_Run_1	Furnace_OM_TI_Run_1.gCS	Original model prediction with time-invariant inputs for the first run of the furnace (feed: pure ethane)
Generate_Measurements_1	Generate_Measurements_1.gCS	Generation of model-based measurements using time-invariant inputs for the first run of the furnace (feed: pure ethane)
Furnace_wrong_TI_Run_1	Furnace_wrong_TI_Run_1.gCS	Wrong model prediction (modified parameters) using time-invariant inputs for the first run of the furnace (feed: pure ethane)
Furnace_S_E_TI_Run_1	Furnace_S_E_TI_Run_1.gCS	State estimation for the first run of the furnace using time-invariant inputs
Furnace_OM_TI_Run_2	Furnace_OM_TI_Run_3.gCS	Original model prediction with time-invariant inputs for the second run of the furnace (feed: pure ethane)
Generate_Measurements_2	Generate_Measurements_2.gCS	Generation of model-based measurements using time-invariant inputs for the second run of the furnace (feed: pure ethane)
Furnace_wrong_TI_Run_2	Furnace_wrong_TI_Run_2.gCS	Wrong model prediction (modified parameters) using time-invariant inputs for the second run of the furnace (feed: pure ethane)
Furnace_S_E_TI_Run_2	Furnace_S_E_TI_Run_2.gCS	State estimation for the second run of the furnace using time-invariant inputs

Table A.4: *gPROMS® process entities and case files used in the fifth chapter (Part 2).*

Process entity	Case file	Description
Furnace_OM_TI_Run_3	Furnace_OM_TI_Run_3.gCS	Original model prediction with time-invariant inputs for the third run of the furnace (feed: pure propane)
Generate_Measurements_3	Generate_Measurements_3.gCS	Generation of model-based measurements using time-invariant inputs for the third run of the furnace (feed: pure propane)
Furnace_wrong_TI_Run_3	Furnace_wrong_TI_Run_3.gCS	Wrong model prediction (modified parameters) using time-invariant inputs for the third run of the furnace (feed: pure propane)
Furnace_S_E__TI_Run_3	Furnace_S_E_TI_Run_3.gCS	State estimation for the third run of the furnace using time-invariant inputs
Furnace_OM_TI_Run_4	Furnace_OM_TI_Run_4.gCS	Original model prediction with time-invariant inputs for the fourth run of the furnace (feed: pure propane)
Generate_Measurements_4	Generate_Measurements_4.gCS	Generation of model-based measurements using time-invariant inputs for the fourth run of the furnace (feed: pure propane)
Furnace_wrong_TI_Run_4	Furnace_wrong_TI_Run_4.gCS	Wrong model prediction (modified parameters) using time-invariant inputs for the fourth run of the furnace (feed: pure propane)
Furnace_S_E__TI_Run_4	Furnace_S_E_TI_Run_4.gCS	State estimation for the fourth run of the furnace using time-invariant inputs

Table A.5: *gPROMS® process entities and case files used in the fifth chapter (Part 3).*

Process entity	Case file	Description
Furnace_OM_TV_Run_1	Furnace_OM_TV_Run_1.gCS	Original model prediction with time-varying inputs for the first run of the furnace (feed: pure ethane)
Generate_Measurements_1	Generate_Measurements_1.gCS	Generation of model-based measurements using time-varying inputs for the first run of the furnace (feed: pure ethane)
Furnace_wrong_TV_Run_1	Furnace_wrong_TV_Run_1.gCS	Wrong model prediction (modified parameters) using time-varying inputs for the first run of the furnace (feed: pure ethane)
Furnace_S_E__TV_Run_1	Furnace_S_E_TV_Run_1.gCS	State estimation for the first run of the furnace using time-varying inputs
Furnace_OM_TV_Run_2	Furnace_OM_TV_Run_3.gCS	Original model prediction with time-varying inputs for the second run of the furnace (feed: pure ethane)
Generate_Measurements_2	Generate_Measurements_2.gCS	Generation of model-based measurements using time-invariant inputs for the second run of the furnace (feed: pure ethane)
Furnace_wrong_TV_Run_2	Furnace_wrong_TV_Run_2.gCS	Wrong model prediction (modified parameters) using time-varying inputs for the second run of the furnace (feed: pure ethane)
Furnace_S_E__TV_Run_2	Furnace_S_E_TV_Run_2.gCS	State estimation for the second run of the furnace using time-varying inputs

Table A.6: *gPROMS® process entities and case files used in the fifth chapter (Part 4).*

Process entity	Case file	Description
Furnace_OM_TV_Run_3	Furnace_OM_TV_Run_3.gCS	Original model prediction with time-varying inputs for the third run of the furnace (feed: pure propane)
Generate_Measurements_3	Generate_Measurements_3.gCS	Generation of model-based measurements using time-varying inputs for the third run of the furnace (feed: pure propane)
Furnace_wrong_TV_Run_3	Furnace_wrong_TV_Run_3.gCS	Wrong model prediction (modified parameters) using time-varying inputs for the third run of the furnace (feed: pure propane)
Furnace_S_E__TV_Run_3	Furnace_S_E_TV_Run_3.gCS	State estimation for the third run of the furnace using time-varying inputs
Furnace_OM_TV_Run_4	Furnace_OM_TV_Run_4.gCS	Original model prediction with time-varying inputs for the fourth run of the furnace (feed: pure propane)
Generate_Measurements_4	Generate_Measurements_4.gCS	Generation of model-based measurements using time-varying inputs for the fourth run of the furnace (feed: pure propane)
Furnace_wrong_TV_Run_4	Furnace_wrong_TV_Run_4.gCS	Wrong model prediction (modified parameters) using time-varying inputs for the fourth run of the furnace (feed: pure propane)
Furnace_S_E__TV_Run_4	Furnace_S_E_TV_Run_4.gCS	State estimation for the fourth run of the furnace using time-varying inputs

Table A.7: *gPROMS® process entities and case files used in the sixth chapter.*

Process entity	Case file	Description
Furnace_OM_Run_1	Furnace_OM_Run_1.gCS	Original model prediction for the first run of the furnace (feed: pure ethane)
Furnace_S_E_Plant_Run_1	Furnace_S_E_Plant_Run_1.gCS	State estimation for the first run of the furnace using real plant data
Furnace_OM_Run_2	Furnace_OM_Run_2.gCS	Original model prediction for the second run of the furnace (feed: pure ethane)
Furnace_S_E_Plant_Run_2	Furnace_S_E_Plant_Run_2.gCS	State estimation for the second run of the furnace using real plant data
Furnace_OM_Run_3	Furnace_OM_Run_3.gCS	Original model prediction for the third run of the furnace (feed: pure ethane)
Furnace_S_E_Plant_Run_3	Furnace_S_E_Plant_Run_3.gCS	State estimation for the third run of the furnace using real plant data
Furnace_OM_Run_4	Furnace_OM_Run_4.gCS	Original model prediction for the first run of the furnace (feed: pure ethane)
Furnace_S_E_Plant_Run_4	Furnace_S_E_Plant_Run_4.gCS	State estimation for the first run of the furnace using real plant data

Acknowledgments

First of all, I would like to thank Process Systems Enterpirse Ltd. for giving me the opportunity to work for six months in a dynamic and collaborative environment. I really enjoyed working with such smart people. Special thanks to prof. Pantelides, Catherine, Trudie, Sreekumar, Stepan and all the new friends that I met during my stay in London: Joao, Joana, Andre, Frederico, Maria, Pierre, JJ, Byeionghil, Yera, Cristobal, Gerardo, Miguel, Niall, Mario, Hugo, Francisco, Mariana.

Thanks to my PSE supervisor, Ying-Sheng Cheng, for his patience, helpfulness and his incredible professionalism. It has been a pleasure working with you during my internship.

I would like to thank prof. Fabrizio Bezzo and prof. Massimiliano Barolo, for their patient and valuable advice.

Finally, thanks to my family, for all the support they gave me during my studies.

Bibliography

Albright, L.F. 2002. Comments on 'Kinetic Modeling of Coke Formation during Steam Cracking'. *Ind. Eng. Chem. Res.*, **41**, 6210–6212.

Arulampalam, M.S., Maskell, S., Gordon, N., & Clapp, T. 2002. A tutorial on particle filters for Online Nonlinear-Non Gaussian Bayesian tracking. IEEE Transactions on Signal Processing.

Biagiola, Silvina I., & Figueroa, Jose' L. 2002. State Estimation in Nonlinear Processes. Aplication to pH Process control. *Ind. Eng. Chem.Res*, **41**, 4777–4785.

Browne, J., Broutin, P., & Ropital, F. 1999. Coke deposition under steam cracking conditions- Study of the influence of the feedstock conversion by micropilots experiments. *Materials and Corrosion*, **49**, 360–366.

Cheng, Y.S., Mongkhonsi, T, & Kershenbaum, LS. 1996. Nonlinear dynamic estimation for a fixed-bed reactor with decaying catalysts. Pergamon-Elsevier Science Ltd.

Cheng, Y.S, Mongkhonsi, T., & Kershenbaum, L.S. 1997. Sequential estimation for nonlinear differential and algebraic systems - Theoretical development and application. *Computers and Chemical Engineering*, **21**, 1051–1067.

Ching, J., Beck, J.L., & K.A.Porter. 2006. Bayesian state and parameter estimation of uncertain dynamical systems. *Probabilistic Engineering Mechanics*, **21**, 81–96.

Curn, J. 2014. *Correlated Estimation Problems and the Ensemble Kalman Filter*. Ph.D. thesis, University of Dublin, Trinity College.

Dyer, M.S. 2000. *A Multiscale Approach to State Estimation with Applications in process operability Analysis and Model Predictive Control*. Ph.D. thesis, Massachussetts Institute of Technology.

Eramo, M. 2013. *Global Ethylene Market Outlook: Low cost feedstocks fuel the next wave of investments in North America and China*. Tech. rept. Technip.

Evensen, Gier. 2003. The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*, **53**, 343–367.

Froment, G., & Bischoff, K. 1980. *Chemical reactor analysis and design*. 2nd edn. Wiley Series in Chemical Engineering.

Froment, G.F. 1992. Kinetics and reactor design in the thermal cracking for olefins production. *Chemical Engineering Science*, **47**, 2163–2177.

Gal, T., & Lakatos, B.G. 2003. Modeling and simulation of steam crackers. *Petroleum and Coal*, **45**, 96–98.

Gillijns, S., Mendoza, O. Barrero, Chandrasekar, J., Moor, B.L.R. De, Bernstein, D.S., & Ridley, A. (eds). 2006. *What is the Ensemble Kalman Filter and How well Does it Work?*

Goosens, A.G, Dente, M., & Ranzi, E. 1978. Improved steam cracker operation. *Hydrocarbon Process*, **57**(9), 227–236.

Haseltine, E.L., & Rawlings, J.B. 2005. Critical Evaluation of Extended Kalman Filter and Moving-Horizon Estimation. *Ind. Eng. Chem. Res*, **44**, 2451–2460.

Ikonen, E. 2013. *Model Predictive Control and State Estimation.* Tech. rept. University of Oulu.

Julier, S.J., & J.K.Uhlmann. 2010. Unscented filtering and Nonlinear Estimation. IEEE.

Knudsen, J., & Katz, D. 1958. *Fluid Dynamics and Heat Transfer.* McGraw-Hill.

Laugier, J-P. 2013. *A wider Ethylene offering.* Tech. rept. Technip.

Laviola, J.J. 2003. A comparison of unscented and extended Kalman filtering for estimating quaternion motion. IEEE.

Lee, J. H. 2012. A stochastic approach to state estimation: current status and open problems. Presented as FIPSE, Olympian Village, Western Peloponnese, Greece.

Luo, X., & Moroz, I.M. 2009. Ensemble Kalman filter with the unscented transform. *Physica D: Nonlinear Phenomena*, **238**, 549–562.

Mahulkar, A.V., Heynderickx, G.J., & Marin, G.B. 2014. Simulation of the coking phenomenon in the superheater of a steam cracker. *Chemical Engineering Science*, **110**, 31–43.

Mandela, R.K., Rengaswamy, R., & Narashimhan, S. (eds). 2007. *Nonlinear State Estimation of Differential Algebraic Systems.*

Nekrasov, B. 1969. *Hydraulics.* Peace Publishers.

Plehiers, P.M., Reyniers, G.C., & Froment, G.F. 1977. Simulation of the run length of an ethane cracking furnace. *Ind. Eng. Chem. Res.*, **27**, 946–951.

Rao, C.V., Rawlings, J.B., & Mayne, D.Q. 2013. Constrained State Estimation for Nonlinear Discrete-Time Systems: Stability and Moving Horizon Approximations. *IEEE Transactions on Automatic Control*, **48**(2).

R.H, Perry, & D.W., Green. 2009. *Perry's chemical engineers' handbook.* 8th edn.

Ristic, B., Arulampalam, S., Gordon, N., & Clapp, T. 2004. Beyond the Kalman Filter: Particle Filters for Tracking Applications.

Romanenko, A., & Castro, J.A.A.M. 1980. The Unscented filter as an alternative to the EKF for nonlinear state estimation: a simulation case study. *Comput. Chem. Eng.*, **28**(3), 347–355.

Romanenko, A., Santos, L.O., & Alfonso, P.A.F.N.A. 2004. Unscented Kalman Filtering of a Simulated pH System. *Ind. Eng. Chem. Res*, **43**, 7531–7538.

Schietekat, Carl M., Cauwenberge, David J. Van, Geem, Kevin M. Van, & Marin, Guy B. 2014. Computational Fluid Dynamics-Based Design of Finned Steam Cracking Reactors. *AIChE Journal*, **60**(2).

Simon, D. 2006. *Optimal State Estimation. Kalman, H_∞ and Nonlinear approaches.* Wiley. Part IV.

Sundaram, K.M., & Froment, G.F. 1977. Modeling of thermal cracking kinetics - I-: Thermal cracking of ethane, propane and their mixtures. *Chemical Engineering Science*, **32**, 601–608.

Sundaram, K.M., & Froment, G.F. 1979. Kinetics of coke deposition in the thermal cracking of propane. *Chemical Engineering Science*, **34**, 635–644.

Sundaram, K.M., & Froment, G.F. 1980. Two-dimensional model for the simulation of tubular reactors for thermal cracking. *Chemical Engineering Science*, **35**, 364–371.

Sundaram, K.M., Damme, P.S. Van, & Froment, G.F. 2004. Coke deposition in the thermal cracking of ethane. *AIChE Journal*, **32**, 601–608.

Technip- Pyrotec Division. *SPYRO® Suite 7.*

Turner, R. 2010. Model based learning of sigma points in Unscented Kalman Filtering. IEEE.

Wan, E.A., & van der Merwe, R. 2000. The Unscented Kalman Filter for Nonlinear Estimation. IEEE.

Wang, Jidong, Reyniers, Marie-Francoise, & Marin, G.B. 2007. Influence of dimethyl disulfide on Coke formation during Steam Cracking of Hydrocarbons. *Ind. Eng. Chem.Res*, **46**, 4134–4148.

Wauters, S., & Marin, G.B. 2002. Kinetic Modeling of Coke Formation during Steam Cracking. *Ind. Eng. Chem. Res.*, **41**, 2379–2391.

Zhang, G., & Evans, B. 2012. Progress of Modern Pyrolysis Furnace Technology. *Advances in Materials Physics and Chemistry*, **2**, 169–172.

Zimmermann, H., & Walzi, R. 2009. Ethylene. *In: Ullmann's Encyclopedia of Industrial Chemistry*, 7th edn. Wiley-VCH.